Transcriptional Response of *Mycobacterium tuberculosis* on Encounter with Neutrophil *In Vitro*

Amin Aziz Bakir

Thesis submitted to The University of Leicester for the degree of Doctor of Philosophy

March 2016

Department of Infection, Immunity and Inflammation
College of Medicine, Biological Sciences and Psychology
University of Leicester
University Road, Leicester, LE1 7RH

Abstract

Background: Neutrophils are reported to be the dominant phagocytes in sputum from tuberculosis (TB) patients and may explain the distinctive pattern of gene expression of *Mycobacterium tuberculosis* (Mtb) in these samples. Moreover, neutrophils play an important role during tuberculosis both in collaboration with macrophages as well as antimicrobial functions such as NETosis. Therefore, understanding the transcriptional changes of mycobacteria in response to neutrophils is of general interest.

Methods: Following preliminary experiments to develop appropriate methods using BCG and a neutrophil cell line, primary human neutrophils were infected with stationary phase H37Rv preconditioned in RPMI overnight then sampled at 4 and 24 hours. Bacterial RNA samples from these experiments and from TB sputum were subjected to RNA-seq analysis by Illumina NextSeq-500. Reads were normalised by two different methods and differentially expressed genes were detected at q-value < 0.01. Selected results were confirmed by qRT-PCR. Transcription factors associated with detected genes were identified.

Results: The colony counts of H37Rv incubated with human neutrophils reduced ~10-fold at 24 hours. Striking up regulation of Mtb *tgs1*, *hspX* and *icl1* was observed during incubation in RPMI relative to exponentially growing bacilli. RNA-seq revealed 90 genes differentially expressed (28 repressed and 62 induced) in neutrophil encounters at 4 hours relative to RPMI controls. DosR and sigK were the most significantly associated transcription factors with the induced genes.

Discussion: Comparison of the neutrophil-stimulated and sputum Mtb transcriptomes did not support the hypothesis that the phagocytes provide a key stimulus underpinning the sputum pattern. SigK regulated expression appears to be a significant aspect of the neutrophil stimulated Mtb transcriptome.

Acknowledgement

First of all, I would like to express my sincere gratitude to my supervisor Prof. Mike Barer for the constant support of my PhD study, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. His deep compassion is what I most admire; I could not have imagined a greater support in hard moments of my life.

My special thanks goes to Dr. Bernard Burke for his valuable supervision at the start of this work and continuous encouragement during my study.

Besides my supervisors, I would like to thank my PRP members Dr. Helen O'Hare and Dr. Galina Mukomolova for their intuitive comments, guidance and also hard questions the hard question which incented me to widen my research from various perspectives. My sincere thanks also goes to Dr. Natalie Garton for her endless help and patience in Cat 3 training as well as other laboratory issues.

I would like to thank my dear friend Ali Yavari who kindly helped me in analyzing RNA-seq data. I thank my fellow labmates Gosia, Abdul, Jonathan, Andrew and Caroline for working together, for the stimulating discussions, and for all the fun we have had in the last four years. I would like to thank Irina for help in providing the material and discussing various issues.

I wish to thank all people in 3i department and University of Leicester who have been involved in this research for giving me the opportunity to fulfill this project.

I would like to express my deepest thanks to my parents for their belief in me and care during every stage of my life. Special thanks go to my wife Lajan for her enormous support during my PhD

Finally, I want to thank the human capacity development program for providing the financial support that made this PhD possible.

List of Abbreviation

ACDP	Advisory Committee on Dangerous Pathogens
ADC	Albumin-dextrose-catalase
AIDS	Acquired immunodeficiency syndrome
ASFD	Alternative sigma factor density
ATP	Adenosine tri-phosphate
ATRA	All-trans retinoic acid
BAL	Bronchoaveolar lavage
BCG	Bacillus Calmette–Guérin
bp	Base pair
BSA	Bovine serum albumin
C/EBPα	CCAAT/enhancer binding protein-α
CD	Cluster of differentiation
CFU	Colony forming unit
CO_2	Carbon dioxide
CR	Complement receptor
CRISP	Cysteine-rich secretory protein
CXCR	Chemokine receptor
DAF	Decay accelerating factor
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid

DosR	Dormancy survival regulator
ECF	Extra cytoplasmic function
Egr1	Early growth response protein-1
ESAT-6	Early secretory antigenic target-6
FBS	Fetal bovine serum
FcγR	Fragment c gamma receptor
g	Gram
G-CSF	Granulocyte colony stimulating factor
Gfi-1	Growth factor independent -1
GM-CSF	Granulocyte monocyte colony stimulating factor
Gp	Glycoprotein
Groβ	Growth related gene product-β
h	Hour
H_2O_2	Hydrogen peroxide
HBSS	Hank's balanced salt solution
HL60	Human leukocyte 60
HOCl	Hypochlorous acid
Ig	Immunoglobulin
IL	Interleukin
iNOS	Inducible nitric oxide synthase
INT	2-4-iodophenyl-3-4-nitrophenyl-5-phenyl-2H-tetrazolium chloride
IS110	*Insertion sequence-110*

ITAMS	Immunoreceptor tyrosine-based activation motif
kb	Kilo base
KD	Kilo Dalton
LAMPs	Lysosome-associated membrane proteins
LIR	Leukocyte immunoglobulin-like receptor
LL-37	Cathelicidin-related antimicrobial peptides
LPS	Lipopolysaccharide
LTBI	Latent tuberculosis infection
Mac-1	Macrophage-1 antigen
MHC-II	Major histocompatibility complex-II
Min	Minute
miscRNA	Miscellaneous RNA
Ml	Millilitre
MMPs	Matrix metaloproteinases
MOI	Multiplicity of infection
MPO	Myeloperoxidase
mRNA	Massenger RNA
Mφ	Macrophage
n	Number of replicates
NA	Not applicable
NADPH	Nicotinamide adenine dinucleotide phosphate
ncRNA	None coding RNA

NE	Neutrophil elastase
NETs	Neutrophil extracellular traps
NFκB	Nuclear factor kappa B
Ng	Nano gram
NO	Nitric oxide
NO_2	Nitric dioxide
NRP.	None replicating persistence
O_2	Oxygen
OADC	Oleic acid-albumin dextrose catalase
OD	Optical density
OH^o	Hydroxyl
ONOO-	Peroxynitrite
PGRS	Glycine-rich proteins
PI3	Phosphoinositide-3
PMN	Polymorphonuclear
ppGpp	guanosine pentaphosphate
PU.1	Protein-rich unit-1
qPCR	Quantitative real-time polymerase chain reaction
RBC	Red blood cell
RBS	Ribosomal binding site
RNA-seq	RNA sequencing
ROS	Reactive oxygen species

RPMI	Rose park memorial institute
rRNA	Ribosomal RNA
SD	Standard deviation
SDF-1	Stromal cell-derived factor-1
SDS	Sodium dodecyle sulphate
Sec	Second
sigK	Sigma factor k
SLPI	Secretory leukocyte protease inhibitor
SP-A	Surfactant protein-A
SP-D	Surfactant protein-D
STAT3	Signal transducer and activator of transcription-3
TB	Tuberculosis
TLRs	Toll like receptors
TPA	12-O-Tetradecanoylphorbol-13-acetate
tRNA	Transfer RNA
TSP	Transcription start point
TFOE	Transcription factor overexpression
uPA	Urokinase plasminogen activator
v/v	Volume to volume
w/v	Weight to volume
µm	Micrometre

List of Figures

List of tables

Table of Contents

CHAPTER ONE

Introduction

1.1. General Introduction

Tuberculosis (TB) is a worldwide health problem with about 1.5 million deaths per year. *Mycobacterium tuberculosis* (Mtb) is the principal causative agent of TB and infects one third of world's population (Zumla *et al.*, 2015). Primarily, the transmission of infection occurs via the respiratory route by inhaling Mtb-containing droplets (Chao *et al.*, 2010).

A microarray study of the Mtb transcriptome in sputum revealed a pattern similar to non-replicating persistence described in *in vitro* growing Mtb (Garton *et al.*, 2008). Attempts were then made in this lab to identify the environmental signals stimulating the sputum transcriptional pattern with limited success. A study by Eum and colleagues found neutrophil leukocytes to be dominant in sputum of TB patients (Eum *et al.*, 2010) thus the Mtb-neutrophil interaction was recognised as a possible candidate providing a stimulus that might contribute to the sputum transcriptome. This provides the basis for the initial hypothesis of the present study.

It was further recognised that there have been no published reports on the transcriptional responses of Mtb to the neutrophil encounter. In fact, it can be appreciated that investigating the interaction between mycobacteria and polymorphs at transcriptome level is of potential general interest. In view of these points, both qPCR and RNA-seq have been deployed to address the central question of this thesis.

1.2. Tuberculosis

TB is still one of the major infectious diseases that annually cause nine million new cases (Zumla *et al.*, 2015). An estimated one third of the world's population is asymptomatically infected with Mtb (Xu et al., 2010). As a result of reactivation or re-infection of initial infection, about one in ten infected individuals will manifest the symptoms of active tuberculosis (Soualhine et al., 2007). The most affected areas

in the world are developing and poor nations due to a variety of factors including the AIDS epidemic, variable efficacy of vaccination with BCG and increasing figure of multi drug resistant Mtb (Nascimento *et al.*, 2005).

1.3. *Mycobacterium tuberculosis* and the *M. tuberculosis* complex (MTBC)

Mtb, originally named *Bacterium tuberculosis,* was discovered in 1882 by Robert Koch. In the classification of mycobacteria proposed in 1986, it was placed in the genus *Mycobacterium,* family Mycobacteriaceae , order Actinomycetales, class Actinomycetes (Shinnick *et al.*, 1994). It is classified as a slow–growing bacillus with a doubling time of 18 – 24 hours under optimal nutrient availability and normal oxygen tension at 37 ^{o}C, producing white to light yellow colonies on agar in 3 – 4 weeks (Gengenbacher *et al.*, 2012).

Mtb bacilli are referred to as acid-fast bacteria. The acid-fastness is due to the unusual composition of cell wall consisting of lipids such as mycolic acid (Skvortsov *et al.*, 2012). For that reason, special staining methods are used to visualise the bacteria by microscopy. Thus, staining with Ziehl–Neelsen stain, Mtb appears as rod shaped red bacilli. The bacilli grow under aerobic condition but can survive in a micro-aerophilic environment for a long time (Starck *et al.*, 2004). Mtb is a non-motile, non-spore forming microorganism with the capacity to develop dormant state in which the metabolic activity is reduced. The GC-rich genome of Mtb is about 4.4 million base pairs encoding 4000 predicted proteins (Gengenbacher *et al.*, 2012)

Bacillus Calmette–Guerin (BCG) is a live attenuated vaccine strain of *Mycobacterium bovis* used in the protection against tuberculosis (Soualhine *et al.*, 2007, Morel *et al.*, 2008). It was derived in 1921 from *M. bovis*, a mycobacterium causing TB in cattle (Kemp *et al.*, 2005), and the only vaccine available against TB since its first administration in 1948 to over three billion individuals (Hilda *et al.*, 2012).

Countless efforts have been undertaken by the major BCG manufacturing labs (e.g. Pasteur, Moreau and Glaxo) to cultivate the bacilli on Dubos solid or pellicle Sauton potato media under highly standardised methods of preparation and limited number of passages. Nonetheless, subjected to the various conditions since its distribution by the Pasteur Institute, different BCG strains have presented high genetic variability and immunogenicity. Such variability in genotype and phenotype of BCG and its diversity in immunological properties have occurred following various protocols of cultivation and storage processes performed by research institutes throughout decades of unlimited sub-culturing of BCG in dispersed liquid cultures (Nascimento et al., 2005).

Due to its lowered virulence, BCG has been used as a model for Mtb in a variety of experimental settings and in many biomedical studies. Indeed, there is a 99.95% sequence identity between *M. bovis* and Mtb. This sequence similarity is also high between Mtb and BCG and they share more than 99.90% genome sequence identity (Carroll *et al.*, 2009). During macrophage infection, BCG and Mtb reside in the phagosomal spaces of macrophages and are able to block the acidification and functional maturation of phagosome by preventing phagosomal membrane from acquisition of ATPase proton pump and inhibiting fusion with hydrolytic lysosomes (Butler *et al.*, 2010). However, within vacuoles BCG bacilli are not able to progress from fused phaglysosomes and remain bound to membranes (Bange *et al.*, 1996).

1.3.1. Pathogenesis of tuberculosis

After inhalation Mtb bacilli enter the host via the mucosal surface of the airways where they encounter and are then internalised by alveolar macrophages (Denis, 1991), polymorphs and dendritic cells (Hedlund *et al.*, 2010). Primarily, Mtb is an intracellular pathogen (Gupta *et al.*, 2012). The internalisation occurs through interaction of bacteria with Fc-γ receptor and complement receptor *in vitro* (Lamhamedi et al., 1999).

Bacteria inside phagocytes enter to the phagosome which normally fuses with lysosome to form phago-lysosome inside which the ingested bacteria are exposed to reactive oxygen and nitrogen species (Skvortsov *et al.*, 2012). Mtb has the ability to survive and multiply in macrophages by inhibiting the fusion of lysosome and phagosome predominantly via suppression of calcium signalling (Dale *et al.*, 2008). Indeed, the ability of slow-growing Mtb bacilli in blocking the maturation of phagosome at early stage of phagocytosis is strongly associated with their success in intracellular survival (Schuller *et al.*, 2001).

The infected macrophages invade the epithelial layer leading to a local inflammatory response followed by recruitment of mononuclear cells from neighbouring blood vessels that provide fresh host cells for the growing population of bacteria (Russell *et al.*, 2010).

Primary TB occurs within a few months of the initial infection as a result of local bacterial growth and dissemination in the respiratory tract and blood circulation. The bacilli can reach various organs and tissues following spread through blood. Secondary or post primary TB commonly develops within two years of infection but may develop many years later as a result of bacillary reactivation often associated with immune defects. The active post primary course of infection has characteristic features including severe destruction of lung with cavitation, involvement of upper lobe and frequently sputum smear positivity. However, such outcomes are not exclusive to post primary TB. The major transmitters of infection are patients that have cavitary lesions whose granulomas have been broken through to the airway (Gideon *et al.*, 2011).

1.3.2. Latent tuberculosis

The spectrum of TB outcomes ranges from acute TB with clinical presentation of coughing and fever, sub-clinical TB and latent TB (Figure 1.1). Although, there is a strong antigen specific immune response after initiation of infection, the ability to

respond becomes markedly reduced in the infected tissues. Following accumulation in tissues the bacilli create an immunologically privileged niche where they can reside and induce a potent immune response. Such extensive immune induction causes tissue damage which is necessary for transmission (Cooper *et al.*, 2012).

In further progress of disease the granulomas are formed resulting from immune response. Although they hinder the occurrence of systemic infection, granulomas promote bacillary proliferation by protecting them from immune system attacks. Since, the pathogen cannot be entirely eliminated from host by protective mechanisms of immune system, the infection remains for longer times while host's

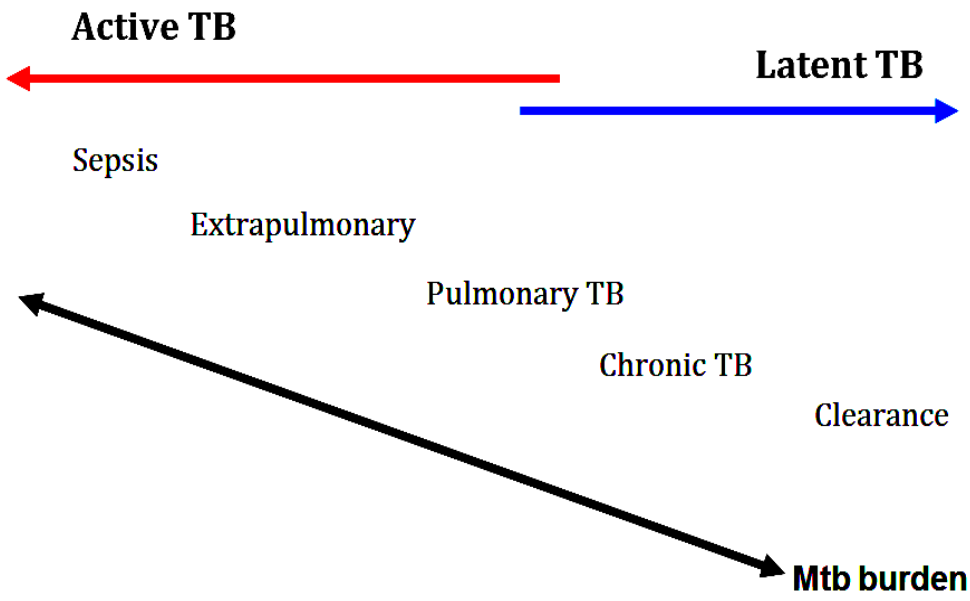


Figure 1.1 Outcomes of Mtb infection. The variability of infection is depicted for subdivisions of clinical outcomes of active (red line) and latent (blue line) infections. The black line indicates for increasing bacterial burden from latent to active TB. Figure redrawn after (Lin *et al.*, 2010).

health is not apparently affected. At this stage it is so – called the state of latent infection which can continue for several years. The infection can be reactivated following reduced immune response leading to active TB (Skvortsov *et al.*, 2012).

Various counterstrategies are employed by Mtb to respond to nutrient deprivation and hypoxia within granulomas (Dorhoi *et al.*, 2011) as well as to detoxify macrophage generated reactive oxygen and nitrogen species (Russell *et al.*, 2010). One approach in response to stresses is that the pathogen adopts a dormant lifestyle via altering gene expression pattern attributed to the non-replicative state called dormancy (Betts *et al.*, 2002). The dormant bacilli are able to survive for decades until the immune system of host is weakened then can develop active TB (Daniel *et al.*, 2004). While discontinuing their replication the bacilli remain viable in the course of latent infection as experiments with RNA Polymerase inhibitors and Rifampicin treatment found to be effective during latency suggesting for the presence of metabolically active dormant cells (Chao *et al.*, 2010). The ability to reside in the asymptomatic host is an indispensable feature that enables Mtb to survive and succeed in the human population (Bartek *et al.*, 2009).

1.4. The genome and gene expression in Mtb

Mtb possesses the largest genome among intracellular bacteria and obligate pathogens of human beings (Rodrigue *et al.*, 2006). The whole chromosome of Mtb (Type strain H37Rv) has 4,411,529 bp of DNA sequence (Figure 1.2) which encodes for 4,018 protein genes, 13 pseudogenes, 45 tRNA genes, 3 rRNA genes, 30 ncRNA genes, 2 miscRNA genes. The protein coding are 4,027,296 bases (about 91.2% of whole chromosome sequence). The gene density is 0.91 genes per kb with average length of 1,002 bases per gene. The genome is a high GC-content amount for 65.9%. (Data collected from Tuberculist version 2.6 release 27 (http://genolist.pasteur.fr/TubercuList).

1.4.1. Regulatory networks

Mtb encounters a broad range of environmental conditions to which it should be able to successfully respond by regulation of gene expression. The distinctive feature of successful Mtb strains is their ability to swiftly adapt to the environmental changes. This is mainly done by rapid alteration in metabolic pathways to acclimatise to the environmental cues through transcriptional regulation (Wang *et al.*, 2011a). The level of expression for a gene is determined by the promoter

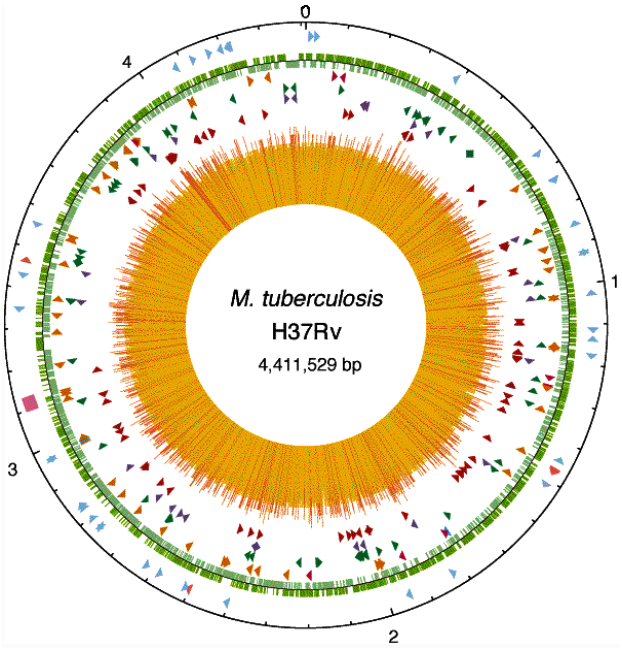


Figure 1.2 Circular map of Mtb H37Rv genome. The outer circle represents the scale in Mb starting with 0 as the origin of replication. The positions of stable RNA genes are denoted by the first ring from exterior (blue for tRNA and pink for others with pink cubes as direct repeat region). Inwards, the second ring displays the coding sequence by strand with dark green as clockwise and light green as anticlockwise. The third ring shows repetitive DNA; orange for insertion sequences, and blue for prophage. The position of the members of the PPE family is shown in the fourth ring (green) and PE family members excluding PGRS are in fifth ring (purple). The sixth ring in dark red depicts PGRS sequences. The center of histogram represents GC content in yellow. Figure adapted from (Cole *et al.*, 1998).

recognised as the sequence of DNA between 10 and 35 bases located upstream of transcription start site where RNA Polymerase binds during transcription (Newton-Foot *et al.*, 2013).

In bacteria, an RNA polymerase holoenzyme mediates the transcription of mRNA from DNA. This enzyme consists of five subunits namely α_2 β β′ ω σ. Prior to initiation of transcription the core complex of RNA polymerase α_2 β β′ ω interacts sigma factor (subunit σ). The DNA promoter sequences are recognised then engaged with one of several sigma factors resulting in a correct positioning of RNA polymerase on promoter (Figure 1.3). This in turn facilitates DNA unwinding followed by RNA polymerisation mediated by the α_2 β β′ ω core complex. The initiation of transcription by all sigma factors occurs via binding to the core RNA polymerase but their recognition of consensus sequences of promoter is different. Therefore, sigma factors are responsible for RNA polymerase specificity to target promoter. Eubacteria have two principal sigma factor families; the RpoN-like σ^{54} family and the RpoD-like σ^{70} family (Wösten, 1998). Only σ^{70} family is present in mycobacteria (Rodrigue *et al.*, 2006).

Based on gene function and its structure, the σ^{70} family is divided into four main phylogenetic groups. Group 1 includes the primary σ factors that are essential for bacterial growth. Group 2 consists of proteins that are closely related to group 1 but not necessary for growth. Group 3 are factors that are often involved in regulon activation in response to environmental signals such as heat shock. Group 3 have more distant relation to σ^{70} and further divided into different functionally related clusters of proteins. Finally, group 4 which encompass the largest number of factors and comprise a highly diverse subfamily of extra-cytoplasmic function (ECF). The majority of the ECF subfamily responds to extra-cytoplasmic stimuli (Paget *et al.*, 2003).

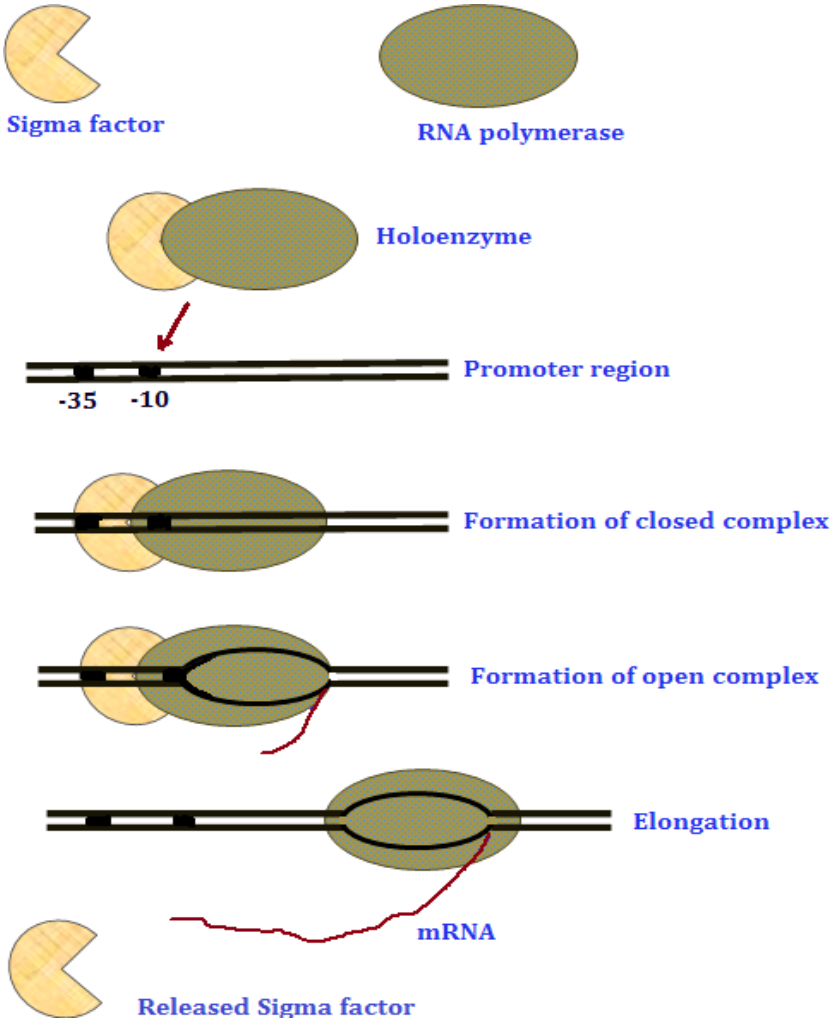


Figure 1.3 The steps for initiation of transcription of RNAP containing σ_{70} or σ_{70} related sigma factor. Binding of RNAP holoenzyme to the promoter forms a closed complex leading to a conformational change that creates an open form at 310 promoter region. RNAP holoenzyme in this open form initiates small RNA molecule production which leads to formation of a ternary complex. The final step is transcription of mRNA molecule by RNAP from which σ factor is dissociated. Figure modified from (Wösten, 1998).

The Mtb genome encodes 13 sigma factors and has an alternative sigma factor

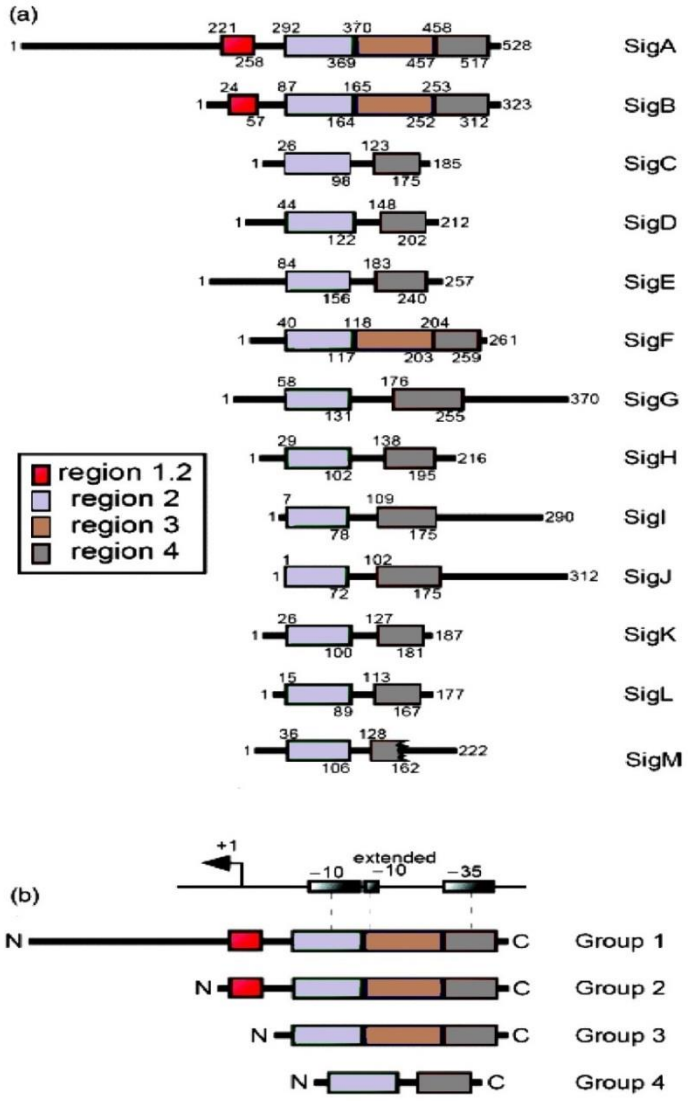


Figure 1.4 Conserved regions in σ factors. (a). The architecture of conserved regions of sigma factors of Mtb H37Rv. The residues at the beginning and end of each conserved region are presented in the top and bottom of colored boxes by numbers. The number of last C-terminus residue is indicated on the right side of each sigma factor and the amino-terminus of each protein is shown as '1' **(b)** The structural organisation is shown for group 1, 2, 3 and 4 sigma factors. Broken lines connect sigma factor regions to their recognised promoter element. The arrow shows transcription start site. Reproduced after (Rodrigue *et al.*, 2006).

density (ASFD) of 2.8 making it the obligate human pathogen with the highest

ASFD (Figure 1.4). These features probably reflect the complicated environmental adaptation strategies and complex growth cycle of Mtb. Because the frequency of ASFD in a genome is a measure of the physiological properties of a microorganism (Rodrigue *et al.*, 2006). Out of 13 sigma factors only σA is essential: it is the primary sigma factor and belongs to group 1. σB is a primary-like sigma factor belonging to group 2. σF is member of group 3 and other sigma factors (σC, σD, σE, σG, σH, σI, σJ, σK, σL and σM) belong to (ECF) subfamily of group 4 (Manganelli, 2014).

The regulation of sigma factors is modulated by anti-sigma factors. These are unconventional transcription factors that bind to specific regions of sigma factor preventing it from binding to promoter DNA or core enzyme of RNA polymerase resulting in inhibition of transcription initiation. Mtb has five anti-sigma factors regulating the activity of σE, σF, σH, σK and σL. The regulation of anti-sigma factors is done in 3 different ways; interaction with extra-cytoplasmic proteins or small effector molecules, sequestration by anti-anti-sigma factor, or secretions from the microorganism (Paget *et al.*, 2003).. The anti-anti-sigma factors further regulate sigma-factors through inhibiting the activity of anti-sigma factors. There are 7 anti-anti-sigma factors in Mtb (Newton-Foot *et al.*, 2013).

1.4.2. Structure of the promoter in Mtb

The same components of classical bacterial promoters are generally present in mycobacteria and they may be intergenic or located within unannotated regions (Dieci *et al.*, 2002).

Transcription initiation start point (TSP)

Transcription starts at a purine base (most often guanine) in the majority of mycobacterial promoters though any base may occur occasionally in the +1 position as TSP in mycobacteria (Bashyam *et al.*, 1996). The position of the majority of TSPs is downstream 5 to 9 bases of the -10 sequence. However, in

some Mtb promoters such as *rel* transcription begins at base 3 of the -10 region and this has implications for transcription because it alters the conformation of the DNA-RNA polymerase complex (Jain *et al.*, 2005).

−10 promoter region

The aligning of sequences for a number of mycobacterial promoters revealed a conserved hexameric region located upstream of the TSP. This region is named −10 promoter region because it is positioned about 10 bases upstream of TSP that corresponds to the Pribnow box in classical bacterial promoters. Indeed, the essentiality of this sequence for transcription is preserved through all bacteria as chloramphenicol resistance tests demonstrated that the −10 region alone was sufficient for transcription whilst its deletion in Mtb H37Rv resulted in no significant transcription (Bashyam *et al.*, 1996).

−35 promoter sequence

The mycobacterial −35 promoter sequence has little similarity to the TTGACA motif in other bacteria. Moreover, not much homology of this sequence occurs within mycobacteria. The −35 sequence is not essential for promoter function as it does not support transcription on its own (Kenney *et al.*, 1996). There is a disparity in the sequence of −35 hexamer causing variation in activity. These diverse −35 promoter sequences are recognised by different sigma factors and additional transcriptional activators are required to allow transcription initiation for promoters with unrecognised −35 motifs (Bashyam et al, 1996).

Intrahexameric region

The region between −35 and −10 in Mtb seems to be conserved at approximately 16 – 19 bp. This distance has no significant transcriptional role in mycobacteria as

variation in this region appears to have no major effects on the strength of promoter (Kremer *et al.*, 1995).

Ribosomal binding site (RBS)

The RBS in mycobacteria is about 6 – 8 bp downstream to TSP and commonly an adenine and guanine rich region. Although, RBS is typically located upstream to start codon, for several genes encoded by leaderless RNA the start codon and TSP are located at the same position. In view of that, the RBS and other transcriptional signals are absent on the mRNA transcript (Bashyam *et al.*, 1996).

Start codon

The position on mRNA where translation into a peptide starts is the start codon. In mycobacteria, similar to most bacteria, ATG is the most dominant start codon from which methionine is translated. Alternative start codons however can occur as about 30% of proteins of Mtb are translated from GTG start codon. This may reflect high GC content of mycobacterial genome (Lew *et al.*, 2011).

Length of promoter

Generally, the minimum length of promoter is considered to be about 40 bp that includes the −10 and −35 hexamers and TSP. While for transcription initiation it seems to be the basic requirement, other components like RBS and DNA binding regulator motifs are necessary for optimal expression. The distance between the start codon and RBS can be about 40 bp and between RBS to the TSP up to 1000 base pairs. A required length of DNA necessary for optimal expression of a polypeptide can be highly variable that can reach up to 2000 bp upstream of the gene (Hunt *et al.*, 2012).

Transcription termination

The two transcription termination mechanisms in bacteria are factor-dependent and intrinsic termination. In factor-dependent the binding of proteins to RNA polymerase dissociates it from DNA, whereas, intrinsic termination is the result of specific feature of mRNA that cause transcription termination. In typical bacteria, the intrinsic terminator has a CG-rich palindromic sequence which is followed by a poly adenine sequence. Following transcription of this region a secondary hairpin structure forms and is continued with a uracil-tail. Formation of hairpin results in RNA polymerase dissociation and consequently termination of transcription (Unniraman *et al.*, 2002).

The putative terminators in the Mtb genome are positioned within 50 base pairs of the stop codon. However, majority of terminators consist of a hairpin structure but lack the uracil tail which might again reflect the high GC content. Indeed, high GC may provide an alternate termination mechanism and a less complicated terminator structure compared to other well-studied bacteria (Mitra *et al.*, 2008).

Expression from multiple promoters

The majority of genes in mycobacteria are expressed from a single promoter upstream of the gene within the intergenic region. However, under different conditions the multiple promoters can express the same gene. The multiple promoter system may provide the ability for a gene to be differentially expressed in basal and inducing conditions as well as for transcription in a reverse orientation. Therefore, multiple promoters contribute to regulating the expression of mycobacterial genes in a variety of conditions in order to optimize the transcriptional pattern in the cell (Gopaul *et al.*, 2003).

Expression from insertion sequences

Insertion sequences are compact, small DNA molecules that are flanked by 10 – 40 base pair terminal inverted repeat sequences. They are mobile DNA elements

that encode factors responsible for their own transposition (Mahillon *et al.*, 1998). The insertion sequence *IS110* is only found in the Mtb complex and is involved in its evolution. It has the ability to transpose to different parts of the genome. Depending on the insertion site this may lead to altered or loss of specific gene expression. OP6110 is an example of insertion sequence mediated up-regulation in the late phase of H37Rv growth in human macrophages. The insertion leads to up-regulation of some downstream genes including *Rv2080*. In this case, the presence of a mobile genetic element capable of transcriptional regulation in response to macrophage environment appears to be essential for the intracellular survival of Mtb (Brosch *et al.*, 2001).

Regulation of transcription by inducible promoters

In promoter region the induction and repression of transcription can be through the binding of sigma factors as well as activator and repressor proteins to specific DNA sequences. Alternatively other transcription factors such as anti-sigma factors and ppGpp that bind to RNA polymerase can influence the regulation of transcription. Using these mechanisms the mycobacteria can adapt to diverse conditions and survive (Haugen *et al.*, 2008).

1.5. Human pulmonary immunity against Mtb

The success or failure of protection against Mtb within lung tissue and bronchoalvelar cells hinges on a collective contribution of innate, adaptive and regulatory mechanisms of immunity in the bronchoalveolar microenvironment. The lung immune response to Mtb exposure characterizes the outcomes of early interaction between host and pathogen that may result in eliminating the pathogen, developing latent infection or progressing to active TB.

Table 1.1 Innate immunity: Cell types, Receptors and effector mechanisms

---

Receptors involved in Mtb uptake and signaling

- **Complement receptor-3 (CR3)**
- **C-type lectin Dectin-1**
- **Fc receptors**
- **Scavenger receptors**
- **Chemokine receptors**
- **Mannose receptors**
- **DC-SIGN**
- **Adenosine receptor**
- **Toll-like receptors (TLRs) 2, 4, and 9**
- **Nucleotide oligomerisation domains (NODs)**

TLR-mediated effector mechanisms

- **NF-κB activation**
- **Antimicrobial peptides**
- **iNOS, nitric oxide**
- **proinflammatory cytokines**
- **IL-12, IL-18**

Mtb TLR agonists

- **Lipoarabinomannan (TLR2)**
- **Heat shock proteins 65 and 71 (TLR2, TLR4)**
- **Mtb DNA (TLR9)**

Lung collectin-mediated host mechanisms

- **NF-κB activation**
- **Expression of TLR2 and TLR4 (surfactant protein A [SP-A])**
- **Increase in Mtb adherence and phagocytosis (SP-A)**
- **Decrease in phagocytosis (SP-D)**
- **Increased expression of phagocytic receptors, scavenger receptor A, and mannose receptor**

Definition of abbreviations: DC-SIGN = dendritic cell-specific intracellular adhesion molecule-3-grabbing non-integrin; iNOS = inducible nitric oxide synthase; NF-κB = nuclear factor-κB. Table reproduced after (Bermudez *et al.*, 1996)

Transmission Dynamics and Host Immunity

The majority of Mtb infections occur through inhalation of aerosolised droplets released from individuals with active pulmonary TB. Rarely, exposure to laboratory generated aerosols and contact with infected animals can also transmit the pathogen. The transmission of Mtb in aerosol depends on several factors including the duration of exposure and shared air space, distance from source, the amount of bacteria in the source TB case's respiratory secretions, and the immune status of the new host. Other factors that can also affect the efficiency of transmission are the amount and rheological characteristics of sputum in addition to the strength of cough (Fennelly *et al.*, 2004).

A wide range of pattern recognition receptors interact with Mtb on inhalation into the pulmonary area (Table 1.1). In the lungs, Mtb bacilli interact with alveolar type I and II epithelial cells and are ingested by alveolar macrophages and dendritic cells (Bermudez *et al.*, 1996).

During interaction, the Mtb products can activate neutrophils that are the dominant phagocytes among bronchoalveolar cells in active TB. Such neutrophil interactions results in expression of a variety of receptors and production of antibacterial effector molecules (Neufert *et al.*, 2001). Moreover, the activated polymorphs have the capability of restricting the growth of Mtb in vitro (Martineau et al, 2007) and can contribute in activation of macrophage in vitro (Tan *et al.*, 2006).

The early initiation of innate immunity against TB involves active contribution of dendritic cells and subsets of bronchoalveolar cells such as αβ and γδ T cells, natural killer cells in addition to active soluble molecules including nitric oxide, surfactants and antimicrobial peptides. Two groups of pattern recognition receptors can be activated by Mtb; the Toll-like receptors (TLRs) and the nucleotide oligomerisation domain receptors (Liu *et al.*, 2008). Depending on the type of antigen, cytokine and inflammatory cells, the TLR engagement can be stimulatory or suppressive to the immune response. The immune-stimulatory TLR2, TLR4 and

TLR9 on macrophage and dendritic cells are essential in innate resistance against mycobacteria (Pompei *et al.*, 2007). As a result chemokine are released and different immune cells are accumulated leading to production of antimicrobial effector molecules (Korbel *et al.*, 2008). However, potent immunosuppressive responses can be mediated by TLR engagement resulting in evasion of Mtb from host immunity. On example can be the anchoring of mycobacterial glycosyl phosphatidyl inositol to phosphatidyl myoinositol hexamonosides that blocks the binding of myeloid differentiation protein 88 and TLR4. As a result, the release of nitric oxide and cytokines such as TNF-α and IL-12 is strongly inhibited (Doz *et al.*, 2009). In a similar way, the activation of NFκB and interferon regulatory factors are inhibited by the early secreted antigenic target 6-KD protein (ESAT-6) after TLR2 – ESAT-6 binding.

The TLR2 – ESAT-6 interaction can suppress IFN-γ induced expression of class II major histocompatibility complex (MHC-II) via TLRs mediated mechanism (Pathak *et al.*, 2007). The bronchoaleveolar space is a unique immunological compartment in which the expression of TLR2 is low; TLR4 at similar levels and TLR9 is higher in alveolar macrophages compared to autologous macrophages. In view of that, there is a difference between alveolar and peripheral macrophages in their interaction with surfactant proteins A and D (SP-A and SP-D) which has a significant effect on lung immune responses to Mtb. The expression and signaling of TLR2 and TLR4 in macrophages is regulated by SP-A resulting in decreased phosphorylation of NFκB activity leading to reduced TNF-α secretion. SP-A could increase phagocytosis of Mtb by human alveolar macrophages via enhancing the adherence of bacilli following up regulation of mannose receptor activity on alveolar macrophages (Gaynor *et al.*, 1995). In contrast, SP-D decreases the phagocytic activity of macrophages partly due to a direct Mtb SP-D interaction that limits the intracellular replication of Mtb via enabling phagosome–lysosome fusion (Ferguson *et al.*, 2006).

1.6. Sputum transcriptome of Mtb

An effective anti-mycobacterial immune response able to control the growth of bacteria causes localisation of bacilli in the granulomas and nodules in the lungs. Within the caseous portion of granuloma, a state of non-replicating persistent is likely due to the effects of hypoxic conditions. Recruitment of O_2 even after a prolonged hypoxia–induced dormancy can initiate bacterial resuscitation (Voskuil *et al.*, 2003).

In the Mtb infected microenvironment of the pre-caseous granuloma multiple stimuli are likely variably present in addition to hypoxia including nitric oxide, carbon dioxide, oxygen radicals, low pH and restricted nutrients. In such a diverse set of microenvironments, it can be expected that Mtb is in different growth states including a dormant/replicating equilibrium (Gengenbacher *et al.*, 2012). Thus, during an active TB a mixture of dormant, resuscitating and replicating bacteria are present (Chao *et al.*, 2010).

Among heterogeneous populations of mycobacteria in human pulmonary infections, individual groups can be dominant or disappear depending on their ability to adapt in the presence of host immune response (Bukka *et al.*, 2011). Host immune status ultimately allows replication of bacteria to resume and ultimately leads to the emptying liquefied caseating granuloma contents into airways (Dorhoi *et al.*, 2011). Based on recent studies prior to transmission the final burst in mycobacterial replication occurs upon release of bacilli from granuloma into the sputum (Eum *et al.*, 2010).

The phenotype and staining properties of isolated Mtb bacilli are different from *in vitro* growing bacteria (Betts *et al.*, 2002). In this regard, visualisation of sputum samples from active TB patients revealed characteristics of dormant bacilli containing lipid bodies similar to non-replicating persister Mtb (Garton *et al.*, 2008, Gengenbacher *et al.*, 2012).

The presence of lipid bodies in sputum population of Mtb was demonstrated in parallel with a transcriptome study (Garton *et al.*, 2008). The pattern of gene expression revealed was unique but most closely resembled non-replicating persistent Mtb in vitro. Looking into functional categories, the expression of genes involved in ribosomal function and aerobic respiration were found to be significantly reduced compared to aerobic growth in vitro. Whereas, the genes required for cholesterol utilisation and fatty acid metabolism had a significant rise in their transcript. Notably mentioned in the same study, similarities in transcript abundance were observed between sputum transcriptome and a model of chronic murine infection. Thus, genes repressed during bacillary stasis in murine model including *ctaD*, *nuoB*, *atpA*, *atpD*, and *qcrC* were also found to be repressed in sputum. Likewise, *narK2* was induced as was the case in murine infection (Garton *et al.*, 2008).

The most prominent pattern of expression observed in sputum was related to the DosR regulon This is interesting as activated DosR has been associated with long term survival of bacteria during anaerobiosis (Honaker *et al.*, 2009). DosR is also essential in bacterial recovery in conditions where aerobic respiration is limited (Leistikow *et al.*, 2010). Up regulation of the isocitrate lyase (*icl1*) gene is an indication that lipids are utilised as a carbon and energy source which has been consistent with the findings in an in vivo study. Moreover, evidence of an activated KstR regulon in pulmonary exudates indicates signals related to cholesterol utilization and occurrence of this type of lipid in sputum (McKinney *et al.*, 2000).

The patterns seen in sputum transcriptome such as utilization of lipids, activation of DosR, and a signature of slow growth are similar to many experimental conditions that have previously been studied in macrophage and animal studies (Garton *et al.*, 2008).

1.7. Neutrophils

Neutrophils are one of key elements of immunity against invading pathogens. High numbers are generated in the bone marrow and they comprise ~60% of all white blood cells in circulation in human (Dale *et al.*, 2008). Circulating neutrophils have a half-life of 6 to 10 hour before their spontaneous apoptosis. Their lifespan can be prolonged by pro-inflammatory signals in inflamed tissues (Corleis *et al.*, 2012). When the skin and mucus membrane fail to prevent microorganisms from entering into the tissues, the local endothelial cells become activated by signals generated from microbes and resident macrophages. Following the activation of endothelial cells at the site of infection passing neutrophils are captured and guided towards the microbes. The neutrophil–microorganism encounter results in phagocytosis of microorganism followed by production of bactericidal substances in the phagocytic vacuole. The products of neutrophils in tissues attract further neutrophils and other immune cells. Ultimately, neutrophils extracellular traps are produced by the neutrophil encounter with microbes (Summers *et al.*, 2010).

1.7.1. Production of Neutrophils

Neutrophil production is the dominant activity in myelopoisis (i.e. collective production of granulocytes and monocytes) within venous sinuses of bone marrow. Granulocytes and monocytes differentiate from a common progenitor cell. The regulation of granulocyte differentiation occurs via selective expression of a subset of transcription factors such as HoxB7, Egr1 and STAT3 as well as overexpression of receptors for granulocyte – macrophage colony stimulating factor (GM-CSF) and N-formylmethionylleucyl-phenylalanine (Sasmono *et al.*, 2007).

In a normal adult person the neutrophil production is extensive with $1 – 2 \times 10^{11}$ cells produced per day. Granulocyte colony stimulating factor (G-CSF) plays an important role in producing higher number of neutrophils during infections (Lieschke *et al.*, 1994). In bone marrow the population of neutrophils can be

classified into three subdivisions. 1) The stem cell pool consisting of undifferentiated hematopoietic stem cells. 2) Mitotic pool consisting of committed progenitor cells which are granulocytes with active proliferation and differentiation. 3) Post-mitotic pool consisting of fully differentiated mature granulocytes which are considered as the neutrophil storage of bone marrow ready for release. Based on studies the transit time from stem cell pool to post-mitotic pool is about 4 to 6 days (Summers *et al.*, 2010).

1.7.2. Terminal Granulocytopoiesis

Mature granulocytes are produced from committed progenitors in a process called terminal granuloctyopoiesis (Figure 1.5).

Two essential transcription factors are required for the commitment of myeloid lineage; these are PU.1 and CCAAT/enhancer binding protein-α (C/EBPα). The balance between PU.1 and C/EBPα determines the subsequent decision between monocyte and granulocyte commitment (Nerlov *et al.*, 1998, Reddy *et al.*, 2002). High expression of C/EBPα derives granulocytopoiesis (Zhang *et al.*, 1997) especially during fungal infections (Hirai *et al.*) whereas PU.1 induces monocytic differentiation (Radomska *et al.*, 1998). Further differentiation of neutrophils requires growth factor independent-1 (Gfi-1) a transcription factor that overexpresses during the commitment of stem cells to promyelocytes (Velu *et al.*, 2009). Beyond promyelocytes, a number of transcription factors are involved for further differentiation. One of the key transcription factors expressed at the stage of myelocyte is C/EBPε (Morosetti *et al.*, 1997). C/EBPε regulates the transition of promyelocyte into myelocyte which is associated with the termination of cell cycle. This is done via binding of Rb and E2F1 mediated by C/EBPε which inhibits the activation of genes associated with cell cycle (Gery *et al.*, 2004). During the transition from myelocytes into metamyelocytes C/EBPβ, C/EBPγ, C/EBPδ, and C/EBPζ are also expressed (Bjerregaard *et al.*, 2003). Granule proteins are highly expressed by C/EBPs during terminal granulocytopoiesis (Chumakov *et al.*, 2007).

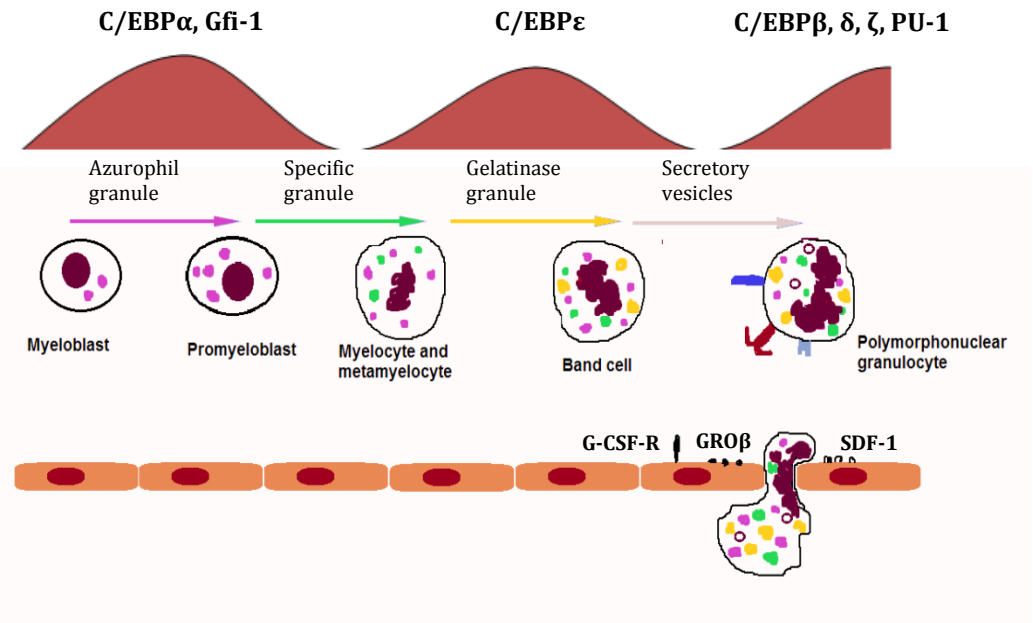


Figure 1.5 Granulocytopoiesis in the Bone Marrow. Osteoblast and endothelial cells provide localisation of stem cells. The transcription factors (top of figure) indicating for different subsets of granules in neutrophil; azurophilic (red), specific (green), gelatinas (yellow), secretory vesicles (empty) that are sequentially formed during promyelocyte maturation. Retention of cells is favored by CXCR4 and SDF-1 ligand whereas the release of cells is favored by CXCR2 and its ligands KC and Groβ. Release of neutrophil is directly stimulated by G-CSF via direct effect on neutrophil and indirect effect on decreasing SDF-1 while increasing Groβ expression on endothelial cells. Figure modified from (Borregaard, 2010).

1.7.3. Release of Neutrophils from Bone Marrow

From early developmental phases the G-CSF receptors are present at high levels on neutrophils surfaces. Correspondingly, there is low expression level of CXC chemokine receptor 4 (CXCR4) on mature neutrophil surfaces (Hübel *et al.*, 2002). Expression of CXCR4 is essential factor for neutrophils to retain in the bone marrow. Stromal-derived factor 1 (SDF-1), a CXC chemokine constitutively

produced by stromal cell of bone marrow is the main ligand for CXCR4. The CXCR4 and SDF-1 interaction can result in retaining of neutrophils in marrow (Hernandez *et al.*, 2003).

After maturation, neutrophils migrate from the haematopoietic compartment to circulation through sinusoidal endothelium by a unique process of trans-cellular migration. This occurs when CXCR4 is significantly reduced during differentiation allowing the mature neutrophils in bone marrow to move into circulation. However, the life span of circulating neutrophils is not affected by the release of mature neutrophils into circulation (Eash *et al.*, 2009).

1.7.4. Circulating and marginated pools of granulocytes

Based on a study done on healthy volunteers, roughly half of the neutrophils given by infusion disappear from circulation (Mauer *et al.*, 1960). Addition of adrenaline to the infusion caused an increase in the number of granulocytes remaining in the circulation. This recoverable number of neutrophils is named the marginated pool. The total granulocyte pool in blood is about 6.5×10^7 cells per kilogram body weight, of which 51% are in the marginated pool and 49% are in the circulating pool (Athens *et al.*, 1961).

1.7.5. The pulmonary marginated pool of granulocytes

The size of a marginated pool is determined by the passing time of neutrophil through capillary bed. There is controversy on the size of pulmonary marginated granulocyte pool. A number of studies suggest that the predominant site for neutrophil margination is the lung as the entire cardiac output passes through the lungs (Hogg *et al.*, 1995). Moreover, the maturation and activation status of neutrophils determine their biodistribution. Thus young circulating granulocytes home back to bone marrow and mature peripheral blood granulocytes mobilise

back to the bone marrow and liver whereas inflammatory mature neutrophils predominantly mobilise back to lungs and liver (Peters, 1998).

1.7.6. Neutrophil-Endothelial Cell Interactions

To arrive into the site of microbial entry, neutrophils should cross the vascular wall and this largely occurs in postcapillary venules (Figure 1.7). The neutrophils can easily come to contact with the walls of blood vessel as the diameter is small and

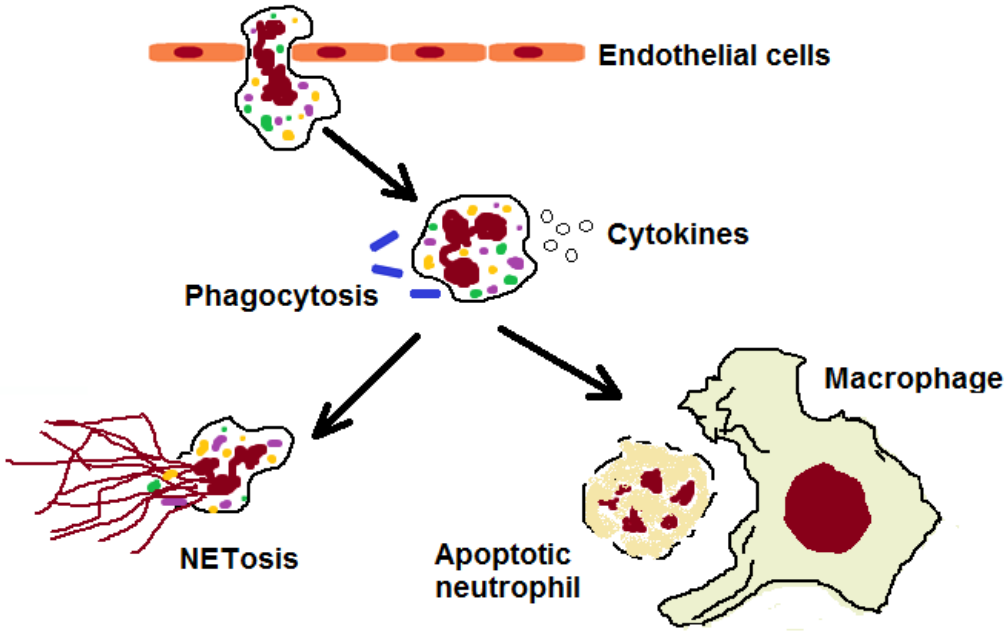


Figure 1.6 Neutrophils in Tissues. The endothelial cells P-selectins capture neutrophils. The rolling and activation of integrins on neutrophil is mediated by binding of endothelial cells to PSGL-1, L-selectin, and CD44 on neutrophil. Thus the cell interacts with expressed ICAMs on endothelial cells. The crossing of neutrophil through endothelial membrane can be paracellular or transcellular. Phagocytosis of microbes may lead to apoptosis or NETs formation. Figure redrawn after (Borregaard, 2010).

the wall is very thin; but neutrophils are too large to cross the endothelium. Nevertheless, leukocytes are recruited to the sites of infection in systemic organs via diapediesis. This mechanism of transendothelial migration is unique to neutrophils (Woodfin *et al.*, 2010).

1.8. Pathogen killing by neutrophils

Three major mechanisms are involved in the process of killing invading pathogens by neutrophils. Firstly, the pathogen is engulfed via receptor mediated uptake into the phagocytic vacuole. Secondly, the vacuole produces highly toxic reactive oxygen species (ROS). Finally, the neutrophil granules containing different antimicrobial elements are fused with the vacuole (Mayadas *et al.*, 2014).

1.8.1. Phagocytosis

As an immunological response, a massive recruitment of circulating neutrophils occurs during microbial challenges. At the sites of infection neutrophils play a key role in phagocytosis and killing of microorganisms. Two different classes of receptors are involved in the neutrophil phagocytosis; complement receptors including CR1 (CD34) and CR3 (CD11/CD18 integrin) and Fcγ receptors including FcγRIIA (CD32) and FcγRIIIB (CD16). However, the major functional receptors for phagocytosis are CR3 and FcγRIIA whose function is facilitated by co-receptors FcγRIIIB and CR1. These two receptor classes are different in their phagocytosis process as well as signalling pathway.

In the FcγRIIA receptors mediate phagocytosis and activated Src-tyrosine kinases phosphorylate immunoreceptor tyrosine-based activation motifs (ITAMS), the cytoplasmic motif of FcγRIIA receptor. The phosphorylation of ITAMS and aggregation of FcγRIIA receptor prompts the internalisation of IgG-coated targets. The engagement of phosphorylated ITAMS with Syk tyrosine kinase activates PI3-kinase and Rho proteins resulting in extension of membrane protrusions. This

serves as the phagocytic cup that engulfs the opsonised particle (Massol *et al.*, 1998). The proposed role of FcγRIII receptors in neutrophil phagocytosis is recruiting the FcγRIIA receptors resulting in aggregation of ITAMS (Chuang *et al.*, 2000).

In CR3 mediated phagocytosis the complement opsonised target is sunk into the neutrophil membrane producing minor protrusions. In addition to binding of CR1 and CR3 to C3b/iC3b the activation of neutrophils is required to achieve complete phagocytosis. Thus, the stimuli in activated neutrophils phosphorylate CR1 that promotes outward signals which in turn enhance the binding capacity of CR3. Unlike, FcγR-mediated phagocytosis, C3b dependent phagocytosis is not accompanied by the production of cytokines and arachidonic metabolites during activation of respiratory burst. Indeed the internalisation of opsonised targets promotes a complex signalling pathway leading to fusion of granules rich in protease within the phagosome that trigger the oxidative burst (Yamamoto *et al.*, 1984).

1.8.2. Granule biogenesis

Upon neutrophil activation the granules are filled with microbicidal molecules. The biogenesis of granule occurs parallel to differentiation pathway of granulocyte (Borregaard, 2010). At the promyelocyte stage the azurophilic granules emerge containing serine proteases, antibiotic proteins and myeloperoxidase (Fouret *et al.*, 1989). Therefore, mobilised upon phagocytosis, the azurophilic granules are considered as compartment with true microbicidal properties. Azurophilic granules are not classified as lysosomes due to the absence of lysosome-associated membrane proteins (LAMPs). Nonetheless, their functional characteristics resemble secretory granules (Cieutat *et al.*, 1998). At the metamyelocyte stage specific granules containing collagenase and lactoferrin appear followed by a population of gelatinase containing tertiary granules. Finally, at the stage of mature neutrophils the fourth type of granules appears; these are known as secretory

vesicles. Since plasma proteins such as albumin are present in secretary vesicle granules, they might be originated from the endocytic environment of neutrophils. The mechanism of secretion the four morphologically distinct granule populations may be separately controlled. The exocytosis resulting from a rise in cytosolic calcium occurs first in secretory vesicles followed by gelatinase granules then specific granules and finally azurophilic granules (Sengeløv *et al.*, 1993).

1.8.3. Granule proteins

Antimicrobial Proteins in granule are an array of peptides and proteins used by neutrophils to eliminate invading microbes (Table 1.2). The majority of antimicrobial proteins are in the azurophilic granule that should be delivered into the phagolysosome.

Bactericidal–permeability increasing protein (BPI) is stored in azurophilic granules and is one of the most active elements; it can also be expressed at the neutrophil plasma membrane (Elsbach, 1998).

Defensins, another group of antimicrobial peptides, are small cationic antibiotic peptides known as beta-sheet defensins. They induce membrane permeability in Gram-negative and Gram-positive bacteria. Defensins are suggested to have a role in wound healing as they are mitogenic for fibroblasts (Murphy *et al.*, 1993).

Proteins localised in the matrix and the membrane of neutrophil granules and secretory vesicles. The localisation is inferred from the gene expression profile according to the targeting-by-timing hypothesis but has not been confirmed at the protein level (Borregaard, 2010).

Lactoferrin is an iron-binding protein produced form proteolysis of longer proteins and also has antimicrobial effects (Hwang *et al.*, 1998). Other proteins in specific granules that may contribute to the bactericidal activity of neutrophils include lysozyme and phospholipase A2 (Harwig *et al.*, 1995).

Table 1.2 Neutrophil granule proteins

Membrane Proteins			
Azurophil granules	**Specific granules**	**Gelatinase granules**	**Secretory vesicles**
- - -	CD11b/Cd18, CD66, CD67 Gp91phox/p22phox	CD11b.CD18, CD67 Gp91phox/p22phox MMP25	CD11b.CD18, CD67 Gp91phox/p22phox MMP25
-	TNFR, uPAR	TNFR	LIR1-4, 6-7, 9; CD35; CD16; C1q-R; IFN-αR1 and IFN-αR2; IFN-γR1 and IFN-γR2; TNFR1 and TNFR2
CD63, CD68, presenilin	SNAP-23, VAMP-2, Stomatin, PGLYRP	SNAP-23, VAMP-2, Nramp1	SNAP-23, VAMP-2, Nramp1, alkaline phosphatase, DAF, CD10, CD13
Matrix Protein			
Elastase, Cathepsin G, Proteinase 3	Collagenase, Gelatinase, uPA, cystatin C, cystatin F	Gelatinase, Arginase1	Plasma proteins
Defensins, BPI, MPO, lysozyme	hCAP18, NGAL, B12BP, lysozyme, lactoferrin, haptoglobin, pentraxin 3, prodefensin	Lysozyme	-
Sialidase, Azurocidin, β-glucoronidase, azurocidin	A-1-anti-trypsin, SLPI, orosomucoid, heparanase, β2-microglobulin, CRISP3	β2-microglobulin, CRISP3	-

Abbreviations: B12BP, vitamin B12 binding protein; CRISP, cysteine-rich secretory protein; DAF, decay-accelerating factor; Gp, granule protein; LIR, immunoglobulin-like receptor; uPA, urokinase plasminogen activator. Reproduced after (Elsbach, 1998)

Secretory leuko-proteinase inhibitor (SLPI) is a production of mucosal surface cells and is present in neutrophils. It is a two domain polypeptide that expresses anti-proteinase activity at its carboxy terminal domain while has a broad antibacterial properties at its amino terminal (Tomee *et al.*, 1997). SLPI synthesis can be induced in murine macrophages upon exposure to lipopolysaccharide of Gram-negative and lipoteichoic acid of Gram-positive bacteria (Jin *et al.*, 1998).

Protegrins demonstrate a strong antibacterial and antifungal activity as well as toxic effects against enveloped viruses

Proteases play an essential role in physiological processes as the most components of extracellular matrix are degraded by neutrophil-derived proteases. Based on the chemical properties of their active sites the proteinases are classified into four distinct groups. Metalloproteases and serine proteases are involved in degradation of extracellular proteins whereas aspartate proteases and thiol-proteases play a role in digestion of intracellular proteins (Owen *et al.*, 1999).

Serine proteinases are found in azurophilic granules. They belong to a large family of enzymes termed the catalytic triad formed from aspartic acid, histidine and serine. The three neutral serine proteases include cathepsin G, elastase, proteinase 3 that are homologs of azurocidin (Almeida *et al.*, 1991). Cathepsin G mediates the aggregation of platelets (Chignard, 1993). Elastase has the ability to cleave CD14 on monocytes and this inhibits LPS-mediated activation of cells (Le-Barillec et al, 1999). Proteinase 3 can induce IL-8 synthesis in endothelial cells (Berger et al, 1996). Finally, Azurocidin has the ability to stimulate phosphate kinase C in endothelial cells (Pereira et al, 1996).

Matrix metaloproteinases (MMPs) a family of enzymes involved in many physiological processes of degradation of matrix (Middelhoven *et al.*, 1997). Collagenase (MMP-8) in specific granules has a specific ability of cleaving type I collagen while, 92 kDa gelatinase (MMP-9) in secretory vesicles can degrade type V collagen (Elkington *et al.*, 2011).

1.8.4. Mechanisms of degranulation

The key mechanisms in the neutrophil activity include intracellular transportation of proteins, delivery of proteins to different compartments, and finally secretion of proteins to extracellular spaces (Figure 1.8). One of the major events during microbicidal activity of neutrophils is degranulation of vesicles into extracellular space or phagolysosomes.

Apart from the secretory vesicles that have endocytic origin the fusion of plasma membrane and neutrophil granules is considered as a heterotypic fusion. Such fusions are protein–protein interactions leading to the final destination of the vesicle that favours the interaction of vesicle phospholipid bilayer with the target membrane (Berton, 1999).

1.8.5. Respiratory burst in neutrophils

Two parallel events in the phagolysosme of stimulated neutrophils determine the antimicrobial efficiency of neutrophils. The release of antimicrobial and enzymatic proteins from granules as well as the ROS production from activated NADPH-dependent oxidase. A variety of receptors as well as phagocytic targets are involved in the triggering these responses.

One of the mechanisms involved in eliminating the invading pathogen is the generation of ROS by activated neutrophils. Respiratory burst is initiated from NADPH oxidase which is a complex enzyme of cytosolic and membrane proteins. The cytosolic parts of NADPH oxidase include p40phox, p47phox, and p67phox whereas membrane proteins are p22phox and glycoprotein-91phox (Babior, 1999).

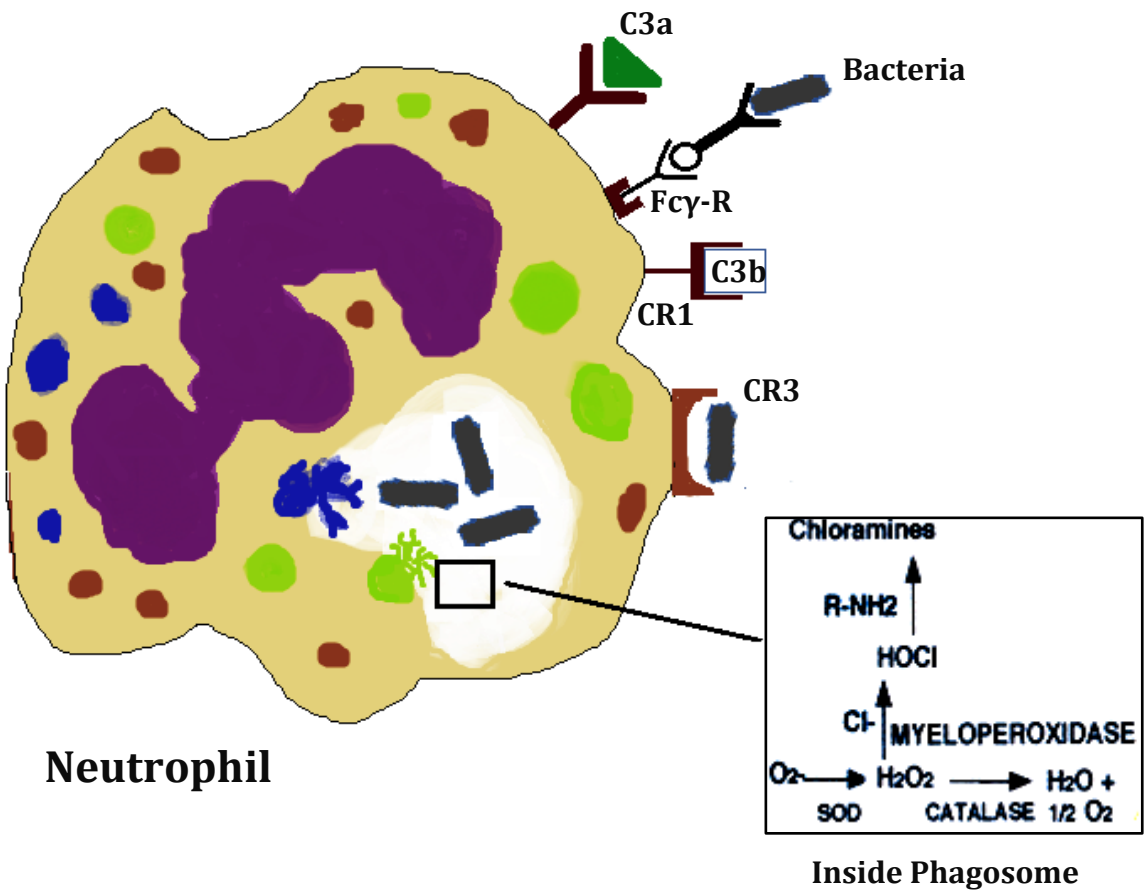


Figure 1.7 The effector mechanisms of neutrophils during inflammation and defense against pathogenic microorganisms. Following phagocytosis of pathogens the effector systems of neutrophil are mobilised. CR1 and CR3 recognise C3b and C4b. Ig receptors (FcγR) recognize IgG opsonins. The oxidative response is the first microbicidal pathway consisting of radical oxygen species resulted from activation of NADPH-oxidase complex which includes superoxide anion (O_2-), hydrogen peroxide (H_2O_2), hypochlorous acid (HOCl) and chloramins. The second pathway is a non-oxygen dependent consisting of the product of phagolysosome or granule proteins. The azurophilic granules contain myeloperoxidase, antibiotic proteins and serine proteases. Specific granules contained antimicrobial proteins such as lactoferrin and cathelicidin, as well as metalloproteinases (gelatinase and collagenase). Gelatinase is present in tertiary granules thus termed gelatinase granules. Redrawn after (Witko-Sarsat *et al.*, 2000)

The phosphorylation of p47phox upon neutrophil activation is accompanied with the migration of cytosolic components to the plasma membrane where they bind to cytochrome b558 resulting in assembly of active oxidase. Formation of this complex can produce superoxide anion (O_2) which in turn dismutates into H_2O_2 (Nathan, 1987).

O_2 reduction into hydrogen peroxide H_2O_2 is mediated by three intermediates generated via one electron transfer namely superoxide ($\bar{O}_2$), (H_2O_2) and hydroxyl (OH^-) radicals. The production of O_2 is the basic material to initiate the generation of wide range of reactive oxidants. One example is halogenated oxidants which are produced from myeloperoxidase (MPO) pathway (Klebanoff, 1998). MPO, a haeme protein in azurophilic granules, is released into the phagolysosome and the extracellular space after activation of neutrophil. It produces reactive intermediates that increase toxic potential of H_2O_2. The main MPO product in the plasma is hypochlorous acid (HOCl), a potent oxidant that can chlorinate electron-rich substrates and lyse heme proteins and nucleotides.

A wide range of byproducts are made from MPO substrates. In this regard, the MPO-H_2O_2 interaction system generates chlorination products, tyrosyl radicals, tyrosine peroxide, reactive aldehydes, and the oxidation of lipoproteins and serum proteins (Heinecke, 1999). MPO utilizes substrates such as hydrogen peroxide and nitrite in order to catalyze the nitration of tyrosine in proteins (Sampson *et al.*, 1998) as well as reacting with peroxinitrite (Podrez *et al.*, 1999). These interactions can initiate the connection with the nitric oxide synthase system. Beside their role in host defence, MPO-derived oxidants are also implicated in processes of carcinogenesis, atherosclerosis, and chronic renal failure due to high levels of MPO-derived proteins in plasma leading to inflammatory reactions (Witko-Sarsat *et al.*, 1998). While the main source of MPO is neutrophil, some MPO-mediated biological activities can occur in monocytes.

Nitric Oxide-Synthase-Derived Reactive Nitrogen Intermediates

Among eukaryotic enzymes, nitric oxide synthases are unique as they are dimeric, calmodulin-containing cytochrome P450 hemoproteins with the ability to combine reductase and oxygenase domains in one monomer. Nitric oxide (NO) is a member of reactive nitrogen intermediates that react with oxygen forming potent oxidants such as nitrogen dioxide (NO_2). The direct toxicity of nitric oxide can be profoundly increased by peroxynitrite (ONOO-) produced from its interaction with superoxide (Beckman *et al.*, 1996).

1.9. Neutrophil Extracellular Traps (NETs)

NETs occur when a meshwork of chromatin strands loaded with granule-derived antimicrobial proteins and enzymes such as elastase and MPO are extruded by neutrophils. This process of NETs formation is called NETosis a novel mechanism to entrap and kill extracellular bacteria (Mesa *et al.*, 2013).

1.9.1. Mechanism of NETosis

During NETosis the granule membranes and nucleus of the cell disintegrate resulting in release of granule products and decondensed chromatin fibres into the extracellular space. The co-localisation of granule proteins and chromatin fibres outside the cell can be visualised by immune-staining and is characteristic of NETs (Papayannopoulos *et al.*, 2009).

NETosis can be distinguished from apoptosis by lack of “eat me” cell surface signals. Therefore, NETs are not removed by other phagocytes, instead nucleases disassemble the residual chromatin strands (Fuchs *et al.*, 2007). The stimulation of NETosis can be initiated by a wide range of inflammatory mediators such as immune complexes, tumour necrosis factor and IL-8, as well as microbial proteins from bacteria, protozoa, and fungi. In fact, the extent of NET formation is determined by the combination and strength of such stimuli. NETosis, unlike

oxidative burst and degranulation, is a slow process taking more than 4 hours to occur (Clark *et al.*, 2007).

NADPH oxidase activity is required to stimulate NETosis during infection (Bianchi *et al.*, 2011). Intracellular caspases can be inactivated by ROS leading to inhibition of apoptosis and induction of autophagy and this promotes the breakdown of the cell membrane during NETosis (Remijsen *et al.*, 2011). In the absence of ROS production other stimuli such as soluble immune complexes immobilised to fibronectin via FCγRIIIA can also induce NETosis (Chen *et al.*, 2012). Citrullination of histones (i.e. conversion of arginine into citrulline) is required for chromatin disassembly to facilitate complete dispersion of NETs (Wang *et al.*, 2009a).

Antimicrobial function of NETs

NETs have antimicrobial functions; first, the pathogen is localised and trapped within a chromatin meshwork, then the pathogen is exposed to high concentrations of antimicrobial peptides and enzymes. In addition to MPO and elastase, NETs contain other concentrated antimicrobial substances such as LL-37 and lactoferrin-chelating proteins. Likewise, histones have considerable antimicrobial activity (Papayannopoulos *et al.*, 2010).

The synergistic action is enhanced when both chromatin activity and antimicrobial proteins work together. In fact, some bacteria such as streptococci and staphylococci degrade NETs by expressing nucleases so that they can be released from chromatin network (Beiter *et al.*, 2006). *Streptococcus pneumoniae* provides another example of an escape strategy by modifying its capsule to reduce contact with NETs (Wartha *et al.*, 2007).

Despite the fact that DNA is expelled from the cell to produce NETs, it has been suggested that neutrophils are able to migrate and perform antibacterial activities in a conventional way even in the absence of a nucleus (Yipp *et al.*, 2012). Thus NETosis is a distinct from apoptosis because the latter is associated with the lack

of neutrophil function. The molecular mechanisms that differentiate between NETosis and apoptosis are not particularly clear as both processes are affected by ROS. In this regard, the fate of neutrophil can be determined by the duration and magnitude of ROS production (Chen *et al.*, 2012).

1.9.2. NETs in tuberculosis

Ramos-Kichik and colleagues showed that Mtb can induce NETosis. Formation of neutrophil aggregates was stimulated by both Mtb H37Rv and *M. canetti* in vitro and detected by transmission and scanning electron microscopy. For confirmation, neutrophil elastase, DNA and histones were stained with fluorescent immunohistochemical stains. While, NETosis is induced by Mtb and entraps the bacilli, it is reported that NET effector mechanisms are not lethal to this organism. This is in contrast to another intracellular pathogen, *Listeria monocytogenes,* and may reflect the importance of unique lipid-rich mycobacterial cell wall (Ramos-Kichik *et al.*, 2009).

The interaction between Mtb and components of innate immune system is complex. Macrophages have been found capable of ingesting proteins from azurophil granules of neutrophils and to utilize these proteins to kill mycobacteria (Jena *et al.*, 2012). This effect was initially attributed to neutrophil elastase and cathepsin G. Alveolar macrophages were able to uptake these enzymes and apparently also use them to eliminate invading mycobacteria (Steinwede *et al.*, 2012). Moreover, neutrophil-derived azurophilic enzymes can be taken up by macrophages and subsequently employed for destruction of mycobacterial species including Mtb (H37Rv), *M. bovis* (BCG) and *M. smegmatis.* Thus, acquiring azurophilic enzymes in macrophages can facilitate bacterial cell envelope degradation (Jena *et al.*, 2012).

Interestingly, extracellular traps can also be generated by macrophages on encounter with Mtb. The formation of macrophage extracellular traps mainly depends on the presence of IFN-γ and the ESX-1 system in mycobacteria leading

to the rupture of phagocyte and release of bacilli. The extracellular traps of macrophages are similar to NETs and comprise both DNA and citrullinated histones (Wong *et al.*, 2013).

1.9.3. Host influence on NETosis in pulmonary system

A direct interplay between microorganism and components of the innate and adaptive immune systems is a key to eliminating invasive pathogens. Mtb provides a prime example of a virulent bacterium that breaks down this system by evading the host and eventually killing it (Alimuddin *et al.*, 2013).

The ability to produce NETs by neutrophils depends on external stimuli as well as the local microenvironment and the host pulmonary environment offers all the required components (Figure 1.9). Proteins such as collectin, SP-A, and SP-D are synthesised by alveolar epithelium and these bind carbohydrate moieties on pathogens such as Mtb and fungi (Giannoni *et al.*, 2006). They can also bind to NETs and stimulate augmented NET-mediated antibacterial effects. In some conditions disturbance in alveolar tissue leads to NETosis and a pool of cell-free DNA (Papayannopoulos *et al.*, 2011).

1.10. Neutrophils in TB sputum

Traditionally, macrophages are believed to be the main target of Mtb and play an important role in containment of these bacilli in granulomas. Nevertheless, the role of polymorphs in TB is gaining recognition. A transcriptome study on peripheral blood of TB patients yielded a strong signal indicating the presence of interferon-stimulated neutrophils (Berry *et al.*, 2010).

Neutrophils are one of the initial cellular components in TB granulomas (Gideon et al., 2011). They isolate the infected cells in an organised structure and suppress mycobacterial replication (Gupta *et al.*, 2012).

Inflammatory cells accumulate at the initial sites of infection. The polymorphs die and alveolar macrophages and the surrounding pulmonary tissues are subjected to the effects of toxic immune products. As a result a caseous necrotic core is generated on which extracellular Mtb bacilli live and replicate (Rustad *et al.*, 2008). The necrotic core is surrounded by a mixture of cells including activated macrophages, neutrophils and fibroblasts (Chao *et al.*, 2010). It has long been thought that caseum of liquefying cavities provides the site for rapid growth of Mtb. The bacilli are ultimately released to airways and transmitted to the air. However, Eum and colleagues propose that there is a burst of replication during the exit of bacilli from a liquefying cavity into the sputum and that this involves contact with neutrophils (Eum *et al.*, 2010).

Therefore, neutrophils may play a much important role in the pathogenesis of Mtb than was previously appreciated. Their participation in several functions during TB has been demonstrated and includes the transportation of live mycobacteria from the lung periphery to lymph nodes in mice (Abadie *et al.*, 2005). Neutrophils isolated from TB-susceptible mice appear to have higher capacity for migration and survive longer (Eruslanov *et al.*, 2005). One function of neutrophils may be the transportation of bacilli from lesions into sputum. Eum et al found them to be the dominant phagocytes present in TB sputum. These authors microscopically examined for acid-fast bacilli in sputum, bronchoalveolar lavage (BAL) and cavity caseum samples from TB patients. The majority of neutrophils were found to be associated with acid-fast bacilli. Further microscopic observations suggested that large proportion of bacilli were extracellular mainly in association with a polymorphonuclear phagocyte (Eum *et al.*, 2010).

Neutrophils are abundantly found in the bronchoalveolar fluid of TB patients (Law *et al.*, 1996). While, recruitment of neutrophils as the first defensive cells occurs following acute infection, they are persistently recruited to tissues during chronic mycobacterial infection (Appelberg *et al.*, 1991). Indeed, sputum samples with

abundant neutrophil content are considered as a high-quality sample (McCarter *et al.*, 1996).

Neutrophils appear to be important in providing a microenvironment for Mtb in the pulmonary environment of TB patients. It is well known that Mtb has a metabolism highly adaptable to its microenvironment. Understanding the mechanisms of mycobacterial adaptation in sputum, specifically during its interaction with neutrophils could reveal novel approaches for interfering with the transmission of Mtb.

1.11. Aim and objectives

This study was performed with the aim to investigate the transcriptional responses of Mtb (H37Rv strain) to human neutrophils in an in vitro interaction. The main objectives were:

- To characterise the in vitro interaction of Mtb and neutrophils regarding phagocytosis and survival of ingested mycobacteria
- To establish a method of RNA extraction from H37Rv associated with human neutrophils in vitro
- To test the primary hypothesis that neutrophil encounter may explain the distinctive sputum signature of Mtb
- To determine the pattern of H37Rv transcriptional changes on encounter with neutrophils
- To detect and categorise differentially expressed genes unique to the neutrophil encounter

CHAPTER TWO

Materials and methods

2.1. Sources of polymorphs and mycobacterial strains used in this study

CELLS	DESCRIPTION	SOURCE
HL60 cells	Human Leucocyte cell line	ATCC (CCL-240)
Neutrophils	Primary human neutrophils	Isolated from human venous blood
***M. bovis* BCG (Glaxo)**	Attenuated category II TB vaccine strain	Laboratory stock
***M. tuberculosis* (H_{37}Rv)**	Virulent laboratory strain	Laboratory stock (WR Jacobs lab strain)

2.2. Materials, culture media and reagents

Falcon tubes and other plastic-wares were from Corning (New York, USA). 1.5 ml micro-tubes and 2 ml screw capped cryotubes were from VWR (Lutterworth, Leicestershire, UK). All chemicals and eukaryotic cell culture media were obtained from Sigma-Aldrich (Poole, Dorset, UK) or Fisher Scientific (Loughborough, Leicestershire, UK), unless otherwise indicated. All bacterial culture media were obtained from Becton Dickinson Biosciences (Oxford, UK) unless otherwise stated.

The sterilisation of media and reagents was achieved by autoclaving at 121°C at 15 psi for 15 minutes unless otherwise stated.

2.2.1. Growth media

2.2.1.1. RPMI 1640

For growth of neutrophils RPMI was supplemented with L-Glutamine (final conc. 1%), FBS (Final conc. 10%) and penicillin / streptomycin (final conc. 1%). For infection and washing purposes the antibiotic free RPMI was used.

2.2.1.2. Middlebrook 7H9 broth

In a 1L Duran bottle Middlebrook broth was made by dissolving 4.7g of broth powder in 900ml distilled water containing 2.5g glycerol then the solution was sterilised. The broth was supplemented with 10% w/v Albumin-dextrose-catalase (ADC) and 10% Tween-80 at concentration of 0.05% (v/v).

2.2.1.3. Middlebrook 7H10 agar

Middlebrook 7H10 agar was made by dissolving 19g of agar powder in 900 ml distilled water containing 6.25g glycerol. The agar was heated and stirred until the powder was fully dissolved then sterilised. Prior to solidification the medium was supplemented with Oleic acid-albumin-dextrose-catalase (OADC) at a concentration of 10% (v/v).

2.2.2. Reagents

Albumin – Dextrose – Catalase (ADC) supplement

ADC was made by dissolving 7.5g bovine serum albumin (BSA) fraction (V), 3.0g D-Glucose, 1.28g Sodium Chloride and 6.5mg catalase in 150ml distilled water. The solution was centrifuged at 6000xg for 30 minutes to exclude undissolved particles then filter-sterilised using 0.2µm micor-filter (Nalgene, Hereford, UK) and stored at 4˚C.

Oleic Acid-Albumin-Dextrose-Catalase (OADC) Supplement

Oleic acid is a supplement for mycobacterial growth (Winn *et al.*, 2006). OADC supplement was prepared as per ADC supplement above, with the addition of 8.63 ml of Oleic Acid solution (1% w/v) in 0.2 M NaOH (pre-warmed to 50˚C). The OADC solution was filter sterilised through a 0.2µm filter.

Tween 80

Tween 80 is used to minimise clumping of the mycobacteria species during growth in liquid culture. Tween 80 was prepared by dissolving 10g of stock solution in distilled water to a final volume of 100 ml and final concentration of 10% (w/v). The solution was sterilised by filtration through a 0.2µm filter unit and stored at 4°C.

Amikacin

500mg amikacin powder was dissolved in 10ml distilled water to a final concentration of 50mg / ml. The solution was filter-sterilised by 0.2µm micro-filter. 10 aliquots of 50mg/ml were made in 2 ml screw capped cryotube and stored at -80˚C.

2.3. General methods

2.3.1. Neutrophils

2.3.1.1. Isolation of human neutrophils

Neutrophils were isolated from venous blood of healthy volunteers who had no alcohol consumption for 48 hours prior to blood collection. A consent form was signed by the donor. Using a 50ml syringe containing 2µl/ml of heparin (5000 unit/ml) 50ml of blood was taken and neutrophils were isolated. To 5 ml of isolation medium (Cedarlane, Burlington, Canada) composed of sodium metrizoate and dextran 500 (v/v) in a 50ml Falcon centrifuge tube 5ml of blood was carefully layered with the pipette tip close to the surface to avoid mixing.

The suspension was centrifuged at 500xg for 35 minutes at room temperature. The blood should separate into 6 distinct bands in the sequence: plasma, monocytes, isolation medium, neutrophils, further isolation medium, and red blood cells as a pellet. If these bands were not clear, the separation process was repeated. The top three layers were carefully removed and discarded. The layer of neutrophils and all of the isolation media beneath were carefully aspirated and transferred into a new falcon tube. The neutrophil suspension was diluted to 10ml with Hank's balanced salt solution (HBSS) without Ca2+/Mg2+ and the tube was inverted a few times to mix.

The suspension was centrifuged at 350xg for 10 minutes producing a red pellet containing neutrophils and residual red blood cells (RBCs). After removing the supernatant 2ml Red Cell Lysis Buffer was added and the suspension vortexed at setting 3000 RPM followed by centrifugation at 250xg for 5 minutes and discarding the supernatant with a pipette. The lysing process was repeated as required. 500µl HBSS without Ca2+/Mg2+ was added to each tube and dilution was done by adding 10ml Ca2+/Mg2+ free HBSS. The tubes were then centrifuged at 250xg for 5 minutes then supernatant was discarded. The pellet was re-suspended in 250µl HBSS Solution (2% HSA). Cells were then counted and adjusted to the desired concentration.

2.3.1.2. Maintenance of HL60 cells

The stock of HL60 cells growing at passage 5 was kindly provided by Dr C. Bunce, University of Birmingham, and was originally acquired from ATCC (CCL-240). The cells were maintained in RPMI-1640 supplemented with 10% heat inactivated FBS and 1% of Streptomycin (final conc. 100 µg / ml) – Penicillin (final conc. 100 U/ml). The cultures were maintained at 37˚C in 5% CO_2. Feeding of cells with fresh RPMI was done every two days to preserve cell density between 2×10^5 to $9x10^5$ cells / ml. This was determined by microscopic observation of Trypan blue stained preparations (See below) which should show a single-cell suspension with no

tendency to adhere to the flask. The cells were only maintained up to a maximum of 35 passages with the optimal passages being 10 – 20 for infection purposes (Fleck *et al.*, 2005).

2.3.1.3. Differentiation of HL60 cells

All-trans retinoic acid (ATRA) was added to primitive HL60 cells in RPMI at concentration of 30ng / ml in order to differentiate them into mature HL60 cells. The cell suspension was incubated at 37˚C in 5% CO_2 for 5 days. After differentiation, cells should remain non-adherent. The accomplishment of cell differentiation was checked by Giemsa staining of treated and untreated cells.

2.3.1.4. Cryopreservation of HL60 cells

For long term storage of cells the cryopreservation of undifferentiated HL60 cells at their third passage was accomplished in liquid nitrogen. In 1.5ml screw cap tube 2 × 10^6 cells / ml in RPMI based medium containing 20% DMSO were added. Since, DMSO is an active inducer of cellular differentiation and is toxic to cells and this effect was minimised by placing the cell suspension at -20 ˚C for 2h immediately after adding DMSO then the vials were transferred into liquid nitrogen. Thawing was done by placing the vial under in water bath at 37˚C. The cells were washed twice with pre-warmed RPMI via centrifugation at 200xg for 5 minutes in order to remove the preservation medium.

2.3.1.5. Trypan blue viability test

From trypan blue stock (Sigma Aldrich, Dorset, UK) a fresh solution of 0.4% in distilled water was prepared.

- 10µl of neutrophils were added to 10µl trypan blue 1:1 (v/v) and gently mixed
- 10µl was transferred to a haemocytometer
- Cells were counted in total area 1mm^2 at magnification of 20x corresponding to cell density of 10^4 / ml

$$Density = Total\ counts\ per\ area\ x\ 10^4 * 2 = \frac{Cells}{ml}$$

- Dead cells stain blue while viable cells remain unstained and fully transparent

$$Viability = \frac{Total-dead}{Total} * 100$$

2.3.1.6. Preparation of ATRA-treated HL60 cells

Prior to infection the differentiated HL60 cells were washed twice with RPMI at 200xg for 5 minutes then the viability and density of cells were measured with Trypan blue. The density of cells was adjusted according to the experimental design and only HL60 suspensions with more than 80% viability were used. Then the cells were re-suspended in pre-warmed RPMI to be ready for infection. When required IFN-γ (50 units/ml) was added for 1 hour before the infection started.

2.3.1.7. Activation of neutrophils with IFN-γ

When required isolated primary human neutrophils and differentiated HL-60 cells were activated with IFN-γ 50 units/ml and incubated at 37˚C for 1h prior to infection.

2.3.1.8. Washing neutrophils with RPMI

Antibiotic free RPMI was used to remove extracellular bacilli, antibiotics and other reagents from neutrophil - mycobacteria suspensions. Any individual wash was done by transferring the cell suspension to a 50ml propylene conical tube and topped up to 40 ml with antibiotic free pre-warmed RPMI then centrifuging at 200xg for 5 minutes. The supernatant was removed and the pellet re-suspended in the desired volume of RPMI.

2.3.2. Mycobacteria

2.3.2.1. Cultivation of BCG

BCG was routinely grown in 7H9 broth supplemented with OADC and Tween 80. The culture was incubated statically in 50 ml conical flasks at 37˚C. The cell suspension was maintained at log phase of growth via mixing 1ml of cell culture at OD = 0.5 – 0.9 in 49ml fresh 7H9 broth. The OD of the culture was regularly measured at λ = 580nm.

2.3.2.2. Cryopreservation of BCG

20 ml BCG in 7H9 broth at OD = 0.5 – 0.8 was transferred into 50ml propylene conical tube and centrifuged at 1000xg for 10 minutes. The supernatant was removed and 10ml RPMI containing 20% human serum was added to re-suspend the pellet. 20 aliquots of 0.5ml were made at density of 2.5 $x10^7$ bacilli per ml in 2 ml screw capped cryotube then frozen at -80°C for future experiments.

2.3.2.3. Cultivation of H37Rv

Based on Advisory Committee on Dangerous Pathogens (ACDP) H37Rv is classified as category III hazardous pathogen which requires level-3 biohazard containment. Thus, all cultivation and experiments with H37Rv were performed within the containment laboratory suite in Class I and class II microbiological safety cabinets in accordance to the suite's code of practice.

A frozen vial of $H_{37}Rv$ was kindly provided by Dr N J Garton. In a 250ml polypropylene conical flask the bacteria were grown in 50 ml 7H9 broth supplemented with 10% ADC and 1% Tween 80 incubated at 37°C /100 rpm in a shaking incubator. The first subculture was done when the OD of the culture reached 0.32 (early exponential phase growth) by transferring 1ml of growing bacilli into 24ml fresh ADC-Tween 80 supplemented 7H9 broth. The culture was

then further incubated as before. After 3 further sub-cultures at early exponential phase bacteria were prepared for cryopreservation.

2.3.2.4. Cryopreservation of H37Rv

30ml H37Rv in 7H9 broth growing at OD = 0.6 – 0.8 was pelleted in a 50ml propylene conical tube at 1000xg for 10 minutes. The supernatant was removed and pellet was re-suspended in 20ml RPMI containing 20% human serum. 20 aliquots of 1ml containing 5.6 x 10^7 bacilli / ml were made in 2 ml screw capped cryotube then stored at -80^{o}C.

2.3.2.5. Preparation of H37Rv for neutrophil infections

For each infection, an aliquot of 1ml frozen H37Rv was thawed and mixed with 24ml 7H9 broth at room temperature in a class 2 safety cabinet. A 5ml pipette was used to disperse the clumps of bacteria by pipetting up and down 10 times then the suspension was transferred to a 125ml flask and incubated at 37°C 100 RPM shaking incubator. The OD of the culture was measured every day at 580nm. When the OD reached 0.3 the bacterial suspension was sub-cultured by transferring 1ml culture into 24ml fresh 7H9 broth then incubated at 37˚C in shaking incubator at 100 rpm. The OD was checked every day. After 3 subcultures, 1ml of culture was taken into 49ml of fresh 7H9 broth in 250ml flask and incubated at 37˚C / 100rpm in shaking incubator. The OD was checked every day until it reached OD=1.0 at 580nm.

A required volume (40ml) of culture at OD=1.0 was taken into a 50ml propylene conical tube then centrifuged at 1000xg for 10 min. The supernatant was removed and the pellet was re-suspended in RPMI 1640 supplemented with 10% Fetal calf serum and 1% L-glutamine to be used for infection.

2.3.2.6. Measuring optical density

The optical density (OD) was measured by transferring 1 ml of culture into a 1.5ml cuvette which was then double sealed with autoclave tape and Nescofilm (Bando Chemical, Kobe, Japan). The OD of BCG was measured using a Sanyo SP75 UV/Vis spectrophotometer (Watford, UK) in Category 2 laboratory whereas Jenway 6300 spectrophotometer (Stone, UK) was used to measure the OD of H37Rv in the Category 3 laboratory. To insure accuracy, thick cultures (OD > 1.0) were diluted 1:10 prior to measurement. The OD580nm of was measured against blank of 7H9 broth.

2.4. Notes and consideration

In all experiments in this study the following should be considered unless OTHERWISE stated:

1. The number for cells, bacteria and multiplicity of infection (MOI) is the mean of replicates
2. For removing antibiotics and other substances two washes were applied.
3. For removing extracellular bacteria three washes were applied.
4. Biological replicates referred to as “n” were individual experiments that have been performed at different times. A new stock of bacteria and neutrophils was used for each biological replicate.
5. Technical replicates were done when practicable and required by the analysis.

CHAPTER THREE

Establishing Conditions for Studying BCG and H37Rv Interactions with Human Neutrophils

3.1. Introduction

It was necessary to determine the factors affecting the interactions between Mtb and neutrophils in order to progress to studying the mycobacterial transcriptional responses to this encounter. It was also important to appreciate and manage the factors affecting bacterial responses *in vitro* so that these could be controlled in the planned experiments and enable reproducible results to be obtained. The HL60 cell line was used here to develop the experimental protocols and enable the most effective use of freshly isolated neutrophils.

3.1.1. HL60 cells

The Human leukocytic cell line (HL60 cell) was originally derived from peripheral blood of a 36 year old Caucasian woman with acute pro-myelocytic leukaemia (Tsiftsogiou *et al.*, 1985).The cells have been used as a model system for human polymorphonuclear phagocytes in view of the fact that terminally differentiated primary human neutrophils have a relatively very short life time and are not suitable for genetic manipulation (Hauert *et al.*, 2002).

3.1.1.1. Characteristics of HL60 cells

For decades since its isolation, the HL60 cell line has been extensively characterised. It has been among the first human myeloid leukocyte lines to be established in long term culture. The original wild type cells show several properties of malignant cells and have oncogenic potential enabling induction of tumours in nude mice; these mainly comprise leukemic myeloblasts and promyelocytes (Fleck *et al.*, 2005). HL60 cells maintain continuous proliferation with doubling times of 36 to 48 hours in suspension culture (Collins, 1987). Enhanced by several passages and a number of colony stimulating factors, the cells can grow in agar and semisolid methylcellulose medium as colonies. As a

result, many sublines have been developed with restricted differentiation potential including eosinophilic sublines incapable of differentiation to monocytes and neutrophils.

Cytological studies of HL60 cells have confirmed that they are myeloblastic or promyelocytic with ovoid or round shape but occasionally expressing pseudopods in culture. They are heterogeneous in size ranging from 9 to 25µm in diameter. With large round nuclei, undifferentiated HL60 cells are sometimes binucleate with fine chromatin and distinct margins.

The cytoplasm of HL60 cells has prominent multiple azurophilic granules and is intensely basophilic. The periodic acid-Schiff reagent stains HL60 cells positively; a characteristic of mature *ex vivo* granulocytes. However, unlike naive granulocytes, they do not express characteristics of in vivo-derived granulocyte such as alkaline phosphatase. Additionally, some HL60 cells can bear a resemblance to the precursors of erythrocytes or megakaryocytes (Fleck *et al.*, 2005).

3.1.1.2. Differentiation of HL60 cells

Multiple environmental conditions including chemicals and pH enable HL60 cells to be differentiated into various myeloid lineages via arresting their cell cycle. The optimal differentiation can be affected by several factors such as concentration of the inducer, the relative proportion of cells in different stages of the cell cycle and the length of exposure. However, even a short exposure to these inducers may be sufficient with no necessity to maintain continuous exposure.

Despite the fact that cultivation of HL60 cells is relatively simple, careful handling, passaging and culturing is needed as the cells can easily be provoked to differentiate into a subline or to a non-proliferating cells. The cell density should be maintained at less than 10^6 cells/ml (Fleck *et al.*, 2005).

Differentiation of HL60 cells into mature polymorphonuclear phagocytes can be induced with a range of chemical substances including All Trans-Retinoic Acid (ATRA) and Dimethyl Sulphoxide (DMSO). They can also be induced to differentiate into macrophage-like cells with 12-O-Tetradecanoylphorbol-13-acetate (TPA) or 1α, 25-dihydroxyvitamin D3 (Tsiftsogiou *et al.*, 1985). During differentiation, HL60 cells begin to change their morphology, express new genes then undergo apoptosis.

Various phenotypic changes occur in the differentiated granulocytes including the expression of surface antigens such as CD11b and CD66 while CD71 undergoes down regulation. The cell can recognise deposited C3b and iC3b and the Fc fragment of bound antibodies on bacterial surfaces via specific receptors (Nordenfelt *et al.*, 2009). The expression of insulin and transferrin receptors is crucial for proliferation. Therefore, even in the absence of serum in nutrient media such as RPMI, the cells maintain their proliferation capacity provided that the medium is supplemented with insulin and transferrin. The requirement for insulin and transferrin is absolute. The expression of transferrin receptors significantly decreases while cells run through terminal differentiation (Collins, 1987).

3.1.1.3. HL60 cells as a model for Primary human neutrophils

Owing to their high functional resemblance, terminally differentiated HL60 granulocytes have been used widely as a model system for primary human neutrophils in a variety of studies including those concerned with haematopoiesis and neoplasia (Tsiftsogiou *et al.*, 1985), responses to chemokines and motility properties (Hauert *et al.*, 2002, Wang *et al.*, 2009b), and in the process of NET formation (Kawakami *et al.*, 2014). In phagosome studies of neutrophils, however, due to the absence of specific granules, the HL60 cells can be viewed as granule poor neutrophils.

The respiratory burst in the phagosome of HL60 cells is relatively lower than in native neutrophils because NADPH oxidase deposition on the phagosome is incomplete and associated with lower phagosome antibacterial efficiency. Such characteristic phagosome inefficiency in HL60 cells has made them useful in the studies of killing capacity of plasma membrane localised respiratory burst (Timm *et al.*, 2008, Nordenfelt *et al.*, 2009). Nonetheless, the use of HL60 cells for human neutrophil studies provided a number of scientific and practical advantages, for example by minimising experimental variation due to heterogeneity of primary human neutrophils.

3.1.2. Aim and objectives

The primary aim of the work described in this chapter was to characterise neutrophil interactions with BCG and Mtb in order to establish a robust model of the encounter between Mtb and these cells in the lower respiratory tract at the time of expectoration in sputum. The main objectives were:

- To quantify the uptake of mycobacteria by neutrophils
- To study the survival of BCG and Mtb in their encounter with neutrophils
- To determine the time course of oxidative burst in neutrophils infected with Mtb

3.2. Methods

3.2.1. Materials and preparations

3.2.1.1. Cell culture plates and flasks

Cell culture plates 12 and 24 well plates were purchased from Nunclon (Roskilde, Denmark).Vent cell culture flasks were provided from Corning (New York, USA).

3.2.1.2. Giemsa Stain

Giemsa stain was used to stain neutrophils. Slides were prepared by spreading 50µl cell suspension on a slide then air-drying at room temperature. The slides were fixed with 100% methanol for 1 minute then rinsed with tap water. Fresh 10% Giemsa stain (v/v) was made in distilled water and this was applied to slides for 30 minutes then rinsed off with tap water and air-dried.

3.2.1.3. Microscopy

For visualisation of slides a Nikon Ti-E eclipse inverted microscope with an Intensi-light C-HGFIE pre-centred fibre optic light source was used. The images were produced using ImageJ software (National Institutes of Health, Besthda, Maryland).

3.2.1.4. Tetrazoluim solution

2-4-iodophenyl-3-4-nitrophenyl-5-phenyl-2H-tetrazolium chloride (INT) solution was made by dissolving 500mg INT powder in 20ml distilled water (25 mg/ml). The solution was incubated overnight at 37˚C in order to completely dissolve the powder then sterilised by passing through a 0.2µm filter.

3.2.1.5. Colony forming unit (CFU) counts

Sodium Dodecyl Sulphate SDS (0.15% final concentration in 7H9 broth) was added to pellet in order to get lysate of host cells and release of internalised bacilli. The lysate was then diluted at 10-fold serial dilution (10^{-1} to 10^{-6}) in 7H9 broth. From each dilution 20µl in triplicate was dropped on Middlebrook agar. Plates were incubated at 37°C for 10 – 15 days after which colonies of mycobacteria were counted. Two plates were made for each sample. On a plate only one section with maximum of 20 colonies was taken for calculation. Counts were based on aggregate of reads from all biological replicates each with two technical replicates.

3.2.2. General Methods

3.2.2.1. Infecting neutrophils with BCG and H37Rv

Infections of HL60 cells and primary human neutrophils were done in 1ml RPMI in 12 well cell culture plates unless otherwise stated. The BCG inoculum was prepared in RPMI by removing 7H9 broth from exponential phase bacilli using centrifugation at 1000xg for 10 minutes. For CFU counting at the end of incubation Sodium Dodecyl Sulphate (SDS; 0.15% final concentration in 7H9 broth) was added to the cell suspension in order to lyse the host cells and release internalised bacilli. Clumps of bacilli were dispersed via pipetting the cell suspension up and down 10 times.

3.2.2.2. Removing extracellular bacilli via multiple washes

The effect of washing on removing extracellular bacilli was measured as:

$$Efficiency\ of\ Wash = \frac{(CFU\ of\ Wash)\ *\ (100)}{(CFU\ of\ Inoculum)}$$

The viability of HL60 cells was assessed after each wash.

3.2.2.3. Measuring oxidative burst in HL60 cells infected with H37Rv

Tetrazolium salts were used to measure the oxidative burst of neutrophils in their encounter with Mtb. Thus, the formazan produced by reduction of INT was measured using a colorimetric assay. 2.6×10^5 IFN-γ treated HL60 cells in 150µl phenol red-free RPMI were added per well in a 24 well cell culture plate. 50µl H37Rv was added to the HL60 cell suspension at MOI = 5.7 in total volume of 200µl then incubated for 0, 30, 60, 90, 120, and 150 minutes. 5µl of INT was added to the sample at the end of incubation (final conc. 2mg / ml). 30 minutes after adding INT the phagocytes were lysed with 5µl of 20%SDS to release the produced formazan. 550µl DMSO was added to the suspension to solubilize the formazan then the contents of the well were transferred into a micro-cuvette to measure the absorbance at 580nm. The blank was set as all reagents without neutrophils and bacilli.

3.3. Results

3.3.1. Efficacy of amikacin for killing extracellular bacteria

The initial focus of this study was to confirm phagocytosis was occurring. The extracellular bacilli were excluded from being counted because of their sensitivity

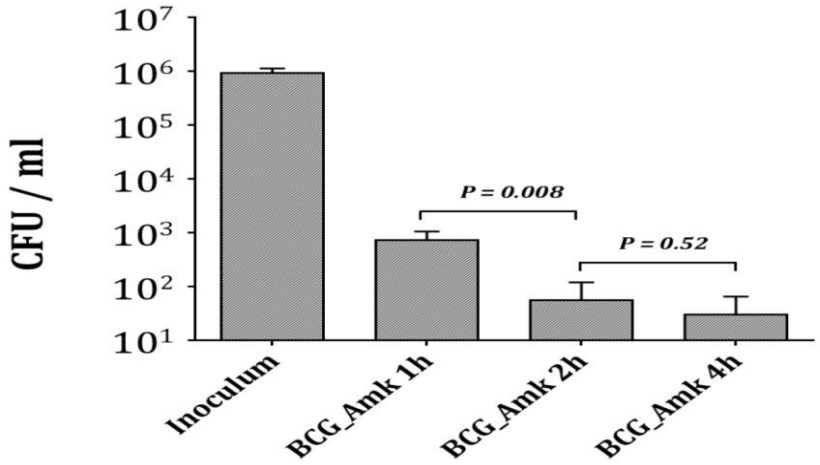


Figure 3.1 CFU count of BCG was significantly reduced at 2h exposure to 200µg/ml amikacin. 9.2±1.2 x 10^6 exponential phase BCG bacilli in RPMI were treated with 200µg/ml for 1h, 2h and 4h then washed with RPMI and CFU was measured from pellet.
(Error bars: SD n = 4)

to the bactericidal effect of amikacin. Figure 3.1 demonstrates a four log reduction in the CFU count of BCG after 1h of amikacin exposure and five logs after 2h and 4h. The difference between 2h and 4h was not significant (Paired t-test p = 0.52). Therefore, 2h amikacin treatment was assigned as the standard exposure time.

3.3.2. Differentiation of HL60 cells

Primitive HL60 cells were treated with ATRA for 5 days to differentiate them into mature neutrophils. Giemsa staining showed cells with no distinct nuclei before

treatment (**Figure3.2A**) and about 97.8% cells with segmented nuclei clearly distinguishable from cytoplasm after treatment with ATRA (**Figure 3.2B**).The trypan blue test revealed that 80% of the cells were viable 72h after differentiation. Viability then decreased substantially to around 30% after 120h (**Figure 3.2C**).

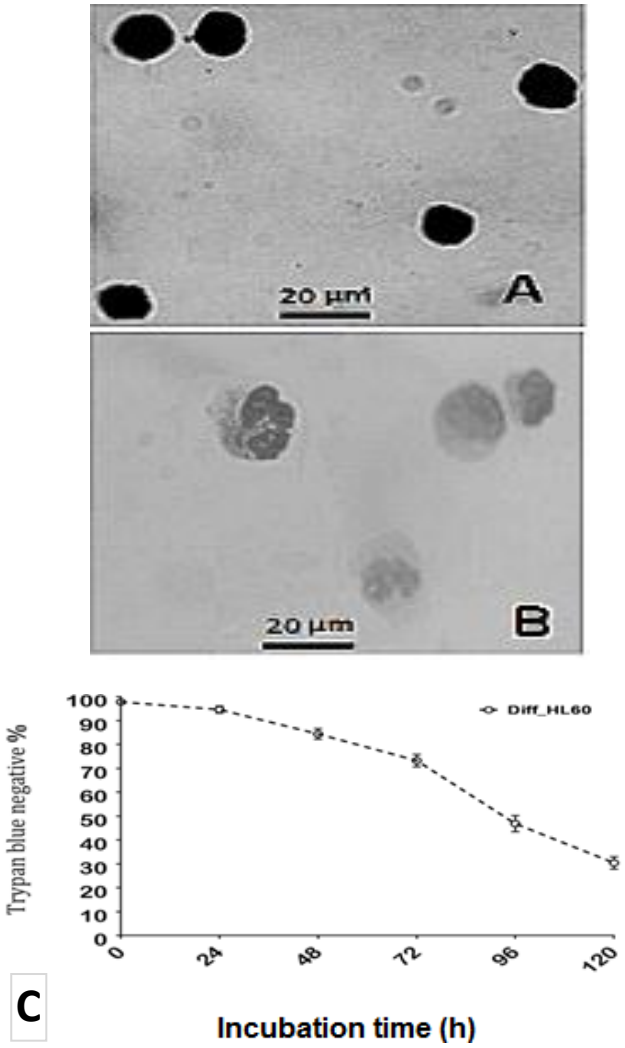


Figure 3.2 HL-60 cells were differentiated with ATRA for 5 days and the majority remain viable for a further 72h. Naïve HL-60 cells were incubated with 30ng/ml ATRA for 5 days then cells were stained with Giemsa. Non-ATRA HL-60 cells (**A**) and ATRA treated HL-60 cells (**B**). $5x10^5$ differentiated HL60 cells in RPMI were incubated for120h then viability was assessed every 24h (**C**) (Error Bars: SD n = 3)

3.3.3. Role of differentiation in phagocytic efficiency

To investigate the influence of differentiation on the phagocytic capacity of HL60 cells, primitive (pHL60) and differentiated (dHL60) cells were infected with exponential growing BCG (**Figure 3.3**). The numbers of phagocytosed and phagocyte-associated BCG were significantly higher in differentiated HL60 cells (**A**). After differentiation there were threefold and tenfold increases in the CFU

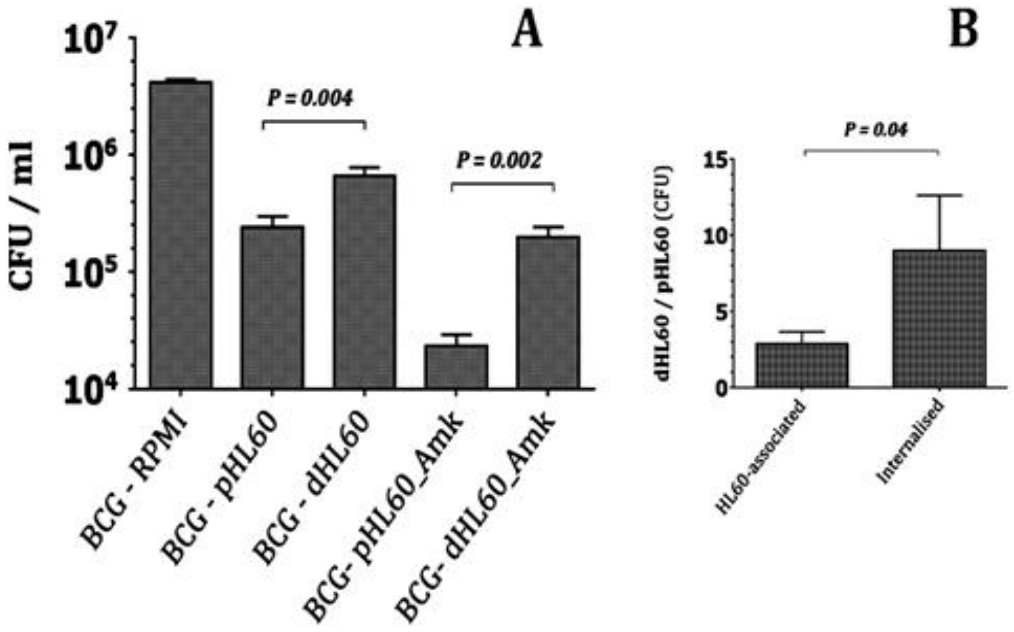


Figure 3.3 Differentiated HL60 cells are more efficient in capturing and internalising BCG. Exponentially growing BCG bacilli at MOI=9.2±3 were incubated with 5.2±1.4 x 10^5 HL60 cells (dHL60: differentiated and pHL60: primitive) in RPMI for 18h. Amikacin was used to kill extracellular bacilli. Three washes with RPMI were applied to remove extracellular bacteria and antibiotics. The CFU of samples (**A**) and the difference in CFU count between associated and internalised bacilli with regard to the differentiation of HL60 cells (**B**). Paired t-test was done to determine statistical significance of difference between conditions. (Error Bars: SD n = 3)

counts for phagocyte-associated and internalised BCG respectively (**B**). Therefore, the use of differentiated HL60 cells was established routinely in subsequent experiments.

3.3.4. Effect of centrifugation on uptake of mycobacteria

Centrifugation can press phagocyte and mycobacteria therefore it was applied to

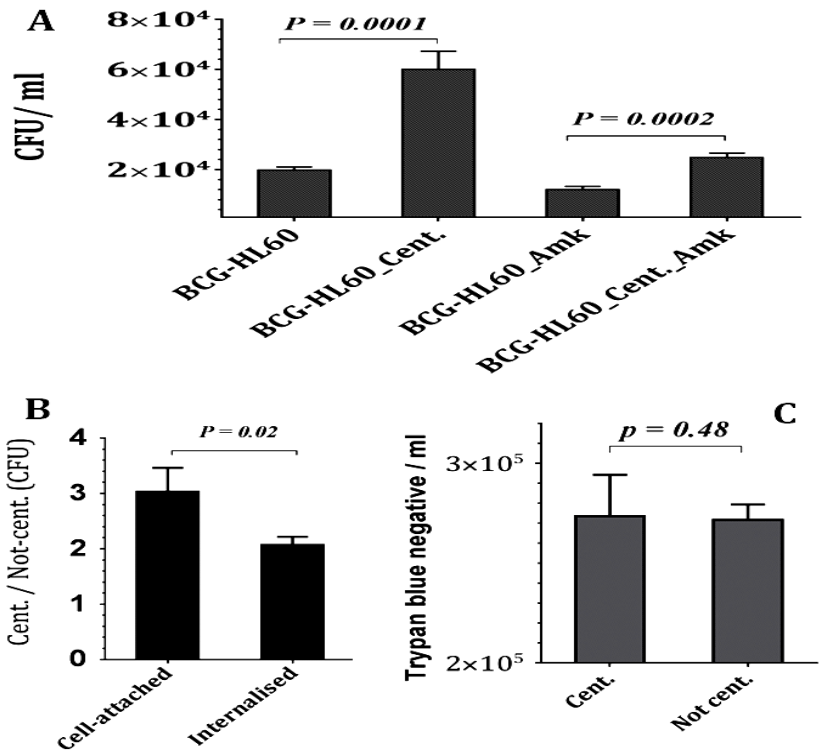


Figure 3.4 Centrifugation increases phagocyte-bound and internalised BCG in HL60 cells. In RPMI 8.4 x 10^5 HL60 cells were incubated with BCG at MOI=5.5 with and without centrifugation (1000xg for 10min) for 18h. Three washes with RPMI were done. The CFU of samples (**A**) the difference in CFU counts between associated and internalised bacilli with regard to the centrifugation of samples (**B**) and Trypan blue test of viability for HL60 cells after incubation (**C**).

Paired t-test was done to determine statistical significance of difference between conditions. (Error Bars: SD n = 3)

improve their interaction. A significant increase in the number of phagocyte-associated and internalised BCG cells occurred when samples were centrifuged at 1000xg for 10 minutes as compared to non-centrifuged samples (**Figure 3.4A**).After centrifugation of samples there was a 3-fold increase in the number of HL60-associated bacilli while the number of internalised bacilli increased by twofold and this difference was noted to be statistically significant (**Figure 3.4B**). The viability of infected HL60 cells was not significantly different after centrifugation and 18h incubation (**Figure 3.4C**). Therefore, prior to incubation, centrifugation was applied in subsequent experiments to increase the proportion of intracellular bacilli.

3.3.5. Removing extracellular bacilli via multiple washings

The initial target in this study was to investigate intracellular Mtb signals only. To this end a series of washes were applied to the bacteria – phagocyte suspensions in order to remove extracellular bacteria. This was done by separation of non-ingested bacteria through differential centrifugation and washing. The results showed a significant difference in the number of removed bacilli between first and second and between second and third washes but no further significant differences between subsequent washes (**Figure 3.5**). Thus, 98.5% of the bacilli were removed after the third wash and no significant effect was achieved with further washes. Trypan blue demonstrated a gradual decrease in neutrophil viability corresponding to the number of washes.

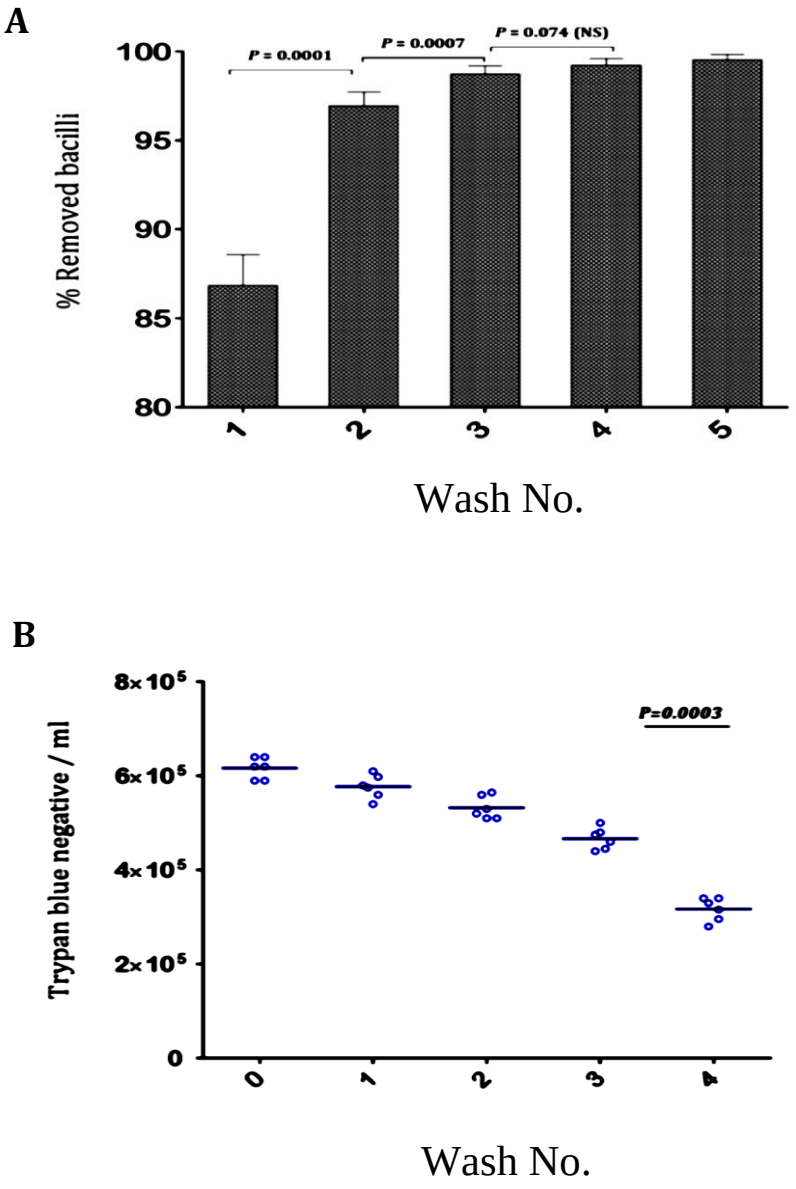


Figure 3.5 T**he Majority of extracellular bacilli are removed after 3 washes**. In RPMI 7x10^5 HL60 cells were incubated with BCG at MOI=7.5 for 18h. Each sample was washed either 1 to 5 times with pre-warmed RPMI. CFU count for BCG (**A**) and viability test for HL60 cells (**B**) was done to each wash. Paired t-test was done to determine statistical significant differences between washes.
(Error Bars: SD n = 3)

3.3.6. The viability of HL60 cells exposed to different MOIs of BCG

To determine the most suitable MOI for infection, samples were prepared with different ratios of mycobacteria to neutrophils. The target was to get the highest number of intracellular bacteria with minimum death of infected neutrophils. The results (**Figure 3.6**) showed that the viability of HL60 cells at all MOIs was above 80% until 12h then it decreased to 60% for MOI = 50 and 100 while remaining above 75% for MOI = 5 and 10. Therefore, for further experiments an MOI of 5 to 10 was prepared.

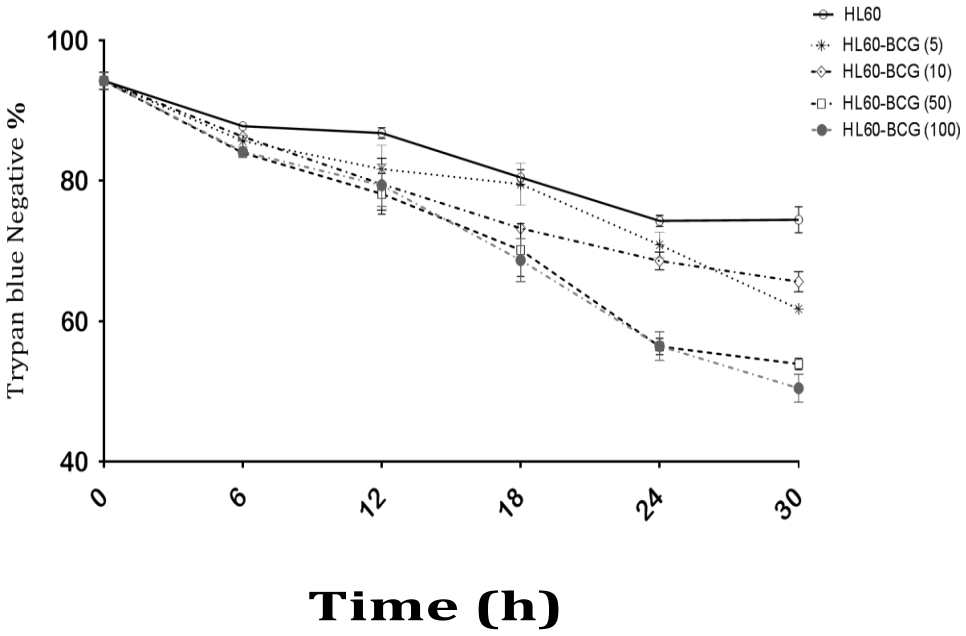


Figure 3.6 HL60 cells show higher viability at lower MOI.
$5x10^5$ ±2HL60 cells incubated in RPMI with BCG at MOI = 5.3, 11.5, 56.6 or 95.2 for 18h. (Error Bars: SD n = 3)

3.3.7. Time-course of the HL60-BCG interaction over 24h

HL60 cells were incubated with BCG to investigate the uptake and phagocyte adhering of BCG within 24h (**Figure 3.7**). The CFU counts revealed a pattern of about 3 fold decrease from 2h to 24h in BCG-HL60 sample. The ANOVA test to five consecutive time points demonstrated an overall statistically significant difference among them. At 2h the ratios of HL60-associated and internalised bacilli relative to BCG-HL60 were about 10% and 1% respectively. From 2h to 24h the successive time points in sample of HL60-associated displayed a gradual dropping down to 10 fold decrease in CFU counts. Notwithstanding, the internalised bacilli were increased from 2h to 4h and remained at a plateau up to 12h then showed a substantial decrease of nearly 10 fold at 24h.

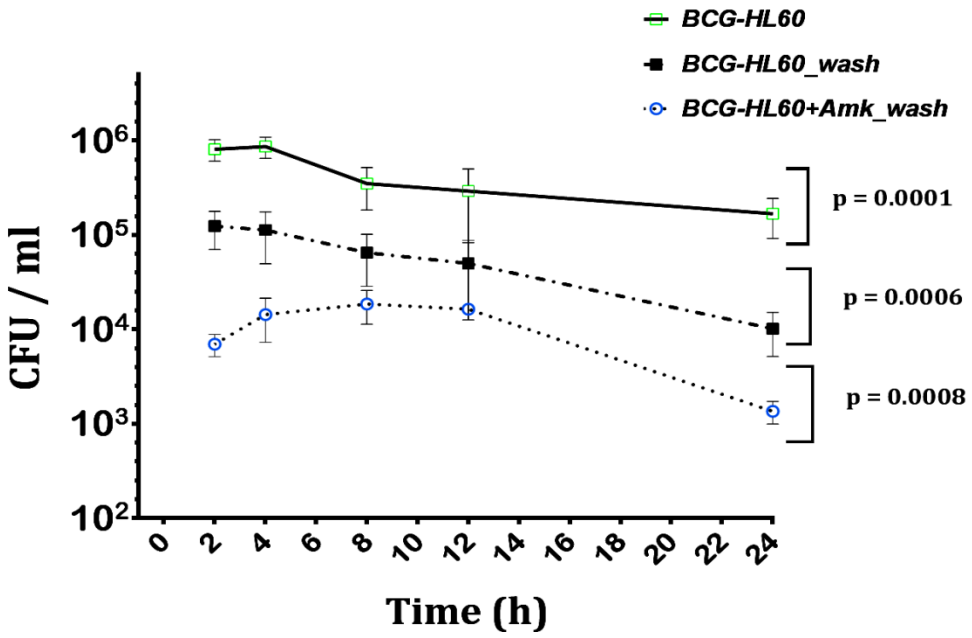


Figure 3.7 CFU counts of BCG during 24h incubation with HL60 cells. In RPMI 6 x 10^5 HL60 cells were incubated with BCG at MOI = 10.6±1.2 for 2, 4, 8, 12 and 24 hours. One set of samples was treated with amikacin for 2h. Statistical analysis was performed to each sample by taking all time points into One Way ANOVA. (Error Bars: SD n = 4)

3.3.8. H37Rv uptake into HL60 cells

HL60 cells were incubated with H37Rv in order to assess their ability to phagocytosis pathogenic MTBC bacilli. The CFU counts showed that less than 5% of H37Rv were internalised after 2h which was about half of total counts for HL60-bound bacilli (**figure 3.8**) while no uptake took place at 4˚C.

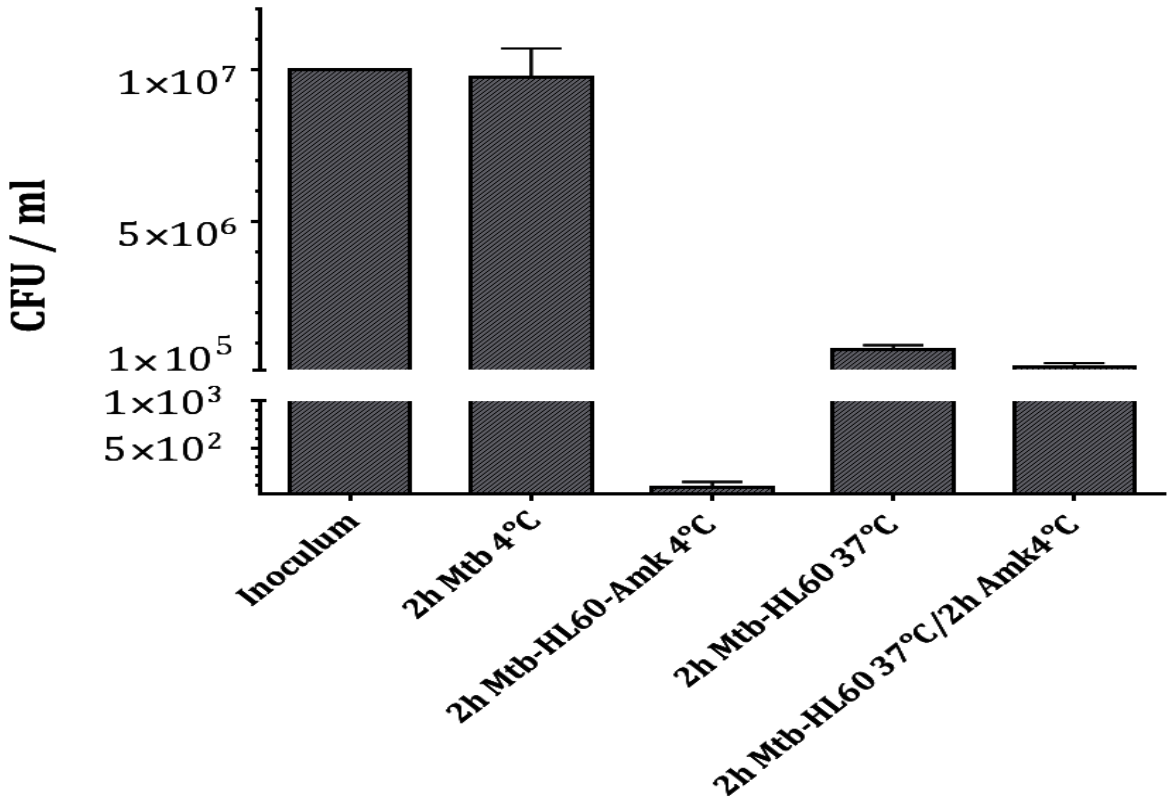


Figure 3.8 Phagocytosis of H37Rv during 2h of HL60 cell infection. 1.2 x 10^6 HL60 cells were incubated with H37Rv at MOI =9.2 in RPMI for 2h at either 4˚C or 37˚C with or without amikacin.
(Error bars: SD n = 3)

3.3.9. Time-course of the HL60-H37Rv interaction over 120h

At this point, for reasons discussed fully in chapter 4, it was realised that monitoring of the interactions between neutrophils and both intracellular and extracellular bacilli was desirable and that mycobacteria within NETs would be likely to represent the true situation in vivo. Therefore H37Rv was incubated with IFN-γ treated-HL60 cells in RPMI with no washing and no amikacin treatment for 120h (**Figure 3.9**). The results demonstrated stable CFU counts in all RPMI samples. While for HL60 the counts were stable up to 24h and then significantly declined between 24h and 120h.

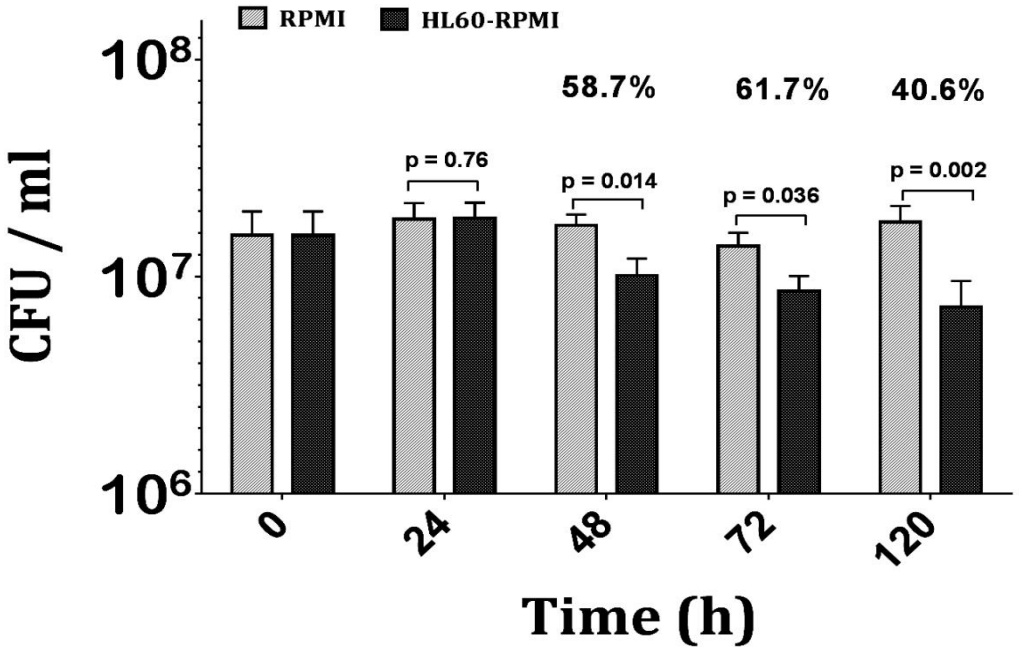


Figure 3.9 RPMI and HL60 cells do not support the growth of H37Rv. In 5 ml RPMI $8x10^6$ IFN-γ-treated HL60 cells were incubated with H37Rv at MOI = 10.5. Samples for CFU count were taken at every 24h without removing extracellular bacilli. Percentages are H37Rv CFU counts with HL60 cells relative to H37Rv alone. Paired t-test was done to determine statistically significant difference between conditions.
(Error Bar = SD n = 4)

3.3.10. The oxidative burst of HL60 cells on encounter with H37Rv

As evidenced in the previous experiments the phagocytosis of Mtb starts in the first 2h. To determine when the response of polymorphs against mycobacterial they encounter occurs, the oxidative burst was measured in HL60 cells incubated with H37Rv for 150 minutes using INT tetrazolium salt (**Figure 3.10**).The OD measurements were constant up to 60 min. While, the OD was relatively unchanging (OD range = 0.16 to 0.17) in the absence of bacteria it increased significantly between 60 and 90 min in the HL60-H37Rv samples.

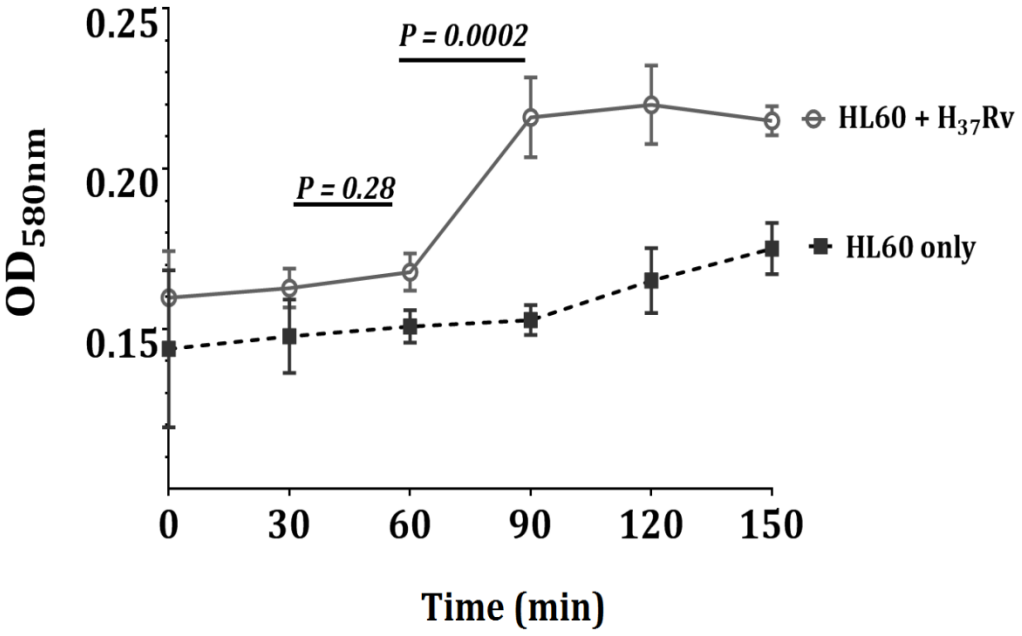


Figure 3.10 Concentration of ROS increases in HL60 cells after 60 minutes of infection with H37Rv. In 200µl Phenol red free RPMI 2.6 x 10^5 IFN-γ treated HL60 cells infected with H37Rv at MOI = 5.7 then oxidative burst was measured using INT. Paired t-test was done to determine statistical significant differences between time points.
(Error Bars: SD n = 4)

3.3.11. Isolation of primary human neutrophils

Giemsa staining was performed to determine the efficiency of the method used to isolate primary human neutrophils (**Figure 3.11**); 93.5% of cells in the sample shown were neutrophils with segmented nuclei and the remainders were mainly monocytes with a one nucleus. Serial viability tests on isolated cells showed 75% trypan blue negativity at 18h but 100% positivity after 36h.

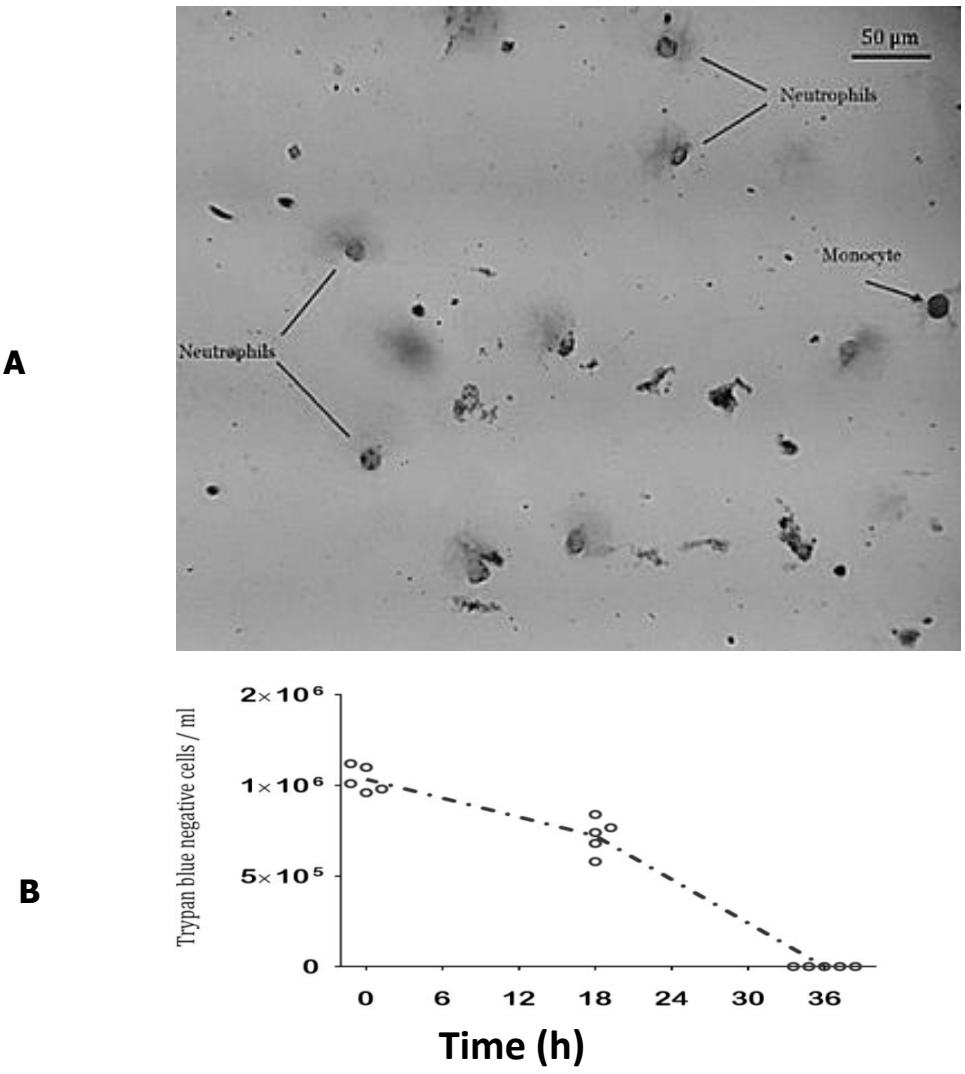


Figure 3.11 Human neutrophils lose viability 36h after isolation.
Giemsa stain of isolated human neutrophils (**A**). Trypan blue test (**B**). After isolation and IFN-γ treatment, $1x10^6$ primary human neutrophils were incubated for 36h in 1ml RPMI. (n = 5)

3.3.12. Interaction of H37Rv with primary human neutrophils

Although HL60 cells are considered as a standard model for granulocytes, they may not fully represent primary human neutrophils (Timm et al., 2008). To investigate the differences and similarities, isolated neutrophils were challenged with H37Rv to investigate their phagocytosis and bactericidal abilities compared to HL60 cells. Primary human neutrophils from four different donors were primed with IFN-γ and incubated with H37Rv (**Figure 3.12**). The CFU counts demonstrated a modest but significant decline at 24h compared to the RMPI samples.

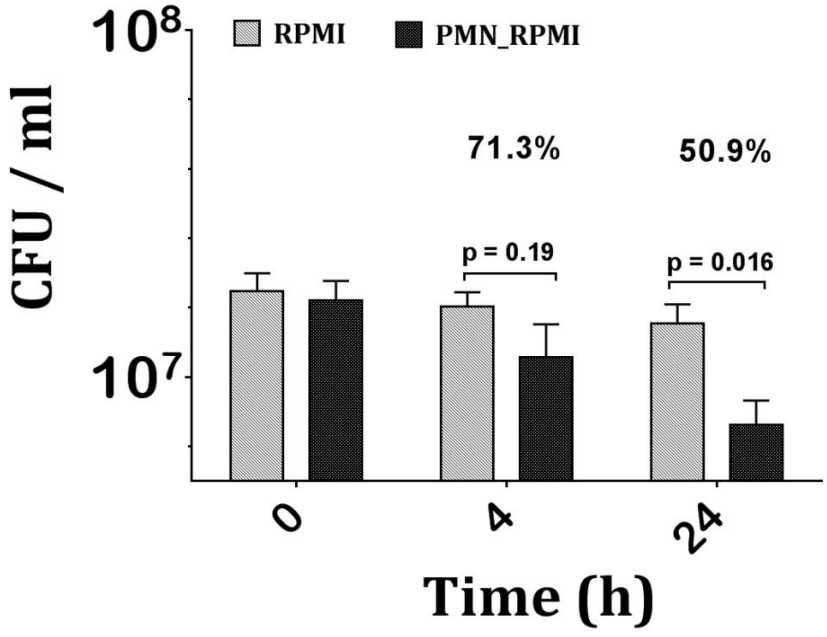


Figure 3.12 Primary human neutrophils reduce the number of H37Rv bacilli within 24h of infection. 2 x 10^6 / ml primary human neutrophils were incubated with H37Rv at MOI = 10 in 10 ml RPMI for 24 hours. Percentages are H37Rv CFU counts with HL60 cells relative to H37Rv alone. Paired t-test was done to determine statistical significant difference between conditions. Blood samples were collected from 4 donors on 5 separate occasions. Error Bars: SEM

3.4. Discussion

This series of experiments was performed prior to establishing a method of RNA extraction for the transcriptional studies. Therefore, the design of experiments was done aimed at stepwise characterisation of the in vitro interactions of mycobacteria and neutrophils.

3.4.1. Effect of amikacin on reducing mycobacterial CFU counts

At the start this study was intended to focus on intracellular bacteria for which an anti-mycobacterial drug was used to kill extracellular bacteria. The idea of using antibiotic came after several attempts to confirm the internalisation of bacteria.

The use of amikacin was adopted from a study in which 200µg amikacin was applied for one hour to kill extracellular mycobacteria associated with infected macrophages (Fortune *et al.*, 2004). However, the results in the current study revealed that a proportion of bacilli even after amikacin exposure were able to produce CFU. To tackle the exposure time, the incubation time was extended to 2h and 4h which demonstrated greater bactericidal results after 2h with no significant increase in 4h. Considering the view that the time for using neutrophils in experiments was very limited, the appropriate timing for amikacin in this study was set at 2h exposure.

3.4.2. Differentiation of HL60 cells

Prior to using primary human neutrophils it was envisaged that the variability in blood samples and delicacy of isolated polymorphs in vitro could make the results prone to irreproducibility. With such difficulties and considering the short life time of primary neutrophils found subsequently, HL60 cells were used in the preliminary work. Several compounds are known to induce differentiation in HL60 cells; all-

trans retinoic acid (ATRA) was used in this study. ATRA was found to give quicker results compared to other reagents. For instance, DMSO treatment needed 7 days to achieve the same level of differentiations seen at 5 days with ATRA (data not shown). ATRA-treated cells were well differentiated polymorphs with multi-lobulated nuclei and viability sustained for 3 days after differentiation (**Figure 3.2C**). This is considerably longer than the survival of primary human neutrophils after isolation (**Figure 3.8B**) and may indicate a relative lack of maturity.

3.4.3. Role of differentiation in phagocytic efficiency

The CFU counts for phagocyte-associated and internalised BCG were significantly higher in differentiated compared to non-differentiated HL60 cells (**Figure 3.3**). This might be due to the more abundant expression of receptors on neutrophils after maturation. Indeed, during ATRA-induced differentiation more than half of HL60 cells become Fcγ receptor-positive which in turn increases the capability of cells for phagocytosis (Nordenfelt *et al.*, 2009). But the larger proportion of internalised bacilli relative to HL60-associated signposts a greater effect on phagocytosis than on binding of bacteria. This might be due to profound expression of key receptors or in general the overall changes occurred during differentiation are more in favour of phagocytosis than binding to extracellular bacilli.

3.4.4. Effect of centrifugation on uptake of mycobacteria

This method was adopted from a study in which centrifugation was used with macrophage infection *(*Rohde *et al.,* 2007*)*. The attachment properties of neutrophils in vitro are different from macrophages as monocytes become surface adherent while both naïve and differentiated neutrophils remain floating in the medium. Therefore, centrifugation could accelerate the settlement of neutrophils

and subsequently bacilli will be pelleted on them. In this regard, it is essential to consider the difference in size and sedimentation speed of cells. Neutrophils, possess buoyant density of 1.1 g/ml (Pember *et al.*, 1983) compared to mycobacteria with buoyant density of 1.01 g/ml (Hertog *et al.*, 2009). Thus, in addition to minimising the depth of suspension an appropriate centrifugation speed is required to achieve mycobacterial sedimentation with consideration of neutrophils fragility. Accordingly, HL60 cells mixed with BCG were centrifuged at 1000xg for 10 minutes (**Figure 3.4**).

The volume of suspension was made as 1ml in a 50ml conical tube that created a 1cm distance between top and bottom of suspension aimed at decreasing the time of sedimentation for BCG. A larger proportion of HL60-associated to internalised bacteria after centrifugation indicated that bacterial binding was more affected than phagocytosis. This was different to the previous observations with differentiated HL60 cells. This higher level of binding might reflect that the capacity for phagocytosis in differentiated neutrophils was close to saturated. It is also possible that centrifuge-effected contact to phagocyte receptors might reduce internalisation. An obvious concern here is that such contact could be destructive to neutrophils. Nonetheless, the trypan blue assessment of neutrophils indicated no significant differences in the viability between centrifuged and non-centrifuged samples.

3.4.5. Removing extracellular bacilli via multiple washings

Three washes were determined to be optimum. A viability of less than 80% is considered unacceptable for phagocyte studies in general and neutrophils in particular(Fleck *et al.*, 2005). Therefore, a trade-off of washing times was needed in which maximum extracellular bacilli would be removed with least neutrophils being lost. Thus a maximum of three washes were applied for the subsequent experiments to remove extracellular bacilli.

3.4.6. Viability of HL60 cells under different MOIs

Another option to increase the number of intracellular bacteria was to increase the multiplicity of infection (MOI). MOIs of 5 and 10 were found to be more suitable ratios for neutrophil infection in this study as the number of trypan blue negative HL60 cells was highest compared to other ratios in 24h of infection. Beside concerns on the viability of phagocytes, choosing lower MOI was advantageous because of increased likelihood of bacterial clumping with increasing MOIs. This could affect both the CFU counts and the interaction with the phagocyte. Moreover, clumped mycobacteria may experience a different phagocytic pathway than individual isolated bacteria, especially over longer exposure times (Schuller *et al.*, 2001).

3.4.7. Time-course of the HL60-BCG interaction over 24h

Similar to earlier experiments, the overall number of internalised BCG was about 1%. However, this level is low compared to macrophage studies in which the phagocytosis of BCG by THP-1 cells has been reported to be between 30% (Reyrat *et al.*, 1996) and 70% (Soualhine *et al.*, 2007). An overall 10% CFU count for washed relative to unwashed samples indicates that the majority of bacilli were bound to the cell membrane of HL60 cells and only a small proportion were internalised.

The CFU counts showed a significant decline in unwashed samples from 2h to 24h. This could indicate a modest capacity of HL60 cells in killing of BCG or perhaps such small reduction in colony forming ability of BCG might be due to a gradually increasing the amount of NET formation and consequently increasing the likelihood of bacterial clumping. Looking into other samples, the trend of CFU counts supports the view that some killing took place. In both samples of washed and amikacin treated preparations, the number of BCG reduced from 2h to 24h at

a statistically significant level though with differing patterns of reduction. In washed samples a constant reduction at each time point indicates a continuous process of BCG elimination which might be explained by two processes; a combination of continued internalisation and killing. Continued internalisation can be recognised from the increased CFU counts in amikacin treated samples at 4h, 8h and 12h. The marked reduction of CFU in the amikacin treated samples at 24h seems to be consistent with killing of bacilli by HL60 cells.

There is no published data on HL60 and BCG interaction. Nonetheless, this result seems to be similar to the pattern seen in macrophage studies. Results from macrophage studies indicate intracellular killing of BCG (Jordao *et al.*, 2008).

3.4.8. H37Rv uptake into HL60 cells

The phagocytosis of H37Rv started in the first two hours though at levels less than 1%, similar to BCG. However, the percentage of internalised relative to HL60-associated bacilli was about 60% which reflects a higher efficiency of HL60 cells in ingesting H37Rv than BCG at 2h.

The attempt to apply low temperature in this experiment to stop phagocytosis is worth considering because no CFU were obtained after 2h amikacin exposure at 4˚C. This further indicates that the antibacterial activity of amikacin is preserved within 4˚C – 37˚C while, the process of internalisation was stopped at 4˚C.

3.4.9. Time-course of the HL60-H37Rv interaction over 120h

In order to investigate the survival of Mtb in neutrophils, H37Rv was incubated with IFN-γ activated HL60 cells for 120h. The presence of this cytokine has been found to increase the half-life of neutrophils in vitro (Tak et al., 2013, Hilda *et al.*, 2012).

Although many bacilli are internalised by HL60 cells early during co-incubation, exposure to the hostile environment resulting from NET formation may be an equally significant part of the neutrophil-mycobacterium interaction. This view combined with the fact that the composite intracellular and extracellular interaction of mycobacteria with neutrophils seems to be more representative of in vivo situation, led us to consider both internalised and extracellular bacilli as equally important. Indeed, since the reason for performing these infection experiments was to establish a neutrophil – mycobacteria system from which bacterial RNA could be collected for transcriptome studies, it was appreciated that the presence of amikacin killed extracellular bacilli in the preparation might distort the picture obtained.

Therefore, it was decided to take both intracellular and extracellular bacilli into account and the need to apply antibiotic was removed. This development also reflected our recognition of the significance of NETs in vivo.

3.4.10. Oxidative burst of HL60 cells on encounter with H37Rv

A key neutrophil response to invading organisms is activation of the respiratory burst. The magnitude of this response can be measured using tetrazolium salts (Tan *et al.*, 2000). There was a significant rise in formazan formation in HL60 cells after 90 minutes encounter with H37Rv (**Figure3.10**). Thus HL60 cells respond to Mtb within the first hour of infection. Such early response indicates that differentiated HL60 cells have capacity for induction of respiratory burst similar to primary human neutrophils. This is consistent with likely up-regulation of the Fcγ receptor CD64 (FcγRI) in IFN-γ activated HL60 cells, a response associated with the respiratory burst in primary H37Rv-infected neutrophils (Hilda *et al.*, 2012). Demonstration of the oxidative burst here supports the view that HL60 cells achieve killing of H37Rv bacilli as also appear to be the case for for primary human neutrophils (Figure 3.12).

3.4.11. Isolation of primary human neutrophils

The use of primary human neutrophils was intended after the preliminary experiments with HL60 cells. Giemsa staining (Figure 3.11A) demonstrated that 93.5% of total isolated cells were polymorphs. Although a small proportion of isolated cells appeared to be either monocytes (about 6%) the method of neutrophil isolation was satisfactory and loss of viability within 36h was consistent with their recognised half-life of neutrophils.

3.4.12. Interaction of H37Rv with primary human neutrophils

Infection of primary human neutrophils with H37Rv for 4h and 24h (Figure 3.12) appeared comparable to that observed with HL60 cells. The CFU counts of H37Rv were found to decrease after 24h of encounter with IFN-γ activated primary human polymorphs. Reduction of CFU counts with HL60 cells did not reach statistical significance until 48h. This indicates a more rapid response of IFN-γ activated primary human neutrophils and may also reflect quicker NET formation.

Overall, results from this series of experiments indicate the suitability of HL60 cells as comparable to primary human neutrophils in this study. HL60 cells were therefore useful in establishing the experimental setup for the transcript analyses discussed in the next chapter.

3.5. Concluding remarks

- Stages in the HL60 neutrophil–MTBC interaction were elucidated by: a) differentiation of cells with ATRA; b) treatment with amikacin; c) centrifugation of phagocyte-bacteria suspensions; d) removal extracellular bacilli via washing; and e) activation of neutrophils with IFN-γ.
- Uptake of BCG and H37Rv started within 2h with HL60 cells

- The respiratory burst in HL60 cells starts after 60 minutes encounter with H37Rv.
- The CFU counts of H37Rv decreased significantly at 24h with primary human neutrophils and at 48h with HL60 cells.
- The growth of BCG and H37Rv are not supported by RPMI for at least 120h.
- Differentiated HL60 cells can be considered as a model for the study of interactions between human neutrophils and the MTBC.

CHAPTER FOUR

Development of conditions for studying transcriptional responses of *Mycobacterium tuberculosis* provoked by the encounter with Neutrophils *in vitro*

4.1. Introduction

The precision and value of transcriptome assessments are profoundly influenced by the quality and quantity of initial RNA isolated. Given the time and financial investment required, it is vital that preliminary evidence of high quality RNA and reproducible signals are obtained for biological and diagnostic applications such as qPCR, microarrays and next generation sequencing (NGS) (Fleige *et al.*, 2006, Akhtar *et al.*, 2011).

4.1.1. Mycobacterial RNA processing

RNA is under regulation of various enzymes that control its content and modulate its spatial and temporal expression. Nearly 20% of dry weight of prokaryotes is RNA. There are multiple species of this molecule with different functions (Licatalosi *et al*., 2010). In prokaryotes, ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA) undergo numerous modification and cleavage events during maturation and degradation. In this regard, mRNA as the mediator between DNA genome and protein synthesis relies on tRNA and rRNA. Additionally, non-coding RNAs can also be involved in such processes and their influence on gene expression has been recognised in many prokaryotic species (Evguenieva-Hackenberg *et al*., 2011).

In general the half-life of mRNA is short relative to bacterial doubling time, while stable RNA species including tRNA and rRNA can remain functional for a number of generations. One of the most perplexing features in the metabolism of bacterial RNA is that, in order to be functional, some regulatory small RNAs (sRNA) and mRNAs must be processed and ultimately degraded by the same enzymes that catalyze rRNA and tRNA maturation and degradation (Mackie, 2013).

4.1.1.1. RNA turnover in Mtb

The firm control of mRNA enables bacteria to accomplish selective transcription and consequent decoding of genetic information appropriate to certain physiological states. The coupled nature of gene transcription and translation in bacteria regulates the access of RNA degrading enzymes to individual transcripts and thus influences the turnover rate of mRNA (Kaberdin et al., 2011).

An increase in abundance of mRNA for a significant proportion of the genome during bacterial responses to environmental stimuli may be achieved either by increasing the transcription rate or by decreasing the degradation rate and the converse applies to reduced transcript levels (Shalem et al., 2008). In response to environmental stimuli, mycobacteria change their metabolic activities. One way to make this change is via alteration of gene expression. These adaptive processes in the metabolism of RNA expressed by responding mycobacteria yield a differential influence on the steady state levels of various transcripts, on the biosynthesis of ribosomes and on RNase E/G levels. A variety of elements and enzymes participate in mRNA metabolism of which RNase E is particularly prominent (Zeller *et al.*, 2007).

Mtb demonstrates a slower mRNA degradation rate than the majority of other bacteria. A number of mechanisms may contribute to this including apparently lower abundance of or slower kinetics of RNA degrading enzymes in Mtb (Rustad *et al.*, 2013). Cleavage site specificities of enzymes together with post-transcriptional modifications can also directly result in a slower global rate of mRNA degradation. For instance, the RNase E enzyme of *E. coli* has a homolog gene in Mtb which despite its kinetically equivalent function displays a limited binding capacity to cleavage sites in Mtb. Thus the impact of degradation on transcript levels must be taken into account in the development of valid models of the regulatory networks in Mtb (Kovacs *et al.*, 2005).

4.1.2. RNA isolation from Mtb

4.1.2.1. Extraction of mycobacterial RNA from infected neutrophils

The currently proposed mycobacterial RNA isolation methods include enzymatic cell wall lysis and mechanical shearing by pressure, sonication, bead beating or nitrogen decompression. Each of these methods has potential deficiencies. Enzymatic lysis is protracted and may compromise the integrity of RNA while mechanical shearing and nitrogen decompression need expensive equipment and may generate aerosols (Akhtar *et al.*, 2011, Mangan *et al.*, 1997).

It poses a serious challenge to design a system of mycobacteria – neutrophil encounter in vitro from which a representative RNA sample could be collected for transcriptional study. The greatest task appears to be controlling the reproducibility of experiments due to the variation between biological replicates. There are a number of factors causing such variability; first of all, neutrophils used in experiment can be different due to the difference in their donors. Although, the variability in neutrophil source can be minimised using a neutrophil cell line for example HL60 cells, there is still a little likelihood of variation in the results. In addition, the discrepancy may be originated from variability in inoculum used in subsequent biological replicates. This is mainly attributed to the experimental settings especially for a slow growing bacterium such as Mtb. In this regard, the interval between two experiments may take days or weeks during which controlling the growth condition is hard to achieve.

The fragility of RNA in bacteria and a fast turnover of their transcripts add to the difficulty of RNA extraction from Mtb in infected neutrophils. Considering all these factors and the presence of all variations inherent to the experiments it seems to be a cumbersome process demanding for a careful design in order to maintain the reproducibility and attain the biologically meaningful results.

4.1.2.2. Purification and quantifying extracted RNA

One approach to the removal of host contaminants in RNA-directed phagocyte–mycobacteria studies is to use chaotropic substances such as guanidium thiocyanate (GTC) which selectively lyse eukaryotic cells and inactivate RNases and other enzymes while leaving the mycobacteria intact. Trizol provides a convenient formulation for RNA purification in conjunction with cell lysis by Fast-prep treatment (Alli. *et al.*, 2009). The residual proteins and other adventitious substances can be removed through several steps of washing with chloroform and dilutions of ethanol.

4.1.3. Aims and objectives

The overall aim was to establish an appropriate protocol for determining the transcriptional responses of Mtb to encounter with neutrophils in vitro and a preliminary view of the relationship of this response to the sputum transcriptome.
The objectives were:

- To establish a reliable method of RNA extraction from the neutrophil experiments
- To determine the transcriptional responses of Mtb in various physiological conditions when exposed to neutrophils in vitro.
- To determine the effect of preparative steps such as centrifugation and RPMI exposure on selected transcripts

4.2. METHODS

4.2.1. Reagents and chemicals

All chemicals and reagents were obtained from QIAGEN (Crawley, West Sussex, UK), Sigma-Aldrich (Poole, Dorset, UK) or Fisher Scientific (Loughborough, Leicestershire, UK), unless otherwise stated.

4.2.1.1. 5M Guanidinium thiocyanate (GTC) solution

GTC it is a chaotropic substance which denatures proteins and therefore stabilises nucleic acids (Mason *et al.*, 2003). To prepare 500 ml of 5M GTC the following materials were added in a Duran bottle:

- Guanidine thiocyanate (Promega, Hampshire, UK) 294.4g
- N-lauryl Sarcosine 2.5g
- Tri-Sodium Citrate 12.5g
- Tween 80 10% (w/v) 25 ml

For preparation of 4M GTC the amount of guanidine thiocyanate was 236.3 g in 500ml water.

200ml distilled water was added and the materials were dissolved via incubation at 37°C overnight shaking at 100 RPM. Distilled water was added to the solution to get the final volume of 500ml.

Aliquots of 20ml 5M GTC were made in 50ml polypropylene conical tubes. The aliquots were kept away from light and heat. Prior to use, 140µl of β-mercaptoethanol was added to each aliquot to reach a final concentration of 7µl / ml.

4.2.1.2. DNase I solution

DNase I powder was dissolved in 550µl DNase-RNase-Free water. Aliquots of 30µl were made and stored at -20˚C. When required 10µl of **DNase I** solution was mixed with 70µl **Buffer RDD** which was provided in the Turbo-DNase kit.

4.2.1.3. Rotor-Gene SYBR Green (2X) master mix

A 2X SYBR Green master mix was provided by the manufacturer containing DNA polymerase isolated from *Thermus aquaticus,* SYBR Green buffer (TrisCl, KCl, NH_4Cl, $MgCl_2$ and Q-Bond), dNTP mix, and SYBR Green I.

4.2.1.4. Preparing qPCR reactions

For each reaction 12.5µl SYBR Green (2x) master mix was added to 0.1ml PCR tube then 1µl of 5mM forward and 1µl of 5mM reverse primer was mixed. The content was brought to 24µl by adding 9.5µl RNase-DNase free water. 1µl of sample cDNA was added to get the final volume of 25µl.

4.2.1.5. Target Genes and primers

The expression levels of a number of genes were investigated using qPCR. Five genes were chosen from the list of differentially expressed genes in sputum transcriptome profile including three up regulated (*icl1*, *tgs1* and *hspX*) and two down regulated (*mce1A* and *rpsL*). *16s rRNA* and *sigA* were also quantified for their expression as housekeeping genes. The list of primers including their sequence, length and melting temperature is available in the appendices.

4.2.2. General methods

The overall process of RNA extraction is displayed in Figure 4.1

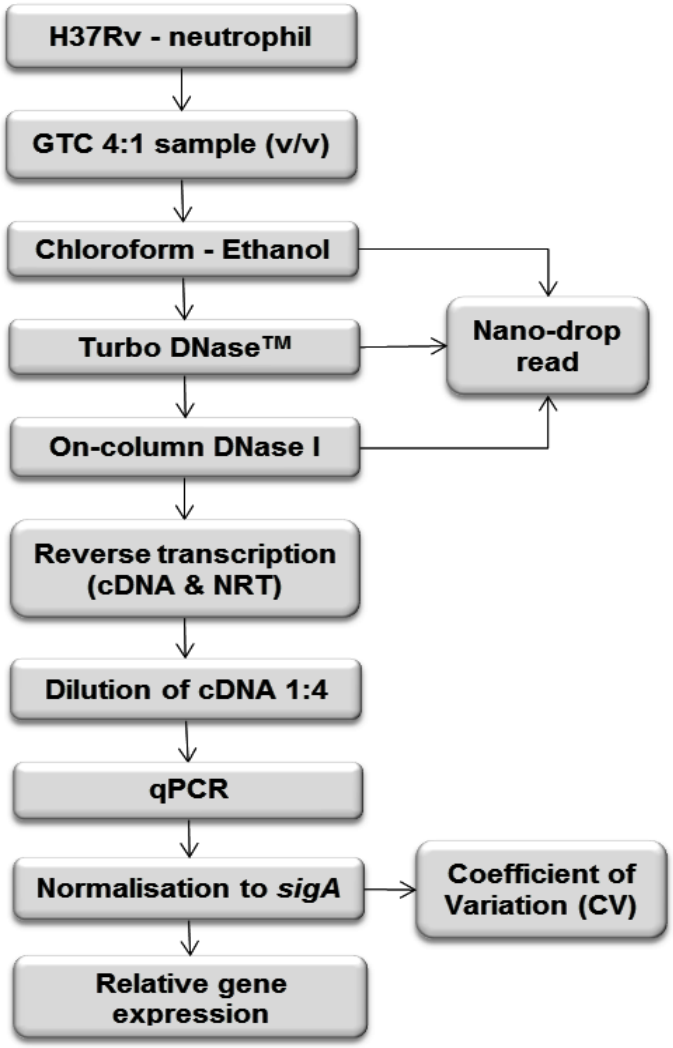


Figure 4.1 Overview of RNA extraction process from H37Rv in infected neutrophils.

4.2.2.1. RNA stabilisation

To stabilize RNA in the process of isolation from mycobacteria in infected neutrophils, 5ml of sample was added to 20ml of 5M GTC 1:4 (*v/v*) in a 50ml propylene conical tube. The viscosity of the suspension was reduced by vortexing at 2000 RPM for 30 seconds to break up the long strands of host DNA. The sample was then centrifuged at 1500xg for 20 minutes then the supernatant was removed. The pellet was re-suspended in 20ml of 4M GTC and centrifuged again at 1500xg for 20 minutes and supernatant discarded. (Note: when mycobacteria alone were processed the second GTC treatment was not performed). The pellet was then re-suspended in 1ml Trizol and transferred into lysing matrix B tube for ribolysing in the Fast-prep set at 6.5 m/s for 45 seconds. Tubes were placed on ice for 10 minutes before the subsequent chloroform – ethanol treatment.

4.2.2.2. Chloroform – Ethanol treatment

After ribolysing and maintenance on ice, 200µl chloroform and 200µl water were added to the tube and thoroughly mixed by vortexing as above. The sample was then centrifuged at 13000xg at 4˚C (centrifugation condition A) for 15 minutes. The upper aqueous phase was taken into a 1.5ml microfuge tube and an equivalent volume of chloroform was added and mixed by vortexing. Following centrifugation (A) for 5 minutes the upper aqueous phase was transferred into a new microfuge tube. Isopropanol at 0.8 v/v was added and mixed gently then the sample was frozen at -80˚C for 2h to precipitation the RNA. After thawing the sample was centrifuged at 17000xg for 15 minutes at 4˚C. The supernatant discarded and the pellet re-suspended in 1ml of 70% ethanol then subjected to centrifugation (A) for 5 minutes. After removing supernatant 200µl of 95% ethanol was added and the preparation again subjected to centrifugation (A) for 5 minutes, the supernatant discarded and the pellet dried at room temperature for 10 minutes. The pellet was re-suspended in 50µl RNase free H_2O.

4.2.2.3. Turbo DNase treatment

5µl Turbo DNase buffer and 1µl Turbo DNase reagent were added to the extracted RNA sample and mixed gently. The sample was incubated at 37˚C for 30 minutes in a heat block then 1µl Turbo DNase was added and incubated for a further 30 minutes. 10µl inactivation reagent was added then centrifuged (A) for 5 minutes. To avoid carryover of the inactivation reagent, the aqueous phase was transferred to a new tube and measured for RNA content.

4.2.2.4. On-column DNase digestion

For further refinement of extracted RNA the sample was treated with on column DNaseI digestion.

The materials provided in the kit included:

- Mini-column
- Collection tube
- Buffer RLT
- 100% Ethanol
- Buffer RW1
- Buffer RPE

The volume of sample was increased to 100µl with RNase free water. 350µl buffer RLT and 250µl 100% ethanol were added and the sample transferred to the mini-column and centrifuged at 13000xg for 20 seconds. The flow through was discarded and a further 350µl buffer RW1 added and the preparation centrifuged at 13000xg for 20 seconds. After removing the flow through 80µl of DNase I solution was added and incubated for 15 minutes at room temperature. 350µl buffer RW1 was added and the preparation centrifuged at 13000xg for 20 seconds, flow through discarded and 500µl buffer RPE was added two times more with centrifugation as before for 20 seconds for first and 2 minutes for the second time. After removing flow through and replacing the collection tube the sample was centrifuged at 20000xg for 1 minute. The collection tube was replaced with a 1.5

ml capped tube then 50µl water was added and incubated for 10 minutes at room temperature. The sample was then centrifuged at 13000xg for 1 minute to elute RNA.

4.2.2.5. RNA quantification

A NanoDrop 1000 Spectrophotometer was used to quantify the extracted RNA. The program was set at λ= 260nm to read the samples. 1µl water was placed on lens as blank. After cleaning the lens 1µl of RNA suspension was dropped on it. The concentration (ng/µl) and ratio (260/280) of sample were recorded from the output.

4.2.2.6. Reverse Transcription

Complementary DNA (cDNA) and no-reverse transcription (NRT) control reactions were made for each RNA sample using Superscript III kit protocol. In a PCR tube 0.5µg RNA was incubated with 1µl of 10µM dNTP and 1µl of genomic directed primers (alternatively 172.5 µM random primers were used) at 65˚C for 5 minutes. 1µl RNase inhibitor reagent, 1µl of 0.1mM DTT and 4µl of 5X first-strand buffer were added to the reaction. 1µl Superscript III (M-MLV reverse transcriptase) was added only to the cDNA reaction. 1µl water was added to the NRT reaction. A no template control (NTC) reaction was also set up using all reagents in the Superscript III kit without the RNA sample to check for DNA contamination. All reactions were set to final volume of 20µl by H_2O addition then incubated as:

- 50°C for 45 minutes
- 70°C for 15 minutes
- 4°C storage

4.2.2.7. Preparation of cDNA samples for qPCR

The cDNA samples were diluted 1:4 with RNase-DNase free H_2O to minimise the potential inhibitory effect of upstream reagents on amplification in qPCR. The test and controls reactions were set as:

- Negative control (H_2O)
- Seven standards (series of 10-fold serial dilution of genomic DNA in order of 10^7 - 10^2 and a $5x10^6$ copies / µl)
- No template control (NTC)
- One non reverse transcribed (NRT) per sample cDNA
- Three test reactions per cDNA

The standards were made from mycobacterial genomic DNA and serially diluted with H_2O as follows:

1. Standard 1 = 10^7
2. Standard 2 = $5x10^6$
3. Standard 3 = 10^6
4. Standard 4 = 10^5
5. Standard 5 = 10^4
6. Standard 6 = 10^3
7. Standard 7 = 10^2

Of each sample 1µl was mixed with 24µl of Master Mix (4.2.1.3).

Rotor Gene 6000 software on Corbett PCR machine was used for quantitative real-time PCR with setting as:

- 95°C for 15min Hold (for activation of Taq polymerase)
- 95°C for 10 sec Denaturing
- 60°C for 15 sec Annealing
- 72°C for 20 sec Extension with acquisition at cycling A
- 82°C for 20 sec Extension with acquisition at cycling B

The green channel was set up to gain at 8 with 470nm and 510nm as source and detector respectively. The acquiring fluorescence intensity was set at 72°C and 82°C to differentiate nonspecific products. Amplification was repeated for 40 cycles and melt curve was fixed as rising 1°C degree from 60°C to 99°C. When needed high resolution melt (HRM) was applied at the end of run.

4.2.2.8. Normalisation of gene expression data

The qPCR concentration values for target genes were normalised against one of mycobacterial housekeeping gene namely *sigA*.

The relative gene expression (as fold change) was measured using the following equation:

Equation 4.1 *Fold Change of gene* = $$\frac{Test\ target\ gene\ /\ Test\ reference\ gene}{Control\ target\ gene\ /\ Control\ reference\ gene}$$

4.2.3. Statistical analysis and visualisation of data

The statistical analysis was done using Graph Pad Prism software (Prism V6.0 for Windows). Rotor-Gene 6000 (Corbett Life Science – Qiagen) machine and software were used to run the qPCR and analyse the data respectively. The Rotor-Gene 6000 and Graph Pad prism were used for visualisation of data. Relevant analysis tests were performed to compute the statistical significance. A $p < 0.05$ was considered statistically significant. Coefficient of Variation was calculated using the equation:

Equation 4.2 $$\boldsymbol{Cv = \frac{\sigma}{\mu}}$$

Where ***Cv*** is defined as the ratio of the standard deviation $\boldsymbol{\sigma}$ to the mean $\boldsymbol{\mu}$.

4.2.4. RNA isolation from mycobacteria in infected neutrophils

All neutrophil infections were performed in 25cm^2 or 75cm^2 vent cell culture flasks in 5ml or 10ml respectively. An aliquot of frozen H37Rv was taken from -80˚C thawed at room temperature for 20-30 minutes then washed twice with pre-warmed RPMI prior to infection. GTC treatment was done immediately after thawing to get RNA representative of the inoculum. When ct value in qPCR results of a given sample showed a difference of less than 10 cycles between NRT and cDNA then the RNA sample was treated again with on column DNase I.

4.3. RESULTS

4.3.1. Gene expression of H37Rv in infected HL60 cells using inocula from recently thawed bacterial preparations shows unacceptable levels of variation.

Frozen aliquots of H37Rv were used to achieve uniformity in inocula for biological

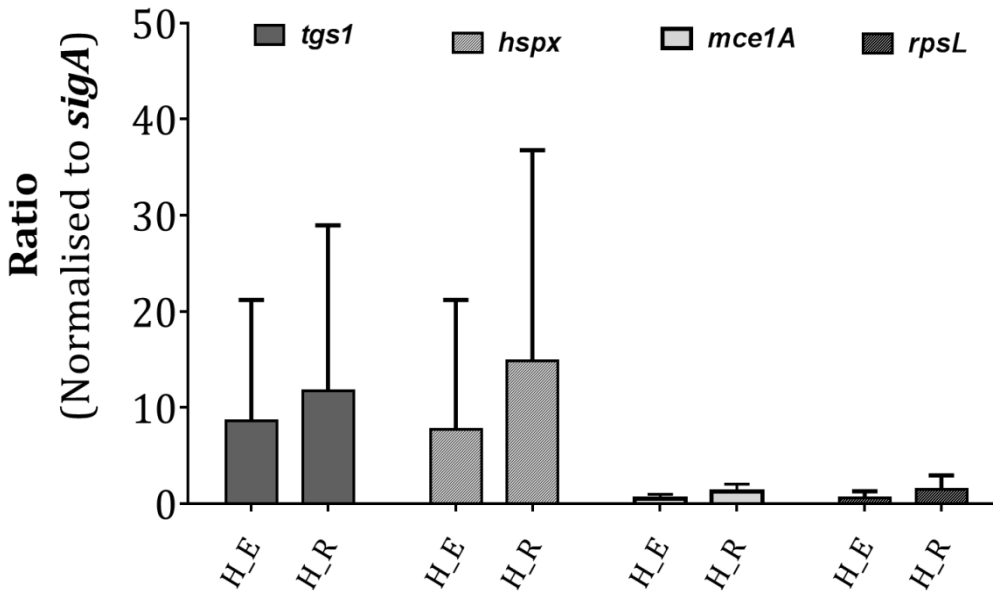


Figure 4.2 Variability was substantial in relative gene expression of frozen H37Rv. In 5ml RPMI 9.6 x 10^6 IFN-γ activated HL60 cells were incubated for 4h at MOI = 5.4 with H37Rv taken from frozen stock then washed twice with RPMI to remove extracellular bacilli and incubated in 5ml RPMI for 2h. H_E (HL60 cells relative to Exponential) and H_R (HL60 cells relative to control RPMI)
(Error Bars: SD n=3)

replicates. Based on the primary hypothesis that the results of this in vitro study

should be compared with sputum transcriptome, four genes were selected from sputum profile including two up regulated (*tgs1* and *hspX*) and two down regulated (*mce1A* and *rpsL*). The genes were investigated for their expression on encounter with HL60 cells relative to exponentially growing H37Rv and RPMI control (Figure 4.2). The mean of fold changes for *tgs1* on HL60 encounter relative to exponential phase in 7H9

broth and RPMI were 8.5 and 12.5 respectively. Fold induction of *hspX* in HL60 cells was 7.5 relative to exponential phase and 15 relative to RPMI. The gene expression of *mce1A* and *rpsL* in HL60 cells showed no significant changes

Table 4.1 Calculated Coefficients of Variation for relative expression of the test genes replicates.

	Coefficient of Variation %			
	tgs1	*hspX*	*mce1A*	*rpsL*
HL60 / Expo.	145.5	175.6	94.5	111.9
HL60 / RPMI	146.5	147.5	61.1	99.3

relative to exponential phase and RPMI. The values of Coefficient of Variation (CV) for the transcript ratios obtained across 3 biological replicates are shown in (table 4.1). CVs of less than 20% are desirable for this type of experiment; therefore experimental conditions were explored in order to reduce inter-experimental variation

4.3.2. Variability of results with exponential phase inocula

Mid exponential H37Rv was used for infection of HL60 cells following centrifugation and re-suspension in RPMI and gene expression determined using qPCR (Figure 4.3). As before *tgs1*, *hspX*, *mce1A* and *rpsL* were investigated relative to exponential phase-7H9 broth and RPMI. There was no significant change in expression of *rpsL* in HL60 cells relative to exponential phase-7H9 broth and RPMI conditions. The overall variation between biological replicates showed

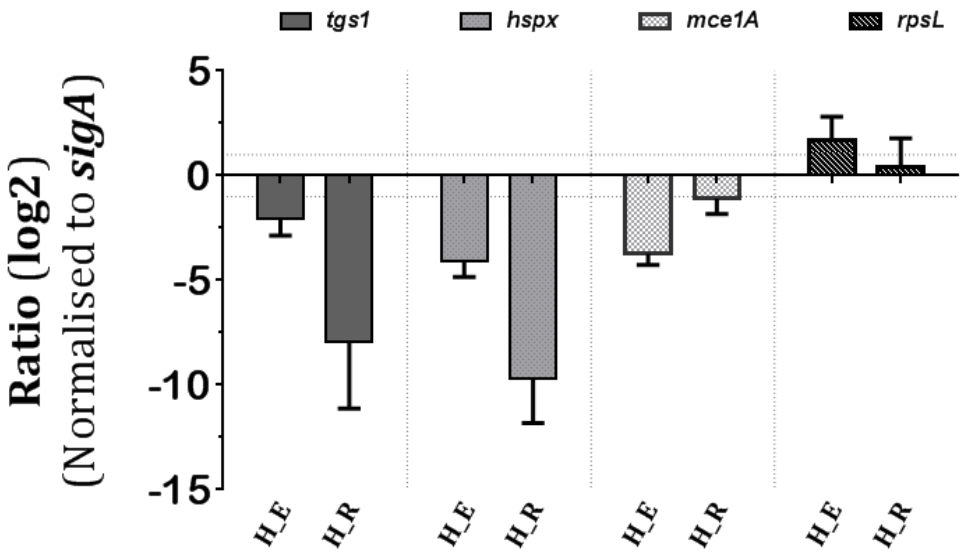


Figure 4.3 Gene expression of exponential *H37Rv* shows down-regulation in HL60 cells relative to 7H9 and RPMI. In 10ml RPMI 8.5 x 10^7 IFN-γ activated HL60 cells were incubated for 4h at MOI = 6.4±1.5 with H37Rv taken from OD= 0.7 then washed twice with RPMI to remove extracellular bacilli and incubated in 5ml RPMI for 2h. H_E (HL60 cells relative to Exponential) and H_R (HL60 cells relative to control RPMI)
(Error Bars: SD n=3)

coefficient of variation less than 25% except for the *tgs1* ratios (CV% = 40.6).

4.3.3. Effects of inoculum preparation on transcription

From the previous experiments it was suspected that centrifugation and re-suspension in RPMI had a significant influence on gene expression and that this effect would be maximal for exponentially growing Mtb. Given that the sputum transcriptome indicated slow or non-growth of the bacilli present it was decided to investigate stationary phase inocula and pre-conditioning in RPMI. Stationary phase H37Rv cells were incubated in RPMI for 120h and expression of *hspX*, *tgs1* and *icl1* monitored to reflect respectively two DosR and one non-DosR regulated genes. The results are shown in Figure 4.4.

More than 1000-fold induction of the DosR regulated genes occurred as a result of centrifugation, while *icl1* was relatively unaffected. When RPMI was added to the pellet (0h RPMI), the expression of *hspX* and *tgs1* fell by 10-fold while *icl1* was induced ~3.5-fold. After a further 2h, the expression of *hspX* and *tgs1* increased again by ~10-fold and *icl1* increased by ~5-fold. Expression appeared relatively stable at 24h then declined out to 120h for *tgs1* and *icl1* but relatively less for *hspX.* The CVs between biological replicates was less than 20% for all time points.

These results were considered to support pre-incubation of inocula in RPMI for 24h prior to neutrophil infection.

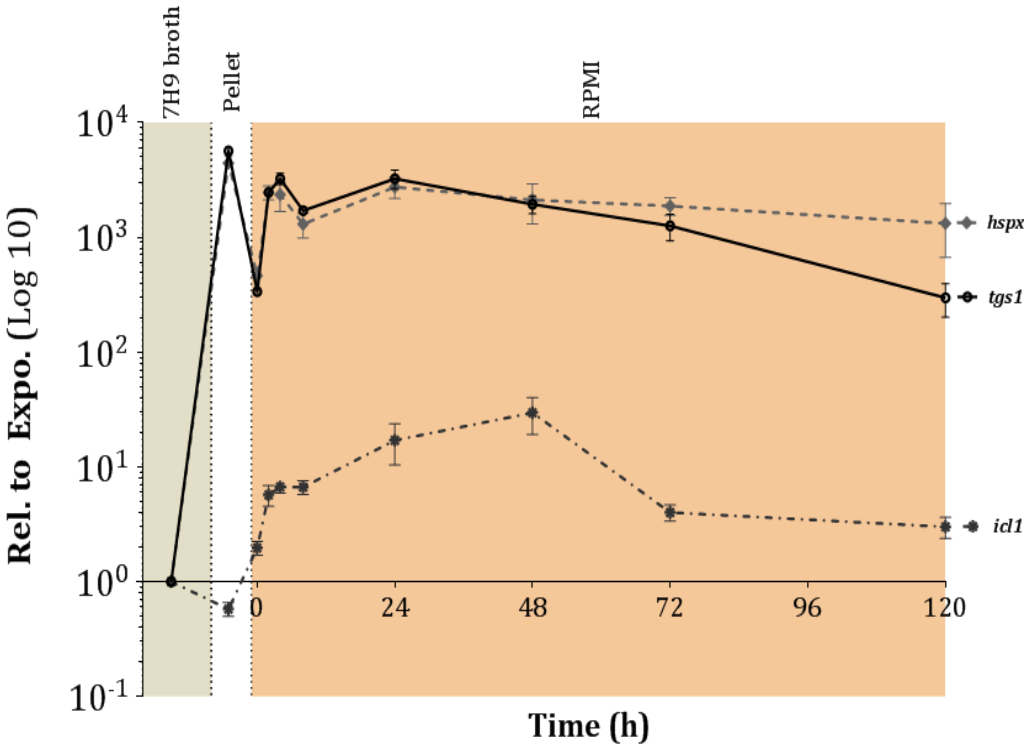


Figure 4.4 Relative expression of target genes stabilises at around 24h in RPMI. In 5ml RPMI 2.2 x10^7 H37Rv taken from OD= 1.8±0.8 were incubated for 120h. Relative gene expression was done to H37Rv in 7H9 at OD=0.7. The samples were taken from pellet (i.e. bacteria in 7H9 broth was centrifuged and supernatant removed), 0h RPMI (i.e. pellet re-suspended in RPMI), 2h, 4h, 8h, 24h, 48h, 72h, and 120h. (Error Bars: SD n=3)

4.3.4. Visualising NETs formation of primary human neutrophils infected with H37Rv

The formation of NETs was investigated via microscopic visualization of Giemsa stained slides. The images suggested an increase in aggregation and formation of cytoplasmic protrusions in samples of Mtb-neutrophil after 24h, while no similar morphological changes were seen in corresponding controls.

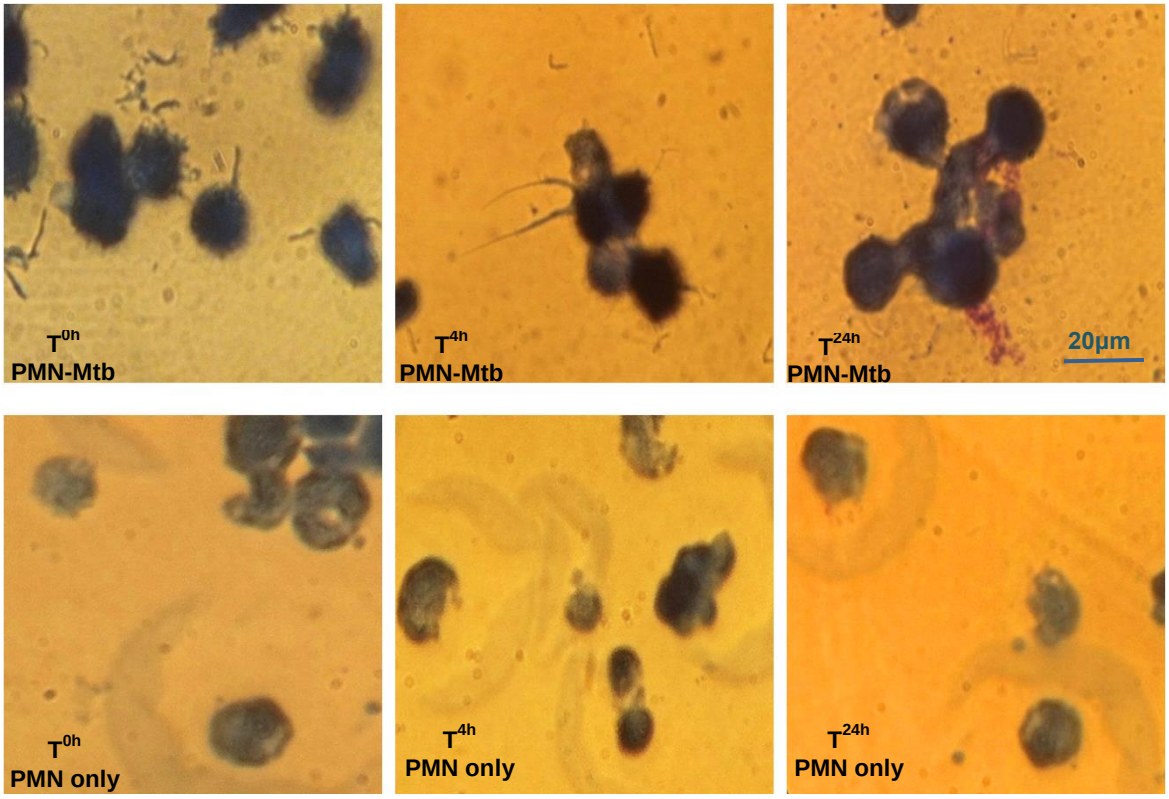


Figure 4.5 Increased aggregation and bridge formation of neutrophils incubated with H37Rv. Slides of primary human neutrophils infected with stationary H37Rv in RPMI at time points of 0h, 4h and 24h were prepared and fixed with 23% formaldehyde overnight. Giemsa stain was applied to the slide and visualised under light microscope at 1000X magnification. Corresponding time points of controls were made from neutrophils without bacilli.
The scale bar applies to all slides

4.3.5. Gene expression in pre-conditioned H37Rv bacilli exposed to HL60 cells

Successive attempts to establish a model of neutrophil infection returned a gradual reduction in inter experimental variability. In addition to achieving stable expression in the inoculum it was realised that processing the preparations to exclude extracellular bacteria from the analysis could further Increase experimental variation. Following recognition of the NET phenomenon it was also realised that this approach may better reflect the interactions taking place in vivo. Thus bacteria suspended in RPMI for 24h were directly mixed with HL60 cells for 4h and RNA extraction was done immediately after incubation (Figure 4.6). The results showed that *tgs1* and *hspX* were up regulated in HL60 cells relative to exponential. While *tgs1* was not significantly changed, *hspX* was statistically different relative to

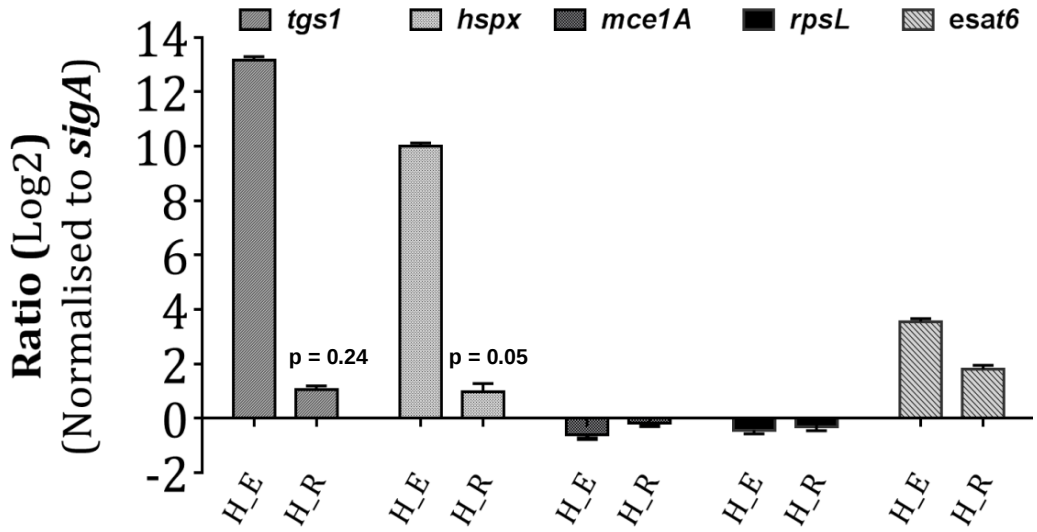


Figure 4.6 Differential expressions of target genes in H37Rv encountering HL60 cells relative to exponential growth and pre-conditioned bacilli in RPMI. In 5ml RPMI 6.5 x 10^6 IFN-γ activated HL60 cells were incubated for 4h at MOI = 11.7±1.2 with stationary H37Rv taken from OD= 1.9 and pre-incubated 24hr in RPMI.

H_E: HL60 cells relative to Exponential; H_R: HL60 cells relative to control RPMI; Paired t-test was done to compare H and R (Error Bars: SEM n=5)

stationary in RPMI. There was no significant change in expression of *mce1A* and *rpsL* in HL60 cells relative to exponential and RPMI. The expressions of *esat6* relative to exponential and RPMI were up regulated. All biological replicates showed a CV of less than 20%.

4.3.6. Gene expression of H37Rv in infected primary human neutrophils

Pre-conditioned H37Rv cells were incubated with IFN-γ treated neutrophils for 4h and RNA extracted. Figure 4.8 demonstrates significant increases in expression of *tgs1* and *hspX* compared to exponential phase growth. While *tgs1* did increase relative to RPMI, *hspX* did not. *RpsL, mce1A* and *esat6* showed no significant change.

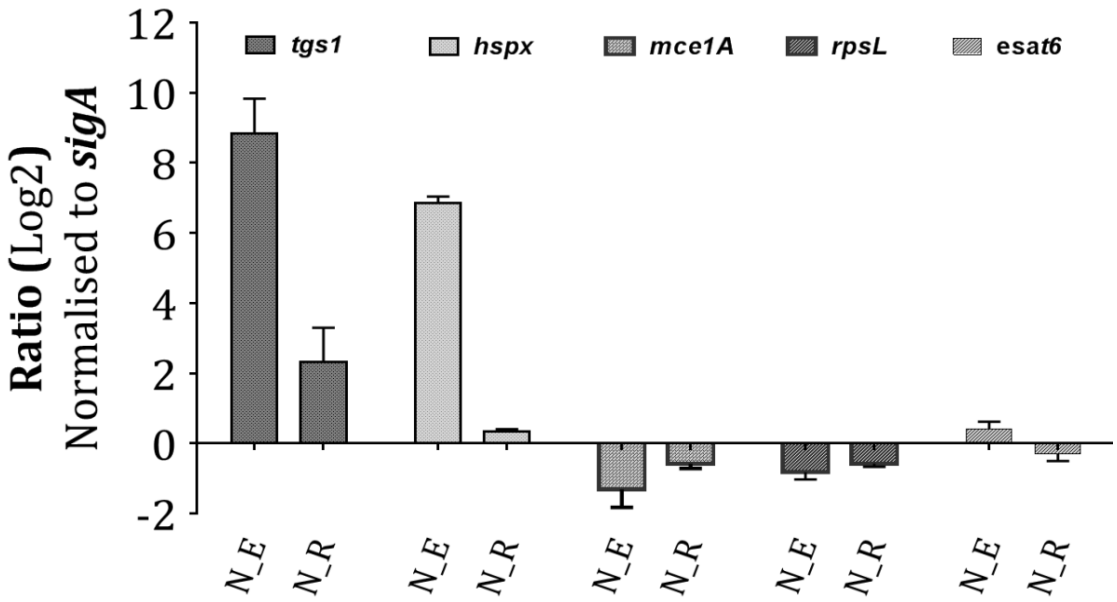


Figure 4.7 The expression of stress genes in stationary H37Rv up-regulated in infected primary human neutrophils relative to 7H9 and RPMI. In 5ml RPMI 5.2 x 10^6 IFN-γ treated neutrophils were incubated for 4h at MOI = 9.2±0.7 with H37Rv taken from OD= 1.8 and pre-incubated 24hr in RPMI. N_E: neutrophils relative to Exponential; and N_R: neutrophils relative to control RPMI (Error Bars: SEM n=5)

4.4. Discussion

Several phases of experimental design were used to establish a model of in vitro neutrophil infection from which a representative RNA sample could be extracted. It was decided to analyze combined intracellular and extracellular responses for the following reasons.

Firstly, because the initial step in RNA extraction experiments was focused on getting intracellular signals from neutrophil infection for which extracellular bacilli were removed via centrifugation and washing (Figure 3.3). The main problem with this approach was difficulties in getting reproducible results due to the influence of a number of stimuli inherent to the washing process. To address this problem, the washing step was omitted in the subsequent series of experiments thus extracellular bacilli were also included in RNA extraction.

Secondly, it was realised that internalised Mtb will eventually be exposed to external environment following NETs formation. In view of that, it seems impossible to differentiate between intracellular and extracellular stimuli that induce transcriptional response in Mtb.

Finally, taking a combined sample of intracellular and extracellular Mtb in vitro would be more representative of in vivo condition where neutrophils and associated Mtb bacilli can be randomly sat at various positions to each other. Altogether, it was recognised that both intracellular and extracellular signals are equally important.

The rationale behind selecting specific target genes for PCR was the fact that the results from this study should be compared with TB sputum transcriptome. Thus, among the up regulated genes in sputum study, *tgs1*, *hspX* and *icl1* were selected in addition to *mce1A* and *rpsL* which were down regulated. There was also qPCR done on *narK2* that was up regulated in sputum but due to the low reproducibility in results it was let off from further investigations (data not shown).

ESAT6 was also included in this study because it is one of the key virulence factors of Mtb causing neutrophil necrosis. This gene has been shown to be involved in NETs formation following necrosis (Francis *et al.*, 2014). Since NETs formation was considered in this study, this gene was included in order to investigate the link between its expression level and NETs.

There were also a number of genes exclusive to intracellular environment namely *echA19*, *mbt1* that were investigated at beginning of this study. Because the direction of study changed towards addressing both intracellular and extracellular, they experiments were stopped later on (data not shown).

4.4.1. HL60 cells infected with recently thawed H37Rv inocula

Although the ultimate target was to work with primary human neutrophils, sourcing from different donations is an obvious cause of variation. Thus the HL60 cell line was used to establish experimental protocols.

HL60 cells were used with less than 3 passages difference between two consecutive samplings. Initially frozen Mtb stocks were used as the inoculum. However, it was rapidly appreciated that the thawing process would lead to changes in gene expression that would be at maximum at the time of inoculation and this probably contributed to the extreme variation observed in this experiment.

Although mycobacteria in different biological replicates came from the same source (frozen stock prepared from exponential H37Rv), the variability in results pointed towards different environmental experiences of bacilli during cryopreservation, thawing and incubation with phagocytes. Other studies have shown that the outcome of experiments can be influenced by the means of bacterial storage (Nascimento *et al.*, 2005). Besides, it seems that other factors such as time to defrost sample at the time of performing experiments were also involved. Therefore, a reasonable suggestion to reduce the variability in gene expression results was to use bacteria at growing stage for infection.

4.4.2. Variability of results with exponential phase inocula

The use of actively growing Mtb was commended from the inconsistencies observed in the results of previous experiments. Thus, the variability was significantly reduced. The majority of samples demonstrated less than 25% coefficient of variation between three biological replicates. One exception was the presence of a relatively large coefficient of variation (CV% = 40.6) in quantification of *hspX* for RPMI samples. This might be due to the fact that expression of *hspX* is essentially induced by abundant environmental stimuli; one of which is the transferring of bacteria from 7H9 broth as bacterial culture into RPMI as eukaryotic medium. Such environmental changes and their impact on transcriptional responses in mycobacteria have been addressed explicitly in the subsequent experiments.

Regardless of slightly improved reproducibility in experiments, the results revealed a new set of problems that were attributed to the use of exponential phase mycobacteria. Thus, the results presented a dissimilarity of results from using exponential phase H37Rv with sputum transcriptome profile. Such differences emphasis, to a great extent, the fact that the majority of transcriptional signals in TB sputum are not derived from exponential phase Mtb populations. As discussed in the Garton study, the transcriptome signature of TB sputum is close to non-replicating persistent bacilli (Garton *et al.*, 2008).

In a population of persisting bacteria, the growth can be slowed as a consequence of low oxygen tension, old age or quorum sensing (Mitchison, 2004). The notion of persister bacteria in the sputum gave rise to the idea that stationary phase Mtb should be used in the subsequent experiments. Essentially mycobacteria at stationary phase seem to bear more resemblance to non-replicating persistent bacilli in sputum.

4.4.3. Transcriptional changes of H37Rv in RPMI

From observation in the previous experiments we recognised a significant influence of environmental stresses including RPMI on gene expression of Mtb. Therefore, H37Rv bacilli were incubated in RPMI for 120h in order to determine the transcriptional responses to the wide range of stimuli in a transition from bacterial culture to centrifugation in test tube then removing the culture and eventually introducing a none bacterial medium such as RPMI (Figure 4.3).

The starting point of the journey was pelleting via centrifugation in order to suspend the bacteria in RPMI. This is the commonly used medium for neutrophil culture in vitro particularly for the purpose of infection studies with HL60 cells (Rook *et al.*, 1987) and primary human neutrophils (Perskvist *et al.*, 2002) with Mtb. It was therefore necessary to consider the changes in gene expression during acclimatisation of bacteria to the new surroundings.

The impact of RPMI on transcriptional changes was found to be prominent during the first 24h of exposure. Such RPMI impact has clear potential to explain the variation in the previous experiment where the encounter with neutrophils took place during the period of maximal response to RPMI. The changes in gene expression in response to RPMI become stabilised and bacteria appear to be acclimatised after 24h of incubation. Therefore the bacteria were incubated in RPMI for 24h prior to infection in subsequent experiments.

4.4.4. Visualising NETs formation of primary human neutrophils infected with H37Rv

Microscopic view of primary human neutrophils was indicative for increased aggregation and cellular bridge formation with increasing time of exposure to H37Rv *in vitro*. These features were not observed in neutrophils without Mtb suggesting that polymorphs undergo different morphological changes depending on presence or absence of H37Rv. In view of that, the results support the NETs

formation. However, confirmation of NETs need special staining techniques such as Sytox green which stains the chromatin so that the protruded strand can be seen as green projection out of neutrophil (Kawakami *et al.*, 2014). According to the standard operating procedure in the University of Leicester a Category III microorganism like Mtb must be killed with 23% formaldehyde overnight before taking the sample out of containment lab. As DNA staining could not be appropriately applied to such fixations, it was decided to use Giemsa stain though it is not as specific as DNA staining in detection of nuclear strands. Nevertheless, the morphological changes over extended time points of incubation imply on a meshwork of neutrophils in a close contact with H37Rv bacilli.

4.4.5. RNA extraction from stationary H37Rv in infected HL60 cells

The final experimental set up was accomplished by infecting HL60 cells with stationary phase H37Rv pre-incubated overnight in RPMI then isolation of RNA from both intracellular and extracellular bacilli. The HL60 cells were also primed with IFN-γ for 1h prior to infection to be more representative of in vivo situation of infection. To this end the results seemed promising considering all modifications applied to optimize the infection model. Thus, relative gene expression of Mtb in infected HL60 cells demonstrated a reasonable level of reproducibility in this modified infection system.

As discussed previously, due to NETosis, non-internalised bacteria will be exposed to similar stimuli to the internalised bacteria. This also provided the rationale for inclusion of *esat6* amongst the transcripts to be studied. A significant induction of *esat6* was observed relative to the exponential phase and RPMI control conditions.

One concern with the stationary phase inoculum was the use of *sigA* as a reference gene for normalisation in qPCR as the expression of *sigA* is found to be slightly reduced in the stationary phase H37Rv (Manganelli et al., 1999). Nonetheless, *sigA* has been used extensively in quantification of mycobacterial

gene expression (Singh *et al.*, 2006) and some reassurance can be taken from the observation that its levels were found proportional to Mtb replication inside THP-1 cells (Dubnau *et al.*, 2002).

Overall, a marked improvement in reproducibility of experiments was achieved with the protocol adopted in this section.

4.4.6. Gene expression of H37Rv in infected primary human neutrophils

The infection system established with HL60 cells was next applied to primary human neutrophils. The data from relative quantification of gene expression for *tgs1*, *hspX*, *mce1A*, *rpsL* and *esat6* on encounter with primary human neutrophils showed a satisfactory level of reproducibility similar to the use of HL60 cells.

The pattern of gene expression of H37Rv in primary human neutrophil was different than HL60 cells. In partiicular, *esat6* was not differentially expressed in this system. This might reflect the greater potency of primary human neutrophils noted previously. NET formation recognised by the aggregation of polymorphs was extremely rapid and prominent in primary cells compared to HL60 cells (data not shown). It seems possible therefore that *esat6* expression might have been stimulated earlier with primary cells and declined to baseline by 4 hours.

Quantifying the expression of the other four selected genes demonstrated that encounter with both types of neutrophils induced similar transcriptional resposnes in H37Rv. With regard to the comparison between sputum transcriptome and neutrophil encounter as the initial hypothesis of the study, the results from four target genes can be interpreted in two ways. on one hand, the results of in vitro neutrophil encounter can be comparable to TB sputum when the relative expression was done against exponentially growing H37Rv. As in the majority of previous works, the differential expression of genes in different test conditions including phagocytosis was measured relative to expression of exponential phase bacteria (Schnappinger *et al.*, 2003). On the other hand, relative expression of four

target genes against control RPMI indicates a different pattern of expression between in vitro neutrophil encounter and sputum transcriptome.

The rationale for comparing to control RPMI is a view that only trasncriptional signals exclusive to neutrophil encounter are appropriate for comparision. This approach was also undertaken in a number of macrophage infection studies (Rohde *et al.*, 2007). Overall, the results were reproducible and further validated by RNA-seq data in chapter 5.

4.5. CONCLUDING REMARKS

- A high degree of variability between biological replicates was found when frozen H37Rv was used for infection of HL60 cells.
- Relative gene expression of H37Rv in HL60 cells was substantially different from sputum transcriptome when log-phase bacteria was used for infection
- RPMI was found to have a great impact on gene expression of H37Rv particularly during first 24h of infection
- The reproducibility of experiments was improved with the use of stationary phase H37Rv pre-incubated in RPMI prior to infection.
- Using stationary H37Rv and extracting RNA from intracellular and extracellular bacilli found to be more representative of in vivo condition and yielded a further reproducibility
- The initial hypothesis that neutrophils can be responsible for transcriptome of TB sputum can still be valid when the gene expressions were measured relative to exponential phase H37Rv while the hypothesis was not supported when measured against control RPMI

CHAPTER FIVE

The Transcriptional Changes in *M. tuberculosis* H37Rv on Neutrophil Encounter *In Vitro* studied by RNA-Seq

5.1. Introduction

5.1.1. Studying the bacterial transcriptome

The transcriptome is defined as the totality of the transcripts that are present in an organism at a specific time. In a complete transcriptome analysis the main goal is to characterise and record all expressed transcripts at a particular time and to quantify the differential expression of transcripts in different conditions (Costa *et al.*, 2006)

To understand the functional output of a genome, a careful analysis of transcriptome is essential. In this regard, all aspects of the transcriptome are important including transcript content, determination of transcriptional start sites, mRNA abundance and antisense RNAs (Filiatrault, 2011). A comprehensive examination of global gene expression is hard to achieve in prokaryotes due to high rRNA and tRNA abundance in addition to the instability of transcripts.

Recently developed technologies such as microarrays and next generation sequencing enable a comprehensive assessment of the transcriptome. However, the bacterial whole transcriptome analysis is rather different from eukaryotes to the fact that rRNA constitutes about 90% of total RNA in bacteria. A cDNA library for eukaryotic mRNA can be selectively produced in the presence of poly (A) tail which is very infrequent in bacterial mRNA. Therefore, removing rRNA from isolated RNA samples to enrich mRNA can be a way to tackle this problem. Nevertheless the rapid advance in high-throughput sequencing technology has provided viable tools for the study of microbial transcriptome even in the presence of such difficulties (He *et al.*, 2010).

5.1.2. RNA-Seq

The capability to directly and quantitatively identify all RNA species independent of their being abundant or rare, large or small is essential for an ideal method of transcriptome analysis (Pinto *et al.*, 2011). RNA-seq, a high throughput sequencing technology is a powerful approach for gene expression profiling, discovery of previously non-annotated genes and mapping the architecture of transcriptome in different species of bacteria (Wagner *et al.*, 2013).

A number of advantages can be offered by RNA-seq over microarray such as marked capacity in detection of low abundance transcripts and means for profiling transcripts in a strain which has no available genome sequence or annotated gene. The continuous reduction in sequencing cost, the mounting number of facilities and accessibility of high-throughput sequencing, in addition to publicly available tools of bioinformatics have made RNA-seq a popular and attractive method for bacterial transcriptome study (Haas *et al.*, 2012).

Nevertheless, there are biases inherent to RNA-seq that if not noticed can create problems in the analysis. Such biases are mainly attributable to the transcript size, fragmentation steps, cDNA synthesis and enrichment of mRNA prior to sequencing. Moreover, RNA-seq is complex involving a massive data processing for which a robust analysis is required. Thus, profound knowledge of bioinformatics and high performance servers for computational processing are necessary.

Two major processes take place in an ordinary RNA-seq namely, data generation (Figure 5.1a) and data analysis (Figure 5.1b).

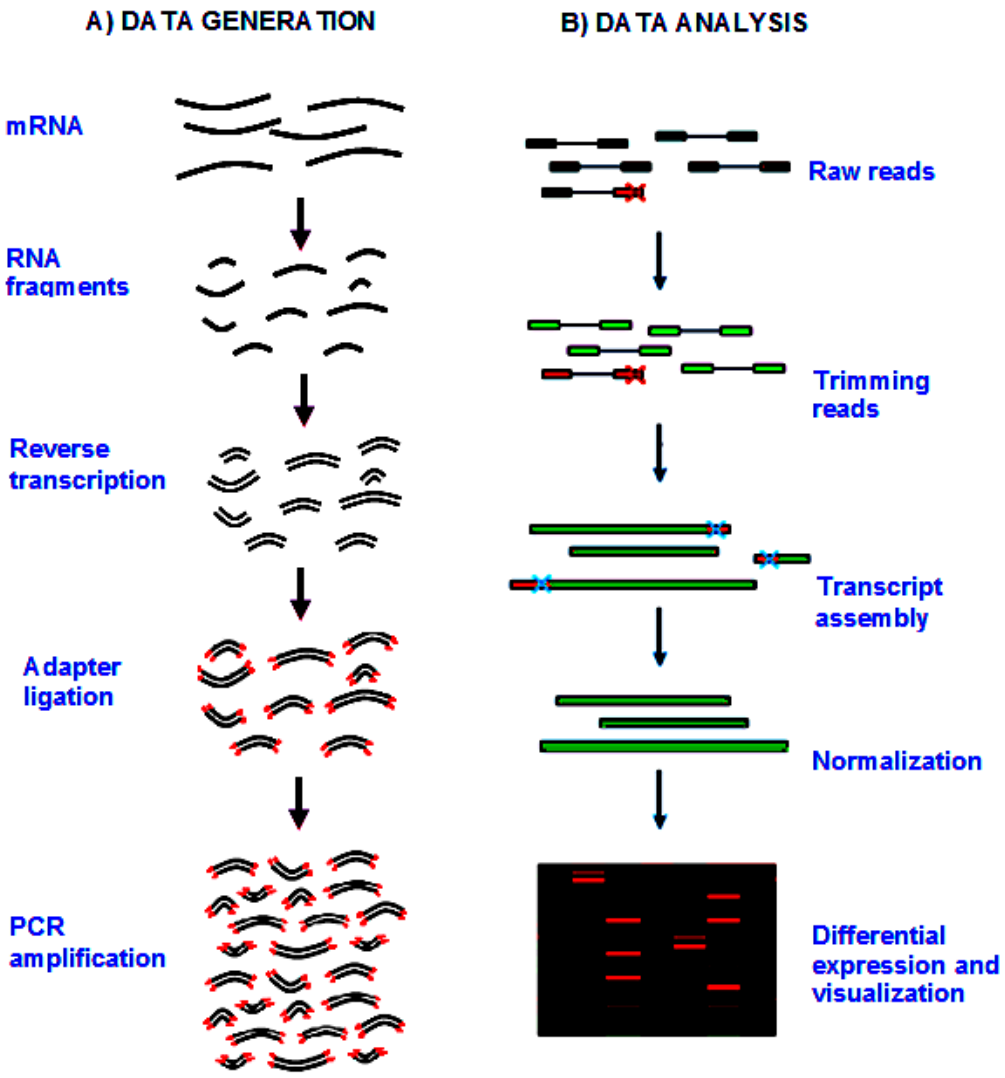


Figure 5.1 A typical RNA-seq experiment. a | Data generation. To generate an RNA-seq data set the RNA is extracted, DNA contamination is removed and short fragments of RNA are made then reverse transcribed into cDNA. The adaptors (red) are ligated and fragment size selection is performed. Finally, the cDNAs are sequenced using NGS technologies to generate short reads (black). **b | Data analysis**. Low quality reads and artefacts such as contaminant DNA (Ewing *et al.*), adaptor sequences and PCR duplicates are removed after sequencing. To improve the quality the sequence errors are optionally removed. The reads are then assembled into transcripts and the errors are removed via post-assembly processes. The transcripts are then post-processed and the expression of each transcript is measured by counting the number of reads that align to each transcript. Figure reproduced after (Martin *et al.*, 2011).

The cDNA library is then sequenced by next generation sequencers to produce millions of short reads using next generation sequencers. In data analysis phase, the pre-processed reads are then assembled into transcripts in order to determine their abundance followed by measuring the differential expression (Martin *et al.*, 2011).

5.1.2.1. RNA preparation

One of the main concerns in RNA library preparation is the RNA integrity. The quality of isolated RNA may vary from one experiment to another and this can be addressed by careful processing and proper handling of RNA after extraction. Often shorter fragments of RNA can occur in the sample due to a rapid digestion of RNA in the presence of ubiquitous RNase enzymes. The conventional approaches to determine RNA integrity are gel OD measurement, Nano-Drop OD measurement and denaturing agarose gel electrophoresis all of which, however, fail to provide a strong correlation with RNA integrity (Schroeder *et al.*, 2006).

Advanced technologies are now available as chip systems that provide a precise measure of RNA integrity. In the Bioanalyser-2100 system for example, minute amounts of sample RNA are separated on channels of a microchip based on molecular weight then detected by laser induced fluorescence. The intensity of fluorescence correlates with the quantity of RNA of a given size. The visualisation of results in an electropherogram allows determining the RNA Integrity Number (RIN) through a neural network-based analysis. This provides a numerical system to classify total RNA in which 1 represents for the most degraded profile and 10 the most intact. The value of RIN is correlated with the validity of downstream gene expression quantification experiments (Fleige *et al.*, 2006).

5.1.2.2. Preparation of cDNA library

A successful RNA-seq analysis demands careful high yield and high quality RNA isolation. Commercially available kits or organic solvents are typically used in most procedures. The method should not bias the sampling across the transcriptome (Croucher et al., 2010) and removal of rRNA prior to generating the cDNA library is a commonly applied approach in bacteriology (Chen *et al.*, 2014).

When RNA-seq was first established, random hexamers were used to make cDNA followed by second strand DNA synthesis (Perkins *et al.*, 2009). Such double stranded DNA loses the strand specific signal in the RNA sample because both sense and anti-sense strands are equal. In the Ilumina RNA-seq protocol, however, the libraries are constructed only from the first strand cDNA. Thus, the DNA strand specificity is maintained in RNA-seq data.

In general, among currently present RNA-seq technologies, Illumina, SOLiD and 454 platforms have been commonly used to study the bacterial transcriptome (Croucher et al., 2010).

5.1.2.3. Sequence library preparation

Prior to sequencing in all current RNA-seq platforms the DNA library needs to be pre-processed by shearing into a suitable platform-specific size range. The next step involves producing a blunt ended DNA fragment after an end polishing process followed by ligation of specific adapters to 3´ and 5´ ends of the fragment. Depending on the sequencing platform an extra set of unique adapters is used which should be compatible with the downstream processing steps. Figure 5.2 demonstrates example of adapters used in different platforms (Buermans et al., 2014).

In the Illumina technology, all sequencing processes take place in a flow cell. The cDNA molecules are attached to the surfaces in a flow cell and are amplified by PCR (Dabney et al., 2012). Each molecule then is sequenced either from one end (i.e. single-end reads) or from both ends (i.e. Pair-end reads).

A typical read, depending on the sequencing technology and design of experiment, is about 30 to 400 bp (Wang et al., 2009c). Longer reads are preferable as they

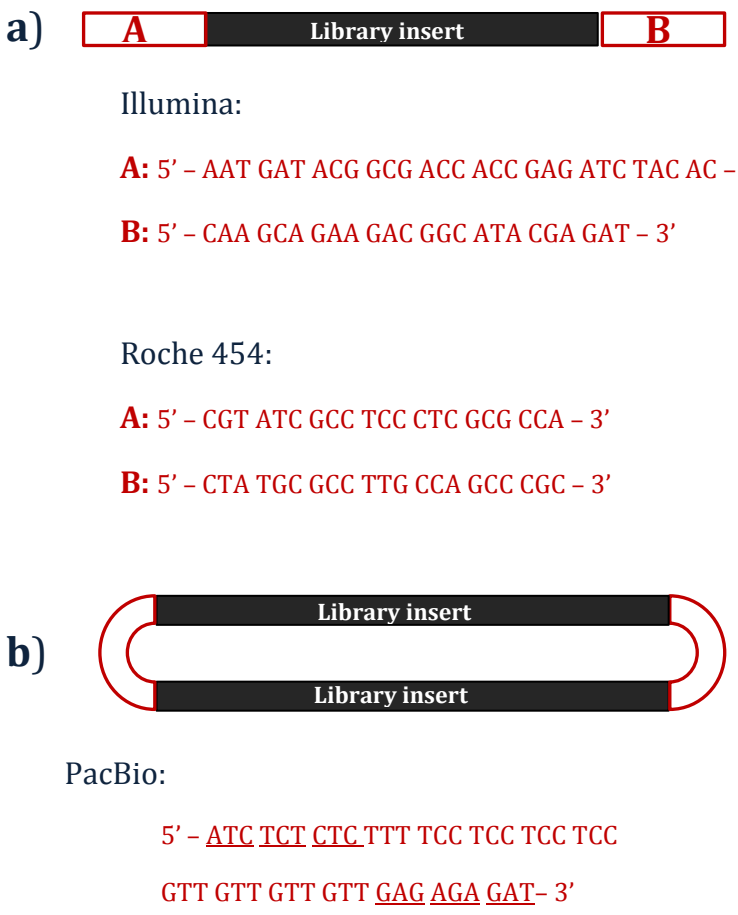


Figure 5.2 Structure of sequence library molecules for the different technologies. Linear library molecules (Panel A) contain different adapter sequences at the 5′ [A] and 3′ [B] ends of the library inserts. Circular library molecules (Panel B) contain identical adapter molecules at both ends of the insert (Buermans *et al.*, 2014).

significantly minimise the complexity of the transcriptome assembly *(Martin et al., 2011).*

The depth of sequencing determines the depth of coverage which is important in detecting and quantifying transcripts with low abundance (Croucher et al., 2010). Thus, in order to enhance the coverage in RNA-seq the total number of reads should be increased *(Chen et al., 2014)*. However, considering the costs of sequencing, the experimental design for an RNA-seq experiment should determine the right balance between the number of reads per sample (i.e. depth of sequencing) and the number of samples needed to be sequenced (i.e. breadth of sequencing). In this regard, for detection of rare transcripts the depth is critical whereas the breadth seems to be important when statistical confidence is required. For example, in order to statistically ensure that a gene is differentially expressed in different conditions, several time-points and biological replicates are needed (Haas et al., 2012). Moreover, there is a limit for increasing the depth of sequencing as genes with no transcript are most probably not expressed. Likewise, for a gene with one or few detected reads among millions of reads it is biologically not meaningful to identify it as expressed gene (Wagner et al., 2013).

5.1.2.4. Base calling and quality score

Individual bases are predicted during sequencing through a base calling algorithm. The mechanistic details are different across sequencing technologies but the overall pipeline is similar (Ledergerber *et al.*, 2011). In the Illumina platform for example, fluorescently labeled bases with an attached reversible terminator are incorporated in each cycle during sequencing. The terminators provide a sequential incorporation of fluorescently labeled bases one by one. The whole cluster is excited by a laser then the emission from individual incorporated bases is captured by an imaging device. After removing reversible terminator the incorporation of next base is allowed for the next sequencing cycle. Thus, four

images with optimal wavelengths are generated at the end of each sequencing cycle for each of four bases separately labelled with fluorescent dyes and nucleotide channels A, T, C and G are represented by a corresponding intensity value (Cacho *et al.*, 2015).

The images translated into base calls are represented in a FASTQ format. In the FASTQ file the nucleotides are designated with an ASCII encoded quality score named as Phred score. Phred score represents the probability that the corresponding base calling is incorrect by the equation:

$$P = 10^{-(Q/10)}$$

Were Q is the quality score ranging from 0 to 41 therefore the error rate ranges from 1 to 0. 00008 for each base call (Del Fabbro *et al.*, 2013).

The majority of low quality base calls occur at the end of reads particularly when Illumina technology is used (Ledergerber *et al.*, 2011). An essential step before downstream processing is to distinguish low quality and good quality bases. Thus, FASTQ files are taken through an extensive computerised process to detect nucleotides with poor phred scores and these are consequently deleted. This is routinely performed via a process called ´trimming´ by which specific and universal adapters, low quality bases in addition to repeat sequences are removed.

5.1.2.5. Alignment and annotation of transcripts

Downstream to quality control of reads is the detection and quantification of transcripts. The overall process falls into three phases starting with alignment of reads followed by assembly of transcripts and then quantification of transcripts. This is a demanding bioinformatics process and multiple programs and approaches are available. Even a small RNA-seq experiment can produce more

than a hundred gigabytes of raw sequencing reads in a single run (Trapnell *et al.*, 2012).

The reconstructing of full length transcripts from reads, except for small classes of RNA, is performed via assembly of reads which is done in two main approaches (Martin *et al.*, 2011). In the reference-based assembly the sequencing reads are aligned against a sequenced reference genome. In general, this approach is preferable because it is relatively precise and fast though it requires a high quality sequence of the reference genome; de novo assembly is required if this is not available. The latter involves assembly of transcripts by combining overlapping reads (Tjaden, 2015). Different mapping algorithms produce varying rates of false positives and negatives. These may be due to errors in sequencing, deleted bases in the referenced genome or structural rearrangements in the reads (Sims *et al.*, 2014).

5.1.2.6. Normalisation methods

One of challenges in performing RNA-seq is measuring the concentration of original RNAs. The question is how quantification of transcripts can represent the absolute RNA concentration (Mortazavi *et al.*, 2008). There are a variety of approaches proposed for normalisation of quantified reads:

Total count (TC): is the measure of all reads divided by the total mapped reads.

Upper quartile (UQ): instead of total count the upper quartile of counts different from 0 are used in the computation (Bullard *et al.*, 2010).

Median (Med): similar to UQ but the median is considered instead of upper quartile.

DESeq normalisation: with the assumption that majority of genes are not differentially expressed the median of ratio for a given gene is divided by the

geometric mean of total read counts. DESeq Bioconductor package is used for this type of normalisation (Anders *et al.*, 2010).

Trimmed Mean of M values (TMM): it is also based on hypothesis that most genes are not differentially expressed. TMM has been used with Illumina data. It is a factor used for computation of lanes results where one lane is considered as reference and others as test samples. TMM implements edgeR Bioconductor to calculate weighted mean of log ratio between reference lane and test lanes (Robinson *et al.*, 2010).

Quantile (Q): this was originally used for microarray data. Q is the distribution of gene counts across lanes for which Limma package in Bioconductor is implemented (Bolstad *et al.*, 2004).

Reads per Kilobase per Million mapped reads (RPKM): originally was used for comparison between genes within a sample. RPKM combines within and between sample normalisation to correct for differences in gene length and library sizes (Mortazavi *et al.*, 2008). RPKM for a gene is calculated using the following equation:

$$RPKMg = \frac{r_g * 10^9}{fl_g * R}$$

r_g : The number of reads mapped to a particular gene

fl_g (feature length): the number of nucleotides in a mapped region of gene

R: Total number of reads in the sequencing run of a sample (Wagner *et al.*, 2013).

However, there are different versions of RPKM including a normalisation approach by Rockhopper software, a program for prokaryotic RNA-seq analysis, which uses upper quartile of gene expression level after excluding the genes with 0 counts

(McClure *et al.*, 2013). Although, RPKM introduces biases due to gene length variances, it is a popular normalisation method in many RNA-seq applications (Dillies *et al.*, 2013).

Transcript per Million (TPM): is a modified normalisation method to measure the transcript abundance.

$$TPMg = \frac{10^6}{N}$$

Where N represents total number of transcripts

It is proportional to RPKM within a sample but slightly more accurate when applied between samples **(Wagner *et al.*, 2013).**

1.11.1.1. Detection of differentially expressed (DE) genes

One of necessities in high-throughput sequencing technology is the adoption of statistical methods for assessing the quantitative differences between experiments. In fact, detecting differentially expressed genes across a number of samples is a complex task.

In comparative transcriptomic data analyses the standard approach is to set a null hypothesis that the logarithmic fold change of expression for a gene is exactly zero between test and control (Love *et al.*, 2014). Based on this, several statistical models have been implemented for detection of DE genes. Well-recognised among these are: likelihood ratio statistics, t-statistics (both based on generalised linear method) and Fisher's exact test statistic (Bullard *et al.*, 2010).

There are parameters inherent to the RNA-seq technique that can introduce biases to the results including differences in flow cell, long transcripts and GC-content (Benjamini *et al.*, 2011).

5.1.3. sigK regulon

All members of Mtb complex possess Sigma factor K (σK). A reduced expression of *mpt70* and *mpt83*, two immunogenic proteins of *Mycobacterium bovis BCG* has been observed. This was recognised to be associated with sigK polymorphism leading to mutation in translation start codon (Manganelli, 2014).

Microarrays applied to investigate the global gene expression of σK impaired BCG strain complemented with *sigK* form a wild type resulting in identification of two loci on the chromosome that contain sigK-regulated genes. The first locus contains *sigK* and *Rv0444* (*rskA*) in addition to four genes including *Rv0447c* (*ufaA1*) that encodes a cyclopropane-fatty-acyl-phospholipid synthase, a putative amino oxidase together with *Rv0446*, *Rv0448c* and *Rv0449c* encoding for functionally unknown proteins. The second locus consists of six genes encoding membrane protein including *Rv2873* (*mpt83*) a surface-associated lipoprotein, Rv2874 (dipZ) an integral membrane C-type cytochrome biogenesis protein, *Rv2875* (*mpt70*) a major secreted immunogenic protein, *Rv2876* a transmembrane protein, *Rv2877c* an integral membrane protein, and *Rv2878c* (*mpt53*) a soluble secreted antigen (Charlet *et al.*, 2005).

The C-terminal domains are similar in *mpt70* and *mpt83* which mediate cells and extracellular matrix interaction. They are proposed to be involved in binding to proteins of host cells (Carr *et al.*, 2003). The *dipZ* gene is also controlled by σL has a conserved CcdA domain and involved in the biogenesis of cytochrome-c in bacteria (Juárez *et al.*, 2001).

The expressions of *mpt83* and *mpt70* are low in Mtb compared to high levels in *M. bovis*. This is largely due to a mutation in the anti-sigma factor *rskA* that controls posttranscriptional regulation of *sigK* leading to high level of *mpt70* and *mpt83* production in *M. bovis* (Saïd-Salim *et al.*, 2006). Whether overexpression of *mpt70* and *mpt83* offers a selective advantage for *M. bovis* is not clear.

The evolutionary analysis of sigK regulon indicates for the presence of a set of genes *mpt83* – *sigK* – *rskA* in a number of bacteria close to mycobacteria. Six genes have been inserted into this region during separation of Rhodococcus and Mycobacteria resulting in separation of *mpt83* from *sigK* and *rskA*. This locus was further separated during the evolution of mycobacteria into two loci; mpt70/83 locus and sigK – rskA locus. An additional gene, *dipZ* was inserted between two paralogs in slow-growing mycobacteria. Although, this regulon considerably varies across species, the sigK/rskA regulatory system across the mycobacteria genus is conserved (Manganelli, 2014, Malkhed *et al.*, 2011)

Expression of sigK regulon genes in macrophage

Different expression levels of genes in s*igK* operon have been observed during intracellular adaptation of Mtb. There is no published data regarding transcriptional changes for members of sigK regulon on interaction with neutrophils. Nonetheless, data from several macrophage infection studies demonstrate no significant expression of *Rv0448c*, *Rv0449c* and *dipZ* while, *mpt83* and *mpt70* have been found to be slightly up regulated (Table 5.1).

Table 5.1 Gene expression of sigK related – operon in macrophages

GENE	Expression in MΦ		Reference
	Differential Expression	Phagocyte Incubation time	
ufaA1	NA	NA	NA
Rv0448c	NS	Murine bone marrow MΦ 2hr	(Rohde *et al.*, 2012)
Rv0449c	NS	Murine bone marrow MΦ 2hr	(Rohde *et al.*, 2012)
mpt83	Significant	Murine bone marrow MΦ 24hr	(Schnappinger *et al.*, 2003) (Fontan *et al.*, 2008)
dipZ	NS	Murine bone marrow MΦ 2hr	(Rohde *et al.*, 2012)
mpt70	Significant	THP-1 cells 24hr	(Saïd-Salim *et al.*, 2006)

- NS: No significant difference
- NA: Not available
- MΦ**:** Macrophage

5.1.4. Aim and objectives

In this chapter the data obtained from RNA-seq are introduced and discussed. The primary aim was to assess the initial hypothesis that interaction with neutrophil can explain the transcriptome of Mtb in sputum. The objectives were:

- To detect differentially expressed genes unique to the Mtb-neutrophil encounter
- To determine the effects of RPMI on gene expression in Mtb
- To validate selected the RNA-seq results by qRT-PCR and also by comparing two normalisation methods namely RPKM and TPM

5.2. METHODS

5.2.1. Experimental design

The design of experiment was done based on optimised system for in vitro infection of neutrophils with Mtb as shown in figure 5.3. Briefly, log phase H37Rv was grown to stationary phase then 5ml suspension containing $5x10^7$ bacilli / ml was centrifuged at 1000xg for 10 minute and supernatant was removed. The pellet was re-suspended in 5ml RPMI containing 5% serum from neutrophil donor and incubated for 24h.

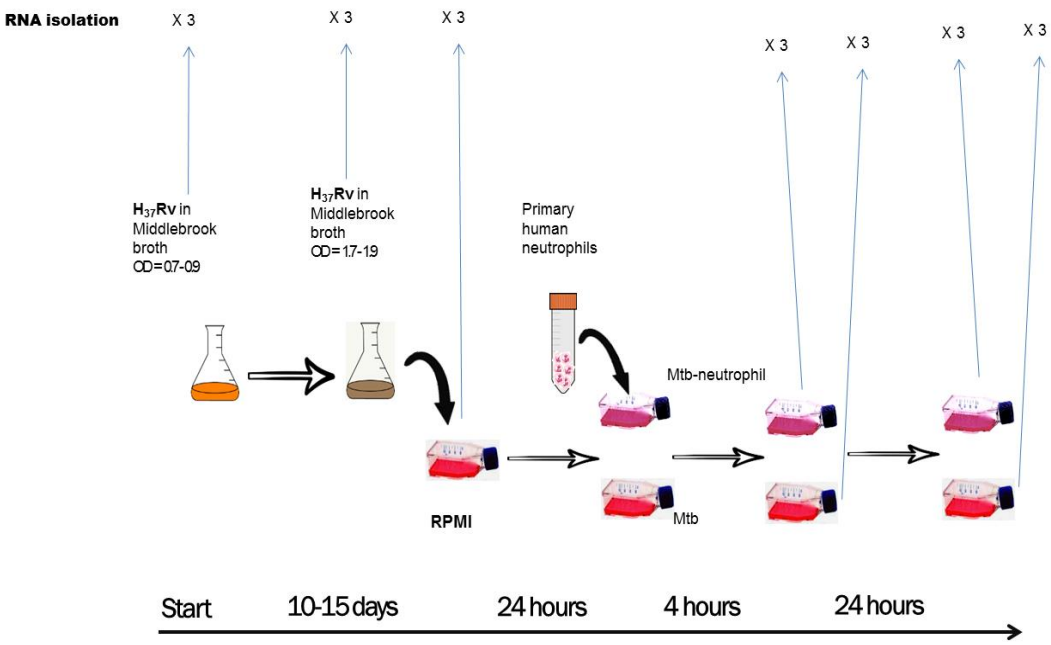


Figure 5.3 Schematic view of experimental design for collection of RNA samples. 24h prior to infection, stationary phase Mtb was pelleted and re-suspended in RPMI containing 5% human serum (isolated from the neutrophil donor). Each replicate was done with serum and neutrophil collected from same donor.

Primary human neutrophils were collected from 3 different individuals with exclusion criteria of TB infection (Table 5.1).

Table 5.2 Contributing neutrophil donors

	Gender	Background
Donor 1	Male	Middle East – Iraq
Donor 2	Female	Middle East – Iran
Donor 3	Male	West Africa – Ghana

Neutrophils were treated with IFN-γ for 1h after isolation from venous blood then mixed with bacteria in 25ml cell culture flask (MOI = 10) and incubated for 4h and 24h.

Controls of bacteria in RPMI only were set up in parallel to neutrophil samples. After incubation the suspension was immediately mixed with 20ml of 5M GTC solution and RNA extraction was carried out.

In addition, a sample of RNA from TB sputum was included in the RNA library for RNA seq. The sample was provided by Dr N. J. Garton and was from an untreated TB patient.

5.2.2. Measuring RNA integrity number (RIN)

The integrity and concentration of extracted RNA was assessed using Agilent RNA 6000 Nano-kit purchased from Agilent Technologies (Kidlington, UK).

5.2.2.1. Chips and reagents

Chips

RNA Nano Chips

Tubes for Gel-Dye Mix

Syringe Kit 4 Spin Filters

Tubes for Gel-Dye Mix

Syringe

Safe-Lock Eppendorf Tubes for gel-dye mix

Reagents

Agilent RNA 6000 Ladder

RNA Nano Dye Concentrate

Agilent RNA 6000 Nano Marker (2 vials)

Agilent RNA 6000 Nano Gel Matrix (2 vials)

5.2.2.2. Setting up the assay equipment and Bioanalyzer

Preparing the Gel

550µl of Agilent RNA 6000 Nano gel matrix was added into the top receptacle of a spin filter then centrifuged for 10 minutes at 4000 RPM. For each run 65µl was filtered into 0.5 ml RNase-free microfuge tube.

Preparing the Gel-Dye Mix

RNA 6000 Nano dye concentrate was mixed thoroughly for 10 seconds on vortex then 1µl was taken to a 65µl of filtered gel and mixed thoroughly. The gel-dye mix was then centrifuged at room temperature for 10 minutes at 14000 RPM.

Loading the Gel-Dye Mix

On a new RNA Nano-chip 9µl of the gel – dye mix was added to the well-marked black **G** by using a plunger that was positioned at 1ml to pressurise the spread of mixture. 9.0 µl of the gel- dye mix was dispensed into each of two wells marked as blue **G**.

Loading the RNA 6000 Nano Marker

5µl of the RNA 6000 Nano marker was pipetted into the well-marked with the ladder symbol and each of the 12 sample wells.

Loading the Ladder and Samples

1µl of the RNA ladder was pipetted into the well-marked with the ladder symbol. From each RNA sample 1µl was pipetted into each of the 12 sample wells. Then the chip was horizontally placed on the adapter of vortex and mixed for 60 seconds at 3000 RPM.

Starting the chip run

The chip was carefully placed into the receptacle of Agilent Bioanlyser 2100. In the list of program in Agilent 2100 expert software prokaryotic – Nano RNA was selected to run the chip. The results from electropherogram can be seen in the appendices.

5.2.3. Quality control of isolated RNA

The isolated RNA was assessed for quality as portrayed in figure 5.4.

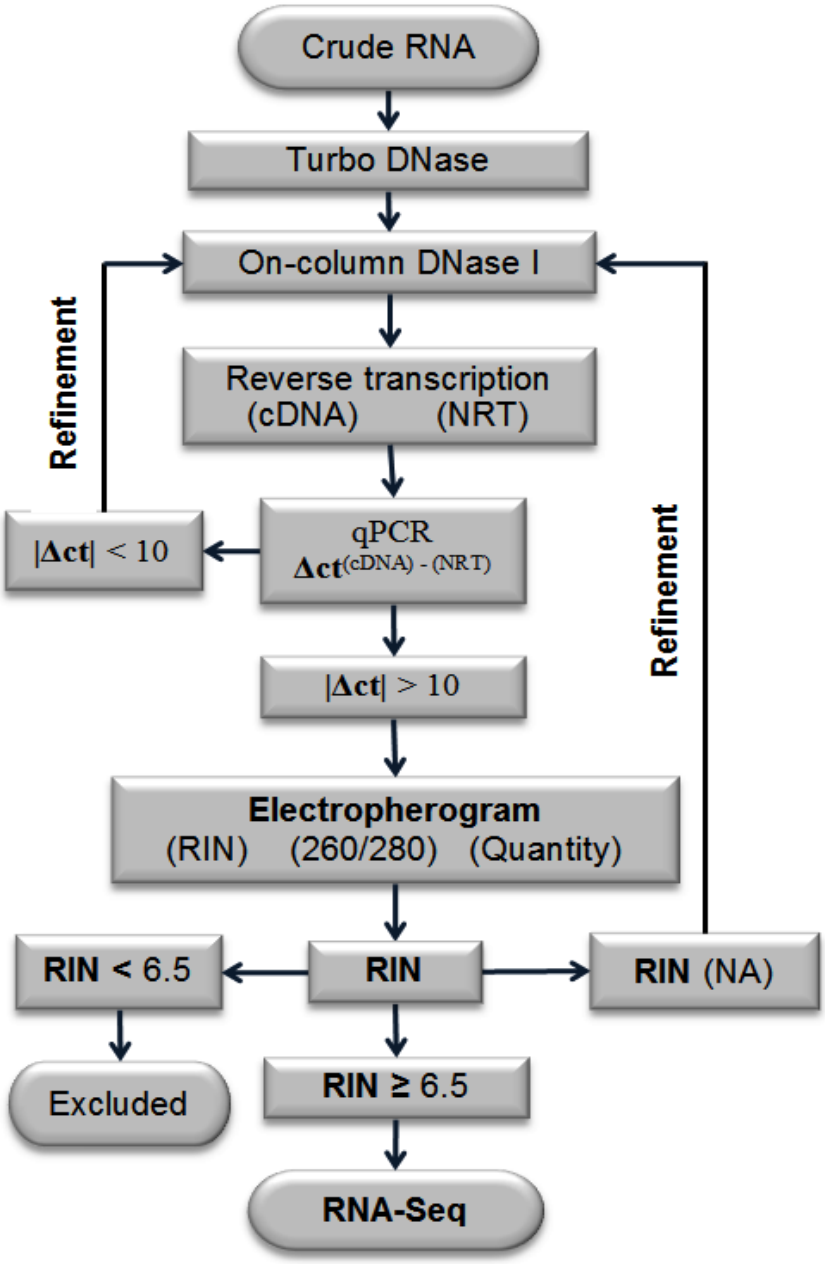


Figure 5.4 Assessment and selection of RNA samples for RNA-seq. The extracted RNA samples were DNase treated and the checked for genomic DNA contamination via qPCR. The RNA integrity was done to determine appropriate RNA samples for RNA-seq. Samples with genomic DNA contamination and no RNA integrity number were re-treated with DNase and re-assessed. Samples with low integrity (RIN < 6.5) were not used for RNA-seq

5.2.4. NextSeq-500 Illumina workflow

The RNA samples were sent for RNA-seq on Illumina platform (vertis Biotechnologie - Germany) recommended by Professor Jay Hinton (University of Liverpool). The sequencing workflow is integrated with BaseSpace software which is the genomics computing environment in Illumina for data analysis, storage, and collaboration. The full information on the platform can be found in NextSeq 500 System Guide (15046563 I) at http://support.illumina.com/.

Starting the sequencing run

RNA samples were reverse transcribed to make cDNA library of samples. The cDNA strands were fragmented into 75nt bases and then universal and specific adaptors were attached to fragment ends. A list of the adapters can be found in appendices. To perform a sequencing run, the libraries of cDNA with adaptors were denatured and diluted to 3pM then loaded onto reagent cartridge in a flow cell.

Cluster generation

Single DNA molecules were bound to the surface of the flow cell and amplified to make the clusters. The amplification ranged from 11 – 13 cycles for in vitro samples and 16 cycles for the sputum sample.

Sequencing

Single end sequencing was performed during which clusters were imaged via 2-channel sequencing chemistry and filter combinations that were specific to each of fluorescent-labelled chain terminators. The process of imaging was repeated for

each cycle of sequencing. The software performed base calling, filtering and quality scoring following image analysis.

Analysis

As the run progressed, the software automatically transferred files of base call (BCL) to BaseSpace for secondary analysis. FASTQ files were generated for each sample containing details of the sequencing including the quality score. These files were transferred to Leicester for further processing.

5.2.5. Trimming low quality reads

Trimming was done with the collaboration from Mr Ali Yavari in the Department of Infection, Immunity and inflammation as well as Dr Mathew Blades in Bioinformatics and Biostatistics Analysis Support Hub (B/BASH) in the University of Leicester.

The universal and Illumina specific adapters were removed in addition to removal of poor quality and repeating sequences using Trimmomatic version 0.32 (Bolger *et al.*, 2014). The quality score was set at Phred > 33 with the minimum acceptable fragment of 25nt. The output files were generated in FASTA and SAM format for downstream processes. The command line for trimming was written in Linux based text. A complete script applied for trimming in this study can be found in appendices. The FASTQ files before and after quality control were visualised using fastqc 0.11.2 program in the ALICE High Performance Computing Facility at the University of Leicester.

5.2.6. Alignment and normalisation of reads

The trimmed reads were aligned and mapped against H37Rv genome using Rockhopper 2.0.2 and RSEM 1.2.9 software.

5.2.6.1. Rockhopper processing of Reads

FASTA files generated after trimming of reads were loaded on Rockhopper software version 2.0.2 for windows (Wellesley – MA, USA). The files were three replicates per condition for in vitro samples while sputum sample was set as a condition with only one replicate. For alignment of reads and TPM analysis Bowtie2 was used to map the reads.

A dynamic programming alignment was applied when the reads did not match exactly to the reference genome based on the Smith–Waterman algorithm (Smith *et al.*, 1981) limited to 15% mismatches. These were corrected via insertion and deletion scores in the dynamic programming table based on the Phred quality score (Ewing *et al.*, 1998).The transcript abundance was measured by normalising the expression levels into RPKM.

5.2.6.2. RSEM processing of Reads

In addition to Rockhopper, RNA-seq by Expectation Maximisation program (RSEM version 1.2.9) was used based on RSEM-calculate-expression script for alignment of the reads and normalisation of transcripts. For alignment of the reads RSEM uses Bowtie2 software (bowtie2/2.2.3) and then normalises the expression levels into Transcript Per Million (TPM) by calculation of relative abundance(Li *et al.*, 2011). The program was run by using a perl script kindly provided by Dr Karsten Hokamp (Trinity College Dublin, Ireland) on FASTA files of samples.

5.2.7. Differential expression of genes

Only Rockhopper software was used for detection of differentially expressed genes. Before calculating the differential expression, the variance in expression of a gene was estimated via calculating the variance of the gene's expression across replicates.

As the variance is affected by the expression level, a highly expressed gene would basically have higher variance across the replicates. Therefore, local regression model (Anders *et al.*, 2010) was applied to the normalised counts to achieve smooth estimation of the variance. Then the differential expressions of genes between conditions were determined by performing statistical test for the null hypothesis. Thus, negative binomial distribution was used as a statistical model in which a two-sided p-value was computed as the probability of observing the expression levels of a gene in two conditions.

To control the false discovery rate due to multiple tests performed across the set of genes, q-values were set using Benjamini – Hochberg procedure (Benjamini *et al.*, 1995). A gene was designated as being differentially expressed if its expression in two different conditions was different at $q < 0.01$.

5.2.8. Visualisation

Intergraded genomics viewer (IGV) was used to visualise the expression levels in Rockhopper software (Thorvaldsdottir *et al.*, 2013).

R for graphics was used to visualise the correlations and multiple scatterplots (Robinson *et al.*, 2010). The packages and scripts used in this study can be found in the appendices.

String version 9.1 software available online at http://string-db.org/ was used to see the connection between Mtb genes as well as visualising the network of functional proteins.

Graph Pad prism was used to graph the data and perform basic statistics.

Raw-density design, an online program http://raw.densitydesign.org/ was used to display the data from long lists of differentially expressed genes.

5.2.9. Target Genes and primers

The expression levels of *esat6*, *Rv0691c*, *Rv0692a*, *mpt70*, *mpt83*, *dipZ* and *sigK* were measure using qRT-PCR. The list of primers including their sequence, length and melting temperature is available in the appendices (Page 188).

5.3. Results

5.3.1. Removing host nucleic acids

Total RNA was measured for samples of stationary H37Rv alone or mixed with primary human neutrophils to investigate the efficiency of GTC method in removing host nucleic acid. Nanodrop measurement following two rounds of DNase treatment showed that H37Rv only and neutrophil-H37Rv were not significantly different (Figure 5.5).

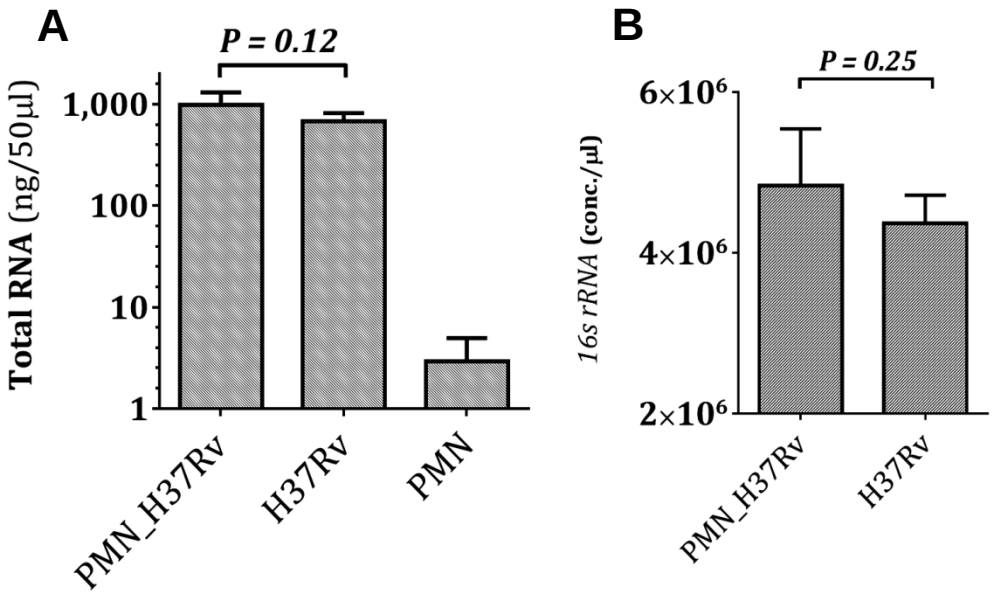


Figure 5.5 Neutrophil RNA was significantly depleted from extracted RNA. In 5ml RPMI 1.2x10^6 IFN-γ activated primary human neutrophils were infected for 4h at MOI=4.5±0.4 with stationary phase H37Rv at OD = 1.8. RNA was extracted and measured (**A**). Then the concentration of *16s rRNA* was quantified (**B**). Paired t-test was performed to determine the statistically significant difference between conditions. (Error Bars: SD n = 3)

5.3.2. Quality Control of Extracted RNA

The RNA samples were assessed for their integrity using both riboanalyser (a.k.a. electropherogram) and Nonodrop1000 spectrophotometer V3.7. Figure 5.6 shows the results from electeropherogram demonstrating the RNA integrity number (RIN) of samples ranged 6.5 – 10 with the median value of 8.3. The highest RNA integrity was found in exponential sample (Mean RIN = 9.8) and the lowest integrity was attributed to Neutrophil-24h sample (mean RIN = 7). The full detail of RIN for each sample is placed in appendix.

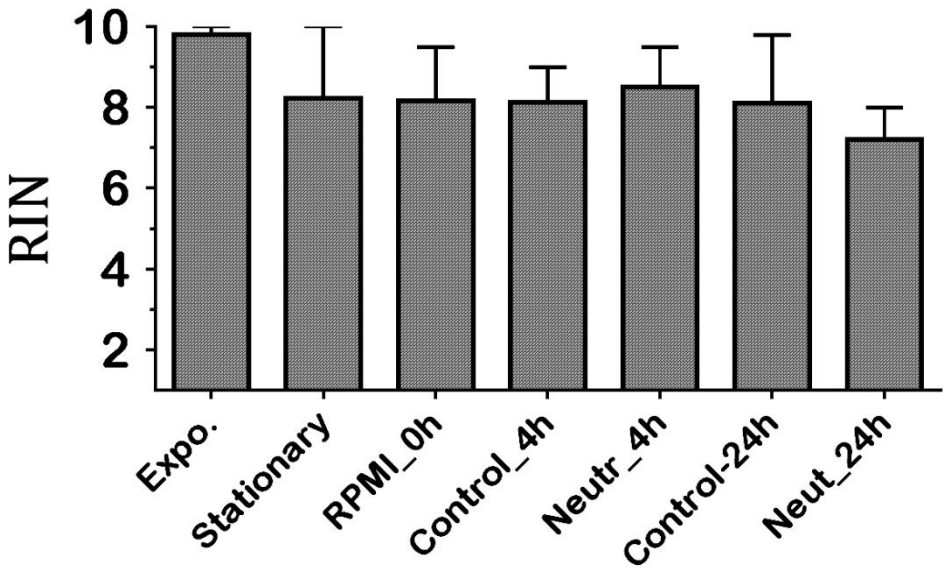


Figure 5.6 Integrity of RNA samples before RNA-seq. RNA Nano-chip was used to assess the integrity of RNA samples using electropherogram prior to RNA-seq

5.3.3. Sequencing results

After quality control, the RNA samples were sent for RNA-seq. The NGS library pool was sequenced on Illumina NextSeq 500 system using 75 bp read length.

Table 5.3 Reads information taken after RNA-seq

Samples	Repl.	Adaptor	PCR Cycles	Total sequences	GC %
Mid Exp. (7H9 broth)	*I*	ATCACG	11	9,284,377	56
	II	CGATGT	11	9,330,053	56
	III	TTAGGC	13	9,403,848	56
Stationary (7H9 broth)	*I*	TGACCA	12	9,435,878	56
	II	ACAGTG	12	9,404,391	57
	III	GCCAAT	12	9,547,909	56
RPMI 0h Stationary	*I*	CAGATC	12	8,481,541	56
	II	ACTTGA	11	12,150,666	56
	III	GATCAG	12	10,069,465	56
RPMI 28h Stationary	*I*	TAGCTT	12	14,246,063	56
	II	CTTGTA	13	11,977,949	56
	III	AGTCAA	13	10,836,780	56
RPMI 4h Stationary Neutrophil	*I*	AGTTCC	12	10,898,947	56
	II	ATGTCA	12	11,458,934	56
	III	CCGTCC	13	9,496,878	56
RPMI 48h Stationary	*I*	GTAGAG	13	11,039,856	56
	II	GTCCGC	13	8,721,199	56
	III	GTGAAA	13	11,537,670	56
RPMI 24h Stationary Neutrophil	*I*	GTGGCC	13	10,516,731	56
	II	GTTTCG	13	5,117,598	56
	III	CGTACG	13	10,116,473	56
Sputum	*I*	ACTGAT	16	12,477,797	49

The sequencing was done and a report was generated containing details of sequencing including sequence of adaptors, number of cycles, total number of sequences and GC content of sequences. The total number of sequences for all samples was 216,551,313 with the average sequence number per sample of 10,252,318. Additionally, the mean GC content of sequences for in vitro samples was 56% whereas the sputum showed 49% GC content. The reads were then assessed for base and sequence quality scores which were detailed in FASTQ files.

5.3.4. Read quality analysis

There are several steps in RNA-seq including preparation of sample, fragmentation, amplification and sequencing which make the technique really complicated. Even a modest error during any of these steps may result in a biased or useless data. Therefore, it is necessary to assess the quality of the RNA-seq data prior to analysis which starts with testing the quality of libraries. The quality scores and relevant information from sequencing output were formatted in FASTQ file. The FASTQ files were visualised using *fastQC v.0.11.2* software, a quality control tool for high throughput sequence data. Overall, less than 1% or all reads were revealed to be of low quality. The average base quality based on Phred score across all reads was 30 which positioned in the green zone designated for good quality scores. However, removing low quality reads, adaptors and K-mers (i.e. repeated sequences) is a prerequisite before alignment.

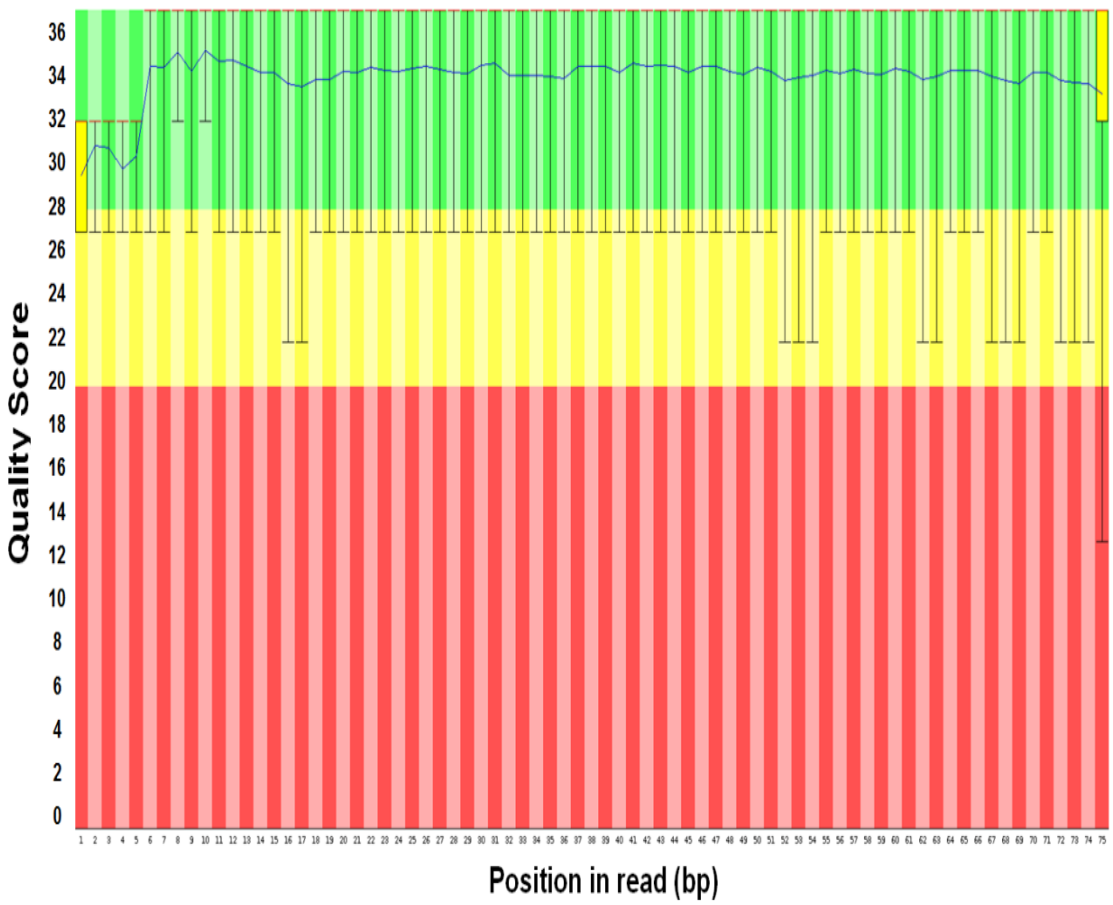


Figure 5.7 An example of per base sequence quality. The reads in FASTQ files were visualised using *fastQC v.0.11.2* software. The X-axis shows nucleotide position and Y-axis shows the quality scores for base call. The colours represent as green for good, yellow for moderate score and the red for low quality scores. The yellow boxes represent the inter-quartile range (25% - 75%). The upper and lower whiskers represent the 10% and 90% points. The blue line represents the mean value of the quality score.

5.3.5. Trimming the adapters and contaminating sequences

In the Illumina platform universal adapters, a sample specific adapter and small RNA adapters were used during clustering and sequencing. Also there is a possibility for the presence of contaminating sequences in such high throughput sequencing mainly originated from poor quality or repetitive sequences.

The adapters and low quality reads were removed from data set using Trimmomatic Software version 0.32. The result showed that for 21 in vitro samples less than 0.1% of total reads were trimmed while about 1% of reads from the sputum sample were deleted.

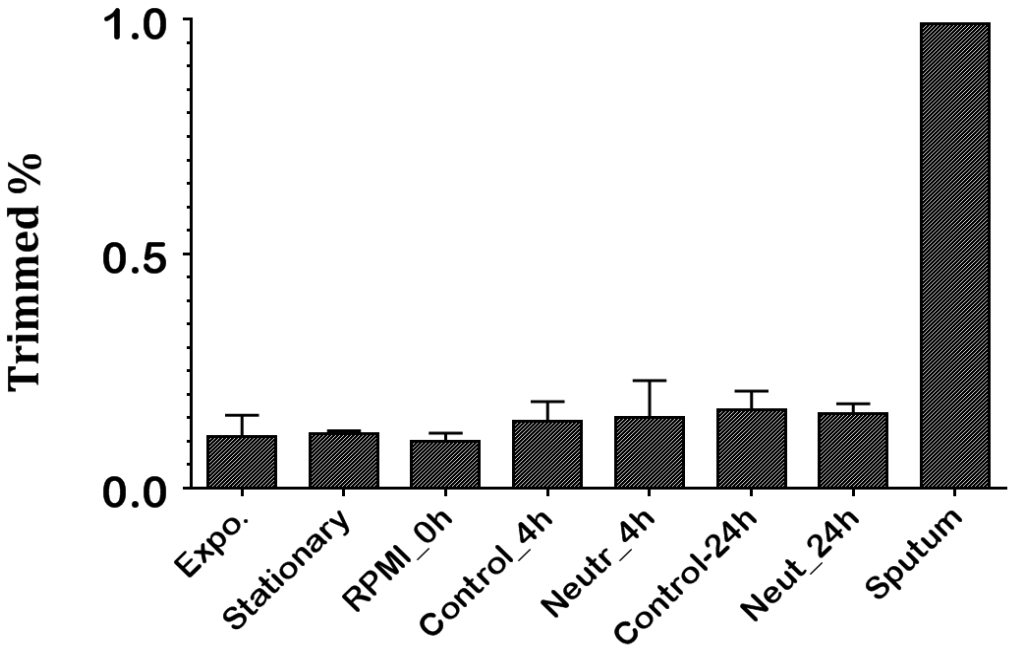


Figure 5.8 Very small amounts of sequences were removed after trimming of reads. Using Trimommatic v.0.32 adapters, K-mers, contaminating sequences and low quality reads below Phred = 33 were removed. (Error Bars: SD)

5.3.6. Alignment of reads

After trimming was completed the good quality reads were aligned against reference genome of Mtb H37Rv. This was done using Rockhopper version 2.0.2 (McClure *et al.*, 2013) that produced a mean mapping percentage of 97% across all the in vitro samples, whereas only 28% of all reads in the sputum sample were mapped to the H37Rv genome.

5.3.7. Genes with no detected expression

Rockhopper does not include zero expression values in its analysis. 90 open

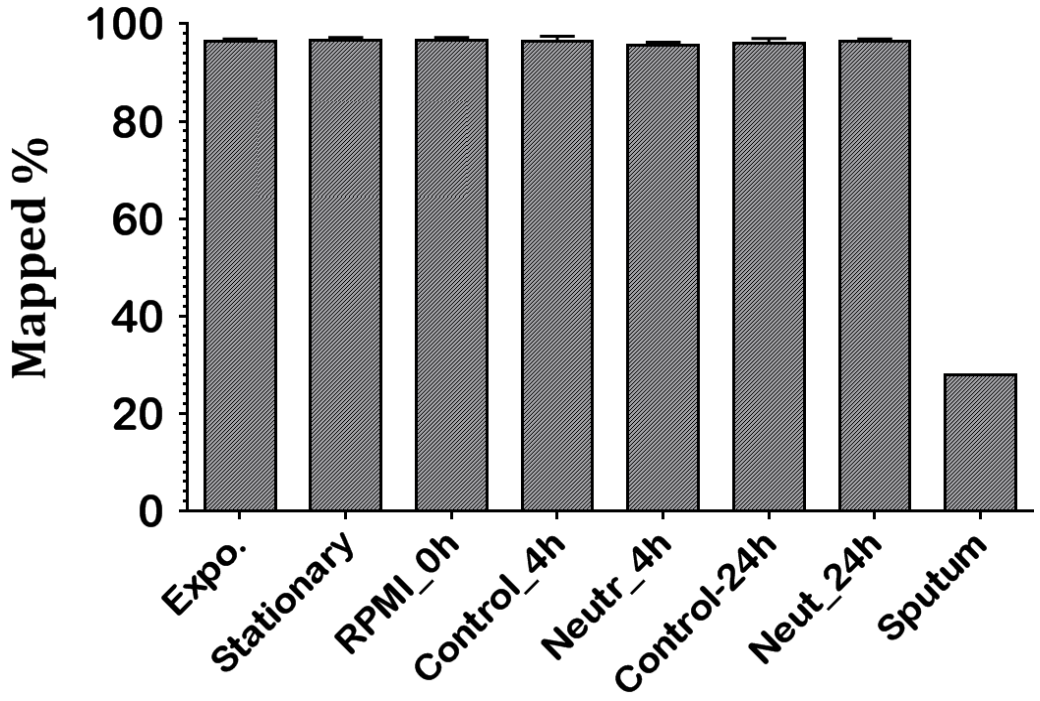


Figure 5.9 About 97% of reads were mapped to H37Rv genome while only 28% of sputum reads were mapped. Trimmed reads with good quality were aligned against reference genome of H37Rv using Rockhopper software version 2.0.2 and the proportion of mapped reads was presented in percentage. (Error Bars: SD)

reading frames (ORFs) were found to have absolutely no expression values in either in vitro or sputum samples. Subsequently, the genes were organised based on functional categories (Table 5.4). Conserved hypothetical, PE/PPE, and Intermediary metabolism and respiration had 29, 23 and 14 genes without expression respectively. While for Information pathways, lipid metabolism, and regulatory pathway only two genes were found to be not expressed.

Table 5.4 Genes with no expression detected

Functional category	**Genes with no detected expression**
Conserved Hypothetical	29
PE/PPE	23
Intermediary metabolism and respiration	14
Virulence, detoxification, adaptation	7
Cell wall and cell processes	7
Insertion seqs and phages	4
Information pathways	2
Lipid metabolism	2
Regulatory protein	2
Total	**90**

5.3.8. Proportion of ribosomal RNA

The bacterial transcriptome is predominated by ribosomal RNA (rRNA) creating major technical challenges in the study of microbial transcriptomic (He *et al.*, 2010). Since rRNA was not removed from RNA samples, it seemed essential to determine the proportion of rRNA from the mapped reads.

The results based on mapped reads to H37Rv genome revealed that the

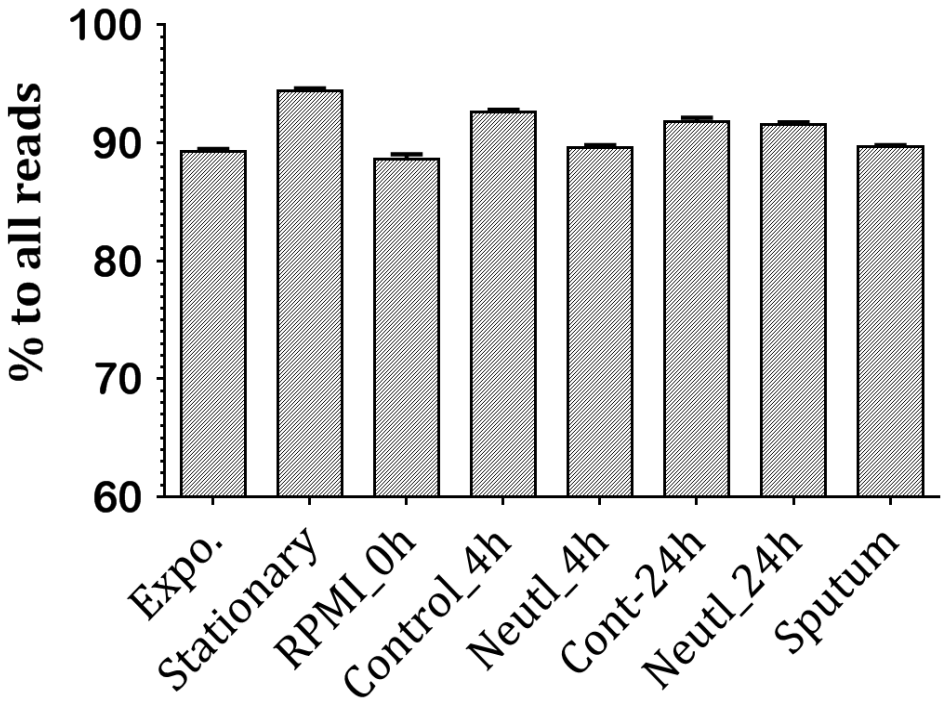


Figure 5.10 The proportion of rRNA is about 90% in total reads. Using Rockhopper software the reads were mapped to H37Rv genome and after computing the expression values the proportion of ribosomal RNA to all reads was calculated from three replicates for in vitro samples in addition to one sputum sample. (Error Bar = SD)

proportion of rRNA to total RNA was ranged between 88% to 95% in overall samples in vitro and in sputum.

5.3.9. Detected non-mRNA species

A considerable fraction of reads were mapped to regions other than predicted or confirmed ORFs. The number of these non-mRNA species were shown in figure 5.6 based on their categorical distribution. Thus, against a total of 4120 annoted

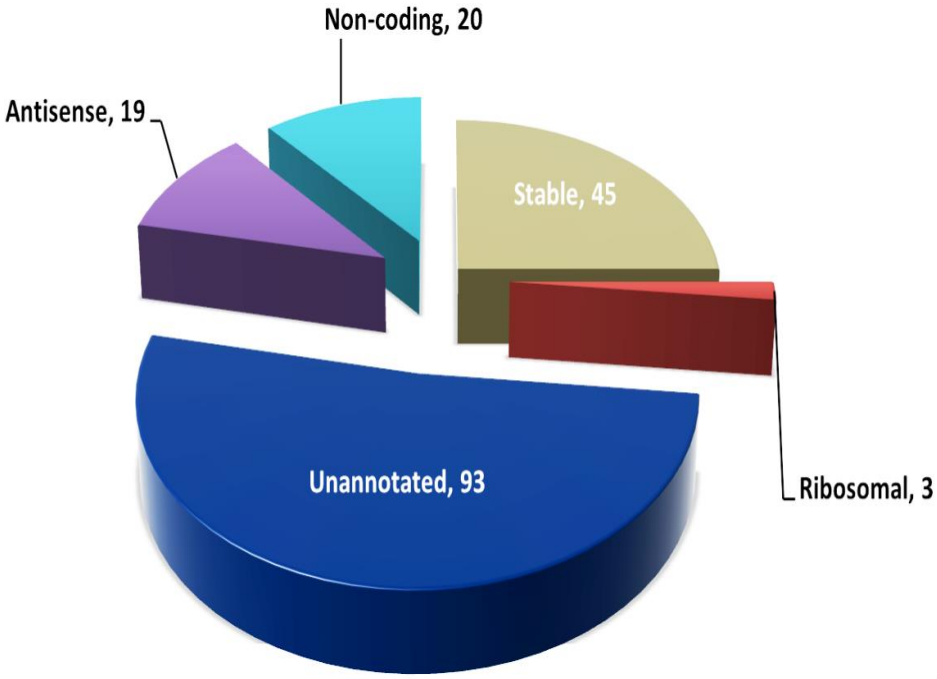


Figure 5.11 The Pie chart of different RNA species detected from RNA-seq data. The trimmed reads were aligned against H37Rv genome then a list of mapped transcripts was made containing annotated and non-annotated H37Rv genes. The non-mRNA species were displayed based on their frequency distribution.

ORFs, 3938 were polypeptide encoding. Of the remaining 182 non-mRNA transcripts; 93 were non-annotated (51%), 19 were anti-sense RNA (10.5%), 45 were stable RNA (25.3%), 20 were non-coding and small RNA (11.5%), and 3 were ribosomal RNA (1.6%).

5.3.10. Anti-sense RNA detection

The presence of anti-sense RNA in a transcriptome can have a significant impact on interpretation of the transcriptome. There were 19 detected anti-sense

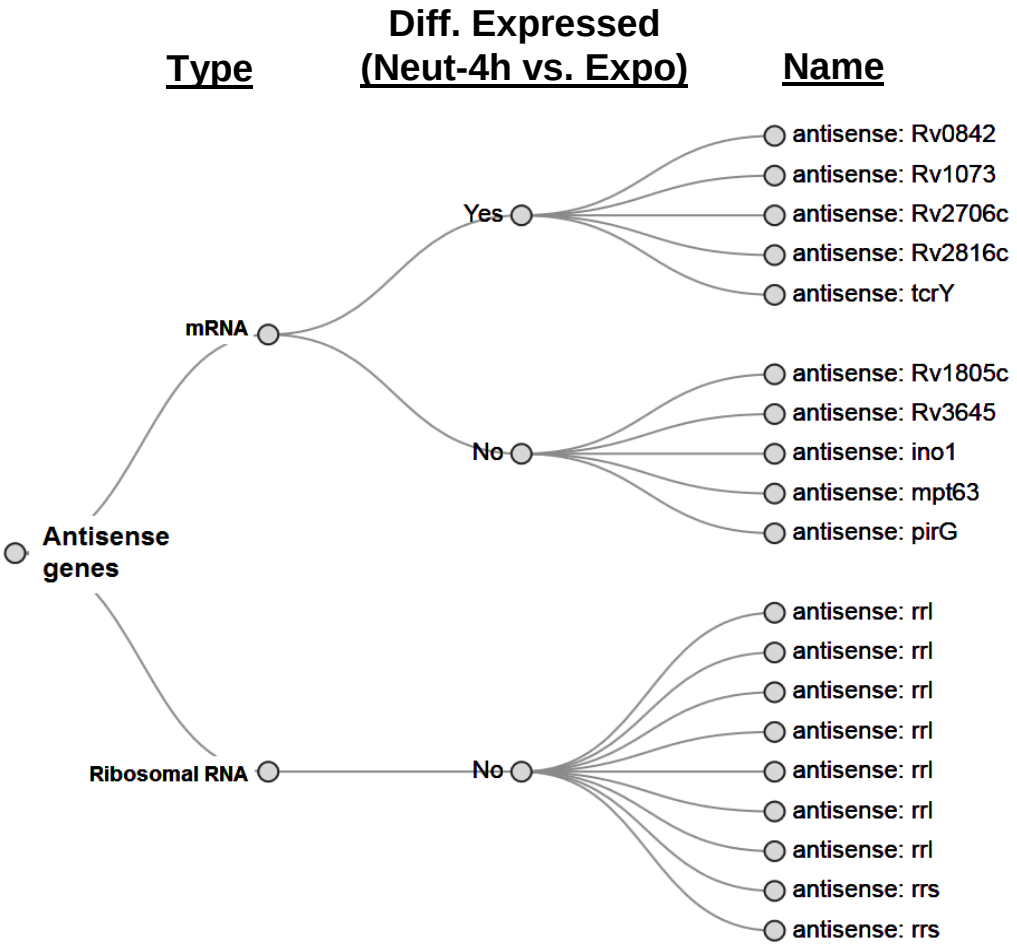


Figure 5.12 Mapping detected 19 antisense transcripts. Mapped transcripts were annotated Rockhopper then displayed according to their type and differential expression.

transcripts of which 10 were antisense-mRNA and 9 antisense-rRNA genes related to *rrl* and *rrs*. These were then investigated for their differential expression in neutrophil 4h versus exponential sample. In the group of antisense-mRNA, 5 were differentially expressed based on q-value < 0.01 and 5 were not differentially expressed. None of the 9 antisense-rRNA transcripts were found to be differentially expressed.

5.3.11. Expressions correlation across all conditions

The first task after performing quality assurance of RNA-seq data was the assessment of primary hypothesis that the transcriptional responses of Mtb in sputum were originated from interaction with neutrophils. Thus, total transcripts from 21 in vitro samples (i.e. 7 conditions in 3 biological replicates) and one TB sputum sample mapped against H37Rv genome were included in a comparison using their TPM values. The TPM values generated for a total of 4132 annotated and un-annotated genes were plotted on scatterplots. To compare the in vitro conditions and sputum sample all together, a multiple scatterplot was generated using R. 3.1.2 software. Spearman rank correlation was applied to each scatterplot which in turn exhibited a colour encoded correlation. Accordingly red indicated for highest, blue for medium and yellow for lowest correlations in the individual scatterplot.

The lowest correlation was found to be between sputum and exponential samples ($R < 0.66$) and the highest correlation was observed between 4h neutrophil and its 4h RPMI control ($R > 0.88$). Noticeably, the correlation between sputum and all in vitro conditions was about $R = 0.66$. Samples of neutrophil 4h, RPMI 0h, and RPMI 4h had correlation value of 0.66 compared to exponential. Whereas stationary sample, RPMI 24h and neutrophil 24h showed correlation of about 0.75 relative to exponential samples. The stationary sample was found to have correlation of 0.75 with individual neutrophil and RPMI samples. The correlation between 4h neutrophil and 24h neutrophil was about 0.88. Likewise, the correlation among 0h,

4h and 24h RPMI samples were 0.88. The test and control samples namely 24h neutrophil and 24h RPMI respectively showed to have a correlation of 0.88 when compared to each other.

5.3.12. Differentially expressed genes in 4h neutrophil encounters

The main question here was the pattern of Mtb transcriptional adaptation in response to the neutrophil encounter. The results showed that 90 genes were differentially expressed at 4h relative to control RPMI; of these 28 were down regulated and 62 were up regulated (Figure 5.14). The majority of genes were conserved hypotheticals followed by cell wall and cell processes. Other categories in order of frequency were intermediary metabolism and respiration, regulatory proteins, insertion sequences and phages, virulence and detoxification, lipid metabolism, PE/PPE, and information pathways.

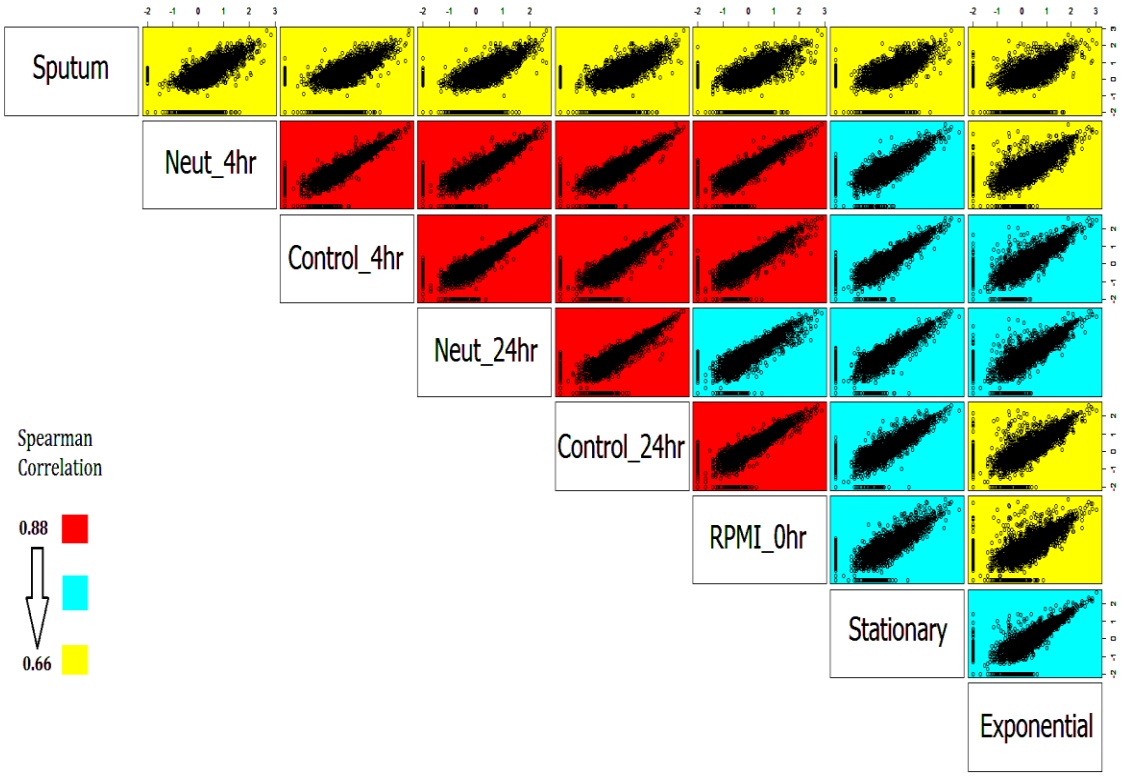


Figure 5.13 The lowest correlation was found between sputum and in vitro samples. All transcripts mapped to H37Rv were normalised as TPM using RSEM v.2.3.1 software, Using R for graphics version 3.2.3 the TPM values from annotated mRNA transcripts were first transformed to a matrix of log2 and then plotted into multiple scatterplots based on spearman rank correlation between samples. The minimum and maximum correlations were ranged as 0.66 to 0.88 respectively and the level of correlation was designated by colour as red indicating for highest blued indicating for medium and yellow indicating for lowest in the range of correlation.

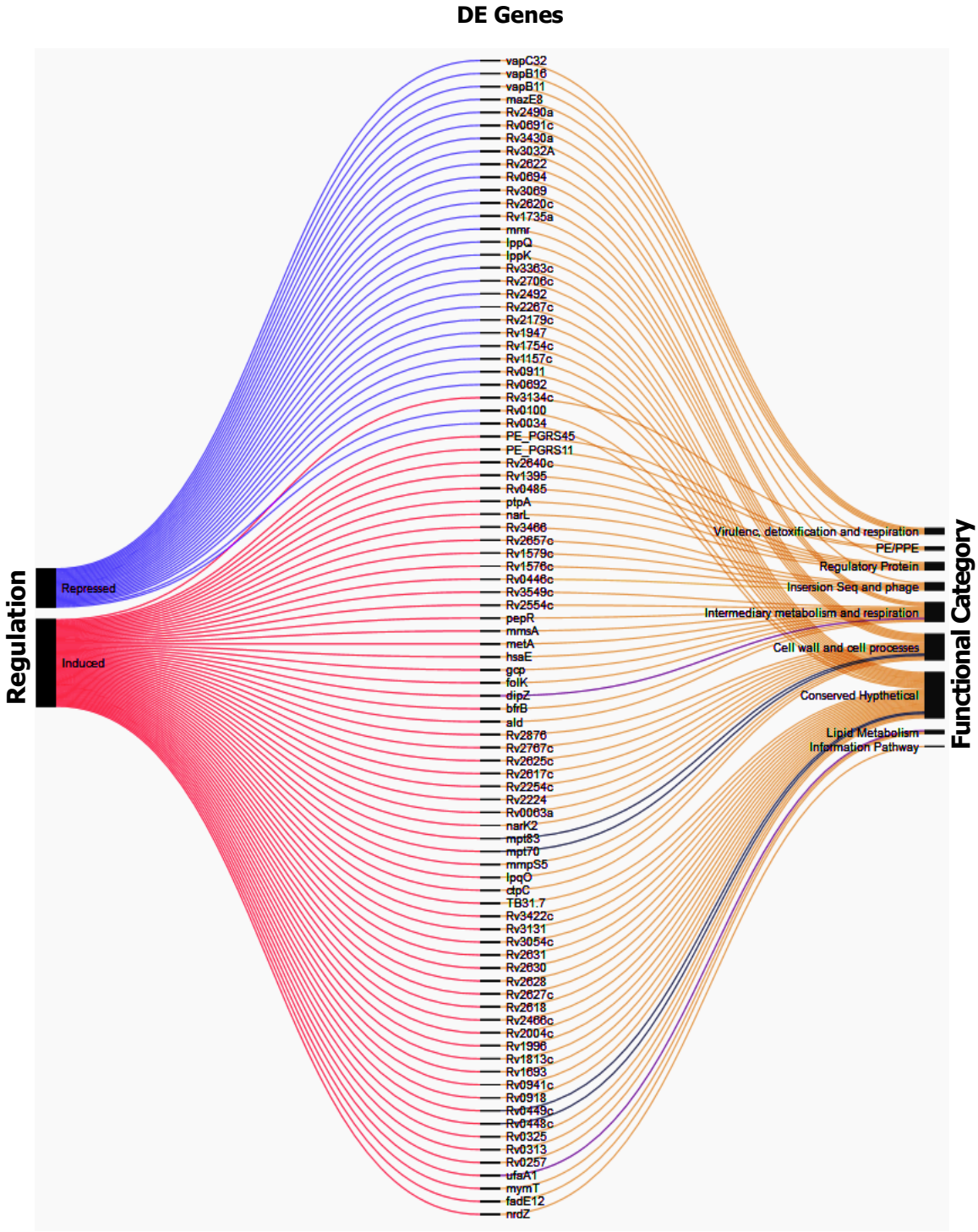


Figure 5.14 The parallel set diagram showing 90 differentially expressed genes detected at 4h neutrophil encounter. A list of 90 genes differentially expressed at 4h neutrophils relative to control RPMI was created based on q-value < 0.01. The genes were classified into induced and repressed then grouped into 9 different functional categories via searching Tuberculist website (http://tuberculist.epfl.ch/).

Following detection of genes exclusive to the neutrophil encounter, attention was drawn to the transcription factors known to affect their expression. Accordingly, a transcriptional regulation network spread-sheet tool made from experiments in which 206 transcription factors were overexpressed (TFOE) and the effects determined by microarrays, was applied to identify all transcription factors that altered the expression of target genes in these experiments (Rustad *et al.*, 2014). The differentially expressed genes were added to the spread-sheet which displayed for a number of regulatory genes significant association ($P < 0.05$) between their overexpression and differential expression of genes (Figure 5.15). Thus, 8 transcription factors were found to be associated with up-regulated genes of which *dosR*, *sigK* and *Rv1024* were dominant. Likewise, 11 transcription factors associated with down regulated genes among them *Rv0091c*, *whiB5* and *vapB40* found to have the highest association. Within three transcription factors of highest significant association *sigK* was found to be unique to the neutrophil encounter as *dosR* and *Rv1024* were recognised to be significantly involved in regulation of differentially expressed genes in other conditions. Therefore, genes under *sigK* control were investigated for their transcriptional regulation.

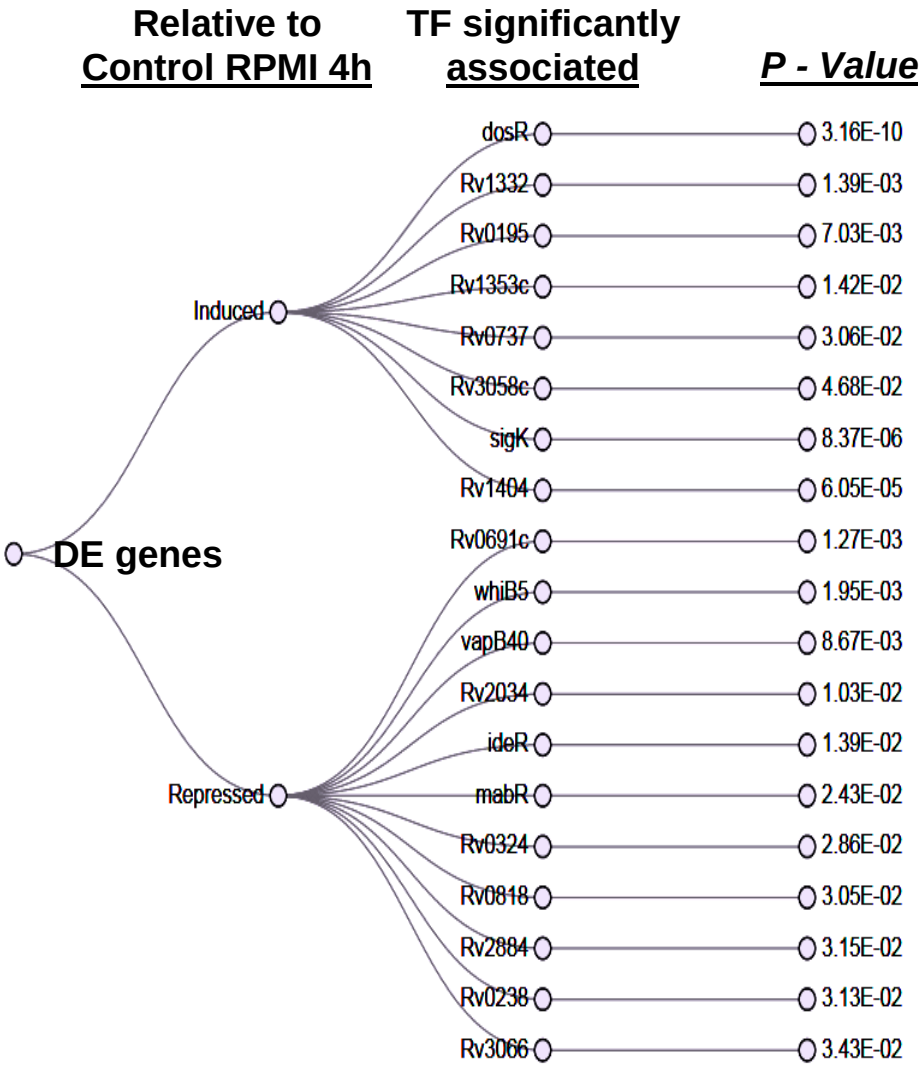


Figure 5.15 Transcription factors associated with 90 differentially expressed genes in neutrophils. 90 genes differentially expressed at 4h neutrophil infection were placed in a spreadsheet of overexpressed transcription factors. The transcription factors were assigned based on significant association of their overexpression with differentially expressed genes.

5.3.14. The sigK regulon

Six genes in the *sigK* regulon were found to be up regulated in neutrophils compared to the RPMI control. In an attempt to elucidate the connections, a total of 8 genes in sigK regulon were visualised using String v.2 software. The image (Figure 5.15) illustrates two loci in the regulon. The mpt70/83 locus composed of

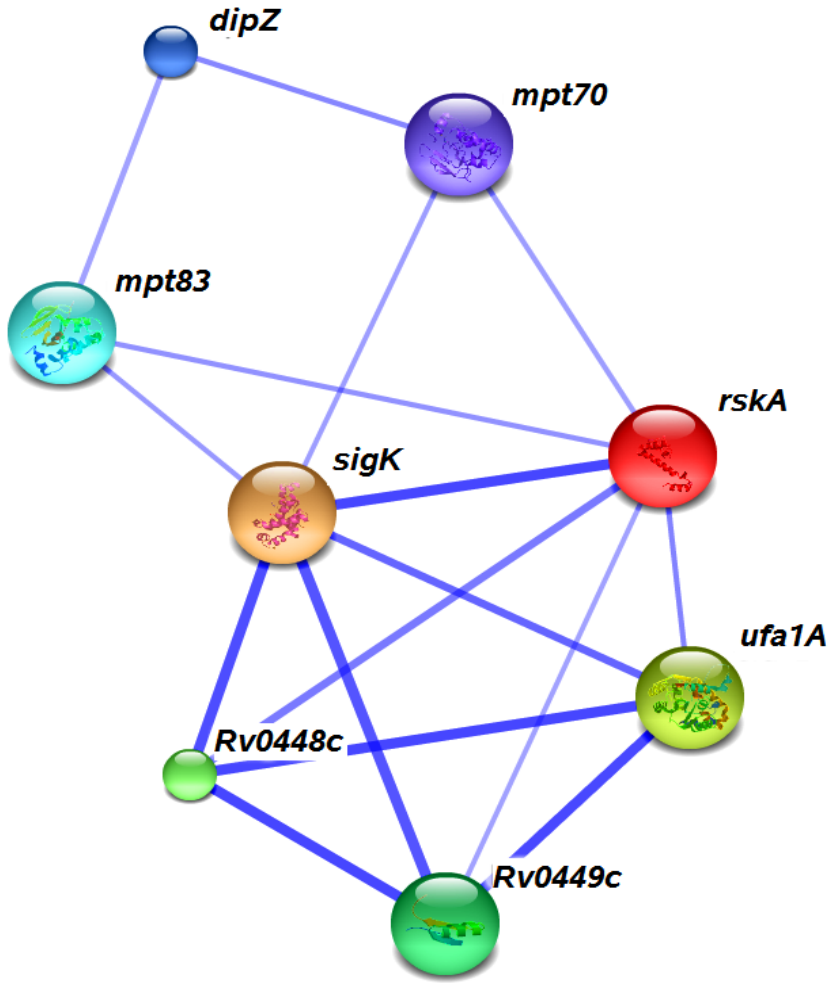


Figure 5.16 String view illustrating the connection between genes in the sigK regulon. Six differentially expressed genes at 4h neutrophil encounter, their main transcriptional regulator (i.e. *sigK*) and *sigK* regulator (i.e. *rskA*) were inserted into String to visualise their connections. Thicker lines represent stronger associations between genes.

dipZ, *mpt70* and *mpt83* which were in turn linked to *sigK* locus containing *sigK*, *Rv0448c*, *Rv0449c* and *ufaA1* genes. The regulation of sigK gene was found to be associated with *rska* which was also depicted in the string image.

5.3.15. Comparing TPM and RPKM of sigK genes

The basis for identifying differentially expressed genes was the use of q-values generated from Rockhopper RPKM values. RSEM software was also used for alignment of initial reads to H37Rv genome followed by normalisation of transcripts

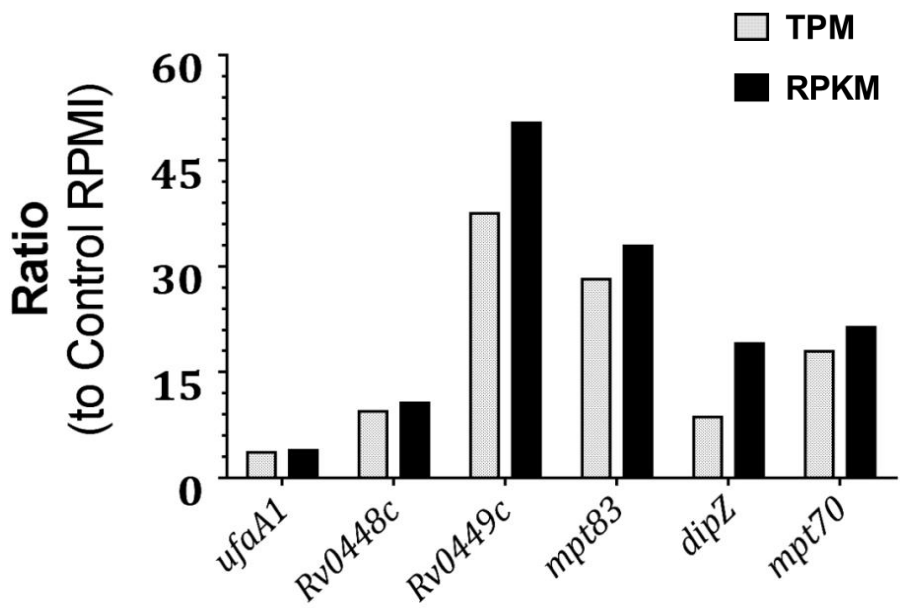


Figure 5.17 A strong correlation was found between TPM and RPKM values of differentially expressed genes. Spearman rank correlation was done to TPM and RPKM values of six differentially expressed genes at 4h neutrophil infection. The ratio between neutrophil encounter and control RPMI at 4h was illustrated.

in TPM. Spearman rank correlation was done to the normalised values of six differentially expressed genes in 4h neutrophil encounter in order to investigate the correlation between their RPKM and TPM values (Figure 5.17). The results indicated that RPKM and TPM values were correlated at R = 0.94.

5.3.16. Changes in differentially expressed genes from 4h to 24h of neutrophil encounter

Attempts were made to investigate the transcriptional changes of H37Rv between 4 and 24h during which period NET formation became complete. In this regard, 90 differentially expressed genes in 4h neutrophil infection relative to control were individually studied in the sample of 24h neutrophil infection in order to see the pattern of transcriptional remodelling of H37Rv during extended encounter with neutrophil. Thus, the list of differentially expressed genes was made on the same basis of q-value < 0.01 in 4h and 24h of neutrophil encounter.

The results indicated that out of 28 down regulated genes in 4h neutrophil sample none of them were further repressed, 16 had no change and 12 were induced; Whereas, in 62 up regulated genes at 4h genes, none were further induced, 18 had no change and 44 genes were repressed (Figure 5.18).

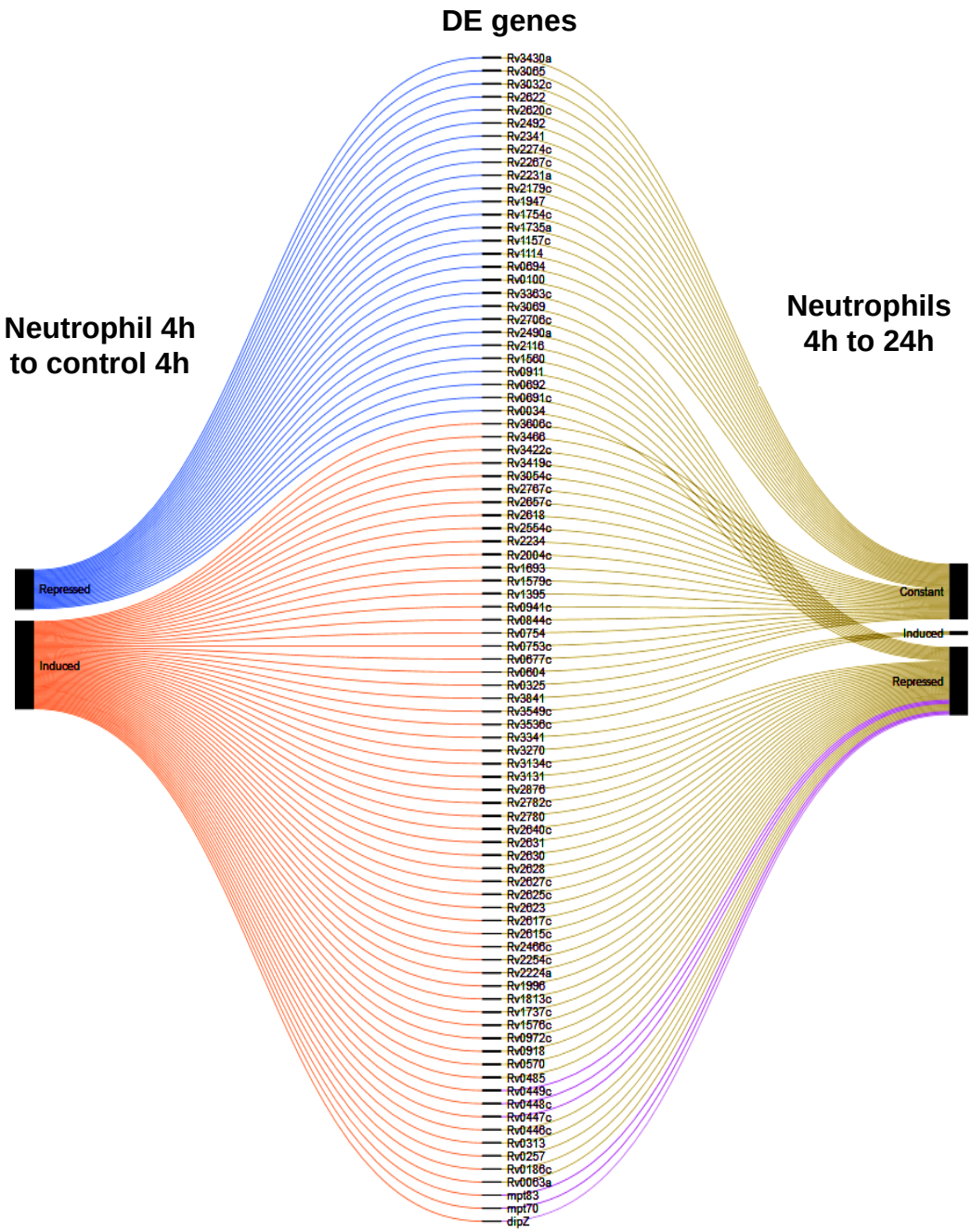


Figure 5.18 Majority of induced genes at 4h neutrophil infection become down regulated at 24h infection. 90 genes differentially expressed at 4h neutrophil infection were investigated for their expression after 24h of infection. The genes were then grouped into constant, induced and repressed clusters based on q-value < 0.01.

5.3.17. Confirmation of RNA-seq results by qPCR

RNA-seq has proven to be a powerful tool for bacterial transcriptome assessment. However, one of main concerns in analysis of RNA-seq results can be the robustness of data due to the complexities and multi-step processing inherent to the technique. Traditionally, qPCR has been the method of choice for validation of results in high-throughput gene expression profiling (Wang *et al.*, 2009c).

Therefore, qPCR was applied on 10 targeted genes with the aim to authenticate the results obtained from RNA-seq (Figure 5.19). The genes were selected from samples of 4h neutrophil encounter and its control of 4h RPMI on a ground that they should have a reasonably abundant expression levels (i.e. TPM > 10) from RNA-seq data. Out of 10 targeted genes, five were differentially expressed in 4h neutrophil relative RPMI control of which *dipZ*, *mpt70* and *mpt83* were up regulated whereas, *Rv0691a* and *Rv0692c* were down regulated. The other five genes namely *tgs1*, *hspX*, *icl1*, *esat6* and *sigK* were not differentially expressed (i.e. q value > 0.01).

The results revealed that *dipZ* was 3-fold induced in qPCR versus 5-fold induction in RNA-seq. Likewise, *mpt70* had 7-fold induction in qPCR as to 4-fold induction in RNA-seq. In other genes, RNA-seq and qPCR displayed a ratio of nearly 4-fold induction for mpt83, 2-fold and 3-fold repression for Rv0691a and Rv0692c respectively with slightly more reduction found in RNA-seq. There was no significant difference in expression of *esat6*, *tgs1*, *hspX*, *icl1* and *sigK* in the data from qPCR and RNA-seq. Spearman rank correlation demonstrated that there was a strong statistical correlation between data from qPCR and RNA-seq (R = 0.98)

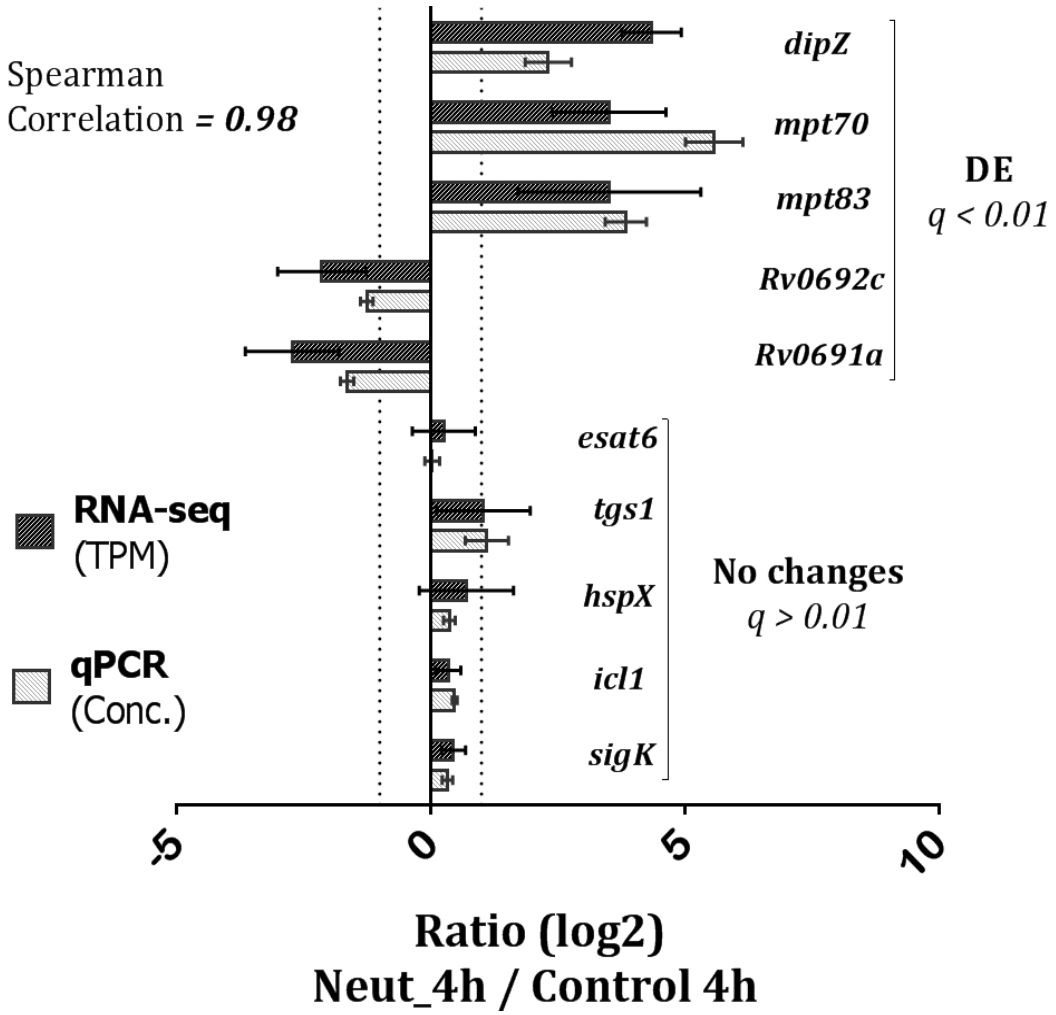


Figure 5.19 A strong correlation was found between RNA-seq and qPCR results. qPCR was done to 10 genes selected from RNA-seq data based on their differential expression in 4h neutrophil encounter relative to 4h RPMI control. qPCR was done to cDNA of RNA provided that the targeted genes had TPM > 10. The normalisation was done against *sigA* for both TPM and qPCR concentration values.

(Error Bars: SD n = 3 and further each qPCR was done with 3 technical replicates)

5.3.18. Expression of sigK regulon genes in RPMI and in neutrophils

In the list of genes differentially expressed in neutrophils the relative expression was calculated against RPMI control. As previously shown, RPMI has a substantial effect on transcription in Mtb. In this regard, 33 genes found to be associated with overexpression of *sigK* (Rustad et al., 2014) including six differentially expressed genes were investigated for their expression in neutrophils and control RPMI relative to exponential phase (Figure 5.20).

The results indicated that 22 genes were up regulated in RPMI as well as in neutrophil of which four genes namely *Rv0448c*, *Rv0449c*, *mpt70* and *mpt83* were found to be markedly overexpressed in neutrophil relative to RPMI. The expressions of *ufaA1* and *dipZ* were found to be constant in RPMI versus exponential; whereas they were significantly up regulated in neutrophils relative to RPMI.

5.3.19. Transcriptional changes of H37Rv at different RPMI time points

Prior to sampling from 4h and 24h of neutrophil infection the Mtb experiences vast transcriptional changes in response to RPMI. To address such effect of environmental changes on transcriptional response of bacilli, samples of RNA were extracted from stationary phase H37Rv and 3 subsequent RPMI exposure times (Figure 5.21). Thus, from 200 genes differentially expressed in RPMI 0h relative to stationary phase. 108 were induced and 92 were repressed. While, after 28h within the same 200 genes, 134 remained constant, 34 became up regulated and 32 were down regulated. At 48h RPMI the number of constant genes becomes 186 whereas only 9 genes were induced and 5 were repressed.

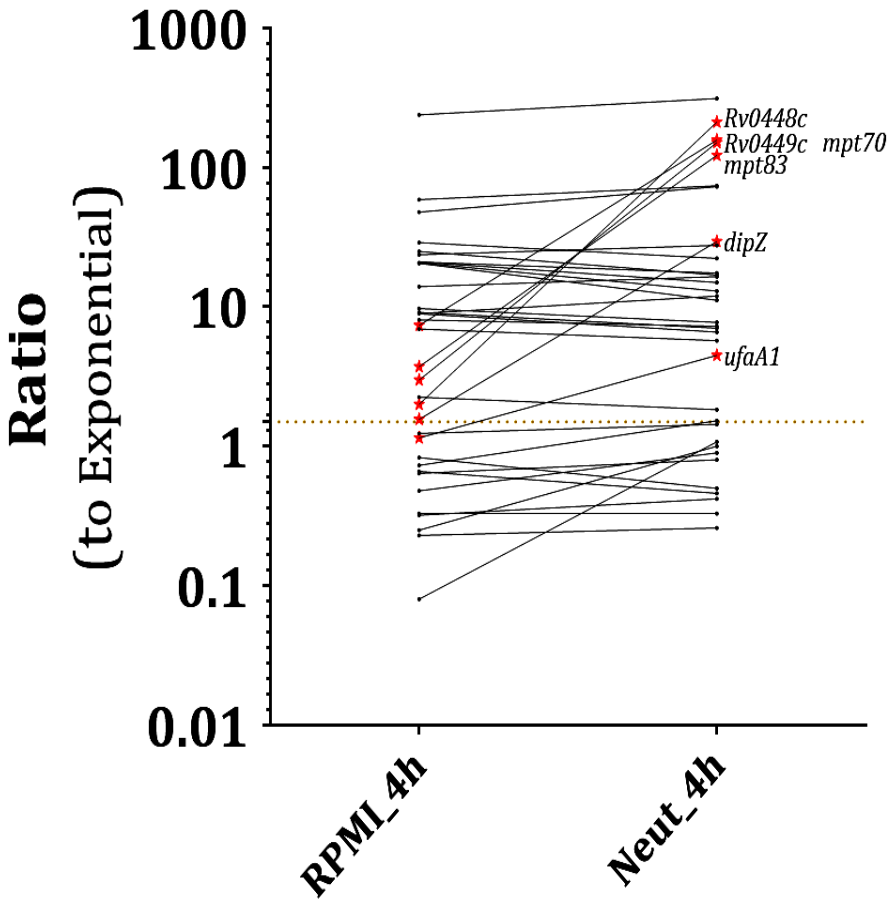


Figure 5.20 Majority of genes induced in neutrophils were also induced in RPMI relative to exponential condition. 33 genes associated with the overexpression of *sigK* were investigated for their relative expression in 4h neutrophils infection and 4h RPMI control against exponential growing H37Rv. Six differentially expressed genes in the sigK regulon were highlighted as further differential expression in neutrophils compared to RPMI

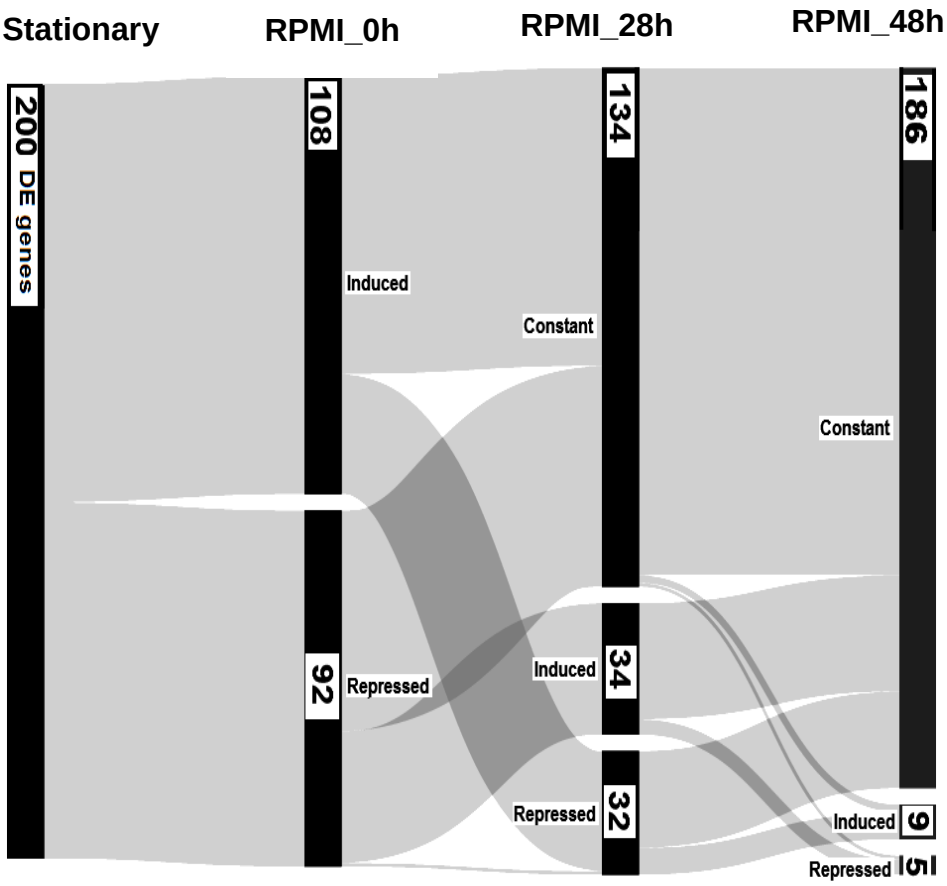


Figure 5.21 Effect of RPMI on expression was gradually reduced. 200 genes in stationary phase differentially expressed relative to log phase Mtb were investigated for their relative expression at 0h RPMI 28h and 48h.

5.4. Discussion

RNA-seq was chosen in this study because of its outstanding capacity compared to other technologies for deducing and quantifying a transcriptome. Basically this approach is capable of cataloguing all transcript species including mRNA, small RNAs and non-coding RNAs. Additionally, using RNA-seq, researchers have been enabled to determine the transcriptional structure of genes and post transcriptional modification as well as quantifying the expression levels of each transcript that change under various conditions (Wang *et al.*, 2009a).

The main task of this part of study was to identify differences in gene expression among multiple samples. Challenges were mainly related to the experimental issues such as integrity of the input RNA, presence of rRNA in the sample and selection of sample size that can influence downstream analysis. Therefore, efforts were made in order to tackle these issues via an experimental design to obtain a representative transcriptome of Mtb from in vitro infection of neutrophils.

5.4.1. Removing host nucleic acids

Attempts were placed to remove host nucleic acids by using GTC solution. The results demonstrated a four-log difference in the concentration of mycobacterial rRNA between H37Rv containing samples and neutrophil only samples. The results propose only a minor possibility for neutrophil-related RNA remained in samples. Also the difference in recovery of RNA from H37Rv and neutrophil-H37Rv samples was not statistically significant ($P = 0.12$).

5.4.2. Quality Control of Extracted RNA

Samples with RINs of less than 6.5 were excluded from the submitted RNA library and samples with genomic DNA contamination or no available RIN value were refined then re-assessed. Altogether, only three out of five separate experiments

were selected as to be appropriate for RNA-seq. Figure 5.4 illustrates the processes of assessment and selection of extracted RNA.

5.4.3. Sequencing results

Technical assessments of RNA-Seq indicated for a bias selection of sequences that depends not only on read length but also on CG-content where GC-poor and GC-rich sequences tend to be under-represented. Such bias can affect the accuracy of differential expression and also downstream analysis (Benjamini *et al.*, 2011). If this is the case, the results in this study may raise the issue of bias when comparing in vitro samples with sputum; as the GC content was about 56% for in vitro samples against 49% in sputum sample.

Two problems arise from the effects of GC-content on read counts. Firstly, within lane gene-specific effects in which the length of genes and their GC-content are specific to samples. Secondly, the between lane distribution differences (i.e. difference in depth of sequencing) can affect the read count (Risso *et al.*, 2011).

The first problem is moderated by proper normalisation of read counts by RPKM and TPM. Nonetheless, it is recognised that it is not possible to completely eliminate the bias resulting from length and GC-content (Dillies *et al.*, 2013). However, for the purposes of this study the TPM values were confirmed by qPCR and showed very good correlation (Figure 5.19).

The second type of problem in this regard seems not to be an issue as the depth of sequencing is relatively consistent for sputum and in vitro samples.

The single-ended sequencing approach was enough to attain the main goal of this part of study in identifying differentially expressed genes. In fact the adequacy of both depth and extent of sequencing allow achieving that. However, paired-end sequencing would have enabled more precise information on transcripts.

Nonetheless, the 75nt fragment was chosen here to suit the Illumina NextSeq-500 platform and complies with the recommended range of 30–400nt (Wang *et al.*, 2009c). Although the complexity of transcriptome assembly can be reduced by using longer reads (Martin et al., 2011), *a shorter length can reduce the number of low quality reads*.

5.4.4. Read quality analysis

Deletion of adapters or any contaminant sequence within a dataset is essentially a trade-off between confirming that all contaminating sequences are eliminated (i.e. sensitivity) while all non-contaminant sequence data remain (i.e. specificity). There is possibility that some contaminant fragments occur due to sequencing errors. Even a small fragment of a contaminant sequence within the read can create serious problem making the downstream analysis more complicated. In order to avoid that, the reads were assessed for their quality in a bioinformatics based approach (Figure 5.7). The overall quality of reads was acceptable with an average Phred score of 33. Nevertheless, trimming was absolutely necessary.

5.4.5. Trimming the adaptors and contaminating sequences

The percentage of removed sequences indicated that only a trivial amount of contaminants were present in the whole reads. Thus, 0.1%trimming for in vitro samples indicated the strength of Nextseq-500 in base calling and producing high quality reads.

As to the sputum sample, the amount of trimmed sequences was about 10-fold higher than in vitro samples. However, this amount of removal was expected due to the presence of other microorganisms in the sample.

5.4.6. Alignment of reads

The percentage of mapped reads was satisfactory. Rockhopper uses Bowtie2 as integrated alignment software which has been attested for its superior capacity compared to other alignment tools (Tjaden, 2015). Indeed, the results from mapping of reads by Rockhopper and RSEM in this study were comparable as both software use Bowtie2. An average of 97% mapped reads for the in vitro samples and 28% for sputum reasonably well represents their abundance of Mtb transcripts.

5.4.7. Genes with no detected expression

From a total of 4032 genes only 90 were found to have no expression in any of in vitro and in vivo samples. This small proportion reflects a high mapping capacity of Bowtie2 software. Nonetheless, there might be reasons other than mapping capacity for the absence of transcripts in the output.

A gene might have no detectable RNA before doing RNA-seq either due to low concentration at the time of sample collection or because of rapid degradation during preparation. Another reason could be technical issues in RNA-seq. For instance, it is possible that base calling of some sequences from a particular gene was particularly poor or generated overall low quality reads and resulted in their removal after trimming. However, undetected transcripts did not map differentially to particular functional categories. Indeed, all functional categories had at least 2 genes with no expression. In this regards the size and GC-content of genes were also examined but revealed no correlation with the absence of transcript (data not shown).

5.4.8. Proportion of ribosomal RNA

The prevalence of rRNA in samples representing 90% of mapped reads can also be a good indicator for the integrity of initial RNA. In fact, there is a difference in the stability of mRNA and rRNA based on structural differences. Under some circumstances RNases eliminate rRNA while mRNA remains intact (Mayne *et al.*, 1999).

Depletion of rRNA prior to preparation of cDNA libraries has been included in many RNA-seq protocols (Chen *et al.*, 2011, Filiatrault, 2011). The ratio of rRNA to total mapped reads was similar for in vitro samples and sputum. This indicates the relatively low quantity of mRNA in the sputum. It was for this reason that rRNA depletion was not applied in order to avoid further depletion of the already low mRNA. To address the initial hypothesis the in vitro samples were compared to sputum, thus the in vitro samples were also not rRNA-depleted.

5.4.9. Detected non-mRNA species

Detection of different types of RNA by Rockhopper in this study reflects a good mapping capacity of this software. In the same way, it can undoubtedly confirm the coverage of RNA-seq since a considerable number of non-annotated and non-coding species were also among mapped transcripts. Adequacy of coverage and sufficiency in depth of sequencing can also be validated from previous observation that only 90 out of 4032 total Mtb genes were found with no detected expression (Table 5.4). However, the coverage of RNA-seq has been shown to be extremely controversial in transcriptome study (Wang *et al.*, 2009c) and the depth of sequencing does not necessarily guarantee adequate coverage.

5.4.10. Anti-sense RNA detection

The anti-sense RNA detected in this study comprised antisense-rRNA and antisense-mRNA. The expression of protein by a transcript can be inhibited in the presence of respective anti-sense sequence either by steric interferences with mRNA complex or through mRNA cleavage under RNase activity (Worley-Morse *et al.*, 2015). Comparing neutrophil 4h condition against exponential phase H37Rv revealed that only some anti-sense mRNA genes were differentially expressed. The absence of differentially expressed in anti-sense rRNA may reflect the relatively constant expression pattern of rRNA in both conditions. Therefore, it seems that even in the presence of anti-sense sequences the ribosomal RNA had a steady-state expression.

A reverse direction of expression between differentially expressed anti-sense and its corresponding gene was observed. Thus, for up regulated anti-sense the sense mRNA demonstrated down regulation and the opposite occurred for down regulated anti-sense mRNA.

5.4.11. Expression Correlation across all samples

TPM values from all 4032 protein coding genes were taken into account to determine the correlation between different conditions in this study (Figure 5.13). As the vast majority of genes have constant expression across all samples, a high level of correlation was achieved. Thus, the correlation levels were significant in all comparisons. The correlation levels have been divided into high, moderate and low correlation in this particular case. The correlations between sputum and in vitro conditions were found to be at the lowest compared to the correlation among other samples.

Since the sputum result was found to be statistically different from the neutrophil encounter transcriptome, the results do not substantiate the primary hypothesis of

this study. Thus, in vitro exposure to neutrophils has failed to explain the distinctive transcriptome in TB sputum. A limitation here is the comparison between in vitro and in vivo samples. In fact, an ideal comparison would be an in vivo sample of neutrophil infection similar to the lung environment where antibodies and other immune mediators were present.

Moreover, detection of sputum DE genes in microarray study was done relative to exponentially growing H37Rv in vitro. While looking into 90 neutrophil DE genes in this study (Figure 5.14) the basis for detection was relative to control RPMI (i.e. stationary H37Rv in RPMI). Nonetheless, the pattern of expression neutrophil sample was found to be significantly different form sputum by RNA-seq and sputum by microarrays (Table 5.5).

Table 5.5 Comparing the expression of 90 neutrophil DE genes in sputum.

RNA-seq (this study) and microarrays (Garton *et al.*, 2008)

DE* genes in neutrophil 4h	#Sputum (RNA-seq)			#Sputum (Microarrays)		
	Up	Down	No change	Up	Down	No change
Up regulated (62 genes)	14	3	45	7	2	53
Down regulated (28 genes)	3	1	25	1	1	26

*Neutrophil 4h gene expression relative to RPMI 4h control

#sputum gene expressions relative to exponentially growing Mtb in vitro

Remarkably, the correlation of sputum to both stationary and log phase was similar. On the other hand, stationary H37Rv demonstrated more correlation with RPMI samples than exponential in general. This might be due to the use of stationary bacilli for neutrophil infection in RPMI samples. Noticeably, the RPMI samples were only moderately correlated with the stationary sample, indicating the substantial impact this medium has on transcription in H37Rv. This effect of RPMI was investigated by qPCR in the previous chapter and was further assessed in the RNA-seq data.

The multiple scatterplots demonstrated a high correlation between neutrophil time points and their corresponding RPMI control indicating that controls were representative for test samples. Additionally, the high correlation between neutrophil 4h and 24h indicates similarity of stimuli for H37Rv gene expression in neutrophils at both time points. Moreover, as neutrophil 4h was highly correlated with all RPMI conditions it can be assumed that any differentially expressed gene found relative to its RPMI control should be unique to the 4h neutrophil encounter.

5.4.12. Differentially expressed genes in 4h neutrophil encounter

The distribution of 90 differentially expressed genes from neutrophil 4h sample relative to control RPMI was proportional to the distribution of functional categories for all Mtb genes. This list of differentially expressed genes would be larger if comparison was done against log phase as the effect of RPMI is substantial on Mtb transcriptome.

In similar studies working with macrophages, the change in expression has been calculated relative to log phase (Dubnau *et al.*, 2002, Schnappinger *et al.*, 2003). Also, in Garton study which is the basis for the first hypothesis, detection of differentially expressed genes in TB sputum were based on their relative expression to log phase. Since this is the only study on *in vitro* transcriptome of Mtb in neutrophils, primarily the transcriptional modulations were based on

expression changes of H37Rv in neutrophil relative to neutrophil-free RPMI. This is very comparable to a study done by Rohde et al, who determined the differentially expressed genes in Mtb (CDC5115 strain) in macrophages against no macrophage (Rohde *et al.*, 2007).

In spite of the fact that sputum transcriptome seems to be more correlated with stationary phase in neutrophil than exponentially growing H37Rv, none of the comparison approaches can support the primary hypothesis that TB sputum is driven from neutrophil encounter. Moreover, there are a number of genes in the list of 90 differentially expressed genes that are also changed in variety of conditions. For instance, several genes are attributed to DosR regulon family found to be differentially expressed during hypoxia (Voskuil *et al.*, 2004), intracellular growth (Wang *et al.*, 2011b, Rohde *et al.*, 2012) and in *in vivo* conditions (Converse *et al.*, 2009). However, the only cluster unique to neutrophil encounter is a group of 6 genes in the family of *sigK* regulon. Thus, from this point the focus of further investigation was centered on *sigK* regulon as the particular transcriptional response of H37Rv on encounter with neutrophils.

5.4.13. Regulatory network of differentially expressed genes

The list of transcription factors was made by Rustad and colleagues on the grounds that overexpression of transcription factor was associated with induction or repression of a target gene (Rustad *et al.*, 2014). Among the key transcription factors significantly associated with induced genes by neutrophils at 4h were *DosR*, *Rv1404* and *sigK*.

As has been previously mentioned DosR induction can be attributed to a variety of conditions and is unlikely to be exclusive to neutrophils. Likewise, the genes associated with Rv1404 were too wide ranging to be designated as unique to neutrophil with the majority of them not detected as differentially expressed. It was *sigK* that associated with six differentially expressed genes and found to be

exclusive to neutrophil encounter. Although, the associated genes are located on sigK regulon, *sigK* itself was not differentially expressed in the neutrophil encounters.

The p-value distribution for all down regulated genes was relatively similar across all eleven regulatory genes. Such lack of hierarchy makes it difficult to designate specific transcription factors that might have most contribution in repressing genes during neutrophil encounter.

5.4.14. sigK regulon

The group of genes in *sigK* regulon that demonstrated overexpression unique to the neutrophil encounter are generally classified as non-essential for growth of H37Rv. Indeed, the exact function of sigK-encoding operon (*Rv0444 – Rv0449*) is unknown (Husson, 2006). However, a number of studies have been done to determine the function of these genes. The majority of them have been categorised as cell wall related and some have been considered as vaccine targets.

In this respect, *Rv2873* (*mpt83*) codes for a lipoprotein antigen (Hewinson *et al.*, 1996) and *Rv2875* (*mpt70*), codes for a major secreted immunogenic protein (Matsumoto *et al.*, 1995). They are membrane proteins that proposed as vaccine candidates (Chen *et al.*, 2012, Wiker, 2009). *Rv2874* (*dipZ*) codes for a membrane-bound protein involved in cytochrome-c biogenesis (Juárez *et al.*, 2001). *Rv0447c* (*ufa1A*) involved in cell membrane function and has methyl transferase activity; it has been proposed as a candidate in developing vaccines against both Mtb and *M. bovis* (Meena *et al.*, 2013). *Rv0448c* and *Rv449c* are functionally categorised as involved in production of conserved hypothetical proteins. While there are several immunogenic proteins encoded within the SigK regulon, their role in immunity and virulence is not clear.

One explanation for overexpression of immunogenic transcripts that are not necessary for growth may reflect an attempt by bacilli to recruit immune cells such as macrophages to the site of infection. Perhaps, it is more advantageous for intracellular bacilli like Mtb to interact with phagocytes than avoiding them. It seems that Mtb not only adapts to the interaction with polymorphs but also tends to provide an opportunity to extend its involvement with other immune cells. This appears to be happening through neutrophils which are dominant phagocytes in active pulmonary tuberculosis (Eum *et al.*, 2010). However, such strategy of Mtb seems to be not the case with macrophage interaction. Reviewing the published literature some of *sigK* regulon genes were found to be differentially expressed but at modest levels (Table 5.1) suggesting a different pattern of transcriptional response of Mtb in neutrophil and macrophage.

Despite the induction of sigK-operon genes at 4h interaction their main transcription factor was not differentially expressed. One possibility can be the up regulation of attributed genes after a transient induction of *sigK* which then gradually decreased to the level of no differential expression by 4h.

Rv0444 (rskA), located within the regulon and adjacent to *sigK* is the negative regulator for *sigK* (Saïd-Salim *et al.*, 2006). The nature of environmental signal to which *rskA* responds is not clear and the mechanism by which *sigK* is controlled is unknown (Husson, 2006). Based on RNA-seq results there is no sign of up or down regulation of *rskA* suggesting for the likelihood that the transcription of *sigK* – anti-sigK events might have been started at earlier times. Thus, a transcriptome determined within 2h of infection might better determine the interplay between *sigK* and *rskA*.

5.4.15. Comparing TPM and RPKM of sigK regulon genes

The results from measuring differential expression of *sigK* regulon genes were validated by comparing RPKM values against TPM. This comparison was done aimed at achieving two decisive conclusions.

Firstly, it was necessary to assess the accuracy of normalisation method used as base for detection of differentially expressed genes. It is generally believed that TPM normalisation is more accurate than RPKM (Wagner *et al.*, 2013). Notwithstanding this, RPKM was the chosen method due to the fact that the calculation has been incorporated in Rockhopper software. Rockhopper software was used because it is a specific tool for analysis of prokaryotic transcriptome while RSEM is primarily employed for normalisation of eukaryotic transcripts in TPM. To identify their correlation, RPKM values generated by Rockhopper were compared against TPM values produced in RSEM software. However, the ability of Rockhopper in estimation of transcript abundance has already been validated for gene expression of *N. gonorrhoeae*. This was done through comparing RNA-seq data in RPKM with qPCR quantitation (McClure et al., 2013).

Secondly, this comparison was done in order to re-confirm the efficiency of RNA-seq in detection of differentially expressed genes. Thus, significant correlation in identifying differential expression was observed when both RPKM and TPM were assessed. Some results were also confirmed by qPCR (Figure 5.19)

5.4.16. Changes in differentially expressed genes from 4h to 24h neutrophil encounter

A large number of neutrophils were found to be dead 24h after isolation; moreover, NET formation requires this time frame to be complete. The results from assessment of 90 differentially expressed genes in neutrophils at 4h revealed that

the majority of genes either remained constant or were repressed at 24h. This might be an indicative of adaptation of bacilli to neutrophil exposure after the initial oxidative burst stimulus and subsequent NET formation. Because of the restrictions in the Category III facility, a detailed time course of NET formation was not investigated here. However, slides prepared at 4h and 24h demonstrated more aggregated neutrophils at later time points, indicative of NET formation. This was evidently supported by the trypan blue test (Figure 3.6) which indicated no live neutrophils after 36h.

During in vivo infection, the trapped bacteria in NETs are eventually ingested by macrophages (Silva, 2011). The bacteria then have to further adapt to the new environment. In contrast, in the experiments reported here, the oxidative burst, degranulation and NET formation by primary cells affect Mtb in the absence of further influx of polymorphs or macrophages and it may be expected that the dominant environment will be that of RPMI in due course. Thus the effect of RPMI tends to increase when polymorphs are dead and their stimuli diminished. One more evidence that 24h is more RPMI relevant than neutrophil can be down regulation of *sigK* genes (Figure 5.19). As *sigK* regulon genes were found to be exclusive to 4h neutrophil encounter and up regulated relative to RPMI, their down regulation at 24h can be indicative for decreased expression relevant to RPMI exposure.

5.4.17. Confirmation of RNA-seq results by qPCR

Routinely, qPCR has been used for confirmation of result from global gene expression (Rienksma *et al.*, 2015). The results showed same direction of regulation in expression of genes with high degree of correlation between TPM values and qPCR concentrations. TPM values were selected because the comparison between conditions required a manual execution of statistics on raw data instead of the expression values produced by Rockhopper. Nevertheless, TPM values were found to be comparable to RPKM as shown in Figure 5.19.

Since confirmation of RNA-seq results was done on all three categories of induced, constant and repressed genes, it can be considered as representative for all detected genes; it might also have been desirable to further validate the RNA-seq by assaying the non-expressed genes by qPCR.

5.4.18. Expression of sigK genes in RPMI and neutrophil

The differential expression of a gene in neutrophil was assessed against control incubation in RPMI on the grounds that substantial changes occur during transition of H37Rv from log phase in 7H9 broth to the point of infection in RPMI. This was shown qPCR (chapter 3) and confirmed by RNA-seq (Figure 5.20). Reviewing 33 genes that are associated with the overexpression of sigK, it can be seen that majority of them were up regulated relative to log phase. Whereas, with the exception of sigK regulon members, other genes had no significant change relative to RPMI control. This adds weight to the view that activation of the sigK regulon is responsive to the neutrophil encounter.

5.4.19. RPMI time points

In addition to the points noted above it appears that Mtb transcriptional changes may never entirely discontinue in RPMI even with more extended time. However, the fact that the majority of genes show constant expression levels further supports the rationale for overnight incubation of bacilli in RPMI prior to infection.

5.5. CONCLUDING REMARKS

- The overall trimming of low quality sequences and adaptors was less than 1% of total reads
- The efficiency of mapping of whole transcripts to the reference genome of H37Rv was about 97% for neutrophil samples and 28% for sputum sample
- The frequency of genes with no detected expression was 90 from total of 4032 annotated genes
- The proportion of rRNA is about 90% in total reads
- 183 non-mRNA transcripts were detected 93 of which were non-annotated
- The primary hypothesis that neutrophils can explain sputum transcriptome of Mtb was not supported by the RNA-seq data
- 90 genes were detected as differentially expressed in neutrophil 4h sample of which six genes in sigK regulon were recognised as unique to neutrophil encounter
- Strong correlations were observed between RNA-seq and qPCR as well as between RPKM and TPM normalisation.
- A large proportion of differentially expressed genes at 0h of RPMI contact had constant expression after 24h.

CHAPTER SIX

General Discussion

6. General discussion

This study was done primarily to address two main questions. Firstly it was hypothesized that the interaction with neutrophil might be a key influence stimulating the reported sputum transcriptome of Mtb. Secondly, a general interest in understanding the pattern of global gene expression in Mtb cells on their encounter with neutrophils. In this regard, it would be better to isolate neutrophils from TB sputum to compare with sputum transcriptome in the first hypothesis. Nevertheless, due to the short half-life of bacterial transcripts combined with the difficulties of neutrophil isolation from TB sputum, the neutrophil – Mtb interaction experiments were designed in vitro. To this end, the initial work was done to achieve the optimal neutrophil – mycobacterial interaction most representing in vivo condition.

The primary in vitro experiments were performed with HL60 cells in order to control the variations other than those expected to result from using isolated neutrophils from different individuals. Only differentiated HL60 cells were used that were confirmed by microscopic examination of stained ATRA-treated HL60 cells.

It was decided to use Mtb (H37Rv strain) for neutrophil infection. H37Rv is a virulent strain that has been used in a large number of studies representing a standard laboratory strain of Mtb (Bermudez *et al.*, 1996). This is particularly the case in high-throughput sequencing as the genome of H37Rv serves as the reference genome for mycobacterial species (Kohl *et al.*, 2014). It was necessary to conduct experiments in the category 3 containment, making this a cumbersome task. For this reason initial experiments were done with *M. bovis* BCG. Infecting neutrophils with BCG provides insight into the interaction of low virulence mycobacteria with neutrophils.

Efforts were made to enhance the interaction of polymorphs and mycobacteria via adjusting the MOI. Thus, the best MOIs suitable for neutrophil infection were 5 and

10 at which neutrophils had the higher viability compared to MOIs of 50 and 100. Using low MOI would also decrease the variability in results due to the general tendency of bacteria for clumping at dense populations. It should also be noted that clumped mycobacteria may experience a different phagocytic pathway than individual isolated bacteria (Schuller et al., 2001).

With regard to assessing mycobacterial numbers via CFU counting it should be appreciated that this has been a universally applied method particularly in assessment of the antibacterial activity of phagocytes. However, CFU counts for Mtb are a very time consuming procedure and are susceptible to poorly controlled variations due to a number of factors. Firstly, the bacilli have tendency to clump due to their hydrophobicity which may vary in different types of medium and with using different medium components (Kairo et al., 1999). Secondly, the effect of phagocytosis on CFU counts should be considered and this is profoundly relevant to the later experiments in this study.

It is recognised that, in the process of cell lysis for release of intracellular bacilli, several bioactive elements could be introduced that may influence the CFU counts (Lowe et al., 2013), nonetheless, this method was used here and these limitations were kept in mind. This assay can help differentiating between intracellular and extracellular bacilli but does not necessarily discriminate between viable and non-viable bacilli. This can be a substantial concern when the read out is done via CFU counts as under different environmental experiences the bacteria may produce different populations of culturable and non-culturable bacilli (Barer *et al.*, 1999).

One of the first steps in characterising neutrophil-mycobacteria interaction was confirmation of phagocytosis. This step was especially important at the starting point of the project as only intracellular signals were targeted for investigation. Thus, different microscopic approaches were considered such as visualisation of GFP expressing BCG as well as FITC labelled antibody to detect intracellular BCG in infected HL60 cells. However, the results were not confirmatory for localisation

of bacilli as to the phagocyte. Therefore, amikacin treatment was applied to detect internalised bacteria as treatment with this antibiotic is an accepted method for killing extracellular bacteria without affecting those inside phagocytes.

Although, antibiotic treatment was a satisfactory approach for determining the ratio of internalised versus extracellular bacteria, it was not implemented in the subsequent experiments when both groups of mycobacteria whether inside or outside were found to be of equal importance. The main reason in fact was the formation of neutrophil extracellular traps upon which antibacterial effects can be equally presented to bacilli regardless of their location (Hahn *et al.*, 2013). Moreover, the sputum transcriptome to which the in vitro samples will be compared was not specifically taken from intracellular mycobacteria (Garton *et al.*, 2008).

To further simulate conditions thought to pertain in TB sputum and to obtain relatively time stable transcription patterns at initiation of the in vitro neutrophil encounter, a stationary phase Mtb inoculum was used. Whether or not the bacteria present in sputum represent persisters, it is worth noting that such bacilli are considered to be non-replicating (Mitchison, 2004).

While the reproducibility in results of gene expression quantification improved by using stationary phase inoculum, it was realised that RPMI had a significant effect on Mtb transcription. This was observed when H37Rv was incubated for 120h in RPMI leading to a notable increase in expression of *tgs1*, *hspX*, and *icl1* relative to log phase within 24h (Figure 4.3). However, with the extension of incubation from 24 to 120h the transcriptional changes were relatively stabilised. This was further confirmed by RNA-seq data (Figure 5.21). Thus, a large proportion of differentially expressed genes in RPMI at 4h turned to out to be constantly expressed within 24h then the proportion further increased by 48h. As the magnitude of changes is at highest during early hours of incubation in RPMI, it was necessary to consider the acclimatisation time for Mtb. Indeed, very substantial changes in gene expression started before introducing Mtb to RPMI. This transcriptional response

was induced during the transition of bacteria from culture (i.e. 7H9 broth) into a test tube followed by centrifugation then replacement of the bacterial medium with RPMI. Adding these factors together provides a potential explanation for the poor reproducibility of results in early experiments. Thus, in subsequent experiments the inoculum was prepared from stationary Mtb incubated in RPMI for 24h prior to infection of neutrophils.

The potential effect of neutrophils on Mtb was recognisable as early as 90 minutes after initiating the encounter. This was demonstrated by measuring the oxidative burst in infected neutrophils. Accordingly, the first time point for RNA isolation was set at 4h during which period the Mtb transcriptome would be expected to reflect the neutrophil encounter. NET formation as consequence of oxidative burst can be mediated by ROS in activated neutrophils (Remijsen *et al.*, 2011). Therefore, a later time point of incubation was set at 24h in order to investigate the effect of established NETs on Mtb gene expression.

The in vitro neutrophil-Mtb infection system was established after a series of optimisations aimed at achieving good reproducibility. The ultimate goal was to isolate RNA from Mtb on its encounter with neutrophil and prepare a library of RNA suitable for RNA-seq. For this reason, only purified RNA with high integrity number (RIN > 6.5) was considered for RNA-seq.

RNA-seq was used in this study because it is a powerful approach for profiling the global gene expression. Indeed, it become a popular method for bacterial transcriptome study due to the decreasing cost of sequencing, availability of bioinformatics tools to the public, and accessibility of high-throughput sequencing (Haas et al., 2012). Nonetheless, like any method RNA-seq has disadvantages mainly attributed to its complex processing and inherent biases during mRNA enrichment, fragmentation steps and cDNA synthesis (Martin *et al.*, 2011).

Although, depletion of rRNA prior to preparation of cDNA library has been considered in many RNA-seq protocols (Chen *et al.*, 2011, Filiatrault, 2011), there

has been no attempt to remove rRNA in this study. In fact, the removal of rRNA was not recommended due to the low quantity of RNA extracted from sputum samples as loss of mRNA could occur during rRNA removal. Since, based on initial hypothesis, the in vitro samples should be compared to sputum they were also not depleted from rRNA. However, there are examples of similar RNA-seq studies including a study by Cortes et al in which the rRNA depletion has not been performed (Cortes *et al.*, 2013). In view of the general interest in this study that detection of none annotated transcripts was not the area of interest and combined with the fact that the depth of sequencing was reasonably sufficient for coverage of known transcripts, it was realised that rRNA depletion is not necessarily important.

The sequencing was carried out in a single ended approach which is less accurate than paired ended sequencing. Because, in this study the abundance of known transcripts and their differential counts are the main targets rather than detection of new transcripts for which two times of base calling should be applied using paired ended sequencing.

The normalisation of reads for RNA-seq data was done in TPM as well as RPKM. While TPM is the relatively more accurate normalisation method compared to RPKM, their difference is not significant for detection of differentially expressed genes. Therefore, RPKM was used when complicated statistics were applied to detect differential expression whereas to find the correlation between two conditions or comparison between two methods TPM values were used. Nevertheless, a high correlation exists between TPM and RPKM normalisation (Wagner *et al.*, 2013).

The correlation between various conditions based on their TPM values revealed a low correlation between transcriptome of sputum and neutrophil 4h. That means the distinctive pattern of gene expression in sputum is not predominantly driven from neutrophil encounters comparable to those achieved in vitro here. Thus, these data do not support the first hypothesis that the interaction with neutrophil

can explain the sputum transcriptome. However, it should be considered that the neutrophil infection system did not include several potentially important influences expected in vivo such as antibodies and other immune mediators.

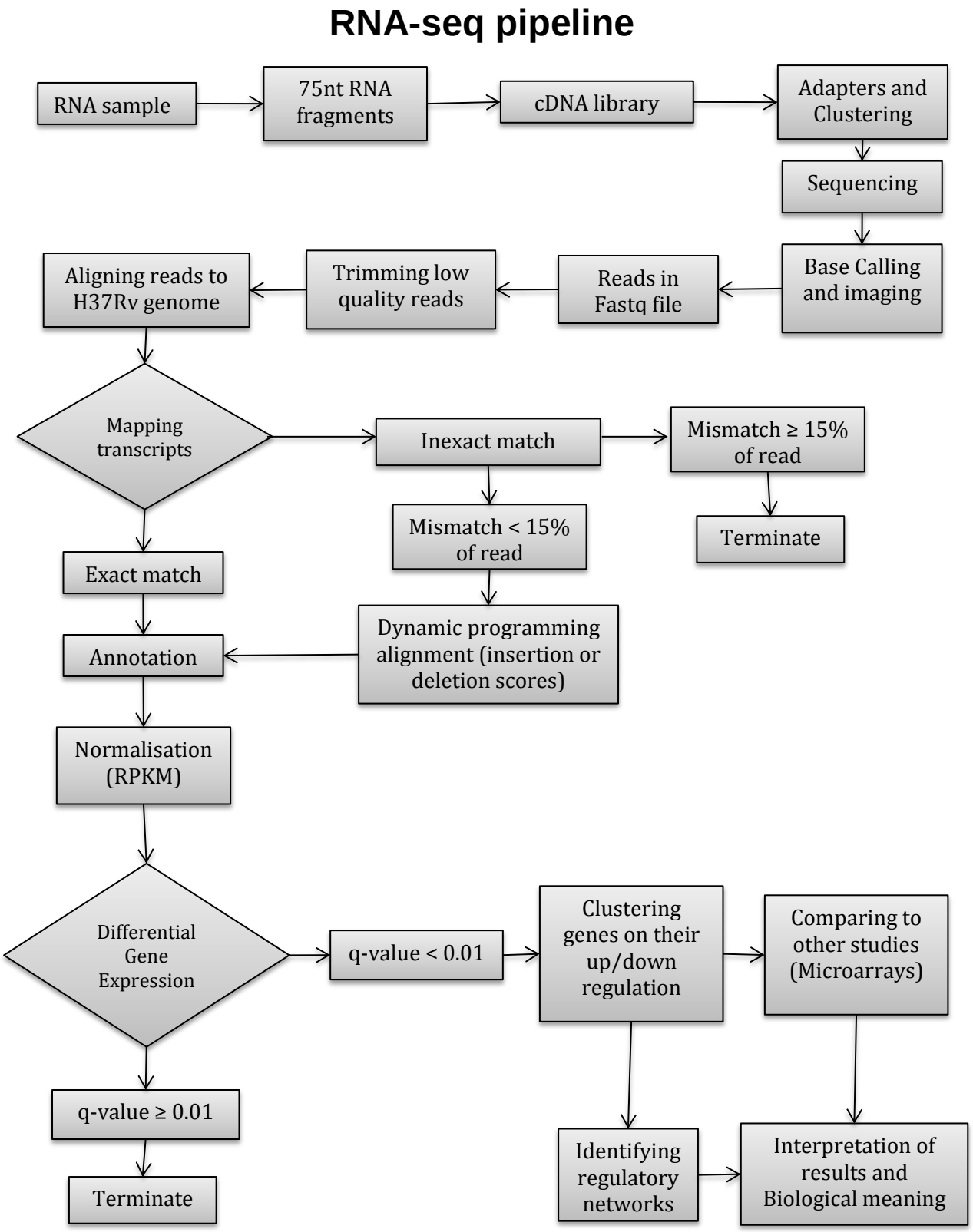
The pattern of Mtb transcriptional adaptation in response to the neutrophil encounter is, however, of general interest. A key part challenge was to prepare appropriate control samples in which the bacteria experience everything but the presence of neutrophils. Thus, the differential gene expression of H37Rv in the sample of 4h neutrophil encounter was quantified relative to H37RV in RPMI without neutrophils. The data from RNA-seq were analysed in Rockhopper software for detection of differentially expressed genes based on q value of less than 0.01 between two compared conditions. Out of 90 differentially expressed genes from neutrophil 4h sample, six were found to be in *sigK* regulon. Although, these genes have been classified as non-essential for growth of Mtb (Lew *et al.*, 2011), a number of studies recognise their immunogenicity (Chen *et al.*, 2012, Wiker, 2009). While they were found to be profoundly overexpressed during neutrophil encounter, some of them have been also found to be overexpressed in macrophage studies though at significantly lower levels under different experimental conditions (Rohde *et al.*, 2012, Schnappinger *et al.*, 2003, Said-Salim *et al.*, 2006).

The confirmation of results from RNA-seq data was done via qPCR which is a routine method of confirmation for high-throughput sequencing (Rienksma et al., 2015). The selected genes for qPCR confirmation fall into three categories based on differential expression relative to RPMI control. Thus, three up regulated (*dipZ*, *mpt83* and *mpt70*), two down regulated (Rv0692c and Rv0691a), and five genes with no change (*tgs1*, *hspX*, *esat-6*, *sigK*, and *icl1*). Overall, the RNA-seq results were confirmed by qPCR. TPM values were used because they seem to be slightly more similar to qPCR quantifications than RPKM.

Altogether this study allows a transcriptome-wide estimation of Mtb adaptation to neutrophil encounter. However, further approaches are required to characterise this process including comprehensive proteomic study. A proteome-wide study would help us to estimate the changes in protein concentration under different conditions. In this regard, in their study Schubert and colleagues found that the level of DosR regulon proteins increased to about 20% of total cell protein content during dormancy while ribosomal proteins remained largely at constant level compared to exponentially growing H37Rv (Schubert *et al.*, 2015). Thus, deploying transcriptome study combined with a complementary technique such as proteomics would provide further quantitative analysis and profoundly serve us in understanding the interaction between Mtb and neutrophil.

6.1. Future work

- More time points of incubation are required in order to precisely identify the transcriptional response of Mtb to neutrophil encounter
- More in vivo samples are required to achieve more representative sample for comparison
- The sigK genes profoundly induced by neutrophil interaction need to be characterised in order to understand the real role of neutrophil in the pathogenesis of Mtb.


RNA-seq pipeline
RNA sample
75nt RNA fragments
cDNA library
Adapters and Clustering
Sequencing
Base Calling and imaging
Reads in Fastq file
Trimming low quality reads
Aligning reads to H37Rv genome
Mapping transcripts
Inexact match
Mismatch ≥ 15% of read
Terminate
Mismatch < 15% of read
Exact match
Dynamic programming alignment (insertion or deletion scores)
Annotation
Normalisation (RPKM)
Differential Gene Expression
q-value < 0.01
Clustering genes on their up/down regulation
Comparing to other studies (Microarrays)
q-value ≥ 0.01
Terminate
Identifying regulatory networks
Interpretation of results and Biological meaning

Appendices

Preparation of 29 random-primed cDNA libraries for Illumina sequencing

Material supplied
Thirty-five RNA samples from *Mycobacterium* or *Mycobacterium* and human as described in Table1, delivered on dry ice.

Table 1: Description of RNA samples

No.	Sample	Description	Conc. (ng/µl)	Total amount (ng)	Conc. (ng/µl)	Total amount (ng)
			customer-specified		own measurement (see Fig.1)	
1	AB1	Mid Exp - 7H9	18	540	64	1.780
2	AB2	Mid Exp - 7H9	24	720	53	1.480
3	AB3	Mid Exp - 7H9	176	5.280	3	90
4	AB4	Stn - 7H9	16	480	119	3.320
5	AB5	Stn - 7H9	28	840	37	1.050
6	AB6	Stn - 7H9	141	4.230	30	870
7	AB7	Rv-RPMI 24hr T0	15	450	43	1.210
8	AB8	Rv-RPMI 24hr T0	36	1.080	39	1.100
9	AB9	Rv-RPMI 24hr T0	30	900	12	330
10	AB10	Rv-RPMI 28hr T4	10	300	26	740
11	AB11	Rv-RPMI 28hr T4	22	660	6	160
12	AB12	Rv-RPMI 28hr T4	26	780	42	1.180
13	AB13	Rv-PMN 4hr	31	930	50	1.400
14	AB14	Rv-PMN 4hr	10	300	31	860
15	AB15	Rv-PMN 4hr	11	330	17	490
16	AB16	Rv-RPMI 48hr T24	27	810	55	1.530
17	AB17	Rv-RPMI 48hr T24	104	3.120	75	2.110
18	AB18	Rv-RPMI 48hr T24	154	4.620	113	3.180
19	AB19	Rv-PMN 24hr	41	1.230	79	2.210
20	AB20	Rv-PMN 24hr	110	3.300	102	2.850
21	AB21	Rv-PMN 24hr	25	750	58	1.610
22	GS1	Patient E043 Aerosol collection 1 direct RNA	-	0	0	1
23	GS2	Patient E043 Aerosol collection 2 direct RNA	-	0	0	4
24	GS3	Patient E201 Aerosol collection 1 direct RNA	-	0	0	9
25	GS4	Patient E202 Aerosol collection 1 direct RNA	-	1	0	2
26	GS5	Patient E043 Aerosol collection 1 amplified RNA	-	9.000	9	550
27	GS6	Patient E043 Aerosol collection 2 amplified RNA	-	27.000	19	1.300
28	GS7	Patient E201 Aerosol collection 1 amplified RNA	-	30.000	0	9
29	GS8	Patient E202 Aerosol collection 1 amplified RNA	-	600	21	1.450
30	GS9	Patient E043 Sputum collection 1 direct RNA	-	13	12	330
31	GS10	Patient E202 Sputum collection 1 direct RNA	-	266	71	2.060
32	GS11	Patient E043 Sputum collection 1 amplified RNA	-	4.500	3	200
33	GS12	Patient E043 Sputum collection 2 amplified RNA	-	21.000	26	1.680
34	GS13	Patient E201 Sputum collection 1 amplified RNA	-	600	0	15
35	GS14	Patient E202 Sputum collection 1 amplified RNA	-	300.000	201	13.900

Appendices

2 Analysis of RNA samples

The RNA samples with enough material available were examined by capillary electrophoresis (Fig. 1).

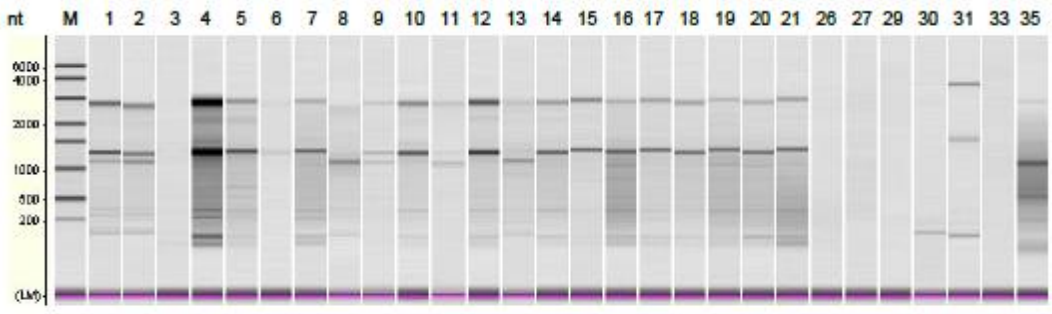


Figure 1: Analysis of RNA samples on a Shimadzu MultiNA microchip electrophoresis system. M = RNA marker

3 cDNA synthesis

RNA samples 1 to 21 and 26 to 34 were chosen for random-primed cDNA library preparation.

The RNA samples were fragmented with ultrasound (4 pulses of 30 sec at 4°C). First-strand cDNA synthesis was primed with a N6 randomised primer. Then, the Illumina TruSeq sequencing adapters were ligated to the 5' and 3' ends of the cDNA. The cDNA was finally amplified with PCR using a proof reading enzyme. Samples no. 28 and 34 (GS7 and GS13) could not be successfully amplified and were not used for sequencing. The number of cycles and the TruSeq barcode sequences which are part of the 3' TruSeq sequencing adapters are included in Table 2. The cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics) and was analyzed by capillary electrophoresis (Figure 2).

4 Description of cDNA samples

The primers used for PCR amplification were designed for TruSeq sequencing according to the instructions of Illumina.

The following adapter sequences flank the DNA insert:

TrueSeq_Sense_primer

5´- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'

TrueSeq_Antisense_ NNNNNN_primer Barcode

5'-CAAGCAGAAGACGGCATACGAGAT-NNNNNN-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT 3'

The total length of the flanking sequences is 122 bp.

The following table indicates the Illumina barcodes of the cDNA samples and numbers of PCR cycles used for cDNA amplification. The cDNA samples were analyzed on a Shimadzu MultiNA microchip electrophoresis system

Appendices

Table 2: Illumina barcodes and number of PCR cycles for cDNA samples

No.	Sample	Barcode	PCR cycles
1	AB1	ATCACG	11
2	AB2	CGATGT	11
3	AB3	TTAGGC	13
4	AB4	TGACCA	12
5	AB5	ACAGTG	12
6	AB6	GCCAAT	12
7	AB7	CAGATC	12
8	AB8	ACTTGA	11
9	AB9	GATCAG	12
10	AB10	TAGCTT	12
11	AB11	CTTGTA	13
12	AB12	AGTCAA	13
13	AB13	AGTTCC	12
14	AB14	ATGTCA	12
15	AB15	CCGTCC	13
16	AB16	GTAGAG	13
17	AB17	GTCCGC	13
18	AB18	GTGAAA	13
19	AB19	GTGGCC	13
20	AB20	GTTTCG	13
21	AB21	CGTACG	13
26	GS5	GAGTGG	21
27	GS6	CAAAAG	17
29	GS8	CAACTA	20
30	GS9	ACTGAT	16
31	GS10	ATGAGC	12
32	GS11	CACCGG	17

33	GS12	ATTCCT	22
35	GS14	GGCTAC	14

Commands to create scatterplots in R

```
# Scatterplot Matrices from the glus Package

library(gclus)

dta <- mtcars[c(1,3,5,6)]   # get data

dta.r <- abs(cor(dta))      # get correlations

dta.col <- dmat.color(dta.r)    # get colors

 # reorder variables so those with highest correlation are closest to the diagonal

 dta.o <- order.single(dta.r)

cpairs(dta, dta.o, panel.colors=dta.col, gap=.5, main="Variables Ordered and Colored by
Correlation" )
```

Appendices

List of PCR primers used in this study

Primer	Sequence	Tm °C
sigK_F	GTG TAT GGA CTG GTG ATG CG	59.4
sigK_R	ACC TTT GGC GGA GTC AAA CT	57.3
Rv0691c_F	CGA TGC CAT ATT GTG TCC AG	57.3
Rv0691c_R	GAC AAG CTC GGT GTC GGT AT	59.4
Rv0692_F	GTG TGG GGT TTA CTG ACC GT	59.4
Rv0692_R	AGG ATG GTG CGA TTT TTC AG	55.2
mpt83_F	ACA ACC CGA TGC TCA GTA CC	59.4
mpt83_R	TTG GCG TCA GTC TTG AGT TG	57.3
dipZ_F	ACA ACT ACG CCA CTT GGA CC	59.4
dipZ_R	GGT GAC GTT GTA ATC GCC TT	57.3
mpt70_F	ATC CGC AAG TAA ACC TGG TG	57.3
mpt70_R	TTG AGG CTG TTA CCC TGA CC	59.4
sigA-F	GAG ATC GGC CAG GTC TAC GGC GTG	65.1
sigA-R	CTG ACA TGG GGG CCC GCT ACG TTG	66.0
16s-F	GAA ACT GGG TCT AAT ACC	46.6
16s-R	ATC TCA GTC CCA GTG TGG	53.7
tgs1-F	AAC GAA GAC CAG TTA TTC GAG	55.0
tgs1-R	CTC ATA CTT TCA TCG GAG AGC	54.9
hspX-F	ACT CCG GCC CAC CTT CGA CA	63.6
hspX-R	AGC ACC TAC CGG CAG CGA CA	64.0
icl1-F	GCG GTG CGG AGG TGC TGT GG	66.4
icl-R	AGG CTC TGG TCG GGG TAG GTG	66.6
rpsL-F	CAGCTGGTCCGCAAGGGTCGTC	64.6
rpsL-R	GAGTTCGGCTTCTTCGGAGTGGTG	62.0

Participant Information Sheet

Neutrophil Infection

Principle Investigator: Amin Bakir PhD Student

Department of Infection, Immunity, and Inflammation, Medical Sciences Building, University of Leicester, LE1 9HN
Tel: 0116 252 2955 email: aab23@le.ac.uk

You are invited to take part in research study. Before you decide to participate, it is important that you understand why the research is being carried out and what it will involve.

What is the purpose of the study?

I am investigating the effect of neutrophils on the transcriptome of *Mycobacterium tuberculosis* (Mtb). Neutrophils are believed to have a significant contribution to the pathogenesis of Mtb. Human blood is required to obtain neutrophils for my experiments. The neutrophils obtained for our experiment are used for carefully controlled experiments in the lab and are then destroyed.

What will be involved if I take part in the study?

An experienced investigator who has been thoroughly trained by the NHS in blood taking will remove a 50ml sample of blood from your arm.

Will information obtained in the study be confidential?

All data will be protected from unauthorised use. The data obtained may be published as part of a scientific paper but the names of the blood donors will not be revealed.

What if I am hurt by the study?

Medical research is covered for mishaps in the same way as for patients undergoing treatment in the NHS, i.e. compensation is available if negligence occurs.

What happens if I do not wish to participate in this study or wish to withdraw from the study?

You may withdraw from this study at any time for any reason, without having to justify your decision in any way.

Participant Consent Form

Neutrophil Infection

Principle Investigator: Amin Bakir

Department of Infection, Immunity, and Inflammation, Medical Sciences Building, University of Leicester, LE1 9HN
Tel: 0116 223 2955 email: aab23@le.ac.uk

This form should be read in conjunction with the Participant Information sheet

I agree to take part in the above study, by providing blood as described in the Participant Information sheet.

I understand that I may withdraw from the study at any time for any reason without having to justify my decision, and that this would not affect any medical care I might receive.

I understand that if I have any reason to believe I am carrying a blood-borne infection such as Hepatitis B or C or HIV, I should not volunteer to give blood.

I understand medical research is covered for mishaps in the same way as for patients undergoing treatment in the NHS, i.e. compensation is only available if negligence occurs.

I have read the Participant Information sheet on the above study and have had the opportunity to discuss the details with the investigator detailed below, and to ask any questions. The nature and purpose of the samples to be taken have been explained to me and I understand what will be required if I take part in the study.

Signature of Participant / Donor ……Date……

(Name in BLOCKLETTERS)……

I confirm I have explained the nature of the study, as detailed in the Participant / Donor Information sheet, in terms which in my judgment are suited to the understanding of the subject.

Signature of investigator…….……..……Date……..

(Name in BLOCK LETTERS)……

Bibliography

Abadie, V., Badell, E., Douillard, P., Ensergueix, D., Leenen, P. J., Tanguy, M., Fiette, L., Saeland, S., Gicquel, B., Winter, N., 2005. Neutrophils rapidly migrate via lymphatics after Mycobacterium bovis BCG intradermal vaccination and shuttle live bacilli to the draining lymph nodes. *Blood,* **106**, 1843-1850.

Akhtar, S., Sarkar, S., Mishra, A., Sarkar, D., 2011. A method to extract intact and pure RNA from mycobacteria. *Anal Biochem,* **417**, 286-288.

Alimuddin, Z., Mario, R., Richard, H., von Reyn, C. F., 2013. Tuberculosis. *N Engl J Med,* **368**, 745-755.

Alli., Terry, O. A., Mangan, J. A., Butcher, P. D., Spreadbury, C. L., 2009. Optimization of RNA extraction in Mycobacterium tuberculosis for studying intracellular gene expression. *African Journal of Clinical and Experimental Microbiology,* **10**, 64 - 79.

Almeida, R. P., Melchior, M., Campanelli, D., Nathan, C., Gabay, J. E., 1991. Complementary DNA sequence of human neutrophil azurocidin, an antibiotic with extensive homology to serine proteases. *Biochem Biophys Res Commun,* **177**, 688-695.

Anders, S. & Huber, W., 2010. Differential expression analysis for sequence count data. *Genome Biol,* **11**, R106.

Appelberg, R., Pedrosa, J., Silva, M., 1991. Host and bacterial factors control the Mycobacterium avium-induced chronic peritoneal granulocytosis in mice. *Clinical & Experimental Immunology,* **83**, 231-236.

Athens, J., Haab, O., Raab, S., Mauer, A., Ashenbrucker, H., Cartwright, G., Wintrobe, M., 1961. Leukokinetic studies. IV. The total blood, circulating and marginal granulocyte pools and the granulocyte turnover rate in normal subjects. *Journal of Clinical Investigation,* **40**, 989.

Babior, B. M., 1999. NADPH oxidase: an update. *Blood,* **93**, 1464-1476.

Bange, F. C., Brown, A. M., Jacobs, W. R., 1996. Leucine Auxotrophy Restricts Growth of *Mycobacterium bovis* BCG in Macrophages. *Infection and immunity,* **64**, 1794-1799.

Barer, M. R., Gribbon, L. T., Whiteley, A. S., Smith, R., Newton, A., O' Donnell, A. G., 1999. Relationships between cellular activity and culturability. *Microbial Biosystems: New Frontiers*.

Bartek, I. L., Rutherford, R., Gruppo, V., Morton, R. A., Morris, R. P., Klein, M. R., Visconti, K. C., Ryan, G. J., Schoolnik, G. K., Lenaerts, A., Voskuil, M. I., 2009. The DosR regulon of M. tuberculosis and antibacterial tolerance. *Tuberculosis (Edinb),* **89**, 310-316.

Bashyam, M. D., Kaushal, D., Dasgupta, S. K., Tyagi, A. K., 1996. A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements. *J Bacteriol,* **178**, 4847-4853.

Beckman, J. S. & Koppenol, W. H., 1996. Nitric oxide, superoxide, and peroxynitrite: the good, the bad, and ugly. *American Journal of Physiology-Cell Physiology,* **271**, C1424-C1437.

Beiter, K., Wartha, F., Albiger, B., Normark, S., Zychlinsky, A., Henriques-Normark, B., 2006. An endonuclease allows Streptococcus pneumoniae to escape from neutrophil extracellular traps. *Current Biology,* **16**, 401-407.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

Benjamini, Y. & Speed, T. 2011. *Estimation and correction for GC-content bias in high throughput sequencing.* Tech Rep Berkeley, CA USA.

Bermudez, L. E. & Goodman, J., 1996. Mycobacterium tuberculosis invades and replicates within type II alveolar cells. *Infect Immun,* **64**, 1400-1406.

Berry, M. P., Graham, C. M., McNab, F. W., Xu, Z., Bloch, S. A., Oni, T., Wilkinson, K. A., Banchereau, R., Skinner, J., Wilkinson, R. J., 2010. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature,* **466**, 973-977.

Berton, G., 1999. Tyrosine kinases in neutrophils. *Current opinion in hematology,* **6**, 51.

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A., Dun, K., 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Molecular Microbiology,* **43**, 717-731.

Bianchi, M., Niemiec, M. J., Siler, U., Urban, C. F., Reichenbach, J., 2011. Restoration of anti-Aspergillus defense by neutrophil extracellular traps in human chronic granulomatous disease after gene therapy is calprotectin-dependent. *Journal of Allergy and Clinical Immunology,* **127**, 1243-1252. e1247.

Bjerregaard, M. D., Jurlander, J., Klausen, P., Borregaard, N., Cowland, J. B., 2003. The in vivo profile of transcription factors during neutrophil differentiation in human bone marrow. *Blood,* **101**, 4322-4332.

Bolger, A. M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.

Bolstad, B., Collin, F., Simpson, K., Irizarry, R., Speed, T., 2004. Experimental design and low-level analysis of microarray data. *International review of neurobiology,* **60**, 25-58.

Borregaard, N., 2010. Neutrophils, from marrow to microbes. *Immunity,* **33**, 657-670.

Brosch, R., Pym, A. S., Gordon, S. V., Cole, S. T., 2001. The evolution of mycobacterial pathogenicity: clues from comparative genomics. *Trends Microbiol,* **9**, 452-458.

Buermans, H. & Den Dunnen, J., 2014. Next generation sequencing technology: advances and applications. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* **1842**, 1932-1941.

Bukka, A., Price, C. T., Kernodle, D. S., Graham, J. E., 2011. Mycobacterium tuberculosis RNA expression patterns in sputum bacteria indicate secreted Esx factors contributing to growth are highly expressed in active disease. *Front Microbiol,* **2**.

Bullard, J. H., Purdom, E., Hansen, K. D., Dudoit, S., 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics,* **11**, 94.

Butler, R. E., Cihlarova, V., Stewart, G. R., 2010. Effective generation of reactive oxygen species in the mycobacterial phagosome requires K+ efflux from the bacterium. *Cellular microbiology,* **12**, 1186-1193.

Cacho, A., Smirnova, E., Huzurbazar, S., Cui, X., 2015. A Comparison of Base-calling Algorithms for Illumina Sequencing Technology. *Brief Bioinform*, bbv088.

Carr, M. D., Bloemink, M. J., Dentten, E., Whelan, A. O., Gordon, S. V., Kelly, G., Frenkiel, T. A., Hewinson, R. G., Williamson, R. A., 2003. Solution Structure

of the Mycobacterium tuberculosis Complex Protein MPB70 FROM TUBERCULOSIS PATHOGENESIS TO INHERITED HUMAN CORNEAL DISEASE. *Journal of Biological Chemistry,* **278**, 43736-43743.

Carroll, M. V., Lack, N., Sim, E., Krarup, A., Sim, R. B., 2009. Multiple routes of complement activation by Mycobacterium bovis BCG. *Molecular immunology,* **46**, 3367-3378.

Chao, M. C. & Rubin, E. J., 2010. Letting sleeping dos lie: does dormancy play a role in tuberculosis? *Annu Rev Microbiol,* **64**, 293-311.

Charlet, D., Mostowy, S., Alexander, D., Sit, L., Wiker, H. G., Behr, M. A., 2005. Reduced expression of antigenic proteins MPB70 and MPB83 in Mycobacterium bovis BCG strains due to a start codon mutation in sigK. *Mol Microbiol,* **56**, 1302-1313.

Chen, K., Nishi, H., Travers, R., Tsuboi, N., Martinod, K., Wagner, D. D., Stan, R., Croce, K., Mayadas, T. N., 2012. Endocytosis of soluble immune complexes leads to their clearance by FcγRIIIB but induces neutrophil extracellular traps via FcγRIIA in vivo. *Blood,* **120**, 4421-4431.

Chen, Y., Lun, A. T., Smyth, G. K., 2014. Differential Expression Analysis of Complex RNA-seq Experiments Using edgeR. *Statistical Analysis of Next Generation Sequencing Data.* Springer 51-74.

Chen, Z. & Duan, X., 2011. Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. *High-Throughput Next Generation Sequencing.* Springer 93-103.

Chuang, F. Y., Sassaroli, M., Unkeless, J. C., 2000. Convergence of Fcγ receptor IIA and Fcγ receptor IIIB signaling pathways in human neutrophils. *The Journal of Immunology,* **164**, 350-360.

Chumakov, A. M., Silla, A., Williamson, E. A., Koeffler, H. P., 2007. Modulation of DNA binding properties of CCAAT/enhancer binding protein epsilon by heterodimer formation and interactions with NFkappaB pathway. *Blood,* **109**, 4209-4219.

Cieutat, A.-M., Lobel, P., August, J., Kjeldsen, L., Sengeløv, H., Borregaard, N., Bainton, D., 1998. Azurophilic granules of human neutrophilic leukocytes are deficient in lysosome-associated membrane proteins but retain the mannose 6-phosphate recognition marker. *Blood,* **91**, 1044-1058.

Clark, S. R., Ma, A. C., Tavener, S. A., McDonald, B., Goodarzi, Z., Kelly, M. M., Patel, K. D., Chakrabarti, S., McAvoy, E., Sinclair, G. D., 2007. Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood. *Nat Med,* **13**, 463-469.

Cole, S., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S., Eiglmeier, K., Gas, S., Barry, C. r., 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature,* **393**, 537-544.

Collins, S. J., 1987. The HL-60 promyelocytic leukemia cell line: proliferation, differentiation, and cellular oncogene expression. *Blood,* **70**, 1233-1244.

Converse, P. J., Karakousis, P. C., Klinkenberg, L. G., Kesavan, A. K., Ly, L. H., Allen, S. S., Grosset, J. H., Jain, S. K., Lamichhane, G., Manabe, Y. C., 2009. Role of the dosR-dosS two-component regulatory system in Mycobacterium tuberculosis virulence in three animal models. *Infection and immunity,* **77**, 1230-1237.

Cooper, A. M. & Torrado, E., 2012. Protection versus pathology in tuberculosis: recent insights. *Curr Opin Immunol,* **24**, 431-437.

Corleis, B., Korbel, D., Wilson, R., Bylund, J., Chee, R., Schaible, U. E., 2012. Escape of Mycobacterium tuberculosis from oxidative killing by neutrophils. *Cell Microbiol,* **14**, 1109-1121.

Cortes, T., Schubert, O. T., Rose, G., Arnvig, K. B., Comas, I., Aebersold, R., Young, D. B., 2013. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in Mycobacterium tuberculosis. *Cell Rep,* **5**, 1121-1131.

Costa, D., Marques, A. P., Reis, R. L., Lima, J. L., Fernandes, E., 2006. Inhibition of human neutrophil oxidative burst by pyrazolone derivatives. *Free Radical Biology and Medicine,* **40**, 632-640.

Croucher, N. J. & Thomson, N. R., 2010. Studying bacterial transcriptomes using RNA-seq. *Curr Opin Microbiol,* **13**, 619-624.

Dabney, J. & Meyer, M., 2012. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *Biotechniques,* **52**, 87-94.

Dale, D. C., Boxer, L., Liles, W. C., 2008. The phagocytes: neutrophils and monocytes. *Blood,* **112**, 935-945.

Daniel, J., Deb, C., Dubey, V. S., Sirakova, T. D., Abomoelak, B., Morbidoni, H. R., Kolattukudy, P. E., 2004. Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in Mycobacterium tuberculosis as it goes into a dormancy-like state in culture. *J Bacteriol,* **186**, 5017-5030.

Del Fabbro, C., Scalabrin, S., Morgante, M., Giorgi, F. M., 2013. An extensive evaluation of read trimming effects on Illumina NGS data analysis.

Denis, M., 1991. Human Neutrophils, Activated with Cytokines or Not, Do Not Kill Virulent *Mycobacterium tuberculosis*. *Journal of Infectious Diseases,* **163**, 919-920.

Dieci, G., Giuliodori, S., Catellani, M., Percudani, R., Ottonello, S., 2002. Intragenic Promoter Adaptation and Facilitated RNA Polymerase III Recycling in the Transcription of SCR1, the 7SL RNA Gene ofSaccharomyces cerevisiae. *Journal of Biological Chemistry,* **277**, 6903-6914.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., 2013. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in bioinformatics,* **14**, 671-683.

Dorhoi, A., Reece, S. T., Kaufmann, S. H. E., 2011. For better or worse: the immune response against *Mycobacterium tuberculosis* balances pathology and protection. *Immunological Reviews,* **240**, 235-251.

Doz, E., Rose, S., Vasseur, V., Charron, S., Gilleron, M., Puzo, G., Fremaux, I., Delneste, Y., Erard, F., Ryffel, B., 2009. Mycobacterial phosphatidylinositol mannosides negatively regulate host Toll-like receptor 4, MyD88-dependent proinflammatory cytokines, and TRIF-dependent co-stimulatory molecule expression. *Journal of Biological Chemistry,* **284**, 23187-23196.

Dubnau, E., Fontan, P., Manganelli, R., 2002. Mycobacterium tuberculosis Genes Induced during Infection of Human Macrophages. *Infection and immunity,* **70**, 2787-2795.

Eash, K. J., Means, J. M., White, D. W., Link, D. C., 2009. CXCR4 is a key regulator of neutrophil release from the bone marrow under basal and stress granulopoiesis conditions. *Blood,* **113**, 4711-4719.

Elkington, P. T., Ugarte-Gil, C. A., Friedland, J. S., 2011. Matrix metalloproteinases in tuberculosis. *European Respiratory Journal,* **38**, 456-464.

Elsbach, P., 1998. The bactericidal/permeability-increasing protein (BPI) in antibacterial host defense. *Journal of leukocyte biology,* **64**, 14-18.

Eruslanov, E. B., Lyadova, I. V., Kondratieva, T. K., Majorov, K. B., Scheglov, I. V., Orlova, M. O., Apt, A. S., 2005. Neutrophil responses to Mycobacterium tuberculosis infection in genetically susceptible and resistant mice. *Infect Immun,* **73**, 1744-1753.

Eum, S.-Y., Kong, J.-H., Hong, M.-S., Lee, Y.-J., Kim, J.-H., Hwang, S.-H., Cho, S.-N., Via, L. E., Barry, C. E., 2010. Neutrophils are the predominant infected phagocytic cells in the airways of patients with active pulmonary TB. *CHEST Journal,* **137**, 122-128.

Evguenieva-Hackenberg, E. & Klug, G., 2011. New aspects of RNA processing in prokaryotes. *Curr Opin Microbiol,* **14**, 587-592.

Ewing, B. & Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res,* **8**, 186-194.

Fennelly, K. P., Martyny, J. W., Fulton, K. E., Orme, I. M., Cave, D. M., Heifets, L. B., 2004. Cough-generated aerosols of Mycobacterium tuberculosis: a new method to study infectiousness. *Am J Respir Crit Care Med,* **169**, 604-609.

Ferguson, J. S., Martin, J. L., Azad, A. K., McCarthy, T. R., Kang, P. B., Voelker, D. R., Crouch, E. C., Schlesinger, L. S., 2006. Surfactant protein D increases fusion of Mycobacterium tuberculosis-containing phagosomes with lysosomes in human macrophages. *Infect Immun,* **74**, 7005-7009.

Filiatrault, M. J., 2011. Progress in prokaryotic transcriptomics. *Curr Opin Microbiol,* **14**, 579-586.

Fleck, R. A., Romero-Steiner, S., Nahm, M. H., 2005. Use of HL-60 cell line to measure opsonic capacity of pneumococcal antibodies. *Clin Diagn Lab Immunol,* **12**, 19-27.

Fleige, S. & Pfaffl, M. W., 2006. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med,* **27**, 126-139.

Fontan, P., Aris, V., Ghanny, S., Soteropoulos, P., Smith, I., 2008. Global transcriptional profile of Mycobacterium tuberculosis during THP-1 human macrophage infection. *Infect Immun,* **76**, 717-725.

Fortune, S. M., Solache, A., Jaeger, A., Hill, P. J., Belisle, J. T., Bloom, B. R., Rubin, E. J., Ernst, J. D., 2004. Mycobacterium tuberculosis Inhibits Macrophage Responses to IFN- through Myeloid Differentiation Factor 88-Dependent and -Independent Mechanisms. *The Journal of Immunology,* **172**, 6272-6280.

Fouret, P., Du Bois, R., Bernaudin, J., Takahashi, H., Ferrans, V., Crystal, R., 1989. Expression of the neutrophil elastase gene during human bone marrow cell differentiation. *J Exp Med,* **169**, 833-845.

Francis, R., Butler, R., Stewart, G., 2014. Mycobacterium tuberculosis ESAT-6 is a leukocidin causing Ca2+ influx, necrosis and neutrophil extracellular trap formation. *Cell Death Dis,* **5**, e1474.

Fuchs, T. A., Abed, U., Goosmann, C., Hurwitz, R., Schulze, I., Wahn, V., Weinrauch, Y., Brinkmann, V., Zychlinsky, A., 2007. Novel cell death program leads to neutrophil extracellular traps. *J Cell Biol,* **176**, 231-241.

Garton, N. J., Waddell, S. J., Sherratt, A. L., Lee, S., Smith, R. J., Senner, C., Hinds, J., Rajakumar, K., Adegbola, R. A., Besra, G. S., Butcher, P. D., Barer, M. R., 2008. Cytological and Transcript Analysis Reveal Fat and Lazy Persister-Like Bacilli in Tuberculosis Sputum. *PLoS Med,* **5**, e75.

Gaynor, C. D., Mccormack, F. X., Voelker, D. R., McGowan, S. E., Schlesinger, L. S., 1995. Pulmonary surfactant protein A mediates enhanced phagocytosis of Mycobacterium tuberculosis by a direct interaction with human macrophages. *The Journal of Immunology,* **155**, 5343-5351.

Gengenbacher, M. & Kaufmann, S. H., 2012. Mycobacterium tuberculosis: success through dormancy. *FEMS Microbiol Rev,* **36**, 514-532.

Gery, S., Gombart, A. F., Fung, Y. K., Koeffler, H. P., 2004. C/EBPϵ interacts with retinoblastoma and E2F1 during granulopoiesis. *Blood,* **103**, 828-835.

Giannoni, E., Sawa, T., Allen, L., Wiener-Kronish, J., Hawgood, S., 2006. Surfactant proteins A and D enhance pulmonary clearance of Pseudomonas aeruginosa. *Am J Respir Cell Mol Biol,* **34**, 704-710.

Gideon, H. P. & Flynn, J. L., 2011. Latent tuberculosis: what the host "sees"? *Immunol Res,* **50**, 202-212.

Gopaul, K. K., Brooks, P. C., Prost, J.-F., Davis, E. O., 2003. Characterization of the two Mycobacterium tuberculosis recA promoters. *J Bacteriol,* **185**, 6005-6015.

Gupta, A., Kaul, A., Tsolaki, A. G., Kishore, U., Bhakta, S., 2012. Mycobacterium tuberculosis: immune evasion, latency and reactivation. *Immunobiology,* **217**, 363-374.

Haas, B. J., Chin, M., Nusbaum, C., Birren, B. W., Livny, J., 2012. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics,* **13**, 734.

2013. *Seminars in immunopathology.* Springer.

Harwig, S., Tan, L., Qu, X., Cho, Y., Eisenhauer, P., Lehrer, R., 1995. Bactericidal properties of murine intestinal phospholipase A2. *Journal of Clinical Investigation,* **95**, 603.

Hauert, A. B., Martinelli, S., Marone, C., Niggli, V., 2002. Differentiated HL-60 cells are a valid model system for the analysis of human neutrophil migration and chemotaxis. *The International Journal of Biochemistry and Cell Biology,* **34**, 838-854.

Haugen, S. P., Ross, W., Gourse, R. L., 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature Reviews Microbiology,* **6**, 507-519.

He, S., Wurtzel, O., Singh, K., Froula, J. L., Yilmaz, S., Tringe, S. G., Wang, Z., Chen, F., Lindquist, E. A., Sorek, R., Hugenholtz, P., 2010. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat Meth,* **7**, 807-812.

Hedlund, S., Persson, A., Vujic, A., Che, K. F., Stendahl, O., Larsson, M., 2010. Dendritic cell activation by sensing Mycobacterium tuberculosis-induced apoptotic neutrophils via DC-SIGN. *Hum Immunol,* **71**, 535-540.

Heinecke, J. W., 1999. Is lipid peroxidation relevant to atherogenesis? *Journal of Clinical Investigation,* **104**, 135.

Hernandez, P. A., Gorlin, R. J., Lukens, J. N., Taniuchi, S., Bohinjec, J., Francois, F., Klotman, M. E., Diaz, G. A., 2003. Mutations in the chemokine receptor gene CXCR4 are associated with WHIM syndrome, a combined immunodeficiency disease. *Nat Genet,* **34**, 70-74.

Hertog, A. L. d., Klatser, P. R., Anthony, R. M., 2009. Buoyant density of Mycobacterium tuberculosis: implications for sputum processing. *The International Journal of Tuberculosis and Lung Disease,* **13**, 466 - 471.

Hewinson, R., Michell, S. L., Russell, W., McAdam, R., Jacobs Jr, W., 1996. Molecular characterization of MPT83: a seroreactive antigen of Mycobacterium tuberculosis with homology to MPT70. *Scandinavian journal of immunology,* **43**, 490-499.

Hilda, J. N., Selvaraj, A., Das, S. D., 2012. Mycobacterium tuberculosis H37Rv is more effective compared to vaccine strains in modulating neutrophil functions: an in vitro study. *FEMS Immunol Med Microbiol,* **66**, 372-381.

Hirai, H. Z., Dayaram, P., Hetherington, T., Mizuno, C., S. et al. 2006 C/EBPbeta is required for'emergency'granulopoiesis. *Nat. Immunol,* **7**, 732-739.

Hogg, J. C. & Doerschuk, C. M., 1995. Leukocyte traffic in the lung. *Annual review of physiology,* **57**, 97-114.

Honaker, R. W., Leistikow, R. L., Bartek, I. L., Voskuil, M. I., 2009. Unique roles of DosT and DosS in DosR regulon induction and Mycobacterium tuberculosis dormancy. *Infection and immunity,* **77**, 3258-3263.

Hübel, K., Dale, D. C., Liles, W. C., 2002. Therapeutic use of cytokines to modulate phagocyte function for the treatment of infectious diseases: current status of granulocyte colony-stimulating factor, granulocyte-macrophage colony-stimulating factor, macrophage colony-stimulating factor, and interferon-γ. *Journal of Infectious Diseases,* **185**, 1490-1501.

Hunt, D. M., Sweeney, N. P., Mori, L., Whalan, R. H., Comas, I., Norman, L., Cortes, T., Arnvig, K. B., Davis, E. O., Stapleton, M. R., 2012. Long-range

transcriptional control of an operon necessary for virulence-critical ESX-1 secretion in Mycobacterium tuberculosis. *J Bacteriol,* **194**, 2307-2320.

Husson, R. N., 2006. Leaving on the lights: host-specific derepression of Mycobacterium tuberculosis gene expression by anti-sigma factor gene mutations. *Mol Microbiol,* **62**, 1217-1219.

Hwang, P. M., Zhou, N., Shan, X., Arrowsmith, C. H., Vogel, H. J., 1998. Three-dimensional solution structure of lactoferricin B, an antimicrobial peptide derived from bovine lactoferrin. *Biochemistry,* **37**, 4288-4298.

Jain, V., Sujatha, S., Ojha, A. K., Chatterji, D., 2005. Identification and characterization of rel promoter element of Mycobacterium tuberculosis. *Gene,* **351**, 149-157.

Jena, P., Mohanty, S., Mohanty, T., Kallert, S., Morgelin, M., Lindstrøm, T., Borregaard, N., Stenger, S., Sonawane, A., Sørensen, O. E., 2012. Azurophil granule proteins constitute the major mycobactericidal proteins in human neutrophils and enhance the killing of mycobacteria in macrophages.

Jin, F., Nathan, C. F., Radzioch, D., Ding, A., 1998. Lipopolysaccharide-related stimuli induce expression of the secretory leukocyte protease inhibitor, a macrophage-derived lipopolysaccharide inhibitor. *Infect Immun,* **66**, 2447-2452.

Jordao, L., Bleck, C. K., Mayorga, L., Griffiths, G., Anes, E., 2008. On the killing of mycobacteria by macrophages. *Cellular microbiology,* **10**, 529-548.

Juárez, M. a. D., Torres, A., Espitia, C., 2001. Characterization of the Mycobacterium tuberculosis region containing the mpt83 and mpt70 genes. *FEMS Microbiology Letters,* **203**, 95-102.

Kaberdin, V. R., Singh, D., Lin-Chao, S., 2011. Composition and conservation of the mRNA-degrading machinery in bacteria. *J Biomed Sci,* **18**, 23.

Kairo, S. K., Bedwell, J., Tyler, P. C., Carter, A., Corbel, M. J., 1999. Development of a tetrazolium salt assay for rapid determination of viability of BCG vaccines. *Vaccine,* **17**, 24232428.

Kawakami, T., He, J., Morita, H., Yokoyama, K., Kaji, H., Tanaka, C., Suemori, S.-i., Tohyama, K., Tohyama, Y., 2014. Rab27a Is essential for the formation of neutrophil extracellular traps (NETs) in neutrophil-like differentiated HL60 cells. *PLoS One,* **9**, e84704.

Kemp, T. J., Ludwig, A. T., Earel, J. K., Moore, J. M., Vanoosten, R. L., Moses, B., Leidal, K., Nauseef, W. M., Griffith, T. S., 2005. Neutrophil stimulation with Mycobacterium bovis bacillus Calmette-Guerin (BCG) results in the release of functional soluble TRAIL/Apo-2L. *Blood,* **106**, 3474-3482.

Kenney, T. J. & Churchward, G., 1996. Genetic analysis of the Mycobacterium smegmatis rpsL promoter. *J Bacteriol,* **178**, 3564-3571.

Klebanoff, S. J., 1998. Myeloperoxidase. *Proceedings of the Association of American Physicians,* **111**, 383-389.

Kohl, T. A., Diel, R., Harmsen, D., Rothgänger, J., Walter, K. M., Merker, M., Weniger, T., Niemann, S., 2014. Whole-genome-based Mycobacterium tuberculosis surveillance: a standardized, portable, and expandable approach. *J Clin Microbiol,* **52**, 2479-2486.

Korbel, D. S., Schneider, B. E., Schaible, U. E., 2008. Innate immunity in tuberculosis: myths and truth. *Microbes and Infection,* **10**, 995-1004.

Kovacs, L., Csandadi, A., Megyeri, K., Kaberdin, V. R., Miczak, A., 2005. Mycobacterial RNase E-associated proteins. *Microbiol. Immunol.,* **49**, 1003 - 1007.

Kremer, L., Baulard, A., Estaquier, J., Capron, A., Locht, C., 1995. Analysis of the Mycobacterium tuberculosis 85A antigen promoter region. *J Bacteriol,* **177**, 642-653.

Lamhamedi, S., De Chastellier, C., Casanova, J. L., 1999. Growth of *Mycobacterium bovis*, Bacille Calmette-Guerin, within human monocytes-macrophages cultured in serum-free medium. *Journal of Immunological Methods,* **225**, 75-86.

Law, K. F., Jagirdar, J., Weiden, M. D., Bodkin, M., Rom, W. N., 1996. Tuberculosis in HIV-positive patients: cellular response and immune

activation in the lung. *American journal of respiratory and critical care medicine,* **153**, 1377-1384.

Ledergerber, C. & Dessimoz, C., 2011. Base-calling for next-generation sequencing platforms. *Brief Bioinform*, bbq077.

Leistikow, R. L., Morton, R. A., Bartek, I. L., Frimpong, I., Wagner, K., Voskuil, M. I., 2010. The Mycobacterium tuberculosis DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J Bacteriol,* **192**, 1662-1670.

Lew, J. M., Kapopoulou, A., Jones, L. M., Cole, S. T., 2011. TubercuList–10 years after. *Tuberculosis (Edinb),* **91**, 1-7.

Li, B. & Dewey, C. N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics,* **12**, 323.

Licatalosi, D. D. & Darnell, R. B., 2010. RNA processing and its regulation: global insights into biological networks. *Nat Rev Genet,* **11**, 75-87.

Lieschke, G., Grail, D., Hodgson, G., Metcalf, D., Stanley, E., Cheers, C., Fowler, K., Basu, S., Zhan, Y. F., Dunn, A., 1994. Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. *Blood,* **84**, 1737-1746.

Lin, P. L. & Flynn, J. L., 2010. Understanding latent tuberculosis: a moving target. *The Journal of Immunology,* **185**, 15-22.

Liu, P. T. & Modlin, R. L., 2008. Human macrophage host defense against Mycobacterium tuberculosis. *Current Opinion in Immunology,* **20**, 371-376.

Love, M. I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* **15**, 550.

Lowe, D. M., Bangani, N., Mehta, M. R., Lang, D. M., Rossi, A. G., Wilkinson, K. A., Wilkinson, R. J., Martineau, A. R., 2013. A novel assay of antimycobacterial activity and phagocytosis by human neutrophils. *Tuberculosis (Edinb),* **93**, 167-178.

Mackie, G. A., 2013. RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol,* **11**, 45-57.

Mahillon, J. & Chandler, M., 1998. Insertion sequences. *Microbiology and Molecular Biology Reviews,* **62**, 725-774.

Malkhed, V., Gudlur, B., Kondagari, B., Dulapalli, R., Vuruputuri, U., 2011. Study of interactions between Mycobacterium tuberculosis proteins: SigK and anti-SigK. *J Mol Model,* **17**, 1109-1119.

Mangan, J. A., Sole, K. M., Mitchison, D. A., Butcher, P. D., 1997. An effective method of RNA extraction from bacteria refractory to disruption, including mycobacteria. *Nucleic Acids Res,* **25**, 675 - 676.

Manganelli, R., 2014. Sigma Factors: Key Molecules in Mycobacterium tuberculosis Physiology and Virulence.

Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F. R., Smith, I., 1999. Differential expression of 10 sigma factor genes in Mycobacterium tuberculosis. *Molecular Microbiology,* **31**, 715-724.

Martin, J. A. & Wang, Z., 2011. Next-generation transcriptome assembly. *Nature Reviews Genetics,* **12**, 671-682.

Mason, P. E., Neilson, G. W., Dempsey, C. E., Barnes, A. C., Cruickshank, J. M., 2003. The hydration structure of guanidinium and thiocyanate ions: implications for protein stability in aqueous solution. *Proc Natl Acad Sci U S A,* **100**, 4557-4561.

Massol, P., Montcourrier, P., Guillemot, J. C., Chavrier, P., 1998. Fc receptor-mediated phagocytosis requires CDC42 and Rac1. *The EMBO Journal,* **17**, 6219-6229.

Matsumoto, S., Matsuo, T., Ohara, N., Hotokezaka, H., Naito, M., Minami, J., Yamada, T., 1995. Cloning and Sequencing of a Unique Antigen MPT70 from Mycobacterium tuberculosis H37Rv and Expression in BCG Using E. coli-Mycobacteria Shuttle Vector. *Scand J Immunol,* **41**, 281-287.

Mauer, A., Athens, J., Ashenbrucker, H., Cartwright, G., Wintrobe, M., 1960. Leukokinetic studies. II. A method for labeling granulocytes in vitro with radioactive diisopropylfluorophosphate (DFP32). *Journal of Clinical Investigation,* **39**, 1481.

Mayadas, T. N., Cullere, X., Lowell, C. A., 2014. The multifaceted functions of neutrophils. *Annu Rev Pathol,* **9**, 181.

Mayne, M., Shepel, P. N., Geiger, J. D., 1999. Recovery of high-integrity mRNA from brains of rats killed by high-energy focused microwave irradiation. *Brain Research Protocols,* **4**, 295-302.

McCarter, Y. S. & Robinson, A., 1996. Quality evaluation of sputum specimens for mycobacterial culture. *American journal of clinical pathology,* **105**, 769-773.

McClure, R., Balasubramanian, D., Sun, Y., Bobrovskyy, M., Sumby, P., Genco, C. A., Vanderpool, C. K., Tjaden, B., 2013. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res,* **41**, e140-e140.

McKinney, J. D., zu Bentrup, K. H., Muñoz-Elías, E. J., Miczak, A., Chen, B., Chan, W.-T., Swenson, D., Sacchettini, J. C., Jacobs, W. R., Russell, D. G., 2000. Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature,* **406**, 735-738.

Meena, L. S. & Kolattukudy, P. E., 2013. Expression and characterization of Rv0447c product, potentially the methyltransferase involved in

tuberculostearic acid biosynthesis in Mycobacterium tuberculosis. *Biotechnology and applied biochemistry,* **60**, 412-416.

Mesa, M. A. & Vasquez, G., 2013. NETosis. *Autoimmune Dis,* **2013**, 651497.

Middelhoven, P., Ager, A., Roos, D., Verhoeven, A., 1997. Involvement of a metalloprotease in the shedding of human neutrophil FcγRIIIB. *FEBS letters,* **414**, 14-18.

Mitchison, D. A., 2004. The search for new sterilizing anti-tuberculosis drugs. *Frontiers in bioscience: a journal and virtual library,* **9**, 1059-1072.

Mitra, A., Angamuthu, K., Nagaraja, V., 2008. Genome-wide analysis of the intrinsic terminators of transcription across the genus Mycobacterium. *Tuberculosis (Edinb),* **88**, 566-575.

Morel, C., Badell, E., Abadie, V., Robledo, M., Setterblad, N., Gluckman, J. C., Gicquel, B., Boudaly, S., Winter, N., 2008. Mycobacterium bovis BCG-infected neutrophils and dendritic cells cooperate to induce specific T cell responses in humans and mice. *Eur J Immunol,* **38**, 437-447.

Morosetti, R., Park, D. J., Chumakov, A. M., Grillier, I., Shiohara, M., Gombart, A. F., Nakamaki, T., Weinberg, K., Koeffler, H. P., 1997. A novel, myeloid transcription factor, C/EBPε, is upregulated during granulocytic, but not monocytic, differentiation. *Blood,* **90**, 2591-2600.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods,* **5**, 621-628.

Murphy, C. J., Foster, B. A., Mannis, M. J., Selsted, M. E., Reid, T. W., 1993. Defensins are mitogenic for epithelial cells and fibroblasts. *Journal of cellular physiology,* **155**, 408-413.

Nascimento, I. P. & Leite, L. C., 2005. The effect of passaging in liquid media and storage on Mycobacterium bovis--BCG growth capacity and infectivity. *FEMS Microbiol Lett,* **243**, 81-86.

Nathan, C. F., 1987. Neutrophil activation on biological surfaces. Massive secretion of hydrogen peroxide in response to products of macrophages and lymphocytes. *Journal of Clinical Investigation,* **80**, 1550.

Nerlov, C. & Graf, T., 1998. PU. 1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes & development,* **12**, 2403-2412.

Neufert, C., Pai, R. K., Noss, E. H., Berger, M., Boom, W. H., Harding, C. V., 2001. Mycobacterium tuberculosis 19-kDa lipoprotein promotes neutrophil activation. *The Journal of Immunology,* **167**, 1542-1549.

Newton-Foot, M. & van Pittius, N. C. G., 2013. The complex architecture of mycobacterial promoters. *Tuberculosis (Edinb),* **93**, 60-74.

Nordenfelt, P., Bauer, S., Lonnbro, P., Tapper, H., 2009. Phagocytosis of Streptococcus pyogenes by All-Trans Retinoic Acid-Differentiated HL-60 Cells: Roles of Azurophilic Granules and NADPH Oxidase. *PLoS One,* **4**, e7363.

Owen, C. A. & Campbell, E. J., 1999. The cell biology of leukocyte-mediated proteolysis. *J Leukoc Biol,* **65**, 137-150.

Paget, M. & Helmann, J. D., 2003. The sigma70 family of sigma factors. *Genome Biol,* **4**, 203.

Papayannopoulos, V., Metzler, K. D., Hakkim, A., Zychlinsky, A., 2010. Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps. *J Cell Biol,* **191**, 677-691.

Papayannopoulos, V., Staab, D., Zychlinsky, A., 2011. Neutrophil elastase enhances sputum solubilization in cystic fibrosis patients receiving DNase therapy. *PLoS One,* **6**, e28526.

Papayannopoulos, V. & Zychlinsky, A., 2009. NETs: a new strategy for using old weapons. *Trends Immunol,* **30**, 513-521.

Pathak, S. K., Basu, S., Basu, K. K., Banerjee, A., Pathak, S., Bhattacharyya, A., Kaisho, T., Kundu, M., Basu, J., 2007. Direct extracellular interaction between the early secreted antigen ESAT-6 of Mycobacterium tuberculosis

and TLR2 inhibits TLR signaling in macrophages. *Nature immunology,* **8**, 610-618.

Pember, S., Barnes, K., Brandt, S., Kinkade, J., 1983. Density heterogeneity of neutrophilic polymorphonuclear leukocytes: gradient fractionation and relationshipt to chemotactic stimulation. *Blood,* **61**, 1105-1115.

Perkins, T. T., Kingsley, R. A., Fookes, M. C., Gardner, P. P., James, K. D., Yu, L., Assefa, S. A., He, M., Croucher, N. J., Pickard, D. J., 2009. A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus Salmonella typhi. *PLoS Genet,* **5**, e1000569.

Perskvist, N., Long, M., Stendahl, O., Zheng, L., 2002. Mycobacterium tuberculosis promotes apoptosis in human neutrophils by activating caspase-3 and altering expression of Bax/Bcl-xL via an oxygen-dependent pathway. *The Journal of Immunology,* **168**, 6358-6365.

Peters, A., 1998. Just how big is the pulmonary granulocyte pool? *Clinical Science,* **94**, 7-19.

Pinto, A., Melo-Barbosa, H., Miyoshi, A., Silva, A., Azevedo, V., 2011. Application of RNA-seq to reveal the transcript profile in bacteria. *Genet Mol Res,* **10**, 1707-1718.

Podrez, E. A., Schmitt, D., Hoff, H. F., Hazen, S. L., 1999. Myeloperoxidase-generated reactive nitrogen species convert LDL into an atherogenic form in vitro. *Journal of Clinical Investigation,* **103**, 1547.

Pompei, L., Jang, S., Zamlynny, B., Ravikumar, S., McBride, A., Hickman, S. P., Salgame, P., 2007. Disparity in IL-12 release in dendritic cells and macrophages in response to Mycobacterium tuberculosis is due to use of distinct TLRs. *The Journal of Immunology,* **178**, 5192-5199.

Radomska, H. S., Huettner, C. S., Zhang, P., Cheng, T., Scadden, D. T., Tenen, D. G., 1998. CCAAT/enhancer binding protein α is a regulatory switch sufficient for induction of granulocytic development from bipotential myeloid progenitors. *Molecular and cellular biology,* **18**, 4301-4314.

Ramos-Kichik, V., Mondragón-Flores, R., Mondragón-Castelán, M., Gonzalez-Pozos, S., Muñiz-Hernandez, S., Rojas-Espinosa, O., Chacón-Salinas, R., Estrada-Parra, S., Estrada-García, I., 2009. Neutrophil extracellular traps are induced by Mycobacterium tuberculosis. *Tuberculosis (Edinb),* **89**, 29-37.

Reddy, V. A., Iwama, A., Iotzova, G., Schulz, M., Elsasser, A., Vangala, R. K., Tenen, D. G., Hiddemann, W., Behre, G., 2002. Granulocyte inducer C/EBPα inactivates the myeloid master regulator PU. 1: possible role in lineage commitment decisions. *Blood,* **100**, 483-490.

Remijsen, Q., Berghe, T. V., Wirawan, E., Asselbergh, B., Parthoens, E., De Rycke, R., Noppen, S., Delforge, M., Willems, J., Vandenabeele, P., 2011. Neutrophil extracellular trap cell death requires both autophagy and superoxide generation. *Cell research,* **21**, 290-304.

Reyrat, J. M., Lopez-Ramirez, G., Oferdo, C., Gicquel, B., Winter, N., 1996. Urease Activity Does Not Contribute Dramatically to Persistence of Mycobacterium bovis Bacillus Calmette-Guerin. *Infection and immunity,* **64**, 3934-3936.

Rienksma, R. A., Suarez-Diez, M., Mollenkopf, H. J., Dolganov, G. M., Dorhoi, A., Schoolnik, G. K., Martins Dos Santos, V. A., Kaufmann, S. H., Schaap, P. J., Gengenbacher, M., 2015. Comprehensive insights into transcriptional adaptation of intracellular mycobacteria by microbe-enriched dual RNA sequencing. *BMC Genomics,* **16**, 34.

Risso, D., Schwartz, K., Sherlock, G., Dudoit, S., 2011. GC-content normalization for RNA-Seq data. *BMC bioinformatics,* **12**, 480.

Robinson, M. D., McCarthy, D. J., Smyth, G. K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* **26**, 139-140.

Rodrigue, S., Provvedi, R., Jacques, P.-E., Gaudreau, L., Manganelli, R., 2006. The σ factors of Mycobacterium tuberculosis. *FEMS Microbiol Rev,* **30**, 926-941.

Rohde, K. H., Abramovitch, R. B., Russell, D. G., 2007. Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe,* **2**, 352-364.

Rohde, K. H., Veiga, D. F., Caldwell, S., Balazsi, G., Russell, D. G., 2012. Linking the transcriptional profiles and the physiological states of Mycobacterium tuberculosis during an extended intracellular infection. *PLoS Pathog,* **8**, e1002769.

Rook, G., Taverne, J., Leveton, C., Steele, J., 1987. The role of gamma-interferon, vitamin D3 metabolites and tumour necrosis factor in the pathogenesis of tuberculosis. *Immunology,* **62**, 229.

Russell, D. G., VanderVen, B. C., Lee, W., Abramovitch, R. B., Kim, M. J., Homolka, S., Niemann, S., Rohde, K. H., 2010. Mycobacterium tuberculosis wears what it eats. *Cell Host Microbe,* **8**, 68-76.

Rustad, T. R., Harrell, M. I., Liao, R., Sherman, D. R., 2008. The enduring hypoxic response of Mycobacterium tuberculosis. *PLoS One,* **3**, e1502.

Rustad, T. R., Minch, K. J., Brabant, W., Winkler, J. K., Reiss, D. J., Baliga, N. S., Sherman, D. R., 2013. Global analysis of mRNA stability in Mycobacterium tuberculosis. *Nucleic Acids Res,* **41**, 509-517.

Rustad, T. R., Minch, K. J., Ma, S., Winkler, J. K., Hobbs, S., Hickey, M., Brabant, W., Turkarslan, S., Price, N. D., Baliga, N. S., Sherman, D. R., 2014. Mapping and manipulating the Mycobacterium tuberculosis transcriptome using a transcription factor overexpression-derived regulatory network. *Genome Biology,* **15**, 502.

Said-Salim, B., Mostowy, S., Kristof, A. S., Behr, M. A., 2006. Mutations in Mycobacterium tuberculosis Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in Mycobacterium bovis. *Mol Microbiol,* **62**, 1251-1263.

Saïd-Salim, B., Mostowy, S., Kristof, A. S., Behr, M. A., 2006. Mutations in Mycobacterium tuberculosis Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in Mycobacterium bovis. *Mol Microbiol,* **62**, 1251-1263.

Sampson, J. B., Ye, Y., Rosen, H., Beckman, J. S., 1998. Myeloperoxidase and horseradish peroxidase catalyze tyrosine nitration in proteins from nitrite and hydrogen peroxide. *Archives of Biochemistry and Biophysics,* **356**, 207-213.

Sasmono, R. T., Ehrnsperger, A., Cronau, S. L., Ravasi, T., Kandane, R., Hickey, M. J., Cook, A. D., Himes, S. R., Hamilton, J. A., Hume, D. A., 2007. Mouse neutrophilic granulocytes express mRNA encoding the macrophage colony-stimulating factor receptor (CSF-1R) as well as many other macrophage-specific transcripts and can transdifferentiate into macrophages in vitro in response to CSF-1. *J Leukoc Biol,* **82**, 111-123.

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. D., Nathan, C., 2003. Transcriptional adaptation of Mycobacterium tuberculosis within macrophages insights into the phagosomal environment. *The Journal of experimental medicine,* **198**, 693-704.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., Ragg, T., 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol,* **7**, 3.

Schubert, O. T., Ludwig, C., Kogadeeva, M., Zimmermann, M., Rosenberger, G., Gengenbacher, M., Gillet, L. C., Collins, B. C., Röst, H. L., Kaufmann, S. H., 2015. Absolute proteome composition and dynamics during dormancy and resuscitation of Mycobacterium tuberculosis. *Cell Host Microbe,* **18**, 96-108.

Schuller, S., Neefjes, J., Ottenhoff, T. H., Young, D. B., 2001. Coronin is involved in uptake of *Mycobacterium bovis* BCG in human macrophages but not in phagosome maintenance. *Cellular microbiology,* **3**, 785-793.

Sengeløv, H., Kjeldsen, L., Borregaard, N., 1993. Control of exocytosis in early neutrophil activation. *The Journal of Immunology,* **150**, 1535-1543.

Shalem, O., Dahan, O., Levo, M., Martinez, M. R., Furman, I., Segal, E., Pilpel, Y., 2008. Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol,* **4**, 223.

Shinnick, T. & Good, R., 1994. Mycobacterial taxonomy. *European Journal of Clinical Microbiology and Infectious Diseases,* **13**, 884-901.

Silva, M. T., 2011. Macrophage phagocytosis of neutrophils at inflammatory/infectious foci: a cooperative mechanism in the control of infection and infectious inflammation. *J Leukoc Biol,* **89**, 675-683.

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., Ponting, C. P., 2014. Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics,* **15**, 121-132.

Singh, A., Singh, Y., Pine, R., Shi, L., Chandra, R., Drlica, K., 2006. Protein kinase I of Mycobacterium tuberculosis: cellular localization and expression during infection of macrophage-like cells. *Tuberculosis (Edinb),* **86**, 28-33.

Skvortsov, T. A. & Azhikina, T. L., 2012. Adaptive changes in gene expression of Mycobacterium tuberculosis during the development of the infection. *Russian Journal of Bioorganic Chemistry,* **38**, 341-353.

Smith, T. F. & Waterman, M. S., 1981. Identification of common molecular subsequences. *Journal of molecular biology,* **147**, 195-197.

Soualhine, H., Deghmane, A., Sun, J., Mak, K., Talal, A., Av-Gay, Y., Hmama, Z., 2007. *Mycobacterium bovis* Bacillus Calmette-Guérin Secreting Active Cathepsin S Stimulates Expressio of Mature MHC Class II Molecules and Antigen Presentation in Human Macrophages. *J Immunol,* **179**, 5137-5145.

Starck, J., Kallenius, G., Marklund, B. I., Andersson, D. I., Akerlund, T., 2004. Comparative proteome analysis of Mycobacterium tuberculosis grown under aerobic and anaerobic conditions. *Microbiology,* **150**, 3821-3829.

Steinwede, K., Maus, R., Bohling, J., Voedisch, S., Braun, A., Ochs, M., Schmiedl, A., Länger, F., Gauthier, F., Roes, J., 2012. Cathepsin G and neutrophil elastase contribute to lung-protective immunity against mycobacterial infections in mice. *The Journal of Immunology,* **188**, 4476-4487.

Summers, C., Rankin, S. M., Condliffe, A. M., Singh, N., Peters, A. M., Chilvers, E. R., 2010. Neutrophil kinetics in health and disease. *Trends Immunol,* **31**, 318-324.

Tak, T., Tesselaar, K., Pillay, J., Borghans, J. A., Koenderman, L., 2013. What's your age again? Determination of human neutrophil half-lives revisited. *Journal of leukocyte biology,* **94**, 595-601.

Tan, A. S. & Berridge, M. V., 2000. Superoxide produced by activated neutrophils efficiently reduces the tetrazolium salt, WST-1 to produce a soluble formazan: a simple colorimetric assay for measuring respiratory burst activation and for screening anti-inflammatory agents. *Journal of Immunological Methods,* **238**, 59-68.

Tan, B. H., Meinken, C., Bastian, M., Bruns, H., Legaspi, A., Ochoa, M. T., Krutzik, S. R., Bloom, B. R., Ganz, T., Modlin, R. L., 2006. Macrophages acquire neutrophil granules for antimicrobial activity against intracellular pathogens. *The Journal of Immunology,* **177**, 1864-1871.

Thorvaldsdottir, H., Robinson, J. T., Mesirov, J. P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform,* **14**, 178-192.

Timm, M., Bartelt, S., Moesby, L., Hansen, E. W., 2008. Cryopreservation of differentiated HL-60 cells for pyrogen testing. *J Immunol Methods,* **339**, 146-152.

Tjaden, B., 2015. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol,* **16**, 1.

Tomee, J. C., Hiemstra, P. S., Heinzel-Wieland, R., Kauffman, H. F., 1997. Antileukoprotease: an endogenous protein in the innate mucosal defense against fungi. *Journal of Infectious Diseases,* **176**, 740-747.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols,* **7**, 562-578.

Tsiftsogiou, A. S., Wong, W., Hyman, R., Minden, M., Robinson, S. H., 1985. Analysis of Commitment of Human HL-60 Cells to Terminal Granulocytic Maturation. *Cancer Research,* **45**, 2334-2339.

Unniraman, S., Chatterji, M., Nagaraja, V., 2002. DNA gyrase genes in Mycobacterium tuberculosis: a single operon driven by multiple promoters. *J Bacteriol,* **184**, 5449-5456.

Velu, C. S., Baktula, A. M., Grimes, H. L., 2009. Gfi1 regulates miR-21 and miR-196b to control myelopoiesis. *Blood,* **113**, 4720-4728.

Voskuil, M. I., Schnappinger, D., Visconti, K. C., Harrell, M. I., Dolganov, G. M., Sherman, D. R., Schoolnik, G. K., 2003. Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. *J Exp Med,* **198**, 705-713.

Voskuil, M. I., Visconti, K., Schoolnik, G., 2004. Mycobacterium tuberculosis gene expression during adaptation to stationary phase and low-oxygen dormancy. *Tuberculosis (Edinb),* **84**, 218-227.

Wagner, G. P., Kin, K., Lynch, V. J., 2013. A model based criterion for gene expression calls using RNA-seq data. *Theory in Biosciences,* **132**, 159-164.

Wang, X., Wang, H., Xie, J., 2011a. Genes and regulatory networks involved in persistence of Mycobacterium tuberculosis. *Science China Life Sciences,* **54**, 300-310.

Wang, X., Wang, H., Xie, J., 2011b. Genes and regulatory networks involved in persistence of Mycobacterium tuberculosis. *Sci China Life Sci,* **54**, 300-310.

Wang, Y., Li, M., Stadler, S., Correll, S., Li, P., Wang, D., Hayama, R., Leonelli, L., Han, H., Grigoryev, S. A., 2009a. Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol,* **184**, 205-213.

Wang, Y., Li, M., Stadler, S., Correll, S., Li, P., Wang, D., Hayama, R., Leonelli, L., Han, H., Grigoryev, S. A., Allis, C. D., Coonrod, S. A., 2009b. Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol,* **184**, 205-213.

Wang, Z., Gerstein, M., Snyder, M., 2009c. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews/Genetics,* **10**, 57-.

Wartha, F., Beiter, K., Albiger, B., Fernebro, J., Zychlinsky, A., Normark, S., Henriques-Normark, B., 2007. Capsule and d-alanylated lipoteichoic acids protect Streptococcus pneumoniae against neutrophil extracellular traps. *Cell Microbiol,* **9**, 1162-1171.

Wiker, H., 2009. MPB70 and MPB83–major antigens of Mycobacterium bovis. *Scand J Immunol,* **69**, 492-499.

Winn, W., Allen, S., Janda, W., Koneman, E., Procop, G., Schreckenberger, P., Woods, G., 2006. *Color atlas and textbook of diagnostic microbiology. 6th edition.* Philadelphia: Lippincott-Raven Publishers.

Witko-Sarsat, V., Friedlander, M., Khoa, T. N., Capeillère-Blandin, C., Nguyen, A. T., Canteloup, S., Dayer, J.-M., Jungers, P., Drüeke, T., Descamps-Latscha, B., 1998. Advanced oxidation protein products as novel mediators of inflammation and monocyte activation in chronic renal failure1, 2. *The Journal of Immunology,* **161**, 2524-2532.

Witko-Sarsat, V., Rieu, P., Descamps-Latscha, B., Lesavre, P., Halbwachs-Mecarelli, L., 2000. Neutrophils: molecules, functions and pathophysiological aspects. *Laboratory investigation,* **80**, 617-653.

Wong, K.-W. & Jacobs, W. R., 2013. Mycobacterium tuberculosis exploits human interferon γ to stimulate macrophage extracellular trap formation and necrosis. *Journal of Infectious Diseases,* **208**, 109-119.

Woodfin, A., Voisin, M.-B., Nourshargh, S., 2010. Recent developments and complexities in neutrophil transmigration. *Current opinion in hematology,* **17**, 9.

Worley-Morse, T. O. & Gunsch, C. K., 2015. A computational analysis of antisense off-targets in prokaryotic organisms. *Genomics,* **105**, 123-130.

Wösten, M., 1998. Eubacterial sigma-factors. *FEMS Microbiol Rev,* **22**, 127-150.

Xu, Y., Liu, W., Shen, H., Yan, J., Yang, E., Wang, H., 2010. Recombinant Mycobacterium bovis BCG expressing chimaeric protein of Ag85B and ESAT-6 enhances immunostimulatory activity of human macrophages. *Microbes Infect,* **12**, 683-689.

Yamamoto, K. & Johnston, R., 1984. Dissociation of phagocytosis from stimulation of the oxidative metabolic burst in macrophages. *J Exp Med,* **159**, 405-416.

Yipp, B. G., Petri, B., Salina, D., Jenne, C. N., Scott, B. N., Zbytnuik, L. D., Pittman, K., Asaduzzaman, M., Wu, K., Meijndert, H. C., 2012. Infection-induced NETosis is a dynamic process involving neutrophil multitasking in vivo. *Nat Med,* **18**, 1386-1393.

Zeller, M. E., Csanadi, A., Miczak, A., Rose, T., Bizebard, T., Kaberdin, V. R., 2007. Quaternary structure and biochemical properties of mycobacterial RNase E/G. *Biochem J,* **403**, 207-215.

Zhang, D.-E., Zhang, P., Wang, N.-d., Hetherington, C. J., Darlington, G. J., Tenen, D. G., 1997. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein α-deficient mice. *Proceedings of the National Academy of Sciences,* **94**, 569-574.

Zumla, A., George, A., Sharma, V., Herbert, R. H. N., of Ilton, B. M., Oxley, A., Oliver, M., 2015. The WHO 2014 Global tuberculosis report—further to go. *The Lancet Global Health,* **3**, e10-e12.