



UNIVERSITY OF <sup>TM</sup>  
KWAZULU-NATAL  
—  
INYUVESI  
YAKWAZULU-NATALI

The role of *Mycobacterium tuberculosis* pili in  
pathogenesis: growth and survival kinetics,  
gene regulation and host immune response,  
and *in vitro* growth kinetics.

By

Georgina Nyawo

Submitted in fulfilment of the requirements for the degree of Master of Medical Science  
(Medical Microbiology)

in the

Discipline of Medical Microbiology and Infection Control

School of Laboratory Medicine and Medical Sciences

College of Health Sciences

University of KwaZulu-Natal

South Africa

2016

Supervisor: Prof. M. Pillay

## **Declaration**

I, Georgina R. Nyawo (Student number: 212562563) declare that this dissertation contains my own work, which and has not been previously accepted for any degree, and is not being currently considered for any other degree, at this, or any other University. I declare that I have specifically acknowledged where I have used the work of others by following to the appropriate referencing agreements.



---

**Georgina Rumbidzai Nyawo**

September 2016

**Presentations emanating from this work:**

**Nyawo GR**, Moodley C, Mvubu NE, Pillay B, Christoffels A, Grosset J, and Pillay M.

GLOBAL TRANSCRIPTOME ANALYSIS REVEALS UP-REGULATION OF HOST CYTOKINE GENES, AND CYTOKINE AND CHEMOKINE RECEPTOR BINDING IN BALB/C MICE BY *MYCOBACTERIUM TUBERCULOSIS* CURLI PILI (MTP).

TB 2016 Conference (ICC, Durban, South Africa), 16-17 July 2016

Poster presentation

Winner of the BEST POSTER AWARD

**Nyawo GR**, Moodley C, Mvubu NE, Pillay B, Christoffels A, Grosset J, and Pillay M.

GLOBAL TRANSCRIPTOME ANALYSIS REVEALS UP-REGULATION OF HOST CYTOKINE GENES, AND CYTOKINE AND CHEMOKINE RECEPTOR BINDING IN BALB/C MICE BY *MYCOBACTERIUM TUBERCULOSIS* CURLI PILI (MTP).

University of KwaZulu-Natal, College of Health Sciences Research Symposium (UKZN, Durban, South Africa), 8-9 September 2016

Winner of the 2<sup>nd</sup> place AWARD for oral presentation.

## **Dedication**

To my family for their support, encouragement and love.

*“Trust in the Lord with all your heart, and lean not on your own understanding, in all your ways acknowledge Him, and He shall direct your paths.” Proverbs 3:5-6*

## **Acknowledgements**

I would like to thank my supervisor, Professor Manormoney Pillay for your continuous guidance during my studies. I am truly indebted to your support and encouragement. Your insights into the project have always been inspiring, and I am truly grateful to have learned so much from you.

The  $\Delta mtp$ -mutant strain used in this strain was kindly donated by the Medical Microbiology and Infection Control Department, UKZN.

I would like to thank the National Research Foundation (Prof. M. Pillay's research grant) and the South African Medical Research Council (SAMRC) (Prof Alan Christoffels' flagship grant) for the funding that allowed for this research. I would also like to thank the SAMRC, and the College of Health Science of the University of KwaZulu-Natal for the Masters bursaries.

Professor Bala Pillay, thank you for your much needed assistance and support. I thank Dr. Nonto Mvubu for her tremendous help with my bioinformatics analysis and writing. Thank you for your advice and devotion to the project. You have been a teacher, mentor and friend throughout this journey. Thank you. Chivonne Moodley, for your training and constant guidance, especially with the animal work. Your constant support and encouragement throughout. Your encouragement and advice on how to deal with the situation was invaluable. You went above and beyond your duties. I would also like to thank my colleagues, in the IPC department for their friendship, help and support.

To my friends and Akkas, thank you for keeping me sane throughout this project. Thank you for listening to me, advising me, and praying for me.

To my parents and family, thank you for your support and lessons to never give up. Thank you for the prayers.

To my nephews, NJ and Matthew, I thank God for your souls. Thank you for the laughs, the smiles and the tears. As you grow I hope I return the favor. I read that life is never as good as you imagine it to be nor as bad as you think it is. It will be just okay. Remember that.

And lastly, thank you to my God Almighty.

## Table of Contents

Declaration.....	ii
Presentations emanating from this work:.....	iii
Dedication.....	iv
Acknowledgements.....	v
Table of Contents.....	vi
List of Figures.....	x
List of Tables.....	xiii
List of Acronyms.....	xiv
Abstract.....	xv
<b>CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW.....</b>	<b>1</b>
<b>1.1 INTRODUCTION.....</b>	<b>2</b>
<b>1.2 LITERATURE REVIEW.....</b>	<b>4</b>
<b>1.2.1 Tuberculosis: Historical perspective.....</b>	<b>4</b>
<b>1.2.2 Epidemiology of tuberculosis.....</b>	<b>4</b>
<i>1.2.2.1 TB and HIV co-infection.....</i>	<i>5</i>
<i>1.2.2.2 M. tuberculosis and drug resistance.....</i>	<i>6</i>
<i>1.2.2.3 TB in South Africa.....</i>	<i>6</i>
<i>1.2.2.4 M. tuberculosis strain families.....</i>	<i>8</i>
<b>1.2.3 Pathophysiology of Tuberculosis.....</b>	<b>9</b>
<i>1.2.3.1 Active Tuberculosis.....</i>	<i>11</i>
<i>1.2.3.2 Immune responses against tuberculosis.....</i>	<i>11</i>
<i>1.2.3.3 Host and Pathogen interaction.....</i>	<i>12</i>
<i>1.2.3.4 Pattern recognition receptors in Mycobacterium tuberculosis infection.....</i>	<i>13</i>
<i>1.2.3.4.1 Membrane bound PRRs - Toll-like Receptors.....</i>	<i>15</i>
<i>1.2.3.4.2 Cytosolic PRRs - Nucleotide-binding Oligomerization Domain (NOD)-like receptors (NLRs).....</i>	<i>16</i>
<b>1.2.4 Adhesins.....</b>	<b>17</b>
<i>1.2.4.1 Curli pili and M. tuberculosis pili (MTP).....</i>	<i>18</i>
<b>1.2.5 Mouse models.....</b>	<b>20</b>
<b>1.2.6 Gene expression.....</b>	<b>20</b>
<b>1.3 Significance of work.....</b>	<b>21</b>
<b>1.4 Hypothesis.....</b>	<b>222</b>
<b>1.5 Aim.....</b>	<b>222</b>
<b>1.6 Objectives.....</b>	<b>22</b>
<b>1.7 Study Design.....</b>	<b>22</b>

1.8	Dissertation Layout.....	22
<b>CHAPTER 2: MATERIALS AND METHODS .....</b>		<b>23</b>
2.1	Ethics Approval.....	24
2.2	Growth of <i>Mycobacterium tuberculosis</i> strains.....	24
2.3	<i>In vitro</i> growth studies. ....	24
2.4	<i>In vivo</i> experiments .....	24
2.4.1	<i>Animals</i> .....	24
2.4.2	<i>Mice Aerosol infection (Day 0) and inoculum preparation</i> .....	25
2.4.3	<i>Organ harvesting and storage</i> .....	25
2.4.3.1	<i>Gross organ morphology</i> .....	25
2.4.4	<i>Determining Colony Forming Units (CFUs)</i> .....	25
2.4.5	<i>Extraction of RNA</i> .....	26
2.4.5.1	TriZol extraction .....	26
2.4.5.2	RNeasy purification .....	26
2.4.6	<i>RNA Sequencing</i> .....	26
2.4.7	<i>Bioinformatics and Data Analysis</i> .....	27
2.4.7.1	Raw data pre-processing - Quality assessment .....	27
2.4.7.2	Alignment of sequences to the mouse reference genome .....	28
2.4.7.3	Differential Expression Analysis of genes .....	28
2.4.7.4	Visualization of data and Global Transcriptomics.....	28
2.4.7.5	Identification of Significantly Differentially Expressed genes (SDEG) .....	28
2.4.7.6	Gene Ontology (GO) enrichment Analysis.....	29
2.4.7.7	Pathway and network analysis .....	29
<b>CHAPTER 3: RESULTS .....</b>		<b>30</b>
3.1	<i>In vitro</i> growth kinetics.....	31
3.2	Growth kinetics of <i>M. tuberculosis</i> in lungs and spleen during early infection.....	32
3.3	Gross organ pathology.....	33
3.4	<i>M. tuberculosis</i> burden in lungs 14days post infection.....	35
3.5	RNA Concentration and Integrity .....	36
3.6	Processing of RNA-Seq reads.....	37
3.6.1	<i>Pre-processing - Read Quality Control</i> .....	37
3.6.2	<i>Alignment and comparison with the reference genome</i> .....	39
3.7	Global changes in the lung transcriptional response of Balb/C mice to infection by <i>M. tuberculosis</i> .....	40
3.7.1	<i>Significantly differentially expressed genes (SDEGs) and Gene Ontology (GO) Analysis</i> .....	42
3.7.1.1	<i>Unique SDEGs in response to WT and <math>\Delta</math>mtp-mutant <i>M. tuberculosis</i> infection</i> .....	44
3.7.1.2	<i>Shared SDEGs in response to both WT and <math>\Delta</math>mtp-mutant infection</i> .....	45

3.7.1.3	<i>Upregulated genes</i> .....	45
3.7.1.4	<i>Downregulated genes</i> .....	48
<b>3.8</b>	<b>Ingenuity Pathway Analysis (IPA) of SDEGs</b> .....	<b>50</b>
<b>3.8.1</b>	<b>Canonical pathways enriched by all SDEGs</b> .....	<b>51</b>
3.8.1.1	<i>Canonical pathways enriched by both WT and <math>\Delta</math>mtp-mutant strains</i> .....	51
<b>3.8.2</b>	<b>Unique canonical pathways enriched by the SDEGs elicited by the host after WT and <math>\Delta</math>mtp-mutant <i>M. tuberculosis</i> infection</b> .....	<b>58</b>
3.8.2.1	<i>Enrichment of canonical pathways does not necessarily lead to their activation</i> .....	62
3.8.2.1.1	<i>Dendritic cell maturation pathway and Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathways</i> .....	64
<b>3.8.3</b>	<b>Network Analysis of SDEGs</b> .....	<b>65</b>
3.8.3.1	<i>Network Analysis of shared SDEGs</i> .....	65
3.8.3.2	<i>Network Analysis of unique SDEGs</i> .....	68
<b>3.8.4</b>	<b>The effect of MTP on specific host-pathogen interactions</b> .....	<b>71</b>
3.8.4.1	<i>Pattern Recognition Receptors</i> .....	71
3.8.4.1.1	<i>Toll-like receptors (TLRs), and Complement system receptors</i> .....	71
3.8.4.2	<i>Analysis of specific genes associated with host-pathogen interactions</i> .....	74
3.8.4.2.1	<i>Top 5 canonical pathways associated with host-pathogen interaction SDEGs</i> .....	76
3.8.4.2.2	<i>Networks associated with host-pathogen interaction SDEGs</i> .....	84
<b>3.8.5</b>	<b>The effect of MTP on host innate and adaptive immune response</b> .....	<b>85</b>
3.8.5.1	<i>The effect of MTP on upstream regulators (Cytokines and Transcriptional factors)</i> .....	86
3.8.5.2	<i>The effect of MTP on the regulation of the host immune response.</i> .....	90
<b>CHAPTER 4: DISCUSSION</b>	.....	<b>95</b>
<b>4.1</b>	<b>The effect of MTP on growth of <i>M. tuberculosis</i></b> .....	<b>96</b>
4.1.1	<i>MTP is vital for growth of <i>M. tuberculosis</i> in vitro</i> .....	96
4.1.2	<i>MTP is essential for growth in vivo</i> .....	96
<b>4.2</b>	<b>Transcriptome profiling post <i>M. tuberculosis</i> infection</b> .....	<b>97</b>
4.2.1	<i>The WT and <math>\Delta</math>mtp-mutant strains induce common as well as unique genes in Balb/C mice</i> ..	99
4.2.1.1	<i>Gene Ontology (GO) analysis reveals the association of MTP with host immunity</i> .....	99
<b>4.3</b>	<b>IPA Pathway analysis reveals host cellular pathways regulated by MTP during <i>M. tuberculosis</i> infection of Balb/C mice</b> .....	<b>100</b>
4.3.1	<i>Pattern recognition receptors (PRRs)</i> .....	100
4.3.1.1	<i>MTP affects the Toll-like Receptor (TLR) pathway</i> .....	101
4.3.1.2	<i>MTP induces the Complement System</i> .....	102
<b>4.3.2</b>	<b>The effect of MTP on host-pathogen interactions</b> .....	<b>103</b>
4.3.2.1	<i>TREM1 Signalling pathway</i> .....	103
4.3.2.2	<i>DC response</i> .....	104

4.3.2.3	<i>Role of NFAT in Regulation of the Immune Response, NF-KB Signalling pathway and iCOS-iCOSL Signalling in T Helper Cells</i> .....	104
<b>4.3.3</b>	<b><i>MTP plays a role in the host immune response</i></b> .....	<b>105</b>
4.3.3.1	<i>Phagosome function pathway</i> .....	105
4.3.3.2	<i>Antigen Presentation</i> .....	106
4.3.3.3	<i>Upstream regulators: Transcription factors</i> .....	106
4.3.3.3.1	The <i>M. tuberculosis</i> DosR transcription factor .....	107
4.3.3.4	<i>Cytokines and chemokines</i> .....	107
<b>4.4</b>	<b>MTP is associated with immune related networks and functions</b> .....	<b>109</b>
<b>4.5</b>	<b>Conclusions</b> .....	<b>110</b>
<b>4.6</b>	<b>Limitations of study</b> .....	<b>110</b>
<b>4.7</b>	<b>Recommendations for future work</b> .....	<b>111</b>
	<b>References</b> .....	<b>113</b>
	<b>Appendices</b> .....	<b>126</b>

## List of Figures

### Chapter 1:

<b>Figure 1.2.2.1.1</b>	An estimate of the HIV prevalence in TB cases in 2014 worldwide.	<b>5</b>
<b>Figure 1.2.2.2.1</b>	Estimated TB incidence including MDR and XDR in 2014.	<b>6</b>
<b>Figure 1.2.2.3.1</b>	Estimated TB incidence in 2014.	<b>8</b>
<b>Figure 1.2.2.4.1</b>	Global Lineages of <i>Mycobacterium tuberculosis</i> represented in colour.	<b>9</b>
<b>Figure 1.2.3.1.</b>	Pathophysiology of Tuberculosis.	<b>10</b>
<b>Figure 1.2.3.4.1</b>	Cell biology of the recognition of <i>M. tuberculosis</i> through pattern recognition receptors (PRRs)	<b>14</b>
<b>Figure 1.2.3.4.1.1</b>	Toll-Like Receptor Signaling Pathways. The best-characterized PRR family is the TLR family.	<b>16</b>

### Chapter 2:

<b>Figure 2.4.7.1</b>	Summary of the RNA-Seq analysis work flow that was followed.	<b>27</b>
-----------------------	--	-----------

### Chapter 3:

<b>Figure 3.1.1</b>	<i>In-vitro</i> growth kinetics of $\Delta mtp$ -mutant, <i>mtp</i> -complemented and wild type strains of <i>M. tuberculosis</i> .	<b>31</b>
<b>Figure 3.2.1</b>	<i>In vivo</i> growth and dissemination of wild type, $\Delta mtp$ -mutant and complemented strains of <i>M. tuberculosis</i> in BALB/c mice.	<b>32</b>
<b>Figure 3.3.1</b>	Mean lungs and spleen sizes after aerosol inoculation, at week 4 post <i>M. tuberculosis</i> challenge.	<b>33</b>
<b>Figure 3.3.2</b>	<i>In vivo</i> <i>M. tuberculosis</i> burden at day 28.	<b>34</b>
<b>Figure 3.3.3</b>	Gross appearance of lungs and spleens from aerosol infected mice at 4 weeks after challenge with <i>M. tuberculosis</i> strains.	<b>34</b>
<b>Figure 3.4.1</b>	<i>M. tuberculosis</i> burden in infected mice lungs at day 14 post-infection.	<b>35</b>
<b>Figure 3.6.1.1</b>	Sample FastQC output of the Illumina reads from the 6 libraries that were generated from the RNA obtained from the WT and $\Delta mtp$ -mutant infected lungs.	<b>37</b>
<b>Figure 3.6.1.2</b>	FastQC output of the reverse sequence of the $\Delta mtp$ -mutant infected lung sample.	<b>38</b>
<b>Figure 3.7.1</b>	Density plot of expressed genes.	<b>40</b>
<b>Figure 3.7.2</b>	Dispersion plot of expressed genes.	<b>41</b>
<b>Figure 3.7.1.1</b>	Significantly Differentially Expressed Genes (SDEGs).	<b>42</b>

<b>Figure 3.7.1.3.1</b>	Differential expression of host genes in response to <i>M. tuberculosis</i> infection 14days post-infection.	<b>45</b>
<b>Figure 3.7.1.4.1</b>	Differential expression of host genes in response to <i>M. tuberculosis</i> infection 14days post-infection.	<b>47</b>
<b>Figure 3.8.1.1.1</b>	IPA of the top 10 canonical pathways enriched from all genes, in response to the WT and $\Delta mtp$ -mutant infection relative to the uninfected.	<b>50</b>
<b>Figure 3.8.1.1.2</b>	Agranulocyte adhesion and Diapedesis Pathway.	<b>53</b>
<b>Figure 3.8.1.1.3</b>	LXR/RXR Activation Pathway.	<b>55</b>
<b>Figure 3.8.2.1</b>	Top canonical pathways enriched from the unique $\Delta mtp$ -mutant infection genes as shown in IPA.	<b>58</b>
<b>Figure 3.8.2.2</b>	IPA pathway and SDEGs associated with The role of Pattern Recognition Receptors (PRRs) in recognition of bacteria and viruses signalling.	<b>58</b>
<b>Figure 3.8.2.3</b>	GABA Receptor Signalling pathway	<b>59</b>
<b>Figure 3.8.2.4</b>	Glutamate Receptor Signaling pathway.	<b>60</b>
<b>Figure 3.8.2.1.1</b>	Activation z-scores of the canonical pathways enriched from the unique genes induced by each strain as shown in IPA.	<b>61</b>
<b>Figure 3.8.2.1.2</b>	Top canonical pathways enriched from the unique genes as shown in IPA.	<b>62</b>
<b>Figure 3.8.2.1.1.1</b>	Differential expression of host genes in response to <i>M. tuberculosis</i> infection 14days post-infection.	<b>63</b>
<b>Figure 3.8.3.1.1</b>	Wild type and $\Delta mtp$ -mutant specific Network 4 enriched by shared SDEGs.	<b>65</b>
<b>Figure 3.8.3.2.1</b>	Top 2 Wild type and $\Delta mtp$ -mutant specific Networks enriched by SDEGs.	<b>68</b>
<b>Figure 3.8.4.1.1.1</b>	Differential expression of host genes in response to <i>M. tuberculosis</i> infection 14days post-infection.	<b>71</b>
<b>Figure 3.8.4.1.1.2</b>	Complement Signalling pathway.	<b>72</b>
<b>Figure 3.8.4.2.1</b>	Differential expression of host genes involved in the host-pathogen interaction, in response to <i>M. tuberculosis</i> infection 14days post-infection.	<b>74</b>
<b>Figure 3.8.4.2.1.1</b>	The top 5 canonical pathways associated with the host-pathogen interaction genes.	<b>75</b>
<b>Figure 3.8.4.2.1.2A</b>	TREM1 Signalling pathway.	<b>76</b>
<b>Figure 3.8.4.2.1.2B</b>	The Dendritic Cell (DC) Maturation pathway.	<b>77</b>

<b>Figure 3.8.4.2.1.2C</b>	The Role of NFAT in Regulation of the Immune Response pathway.	<b>78</b>
<b>Figure 3.8.4.2.1.2D</b>	The NF- $\kappa$ B Signalling pathway.	<b>79</b>
<b>Figure 3.8.4.2.1.2E</b>	The iCOS-iCOSL Signalling in T Helper Cells pathway.	<b>80</b>
<b>Figure 3.8.4.2.1.3</b>	IPA Heat maps for top pathways enriched from the genes associated in host-pathogen interaction:	<b>81</b>
<b>Figure 3.8.5.1</b>	Antigen Presentation Pathway.	<b>83-84</b>
<b>Figure 3.8.5.1.1</b>	Upstream regulator analysis.	<b>86</b>
<b>Figure 3.8.5.1.2</b>	Interferon signalling pathway and the associated genes.	<b>86-87</b>
<b>Figure 3.8.5.1.3</b>	STAT1 network.	<b>88</b>
<b>Figure 3.8.5.2.1</b>	Crosstalk between Dendritic cells and Natural Killer Cells pathway.	<b>89</b>
<b>Figure 3.8.5.2.2</b>	Genes involved in the Macrophage migration inhibitory factor (MIF) Regulation of Innate Immunity pathway.	<b>90</b>
<b>Figure 3.8.5.2.3</b>	Production of Nitric Oxide and Reactive Oxygen Species in Macrophages.	<b>90-91</b>
<b>Figure 3.8.5.2.4</b>	Phagosome formation pathway.	<b>92</b>

## List of Tables

### Chapter 1:

<b>Table 1.2.4.1</b>	A list of the currently known adhesins of <i>Mycobacterium tuberculosis</i> .	<b>18</b>
----------------------	---	-----------

### Chapter 3:

<b>Table 3.5.1</b>	RNA concentrations post extraction.	<b>36</b>
<b>Table 3.6.2.1</b>	Summary of the number of the 100bp paired ended sequences and percent of mapped reads obtained for the 6 libraries mapped to the <i>mus musculus</i> mm10 mouse genome.	<b>39</b>
<b>Table 3.7.1.1</b>	Top ten up-regulated and down regulated SEDGs elicited by the host in response to each <i>M. tuberculosis</i> infection relative to uninfected mice.	<b>42-43</b>
<b>Table 3.7.1.1.1</b>	The GO functional categories (FCs) of the GO molecular functions from the unique gene datasets.	<b>43-44</b>
<b>Table 3.7.1.3.1</b>	Gene ontology enrichment of the 95 common up-regulated genes.	<b>46</b>
<b>Table 3.7.1.4.1</b>	Gene ontology enrichment of the 33 common down regulated genes.	<b>48</b>
<b>Table 3.8.1</b>	A summary of the numbers of significantly differentially expressed genes, and their associated networks and pathways.	<b>49</b>
<b>Table 3.8.1.1.1</b>	List of the top ten most significantly enriched canonical pathways by all genes elicited in both infections.	<b>51</b>
<b>Table 3.8.2.1</b>	Top 10 IPA pathways significantly enriched by the WT infection SDEGs.	<b>57</b>
<b>Table 3.8.3.1.1</b>	Diseases and functions associated with networks enriched by shared genes.	<b>64</b>
<b>Table 3.8.3.2.1</b>	Top diseases and functions associated with unique genes from top 10 IPA networks.	<b>66</b>
<b>Table 3.8.4.2.1</b>	Top 10 most upregulated and down regulated genes associated with host-pathogen interactions.	<b>73</b>
<b>Table 3.8.4.2.2.1</b>	IPA network analysis of host-pathogen interaction genes.	<b>82</b>

## **List of Acronyms**

<b>Abbreviation</b>	<b>Abbreviated term</b>
APC	Antigen presenting cell
BCG	Bacille Calmette-Guerin
CFU	Colony forming units
DC	Dendritic Cell
DC-SIGN	DC-specific intercellular-adhesion-molecule-3-grabbing non-integrin
DE	Differential Expression
DNA	Deoxyribonucleic Acid
IFN- $\gamma$	Interferon-gamma
IL	Interleukin
IRF	Interferon regulatory factor
LPS	Lipopolysaccharide
ManLAM	mannose-capped LAM
MDR	Multi Drug Resistant
MHC	major histocompatibility complex
MR	mannose receptor
mRNA	messenger RNA
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
NK	Natural Killer
NLR	NOD-like receptor
NOD	Nucleotide-binding oligomerization domain
OD <sub>600</sub>	Optical density at 600nm
PAMPs	pathogen-associated molecular patterns
PRR	pathogen recognition receptors
QC	Quality Control
RNA	Ribonucleic Acid
RNA-Seq	RNA sequencing
RNI	reactive nitrogen intermediates
ROS	reactive oxygen species
SDE	Significantly differentially expressed
TB	Tuberculosis
TNF- $\alpha$	Tumor necrosis factor-alpha
TLR	Toll-like receptor
XDR	Extensively Drug Resistant
WHO	World Health Organization

## **Abstract**

**Background:** There is an urgent need to identify novel drug and diagnostic targets for *Mycobacterium tuberculosis*, as well as vaccine candidates because of the lack of biomarkers that have been adequately characterised for this purpose. Improving our current understanding and knowledge of immune-pathogenesis of TB can assist in the design of these new vaccines and drug candidates, and facilitate the evaluation of their effectiveness. It has been increasingly demonstrated that biosignatures, rather than individual biomarkers, would be more beneficial for this purpose due to the multifactorial complexity of TB. *Mycobacterium tuberculosis* curli pili (MTP), encoded by the *mtp* gene was recently reported as a potential candidate biomarker for diagnostic, drug and vaccine development. Therefore, this study aimed to further the understanding of the role of MTP in pathogenesis by evaluating its role in, host-pathogen interactions and host immune response in a mouse model infected with  $\Delta mtp$ -mutant, *mtp*-complemented and wild type (WT) strains of the V9124 strain of *M. tuberculosis*.

**Methods:** Female Balb/C mice were infected with the WT,  $\Delta mtp$ -mutant and *mtp*-complemented strains. Bacterial growth kinetics were carried out to determine the effect of the gene on growth by observing *in vitro* CFU assays, and *in vivo* CFUs from *M. tuberculosis* infected tissues. Global changes in gene expression were analysed by RNA-sequencing. The transcriptional responses of the mice lungs after infection with the WT and  $\Delta mtp$ -mutant strains, relative to that of uninfected mice were investigated 14 days post infection, using the mm10 mouse reference genome. The transcriptional profiles of lungs infected with the  $\Delta mtp$ -mutant, and WT strains for fourteen days were compared using Tophat, Cuffdiff and various other bioinformatics tools to further analyse the effect of MTP.

**Results and discussion:** The growth assays revealed that the  $\Delta mtp$ -mutant strain grew at a significantly decreased rate compared to that of the WT during the log phase in broth culture and in the lungs and spleen of infected mice. This suggests that the deletion of the *mtp* gene results in a slower growth rate of the bacterium, and hence, MTP has an effect on the growth of *M. tuberculosis* *in vitro* and *in vivo*. A total of 512 and 1 059 genes were significantly differentially expressed due to infection with WT and  $\Delta mtp$ -mutant respectively. Only 128 genes were common in both infections, and of those, 74.2% and 25.8% were upregulated and down regulated respectively. The number of differentially expressed genes uniquely induced by WT and  $\Delta mtp$ -mutant strains were 384 and 930 respectively. Analysis of fold changes in Gene Ontology functional categories (FCs) revealed that over-represented FCs from WT infection included functions associated with host immune response such as chemokine and cytokine receptor binding and cytokine activity ( $p < 0.05$ ). In contrast, the  $\Delta mtp$ -mutant did not elicit these host functions. Specific pathways were identified to be affected by the absence of MTP from the bacilli cell wall. This was seen by the lower enrichment of pathways, molecular networks and cytokines involved in the host immune response to the  $\Delta mtp$ -mutant strains compared to the WT. Pathways involved in host-pathogen interaction (PRR pathway, complement pathway and TLR signalling pathway), host

immune response (antigen presentation pathway and phagosome maturation pathway), as well as regulation of the immune response (MIF (Macrophage Migration Inhibitory Factor) Regulation of Innate Immunity pathway, Crosstalk between Dendritic cells and Natural Killer Cells pathway) were positively enriched only during WT infection, and negatively enriched in the  $\Delta mtp$ -mutant infection. Our current results support findings from other studies that infection with the WT strain of *M. tuberculosis* initiated a cascade of immune responses and inflammatory signals from the host. WT associated networks were associated with immunologically related functions such as Immunological Disease, and antimicrobial inflammatory response, whereas the top  $\Delta mtp$ -mutant networks were not related to these functions. Overall, the gene ontology, canonical pathway and network analysis in this study suggests that MTP has a significant impact on the biological functions, and pathways that are essential for host immunity during *M. tuberculosis* infection.

**Conclusions:** MTP was proven to play a significant role in the specific host-pathogen interactions following *M. tuberculosis* infection, resulting in host immune responses essential to the hosts' defence by triggering the innate immune response and inflammatory response. Further, it is associated with the regulation of the immune response by transcription factors and cytokines. Taken together, transcriptome analyses of lung tissue infected with the MTP-deficient strain of *M. tuberculosis* has shown MTP to be a strong immunogen. These findings provide further supporting evidence to previous studies that suggested that MTP is a strong therapeutic and vaccine candidate.

## **CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW**

## 1.1 INTRODUCTION

More than a century since its discovery, TB still continues to be the major cause of mortality in the world [1,2] together with the Human Immuno-deficiency Virus (HIV) [2]. Tuberculosis (TB) is believed to be the leading infectious disease globally [2], with an estimated mortality of 1.5 million people in 2014, about 27% of whom were HIV-positive. About a third of the world's population is thought to be latently infected [2]. The highest TB incidence was reported in Asia and Africa, with South Africa documenting the highest incident burden in the latter continent [3].

*Mycobacterium tuberculosis*, the causative agent of TB in humans belongs to the *Mycobacterium tuberculosis* complex (MTBC) that includes also *M. bovis*, *M. bovis* BCG, *M. bovis* subsp. *caprae* comb, *M. africanum*, *M. microti* and *M. canettii*, that cause TB in different hosts [4]. It is a facultative intracellular bacterial pathogen a slow growing, non-motile, non-spore forming bacillus, that flourishes in the oxygen rich apex of lungs. TB is a chronic airborne infectious disease, with the source being an infected patient with cavitary pulmonary TB [3].

Of the estimated 9.6 million new TB cases in 2014, only 6 million were recorded and documented by the WHO. This translates into a staggering 37% undocumented, or most importantly, undiagnosed cases. In order to reduce these major gaps in the detection and treatment of TB, the development of new tools is required [2]. To date, there is no TB vaccine that is completely protective against this deadly disease, and the development of new drugs is very slow. This is worsened by the long therapy regimens (up to six months) of the current therapy regimens, which makes patient compliance difficult, resulting in the emergence of resistant strains of the organism, that significantly threaten TB control programmes [5].

The WHO have implemented the “The End TB Strategy” in order to end the TB epidemic (REF), by decreasing TB mortality by 90% by the year 2030 [2]. In order to achieve this goal, the research and development pipeline would need key breakthroughs including a point-of-care diagnostic test that is capable of distinguishing latent TB infection from active TB disease, a short and effective treatment for latent TB, and an effective post-exposure vaccine [2]. Therefore, there is a need to identify novel drug targets for *M. tuberculosis* because of the lack of biomarkers that have been adequately characterised for this purpose. Currently, there are 7 vaccines that are in the Phase I of clinical trials, and 8 that in the Phase II or Phase III of clinical trials [2]. The immune responses elicited early in the lung during TB infection have been shown to be vital in the TB pathogenesis, and therefore possess great potential in drug and vaccine discovery [6].

Adhesion to host cells is essential for most bacterial pathogens. Microbial adhesins mediate the essential processes that determine bacterial attachment to host tissue surfaces [7]. The attachment and invasion

capacity of *M. tuberculosis* is important to its pathogenesis as it enables entry into the host cells, i.e. macrophages and other host cell types. This provides the bacilli with a chance to resist killing in the extra-cellular environments, and to replicate in the intracellular sites [7]. *M. tuberculosis* has been shown to adhere to and invade the alveolar epithelial cells [8], possibly due to the presence of adhesins that interact with receptors present on these cell surfaces, since bacterial adhesion to host cell surfaces is specific [9]. There is increasing evidence that *M. tuberculosis* curli pili (MTP) is one of the major adhesins involved in pathogenesis of this organism. MTP encoded by the Rv3312A (*mtp*) gene [10] play a significant role in biofilm production *in vitro* [11], as well as in adhesion and invasion of epithelial cells [12,13] and macrophages [14,15]. Bacterial pili have also been reported to induce host immune response via cytokine/chemokine production [16]. *M. tuberculosis* has been documented to elicit the production of certain cytokines and chemokines in different types of host cells including microglia, macrophages [17], and epithelial cells [18]. However, Ramsugit and Pillay, 2015 have shown surprisingly minimal contribution of MTP to the overall cytokine and chemokine induction in the latter non-phagocytic host cell type, suggesting this to be one of the mechanisms used to bypass the host defence [18]. Therefore, there is accumulating evidence in recent years of the utility of MTP as a biomarker with potential to be targeted for diagnostic and therapeutic interventions.

Upon infection, different signals induce the alteration of gene expression, resulting in up-regulation and/or down regulation of specific genes [19]. The number of transcriptomic studies has increased over the years, resulting in the discovery and listing of functionally categorized genes [20]. In the past decade, Next-generation RNA sequencing (RNA-Seq) of microbes has evolved into a standard technique for quantifying and annotating microbial transcripts [21,22] and recently, has allowed for the positive identification of both lowly and highly expressed genes in a single experiment [21]. However, the application of this technique in whole transcriptomic studies in *M. tuberculosis* infection models is still in its infancy, and have explored the pathogen and its interactions with experimental hosts in order to advance our knowledge TB pathology [20]. In addition, transcriptomic research has also elucidated molecular mechanisms associated with virulence and resistance to drugs [20].

In order to complement the growing body of knowledge on the role of MTP in TB pathogenesis, and to identify potential biosignatures for therapeutic interventions, RNA-Seq was used in the current study to elucidate and quantify significantly differentially expressed genes induced by MTP in the lungs of Balb/C mice during early infection with a *mtp* knockout mutant strain of *Mycobacterium tuberculosis*.

## **1.2 LITERATURE REVIEW**

### **1.2.1 Tuberculosis: Historical perspective**

The causative agent of the deadly disease Tuberculosis (TB), *M. tuberculosis* was discovered in 1882, by Robert Koch [1,23]. This was however, thousands of years after the disease had been documented to be a problem, and dates as far back as the writing of the Biblical books such as Deuteronomy and Leviticus, more than two thousand years ago [1], and the ancestor from whom it is thought to have evolved from was found in East Africa more than 3 million years ago [24]. Koch's discovery paved the way for research, enabling Scientists the opportunity to fight against TB with an understanding of how the pathogen causes disease, leading to therapeutic regimens against the disease, and efforts of preventing and controlling its spread [1].

The World Health Organisation (WHO) has been at the forefront in the fight against TB. In 2015, the WHO concluded their Millennium Development Goals (MDGs) where they had set global TB targets to stop and decrease TB incidence worldwide [2,25]. The positive achievements were highlighted, including the overall decrease of TB mortality by 47 percent, and the continuous decrease in TB incidence by 1.5 percent annually, since 2000 [2]. Recently, the Sustainable Development Goals (SDGs), and simultaneously, the Stop TB Strategy transitions to the End TB Strategy had been initiated.

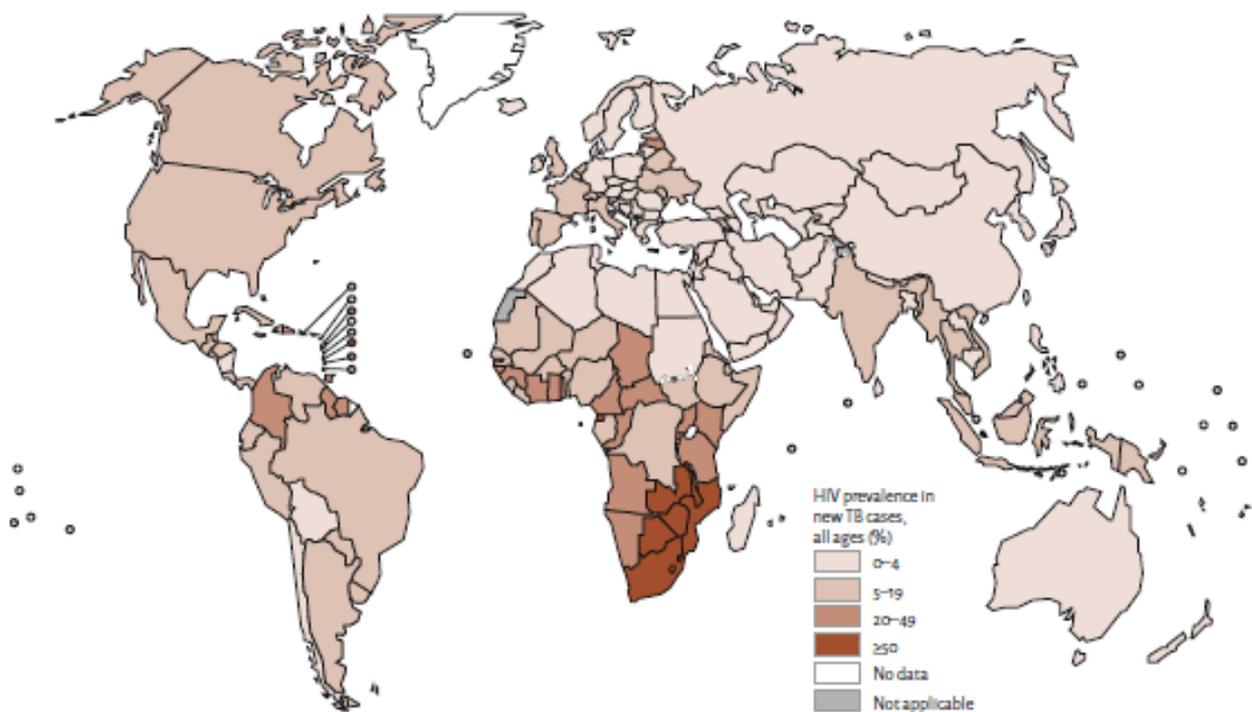
It has been twenty years since the establishment of the TB global monitoring system was established by the World Health Organisation, and the annual collection of data and statistics have shown that advances in prevention, treatment and diagnosis have failed to see an end to this disease [2]. The development of the only TB vaccine, the attenuated strain Bacille Calmette-Gue´rin (BCG), by Calmette and Gue´rin, was a major step forward in the fight against TB. However, the vaccine has proven to be not as effective in adults. The exploitation of immune-pathogenesis of tuberculosis using global transcriptomics is promising in the pursuit of a proper vaccine as well as immunotherapy, especially in patients infected with drug resistant strains of tuberculosis.

### **1.2.2 Epidemiology of tuberculosis**

TB has been deemed an international public health problem. A decade ago, the WHO documented almost a third of the world's population to have latent TB infection [26]. With an estimate of 9.6 million new cases (12 percent being HIV-positive), TB has claimed the lives of about 1.5 million people in 2014, 37% of them being HIV positive [2]. Of the new cases, 5.4 million were men, 3.2 million were women, and 1.0 million were children [2]. The TB epidemic was aggravated by co-infection of TB patients with the Human Immunodeficiency Virus (HIV), as well as the rise in drug-resistant TB making it even more difficult to treat the disease [2].

### 1.2.2.1 TB and HIV co-infection

The combination of TB and HIV is a lethal one, which results in the speeding of both conditions in a patient. The increase in the prevalence of TB is related to the increase of HIV infection in Sub Saharan Africa [27]. An individual infected by the HIV virus is more likely to develop active TB when they are exposed to the *M. tuberculosis* bacilli. This lethal combination is the most frequent cause of death amongst immune compromised individuals [28]. This is because HIV makes the immune system weak, and less likely to fight TB infection, leading to the high TB mortality among HIV-positive individuals. About a third of the HIV infected people worldwide are co-infected with TB. In 2014, the number of people living with HIV, and developed TB was estimated to be 1.2 million [2], and only a third of those were documented as receiving antiretroviral therapy (ART). However, there was a 32% decrease in the HIV-associated TB mortality rate between 2004 and 2014 [2].



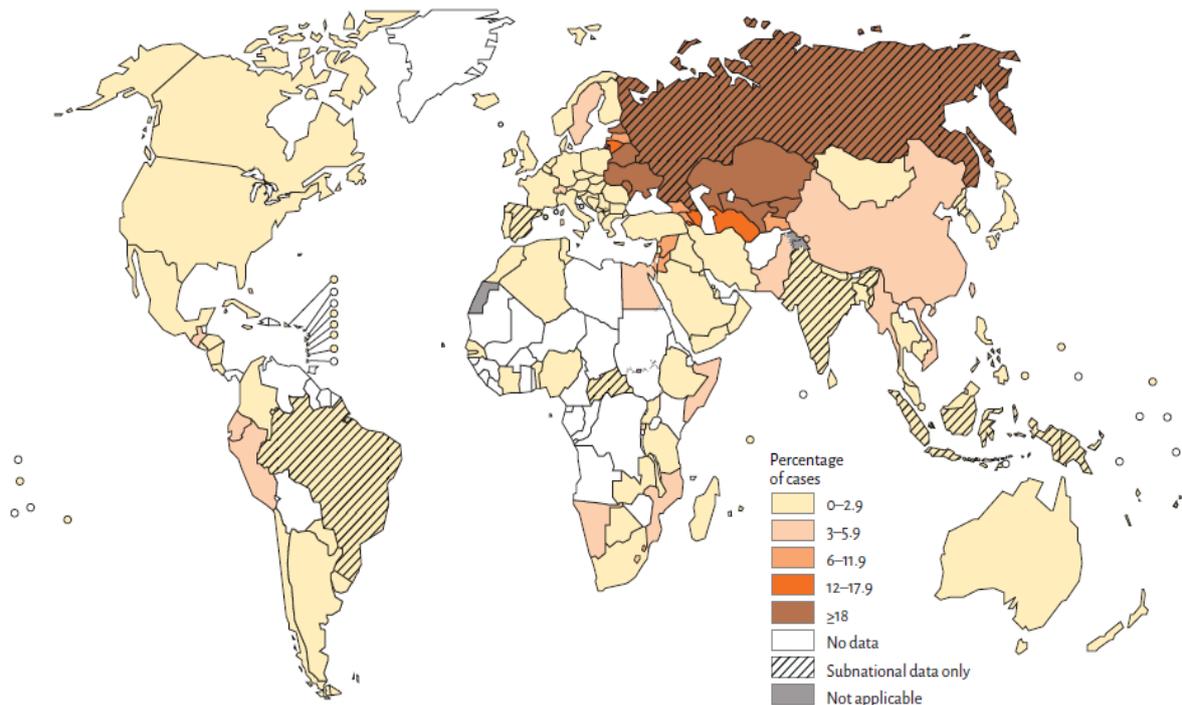
**Figure 2.2.2.1.1: An estimate of the HIV prevalence in TB cases in 2014 worldwide [2].**

A majority of these cases are observed in the sub Saharan region (Figure 1.2.2.1.1). Whilst TB is the leading cause of death amongst the HIV infected population, an overall 32% of TB cases in the African region were estimated to be co-infected with HIV [2]. This proportion ranks Africa the highest with HIV-TB co-infection in the world as it accounts for 74% of all TB cases among HIV-positive people worldwide. Some parts of southern Africa have more than 50% of TB their cases living with HIV [2]. South Africa alone constituted up to 17% of the HIV burden globally in 2007 [29], and the combined HIV and TB epidemics are a key public health concern [30].

### 1.2.2.2 *M. tuberculosis* and drug resistance

In addition to TB-HIV co-infection, the emergence of drug resistant (DR) *M. tuberculosis* strains has become a major problem to cure TB patients using the standard anti-TB drug regimens, increasing the threat to TB control programs [2]. The resistant strains include multidrug-resistant (MDR) (resistant to the first-line-TB drugs, Rifampicin (RIF) and Isoniazid (INH) [2]), and extremely drug-resistant (XDR) strains (MDR-plus resistance to any injectable second-line-TB drugs namely the fluoroquinolones, amikacin, capreomycin or kanamycin [2,5,31]). A further resistant to more drugs results in totally drug resistant TB (TDR-TB) in extreme cases. Totally drug resistant (TDR) strains, are defined as those that are resistant to all the 1<sup>st</sup> and 2<sup>nd</sup> line drugs [32]. TDR were first documented in Italy in 2007 [33], and have been reported to be emerging most recently from South Africa [34] in 2013 and India [35,36] in 2012.

In 2014, 480 000 MDR-TB cases were estimated to have occurred worldwide, and of those, 190 000 died [2]. India, China and Russia were reported to have the largest burden (54%) of cases of MDR-TB (Figure 1.2.2.2.1). A total 9.7% of people with MDR-TB are estimated to have XDR-TB [2].



**Figure 1.2.2.2.1: Estimated TB incidence including MDR and XDR in 2014 [2].**

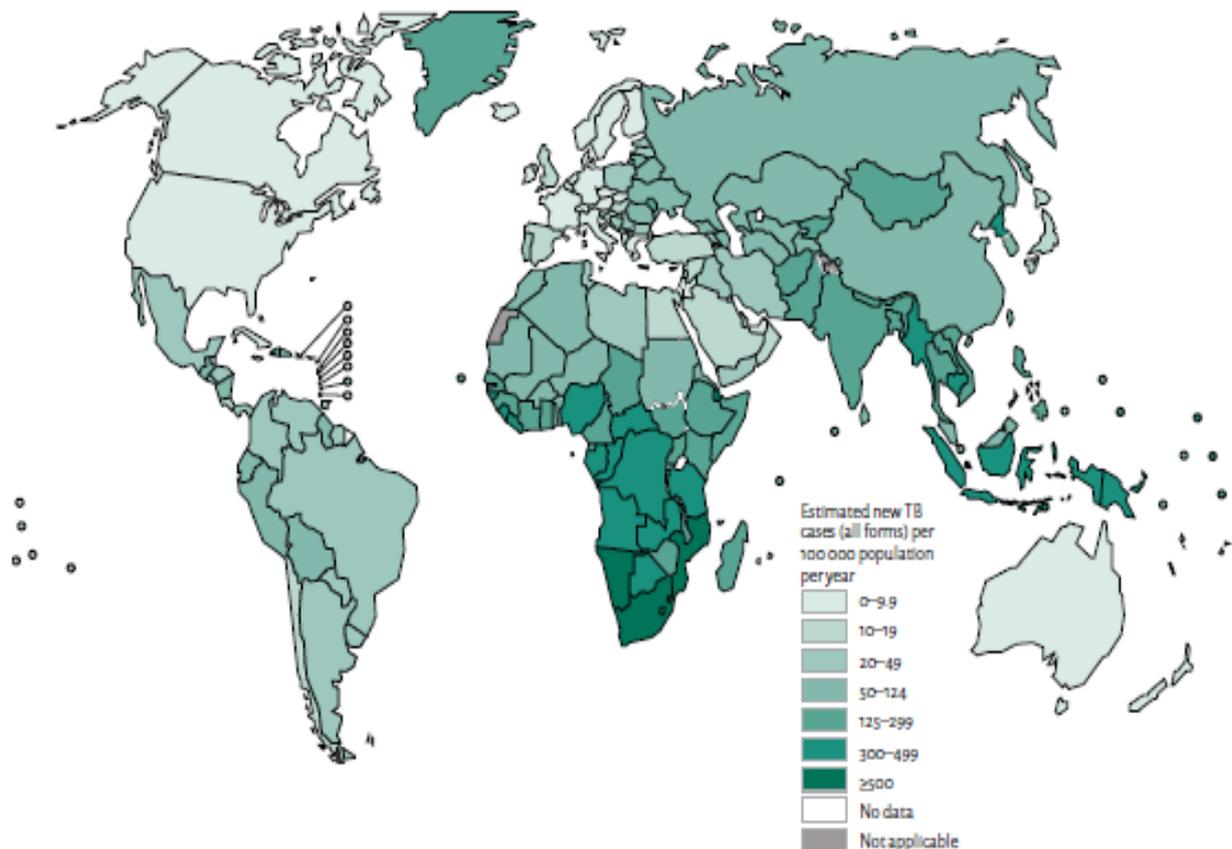
### 1.2.2.3 *TB in South Africa*

The African Region carried 28% of the burden of the estimated number of TB cases in 2014 [2]. In Southern Africa, roughly 1% of the population develops TB annually. Figure 1.2.2.3.1 provides a graphic overview of the TB incidence as estimated by the WHO in 2014 [2]. South Africa is one of the high-burden countries, ranked at 6<sup>th</sup> highest [2,37], accounting for up to 19% of all adult TB cases in

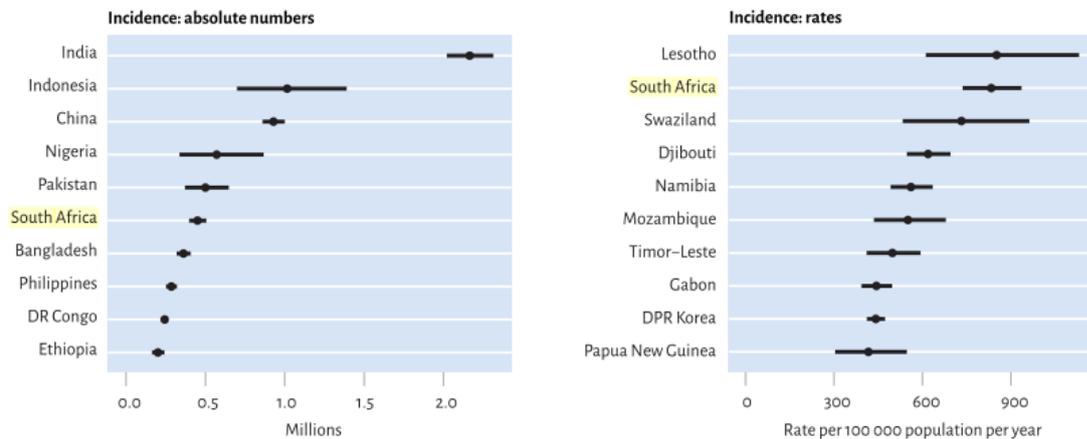
the world [26,29,30] (Figure 1.2.2.3.1). The high mortality rate of MDR and XDR-TB patients in South Africa (47%) could possibly be related to the high prevalence (696 per 100 000) of HIV-TB co-infection [2,5,34].

The KwaZulu-Natal (KZN) province is one of the high ranking provinces with new cases [5]. Currently, the TB incidence rate is 295 cases per 100,000 in this province. In 2006, 31% of the national TB cases in SA were recorded in the KZN province, and the first cases of MDR and XDR-TB were reported in this province [5,29].

A)



B)

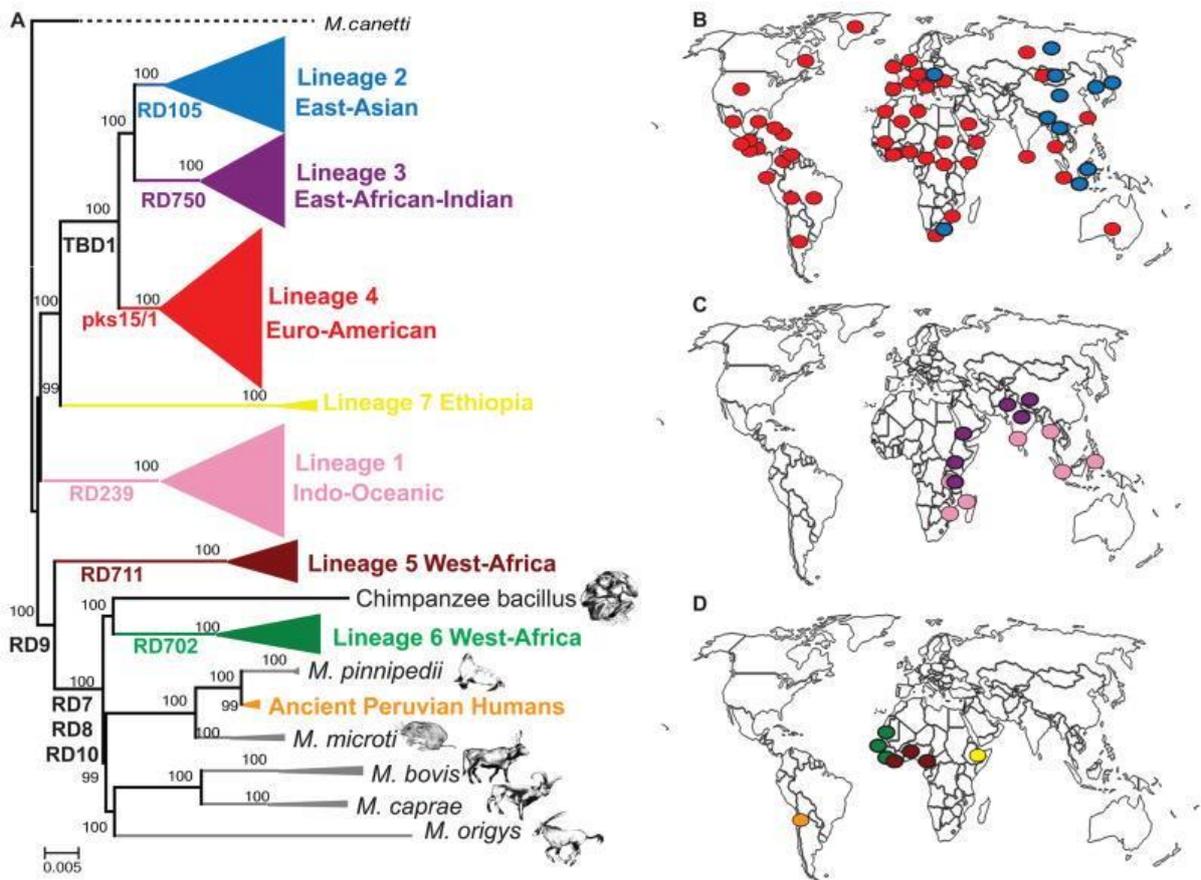


**Figure 1.2.2.3.1: Estimated TB incidence in 2014.** A) Estimated new TB cases of all forms per year. B) Top 10 countries, 2014. The range shows the lower and upper bounds of the 95% uncertainty interval. The bullet mark is the best estimate [2].

#### 1.2.2.4 *M. tuberculosis* strain families

The MTBC demonstrates a robust phylogeographical population structure, and thus, consists of 7 phylogenetically lineages that are distinctly related to different geographical regions (Figure 1.2.2.4.1) [38]. Whole genome sequencing is the current gold standard for strain typing of *M. tuberculosis* [39]. The most broadly spread amongst the 7 lineages are Lineage 2 and 4 (Figure 1.2.2.4.1), also known as the East-Asian and Euro-American lineages respectively [38], are also the most virulent lineages [38,40]. Lineage 2 which includes the Beijing family of strains is predominantly present in Eastern Asia (endemic in China), and also circulates in Russia and South Africa (Figure 1.2.2.4.1).

Different strain families circulate in different areas of South Africa. For example, in the Western Cape, the Beijing, F11, F28 strains were found to drive the TB burden [41], with the Beijing strain dominating amongst drug resistant strains. In contrast, only 2 strains primarily drive the KZN province TB burden, the Beijing and F15/LAM4/KZN (KZN) strains [42]. A large number of resistant TB cases reported in South Africa were observed in the KZN province that saw the Tugela Ferry XDR-TB outbreak [43]. The genotype responsible for this outbreak (KZN) has been since reported to be strongly associated with drug resistance, XDR-TB cases included [42–45].



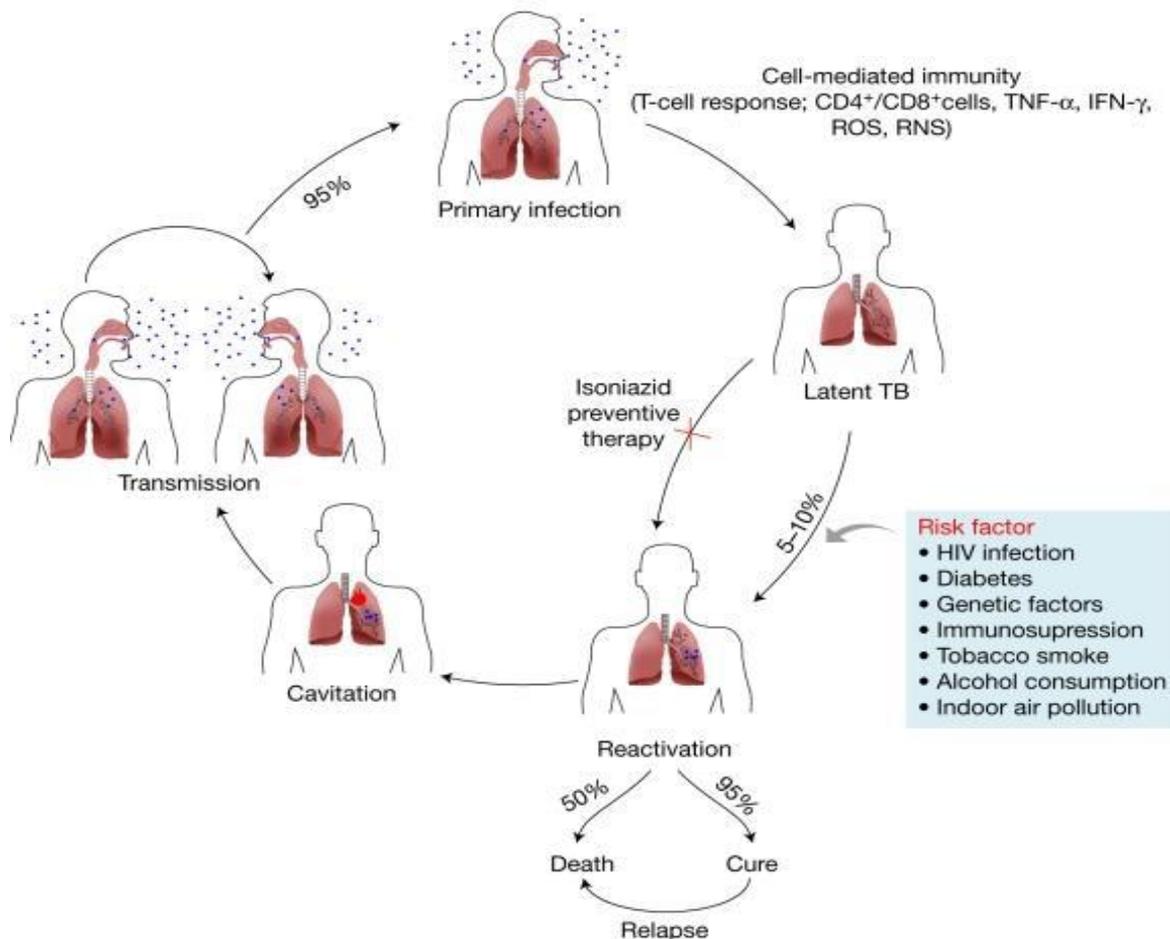
**Figure 1.2.2.4.1: Global Lineages of *Mycobacterium tuberculosis* represented in colour [38].**

### 1.2.3 Pathophysiology of Tuberculosis

*M. tuberculosis* is a rod-shaped, acid fast bacillus, whose cell wall structure is unique, contributing to its virulence. This facultative intracellular pathogen is characterised by the presence of cell wall rich in lipids, like other mycobacteria, and its generation time ranges between 18 to 24 hours, hence it is one of the slow growing mycobacteria [24,46].

Tuberculosis is nearly entirely transmitted by aerosolized droplets that contain infectious *M. tuberculosis* and primary route of infection is the respiratory tract, to the lung [1]. When a person with the pulmonary TB coughs, droplets containing infectious *M. tuberculosis*, are generated, which, when inhaled by an uninfected person, can transmit the infection [1]. The number of bacilli contained in one aerosol droplet ranges between 1 to 400, and the infectious dose (the number of bacilli required to cause infection) is reported to range from 1 and 200 bacilli [24]. The inhaled tubercle bacilli migrate to the alveoli, where they rapidly enter resident alveolar macrophages via phagocytosis, and replicate at the site of infection [1,24,47]. The disease progression thereafter, was described in detail by Dr Arthur Dannenberg using rabbit infection models [47]. There are five stages during infection, namely; the onset; symbiosis; early stages of caseous necrosis; interplay of cell-mediated immunity and tissue

damaging delayed-type hypersensitivity; and liquefaction and cavity formation [48]. As these cellular processes occur in an individual, the development of TB will differ in each patient [47], according to the complex host-pathogen interactions following infection, and the status of the patient's immune system, and this will determine the clinical outcome of infection. Figure 1.2.3.1 provides a schematic summary of the possible outcomes following infection with *M. tuberculosis* [49].



**Figure 1.2.3.1. Pathophysiology of Tuberculosis.** Possible outcomes of infection with *M. tuberculosis*. The clinical manifestations of *M. tuberculosis* infection are either active disease or latent TB infection (LTBI) [49].

Infection with *M. tuberculosis* results in different outcomes ranging from asymptomatic infection known as latent TB infection (LTBI), to fatal disease [1,50]. The host is capable of mounting an effective initial response that will eliminate the bacilli [51]. In the event that the host response is not as effective, about 10% of infected immune competent individuals will develop active, symptomatic disease. However, primary tuberculosis is often asymptomatic as the majority (90%) of infected people develop a LTBI [50], where the bacilli are enclosed but not completely eliminated, and the individuals will not show any signs or symptoms of the disease upon initial infection [47]. This latency can last a

life time, but if an individual's immune system later becomes compromised (Figure 1.2), the risk of disease reactivation increases [47].

### **1.2.3.1 Active Tuberculosis**

Most persons exposed to *M. tuberculosis* maintain a latent state of the infection. The small percentage of the infected community that develops active TB could be explained by the differences in the capability to develop an effective immune response displayed by every individual. The inability of an individual to mount an immune response capable of controlling the initial infection [52] is strongly correlated with the ability to control *M. tuberculosis* infection, resulting in active disease characterised by signs of clinical disease [53]. The infected host has several immunity checkpoints that the bacterium has to pass in order to cause active TB [53], and these include

- (i) avoiding being destroyed by the very early host immune mechanisms,
- (ii) overcoming the initial innate immune control by the host, and presentation to adaptive immunity by host antigen presenting cells (APCs),
- (iii) finally, the bacteria have to overcome the effector mechanisms posed by the subsequent adaptive immunity mounted by the host. Once immune control is lost, the host allows for effective replication of the bacteria [54].

### **1.2.3.2 Immune responses against tuberculosis**

The mammalian immune response to any infective pathogen, *M. tuberculosis* included, consists of the innate and adaptive immune responses. The innate system is a form of natural immunity whereby the immune cells provides the first line of protection against an invading pathogen that the host has never previously encountered. It is facilitated by phagocytes like macrophages and dendritic cells (DCs) that are resident at the site of infection, and are recruited following the establishment of *M. tuberculosis* [55]. Once microbial infection has been sensed, the invading microbe is engulfed, and an inflammatory response is induced. An inflammatory response is necessary in order for the host to control the infection [56]. When strong enough, the innate immune response is capable of eliminating the pathogen. On the other hand, the adaptive immune system is extremely specific, as it is influenced by the immune system's previous interaction with the pathogen, or its immunogenic components (antigens). It is therefore facilitated by specific antigen receptors that are expressed on the surface of T and B lymphocytes, and its immunological memory makes the adaptive response long lasting. Both mediated immunity (Th1) and humoral responses (Th2) are activated during *M. tuberculosis* infection [57,58]. Cytokines such as TNF- $\alpha$ , IL-1 $\beta$ , IL-6, and IL-12 are secreted in the inflammatory Th1 response [58]. In order for host to eliminate invading microbial pathogens efficiently, cooperation of the innate and adaptive immune systems is required [48,56,59].

### ***1.2.3.3 Host and Pathogen interaction***

It is important to understand the clinical characteristics induced by *M. tuberculosis* in its host in order to comprehend its pathogenesis and infection at a molecular level [1]. The host's internal environment is separated from the external by various cell types such as epithelial cells which act as important barriers by lining the mucosal surfaces. Epithelial cells have various functions, including facilitating the transport of ions and acting as the first site of contact for invading pathogens in a host organism [16]. The pathways and outcome of infection within the host depend on the primary interaction between *M. tuberculosis* and the host. Once bacilli reach the alveoli after inhalation, they interact with and are phagocytosed by the resident alveolar macrophages [55], which are the key resident cell population leading the host defence against *M. tuberculosis* [55,60]. Dendritic cells (DCs) and neutrophils, as well as the pulmonary epithelial cells lining the alveoli also take up the bacilli [8,55]. The number of alveolar macrophages present in the alveolar space is small compared to the number of epithelial cells [8]. Therefore, after inhalation, the probability of the bacilli encountering an epithelial cell than a macrophage in the lung is higher [8,9,61]. This initial interaction of *M. tuberculosis* with epithelial cells might result in chemokine secretion [13], which elicits macrophage recruitment to the primary site of infection. Furthermore, the chances of survival within epithelial cells are much higher than in the potential killing environment within the macrophage [62]. Nevertheless, not much attention has been focused on the interaction of *M. tuberculosis* with epithelial cells, while the role of the alveolar macrophages in the pathogenesis of TB has been extensively studied [61]. This is because the bacteria favour alveolar macrophages and other mononuclear phagocytes as hosts, even though they are able to invade the other three cell types [8].

Bermudez and Goodman (1996) were the first investigators to demonstrate that *M. tuberculosis* bacilli binds to and invades alveolar epithelial cells was [8]. The involvement of epithelial cells with host immune has since been of interest and studies have shown that that play a role in inflammatory response by producing chemokines such as IFN- $\gamma$ , MCP-1 and IL-8 after infection with *M. tuberculosis*-infected epithelial [63,64]. Epithelial cells produce collectins surfactant protein A (SP-A) and surfactant protein D (SP-D) that control phagocytosis and opsonisation, as well as increase the adherence of epithelial cells to *M. tuberculosis* [65]. Epithelial cells also produce the glycoprotein fibronectin (FN), which has a role in pathogen opsonisation, and has been shown to enhance the attachment of and direct invasion of epithelial cells by *M. tuberculosis* bacilli [65,66]

During *M. tuberculosis* infection, the bacilli primarily infects macrophages [56], and therefore, they are essential for initiating of an immune response by the host [67]. They internalize the bacteria via a range of multiple cell surface receptors like the mannose receptors (MRs), complement receptors (CRs), and Fc receptors. *M. tuberculosis* contains a cell wall- associated mannosylated glycolipid, (Lipoarabinomannan) LAM, which are able to bind directly to the MRs on macrophages [68]. The

binding of CRs and Fc receptors to the bacilli occurs indirectly [68]. The bacteria then reside within membrane-bound vacuoles inside the macrophages, which become phagosomes. The phagosomes go through maturational procedures that allow lysosomes to fuse with them, exposing the internalized microbes to substances that are cytotoxic, like acid and lytic enzymes [69]. However, *M. tuberculosis* is known to modify the maturation of the phagosomal compartment, and subvert the killing mechanisms of the phagosome [69,70], so as to allow for its own intracellular survival and replication within the macrophage [71,72].

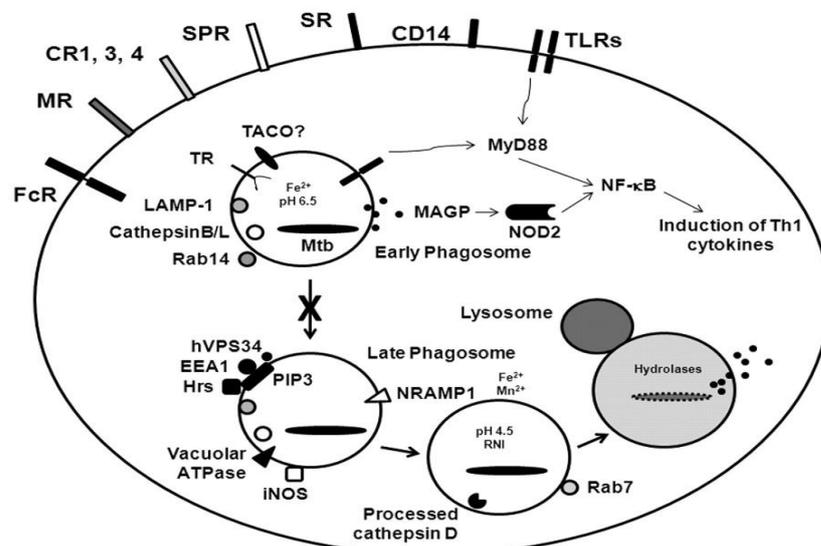
*M. tuberculosis* also infect DCs [73,74] by binding the Dendritic Cell-Specific ICAM-3-Grabbing Non-integrin (DC-SIGN) receptor to its mannose-capped cell-wall component lipoarabinomannan (ManLAM). DCs also possess other antigen receptors like the Fc receptors and Mannose receptors (MRs). These bone marrow immune cells [73] have a major role in the initiation of the early immune response to *M. tuberculosis* infection [51,74]. DCs are the key antigen-presenting cells, better than macrophages [75], and have the unique ability to activate naïve T-cells with specific *M. tuberculosis* antigens [73,74,76], thus they are known as “professional antigen presenting cells” [51]. The expression receptors on DCs differs with their developmental stage. Immature antigen capturing DCs express abundant MRs and DC-SIGN on their surface [77]. These interact with *M. tuberculosis* ligands during phagocytosis and initiate processes that result in the activation and maturation of DCs into APCs [51], and their migration to lymph nodes. The expression of MRs and DC-SIGN decreases with maturation [73,78], and their migration is linked to their activation which ends with the interaction and activation of T-cells [51] when they present the mycobacterial antigens. This ability of DCs cells to express different receptors at different developmental stages allows for the cells to have different levels and forms of activation. This determines the degree and type of immune response induced by *M. tuberculosis* in each individual, which in turn determines whether the infection is controlled, or leads to disease progression.

#### **1.2.3.4 Pattern recognition receptors in *Mycobacterium tuberculosis* infection**

The main defense mechanism of the host after inhalation of *M. tuberculosis* is phagocytosis by resident phagocytic cells. In order to invade and overcome the early innate immunity of the host, *M. tuberculosis* or its components have to be recognized as an invading pathogen [79], prior to phagocytosis. The bacterium possesses various microbial structures known as pathogen-associated molecular patterns (PAMPs), and these initiate the innate immune response by facilitating the recognition of invading cells [80]. Host cells (mainly immune cells) have several pattern recognition receptors (PRRs) that recognize the pathogen’s PAMPs [80], and engulf the bacilli, initiating infection. These PRRs have been associated with *M. tuberculosis* infection [81], and include Toll-like receptors (TLRs), DC-SIGN, Dectin-1, nucleotide-binding oligomerization domain-containing protein 2 (NOD2), C-type lectin receptors (CLRs), and MRs [82,83]. The type of receptor the host uses to internalize *M. tuberculosis*

influences the host cellular response [83], and because the pathogen possesses a number of these receptors, it does not have an exact route of entry into the host phagocytic cells [83]. These receptors mediate cytokine production and other effector molecules that are involved in mounting an effective immune response on antigen presenting cells (APCs). However, the route of entry can also influence *M. tuberculosis* survival within host cells, as some of these receptors might favour the infection of the bacilli [83]. Thus, this PRR- dependent entry of the bacilli into host cells is an important determinant of the fate of infection.

A number of other PRRs expressed on the surface of immune cells (Figure 1.2.2.2) can facilitate phagocytosis of *M. tuberculosis*, besides TLRs [83]. CRs, namely, CR1, CR3, and CR4 are other examples of membrane proteins that are expressed on host phagocytic cells, particularly macrophages. These facilitate *M. tuberculosis* internalization [83] by interacting with complement components that are important for opsonisation of *M. tuberculosis* in preparation of phagocytosis [83]. Phagocytosis of *M. tuberculosis* by monocytes was evidently reduced *in vitro* after the addition of monoclonal antibodies against CR3, showing its significance relative to the other CRs [84]. However, *M. tuberculosis* entrance into host cells via CRs is also an advantageous route for its and pathogenesis. CR3 binding to *M. tuberculosis* was shown to prevent the maturation of phagosomes by inhibiting establishment of respiratory bursts [83,84]. MRs on the other hand, assist the uptake of non-opsonised bacteria by macrophages [83]. They are transmembrane proteins that specifically bind to the sugar mannose, that is found on the outer wall of pathogens. MRs were demonstrated to be associated only with virulent strains of *M. tuberculosis*[85] [83,85], where they interact with the LAM that is specifically capped with a mannose residue, known as the ManLAM [83]. MRs and CRs have a minor part in the interaction of *M. tuberculosis* with DCs, with DC-SIGN playing the major role of mediating the recognition and internalization of *M. tuberculosis* by DCs [51,83]. DC-SIGN also recognizes carbohydrates (ManLAM), and are mostly expressed on DCs [51,83].



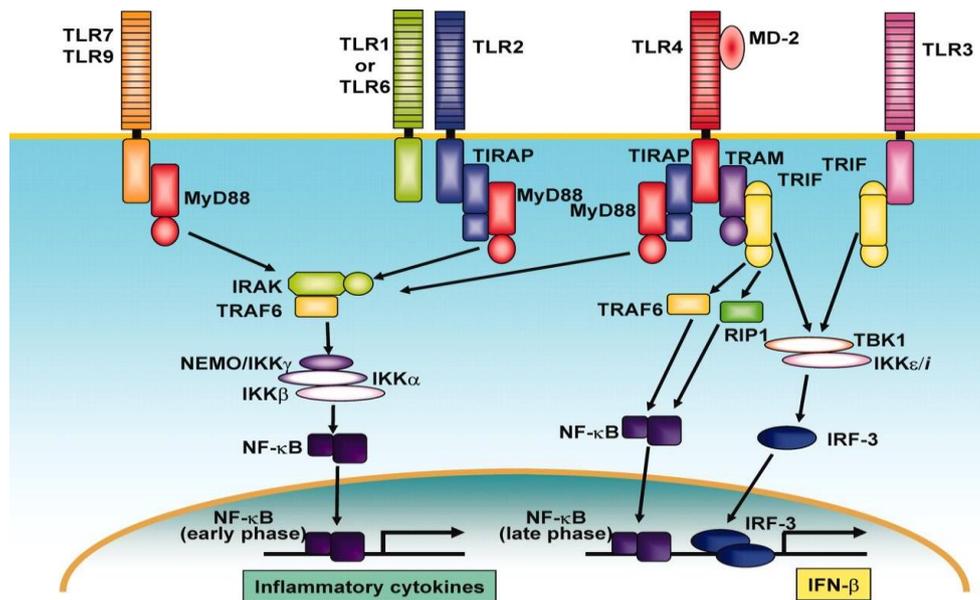
**Figure 1.2.3.4.1. Cell biology of the recognition of *M. tuberculosis* through pattern recognition receptors (PRRs).** The innate immune system detects the presence of pathogens through different PRRs that are germ line-encoded. Specific PRRs are expressed on phagocytic cells and recognize the various PAMPs of *M. tuberculosis*, allowing the bacilli to enter into the cells, initiating and coordinating the immune response. PRRs are expressed in different compartments of the cell like the cell surface, lysosome, endosome, or cytoplasm. The 2 families of PRRs shown here are the transmembrane proteins which include the Fc receptor (FcR), mannose receptor (MR), complement receptors (CR1, CR3, CR4), surfactant protein receptors (SPR), scavenger receptors (SR), CD14, and Toll-like receptors (TLRs). The second family is the cytoplasmic proteins (not shown here), such as the NOD-like receptors (NLRs) and the Retinoic acid-inducible gene (RIG)-I-like receptors (RLRs) [86].

#### *1.2.3.4.1 Membrane bound PRRs - Toll-like Receptors*

Toll-like receptor (TLR) signaling is crucial for immunity against numerous intracellular pathogens [87], and TLRs are believed to be the most significant PRRs during *M. tuberculosis* infection [88]. They are expressed either on the cell surface, for example, TLR-2 and TLR-4, or they are expressed inside the compartments of the cell (TLR-8 and TLR-9) [59,87], and stimulate an intracellular signalling cascade in the host cells. This mediates the secretion of various pro-inflammatory cytokines [82,88] like interferon (IFN)- $\gamma$  and tumour necrosis factor (TNF)- $\alpha$ , which are crucial for eliciting the adaptive immune system, and other anti-bacterial effector molecules, thereby stopping bacterial growth, and clearing infecting *M. tuberculosis* [83,87].

TLR2 can function alone, or form heterodimers with TLR1 or TLR6 (Figure 1.2.2.3) to produce a strong pro-inflammatory response after interacting with mycobacterial cell wall components like the 19 KDa lipoprotein, Lipoarabinomannan (LAM), Lipomannan (LM) and Phosphatidyl-myo-inositol mannoside (PIM) [87,89]. The robust pro-inflammatory response elicited by the host immune system against *M. tuberculosis* infection via TLR2 signalling was determined to be mediated by myeloid differentiation primary-response protein 88 (MyD88), an adaptor protein [80,90], from studies that used mouse macrophages lacking TLR2 and MyD88. Furthermore, MyD88- deficient mice or mice lacking both TLR2 and TLR4 or TLR9 were found to have uncontrolled bacterial growth, and more susceptible to infection [24]. TLRs trigger and mount the recruitment of MyD88, which induces cytokine production via the activation and translocation of transcription factors such as the nuclear factor kappa-light chain-enhancer of activated B-cell (NF $\kappa$ B) signalling cascade (Figure 1.2.2.3) [59]. NF $\kappa$ B activation leads to the transcription of genes that are involved in the activation of the innate immune defences, primarily the production of pro-inflammatory cytokines like TNF, IL1 $\beta$ , and IL-12 and nitric oxide [79,80]. Interestingly, recent *in vivo* studies suggested that the TLR signalling pathway may possibly be one of the many pathways exploited by *M. tuberculosis* [87,90]. Studies have shown that persistent TLR2 signalling due to the over-stimulation of the TLR2 pathway may also be beneficial for *M. tuberculosis*

by decreasing the antigen presentation ability of infected macrophages [87,90]. Thus, an equilibrium in the TLR signalling network must be reached in order to control inflammation and prevent an inflated inflammatory response that may be damaging for the host [87,91]. Cytokines and chemokines are also important for recruiting and activating more phagocytes and immune cells [63,81]. Hence, TLRs also function as an important link connecting the innate immune response to the adaptive immune response against infection with *M. tuberculosis* [83].



**Figure 1.2.3.4.1.1: Toll-Like Receptor Signaling Pathways.** The best-characterized PRR family is the TLR family. Different bacterial components interact with different TLRs [90], but most important in TB immunity is the interaction of TLR2, TLR4, TLR1/6 and TLR9 [90,92]. TLR2 recognizes its ligands and functions alone, or forms heterodimers with either TLR1 or TLR6. This pathway recognizes the 19 KDa lipoprotein, Lipoarabinomannan (LAM), Lipomannan (LM) and Phosphatidyl-myo-inositol mannoside (PIM). Lipopolysaccharide (LPS) is recognized on the cell surface by 2 sets of TLR4/MD2 complexes. Both pathways result in the strong induction of pro-inflammatory cytokine genes [93].

#### 1.2.3.4.2 Cytosolic PRRs - Nucleotide-binding Oligomerization Domain (NOD)-like receptors (NLRs)

NLRs like NOD-1 and NOD-2 also have a role in host immune defence against *M. tuberculosis* infection. They are located in the cytoplasm, therefore, they are intracellular receptors that regulate stimulation of pro-inflammatory cytokine response by *M. tuberculosis*, and control the intracellular growth of the organism [79,94]. NOD2 contains a caspase activation and recruitment domain (CARD) which allows it to form interactions that lead to the recruitment of NF-κB after recognizing bacterial peptidoglycans like Mycolylarabinogalactan-peptidoglycan (MAGP) from the *M. tuberculosis* cell wall

(Figure 1.2.2.2.). Therefore, both TLR and NOD2 signalling pathways use the transcription factor NF- $\kappa$ B to induce the transcription of cytokine genes.

NODs synergize with TLRs to activate the pro-inflammatory cytokine production [92]. NOD2 can cooperate with other signalling pathways like TLRs to enhance pro-inflammation [95]. This occurs through cytosolic vesicular fusion mechanisms controlled in the course of *M. tuberculosis* phagocytosis, as it appears to be associated with intracellular vesicles [96] like the early phagosome (Figure 1.2.2.2.). NOD2-deficient mice IL-1R-deficient mice showed impaired production of pro-inflammatory cytokines and nitric oxide increased susceptibility to infection when infected with *M. tuberculosis*. However, the exact mechanisms, and the capability of NOD-2 associated with networks involved in phagocytic receptor trafficking has not yet been established. The intersection of NOD-2 with these signalling networks still needs to be explored. Several other NLRs have also been found to function in immunity via the formation of an inflammasome, which is a multi-protein complex that plays a vital role in the pathogenesis of chronic disorders [86].

#### **1.2.4 Adhesins**

Adhesion to host cells is necessary for colonizing the host and its immune response [72], and is therefore an essential virulence factor of most bacterial pathogens such as *M. tuberculosis*. The fundamental processes that determine bacterial attachment to host tissue surfaces are mediated by microbial adherence molecules known as adhesins [7]. Adhesins are typically surface-exposed molecules [97] that facilitate cell-to-cell interactions and thus are termed intercellular adhesion molecules (ICAM). The other group of adhesins, substrate adhesion molecules (SAM), facilitate cell-to- extracellular matrix (ECM) adherence [97]. Both bacterial and host cell determinants are involved in the specific interactions that occur between the pathogen and the host which lead to the invasion of host cells by microbes [8], and adhesins are key players that have a vital role in this complex process [8,98]. There is little information on the adhesins associated with the complex interaction that specifically takes place between *M. tuberculosis* and the human host [7]. *M. tuberculosis* has been shown to adhere to and invade the alveolar cells (A549) more efficiently compared to bronchial cells (BBM). This may possibly be due to the presence of multiple adhesins (Table 1.2.3.5.1) that interact with receptors present on the alveolar epithelial cells, since bacterial adhesion to host cell surfaces is specific [9].

It is known that *M. tuberculosis* has a number of adhesins on its cell surface like Heparin binding hemagglutinin adhesin (HbhA), a 19 kDa lipoprotein antigen, Apa, Malate Synthase (MS), and other molecules that are potential adhesins (Table 1.2.3.5.1) which may facilitate entry into epithelial cells and macrophages [66,99]. Of the few adhesins that have been described for *M. tuberculosis*, the HBHA surface-exposed protein is the most described. Most importantly, it is associated with dissemination of *M. tuberculosis* to extra-pulmonary infection sites [100]. The 19 kDa lipoprotein antigen binds to

mannose receptors on macrophages, promoting phagocytosis [101]. MS binds to laminin and fibronectin, and therefore it has adhesin during *M. tuberculosis* infection [102]. Recently, *M. tuberculosis* has been shown to produce pili during human infection [10,13], and this has been proven to have adherence properties towards host cells [12,14,15]. Further insights on the mechanisms used by *M. tuberculosis* to adhere to, and infect host cells may suggest a novel view on rapid diagnostics, drug design and vaccine production.

**Table 1.2.4.1: A list of the currently known adhesins of *Mycobacterium tuberculosis* [103]**

<b>Adhesin</b>	<b>Gene (s)</b>	<b>Mediates adhesion to</b>
19-kDa antigen	<i>Rv3763</i>	Monocytes and macrophages
Alanine- and proline-rich antigen (Apa)	<i>Rv1860</i>	Pulmonary surfactant protein A and macrophages
Antigen 85 complex	<i>Rv0129c, Rv1886c, Rv3803c, and Rv3804c</i>	Fibronectin and macrophages
Cpn60.2 molecular chaperone	<i>Rv0440</i>	Macrophages
Curli pili	<i>Rv3312A</i>	Laminin, <i>M. tuberculosis</i> , macrophages, and epithelial cells
DnaK molecular chaperone	<i>Rv0350</i>	Macrophages
Early secreted antigen ESAT-6	<i>Rv3875</i>	Laminin
Glutamine synthetase A1	<i>Rv2220</i>	Fibronectin
Glyceraldehyde-3-phosphate dehydrogenase	<i>Rv1436</i>	Possibly fibronectin (as occurs in group A streptococci)
Heparin-binding hemagglutinin adhesin	<i>Rv0475</i>	<i>M. tuberculosis</i> and epithelial cells
Laminin-binding protein	<i>Rv2986c</i>	Laminin
L,D-transpeptidase	<i>Rv0309</i>	Fibronectin and laminin
Malate synthase	<i>Rv1837c</i>	Fibronectin, laminin, and epithelial cells
Membrane protein	<i>Rv2599</i>	Collagen, fibronectin, and laminin
<i>Mycobacterium</i> cell entry-1 protein	<i>Rv0169</i>	Epithelial cells
N-acetylmuramoyl-L-alanine amidase	<i>Rv3717</i>	Fibronectin and laminin
PE-PGRS proteins	<i>Rv1759c</i> <i>Rv1818c</i>	Fibronectin <i>M. tuberculosis</i> and macrophages
Protein kinase D	<i>Rv0931c</i>	Brain endothelia and laminin
PstS-1 (38-kDa antigen)	<i>Rv0934</i>	Macrophages
Type IV pili	<i>Rv3654c-Rv3660c</i>	Possibly macrophages and epithelial cells

#### **1.2.4.1 Curli pili and *M. tuberculosis* pili (MTP)**

Like many pathogenic bacteria, *M. tuberculosis* produces polymeric cell-surface adhesive organelles called fimbriae, or pili [10], that assist the initial attachment and subsequent successful colonization of eukaryotic cells [104]. Bacterial pili are polymeric, hydrophobic, proteinaceous structures generally

composed of a major repeating subunit called pilin and, in some cases, a minor tip-associated adhesin subunit [105]. They are involved in many virulence-associated biological functions, such as adherence, cell aggregation, colonization of mucosal surfaces, host cell adhesion and invasion, and biofilm formation [105,106]. Bacterial pili have also been reported to be potent inducers of the host's inflammatory response by inducing host cytokine/chemokine production [16]. They are viewed as virulence factors because of their key role in bacterial pathogenesis, therefore, are important targets for vaccine and drug development [10].

In *M. tuberculosis* infection, production of certain cytokines and chemokines has been shown to be elicited in different types of host cells including microglia and macrophages. In microglia, the following cytokines are elicited, TNF- $\alpha$ , IL-1 $\beta$ , and IL-6, and the following chemokines CCL2, CCL5, CXCL8, and CXCL10 [107]. Infected alveolar epithelial cells have also been shown to produce cytokines such as IL-6, IL-8, IFN- $\gamma$  and TNF- $\alpha$  [18,63,108,109].

Alteri (2005) showed evidence using transmission electron microscopy (TEM) that *M. tuberculosis* produces a compact fibrillar meshwork made of thin coiled, grouped fibres similar to pili, and named these *Mycobacterium tuberculosis* curli pili or MTP [10,13]. Further scanning electron microscopy revealed that *M. tuberculosis* possesses the ability to produce 2 different pili types, the second being a type IVB pili [13]. The pilin subunits of MTP are encoded by the Rv3312A (*mtp*) gene in *M. tuberculosis* [10,13], which is located between the *Rv3312c* gene whose function is still unknown, and the *add*, *deoA* and *cdd* whose function is involved in intermediary metabolism [110]. Alteri demonstrated the presence of IgG antibodies specific to MTP in the blood serum of patients suffering from TB, which suggested that MTP are produced *in vivo* during human infection [10,13]. *In vitro* studies with isolated MTP and the extracellular matrix (ECM) protein, laminin, showed both bind to each other, suggesting that MTP have adherence properties, and therefore may be essential for host colonization by *M. tuberculosis* [10,13]. Further *in vitro* studies conducted by Ramsugit and Pillay (2014), using a *mtp*-deficient strain [11,15], its wild-type and complemented strains, confirmed the findings by Alteri [15]. This study showed that that MTP play a significant role in the adhesion to and invasion of macrophages (THP-1 cells) [15]. MTP were also shown to be involved in biofilm production *in vitro* [11], as well as in adhesion and invasion [15,18] of epithelial cells [10,13,18]. A study by Naidoo *et al* (2014) showed that MTP are distinctive to members of the MTBC and that the *mtp* gene is extremely conserved among the clinical strains of *M. tuberculosis*, [111]. Taken together, these findings suggest that MTP could be considered as a diagnostic, drug and/or vaccine candidate, thus, the importance of MTP during *in vivo* TB infection in a mouse model will be interrogated further in this study.

### 1.2.5 Mouse models

Generally, most bacteria are hardly ever found to exist in environments similar to those in the cultures of a laboratory, i.e. in favourable conditions that permit for exponential growth for extended periods of time. An example of those unfavourable conditions is the deprivation of Oxygen, but it has been however, shown that the bacilli of *M. tuberculosis* stay viable and virulent for long periods of time in the absence of Oxygen *in vitro* [112]. Furthermore, results obtained by Ashiru et al showed that the bacilli are also able to invade epithelial cells when they are grown under oxygen deprivation [113]. This is important as these environments deprived of Oxygen are common in the human body, especially in the environments where TB takes its course, like macrophages and granuloma. Granuloma formation is a typical feature of latent TB infection, conditions of low redox potential which results in Oxygen deprivation in the environment [114]. This makes *in vivo* experiments important in TB research. The correlation between *in vitro* findings and *in vivo* conditions is not clear, but important information is revealed from laboratory studies.

Mice, rabbits and guinea pigs are productively used as models for *M. tuberculosis* studies. All 3 models can be infected with different strains of the bacilli, but they confer different susceptibilities to TB. Generally, the guinea pig is used in studies that define the progressive pathology of tuberculosis, whilst the mouse model is used in host immune response studies [115]. The mouse model is the most widely used for *in vivo* TB research due to the ease of handling and their cost effectiveness compared to other *in vivo* models [56]. They are efficiently infected by different *M.tuberculosis* strains and are usually used to explain and describe the host immune response to *Mycobacterium tuberculosis* infection [115]. The most appropriate route of infection used in mouse infections is the aerosol route of infection as it is more relevant to the human TB disease [115]. Furthermore, the aerosol route of infection is currently the most reproducible technique, with the least degree of variability, compared to other techniques used like the intravenous, intra-nasal or intra-tracheal inoculation techniques. The aerosol infection route involves the generation of a cloud of aerosols, containing the bacilli in very small droplet nuclei, that the mice inhale and deposit in the alveolar areas of the lungs, causing disease in the animal [115]. The mouse lung is therefore often used for many studies, including transcriptomic studies, for these reasons [20].

### 1.2.6 Gene expression

Pathogens express certain gene products that allow them to infect and cause disease upon encountering a host. Various types of genes are expressed that enable the pathogen to grow within the host, counteract the host defences and persist. Different signals induce the alteration of gene expression which results in up- and/or regulation of specific genes [114] in both the host and the pathogen. cDNA libraries synthesized from the RNA are used to analyse transcripts via sequencing by hybridization, for example,

microarrays. In the past decade, Next-generation RNA sequencing (RNA-Seq) of microbes and their host cells has developed into a standard technique for quantifying and annotating microbial transcripts. Understanding and quantifying the changing transcript levels and the dynamics of their gene expression dynamics is important as pathogens react to host conditions in order to determine which genes are up- or down-regulated upon infection, and for understanding host-pathogen interactions of *M. tuberculosis*.

Recently, RNA sequencing methods (that are centered on ultra-high throughput sequencing of total RNA and methodical counts of all transcripts that are expressed) [116] have become more popular as they provide the possibility to overcome many of the restrictions related to the microarray technology [21]. In contrast to the hybridization-based method, RNA-Seq methods are more sensitive, and enable strand-specific identification of common and novel transcripts in an unbiased manner, and allow for more information to be deduced from the RNA. It also allows the positive identification of both lowly expressed and highly expressed genes in a single RNA-Seq experiment [21], enabling a platform for unbiased and fully qualitative and quantitative transcriptomic profiles of host cells to be interrogated, following mycobacterial infection [116].

Global changes in transcriptional response have allowed some enlightening on the molecular events that occur within the host during the establishment of a *M. tuberculosis* infection [117,116,118]. Several studies have described the transcriptional responses to *M. tuberculosis* infection [119,120], and recently global transcriptional changes at different time points have been analysed in the host [121]. Lung gene expression profiles of mice infected with *M. tuberculosis* or vaccinated with BCG prior to *M. tuberculosis* infection have indicated differences in naturally and vaccine induced immune response signatures [122]. Differential whole lung gene expression has identified signature profiles during *M. tuberculosis* disease progression [122] for example those that define different signalling pathways and immunological responses to *M. tuberculosis* infection [120,121].

Therefore, in this study, we have used RNA-Seq to examine the transcriptome of mice lungs following a 14 day *in vivo* infection with *M. tuberculosis* to gain novel understandings of the transcriptional changes and cellular pathways induced by MTP during the early stages of infection.

### **1.3 Significance of work**

Pili are present on the surface of bacterial cells, including that of *M. tuberculosis* [105,106,123]. They have been demonstrated to have roles in bacterial adherence and host colonization, and have been shown to be immunogens, and thus pili make for ideal vaccine candidates [123]. This study will add to the body of knowledge of MTP, exposing its role in pathogenesis and host immune response.

## 1.4 Hypothesis

MTP plays a role in the specific host-pathogen interactions that follow *M. tuberculosis* infection, and thus, the host immune response

## 1.5 Aim

To evaluate the role of *M. tuberculosis* curli pili (MTP) in host gene regulation in a mouse model infected with  $\Delta mtp$ -mutant, *mtp*-complemented and wild-type strains of the V9124 strain of the F15/LAM4/KZN genotype of *M. tuberculosis*.

## 1.6 Objectives

- (i) To conduct *in vitro* growth assays of the WT, complemented and  $\Delta mtp$ -mutant strains by performing colony-forming units (CFU) over a 28-day period.
- (ii) To determine the viability of *M. tuberculosis*  $\Delta mtp$ -mutant strain compared to the complemented and the WT in mice lungs and spleen by CFU enumeration during disease progression until 28 days.
- (iii) To assess visible signs of disease on infected lungs after 28 days of infection by photographic imagery using a NIKON 200 camera
- (iv) To evaluate the lung gene expression at D14 post *M. tuberculosis* infection of the  $\Delta mtp$ -mutant infected lungs compared to WT strain infected lungs by using whole transcriptome analysis.

## 1.7 Study Design

Female BALB/c mice were infected with the  $\Delta mtp$ -mutant, complemented and WT strains. Bacterial growth kinetics were carried out to determine the effect of the gene on growth by observing *in vitro* CFU assays, and *in vivo* CFUs from *M. tuberculosis* infected tissues. The transcriptional profiles of lungs infected with the  $\Delta mtp$ -mutant, and WT strains for fourteen days were compared using various bioinformatics tools to further analyse the effect of *mtp*.

## 1.8 Dissertation Layout

This dissertation is composed of four chapters, in the traditional format as accepted by the University of KwaZulu-Natal. The introduction and literature reviews are included in the first chapter. In Chapter two, the materials and methods used for the experimental work involved in this study are described in detail. Chapter three describes the results obtained, and Chapter four discusses the findings, challenges faced, conclusions and recommendations for future studies.

## **CHAPTER 2: MATERIALS AND METHODS**

## **2.1 Ethics Approval**

The study was ethically approved by the Animal Research Ethics Committee (AREC), University of KwaZulu-Natal (reference number: 053/15/animal).

## **2.2 Growth of *Mycobacterium tuberculosis* strains.**

A MTP-deficient (*Δmtp*-mutant), *mtp*-complemented and Wild Type (WT) strains of the clinical F15/LAM4/KZN genotype of *M. tuberculosis* [11], that were confirmed by PCR were cultured in 7H9 broth (BD, Difco Laboratories, USA) (supplemented with 10% (v/v) Oleic Albumin Dextrose Catalase (OADC) (BD, Difco Laboratories, USA), 0.05% (v/v) Tween-80 and 0.5% (v/v) Glycerol (Sigma-Aldrich, USA) to an optical density (OD<sub>600nm</sub>) of 0.9–1.2 [124,125]. Stock cultures were subsequently stored in 50% supplemented Middlebrook 7H9 broth (BD, Difco Laboratories, USA) and 50% Glycerol (Sigma-Aldrich, USA) at -70°C [124].

## **2.3 *In vitro* growth studies.**

Frozen stocks of the WT, *Δmtp*-mutant and *mtp*-complemented strains were inoculated in triplicate into 100 mL Middlebrook 7H9 liquid broth, supplemented as previously mentioned, and incubated with gentle shaking at 37°C until reaching an OD<sub>600</sub> of 1.6–1.8. The cultures were centrifuged at 3000 rpm for 10 minutes and thereafter, the pellet was washed three times in 8 mL of 10 % Phosphate Buffered Saline (PBS) with 0.05% Tween-80. After the final wash, the pellet was re-suspended in 1mL of supplemented 7H9 broth, and vortexed to achieve a homogeneous suspension. The cultures were back diluted to an OD of 0.015 in 25mL of supplemented 7H9 broth in triplicate [125] and incubated in a shaking incubator at 37°C for 28 days and colony forming units (CFUs) were quantified in triplicate on Middlebrook 7H11 agar plates supplemented with 10 % OADC and 0.5 % glycerol, from 10-fold serial dilutions at 4-day intervals. The Middlebrook 7H11 agar plates were read after incubation at 37°C for three weeks.

## **2.4 *In vivo* experiments**

### **2.4.1 *Animals***

Female Balb/c mice aged between 6-8 weeks were purchased from the KwaZulu-Natal Research Institute for Tuberculosis and HIV/AIDS (K-RITH). The BALB/c mice were bred at the Biomedical Research Unit (BRU), University of KwaZulu-Natal (UKZN). All experiments involving mice were conducted under humane conditions specified by animal ethics guidelines, in Class II bio-safety cabinets in the bio-safety level 3 (BSL-3 facility) at the BRU, and K-RITH. All subsequent experiments on harvested organs were conducted in the TB laboratory, Medical Microbiology and Infection Control, Nelson R Mandela School of Medicine, UKZN.

#### **2.4.2 Mice Aerosol infection (Day 0) and inoculum preparation.**

A Glas-Col Middlebrook Inhalation Exposure System was used to infect the female BALB/c mice with the  $\Delta mtp$ -mutant, complemented and WT strains of *M. tuberculosis* via the aerosol route. The mouse infection inoculum was prepared by culturing 100  $\mu$ L of frozen stocks of each strain in supplemented 7H9 broth (BD, Difco Laboratories, USA) to an OD<sub>600</sub> of 1-1.2 [115], and a 1 in 100 dilution was performed to make the required dose. Colony forming units (CFU/mL) were enumerated by plating out serial dilutions of the inoculum onto 7H11 agar plates that were incubated at 37°C for about 3 weeks, to determine the infection dose. A total of 95 mice i.e. 5 per strain (WT,  $\Delta mtp$ -mutant and complemented) for the 6 time points of sacrifice (day 1; 2; 7; 14; 21; 28) were infected to ensure variability and greater accuracy of results.

#### **2.4.3 Organ harvesting and storage**

Five mice infected with each of the 3 strains were anaesthetized with Isofor and sacrificed by cardiac puncture at the following intervals: 24 hours, 48 hours, 7, 14, 21 and 28 days. Blood was collected from each mouse and stored in Serum separator tubes for future cytokine analysis. The spleens and lungs which were removed under aseptic conditions, were split in half for CFU enumeration to determine the number of infecting mycobacteria at each interval, and RNA extraction. For CFU enumeration, half the organs were transferred to an O-ring tube containing 10% Phosphate PBS with 0.05% Tween 80 and 2um glass beads, and were immediately homogenized using the MagNa Lyser (Roche Molecular Diagnostics, Rotkreuz, Switzerland) and stored on ice. For RNA extraction, the other half of the organs was transferred to an O-ring tube containing TRIzol solution, and immediately snap frozen on dry ice. [126].

##### **2.4.3.1 Gross organ morphology**

Infected lungs were photographed 28 days post-infection using a NIKON 200 camera to assess visible signs of disease, and any differences in size among the organs from mice infected with the different strains.

#### **2.4.4 Determining Colony Forming Units (CFUs)**

All organs were homogenized at 7000rpm for 90s using the MagNa Lyser and homogenates were plated onto Blood and MacConkey Agar plates to detect for contamination. CFU enumeration to determine the number of infecting mycobacteria were performed at each interval, by plating out 100ul of serial dilutions onto Middlebrook 7H11 agar (BD, Difco Laboratories, USA) supplemented with 10% OADC (BD, Difco Laboratories, USA), 0.5% Glycerol and PACT (200.000 units/L of Polymixin B (Sigma-Aldrich, USA); 20 mg/L of Amphotericin B (Sigma-Aldrich, USA); 100 mg/L of Carbenicilin (Sigma-

Aldrich, USA) and 20 mg/L of Trimethoprim (Sigma-Aldrich, USA) antibiotics in 65mm petri dishes. The plates were left to dry and thereafter incubated at 37°C in 5% CO<sub>2</sub> for 21 to 28 days.

#### 2.4.5 *Extraction of RNA*

Extraction of RNA was performed using a combination of the TRIzol reagent (Sigma-Aldrich, USA), and the RNeasy kit (Qiagen, South Africa) protocol.

##### 2.4.5.1 TriZol extraction

Briefly, lung tissue was disrupted and homogenized in the TRIzol reagent (Sigma-Aldrich, USA) it was stored in (to minimize loss of RNA), using a sterile glass homogenizer on ice. Each organ was homogenized with its own homogenizer to prevent cross contamination. To allow for the complete dissociation of nucleoprotein complexes, the lysate was incubated for 5mins at room temperature (RT). Thereafter, 500 µL of Chloroform (Sigma-Aldrich, USA) was added to the tubes, vigorously inverted for 15 secs and incubated at room temperature for 2-3 min. After centrifugation for 15 min at 15 000 rpm at 4°C, the top aqueous layer was carefully transferred into a sterile 1.5 mL Eppendorf tube. An equal volume of 70% cold ethanol (Sigma-Aldrich, USA) was added, and mixed immediately by pipetting.

##### 2.4.5.2 RNeasy purification

The sample was transferred to the RNeasy spin column, washed once with 700 µl of Buffer RW1 and twice with 500 µl of Buffer RPE by centrifuging for 15 s at 12 000 rpm. The RNA was eluted from the RNeasy spin column into a new 1.5 ml Eppendorf tube in 60 µL of RNase-free water by centrifugation for 1 min at 12 000 rpm. The RNA was quantified using the Nanodrop (Thermo Scientific, South Africa), and the quality was checked by electrophoresis in a 3-(N-morpholino)propanesulfonic acid (MOPS) gel electrophoresis (Appendix 2). Aliquots of 25 uL of RNA were stored at -70C. The integrity of the RNA was further analysed in the Agilent 2100 Bioanalyzer (Johns Hopkins Deep Sequencing & Microarray Core Facility, USA), prior to sequencing.

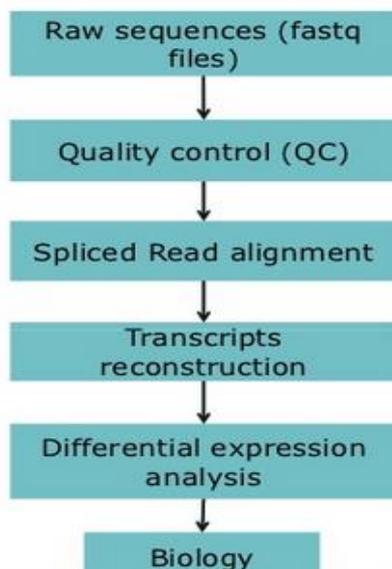
#### 2.4.6 *RNA Sequencing*

Sequencing of the extracted RNA was outsourced to the Johns Hopkins Deep Sequencing and Microarray Core Facility. RNA samples at day 14 from 2 biological replicates with a RNA integrity number (RIN) equal to or above 7.9 were used to prepare single cDNA libraries with the Epicenter ScriptSeq Complete Epidemiology RNA-Seq kit (Illumina, CA, USA) following the manufacturer's procedure. A maximum of 500 ng total RNA was treated with the Ribo Zero kit (Illumina, CA, USA) to remove all rRNAs in a 20 uL reaction. Of this, 9 uL was used to make the cDNA libraries with the ScriptSeq v2 RNA-Seq Library Preparation Kit (Illumina, CA, USA). The library products were used

to sequence 100 – 150 million paired 100bp reads, using the TruSeq Cluster Kits (Illumina, CA, USA) on the Illumina HiSeq 2000, in one sequencing lane per sample since the RNA contained both mammalian and bacterial transcriptomes. The duplicate RNA samples extracted from uninfected lungs as controls were multiplexed in a single flow cell lane. The resulting sequencing reads were provided via BaseSpace (Illumina) as fastq format.

#### 2.4.7 Bioinformatics and Data Analysis

For the CFU assays, each dilution was plated in triplicate for each mouse. For the RNA Sequencing, library preparation was performed for mice infected by each strain, using RNA that was extracted at two independent times. Statistical analysis was performed with SSPS Software 23.0 (IBM) and Microsoft Excel, using a Student's t-test and One-way analysis of variance (ANOVA) from SPSS, with a  $p$ -value  $\leq 0.05$  and a confidence level of 95% being considered statistically significant for growth and organ comparisons. The Illumina fastq sequencing raw reads were analysed by various Bioinformatics tools. All sequence manipulations were performed using in-house Linux shell scripts (Appendix 3), and a RNA-Seq analysis workflow was generated followed (Figure 2.4.7.1). Data visualization was carried out on the R platform with Cumberbund and plotting packages ggplot2 installed.



**Figure 2.4.7.1 Summary of the RNA-Seq analysis work flow that was followed.**

##### 2.4.7.1 Raw data pre-processing - Quality assessment

Pre-processing of the reads included the assessment for quality control (QC) using the FastQC tool [<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>], (version 0.11.3; Babraham Bioinformatics, Cambridge, UK), to perform the following QC checks: per base sequence quality, per base sequence content, sequence length duplication, base GC content, sequence duplication levels, sequence quality scores, sequence GC content, and overrepresented sequences, base N content and Kmer content [127].

#### 2.4.7.2 Alignment of sequences to the mouse reference genome

The principle of mapping of the RNA-Seq reads is to find matches, on the known reference genome for the sequences of the sampled short RNA-Seq reads. Quality controlled sequence reads were mapped to the UCSC mouse reference genome, *Mus musculus* (mm10), using Tophat (version tophat-2.1.0) (<http://tophat.cbcb.umd.edu>) together with Bowtie2 (<http://bowtie-bio.sourceforge.net>) [128,129]. The Tophat alignment was set to the default parameters, which allowed for a maximum of two mismatches at most to be accepted per read, and for mapping to more than one locus. The UCSC mouse reference genome was provided with a pre-built Bowtie2 index in order to create a transcriptome index from the gene transfer format (GTF) data. The Bash commands used for the alignment and the overall analysis pipeline are summarized in Appendix 3.

#### 2.4.7.3 Differential Expression Analysis of genes

Differentially expressed genes were identified using the Cuffdiff package of Cufflinks (version cufflinks-2.2.1, <http://cufflinks.cbcb.umd.edu/>) [130]. Based on the Tophat and Bowtie mapping outputs, cufflinks assembled the mapped reads into transcripts, and Cuffdiff estimated their abundances, and reported the differentially expressed reads and transcripts in a txt file as log<sub>2</sub>-normalized [129]. Reads obtained from the uninfected lungs were used as a control to normalize the gene expression ratio as infected vs. uninfected, and denote the ratio as log<sub>2</sub> values. Therefore, the genes differentially expressed in response to the WT infection were identified by calculating the ratio of gene expression between WT infection and uninfected, control mice, and the same was performed for genes differentially expressed in response to the  $\Delta mtp$ -mutant infection. The Linux commands used for the differential expression are summarized in Appendix 3.

#### 2.4.7.4 Visualization of data and Global Transcriptomics

The differential expression outputs were analysed using the reads obtained from the uninfected lungs as a control. Global statistics and visualisation of differentially expressed genes were conducted and determined using R version 2.4.1 ([www.r-project.org](http://www.r-project.org)), R-studio, and Bioconductor packages such as CummeRbund. The Multi Experimental Viewer (MeV, v. 4.9.0) program was used to visualise the DE gene sets in heat maps.

#### 2.4.7.5 Identification of Significantly Differentially Expressed genes (SDEG)

SDEGs were identified using the following parameters: (i) a log<sub>2</sub> fold-change above 1.5 for up-regulated genes, and less than 1.5 for down-regulated genes; and (ii) false discovery rate (FDR) corrected p-value (q-value) of  $\leq 0.05$  [131].

#### 2.4.7.6 Gene Ontology (GO) enrichment Analysis

GO enrichment analysis of SDEGs was performed with the Gene Ontology Consortium (<http://geneontology.org/page/go-enrichment-analysis>) using text files that contained the Gene ID and expression values as given by Cuffdiff. Under- and over-represented GO terms were extracted from the differentially expressed genes by selecting a p-value <0.05. This was used to study the molecular function (MF) of the regulated genes, the biological processes (BP) they are involved in, and, and the cellular compartments (CC) in which they are found.

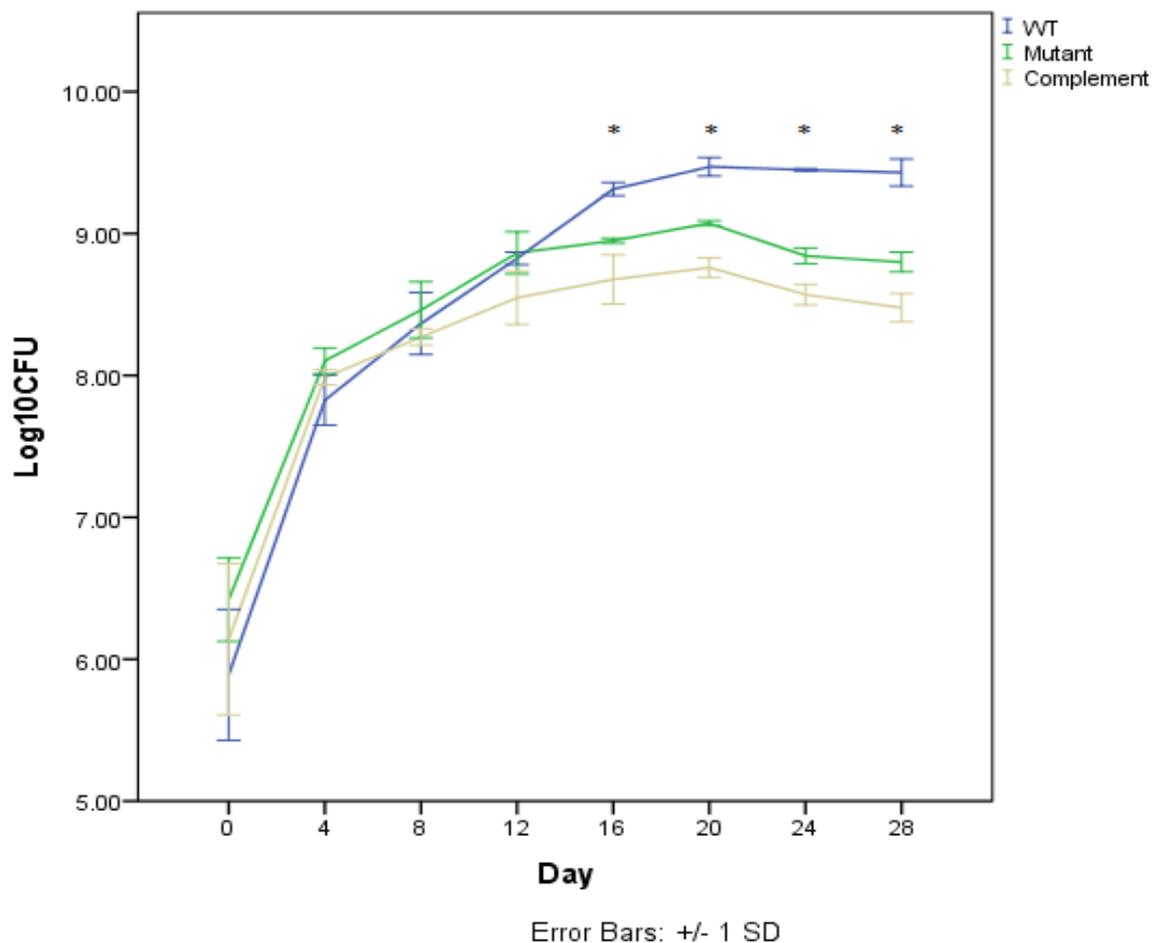
#### 2.4.7.7 Pathway and network analysis

Further analysis was carried out on the SDEGs ( $p \leq 0.05$ ) to determine their functional relevance within the context of networks and pathway enrichment using the Ingenuity Pathway Analysis (IPA) software ([www.ingenuity.com](http://www.ingenuity.com)) [132]. IPA interprets data from differentially expressed genes into known functions, pathways and gene interaction networks that are available on the Ingenuity database. Genetic networks are ranked by scores, transcriptional regulators are assigned activation z-scores which predict whether or not they are activated or inhibited in any given functional pathway/network, and molecules are assigned to functions and canonical pathways after specific p-value calculations. This data analysis by IPA permits for understanding and interpreting the significance of differentially expressed genes, and the gene products involved within a larger biological system [133]. Lists of SDEGs from the WT and  $\Delta mtp$ -mutant infection were uploaded to IPA, and the significance was set at a p-value of 0.05 [134].

## **CHAPTER 3: RESULTS**

### 3.1 *In vitro* growth kinetics

The WT,  $\Delta mtp$ -mutant and *mtp*-complemented strains were cultured in Middlebrook 7H9 broth supplemented with 10% (v/v) OADC, 0.5% (v/v) glycerol and 0.05% (v/v) Tween-80, to assess their *in vitro* growth rate (Figure 3.1.1). The initial inoculum of the  $\Delta mtp$ -mutant at Day 0 of the assay was unintentionally higher ( $3.02E+06$  CFU/mL) than the WT ( $1.09E+06$  CFU/ mL). However, the bacillary load of the latter increased to  $2.07E+09$  CFU/ mL compared to that of the  $\Delta mtp$ -mutant ( $8.91E+08$  CFU/ mL) by Day 16 of the assay. This Day 16 was the first time at which the first significant difference in growth was observed ( $p = 0.01$ ). The WT maintained a significantly higher ( $p < 0.001$ ) bacillary load through the stationary phase until day 28.



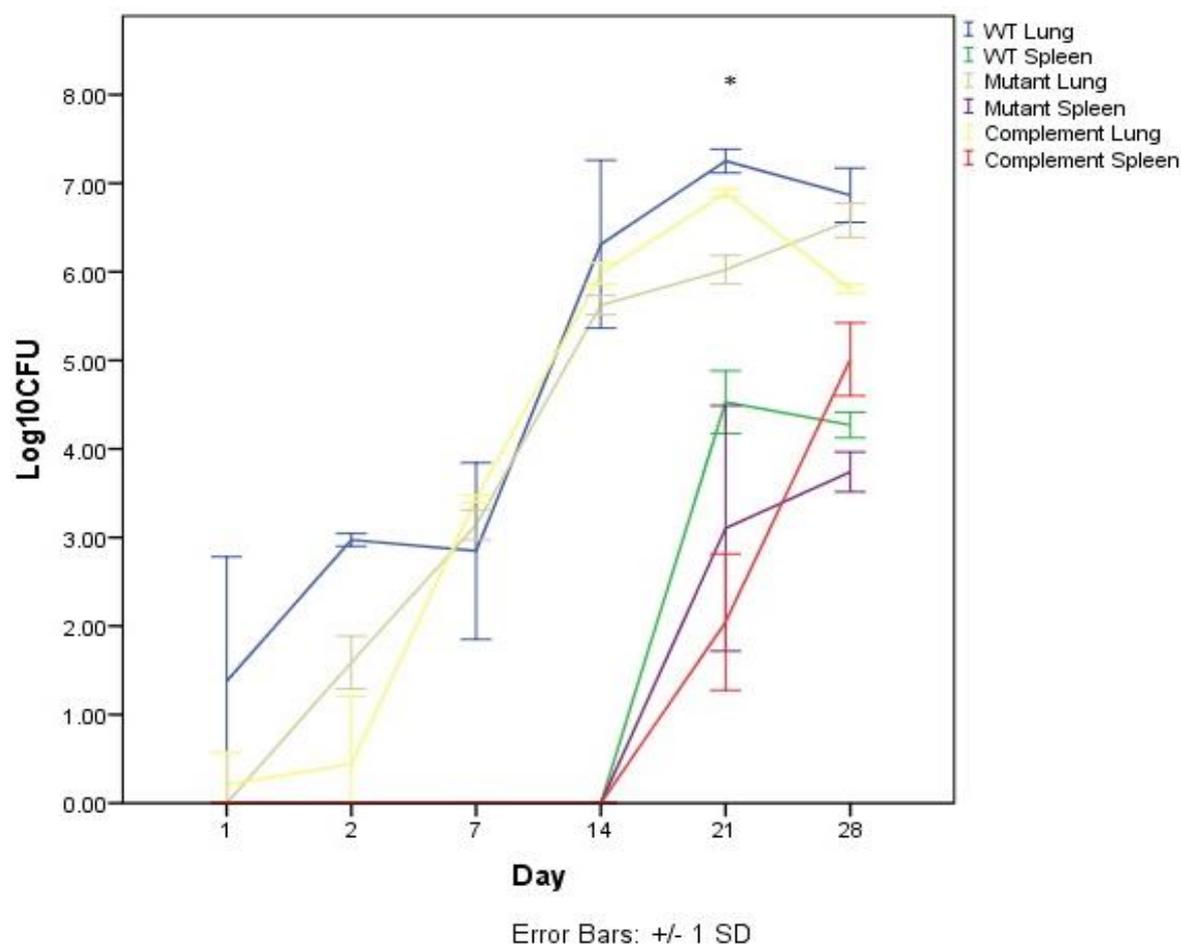
**Figure 3.1.1: *In-vitro* growth kinetics of  $\Delta mtp$ -mutant, *mtp*-complemented and wild type strains of *M. tuberculosis*.** The V9124 wild type of F15/LAM4/KZN family of *M. tuberculosis*, its  $\Delta mtp$ -mutant and *mtp*-complemented strains were grown in broth Middlebrook 7H9 culture for 28 days. CFU were enumerated at every 4-day interval. The  $\Delta mtp$ -mutant strain is deficient of the *mtp* gene that encodes curli pili in *M. tuberculosis*. Data from 3 experiments performed on 3 different days in triplicate. Mean CFU data and standard deviation is plotted.

\* student t-test performed for WT CFU compared to  $\Delta mtp$ -mutant CFU,  $P < 0.05$ .

WT: Wild type infection; Mutant:  $\Delta mtp$ -mutant infection; Complement: *mtp*-Complemented infection

### 3.2 Growth kinetics of *M. tuberculosis* in lungs and spleen during early infection

In the present study, Balb/C mice were infected with the  $\Delta mtp$ -mutant and complemented strains via aerosol inoculation to determine if they differed in their growth ability *in vivo* compared to the WT strain. Enumeration of the initial inocula by CFU/mL demonstrated that at day 0, mice were infected with  $4.17E+06$  CFU/mL,  $1.08E+06$  CFU/mL and  $2.70E+06$  CFU/mL ( $p=0.00018$ ) of the WT,  $\Delta mtp$ -mutant and complemented strains, respectively.



**Figure 3.2.1: *In vivo* growth and dissemination of wild type,  $\Delta mtp$ -mutant and complemented strains of *M. tuberculosis* in BALB/c mice.** *M. tuberculosis* burden in the lungs and spleen of BALB/c mice infected via aerosol inhalation. CFU/mL, expressed as log<sub>10</sub>CFU were enumerated in the lungs and spleen at the different time intervals. The  $\Delta mtp$ -mutant strain is deficient of the *mtp* gene that encodes curli pili in *M. tuberculosis*. Mean CFU data from 5 mice and standard deviation is plotted.

\* student t-test performed for WT CFU compared to  $\Delta mtp$ -mutant CFU,  $P<0.05$ .

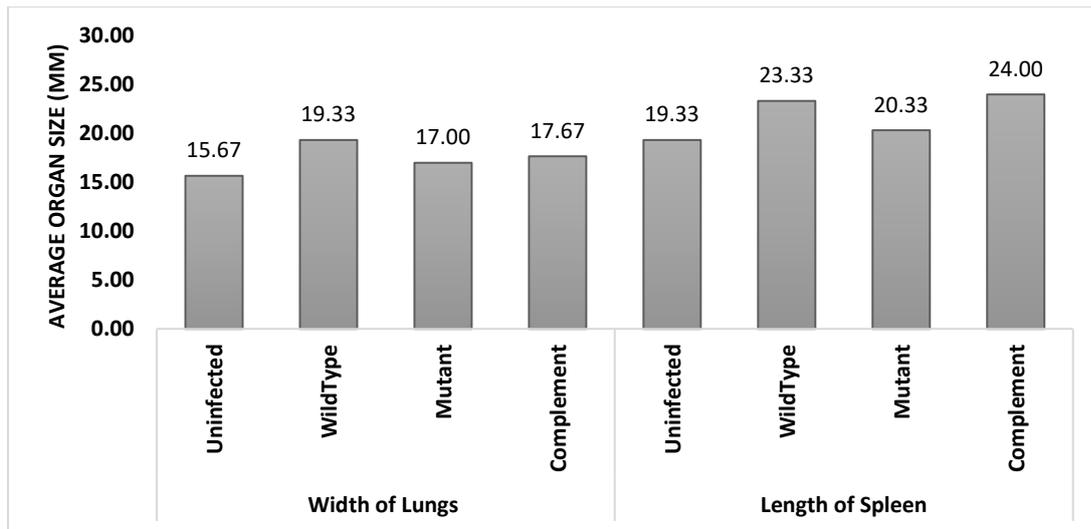
WT: Wild type infection; Mutant:  $\Delta mtp$ -mutant infection; Complement: *mtp*-Complemented infection

On the specified days after infection, mice were sacrificed and their organs were collected, weighed and homogenized in PBS. The bacterial load in the lungs and spleen was determined by serial dilution

and plating of the organ homogenates onto 7H11 agar plates. The mean log CFU/ mL from five mice per group of strains at each time point is shown (Figure 3.2.1). Despite the limitation of unequal infecting inoculums, the initial lung burden implanted into the lungs after 24hours of infection was slightly, but not significantly higher ( $p= 0.36$ ) for WT strain compared to the  $\Delta mtp$ -mutant. The  $\Delta mtp$ -mutant strain was slower to grow in both the lungs, and spleen compared to the WT strain (Figure 3.2.1). Lungs infected with the WT strain displayed a non-significant higher number of infecting bacilli throughout the duration of infection (Figure 3.2.1), except at 7days post infection. By 21days post infection, lungs infected with the WT displayed a significantly higher bacillary load  $1.46E+07$  CFU/mL compared to those infected by the  $\Delta mtp$ -mutant strain  $8.55E+05$  CFU/mL ( $p= 0.0098$ ) (Figure 3.2.1). The *mtp*-complemented strain displayed a slightly lower microbial burden than the WT, but significantly higher than the  $\Delta mtp$ -mutant strain, showing restoration of function. Similarly, spleens infected with the WT strain contained a higher bacterial load ( $1.54E+04$  CFU/mL) compared to the  $\Delta mtp$ -mutant strain ( $4.46E+03$  CFU/mL) after 28 days of infection ( $p= 0.021$ ) (Figure 3.2.1). At day 28, a slight decrease was observed in the number of bacilli burdening the lungs infected with the WT strain ( $p= 0.019$ ), but not in the lungs infected with the  $\Delta mtp$ -mutant strain. The *mtp*-complemented strain displayed a higher bacterial burden than both the WT and  $\Delta mtp$ -mutant strain, again, showing restoration of function.

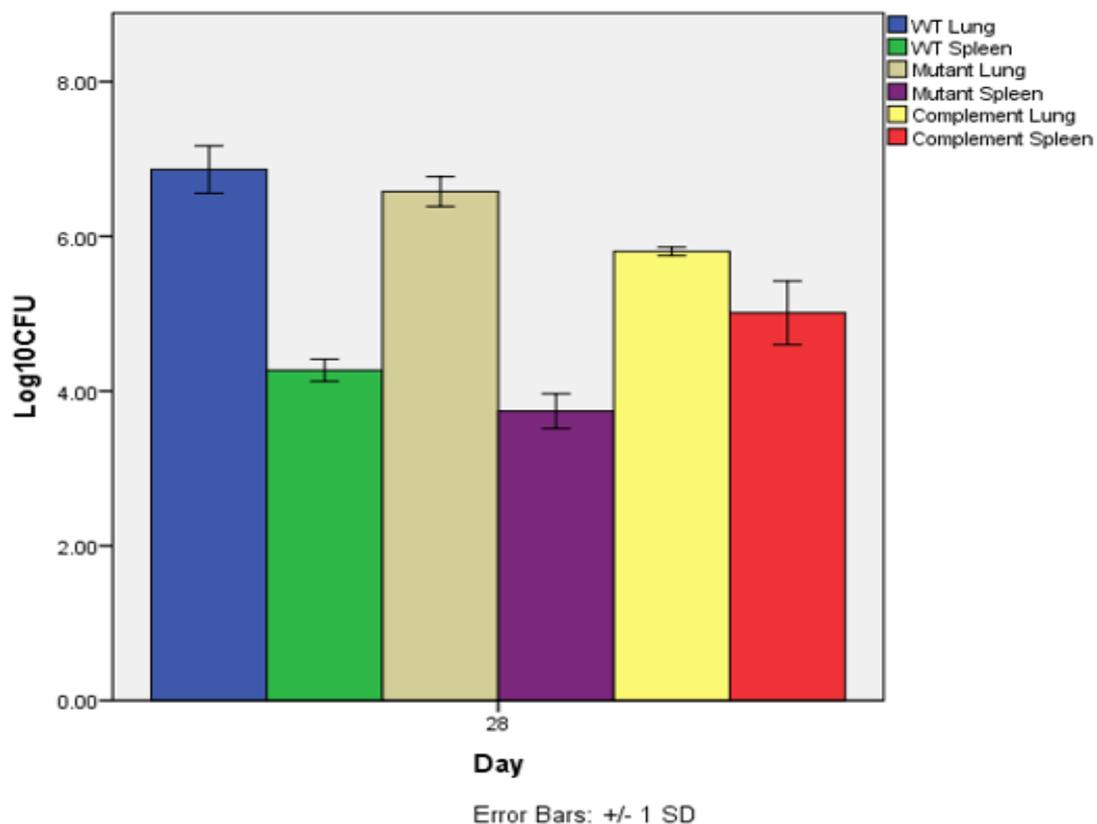
### 3.3 Gross organ pathology

After aerosol challenge with *M. tuberculosis*, all mice survived up until they were sacrificed according to protocol. Photographic images of harvested lungs and spleens were taken 28days post-infection using a NIKON 200 camera to assess visible signs of disease, and any differences in size between the organs infected with the 3 different strains (Figure 3.3.1). The spleen length correlated with the bacillary load of each strain in this organ (Figures 3.3.1 and 3.3.2). In contrast, the width of the lung infected with only the WT strain correlated with the bacillary load at day 28 (Figures 3.3.1 and 3.3.2). In the lungs, the WT strain resulted in the highest increase ( $p=0.04$ ) in width (Figures 3.3.1 and 3.3.3B), whereas infection with the *mtp*-complemented strain resulted in the highest size increase in the spleen (Figures 3.3.1 and 3.3.3D). The bacillary load in the lungs followed a similar pattern, where the number of organisms infecting the lungs was higher in the WT infected lungs than the *mtp*-complemented infected lungs (Figure 3.3.2). Similarly, the *mtp*-complement infected spleen harboured the most infecting bacilli (Figure 3.3.2), followed by the WT ( $6.37E+06$  CFU/mL), the lastly the  $\Delta mtp$ -mutant challenged spleen ( $2.97E+06$  CFU/mL), and the increase in spleen length followed a similar pattern (Figure 3.3.3).



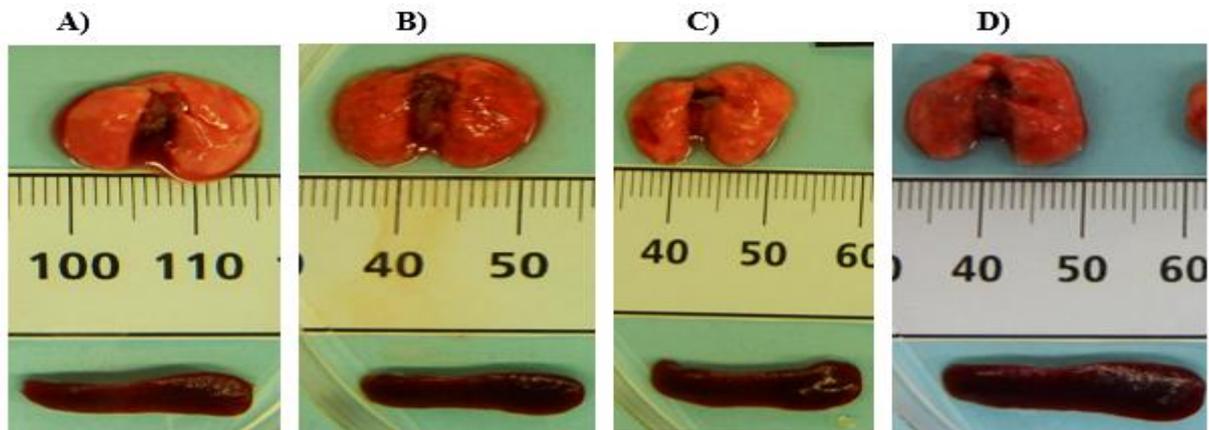
\* Mutant:  $\Delta mtp$ -mutant infection; Complement:  $mtp$ -Complemented infection

**Figure 3.3.1: Mean lungs and spleen sizes after aerosol inoculation, at week 4 post *M. tuberculosis* challenge.**



\* WT: Wild type infection; Mutant:  $\Delta mtp$ -mutant infection; Complement:  $mtp$ -Complemented infection

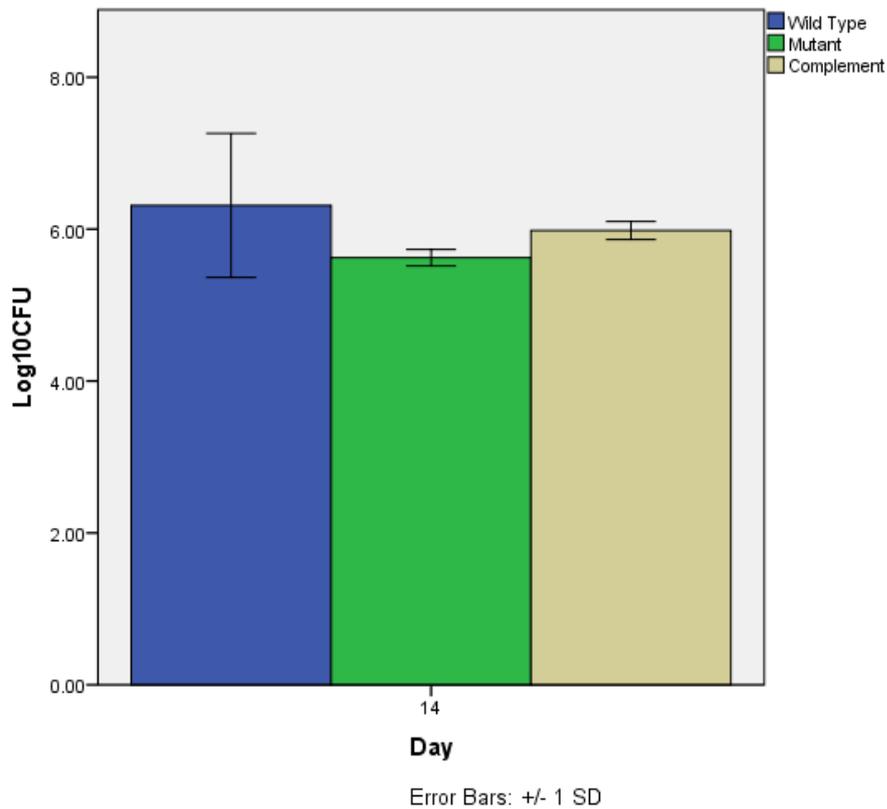
**Figure 3.3.2: In vivo *M. tuberculosis* burden at day 28.** Histogram shows the mycobacterial CFU/mL of the V9124 wild type,  $\Delta mtp$ -mutant and  $mtp$ -complemented strains of *M. tuberculosis* infecting BALB/c mice lungs and spleen 28 days post- infection.



**Figure 3.3.3: Gross appearance of lungs and spleens from aerosol infected mice at 4 weeks after challenge with *M. tuberculosis* strains.** A: Uninfected BALB/c mice organs; followed by organs infected with B: V9124 wild type strain of the F15/LAM4/KZN family of *M. tuberculosis* (C),  $\Delta mtp$ -mutant (D) *mtp*-complemented strain.

### 3.4 *M. tuberculosis* burden in lungs 14days post infection

RNA-Sequencing was performed at 14 days post-infection on the lungs of 6 mice, consisting of two uninfected, and two each infected with the WT and  $\Delta mtp$  strains respectively, for transcriptome analysis. At this time point, the lungs of the infected mice contained fewer  $\Delta mtp$ -mutant bacilli ( $3.21E+05$  CFU/mL) than the WT ( $7.30E+06$  CFU/mL) (Figure 3.4.1). However, this difference was not significant ( $p = 0.3$ ).



\* WT: Wild type infection; Mutant:  $\Delta mtp$ -mutant infection; Complement: *mtp*-Complemented infection

**Figure 3.4.1: *M. tuberculosis* burden in infected mice lungs at day 14 post-infection.** CFU/mL of the V9124 wild type,  $\Delta mtp$ -mutant and *mtp*-complemented strains of *M. tuberculosis* infecting BALB/c mice lungs at 14 days post- infection were enumerated.

### 3.5 RNA Concentration and Integrity

In order to interrogate the transcriptome profile of *M. tuberculosis* infected lungs, and understand the transcriptomic response associated with MTP, RNA was extracted from the WT- and  $\Delta mtp$ -mutant infected lungs. RNA was also generated from uninfected lungs for comparative purposes. The concentration and integrity (Table 3.5.1) of the RNA was verified for its use in library preparation for RNA sequencing with an Illumina HiSeq 2000. For each condition (n=3), the two biological replicates (Table 3.5.1) yielded high RNA concentrations of over 1000ng/  $\mu$ L, high A260/280 and high 260/230 ratios for all 6 samples, and RNA integrity numbers (RIN) between 6.90 and 9.80 (Table 3.5.1). The MOPS gel (Appendix 2) showed 2 distinct RNA bands representing the 2 eukaryotic RNA species (18S and 28S).

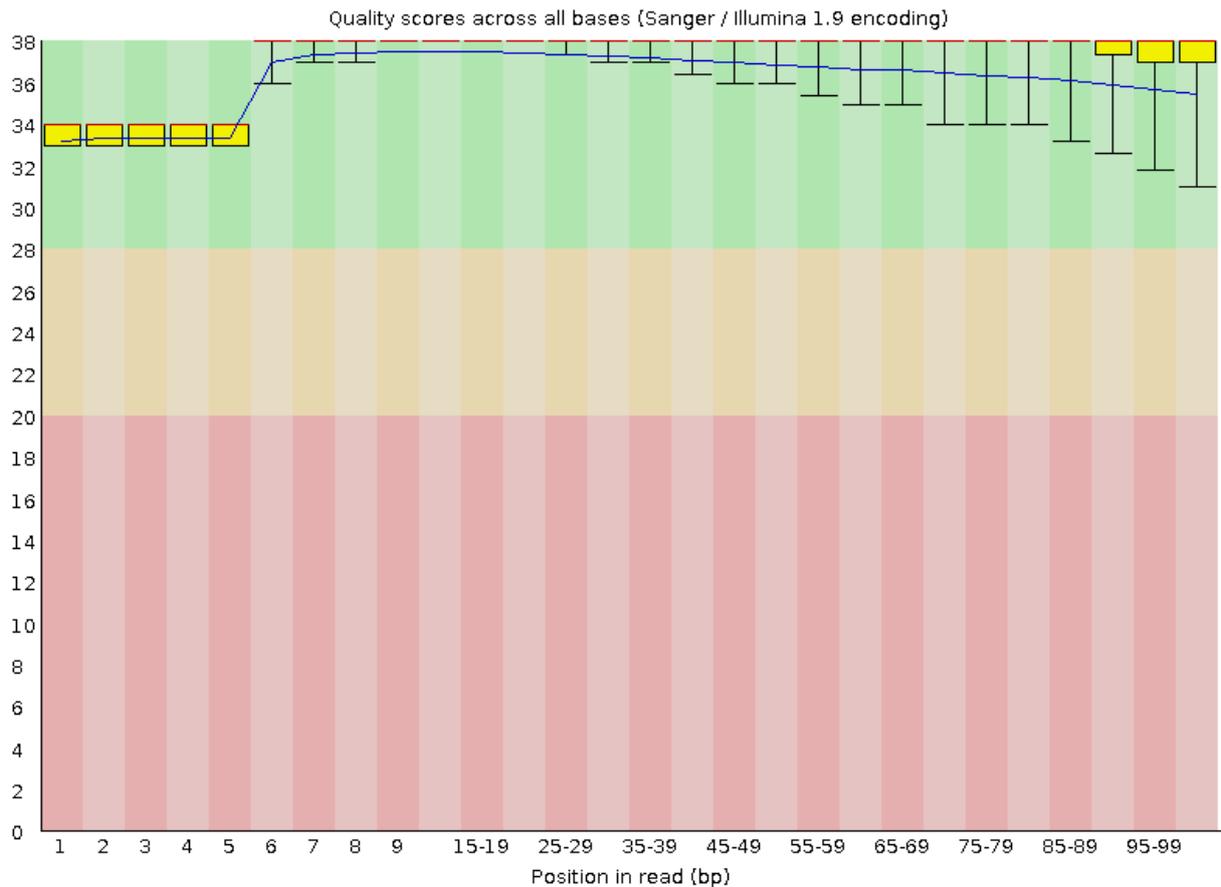
**Table 3.5.1: RNA concentrations post extraction.** RNA extraction was performed using the Qiagen RNeasy kit. RNA concentration was obtained using a Nanodrop, and the RIN was measured using the Agilent Bioanalyzer

Sample ID	Concentration [ng/μl]	A260/280	A260/230	RIN
Uninfected Control 1	1378,7	2,13	2,17	9,80
Uninfected Control 2	1119,4	2,06	2,14	9,30
Day14 Wild Type 4	2252,6	2,12	2,22	7.50
Day14 Wild Type 5	2521,8	2,14	2,22	6.90
Day14 $\Delta mtp$ -mutant 3	1034,2	2,14	2,03	8.00
Day 14 $\Delta mtp$ -mutant 4	2165.3	2,04	2,16	8,00

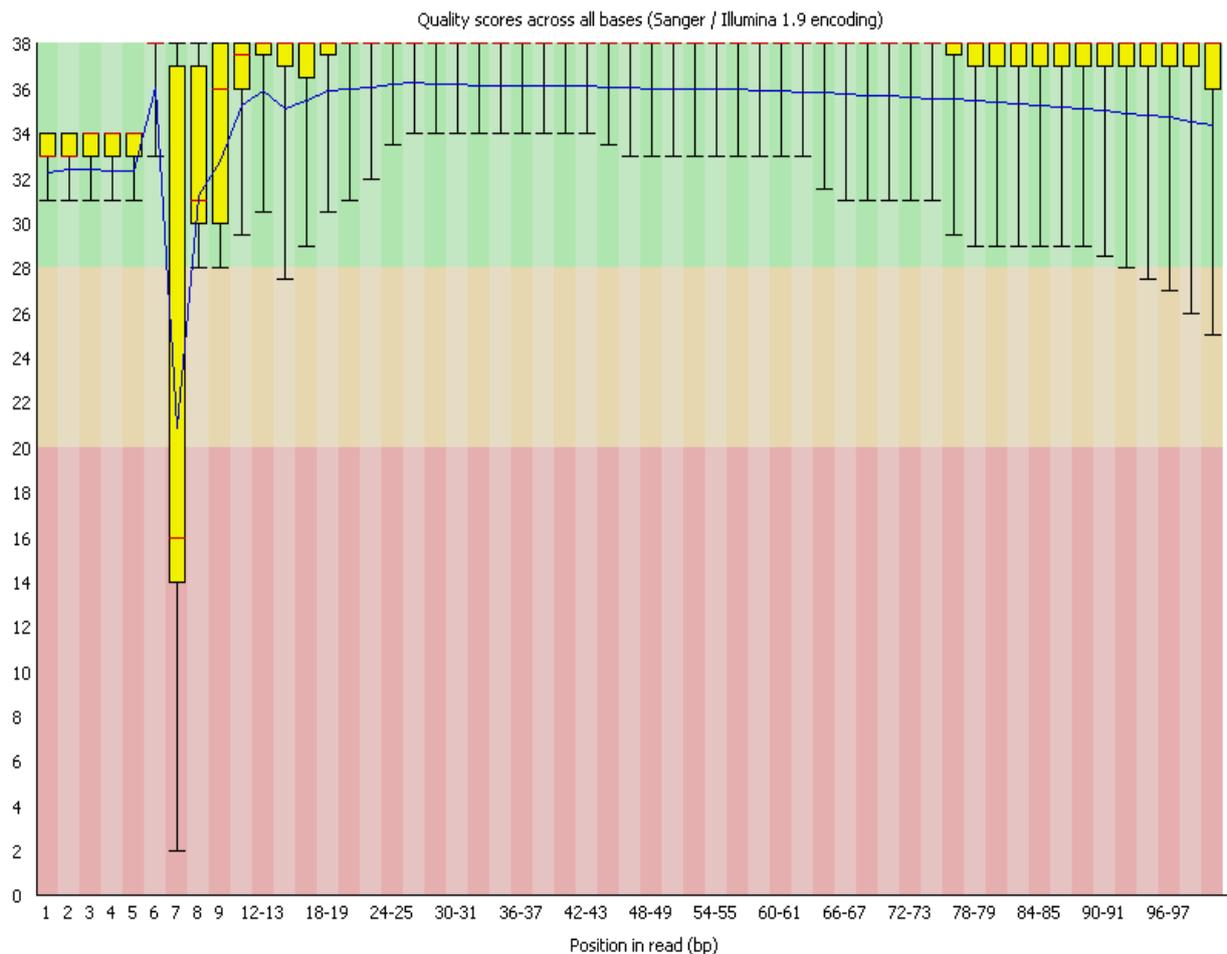
### 3.6 Processing of RNA-Seq reads

#### 3.6.1 Pre-processing - Read Quality Control

The extracted RNA was converted into cDNA, and RNA libraries were created for each condition. The independent biological replicates were sequenced and the resulting reads were processed through our RNA-Seq data analysis pipeline (Appendix 3). The FastQC tool was used to evaluate the read quality of the generated raw data [<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>]. The basic statistics of the quality control (QC) results showed that all reads generated were 100bp long in sequence length, with no sequences flagged as poor quality, and the GC base content was between 42-44%. The mean of the per base sequence quality score was above 20 (example shown in (Figure 3.6.1.1) for all but 1 sample, the reverse sequence of the  $\Delta mtp$ -mutant 3 sample, where the quality score of the read in position 7 fell below 20, to a score of 16 (Figure 3.6.1.2).



**Figure 3.6.1.1: Sample FastQC output of the Illumina reads from the 6 libraries that were generated from the RNA obtained from the WT and  $\Delta mtp$ -mutant infected lungs.** The rest of the graphs from this QC are presented in Appendix 4. The mean quality score across the sequence length is represented by the blue line. The green, orange and pink colours represent areas of good, acceptable and bad quality scores respectively.



**Figure 3.6.1.2: FastQC output of the reverse sequence of the  $\Delta mtp$ -mutant infected lung sample.** The quality score of the read in position 7 is 16. The mean quality score across the sequence length is represented by the blue line. The green, orange and pink colours represent areas of good, acceptable and bad quality scores respectively.

### 3.6.2 Alignment and comparison with the reference genome

The raw Illumina sequence reads generated from *M. tuberculosis* infected and uninfected lungs were mapped to the annotated *Mus musculus* mm10 mouse genome from the UCSC browser ([illumina.com/Mus\\_musculus/UCSC/mm10/Mus\\_musculus\\_UCSC\\_mm10.tar.gz](http://illumina.com/Mus_musculus/UCSC/mm10/Mus_musculus_UCSC_mm10.tar.gz)) using Tophat version tophat-2.1.0 [135]. Between 64.9%-77% of the reads for each sample aligned to the mouse reference genome (Table 3.6.2.1). Of these, 91.7% - 92.5% did not align to multiple locations in the genome (Table 3.6.2.1). The  $\Delta mtp$ -mutant infected lung samples had the highest percentage (72.1%) of reads aligning to the reference genome, whereas the WT infected lung samples had 70.75% of their read aligning to the reference genome (Table 3.6.2.1).

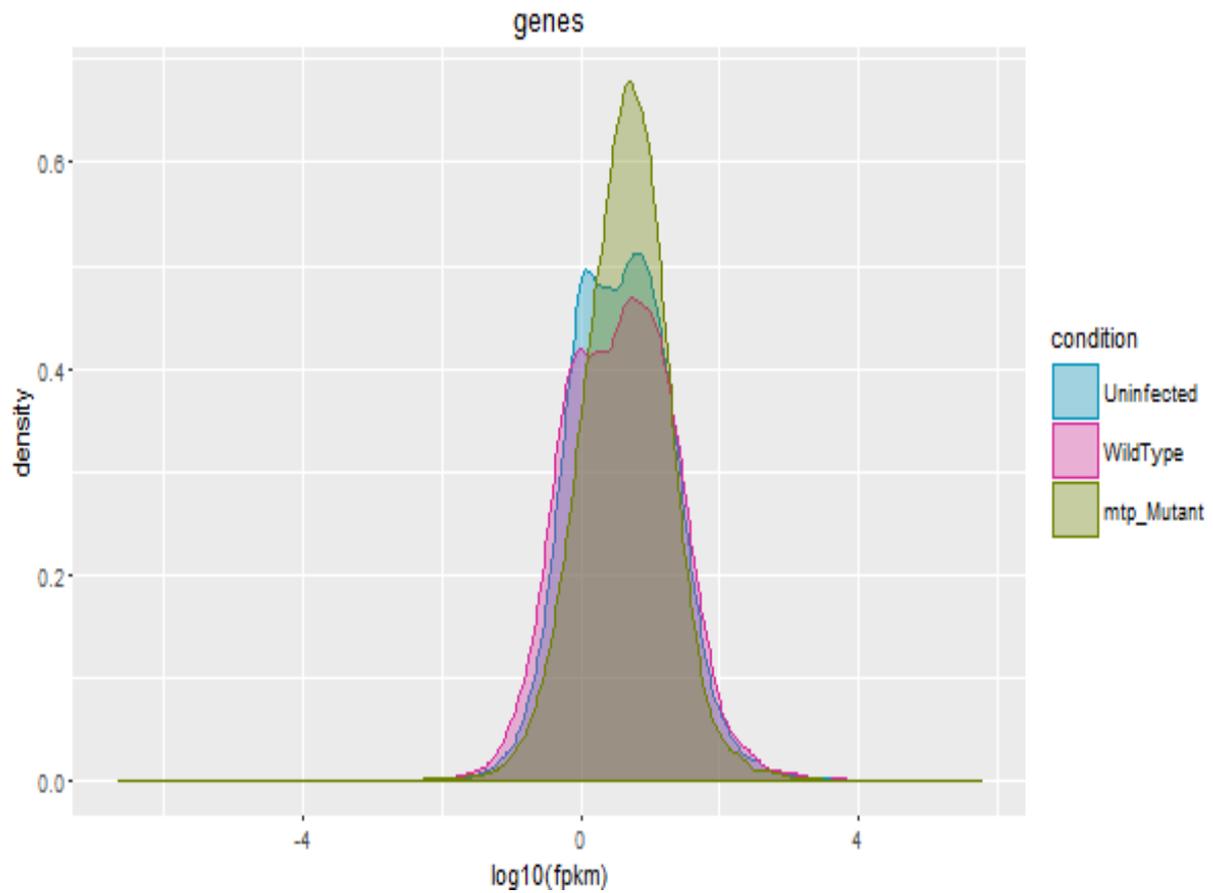
**Table 3.6.2.1. Summary of the number of the 100bp paired ended sequences and percent of mapped reads obtained for the 6 libraries mapped to the *mus musculus* mm10 mouse genome.**

	Mouse #	Total number of reads	% mapped reads	% uniquely aligned reads
<b>Uninfected</b>	1	69,548,356	64.9	92.50
	2	73,775,199	72.5	91.90
<b>WT</b>	1	195,652,780	70	92.55
<b>infected</b>	2	177,887,281	71.5	91.70
<b><i>Δmtp</i>-mutant</b>	1	162,002,242	77	91.95
<b>infected</b>	2	242,575,164	67.2	92.15

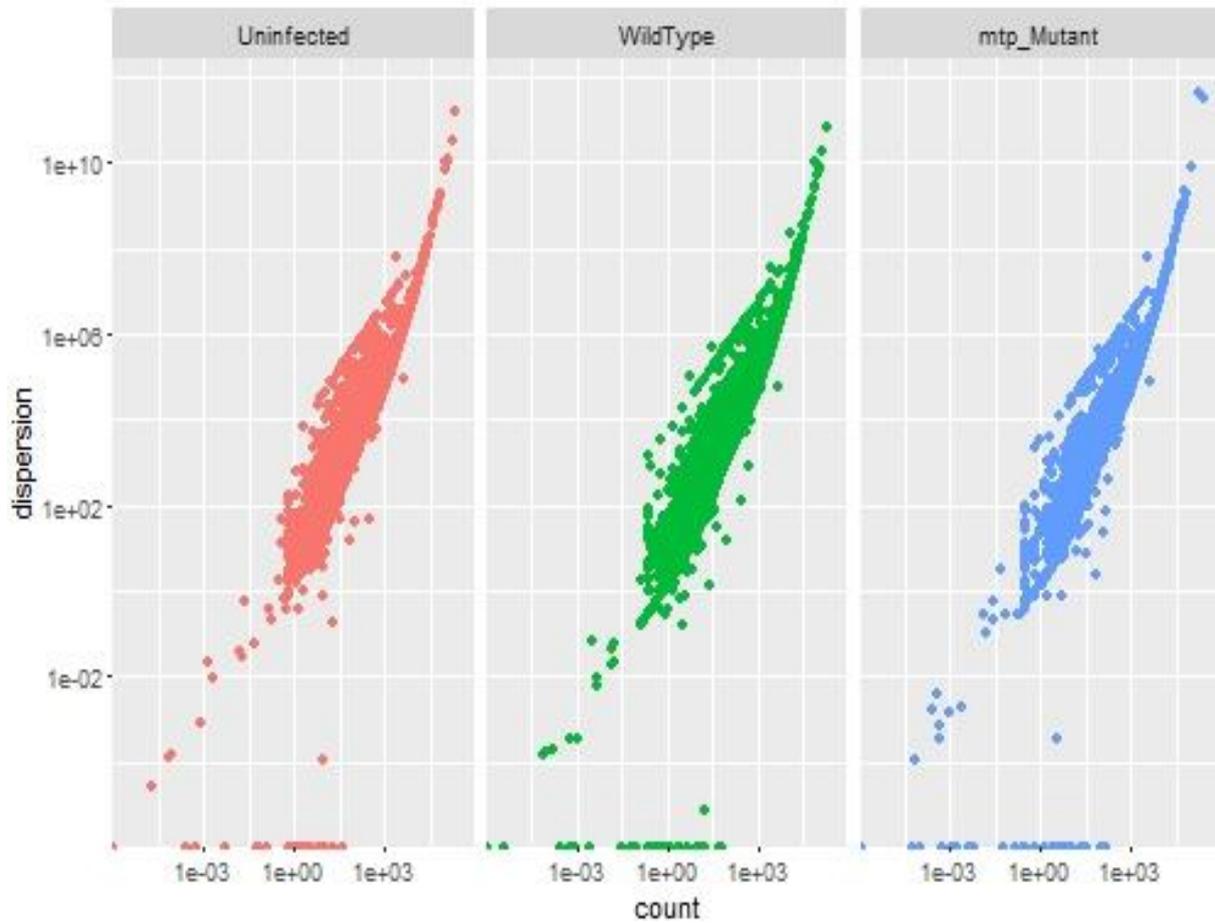
\* WT: Wild type infection

### **3.7 Global changes in the lung transcriptional response of Balb/C mice to infection by *M. tuberculosis***

Expression values of the 24 346 mouse genes were analysed to define global changes in the WT and *Δmtp*-mutant *M. tuberculosis*-infected mouse lung transcriptomes relative to uninfected lungs. The aligned reads were assembled into gene transcripts and Cufflinks was used to calculate the relative abundance of these transcripts. Transcription profiles were obtained using Cuffdiff outputs in order to identify the differentially expressed genes in the two experimental groups (WT and *Δmtp*-mutant infected lungs). Density (Figure 3.7.1) and Dispersion (Figure 3.7.2) plots of host genes were constructed to assess the distributions of fragments per kilobase per million mapped reads (FPKM) scores across all samples. The total reads mapping to the mouse genome from the WT *M. tuberculosis* infection read pool was lower than that of the *Δmtp*-mutant *M. tuberculosis* infection read pool (Table 3.6.2.1). Similarly, the density of transcripts that were differentially expressed in response to *Δmtp*-mutant *M. tuberculosis* infection was greater than that of the WT relative to the transcripts from uninfected mice (Figure 3.7.1). However, the transcripts from the WT and *Δmtp*-mutant infection showed similar patterns of dispersion (Figure 3.7.2).



**Figure 3.7.1. Density plot of expressed genes.** Differential expression analysis of uninfected and *M. tuberculosis*-infected mouse lungs. Differential gene expression data was analysed with the R version 3.2.2. The plot was generated using the CummeRbund package v. 2.0.

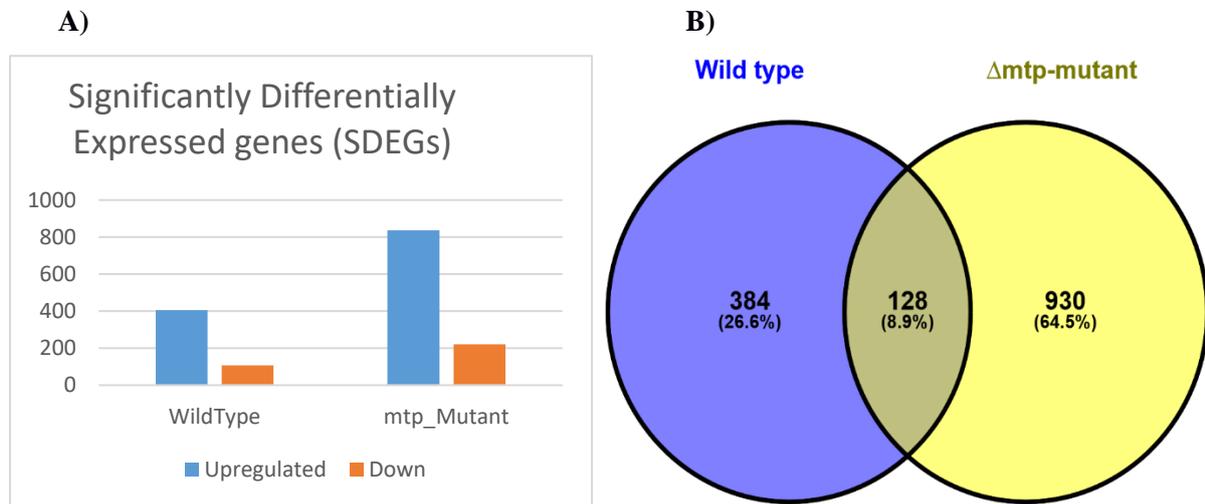


**Figure 3.7.2. Dispersion plot of expressed genes.** Differential expression analysis of uninfected and *M. tuberculosis*-infected mouse lungs. Differential gene expression data was analysed with the R version 3.2.2. Plot was generated using the CummeRbund package v. 2.0.

### 3.7.1 Significantly differentially expressed genes (SDEGs) and Gene Ontology (GO) Analysis

A total of 24 346 genes was analysed, of which 1 262 and 1 523 genes were differentially expressed in the lungs of mice due to infection with the WT and  $\Delta mtp$ -mutant strains of *M. tuberculosis* respectively, relative to the uninfected mice. Of these, only 512 and 1 058 genes were significantly differentially expressed due to the WT and  $\Delta mtp$ -mutant respectively. The  $\log_2$  fold change ratio of Treatment/Control was used to determine significant genes with a  $q$ -value  $< 0.05$ . A specific cut-off was imposed, where the upregulated or down regulated genes with a ratio of  $\geq 1.5$  or  $\leq -1.5$  respectively were selected and considered as significantly differentially expressed genes (SDEGs) in order to identify gene expression changes with the highest potential impact. A total of 1 442 genes was found to be significantly differentially expressed due to *M. tuberculosis* infection with the WT and  $\Delta mtp$ -mutant strain. Of these, the total number of activated genes was significantly higher compared to the total number of repressed genes in both infections (Figure 3.7.1.1). The WT and  $\Delta mtp$ -mutant infection uniquely induced 384 and 930 SDEGs respectively (Figure 3.7.1.1). The WT infection displayed the widest range of expression

(-4.93211 to 10.1523) compared to the  $\Delta mtp$ -mutant infection expression range (-4.54057 to 9.87822). Of the top 10 most upregulated genes (Table 3.7.1.1), 8 were common in both infections. These were the *Krt13*, *Krt4*, *Acta1*, *Defb4*, *Serpib3a*, *Serpib12*, *Mylpf*, and *Tnnt3* genes. In contrast, only 2 genes, *S100a9* and *Hist1h3g* were common in the top 10 downregulated genes.



**Figure 3.7.1.1: Significantly Differentially Expressed Genes (SDEGs).** Changes in the lung transcriptome of WT and  $\Delta mtp$ -mutant *M. tuberculosis* infected mice lungs were expressed as a ratio of  $\log_2$  fold change of gene expression relative to the uninfected lungs. A total of 512 and 1 058 genes were significantly differentially expressed due to the WT and  $\Delta mtp$ -mutant infection respectively. A: Comparison of the up- (blue), and down- (orange) regulated SDEGs. B: Venn diagram depicting the distribution of the SDEGs with the overlapping intersection representing a set of genes evident as significantly differentially expressed in both infections. Wild type: Wild type infection and  $\Delta mtp$ -mutant:  $\Delta mtp$ -mutant infection.

**Table 3.7.1.1: Top ten up-regulated and down regulated SEDGs elicited by the host in response to each *M. tuberculosis* infection relative to uninfected mice.**

Upregulated genes				Downregulated genes			
WT infection		$\Delta mtp$ -mutant infection		WT infection		$\Delta mtp$ -mutant infection	
Gene ID	Expression value ( $\log_2$ fold change)	Gene ID	Expression value ( $\log_2$ fold change)	Gene ID	Expression value ( $\log_2$ fold change)	Gene ID	Expression value ( $\log_2$ fold change)
Krt13	10.1523	Krt13	9.87822	Hist1h3g	-4.93211	Camp	-4.54057
Krt4	8.609	Krt4	8.91507	Terc	-4.53914	Ngp	-3.86736
Defb4	8.15211	Saa3	7.35888	H2-Ea-ps	-4.08168	S100a9	-3.74043

Acta1	8.00758	Acta1	7.30966	S100a9	-4.00378	Stfa2	-3.60231
Serpib3a	7.3584	Defb4	6.75932	Hist1h4i	-3.95731	Snord22	-3.52332
Mylpf	7.14667	Serpib3a	6.72511	Gdpd3	-3.94795	Mcpt8	-3.16553
Serpib12	7.00409	Serpib12	6.4943	H2-Eb1	-3.80649	Prss34	-3.14263
Tnnt3	6.88874	Mylpf	6.3218	Ifitm6	-3.77469	Retnlg	-3.0125
Myh8	6.53424	Tnnt3	6.03244	Wfdc21	-3.70049	Col4a6	-3.00524
Myh1	6.17073	AA467197	5.8065	2410137M14Rik	-3.47148	Hist1h3g	-2.967

### 3.7.1.1 Unique SDEGs in response to WT and $\Delta mtp$ -mutant *M. tuberculosis* infection

The  $\Delta mtp$ -mutant infection induced more than a 2-fold increase in the number of unique SDEGs compared to the WT infection (Figure 3.7.1.1). GO data was interrogated to determine the biological functions and the GO functional categories (FCs) associated with the 384, and 930 unique SDEGs elicited by the WT and  $\Delta mtp$ -mutant *M. tuberculosis* infections respectively. Among the 384 SDEGs genes that were regulated  $\geq 1.5$ -fold following bacterial infection with the WT strain only, a significant enrichment was observed in functions related to chemokine receptor binding e.g. GO:0042379, chemokine activity e.g. GO:0008009, cytokine receptor binding e.g. GO:0005126 and cytokine activity e.g. GO:0005125. Other binding functions e.g. GO:0005102 were also found to be associated with these genes (Table 3.7.1.1.1). The CXCR chemokine receptor binding (GO:0045236) and the CCR chemokine receptor binding (GO:0048020) were amongst the top biological functions enriched by the WT specific genes.

Whilst the WT enriched biological functions were involved in immune response, comparative GO enrichment analysis of the 930 SDEGs uniquely elicited by the host after  $\Delta mtp$ -mutant infection revealed that molecular transducer activity (GO:0060089), receptor activity (GO:0004872), signal transducer activity (GO:0004871), signalling receptor activity (GO:0038023) and ion binding (GO:0043167) were among the top biological functions (Table 3.7.1.1.1) affected by removal of the *mtp* gene. The G-protein was found to be significantly associated with biological functions enriched by SDEGs uniquely elicited by both infections. The G-protein coupled receptor binding (GO:0001664) and G-protein coupled receptor activity (GO:0004930) were enriched by the WT and  $\Delta mtp$ -mutant infections respectively (Table 3.7.1.1.1).

**Table 3.7.1.1.1: The GO functional categories (FCs) of the GO molecular functions from the unique gene datasets.**

Genes unique to WT	P-value	Genes unique to $\Delta mtp$ -mutant	P-value
CXCR chemokine receptor binding (GO:0045236)	1.51E-07	molecular transducer activity (GO:0060089)	3.63E-19

chemokine activity (GO:0008009)	4.68E-18	receptor activity (GO:0004872)	2.56E-22
CCR chemokine receptor binding (GO:0048020)	1.23E-06	signal transducer activity (GO:0004871)	5.48E-19
chemokine receptor binding (GO:0042379)	6.08E-19	signalling receptor activity (GO:0038023)	2.73E-22
cytokine receptor binding (GO:0005126)	2.20E-12	ion binding (GO:0043167)	6.68E-03
G-protein coupled receptor binding (GO:0001664)	1.13E-10	organic cyclic compound binding (GO:0097159)	2.42E-03
protein homodimerization activity (GO:0042803)	1.36E-02	heterocyclic compound binding (GO:1901363)	2.78E-03
receptor binding (GO:0005102)	9.93E-06	catalytic activity (GO:0003824)	2.22E-06
identical protein binding (GO:0042802)	5.21E-04	olfactory receptor activity (GO:0004984)	2.06E-14
carbohydrate derivative binding (GO:0097367)	1.75E-03	G-protein coupled receptor activity (GO:0004930)	3.71E-14
anion binding (GO:0043168)	4.31E-03	odorant binding (GO:0005549)	1.12E-15
hydrolase activity (GO:0016787)	4.06E-02		
protein binding (GO:0005515)	6.44E-12	pheromone receptor activity (GO:0016503)	1.16E-05
binding (GO:0005488)	4.33E-08	pheromone binding (GO:0005550)	5.26E-04

\* WT: Wild type infection

### 3.7.1.2 Shared SDEGs in response to both WT and $\Delta mtp$ -mutant infection

Only 128 (8.90%) of the 1442 genes were common in both infections, and of those, 74.2% and 25.8% were upregulated and downregulated respectively. In both strains, the regulation of the genes was similar, i.e. all genes were up/down regulated in both conditions. However, the extent of the regulation differed. For most of the genes, the magnitude of up-regulation was higher in the WT infection model compared to the  $\Delta mtp$ -mutant infection. Up and down regulated SDEGs common in both infections were enriched for molecular functions, cellular components and biological processes (Tables 3.7.1.3.1 and Table 3.7.1.4.1.).

### 3.7.1.3 Upregulated genes

The heat map (Figure 3.7.1.3) showed that among the upregulated genes, the Keratin 13 (*Krt13*) and Keratin 4 (*Krt4*) genes were the most up-regulated genes in both infections. The *Krt13* displayed a log<sub>2</sub>fold change value of 9.88 and 10.15, whilst the values for *Krt4* were 8.91507 and 8.609 in the WT and  $\Delta mtp$ -mutant infection respectively. These genes both belong to the keratin family, which are intermediate filaments involved in structural molecule activity, structural constituent of cytoskeleton;

protein binding and keratin filament binding. It then follows that the over-represented GO terms from this set of genes are associated with muscle contraction (GO:0006936), regulation of striated muscle contraction (GO:0006942), muscle system process (GO:0003012) and regulation of muscle system process (GO:0090257), and they were mostly found in the myofibril (GO:0030016) and contractile fibers (GO:0043292) (Table 3.7.2.1.1).

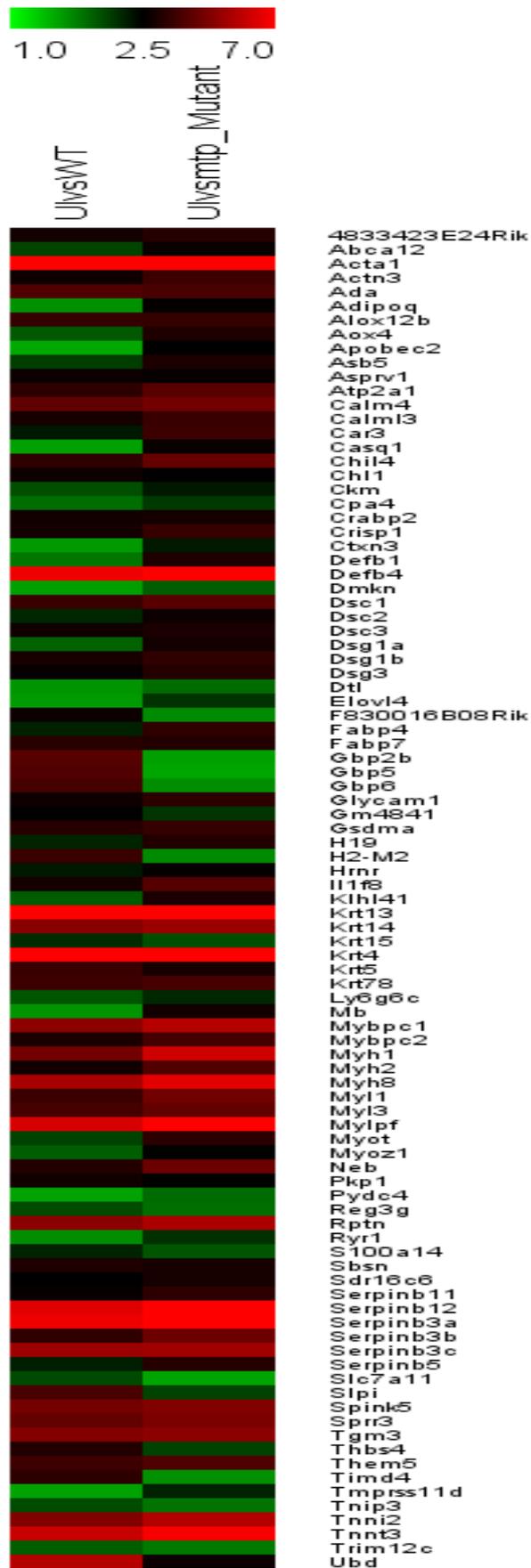


Figure 3.7.1.3.1: Differential expression of host genes in response to *M. tuberculosis* infection 14days post-infection. The 95 commonly up-regulated genes between the WT and  $\Delta mtp$ -mutant

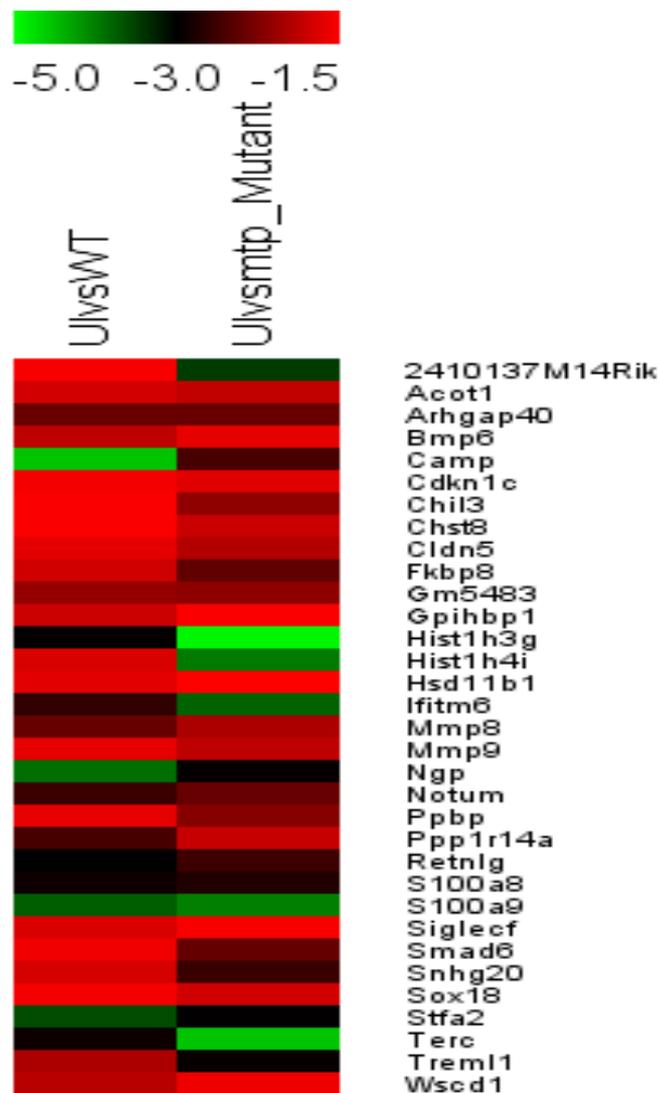
infections of Balb/C mice lungs. UI vs WT: Wild type infection and UI vs mtp Mutant:  $\Delta mtp$ -mutant infection. Colour intensity indicates the magnitude of up-regulation: red represents highly activated genes, and green represents lowly activated genes, relative to the uninfected control ( $p < 0.05$ ).

**Table 3.7.1.3.1: Gene ontology enrichment of the 95 common up-regulated genes.** GO enrichment analysis was performed using the Gene Ontology Consortium.

<b>GO biological process</b>	<b>P-value</b>	<b>GO molecular function</b>	<b>P-value</b>	<b>GO cellular component</b>	<b>P-value</b>
regulation of striated muscle contraction (GO:0006942)	3.72E-03	serine-type endopeptidase inhibitor activity (GO:0004867)	1.44E-04	myofibril (GO:0030016)	5.14E-11
muscle contraction (GO:0006936)	6.08E-07	endopeptidase inhibitor activity (GO:0004866)	3.27E-03	contractile fiber (GO:0043292)	1.26E-10
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	1.48E-02	endopeptidase regulator activity (GO:0061135)	4.12E-03	sarcomere (GO:0030017)	2.01E-10
regulation of muscle contraction (GO:0006937)	4.13E-05	peptidase inhibitor activity (GO:0030414)	4.97E-03	contractile fiber part (GO:0044449)	5.23E-10
muscle system process (GO:0003012)	5.26E-07	peptidase regulator activity (GO:0061134)	1.48E-02	extracellular region (GO:0005576)	2.00E-09
regulation of muscle system process (GO:0090257)	6.71E-04	calcium ion binding (GO:0005509)	1.06E-09	desmosome (GO:0030057)	1.41E-08
		nucleic acid binding (GO:0003676)	1.70E-02	extracellular exosome (GO:0070062)	1.59E-07

#### 3.7.1.4 Downregulated genes

The down regulated genes did not follow a similar trend to the up-regulated genes. The 2 genes that were most repressed due to the WT infection, based on expression fold-change, differed from those that were most repressed in the  $\Delta mtp$ -mutant. In the lungs infected with the WT strain, the cathelicidin antimicrobial peptide (*Camp*) gene was the most repressed, followed by the neutrophilic granule protein (*Ngp*), with log2fold change values of -3.87 and -4.54 respectively (Figure 3.7.1.1). *Camp* is associated with immune system processes like antibacterial and antifungal humoral response, and response to stimulus. The *Ngp* gene is involved in defence response processes as well as negative regulation of endopeptidase and peptidase activity. It is mostly found in the extracellular region. With an expression fold change value of -4.93, Histone cluster 1, H3g (*Hist1h3g*) gene was the most repressed gene in the lungs after infection with the  $\Delta mtp$ -mutant. It plays a crucial role in DNA replication and chromosomal stability, DNA repair and transcription regulation. The 2<sup>nd</sup> most repressed gene with an expression fold change value of -4.54, encodes the telomerase RNA component (*Terc*) in the nucleus. It is involved in cell differentiation and proliferation, cellular component organization, homeostatic and immune system process, and system development.



**Figure 3.7.1.4.1: Differential expression of host genes in response to *M. tuberculosis* infection 14days post-infection.** The heat map shows the 33 commonly downregulated genes between the WT and  $\Delta mtp$ -mutant infections of Balb/C mice lungs. UI vs WT: Wild type infection and UI vs mtp Mutant:  $\Delta mtp$ -mutant infection. Colour intensity indicates the magnitude of up-regulation: Green represents highly activated genes, and Red represents lowly activated genes, relative to the uninfected control ( $p < 0.05$ ).

The GO biological process terms from the downregulated genes are associated with neutrophil aggregation (GO:0070488), leukocyte migration involved in inflammatory response (GO:0002523), and leukocyte migration (GO:0050900). The molecular function involved with these genes is the Toll-like receptor 4 binding (GO:0035662), and the significantly located in the extracellular space (GO:0005615) (Table 3.7.1.4.1).

**Table 3.7.1.4.1: Gene ontology enrichment of the 33 common down regulated genes.** GO enrichment analysis using the Gene Ontology Consortium revealed significantly enriched GO terms.

GO biological processk	P-value	GO molecular function	P-value	GO cellular component	P-value
neutrophil aggregation (GO:0070488)	2.39E-02	Toll-like receptor 4 binding (GO:0035662)	3.72E-02	extracellular space (GO:0005615)	5.01E-03
leukocyte migration involved in inflammatory response (GO:0002523)	5.04E-03				
leukocyte migration (GO:0050900)	1.09E-02				

### 3.8 Ingenuity Pathway Analysis (IPA) of SDEGs

Canonical pathways, networks and upstream regulators associated with the SDEGs in the lungs after infection with the WT and  $\Delta mtp$ -mutant strains were analysed by Ingenuity Pathway Analysis (IPA), in order to elucidate the effect of the *mtp* gene deletion on the host. The IPA knowledgebase was used for functional pathway/network analysis, using our annotated SDEGs. Pathways and networks enriched by both strains, and/or uniquely by either strain were identified and are summarized in Table 3.8.1.

**Table 3.8.1: A summary of the numbers of significantly differentially expressed genes, and their associated networks and pathways.**

	All WT	All $\Delta mtp$ -mutant	Shared	Unique WT	Unique $\Delta mtp$ -mutant
<b>Number of Genes</b>	512	1058	128	384	930
<b>Number of Pathways</b>	109	54	30	79	24
<b>Number of Networks</b>	17	25	6	14	21

WT: genes elicited in the Wild type infection

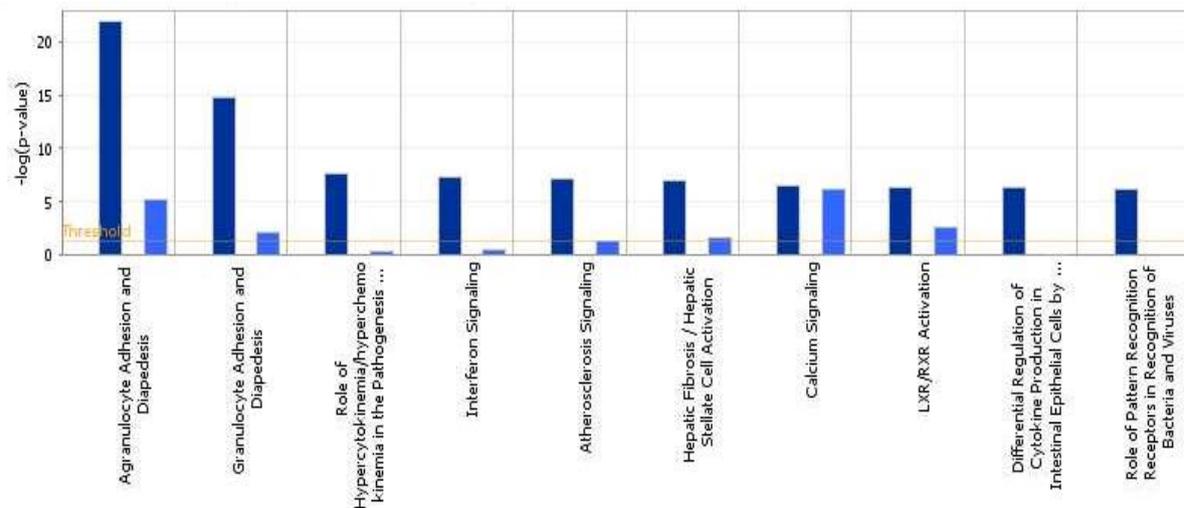
$\Delta mtp$ -mutant: genes elicited in the  $\Delta mtp$ -mutant infection

### 3.8.1 Canonical pathways enriched by all SDEGs

Pathways with  $p$ -values  $<0.05$  were considered significantly over-represented in a specific gene expression data set. Whilst the number of SDEGs elicited by the  $\Delta mtp$ -mutant infection was abundantly higher (2-fold) than the SDEGs elicited by the WT infection, the total number of canonical pathways significantly enriched by the WT genes (109) in IPA was higher than the total number of pathways enriched by the  $\Delta mtp$ -mutant infection genes (54) (Table 3.8.1).

#### 3.8.1.1 Canonical pathways enriched by both WT and $\Delta mtp$ -mutant strains

The pathways resulting from  $\Delta mtp$ -mutant infection had a  $p$ -value significantly lower than those from the WT infection (Figure 3.8.1.1.1), showed by the higher  $-\log(p\text{-values})$  that correspond to higher significance. Of the top 10 most significantly enriched pathways, 8 were common in both infections (Table 3.8.1.1.1). These were the Agranulocyte Adhesion and Diapedesis, Granulocyte Adhesion and Diapedesis, Role of Hypercytokinemia / hyperchemokine in the Pathogenesis of Influenza, Interferon Signalling, Atherosclerosis Signalling, Hepatic Fibrosis / Hepatic Stellate Cell Activation, Calcium Signalling and LXR/RXR Activation. The Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F and Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses pathways were the 2 pathways enriched by the SDEGs elicited by the WT infection from the top 10 list of pathways (Table 3.8.1.1.1). In contrast, the  $\Delta mtp$ -mutant infection SDEGs enriched the Acute Phase Response Signalling pathway and Communication between Innate and Adaptive Immune Cells pathway (Table 3.8.1.1.1).



**Figure 3.8.1.1.1: IPA of the top 10 canonical pathways enriched from all genes, in response to the WT and  $\Delta mtp$ -mutant infection relative to the uninfected.** The rest of the pathways from this comparison analysis by the IPA tool are presented in Appendix 5. Dark blue represents WT and light blue represents  $\Delta mtp$ -mutant ( $p$ -value < 0.05).

**Table 3.8.1.1.1: List of the top ten most significantly enriched canonical pathways by all genes elicited in both infections.**

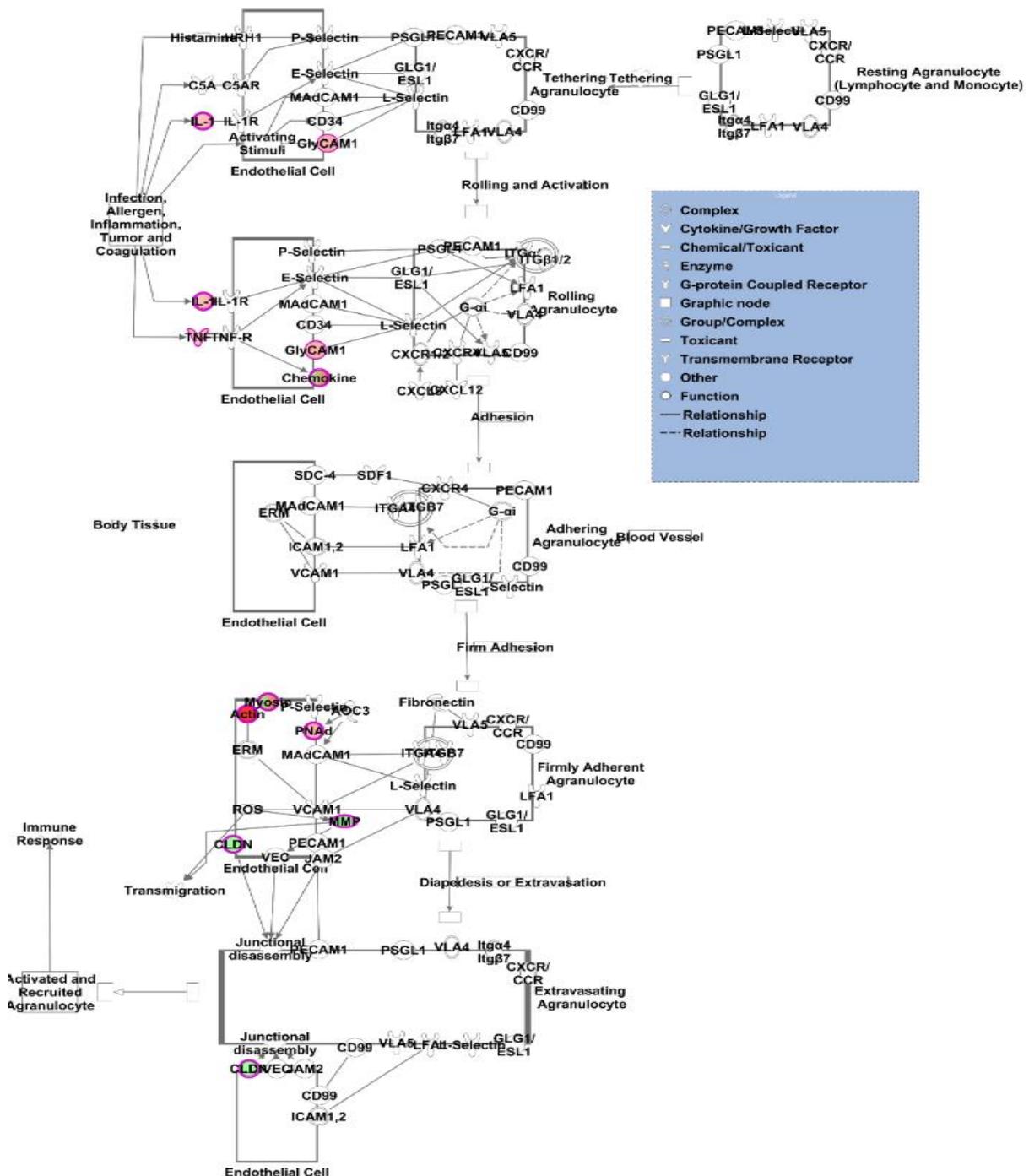
Wild type Infection		$\Delta mtp$ -mutant Infection	
Canonical Pathways	Molecules	Canonical Pathways	Molecules
<b>Agranulocyte Adhesion and Diapedesis</b>	n = 35 IL1A, MYH8, CCL17, CCL20, CCL22, Cxcl9, CXCL10, CXCL3, CCL2, CXCL13, Ccl8, CCL3L3, Ccl2, MMP8, XCL1, Cxcl3, MYL3, ACTA1, CCL19, MYH1, Cxcl11, Ppbp, MYL1, CXCL6, IL36B, MYL7, Glycam1, CCL4, CLDN5, MYH2, IL1RN, CXCL2, TNF, MMP9, Ccl7	<b>Agranulocyte Adhesion and Diapedesis</b>	n = 20 MYH4, Ppbp, MYH8, PF4, MYL1, ITGB7, L36B, SELPLG, Glycam1, MYH2, CLDN5, MMP8, IL36RN, MMP11, Ccl6, MYL3, MMP9, ACTA1, CLDN3, MYH1
<b>Granulocyte Adhesion and Diapedesis</b>	n = 27 IL1A, CCL17, CCL20, CCL22, Cxcl9, CXCL10, CXCL3, CXCL13, CCL2, Ccl8, CCL3L3, MMP8, Ccl2, XCL1, Cxcl3, CCL19, Cxcl11, Ppbp, CXCL6, IL36B, CLDN5, CCL4, IL1RN, CXCL2, TNF, MMP9, Ccl7	<b>Granulocyte Adhesion and Diapedesis</b>	n = 13 Ppbp, PF4, IL36B, SELPLG, CLDN5, MMP8, IL36RN, MMP11, IL1RAPL1, Ccl6, MMP9, CLDN3, HSPB1
<b>Role of Hypercytokinemia/hyperchemokinaemia in the Pathogenesis of Influenza</b>	n = 10 CXCL10, IFNG, IL1A, CCR5, CCL4, CCL2, IL1RN, IL12B, TNF, IL36B	<b>Role of Hypercytokinemia/hyperchemokinaemia in the Pathogenesis of Influenza</b>	n = 2 IL36RN, IL36B

<b>Pathogenesis of</b>			
<b>Influenza</b>			
<b>Interferon Signalling</b>	n = 9 IFIT3, SOCS1, IFNG, STAT2, PSMB8, STAT1, TAP1, IRF1, ISG15	<b>Interferon Signalling</b>	n = 2 IFITM3, IFITM2
<b>Atherosclerosis Signalling</b>	n = 15 PLA2G16, ALOX15, IFNG, ALB, CD40LG, ALOX12B, IL1A, CCL2, PLA2G2D, MSR1, IL1RN, S100A8, TNF, MMP9, IL36B	<b>Atherosclerosis Signalling</b>	n = 8 ALOX12B, IL36RN, ALOX12, S100A8, MMP9, IL36B, SELPLG, APOD
<b>Hepatic Fibrosis / Hepatic Stellate Cell Activation</b>	n = 18 IFNG, CCR5, IL1A, CD40LG, COL4A6, MYH8, MYL1, MYL7, CXCL3, MYH2, CCL2, IGFBP3, CD14, STAT1, TNF, MYL3, MMP9, MYH1	<b>Hepatic Fibrosis / Hepatic Stellate Cell Activation</b>	n = 12 MYH4, COL19A1, MYH2, Agtr1b, MYH8, CYP2E1, COL24A1, IL1RAPL1, MYL1, MMP9, MYL3, MYH1
<b>Calcium Signalling</b>	n = 17 CALML5, TNNI2, TNNT3, MYH8, GRIA1, TNNT2, CREB5, MYL1, ATP2A1, MYL7, MYH2, TRPV6, CASQ1, RYR1, ACTA1, MYL3, MYH1	<b>Calcium Signalling</b>	n = 21 MYH4, GRIN2A, CALML5, TP63, Calm1, TNNI2, TNNT3, MYH8, TRDN, GRIA2, TRPC4, GRIA4, Tpm2, MYL1, ATP2A1, MYH2, CASQ1, RYR1, MYL3, ACTA1, MYH1
<b>LXR/RXR Activation</b>	n = 14 C4A/C4B, ALB, IL1A, MSR1, CCL2, IL1RN, VTN, CD14, S100A8, PTGS2, NOS2, TNF, MMP9, IL36B	<b>LXR/RXR Activation</b>	n = 11 C4A/C4B, KNG1, TTR, ITIH4, IL36RN, CYP7A1, S100A8, IL1RAPL1, MMP9, IL36B, APOD
<b>Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F</b>	n = 7 IFNG, IL1A, CCL4, CCL2, IL12B, LCN2, TNF	<b>Acute Phase Response Signalling</b>	n = 9 C4A/C4B, TTR, F8, ITIH4, IL36RN, CEBPB, TCF3, CRABP2, IL36B
<b>Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses</b>	14 IFNG, IL1A, C1QC, C1QA, C1QB, OAS3, IRF7, IL12B, LTA, CASP1, CLEC6A, EIF2AK2, TNF, C3AR1	<b>Communication between Innate and Adaptive Immune Cells</b>	6 HLA-A, HLA-DRA, IL36RN, FCER1G, HLA-DRB5, IL36B

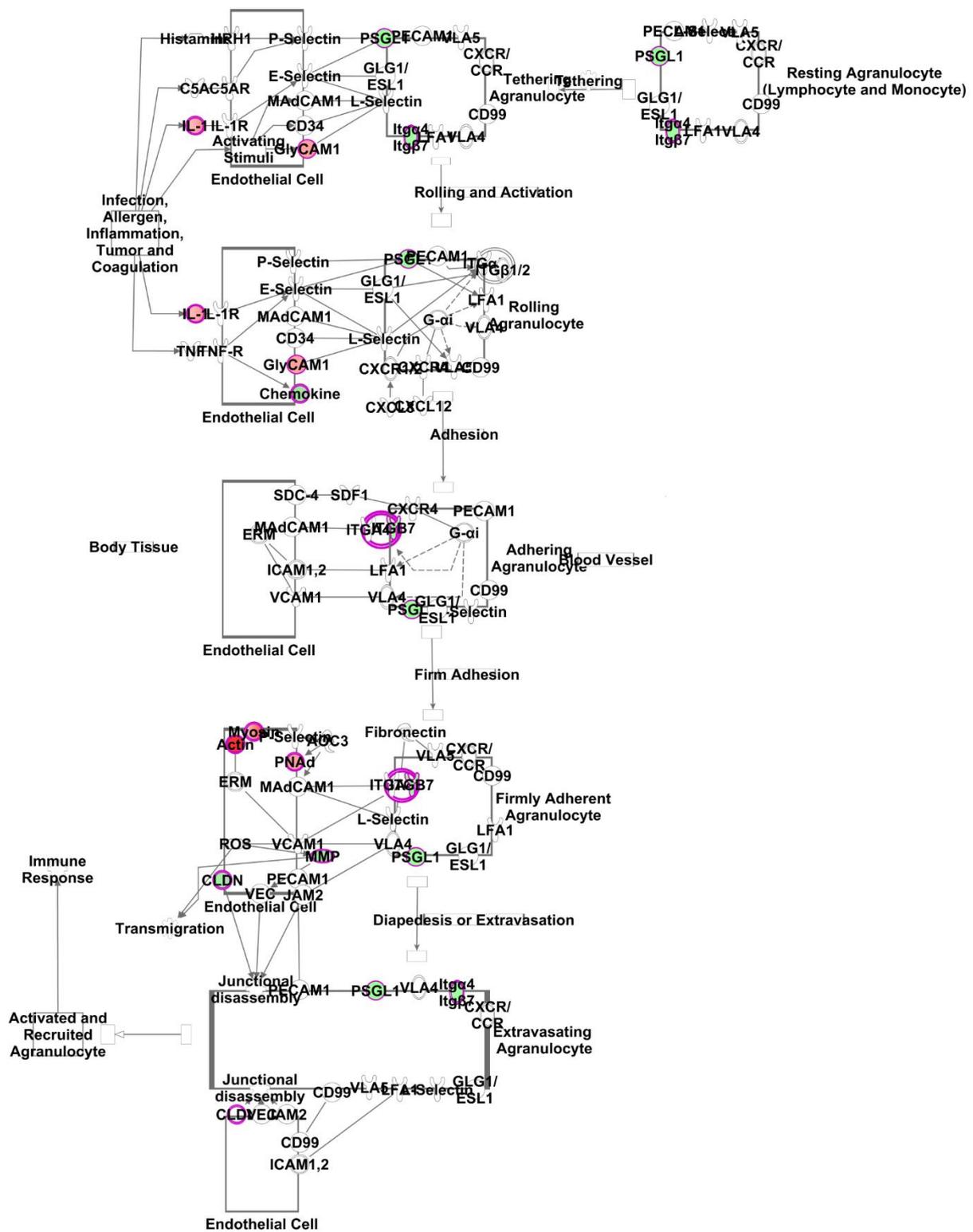
## Pathway affected by MTP

The most significant canonical pathway enriched by genes elicited in both infections was observed to be the Agranulocyte adhesion and Diapedesis pathway, with a  $p$ -value = 1.38E-22 for the WT infection compared to  $p$ -value = 7.11E-06 for the  $\Delta mtp$ -mutant infection. As expected, this pathway was lowly enriched in the  $\Delta mtp$ -mutant infection (Figure 3.8.1.1.2). This pathway is involved in Cell-To-Cell Signalling and Interaction; Haematological System Development and Function; and Immune Cell Trafficking Functions, which are part of the Cellular Immune Response.

A)



B)

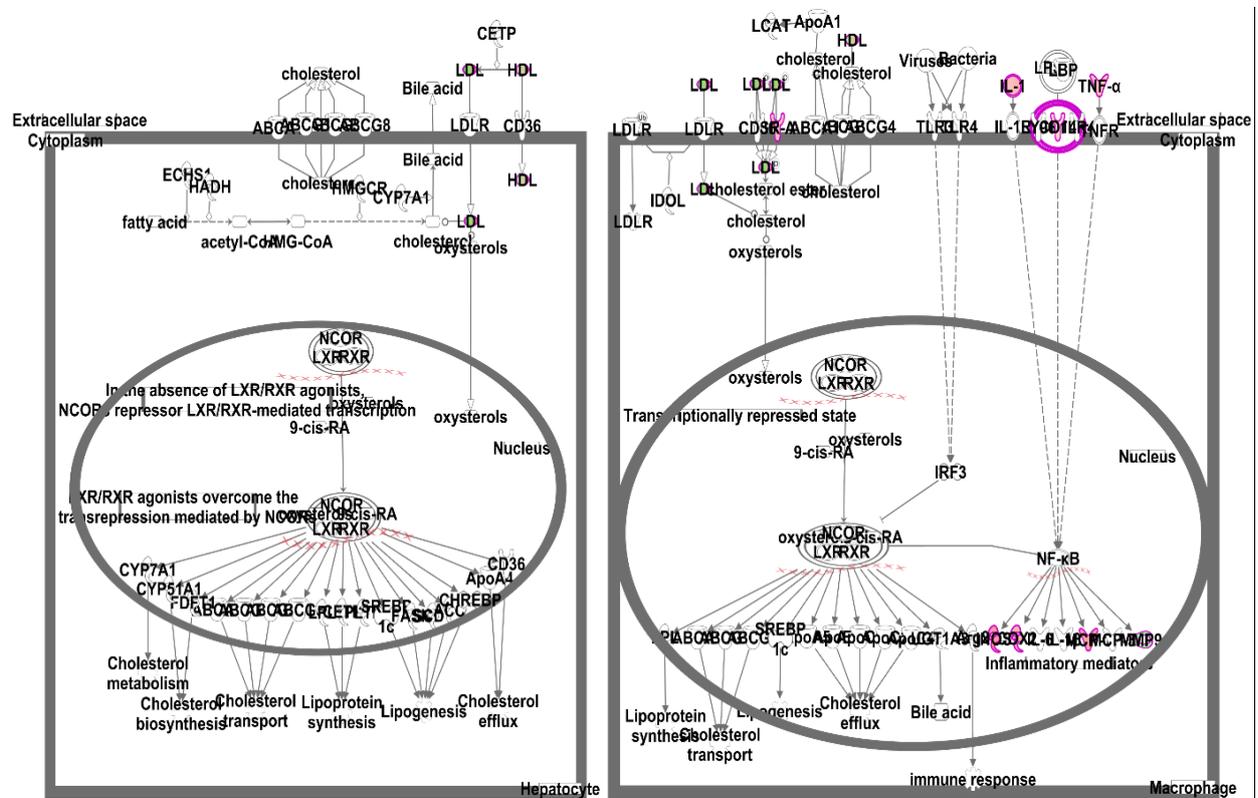


**Figure 3.8.1.1.2: Agranulocyte adhesion and Diapedesis Pathway as predicted by IPA.** A: The pathway was enriched due to the WT infection ( $p$ -value=1.38E-22). B: The enrichment of the same pathway due to the  $\Delta mtp$ -mutant infection ( $p$ -value=7.11E-06). The colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents

genes in the pathway not represented in the dataset. It is clear that the  $\Delta mtp$ -mutant infection negatively enriched the pathway (green areas), whereas the WT infection only positively enriched the same pathway.

With the exception of Calcium Signalling, the WT infection elicited more genes and molecules that were included in all the top significantly enriched IPA pathways compared to the  $\Delta mtp$ -mutant infection (Table 3.8.1.1.1). For example, the WT infection elicited 3 more genes/molecules associated with the LXR/RXR Activation compared to the  $\Delta mtp$ -mutant infection (Table 3.8.1.1.1), and it is clear that the pathway was significantly less enriched during the  $\Delta mtp$ -mutant infection compared to the WT infection as (Figure 3.8.1.1.3).

A)





7.23E-07 in WT and 3.86E-07 in  $\Delta mtp$ -mutant infection) (Figure 3.8.1.1), Graft-versus-Host Disease signalling (p-value = 5.85E-04 in WT and 1.91E-04 in  $\Delta mtp$ -mutant infection) and FXR/RXR Activation (p-value = 6.41E-03 in WT and 1.08E-02 in  $\Delta mtp$ -mutant infection).

### **3.8.2 Unique canonical pathways enriched by the SDEGs elicited by the host after WT and $\Delta mtp$ -mutant *M. tuberculosis* infection.**

Some pathways were significantly enriched by the genes from only one infection, but not the other, with (Figure 3.8.1.1). IPA core analysis of all the SDEGs showed that the genes unique to WT infection were associated with a total of 79 canonical pathways, in comparison to only 24 with  $\Delta mtp$ -mutant infection (Table 3.8.1). The pathways specific to the WT infection were associated with the host cellular immune response functions, categorized into various functional annotations (n=432). The Inflammatory Response functional category was the largest, with 106 associated functional annotations, for example, inflammation, activation, antigen presentation and phagocytosis. Cellular Movement (migration of leukocytes, T cells, antigen presenting cells, dendritic cells, phagocytes and NK cells), Cell-mediated Immune Response and Humoral Immune Response functional categories displayed in Appendix 6 are other functional categories associated with pathways specific to the WT infection.

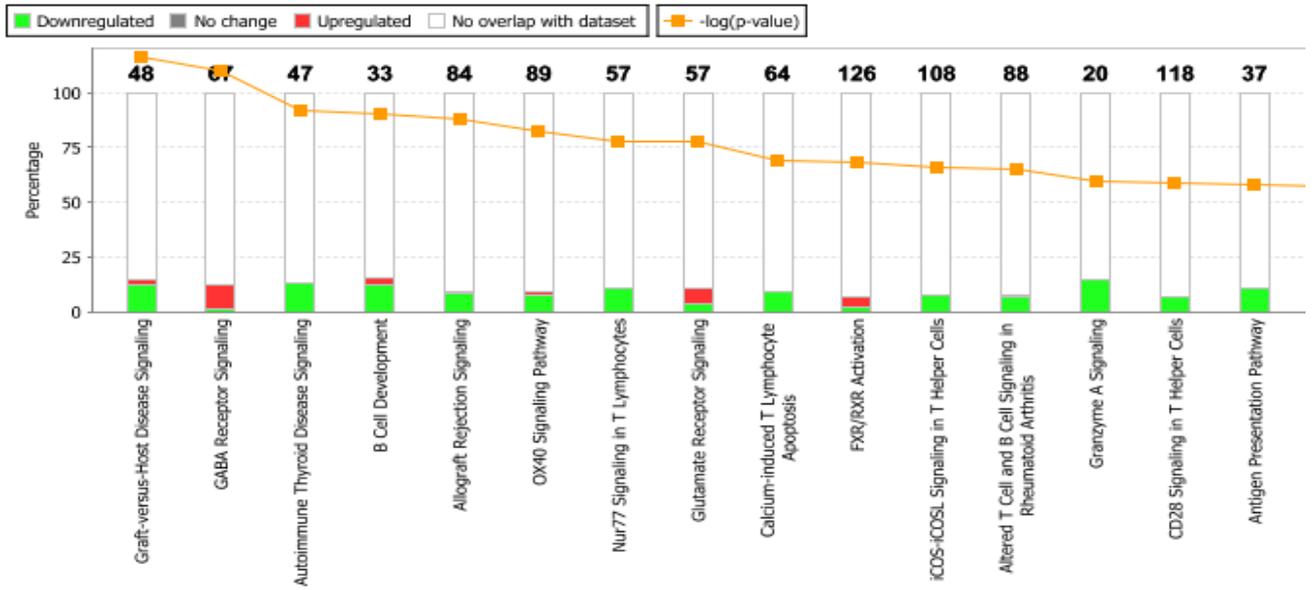
In contrast, only 50 functional annotations associated the  $\Delta mtp$ -mutant pathways were involved with immune response (Appendix 7). The Cell-To-Cell Signaling and Interaction functional category was the largest, with 28 associated functional annotations. The top 10 most significant functional annotations associated with this category were signal transduction, communication of cells, communication, neurotransmission, synaptic transmission, GABA-mediated receptor currents, density of excitatory synapses, density of synapse, binding of lung cell lines and binding of blood platelets (Appendix 7).

The most significantly enriched pathways by the WT infection SDEGs (Table 3.8.2.1) differed from those of the  $\Delta mtp$ -mutant (Figure 3.8.2.1). The most significantly affected pathways during WT infection alone included Agranulocyte Adhesion and Diapedesis (Figure 3.8.1.2), Granulocyte Adhesion and Diapedesis, Interferon Signalling, Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses (Figure 3.8.2.2), Role of Hypercytokinemia/hyperchemokinememia in the Pathogenesis of Influenza, Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F (Figure 3.8.2.3), Complement System, Activation of IRF by Cytosolic Pattern Recognition Receptors, Acute Phase Response Signaling and Dendritic Cell Maturation pathways (Table 3.8.2.1).

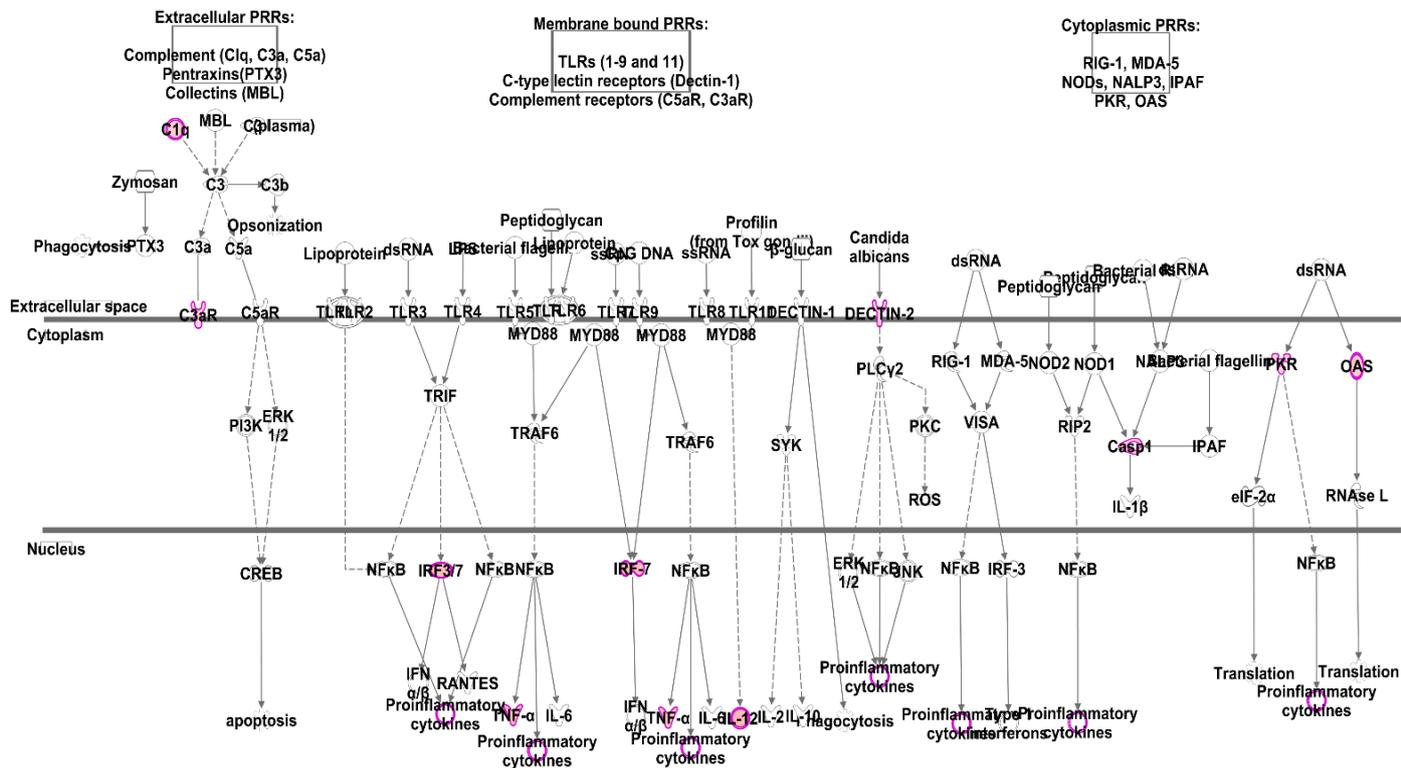
In contrast, SDEGs of the  $\Delta mtp$ -mutant infection enriched the Graft-versus-Host Disease Signaling, GABA Receptor Signalling (Figure 3.8.2.4), Auto-immune Thyroid Disease Signaling, B-Cell Development, Allograft Rejection Signaling, OX40 Signaling Pathway, Nur77 Signaling in T Lymphocytes, Glutamate Receptor Signaling (Figure 3.8.2.5), Calcium-Induced T Lymphocyte Apoptosis, and the FXR/RXR Activation pathways. Our results suggest that in the absence of MTP in the infecting bacilli, the host upregulated genes involved in other signalling pathways enriched by the genes unique to the  $\Delta mtp$ -mutant infection only (Figure 3.8.2.1). These included the GABA receptor signalling pathway, Granzyme A signalling, Signalling by Rho GTPases, Glutamate receptor signalling and FXR/RXR signalling. These pathways were not enriched after infection with the WT.

**Table 3.8.2.1: Top 10 IPA pathways significantly enriched by the WT infection SDEGs.**

<b>Ingenuity Canonical Pathways</b>	<b>-log(p-value)</b>	<b>Ratio</b>	<b>Molecules</b>
<b>Agranulocyte Adhesion and Diapedesis</b>	1.32E+01	1.22E-01	IL1A, Cxcl11, CCL17, CCL20, CCL22, Cxcl9, CXCL6, MYL7, CXCL10, CXCL3, CCL4, CCL2, CXCL13, IL1RN, Ccl8, CCL3L3, Ccl2, XCL1, Cxcl3, CXCL2, TNF, CCL19, Ccl7
<b>Granulocyte Adhesion and Diapedesis</b>	1.28E+01	1.24E-01	IL1A, Cxcl11, CCL17, CCL20, CCL22, Cxcl9, CXCL6, CXCL10, CXCL3, CCL4, CXCL13, CCL2, IL1RN, Ccl8, CCL3L3, Ccl2, XCL1, Cxcl3, CXCL2, TNF, CCL19, Ccl7
<b>Interferon Signalling</b>	8.34E+00	2.50E-01	IFIT3, SOCS1, IFNG, STAT2, PSMB8, STAT1, TAP1, IRF1, ISG15
<b>Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses</b>	7.78E+00	1.12E-01	IFNG, IL1A, C1QC, C1QA, C1QB, OAS3, IRF7, IL12B, LTA, CASP1, CLEC6A, EIF2AK2, TNF, C3AR1
<b>Role of Hypercytokinemia/hyperchemokinaemia in the Pathogenesis of Influenza</b>	7.60E+00	2.09E-01	CXCL10, IFNG, IL1A, CCR5, CCL4, CCL2, IL1RN, IL12B, TNF
<b>Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F</b>	7.26E+00	3.04E-01	IFNG, IL1A, CCL4, CCL2, IL12B, LCN2, TNF
<b>Complement System</b>	6.94E+00	2.16E-01	C1R, C4A/C4B, C1S, CFB, C1QC, C1QA, C1QB, C3AR1
<b>Activation of IRF by Cytosolic Pattern Recognition Receptors</b>	6.17E+00	1.45E-01	IRF7, LTA, NFKBIE, ZBP1, STAT2, STAT1, IFIT2, TNF, ISG15
<b>Acute Phase Response Signaling</b>	6.13E+00	8.28E-02	C1R, C4A/C4B, HMOX1, SOCS1, ALB, IL1A, IL1RN, Saa3, C1S, NFKBIE, CFB, SERPINA3, TNF, FGG
<b>Dendritic Cell Maturation</b>	5.89E+00	7.91E-02	CD1D, IL1A, CD40LG, IL12B, IL1RN, FSCN1, LTA, NFKBIE, CD86, STAT2, STAT1, CREB5, TNF, FCGR3A/FCGR3B

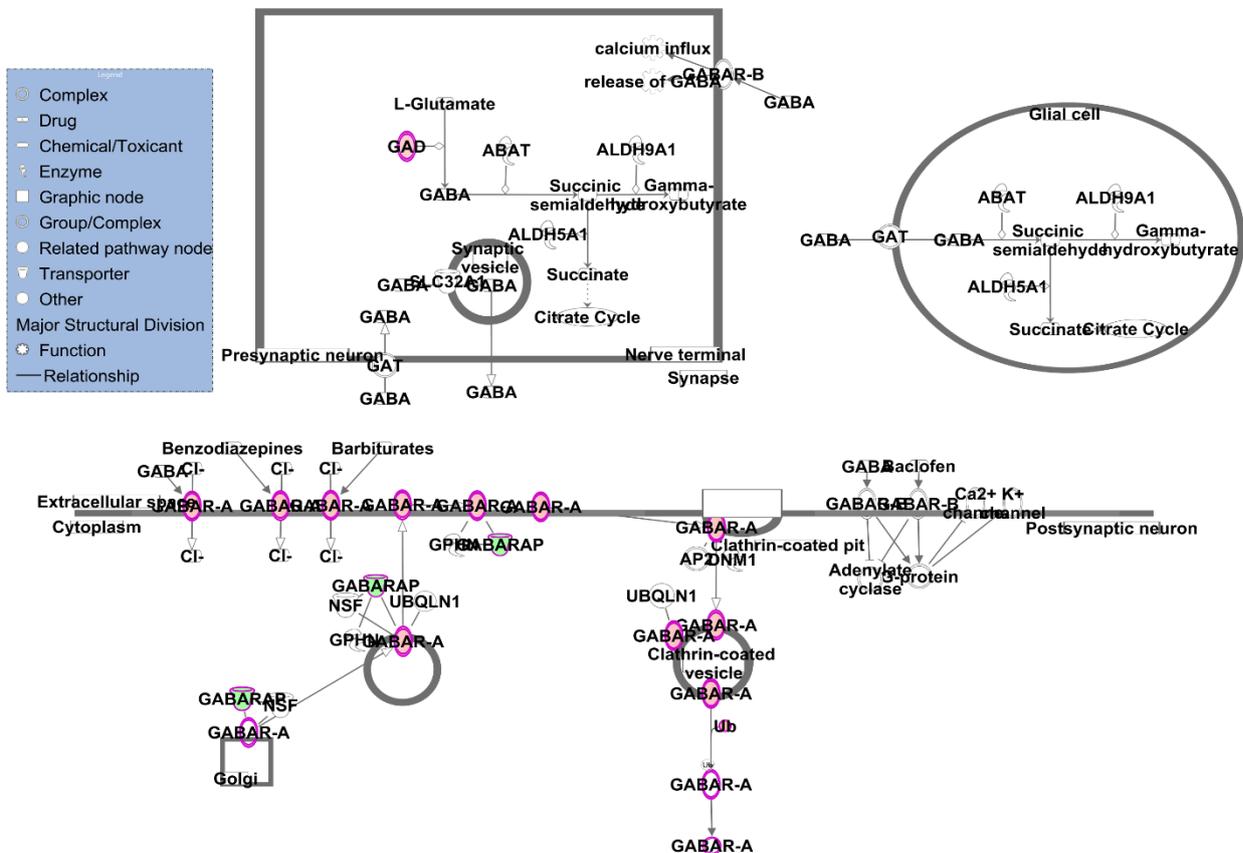


**Figure 3.8.2.1: Top canonical pathways enriched from the unique  $\Delta mtp$ -mutant infection genes as shown in IPA.** The rest of the pathways from this comparison analysis by the IPA tool are presented in Appendix 8. Red represents the genes that were up-regulated and green represents genes that were down regulated in the pathway.



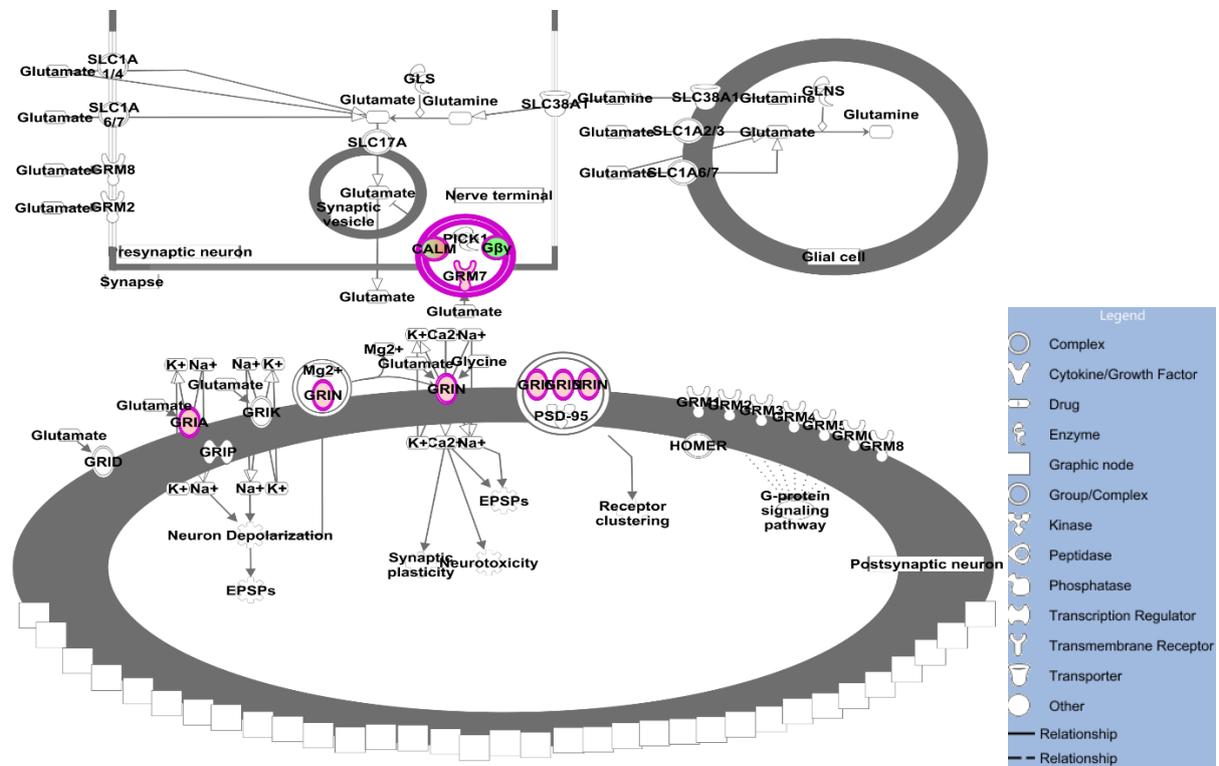


**Figure 3.8.2.2: IPA pathway and SDEGs associated with The role of Pattern Recognition Receptors (PRRs) in recognition of bacteria and viruses signalling.** The role of Pattern Recognition Receptors (PRRs) in recognition of bacteria and viruses signalling pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway.



**Figure 3.8.2.3: GABA Receptor Signalling pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the

uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the  $\Delta mtp$ -mutant infection only. The genes from the WT infection did not enrich this pathway.



**Figure 3.8.2.4: Glutamate Receptor Signaling pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the  $\Delta mtp$ -mutant infection only. The genes from the WT infection did not enrich this pathway.

### 3.8.2.1 Enrichment of canonical pathways does not necessarily lead to their activation

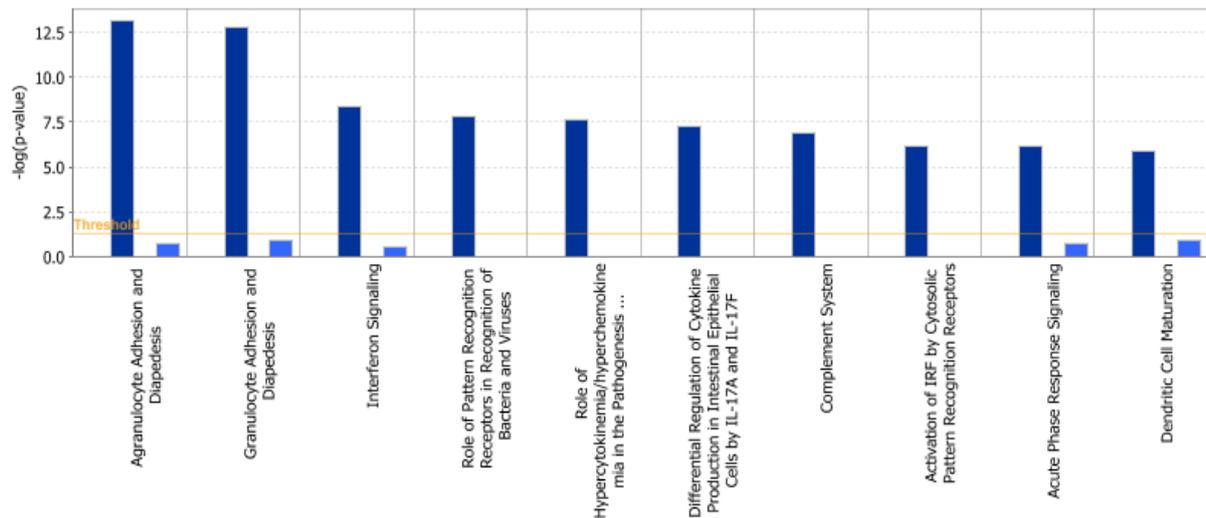
IPA analysis of transcriptomic data indicates whether pathways may be either be activated or inhibited. The activation prediction states associated with the canonical pathways enriched from the unique genes are displayed in the form of a heat map based on their z-scores (Figure 3.8.2.1.1). A positive z-score infers activation, whilst a negative z-score implies inhibition of the pathway in IPA. These results revealed that the most significantly enriched pathways, represented by the  $-\log(p\text{-values})$  (Figure 3.8.2.1.2)) were not necessarily the most activated pathways. An example is represented by the Agranulocyte adhesion and Diapedesis Pathway ( $p=1.38E-22$  and  $p=7.11E-06$  in the WT and  $\Delta mtp$ -mutant infection respectively).

Pathways resulting from the  $\Delta mtp$ -mutant infection had enrichment z-scores significantly less than pathways from the WT infection (Figure 3.8.2.1.1), resulting in a decrease in activation, or de-activation of the pathways. In contrast, the pathways enriched by the WT SDEGs exhibited positive z-scores that resulted in increased activation of the former (Figure 3.8.2.1.1). For example, the Dendritic Cell Maturation pathway was positively activated with a z-score of 3.61 in the WT infection, whilst the  $\Delta mtp$ -mutant infection deactivated the pathway with a negative activation z-score of -1.89 (Figure 3.8.2.1.1). The Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway responded in a similar manner with activation z-scores of 2.65 and -2.24 in the WT and  $\Delta mtp$ -mutant infection respectively (Figure 3.8.2.1.1).



\*Truncated pathways: Production of Nitric Oxide and Reactive Oxygen Species in Macrophages, Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses, Role of IL-17F in Allergic Inflammatory Airway Diseases, and Activation of IRF by Cytosolic Pattern Recognition Receptors.

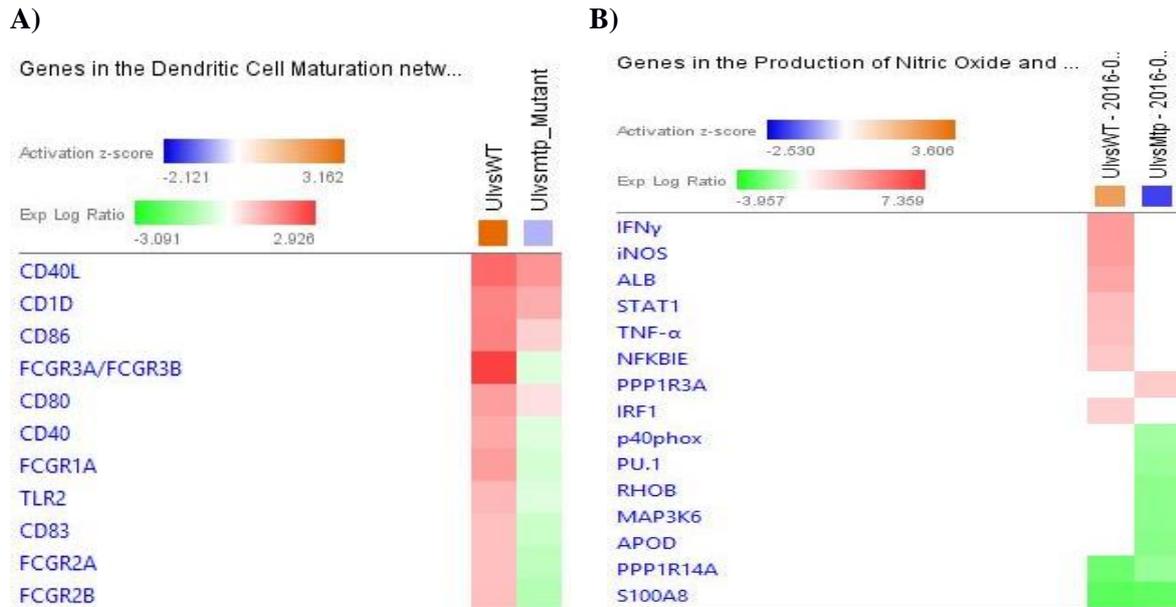
**Figure 3.8.2.1.1 Activation z-scores of the canonical pathways enriched from the unique genes induced by each strain as shown in IPA.** Colour intensity indicates the activity prediction of regulation by IPA: orange represents an overall activation increase of the pathway, and blue represents the overall repression of the pathway activity, relative to the uninfected control ( $p < 0.05$ ). UI vs WT: Wild type infection and UI vs *mtp*:  $\Delta mtp$ -mutant infection.



**Figure 3.8.2.1.2: Top canonical pathways enriched from the unique genes as shown in IPA.** The rest of the pathways from this comparison analysis by the IPA tool are presented in Appendix 9 Dark blue represents WT and light blue represents  $\Delta mtp$ -mutant ( $p$ -value  $< 0.05$ ).

### 3.8.2.1.1 Dendritic cell maturation pathway and Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathways

Exploration of the genes involved in the Dendritic cell maturation pathway (Figure 3.8.2.1.1.A) showed that all of these genes were up-regulated during infection with the WT, whereas only 4 of 11 (CD40L, CD1D, CD86 and CD80) of these genes were up-regulated during the  $\Delta mtp$ -mutant infection, and the rest, namely CD40, FCGR1A, FCGR2A, FCGR2B, FCGR3A/FCGR3B, TLR2 and CD83 were downregulated. However, the same number of different genes, 7 of 15, involved in the Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway (Figure 3.8.2.1.1.B) were upregulated and downregulated during the WT and  $\Delta mtp$ -mutant infection respectively (Figure 3.8.2.1.1.B).



**Figure 3.8.2.1.1.1: Differential expression of host genes in response to *M. tuberculosis* infection 14days post-infection as predicted by IPA.** A) Genes involved in the Dendritic cell maturation pathway, and B) Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control ( $p < 0.05$ ). UI vs WT: Wild type infection and UI vs mtp\_Mutant or UI vs Mtp:  $\Delta mtp$ -mutant infection.

### 3.8.3 Network Analysis of SDEGs

The inter-relationships among the input genes/molecules were analyzed using IPA Path Designer tool and were displayed graphically as networks. These networks were ranked by a scoring system, where the scores are equal to the negative logarithm of the calculated  $p$ -value. This indicates the probability of the SDEGs being found randomly by chance in a given network, and hence, the significance of the network [136].

#### 3.8.3.1 Network Analysis of shared SDEGs

Network analysis revealed a higher number of networks ( $n=25$ ) associated with the abundantly higher (2-fold) SDGEs involved in the  $\Delta mtp$ -mutant infection, than in the WT infection ( $n=17$ ) (Table 3.8.1). Of these, 6 networks were found to be associated with the genes shared by both infections, 2 of which were involved with different molecules (Networks 3 and 4). These different molecules resulted in different diseases and functions associated with these networks between the UI vs WT and  $\Delta mtp$ -mutant (Table 3.8.3.1.1). Network 3 lacked the calcifediol molecule in the  $\Delta mtp$ -mutant infection, and this resulted in its association with Drug Metabolism, Lipid Metabolism, Small Molecule Biochemistry, whilst the WT infection was associated with Cancer, Organismal Injury and Abnormalities, Dermatological Diseases and Conditions. Similarly, the atypical chemokine receptor 1 (ACKR1)

molecule was present in WT infection Network 4 (Figure 3.8.3.1.1A), but absent in the  $\Delta mtp$ -mutant, resulting in a difference in network structure (Figure 3.8.3.1.1B), and the functions this network was associated with in the 2 infections. Network 4 in the WT infection was associated with Cellular Movement, Haematological System Development and Function, and Immune Cell Trafficking, whilst in the  $\Delta mtp$ -mutant, it was associated with Cellular Compromise, Cell Signalling, Molecular Transport.

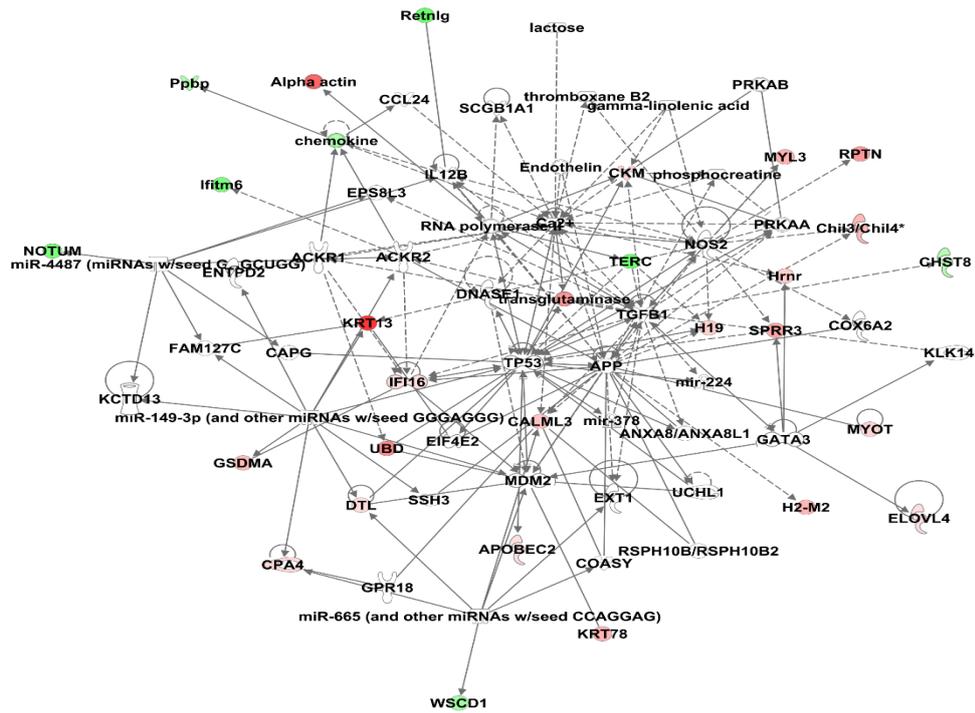
**Table 3.8.3.1.1: Diseases and functions associated with networks enriched by shared genes.**

Network	Analysis	Score	Molecules	Top Diseases and Functions
1	WT	67	35	Cancer, Dermatological Diseases and Conditions, Organismal Injury and Abnormalities
1	$\Delta mtp$ -mutant	67	35	Cancer, Dermatological Diseases and Conditions, Organismal Injury and Abnormalities
2	WT	57	31	Dermatological Diseases and Conditions, Hereditary Disorder, Organismal Injury and Abnormalities
2	$\Delta mtp$ -mutant	57	31	Dermatological Diseases and Conditions, Hereditary Disorder, Organismal Injury and Abnormalities
3	WT	54	30	Cancer, Organismal Injury and Abnormalities, Dermatological Diseases and Conditions
3	$\Delta mtp$ -mutant	52	29	Drug Metabolism, Lipid Metabolism, Small Molecule Biochemistry
4	WT	45	26	Cellular Movement, Haematological System Development and Function, Immune Cell Trafficking
4	$\Delta mtp$ -mutant	43	25	Cellular Compromise, Cell Signalling, Molecular Transport
5	WT	13	10	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry
5	$\Delta mtp$ -mutant	13	10	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry
6	WT	2	1	Organ Morphology, Organismal Survival
6	$\Delta mtp$ -mutant	2	1	Organ Morphology, Organismal Survival

$\Delta mtp$ -mutant:  $\Delta mtp$ -mutant infection

WT: Wild type infection

A)



B)

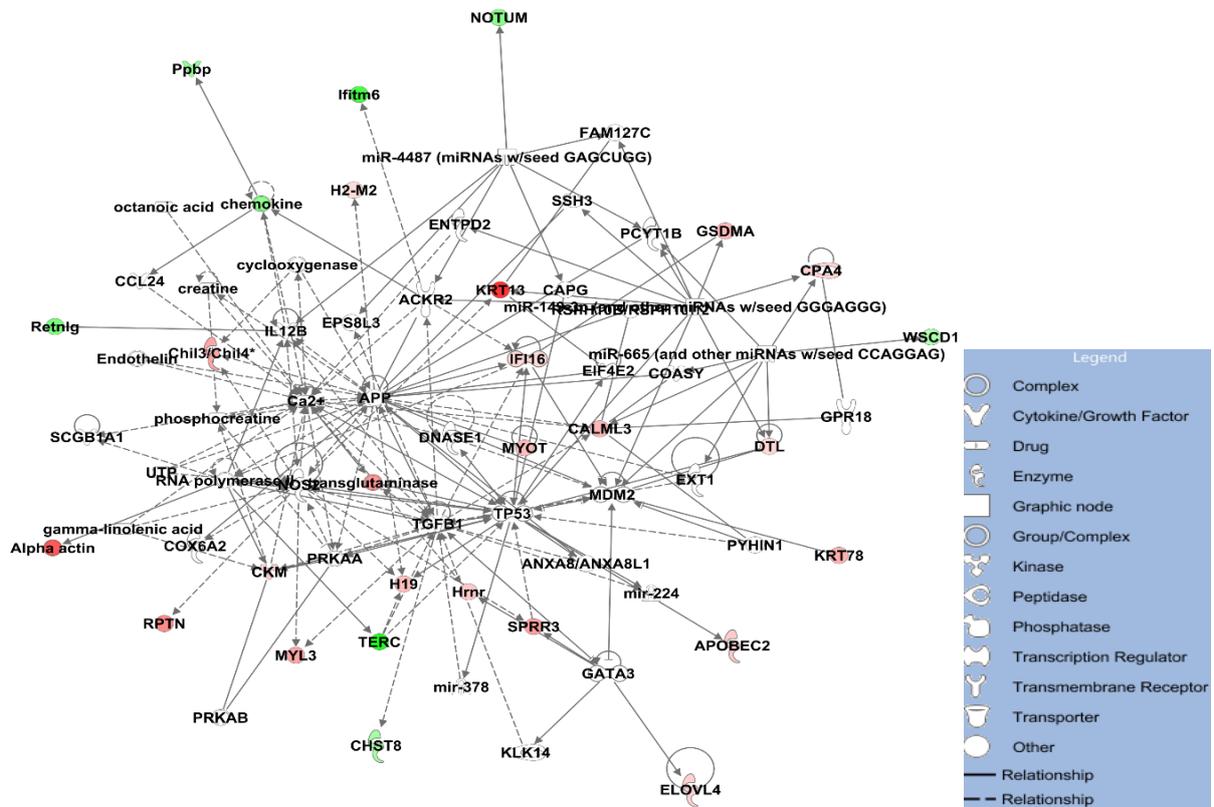


Figure 3.8.3.1 Wild type and  $\Delta mtp$ -mutant specific Network 4 enriched by shared SDEGs in IPA. A: Network 4 was enriched due to the WT infection, and B: Network 4 was enriched due to the  $\Delta mtp$ -mutant infection. IPA Path Designer tool was used for visual presentation. Solid connecting lines

represent directly connected SDEGs, and dotted lines show an indirect connection associated with the SDEGs. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control.

### 3.8.3.2 Network Analysis of unique SDEGs

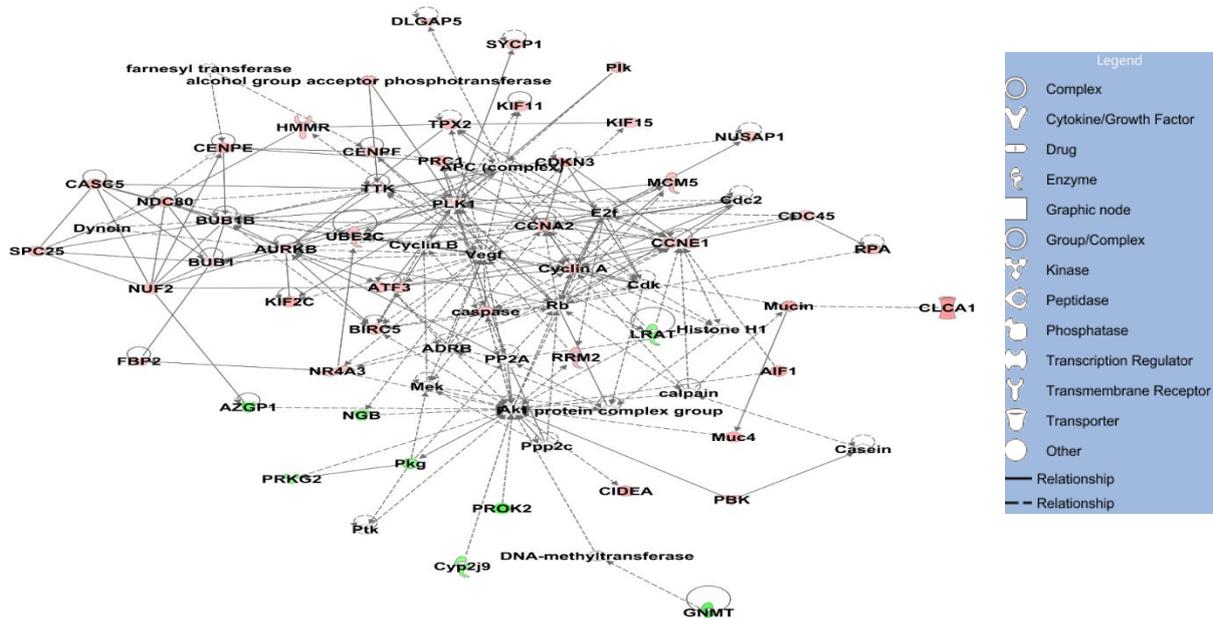
Network analysis of the unique SDEGs showed that all top 10 networks from both infections were different, with distinct diseases and their associated functions (Table 3.8.3.2.1). The 2 highest scoring networks (scores = 66 and 64) induced by the WT strain were associated with Gastrointestinal and Immunological Disease (Figure 3.8.3.2.1A), and Antimicrobial Response, Inflammatory Response and Organismal Injury (Figure 3.8.3.2.1B) respectively. The 2 highest scoring networks (scores = 75 and 73) induced by the  $\Delta mtp$ -mutant strain were associated with Cell-To-Cell Signaling and Interaction, Cellular Function and Maintenance (Figure 3.8.3.2.1C), and Hereditary Disorder, Organismal Injury (Figure 3.8.3.2.1D) respectively.

**Table 3.8.3.2.1. Top diseases and functions associated with unique genes from top 10 IPA networks.**

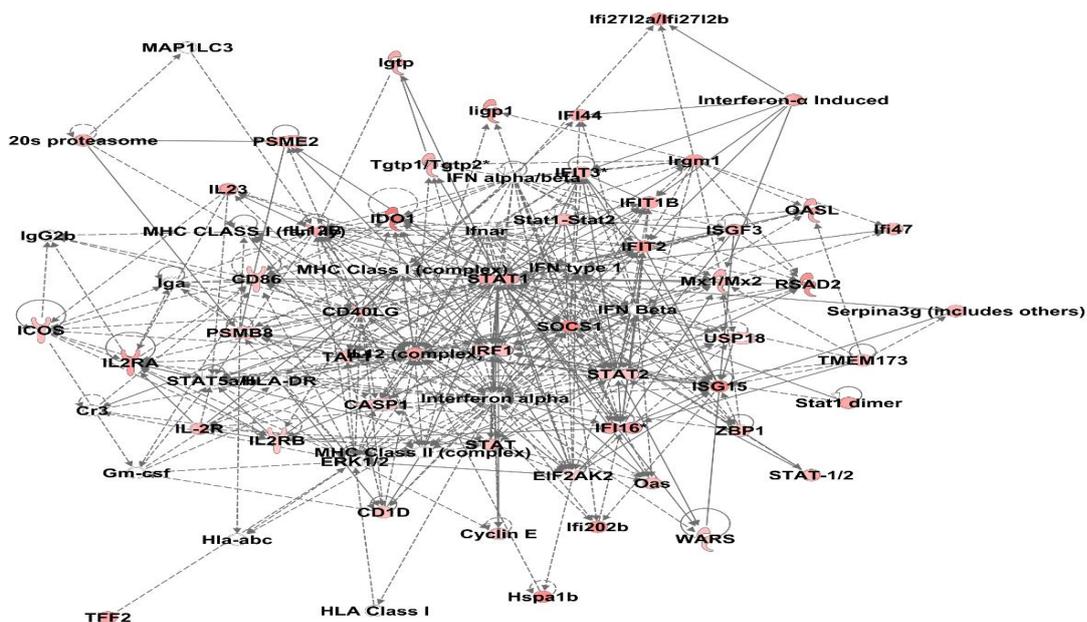
Wild type infection			$\Delta mtp$ -mutant infection		
ID	Score	Top Diseases and Functions	ID	Score	Top Diseases and Functions
1	66	Gastrointestinal Disease, Immunological Disease	1	75	Cell-To-Cell Signaling and Interaction, Cellular Function and Maintenance
2	64	Antimicrobial Response, Inflammatory Response,	2	73	Hereditary Disorder, Organismal Injury
3	62	Cell Cycle, Cellular Assembly and Organization,	3	53	Cell-To-Cell Signaling and Interaction, Nervous System Development and Function
4	50	Tissue Morphology, Organismal Injury and Abnormalities	4	51	Skeletal and Muscular System Development and Function,
5	48	Cell-To-Cell Signaling and Interaction, Cellular Movement, Hematological System Development and Function	5	51	Carbohydrate Metabolism, Small Molecule Biochemistry, Lipid Metabolism
6	48	Cell Morphology, Embryonic Development, Hair and Skin Development and Function	6	42	Cell Cycle, Cellular Growth and Proliferation, Tissue Development
7	41	Inflammatory Response, Cellular Movement, Hematological System Development and Function	7	40	Cellular Assembly and Organization, Cellular Movement, Neurological Disease

8	41	Humoral Immune Response, Protein Synthesis, Cellular Movement	8	40	Embryonic Development, Organ Development, Organismal Development
9	41	Gastrointestinal Disease, Organismal Injury and Abnormalities, Post-Translational Modification	9	39	Embryonic Development, Tissue Development, Cellular Growth and Proliferation
10	37	Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking	10	39	Hereditary Disorder, Neurological Disease, Organismal Injury and Abnormalities

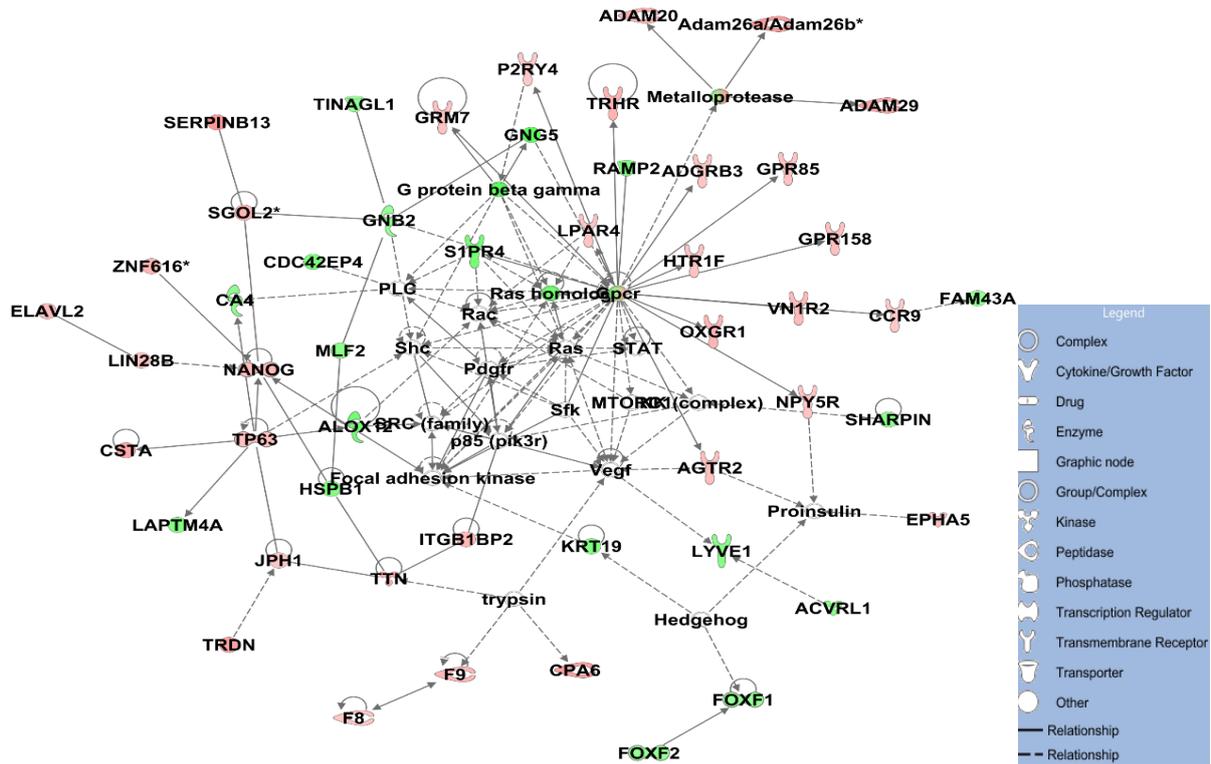
A)



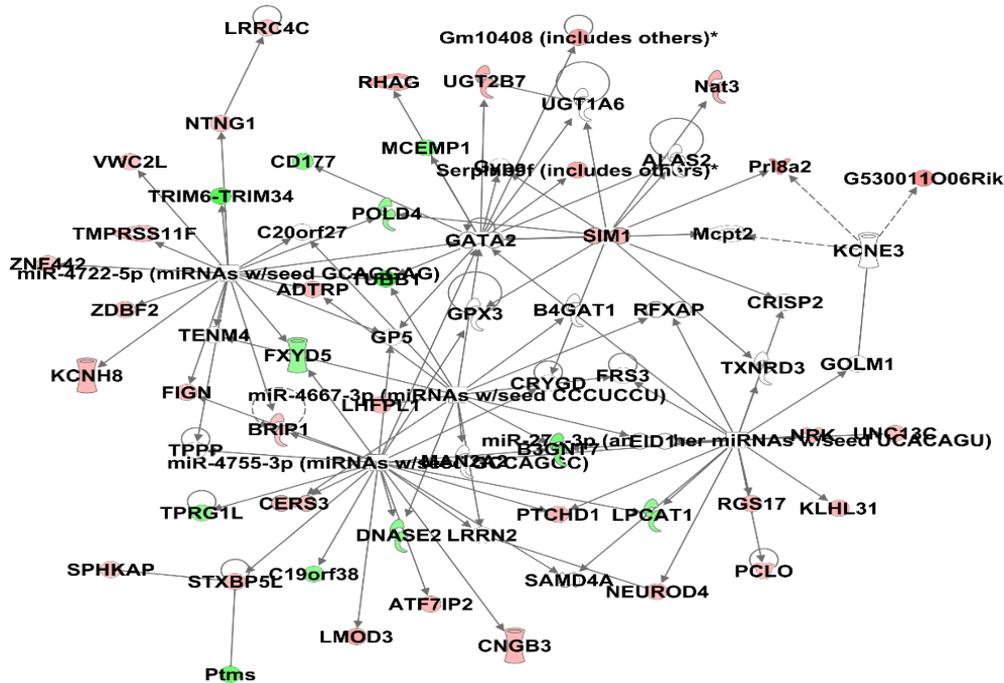
B)



C)



D)



**Figure 3.8.3.2.1: Top 2 Wild type and  $\Delta mtp$ -mutant specific Networks enriched by SDEGs in IPA.**

A and B: Top 2 networks enriched due to the WT infection. C and D: Top 2 networks enriched due to the  $\Delta mtp$ -mutant infection. IPA Path Designer tool was used for visual presentation. Solid connecting lines represent directly connected SDEGs, and dotted lines show an indirect connection associated with

the SDEGs. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control.

### **3.8.4 The effect of MTP on specific host-pathogen interactions**

#### **3.8.4.1 Pattern Recognition Receptors**

The Pattern Recognition Receptors in recognition of bacteria and viruses signalling pathway (Figure 3.8.2.2) ( $p= 6.01E-07$ ) was ranked as the 5<sup>th</sup> most activated pathway. Analysis of the genes involved in this pathway showed that these were up-regulated only during infection with the WT, but not the  $\Delta mtp$ -mutant strain (Figure 3.8.4.1.1.A). Thus, this pathway was enriched in the WT infection only, and not in the  $\Delta mtp$ -mutant infection (Figure 3.8.1.1.1), and is involved in the following signalling pathway categories: Cellular Immune Response; Pathogen-Influenced Signalling, and associated with the following functions: Infectious Diseases; Inflammatory Response; Antimicrobial Response.

##### **3.8.4.1.1 Toll-like receptors (TLRs), and Complement system receptors**

The PRRs that have been associated with *M. tuberculosis* infection [81] include Toll-like receptors (TLRs), and receptors of the complement system. Analysis of the genes involved in the TLR signalling pathway showed that 8 of these 9 genes were up-regulated during infection with the WT, whereas only 3 (*UBD*, *IL36B* and *IL36RN*) were up-regulated during the  $\Delta mtp$ -mutant infection (Figure 3.8.4.1.1.B) in this study. This resulted in the enrichment of the TLR signalling pathway only during WT infection (Figure 3.8.4.1.1.D), and the transcription factors (Figure 3.8.4.1.1.C) and cytokines (Figure 3.8.4.1.1.D) elicited by this pathway were enriched only in the WT infection). Likewise, the pathway associated with the complement system (Complement Signalling pathway,  $p = 1.06E-06$ ) utilised by host cells to engulf infecting bacteria was not activated during infection with the MTP-deficient strain (Figure 3.8.4.1.1.2).

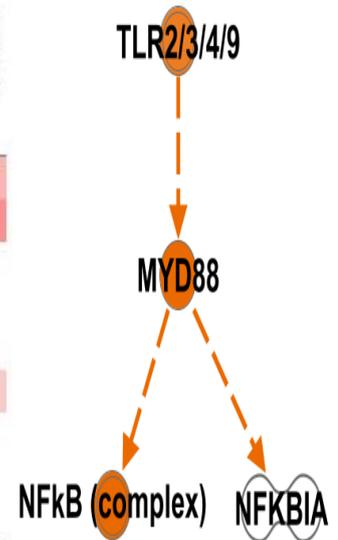
A)



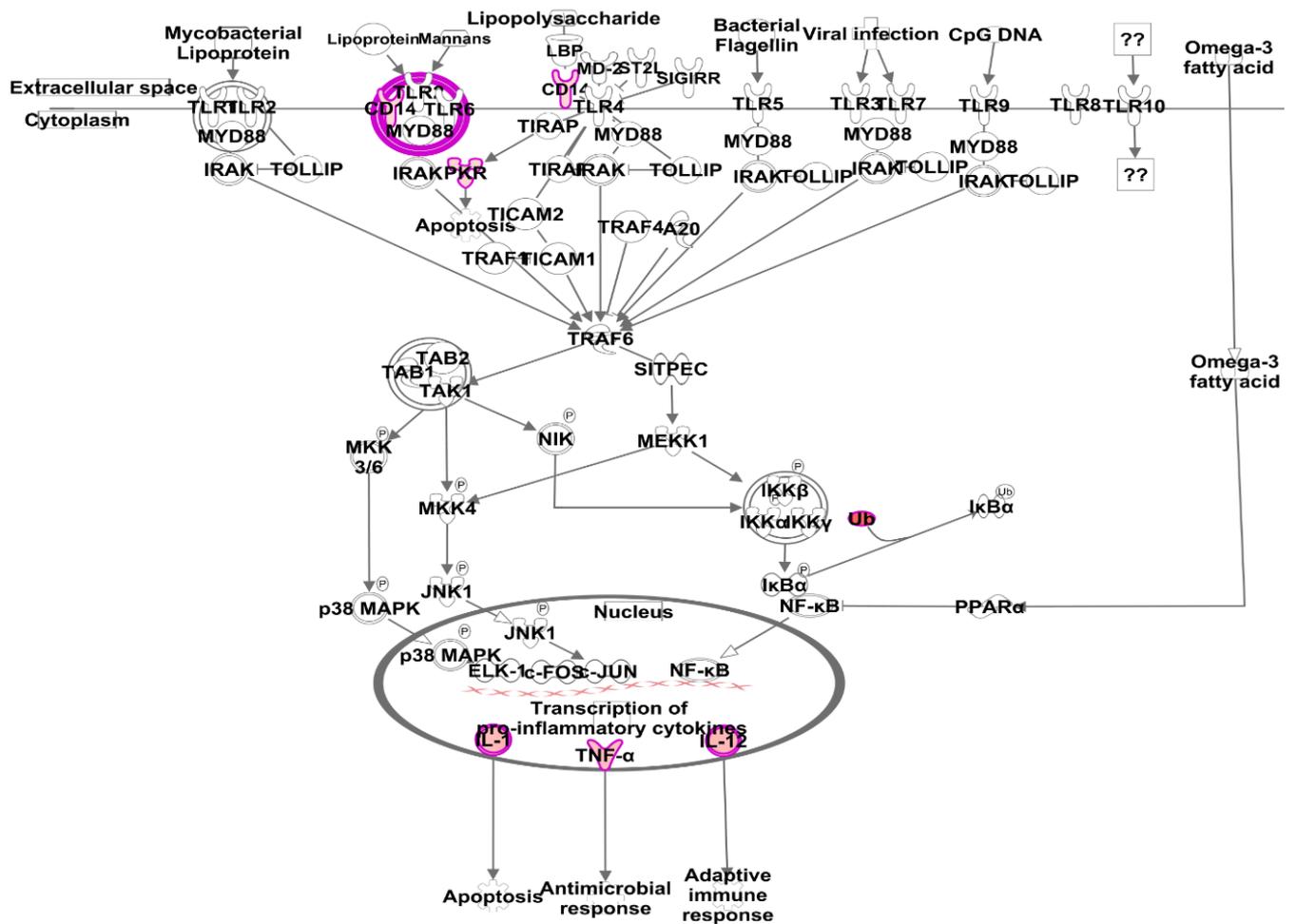
B)

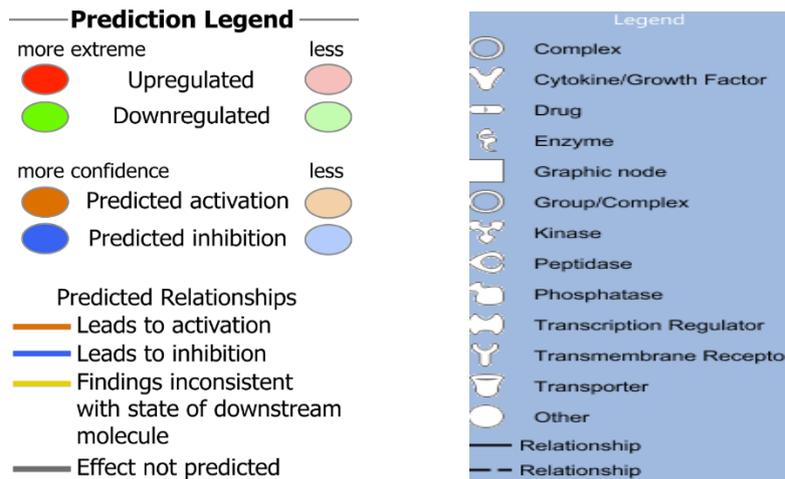


C)



D)





**Figure 3.8.4.1.1.1: Differential expression of host genes in response to *M. tuberculosis* infection 14days post-infection as predicted by IPA.** A) Genes involved in the Pattern Recognition Receptors in recognition of bacteria and viruses signalling pathway, and B) Genes involved in the Toll-like receptor signalling pathway. UI vs WT: Wild type infection and UI vs Mtp:  $\Delta mtp$ -mutant infection. C) TLR 2/3/4/9 Mechanistic pathway of upstream regulators and D) Toll-like receptor signalling pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control ( $p < 0.05$ ). Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset.

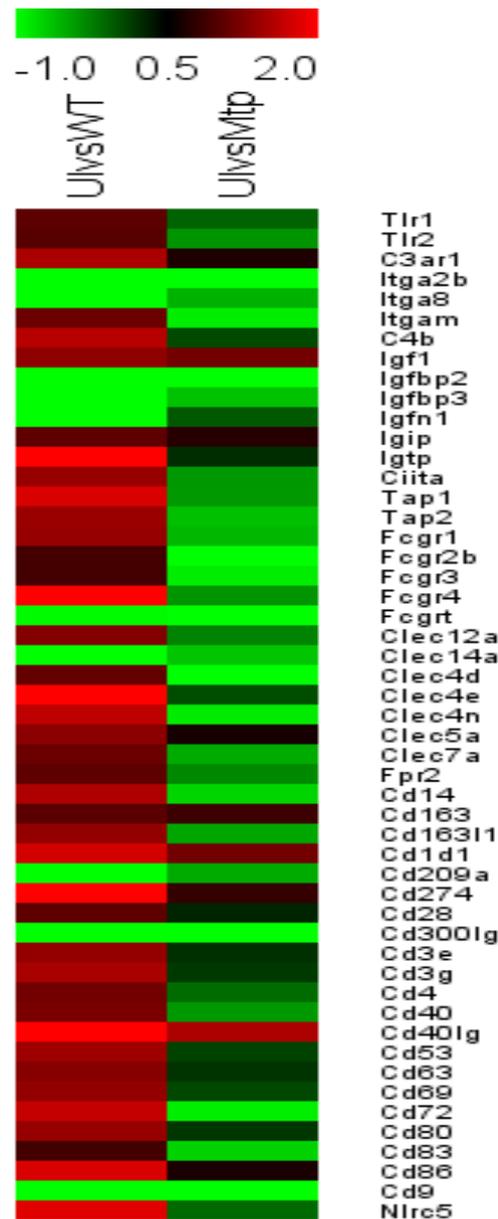


only 10 were downregulated in both infections (*ITGA2B*, *ITGA8*, *IGFBP2*, *IGFBP3*, *IGFN1*, *FCGRT*, *CLEC14A*, *CD209A*, *CD300LG* and *CD9*). The top 10 most upregulated and down regulated genes from this gene set are exhibited in Table 3.8.4.2.1. The most upregulated gene was the *CD274* gene in the WT infection with a log<sub>2</sub>(fold change) value of 2.9262 compared to 0.812163 for the *Δmtp*-mutant infection. The *Δmtp*-mutant strain caused the most downregulation in these genes, with the *ITGA2B* gene having a log<sub>2</sub>(fold change) value of -3.0907.

Out of the selected 51 genes associated with host-pathogen interaction (Appendix 10), 49 clustered together into 8 gene families (Figure 3.8.4.2.1), the C-type lectin domain family, toll like receptor family, integrin subunit alpha family, Fc fragment of IgG receptor family, Insulin Like Growth Factor family, Transporter associated with Antigen Processing family, CD gene family and the complement system family of genes. The largest family, CD gene family, contained 21 genes: *CD14*, *CD163*, *CD163L1*, *CD1D1*, *CD209A*, *CD274*, *CD28*, *CD300LG*, *CD3E*, *CD3G*, *CD4*, *CD40*, *CD40LG*, *CD53*, *CD63*, *CD69*, *CD72*, *CD80*, *CD83*, *CD86* and *CD9*. However, all 51 genes are associated and belong to the “cluster of differentiation molecules” group.

**Table 3.8.4.2.1: Top 10 most upregulated and down regulated genes associated with host-pathogen interactions.**

Upregulated genes			Downregulated genes		
Gene ID	WT	Mtp	Gene ID	WT	Mtp
<b>Cd274</b>	2.9262	0.812163	<b>Igfbp3</b>	-1.69101	-0.64387
<b>Fcgr4</b>	2.73725	-0.38016	<b>Cd300lg</b>	-1.59795	-1.16734
<b>Igtp</b>	2.39848	0.23047	<b>Igfbp2</b>	-1.4404	-1.42984
<b>Clec4e</b>	2.23019	0.04585	<b>Clec14a</b>	-1.42536	-0.66403
<b>Cd40lg</b>	2.14313	1.52679	<b>Itga8</b>	-1.32416	-0.54906
<b>Nlrc5</b>	1.83406	-0.12355	<b>Cd209a</b>	-1.2476	-0.50808
<b>Tap1</b>	1.78121	-0.40511	<b>Igfn1</b>	-1.2349	-0.0225
<b>Cd86</b>	1.7798	0.661663	<b>Cd9</b>	-1.10082	-1.49473
<b>Cd1d1</b>	1.75068	1.17743	<b>Itga2b</b>	-1.06062	-3.0907
<b>Cd72</b>	1.66875	-0.90293	<b>Fcgrt</b>	-0.97318	-1.18799



**Figure 3.8.4.2.1: MeV heat map of differential expression of host genes involved in the host-pathogen interaction, in response to *M. tuberculosis* infection 14days post-infection.** The 51 host-pathogen interaction associated genes shared by both the WT and  $\Delta mtp$ -mutant infections of Balb/C mice lungs clustered into 8 families. UIvsWT: Wild type infection and UIvsMutant:  $\Delta mtp$ -mutant infection. Colour intensity indicates the magnitude of up-regulation: red represents highly activated genes, and green represents lowly activated genes, relative to the uninfected control ( $p < 0.05$ ). The level of regulation (up- or downregulation) was higher in the WT infection genes compared to  $\Delta mtp$ -mutant infection genes.

### 3.8.4.2.1 Top 5 canonical pathways associated with host-pathogen interaction SDEGs

IPA analysis revealed a total of 52 canonical pathways that were enriched from the host-pathogen interaction genes listed in (Appendix 10). All 5 of the high-ranking canonical pathways were associated

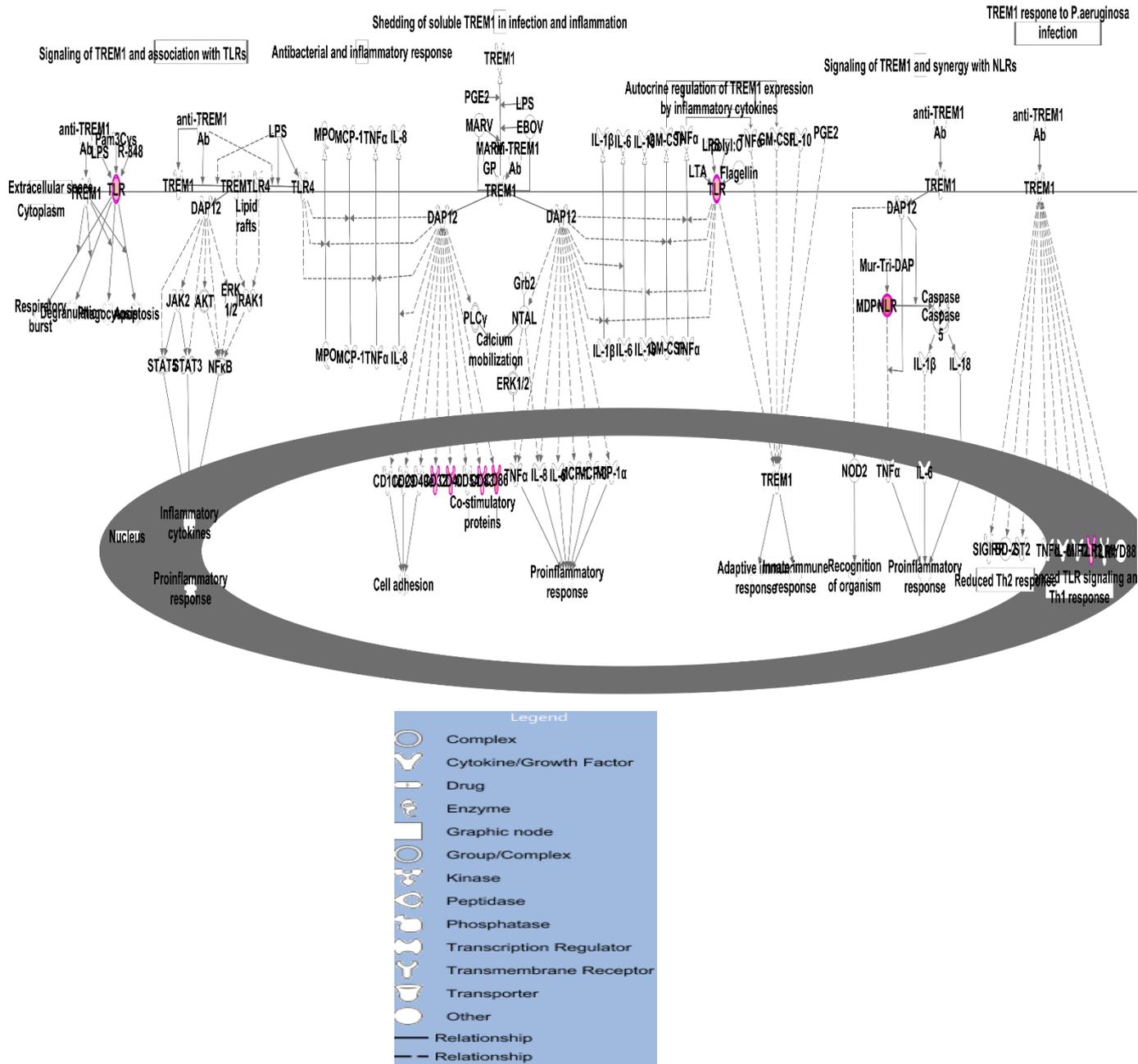
with the host immune response. Whilst the WT infection resulted in a higher enrichment of these pathways, 4 of them were lowly enriched by the  $\Delta mtp$ -mutant infection (Figure 3.8.4.1.1), suggesting that MTP plays a significant role in host response.



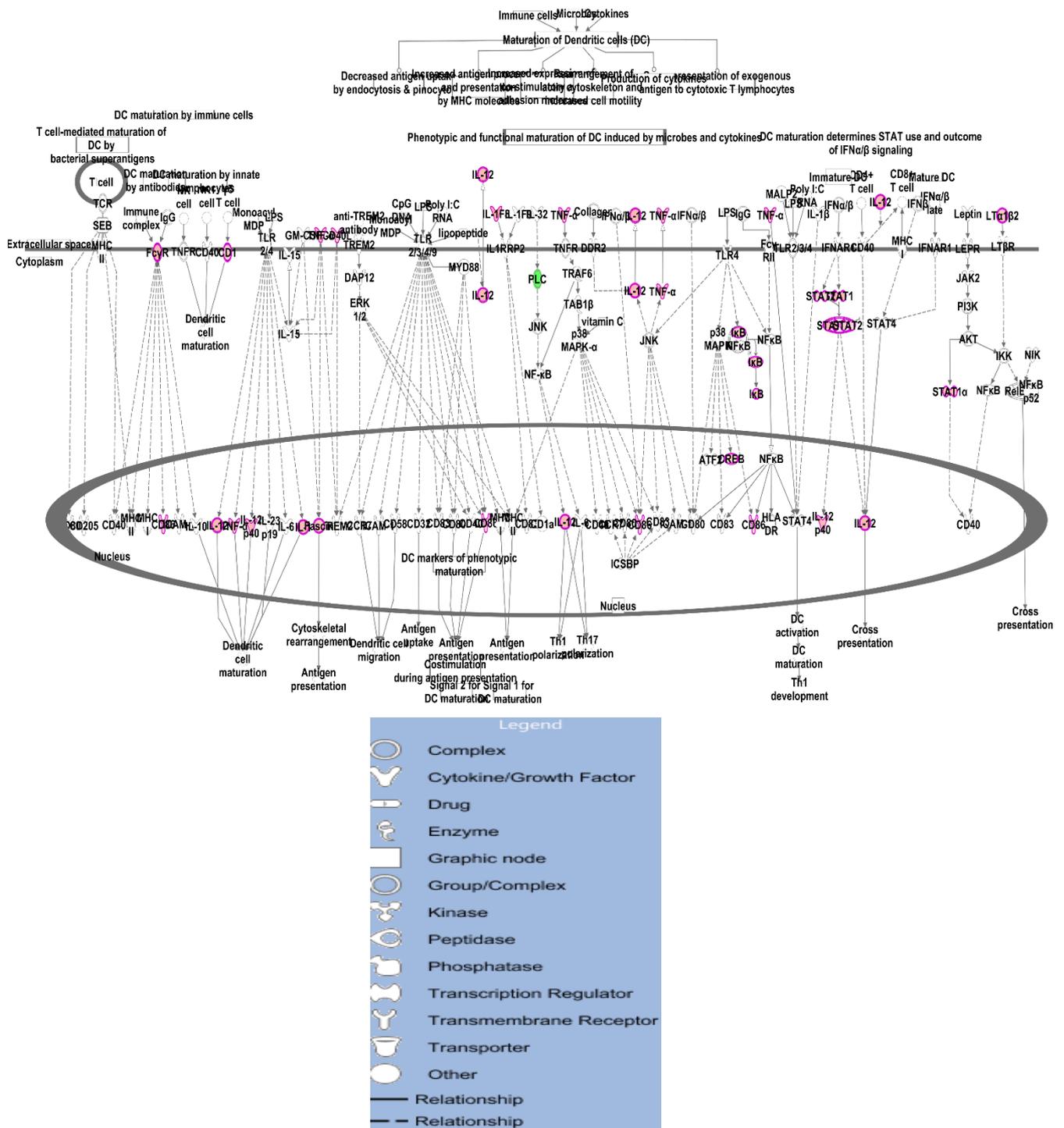
\*line 3: Role of NFAT in Regulation of the Immune Response

**Figure 3.8.4.2.1.1: The top 5 canonical pathways associated with the host-pathogen interaction genes as predicted by IPA.** The heat map shows a comparison of z-score for the 5 top pathways involved in host-pathogen interaction; namely TREM1 Signalling, Dendritic Cell Maturation, Role of NFAT in Regulation of the Immune Response, NF- $\kappa$ B Signalling and iCOS-iCOSL Signalling in T Helper Cells. UivsWT: Wild type infection and Uivsntp\_Mutant:  $\Delta mtp$ -mutant infection. The WT strain resulted in positive differential regulation (shown by orange bars) of all 5 pathways. In contrast, the  $\Delta mtp$ -mutant strain elicited a negative differential regulation in most of these pathways (shown by blue bars).

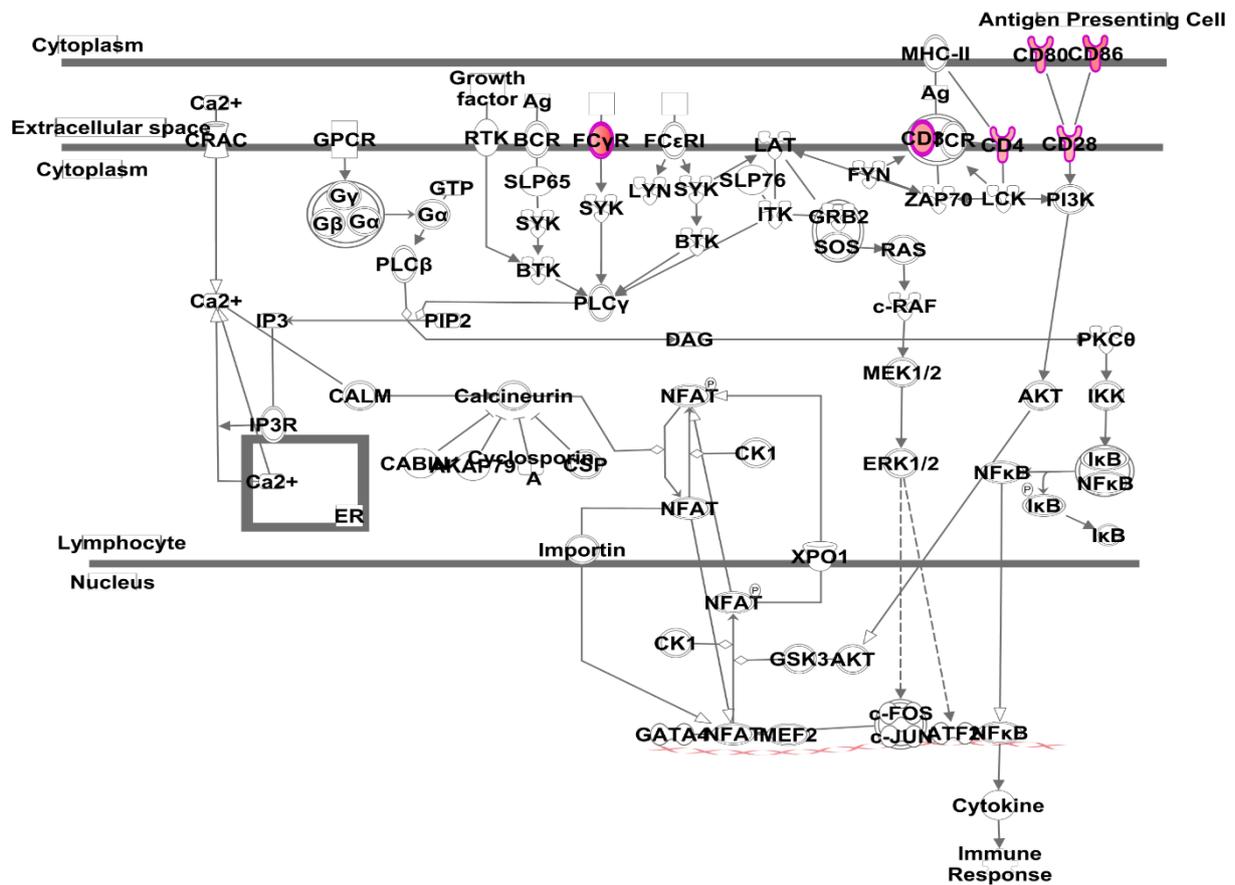
IPA predicted an activation z-score of 2.83 and -2.12 for the enrichment of the TREM1 Signaling pathway (3.8.4.2.1.2A) by the WT and the  $\Delta mtp$ -mutant genes respectively. The Dendritic Cell Maturation pathway (Figure 3.8.4.2.1.2B) showed a similar trend in its activation as previously mentioned (Figure 3.8.2.6), as did the NF- $\kappa$ B Signaling pathway (Figure 3.8.4.2.1.2C), with z-scores of 2.0 and -1.0 for the WT and  $\Delta mtp$ -mutant infection respectively. The role of NFAT in regulation of the immune response (Figure 3.8.4.2.1.2D) pathway with the same activation score as the Dendritic Cell Maturation pathway in the WT infection was neither activated nor deactivated, (z-score of 0.0) during infection with the  $\Delta mtp$ -mutant strain. The iCOS-iCOSL Signaling in T Helper Cells pathway (Figure 3.8.4.2.1.2E) was the only pathway that was activated in both infections, with the WT and  $\Delta mtp$ -mutant infections resulting in z-scores of 2.24 and 0.45 respectively. The magnitude of activation was still greater in the WT infection (Figure 3.8.4.2.1.1).



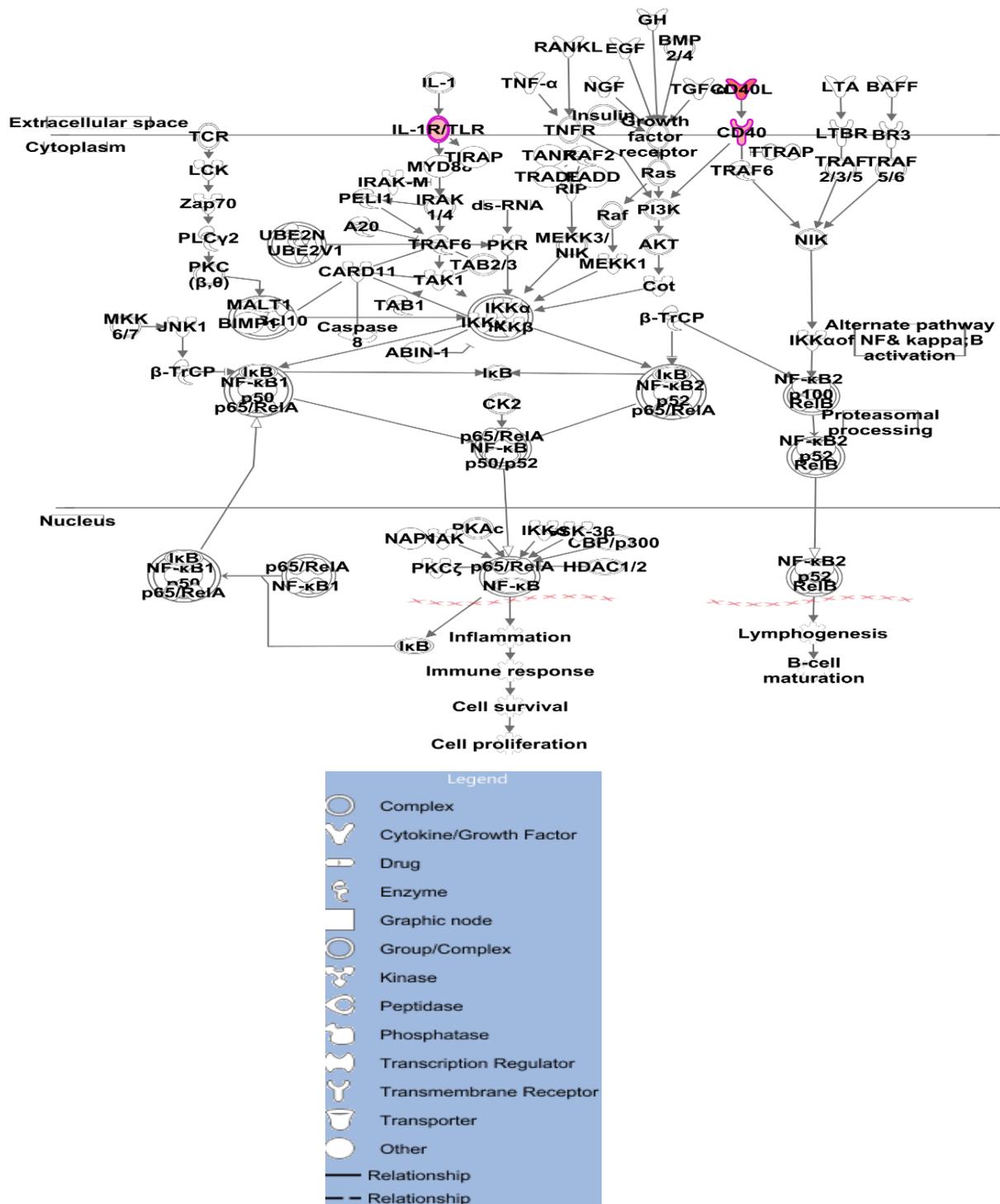
**Figure 3.8.4.2.1.2A: TREM1 Signalling pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway.



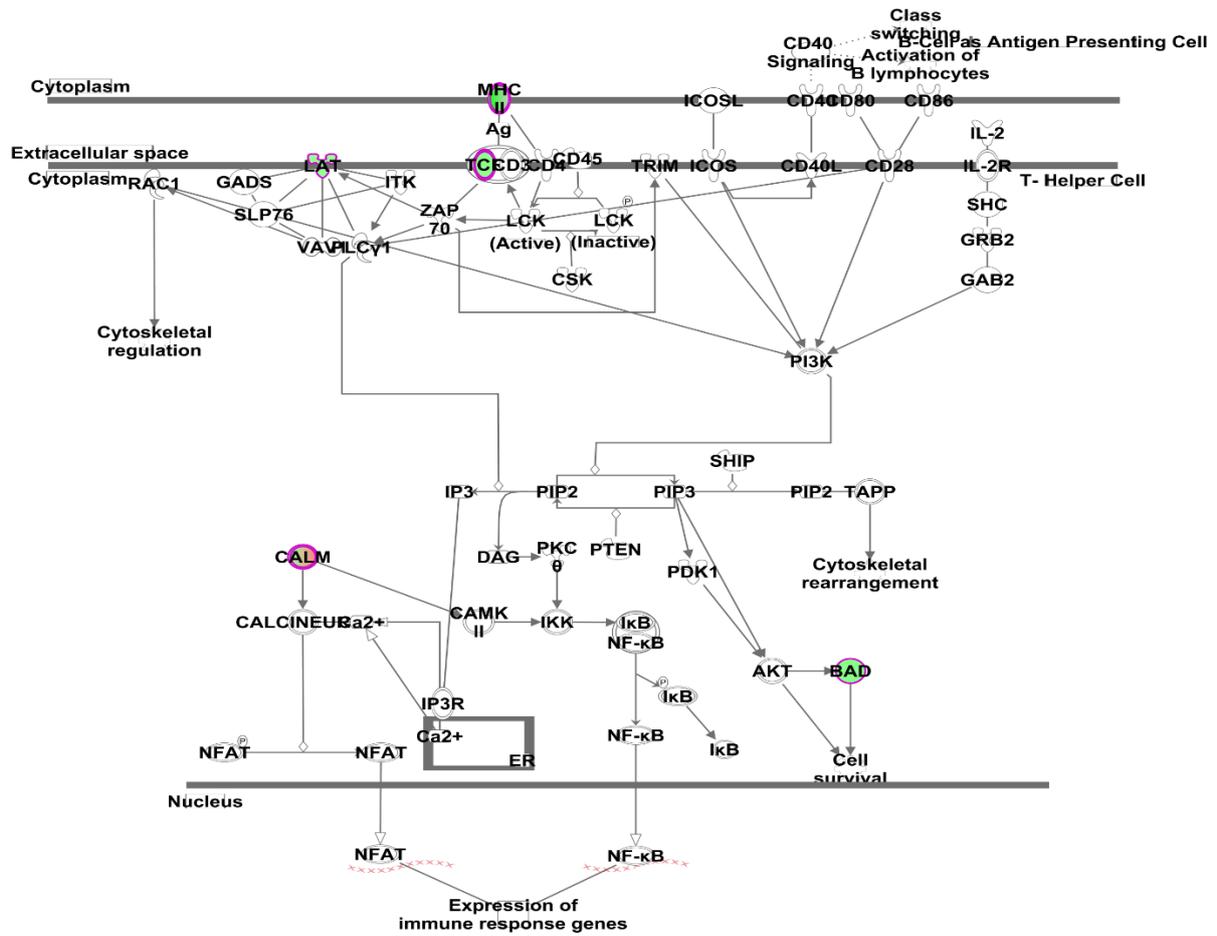
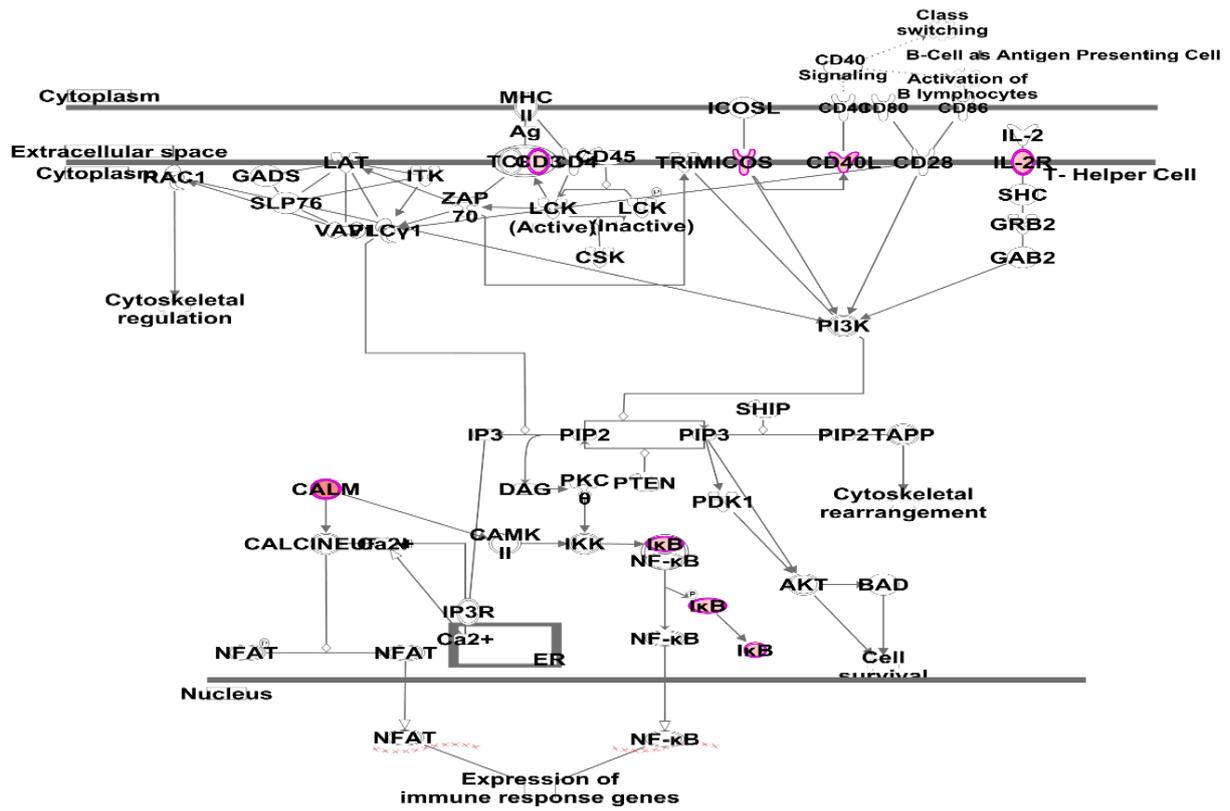
**Figure 3.8.4.2.1.2B: The Dendritic Cell (DC) Maturation pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway.



**Figure 3.8.4.2.1.2C: The Role of NFAT in Regulation of the Immune Response pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway.



**Figure 3.8.4.2.1.2D: The NF-κB Signalling pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway.

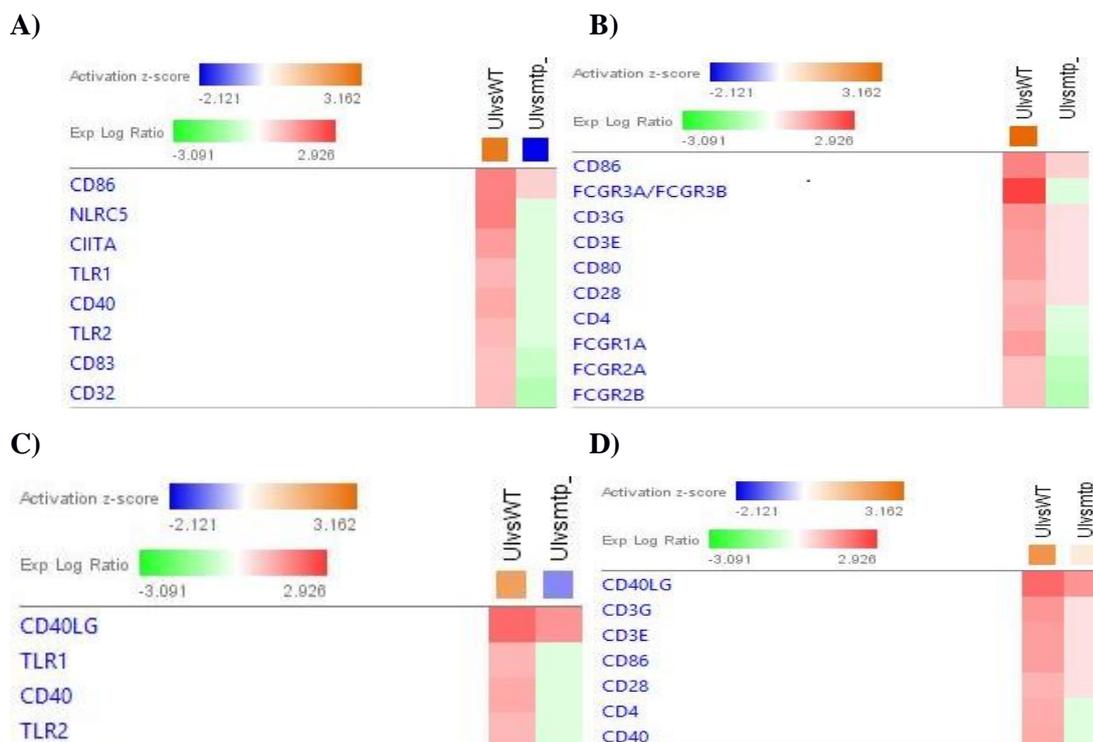




**Figure 3.8.4.2.1.2E: The iCOS-iCOSL Signalling in T Helper Cells pathway as predicted by IPA.**

i: The pathway was enriched due to the WT infection. ii: The enrichment of the same pathway due to the  $\Delta mtp$ -mutant infection. The colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. It is clear that the  $\Delta mtp$ -mutant infection negatively enriched the pathway (green areas), whereas the WT infection only positively enriched the same pathway.

Heat maps (Figure 3.8.2.8A and Figure 3.8.4.2.1.3) of elicited genes associated with the top 5 canonical pathways showed that the WT infection was associated with positively regulated genes in the pathways, in contrast to the largely negative regulation by the  $\Delta mtp$ -mutant infection. As a result, all 5 pathways were graphically depicted for only the WT infection by IPA, as shown in the Dendritic Cell Maturation pathway (Figure 3.8.4.1.3).



**Figure 3.8.4.2.1.3: IPA Heat maps for top pathways enriched from the genes associated in host-pathogen interaction as predicted by IPA:** A: TREM1 Signaling, B: Role of NFAT in Regulation of the Immune Response, C: NF-KB Signaling and D: iCOS-iCOSL Signaling in T Helper Cells. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control ( $p < 0.05$ ). UIvsWT: Wild type infection and UIvsMtp:  $\Delta mtp$ -mutant infection. All heat maps show that the WT infection is associated with genes being positively regulated in the pathways, whereas the  $\Delta mtp$ -mutant infection is associated with negative regulation of most of the genes involved in the pathways.

### 3.8.4.2.2 Networks associated with host-pathogen interaction SDEGs

The 51 host-pathogen interaction genes were found to be involved in only 3 IPA networks, 2 of which were similar in both infections. However, Network 2 was scored 30 in the WT infection and 28 in the  $\Delta mtp$ -mutant infection, resulting in different functions associated with each. The  $\Delta mtp$ -mutant infection network was associated with Cellular Function and Maintenance, Developmental Disorder, and Hereditary Disorder, whilst the corresponding WT infection network was associated with Cellular Function and Maintenance, Inflammatory Response, and Cell Signaling (Table 3.8.4.2.2.1).

**Table 3.8.4.2.2.1: IPA network analysis of host-pathogen interaction genes.**

Network ID	Analysis	Score	Focus Molecules	Top Functions
1	$\Delta mtp$ -mutant	54	25	Humoral Immune Response, Protein Synthesis, Cell-To-Cell Signaling and Interaction
1	WT	54	25	Humoral Immune Response, Protein Synthesis, Cell-To-Cell Signaling and Interaction
2	WT	30	16	Cellular Function and Maintenance, Inflammatory Response, Cell Signaling
2	$\Delta mtp$ -mutant	28	15	Cellular Function and Maintenance, Developmental Disorder, Hereditary Disorder
3	$\Delta mtp$ -mutant	23	13	Cell-To-Cell Signaling and Interaction, Amino Acid Metabolism, Protein Synthesis
3	WT	23	13	Cell-To-Cell Signaling and Interaction, Amino Acid Metabolism, Protein Synthesis

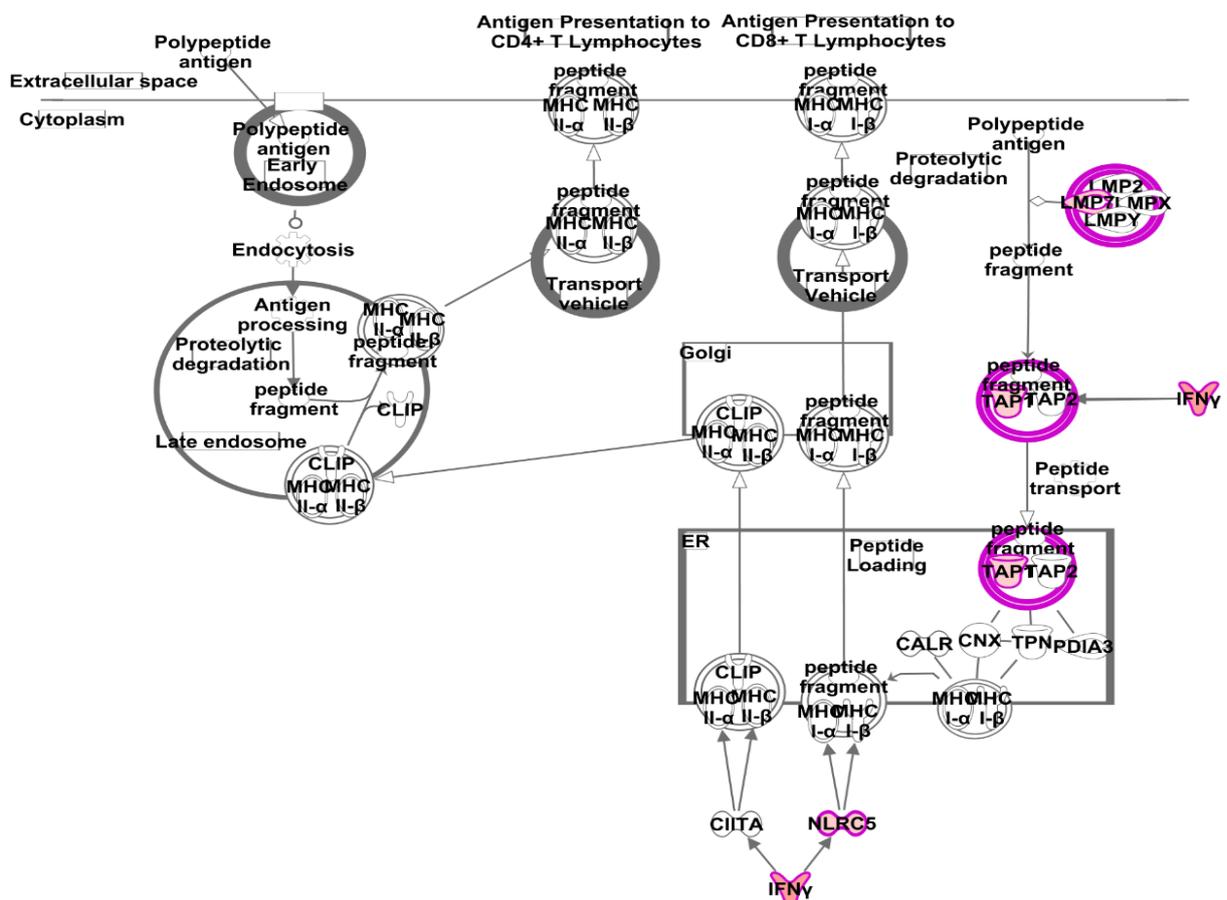
$\Delta mtp$ -mutant:  $\Delta mtp$ -mutant infection

WT: Wild type infection

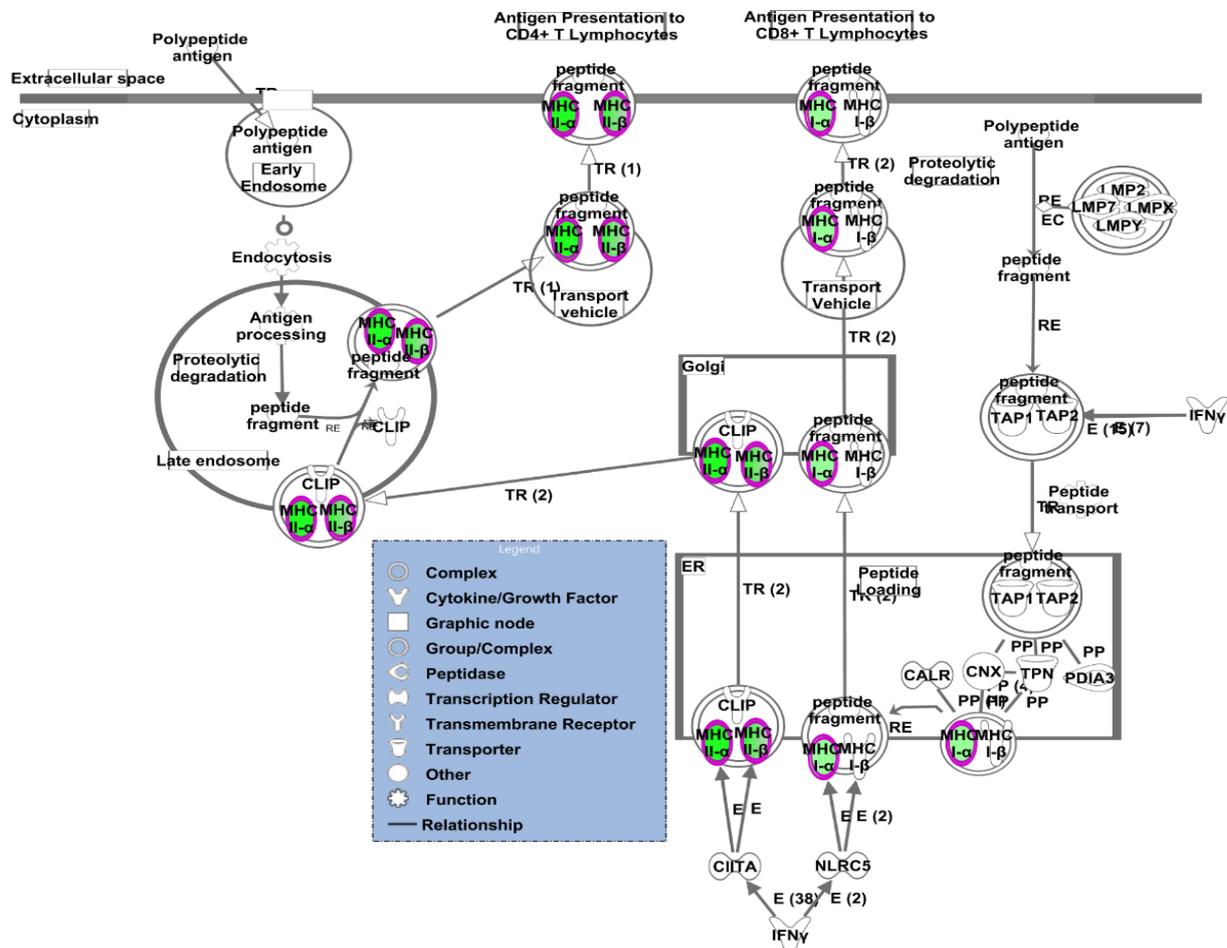
### 3.8.5 The effect of MTP on host innate and adaptive immune response

Infection by *M. tuberculosis* results in the induction of both the innate and adaptive immune responses [137]. The host innate response by the internalization of pathogen via PRRs [90], was shown to be affected by the absence of MTP in this study (Figure 3.8.2.2). Furthermore, the Antigen presentation pathway was lowly enriched in the  $\Delta mtp$ -mutant infection, compared to the WT infection (Figure 3.8.5.1). This pathway impacts the ability of the host to present the antigen to immune cells, directly impacting the host's innate immune response. The antigen presentation pathway enriched by the  $\Delta mtp$ -mutant infection SDEGs downregulated the MHC class II complexes of that pathway (Figure 3.8.5.1B), resulting in downregulation of antigen presentation in both CD4<sup>+</sup> T lymphocytes and CD8<sup>+</sup> T lymphocytes (Figure 3.8.5.1B) by antigen presenting cells (APCs). In contrast, these complexes were not downregulated in the WT infection, in which *TAP1* of the peptide fragment, *NLRC5* and Interferon gamma (*IFNG*) genes were upregulated (Figure 3.8.5.1A).

A)



B)



**Figure 3.8.5.1: Antigen Presentation Pathway as predicted by IPA.** The pathway as enriched due to the A: WT infection, and B:  $\Delta mtp$ -mutant infection. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. It is clear that the  $\Delta mtp$ -mutant infection negatively enriched the pathway (green areas), whereas the WT infection only positively enriched the same pathway.

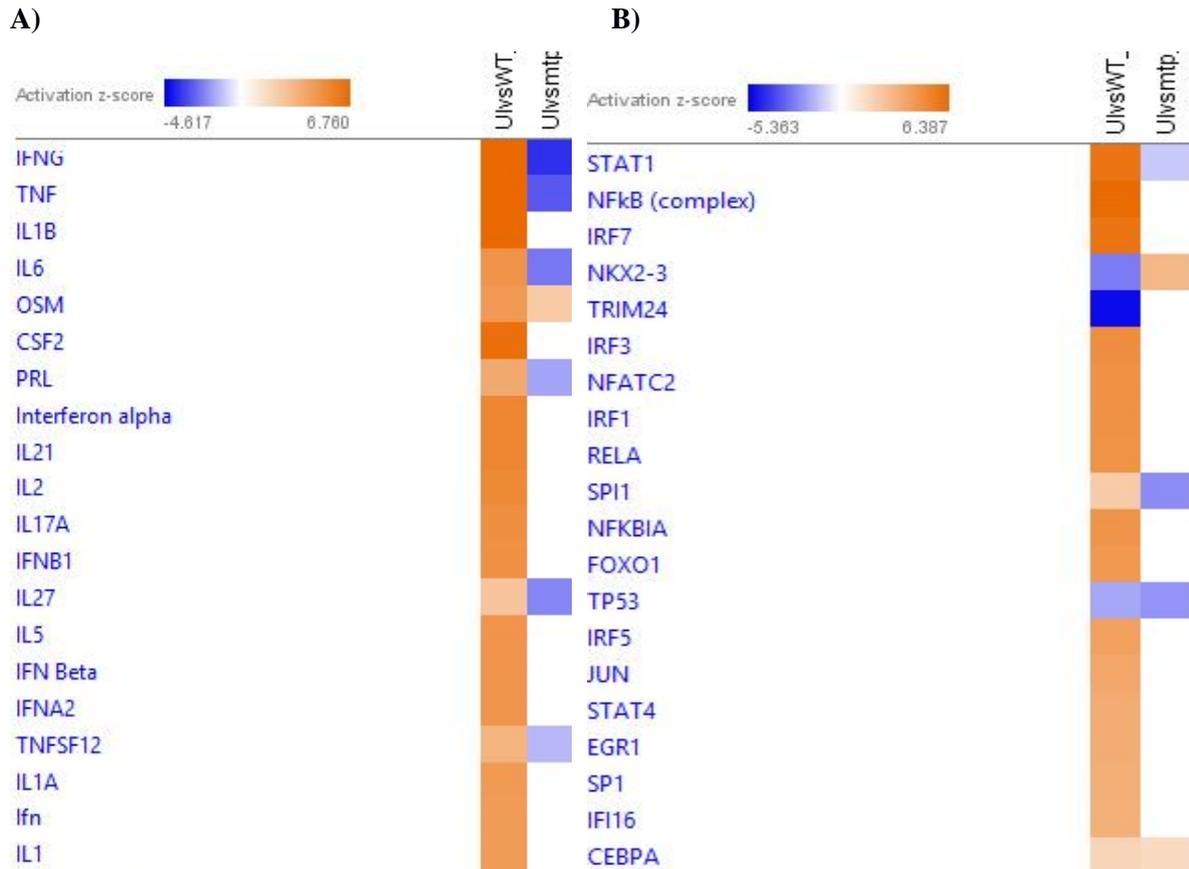
### 3.8.5.1 The effect of MTP on upstream regulators (Cytokines and Transcriptional factors)

The IPA Comparison Analysis default parameters were used to identify and analyse upstream regulators of the SDEGs. Cytokines such as IFNG and Tumour necrosis factor (TNF), and transcriptional factors/regulators (TFs) of the canonical pathways activated by SDGEs, as well as other molecules e.g. lipopolysaccharide (LPS) which affect regulation and thus enrichment of pathways mentioned above, were predicted to be key upstream regulators affected by *M. tuberculosis* on Balb/C mouse lungs. IFNG was the top ranking upstream regulatory molecule, with an activation z-score of 9.18 and -3.78 in the

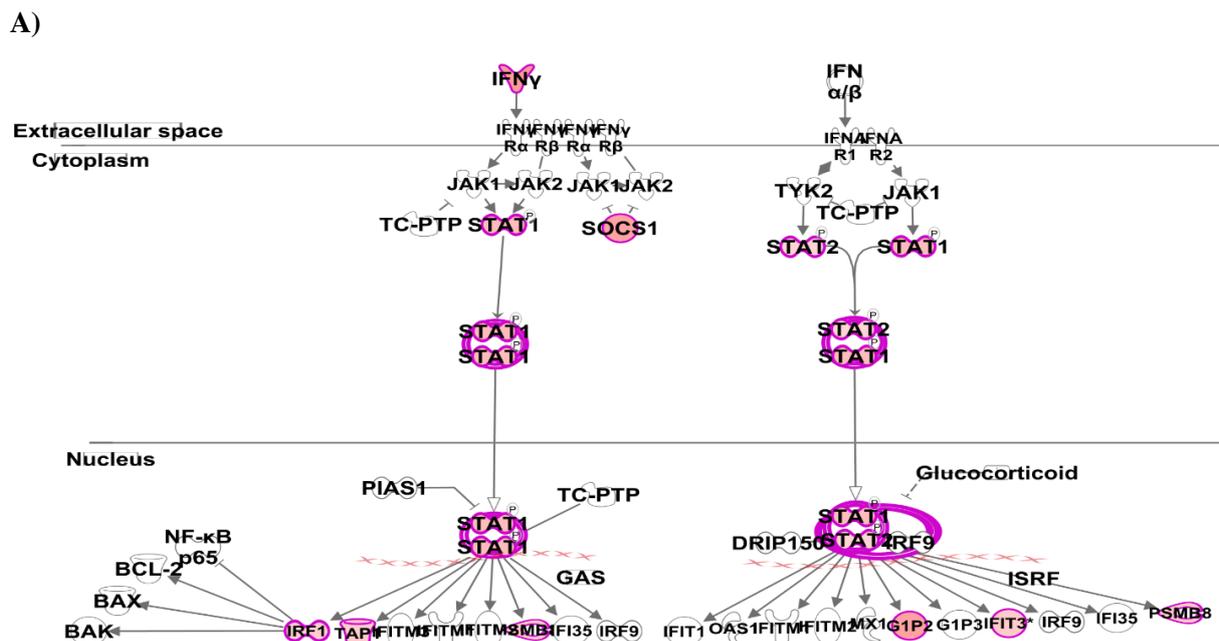
WT and  $\Delta mtp$ -mutant infection respectively. The next most highly ranked key upstream regulator was LPS, with the WT infection exhibiting an activation score of 8.91, whilst the  $\Delta mtp$ -mutant infection negatively regulated this molecule (z-score = -3.56).

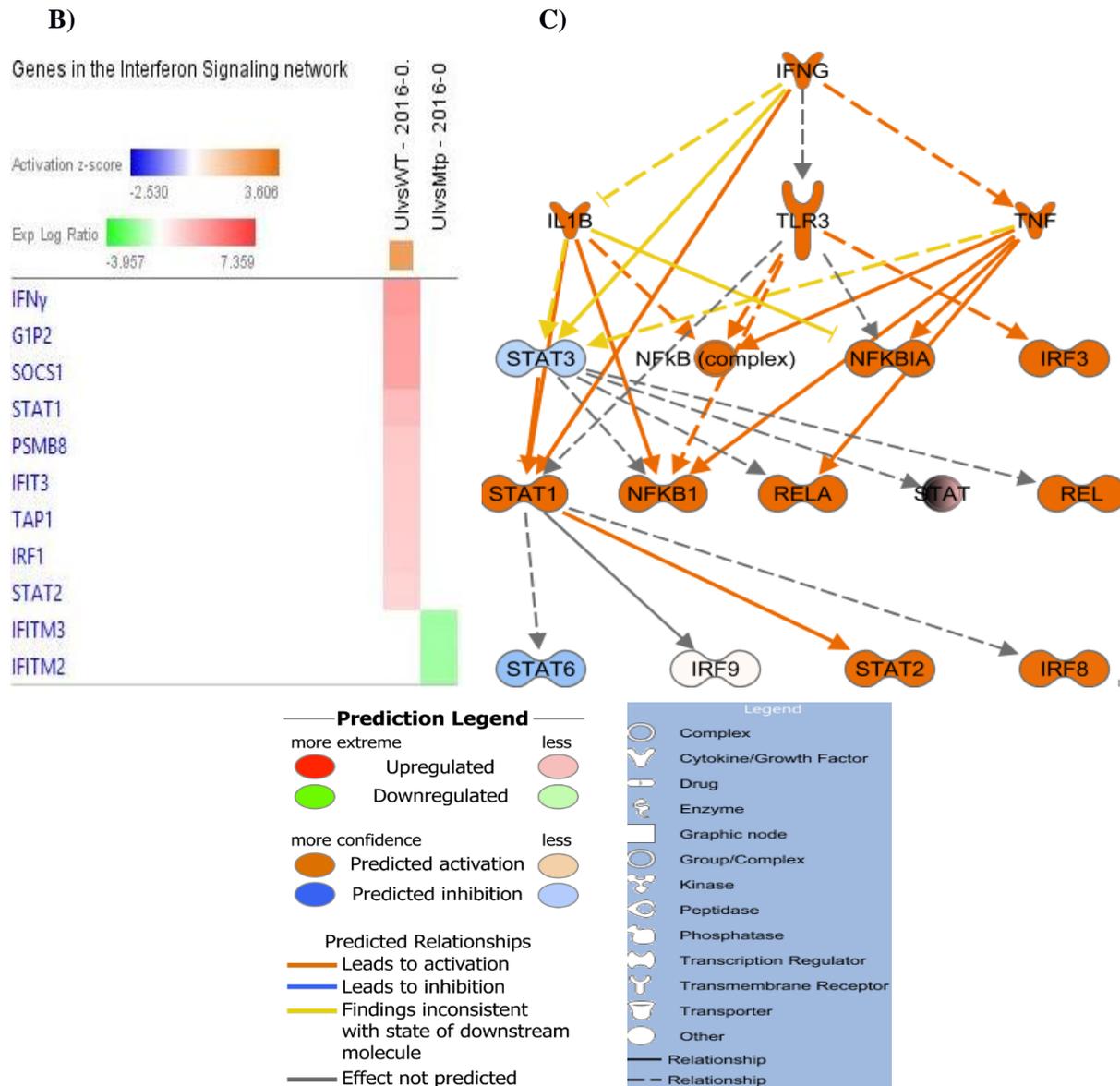
Analysis of cytokines and TFs with IPA default settings revealed that 139 cytokines were regulated by the 2 infection models studied, and only 7 of the top 20 most regulated cytokines were enriched by  $\Delta mtp$ -mutant genes relative to uninfected (Figure 3.8.5.1.1A). In contrast, the WT infection upregulated all 20 of these cytokines (Figure 3.8.5.1.1A). IFNG was the most highly ranked upstream regulatory cytokine. Similar to the cytokines, the WT infection SDEGs positively enriched 9 genes/molecules (*IFIT3*, *SOCS1*, *IFNG*, *STAT2*, *PSMB8*, *STAT1*, *TAP1*, *IRF1* and *ISG15*) (Figure 3.8.5.1.2B) associated with the Interferon Signalling pathway, in contrast to only 2 genes/molecules (*IFITM3*, *IFITM2*) by the  $\Delta mtp$ -mutant infection. Moreover, these genes/molecules were negatively regulated (Figure 3.8.5.1.2B), and hence, the genes from the  $\Delta mtp$ -mutant infection negatively enriched this pathway (Figure 3.8.5.1.2A), even though the pathway was shown to be significantly affected by the  $\Delta mtp$ -mutant infection (Table 3.8.1.1). The mechanistic network of IFNG shows the interaction of IFNG with other upstream regulators (Figure 3.8.5.1.2C). TNF was the next most highly ranked upstream regulatory cytokine with a z-score of 8.11 and -3.04 in the WT and  $\Delta mtp$ -mutant infection respectively.

Comparison Analysis of the transcription factors enriched from the SDEGs showed signal transducer and activator of transcription (STAT)1 as the most upregulated, with a z-score of 5.93 and -1.15 in the WT and  $\Delta mtp$ -mutant infection respectively. The STAT1 mechanistic network in the  $\Delta mtp$ -mutant infection differed from that in the WT infection (3.8.5.1.3). More (13) molecules regulated by SDEGs of the STAT1 mechanistic network were upregulated during WT infection (Figure 3.8.5.1.3A), whilst only 5 were upregulated during  $\Delta mtp$ -mutant infection (Figure 3.8.5.1.3B). The same number of molecules (5) were down regulated during WT and  $\Delta mtp$ -mutant infection (Figure 3.8.5.1.3). The next most highly ranked TF was the NF- $\kappa$ B complex, followed by IRF7 (Figure 3.8.5.1.1B). In the top 20 most regulated TFs, most (17) were upregulated in the WT infection, but not in the  $\Delta mtp$ -mutant infection (Figure 3.8.5.1.1B).

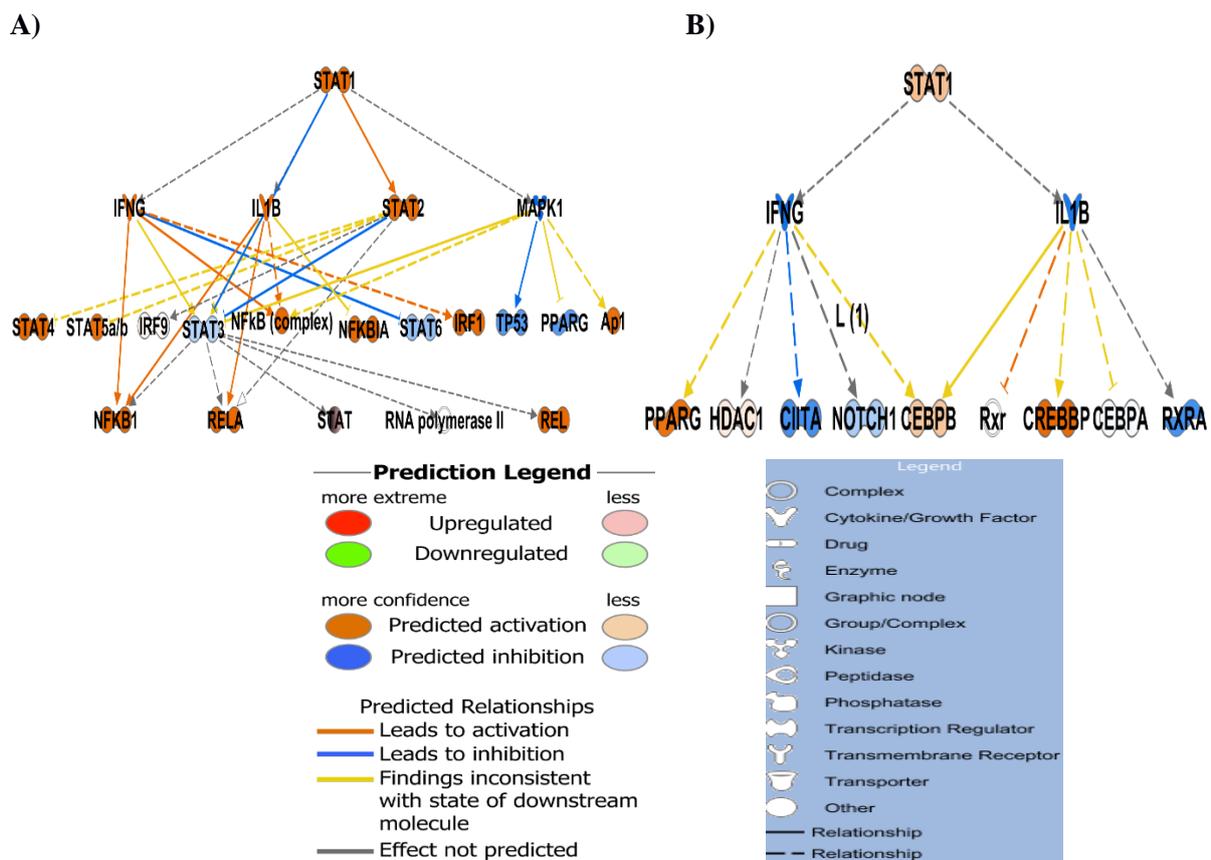


**Figure 3.8.5.1.1: Upstream regulator analysis by IPA.** A: Comparison Analysis of the cytokines enriched from the unique genes as shown in IPA. B: Comparison Analysis of the transcription factors enriched from the unique genes as shown in IPA. Colour intensity indicates the activity prediction of regulation by IPA: orange represents an overall activation increase of the pathway, and blue represents the overall repression of the pathway activity, relative to the uninfected control. UIvsWT: Wild type infection and UIvsmtf:  $\Delta mtp$ -mutant infection.





**Figure 3.8.5.1.2: Interferon signalling pathway and the associated genes as predicted by IPA.** A: Interferon signalling pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only, and genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway. B: The genes involved in the Interferon signalling pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control ( $p < 0.05$ ). UivsWT: Wild type infection and UivsMtp:  $\Delta mtp$ -mutant infection. C: The mechanistic network of IFNG. Colour intensity indicates the activity prediction of regulation by IPA: orange represents an overall activation increase of the pathway, and blue represents the overall repression of the pathway activity, relative to the uninfected control. Solid connecting lines represent directly connected SDEGs, and dotted lines show an indirect connection associated with the SDEGs.



**Figure 3.8.5.1.3: STAT1 network as predicted by IPA. A:** The mechanistic network enriched due to the A: WT infection, and B:  $\Delta mtp$ -mutant infection. Colour intensity indicates the activity prediction of regulation by IPA: orange represents an overall activation increase of the pathway, and blue represents the overall repression of the pathway activity, relative to the uninfected control. Solid connecting lines represent directly connected SDEGs, and dotted lines show an indirect connection associated with the SDEGs.

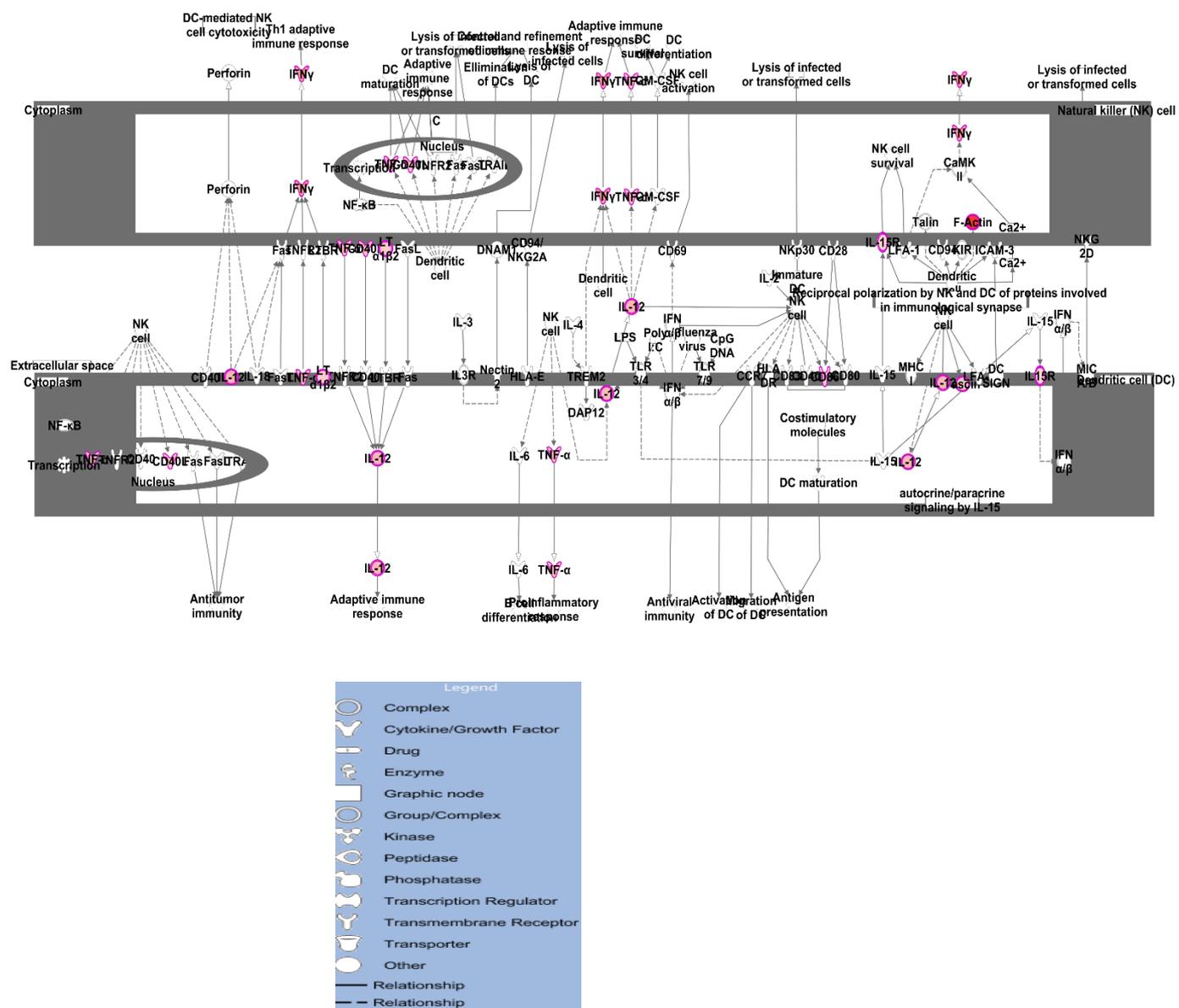
### 3.8.5.2 The effect of MTP on the regulation of the host immune response.

The activation status of specific IPA pathways, represented by z- scores, and the analysis of the involved in the pathways clearly shows the association of the WT infection in eliciting the immune response, whereas the  $\Delta mtp$ -mutant infection did not activate these pathways. For example, the Crosstalk between Dendritic cells and Natural Killer (NK) Cells pathway (Figure 3.8.5.2.1). Based on the expression pattern of the SDEGs that regulated this pathway, it was only enriched in the WT infection. This infers that MTP plays a role in the communication of 2 cell types involved (DCs and NK cells).

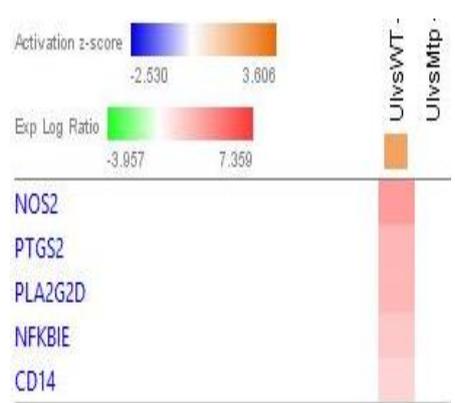
IPA cytokine analysis further corroborated our findings showing that cytokines were more upregulated by the WT infection (Figure 3.8.5.1.1A). The Macrophage migration inhibitory factor (MIF) is a cytokine with an integral role mediating the innate immune system. The heat maps of elicited genes

associated with the MIF Regulation of Innate Immunity pathway showed that the WT infection was associated with positively regulated genes in the pathways, in contrast to the absence of regulation by the  $\Delta mtp$ -mutant infection (Figure 3.8.5.2.2).

The genes associated in the Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway (Figure 3.8.5.2.3A) pathway were upregulated only in the WT infection, whilst the  $\Delta mtp$ -mutant infection depicted a downregulation of the genes (Figure 3.8.5.2.3B). Furthermore, the Phagosome formation pathway was also only enriched by SDEGs of the WT infection only (Figure 3.8.5.2.4). The SDEGs involved in this pathway, were all positively regulated (squares in red) in the WT infection, whereas the  $\Delta mtp$ -mutant infection was associated with negative regulation (green squares) of all of the genes involved in the pathway (Figure 3.8.5.2.4B).

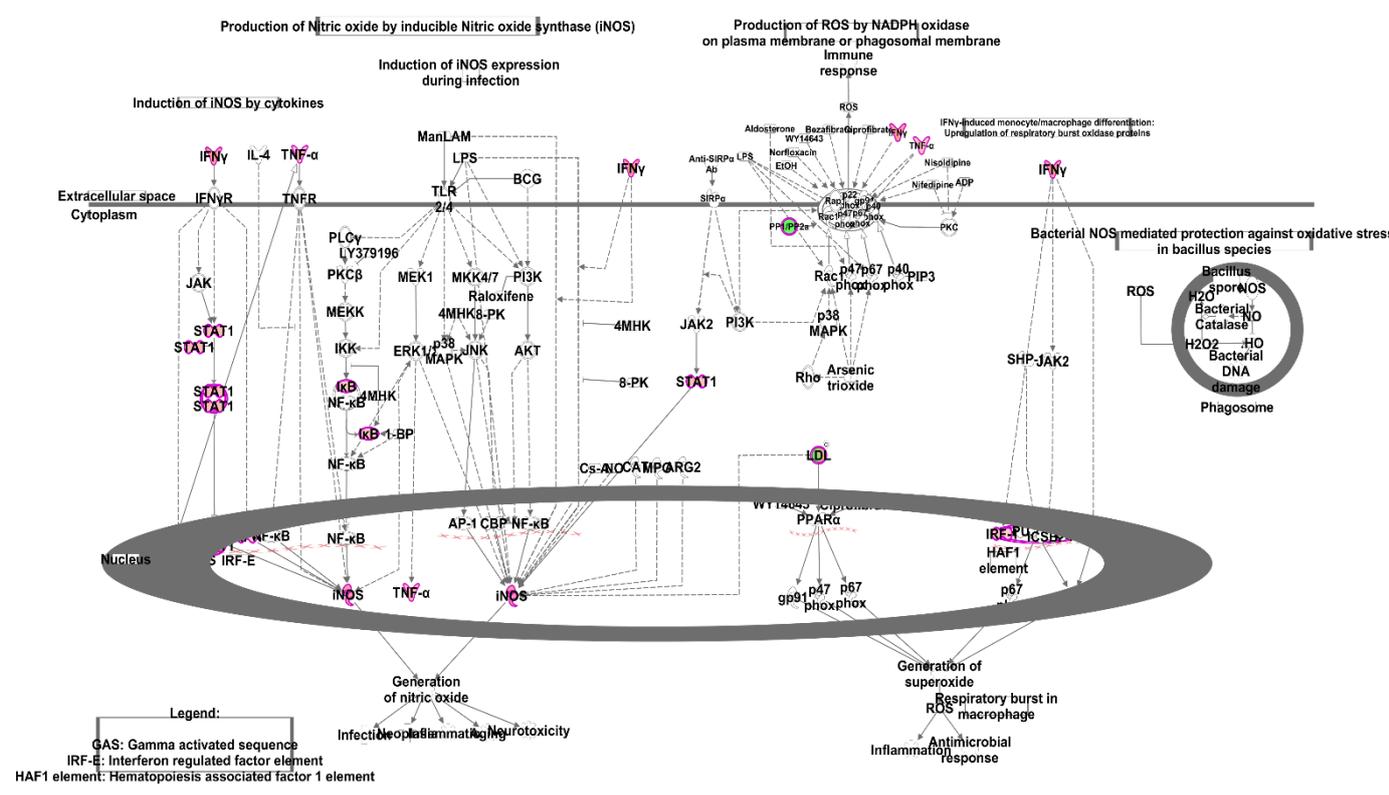


**Figure 3.8.5.2.1: Crosstalk between Dendritic cells and Natural Killer Cells pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway.

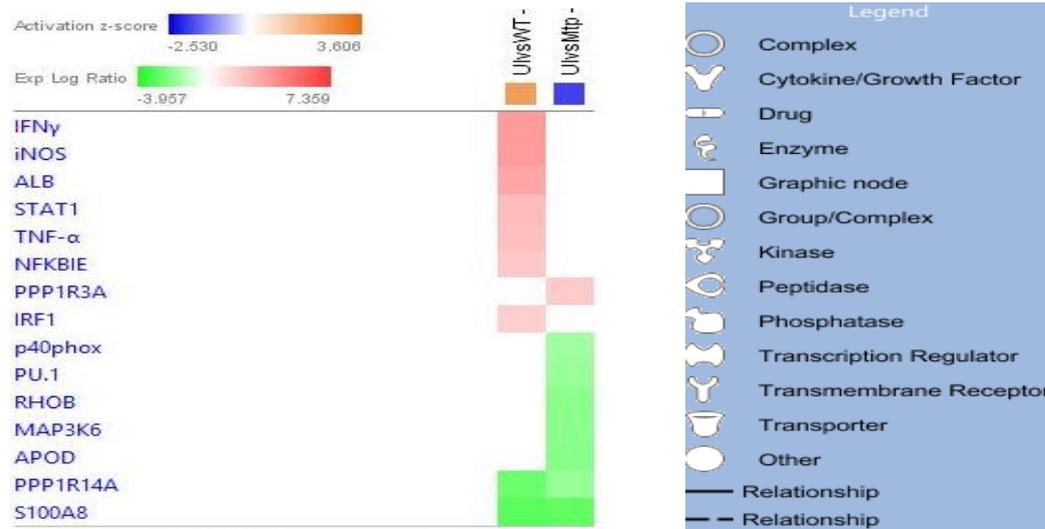


**Figure 3.8.5.2.2: Genes involved in the Macrophage migration inhibitory factor (MIF) Regulation of Innate Immunity pathway as predicted by IPA.** Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control ( $p < 0.05$ ). UI vs WT: Wild type infection and UI vs Mtp:  $\Delta mtp$ -mutant infection.

A)



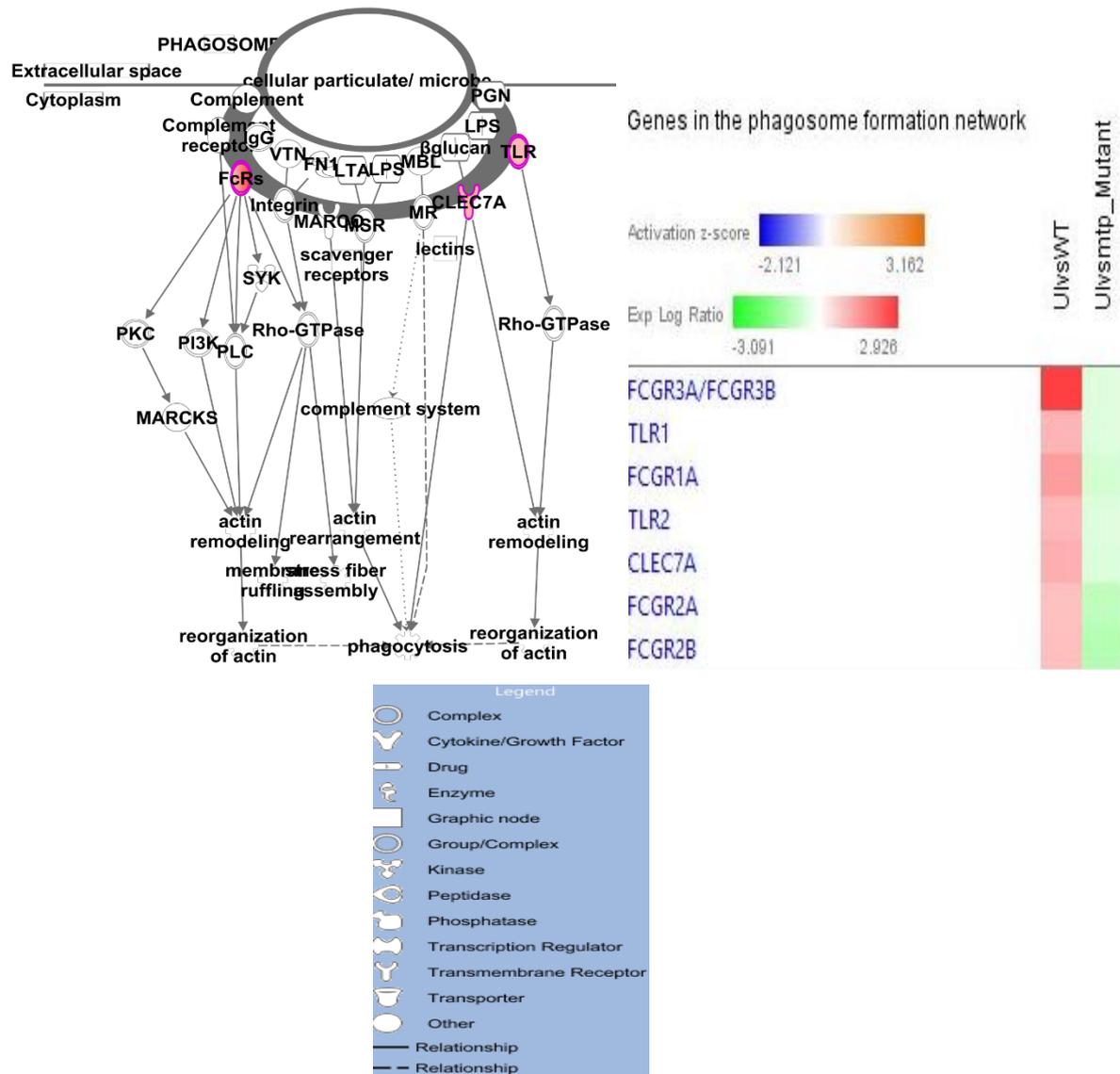
**B)**



**Figure 3.8.5.2.3: Production of Nitric Oxide and Reactive Oxygen Species in Macrophages as predicted by IPA.** A) Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. The genes from the  $\Delta mtp$ -mutant infection did not enrich this pathway. B) Heat map of the genes involved in this pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control ( $p < 0.05$ ). UIvsWT: Wild type infection and UIvsMtp:  $\Delta mtp$ -mutant infection.

**A)**

**B)**



**Figure 3.8.5.2.4: Phagosome formation pathway as predicted by IPA.** A: Phagosome formation pathway. Colour intensity indicates the degree of regulation: red representing activation and green representing repression, relative to the uninfected control. Grey shading indicates genes that were not differentially expressed; white shading represents genes in the pathway not represented in the dataset. This pathway was enriched by the genes from the WT infection only. B. Heat map of the genes involved in this pathway, where all genes from the WT infection are shown as being positively regulated (squares in red) in the pathway, whereas the  $\Delta mtp$ -mutant infection is associated with negative regulation (green squares) of all of the genes involved in the pathway.

## **CHAPTER 4: DISCUSSION**

Since the first report of *M. tuberculosis* curli pili (MTP) [10], research in this field has gained momentum, resulting in a steady increase in the knowledge pertaining to its importance in the host-pathogen interactions [12,14,138]. Curli pili, like other adhesins, have been previously found to facilitate bacterial adherence not only to host cells [10,15,18], but to surfaces [11] and other bacteria as well [7]. Adhesion to host cells is necessary for colonizing the host and induction of the immune response [7,72], and is therefore, an essential virulence factor of most bacterial pathogens, including *M. tuberculosis*. Ramsugit *et al* (2014) showed reduction in adhesion to and invasion of pulmonary epithelial cells [12,18] as well as THP-1 macrophages, by a MTP-deficient strain [12,15]. MTP was also reported to facilitate biofilm production *in vitro* [11], thus targeting this protein has been previously suggested as a potential mechanism to reduce the pathogen's persistence within a host during infection [14].

In this study, we analysed the *in vivo* growth kinetics of a *mtp* gene knockout mutant relative to its wild type and complemented strains, and the effect of MTP on gross organ pathology. In addition, this study elucidates for the first time, by RNA sequencing, the whole transcriptome response to MTP in a Balb/C mouse lung during early infection with the MTP deficient ( $\Delta mtp$ -mutant) clinical strain of *M. tuberculosis*. A time point of 14 days post-infection was chosen to compare and understand the different host responses including host-pathogen interactions and host immune response, as previous studies have shown that early onset of the host adaptive immune response can commence by day 14 [139], and hence the ability to look at both innate and early adaptive immune responses.

#### **4.1 The effect of MTP on growth of *M. tuberculosis***

##### **4.1.1 *MTP is vital for growth of M. tuberculosis in vitro***

The *in vitro* growth assays statistically revealed that the *mtp*-mutant strain grew at a significantly decreased rate compared to that of the WT during the log phase in broth culture (Figure 3.1.1). This suggests that the deletion of the *mtp* gene results in a slower growth rate of the bacterium, and hence, MTP has an effect on the growth of *M. tuberculosis* in broth culture. These findings are contradictory to those of another major adhesin of *M. tuberculosis*, HBHA, which was reported to have no effect on the *in vitro* growth [140] after the  $\Delta hbhA$ -mutant and its parent strain showed similar growth rates [140]. The restoration of the *mtp* gene in the complemented strain did not result the restoration of the growth rate. This observation suggests that the deletion of the *mtp* gene could have concurrently induced and resulted in another change in the genome, whose function is the growth deficit phenotype observed.

##### **4.1.2 *MTP is essential for growth in vivo***

MTP are produced *in vivo* during TB infection [10,13], and play a significant role in the adhesion to and invasion of Eukaryotic cells [15]. An exponential growth of the bacteria was observed in the mice

lungs during the first 21 days of infection with both the  $\Delta mtp$ -mutant and its WT (Figure 3.2.1), in contrast to previous studies where exponential bacterial growth was observed within the first 30 days [137]. This could be associated with the control of bacterial growth that takes place in early innate immunity during lung infection of *M. tuberculosis* [6]. Deletion of MTP affected the growth of the bacilli in the lungs and spleen of infected mice. The lower bacterial load that was noted during the  $\Delta mtp$ -mutant infection over the infection period in this study could be due to the slower *in vivo* growth rate of the  $\Delta mtp$ -mutant, or to a decrease in the ability of the  $\Delta mtp$ -mutant to colonize or persist in the lungs. This serves as supporting evidence that *M. tuberculosis* uses pili to enable colonization of the host. This reduced infectivity was also observed by Mueller-Ortiz *et al* (2002), in the lungs of C57BL/6 mice using a  $\Delta hbhA$ -mutant [140]. However, at day 14 post-infection the WT and *mtp*-mutant strain growth rates showed that there was no growth deficit in the *mtp*-mutant strain regardless of the absence of MTP after 14 days of *M. tuberculosis* infection. Pathway analysis at this time point revealed enrichments of host pathways after infection with the WT strain compared to the *mtp*-mutant strain. This suggests that virulence of the *mtp*-mutant strain is not dependent on the growth rate, and hence proposing that MTP plays a role in the virulence of *M. tuberculosis*, but not on the growth of the bacilli. Studies have shown that the progressive inflammatory response to *M. tuberculosis* in mice is independent of the total number of bacilli that is cultureable [121]. To our knowledge, this is the first report that shows that *mtp*, the gene encoding curli pili in *M. tuberculosis*, is necessary for the complete virulence of *M. tuberculosis in vivo*. Curli pili has been previously shown to influence the virulence of *Streptococcus pneumoniae* (*S. pneumoniae*) *in vivo* after a non-piliated mutant strain of *S. pneumoniae* was found to be less virulent than its WT strain [141].

#### **4.2 Transcriptome profiling post *M. tuberculosis* infection.**

The analysis of the various host transcriptomes in response to *M. tuberculosis* infection has enriched our understanding of the molecular and immunological mechanisms underlying infection, as well as the signaling and cellular pathways that determine the outcome of the infection, whether it be active or latent infection [142–144]. The current study was carried out with the principal objective to distinguish the difference in transcriptome response of the  $\Delta mtp$ -mutant strain compared to the WT after 14 days of infection. Global gene expression analysis revealed that the overall transcriptome response was greater in the lungs infected with the  $\Delta mtp$ -mutant strain (Figure 3.7.1). However, the number of organisms contained in the lungs of the  $\Delta mtp$ -mutant infected mice at the time point that RNA was sequenced was not significantly different ( $p = 0.3$ ) from the infecting WT organisms after aerosol infection (Figure 3.4.1). This suggests that the higher number of reads elicited by the host after  $\Delta mtp$ -mutant infection was due to MTP and not to significant differences in bacterial load between WT and  $\Delta mtp$ -mutant infected mice.

Moreover, with the exception of only 1 read, only high quality read data was obtained from the sequencing method used. Thus, the host differential gene expression (DGE) profiles from these read pools could not have been affected by read quality, which may affect gene expression profiles. Care was exercised at the time of sample processing, both during organ harvesting and RNA extraction, in order to conserve RNA integrity. This contributed to the good RIN numbers obtained (Table 3.5.1), and the high quality reads for all samples (Figure 3.6.1.1). Thus, the higher number of reads elicited in the host after  $\Delta mtp$ -mutant infection was not due to higher read quality for that sample. The read number was shown by the higher read pool density (Figure 3.7.1) which corresponded to the alignment/mapping percentages of the host transcriptome after infection with the  $\Delta mtp$ -mutant strain, compared to the WT strain (Table 3.6.2.1). This suggests that the absence of MTP in *M. tuberculosis* resulted in the induction of an increased host response in the form of transcriptional regulation.

A variation in the millions of reads between the two  $\Delta mtp$ -mutant biological strain was observed. However, the quality check reported good quality reads that were obtained from RNA of the same good integrity (RIN=8 for both replicates). Therefore, variation would not be a limitation factor as the sequencing had enough coverage from each library to perform differential expression analysis [145]. Furthermore, the differential analysis tool used (Cufflinks) takes into account dataset size for comparing gene expression levels by normalizing reads mapping to each gene using gene length and total number of reads, and therefore any downstream analysis was not affected by the difference [145]. Cufflinks also assesses differential expression using an ‘expectation-maximization’ approach that approximates the transcript abundances by taking into account biases such as the non-uniform read distribution. Therefore, the total number of reads for the respective experiments will not alter the outcome of the analysis [130].

For stringent analysis, only transcripts showing a 2-fold or higher change in their gene expression, with a corrected false discovery rate (FDR)  $p < 0.05$  were considered for further analyses as significantly differentially expressed genes (SDEGs) [137]. The number of SDEGs in the  $\Delta mtp$ -mutant infection was significantly higher (2-fold) than that of the SDEGs in the WT. The number of upregulated genes were significantly more than downregulated genes across both infections (Figure 3.7.1.1A).

Analysis executed at the gene level showed that the defensin beta 4 (*DEFB4*) gene was one of the most highly upregulated genes, in both infections (Table 3.7.1.1). This is an immune related gene that codes for an antimicrobial peptide, beta defensin, that is also an important chemo-attractant [146]. Similar to our study, this gene has previously been shown to be induced after *M. tuberculosis* infection and is associated with innate immune response against *M. tuberculosis* [147]. Amongst the most downregulated  $\Delta mtp$ -mutant infection genes in our study, was the CAMP gene which encodes the

cathelicidin antimicrobial peptide (*CAMP*), that has chemotactic activity and has been shown to control growth of *M. tuberculosis* [148]. This correlates with the low growth rate observed *in vivo* in this study.

#### **4.2.1 The WT and $\Delta mtp$ -mutant strains induce common as well as unique genes in Balb/C mice**

A group of 128 genes was significantly either up regulated or down regulated in both WT and  $\Delta mtp$ -mutant infections. However, the magnitude of the expression was different, with the WT strain causing a higher expression in most of the genes, both activated (Figure 3.7.1.3.1) and repressed (Figure 3.7.1.4.1). We postulate that these 128 genes observed in both infections could represent the core host genes necessary for early infection, and are essential in the general pathogenesis of *M. tuberculosis*.

Although the transcripts from the WT and  $\Delta mtp$ -mutant infection showed similar patterns of dispersion (Figure 3.7.2), the gene expression profile of the  $\Delta mtp$ -mutant infected lungs appear to be mostly unique compared to the expression profile of the WT infected lungs 14 days post infection.

##### **4.2.1.1 Gene Ontology (GO) analysis reveals the association of MTP with host immunity**

SDEGs were enriched to analyze the functional categories associated, alongside the concurrent description and comparison of the regulation of the host transcriptome profile after infection with the 2 strains. This was anticipated to reveal the effect of MTP during infection of mammalian lungs. The patterns of transcript fold changes in GO functional categories (FCs) from the groups of shared gene set (n=128) revealed that over-represented upregulated genes were involved in enzyme related functions such as serine-type endopeptidase inhibitor activity (GO:0004867), endopeptidase inhibitor activity (GO:0004866), and endopeptidase regulator activity (GO:0061135) (Table 3.7.1.3.1). The over-represented downregulated genes were associated with only 1 molecular function involved in the Toll-like receptor 4 binding (GO:0035662; p = 3.72E-02) (Table 3.7.1.3.1) and this is associated with host immune responses, and enriched significantly for related molecular functions such as those [80]. Genes for signalling molecules like *S100a8* and *S100a9*, and chemokines e.g. *Ppbp*, featured among the most downregulated genes. This implies that the downregulation of immune related functions is due to the absence of MTP, and hence MTP plays a role in immune response.

GO terms similar to those enriched by SDEGs regulated following bacterial infection with the WT strain only, have been reported to be stimulated by *M. tuberculosis* after 24hours of infection [149]. These include GO terms related to cytokine and chemokine receptor binding and activity. On the other hand, GO enrichment of 930 SDGEs unique to the  $\Delta mtp$ -mutant infection revealed that molecular transducer activity (GO:0060089), receptor activity (GO:0004872), signal transducer activity (GO:0004871), signalling receptor activity (GO:0038023) and ion binding (GO:0043167) were among

the top biological functions (Table 3.7.1.1.1) affected by removal of the *mtp* gene. These results show evidence of the repression of the host immune response by the  $\Delta mtp$ -mutant strain (Table 3.7.1.1.1). Taken together, these findings highlight that the involvement of MTP was host immune response related. A gene expression study of host responses of TB patients after *M. tuberculosis* infection revealed that the GO terms enriched by the gene expression profile specifically related to *M. tuberculosis* infection is mostly related to immune cell activation and differentiation, chemokine receptor activity, as well as regulation of the immune response [150]. Immune system genes in another previous study were also shown to enrich a wide range of GO biological processes, such as immune response activation, effector processes, cell migration, and immune response terms where antigen processing and presentation was found to be dominating [151]. Furthermore, in the same study, a substantial number of genes associated with cytotoxic cell-mediated killing was reported to be significantly upregulated after GO analysis [151].

#### **4.3 IPA Pathway analysis reveals host cellular pathways regulated by MTP during *M. tuberculosis* infection of Balb/C mice**

In order to decipher the effects of MTP on the molecular mechanisms engaged by the host, a comprehensive pathway analysis was performed on the  $\Delta mtp$ -mutant and WT infected mouse lung transcriptome after 14 days of infection. Previous studies have shown that at this time point, both the host's innate and adaptive immune response can be detected. However, the adaptive immune response would not have fully developed [139]. Although the *p*-values of the pathways enriched by all the genes showed significance ( $p < 0.05$ ), a trend of higher *p*-values was observed in the pathways enriched by the  $\Delta mtp$ -mutant strain (Figure 3.8.1.1.1), suggesting lower significance of these pathways. The same trend was observed in pathways enriched by the unique genes, where the pathways from the WT infection were more significant than the  $\Delta mtp$ -mutant strain pathways. This suggests that the deletion of the *mtp* gene affected the regulation of these pathways in mice lungs, resulting in a lower significance of the enriched pathway. The Agranulocyte adhesion and Diapedesis pathway was the most significantly enriched. The activation status of this pathway revealed an increased activity in the WT infection in contrast to that in the  $\Delta mtp$ -mutant infection (Figure 3.8.1.1.2). The IPA database reported that this pathway is part of the Cellular Immune Response pathways, and is involved in Cell-To-Cell Signalling and Interaction; Haematological System Development and Function; and Immune Cell Trafficking functions [152].

##### **4.3.1 Pattern recognition receptors (PRRs)**

The host utilizes a number of PRRs to recognize conserved structures or PAMPS on the bacilli [83]. After evasion by *M. tuberculosis*, modulation of the host immune response via PRRs is required for host protection. Irregularities in the activation of PRR signalling pathways regulated by TB affects

disease pathogenesis and thus, needs to be elucidated [86]. Alteri *et al* (2005) suggested that MTP may mediate the interactions between the bacilli and host cells when IgG antibodies were demonstrated in TB patient sera, indicating that MTP was produced during *in vivo* infection by *M. tuberculosis* [13]. Our study showed that the absence of MTP from the bacilli has a deleterious effect on the activation of the role of PRRs in recognition of bacteria and viruses signalling pathway (Figure 3.8.2.2), that is activated upon *M. tuberculosis* infection in mice lungs after 2 weeks of infection. These PRRs are located extracellularly, on membranes, as well as within phagocytic cells like Macrophages and DCs [77]. The PRR- dependent entry is a key determinant of the fate of *M. tuberculosis* after infection [83], and the induction of this pathway results in gene expression activation, and the production of a wide range of immune related molecules and their receptors [153]. In the current study, all 3 types of PRRs were shown to be affected by MTP. The extracellular C1q protein, the membrane bound C3a receptor, and the cytoplasmic Cosp1 were found to be upregulated in the pathway enriched by the WT infection (Figure 3.8.2.2). This resulted in an upregulation of transcription of the transcription factors *IRF3* and *IRF7*, and subsequently an increase in the transcription and translation of pro-inflammatory cytokines (Figure 3.8.2.2). However, the  $\Delta mtp$ -mutant infection did not enrich any of the genes involved in this pathway, that were seen in the WT infection, Furthermore, this pathway is important in linking the innate immune response to the adaptive immune response of the host [153].

#### 4.3.1.1 MTP affects the Toll-like Receptor (TLR) pathway

TLRs play the most significant role, amongst all the PRRs during *M. tuberculosis* infection [91], as they are known to recognize a wide range of *M. tuberculosis* structures [87]. They induce the adaptive immune response by stimulating the production of cytokines and additional anti-bacterial effector molecules, after interacting with ligands that are pathogen specific [83,87]. They are also involved in arresting the growth of infecting bacteria, and link the innate immune responses to the adaptive immune responses [83]. Curli pili have been categorized as PAMPs, recognized by TLR2 and this interaction results in the activation of an innate immune response via the stimulation of IL-8 [105,154].

In this study, the elimination of the PRRs pathway was accompanied by the exclusion of the TLR Signalling pathway in the host response where *mtp* was deleted (Figure 3.8.4.1.1.1). The genes involved in the TLR pathway were upregulated in the WT infection model while only 3 of them were upregulated in the  $\Delta mtp$ -mutant infection model (Figure 3.8.4.1B). This resulted in the activation of the adapter molecule MYD88, and the NFkB complex proteins, initiating signal transduction [155]. Previously, the expression of nuclear factor proteins such as NFkB was found to be upregulated in *M. tuberculosis* infection, related to the upregulation of TLRs [156]. The enrichment of this pathway in the WT infection resulted in increased transcription of pro-inflammatory cytokines involved in the activation of the antimicrobial response and apoptosis as part of the innate host defence: IL-1 and TNF, and IL-12 as part of the adaptive immune response (Figure 3.8.4.1.1.1D) [79]. Because the TLR signalling pathway

is a pathogen-influenced signalling pathway [155], it strongly supports that MTP has a major role in the signalling of an immune response by the host. Our studies corroborate with earlier studies in mice with inactivated TLR genes that showed the importance of TLR2 controlling *M. tuberculosis* infection and regulating inflammation after infection [157]. The F15/LAM4/KZN strain used in this study was shown to enrich the TLR pathway, and other immune related canonical pathways *in vitro* study previously [132], which is supported by these findings.

#### 4.3.1.2 MTP induces the Complement System

The host relies on complement activation, the humoral arm of the innate immune system [90,158] to recognize and minimize microbial infection at an early stage [153]. The complement system components include receptors, regulators and effector proteins, that have a key role in the host's anti-bacterial defence [159]. The complement system is activated through three distinct pathways: the alternative, classical, and the lectin pathway, subject to the activating surface and the specific recognition molecules involved [160]. Previous studies demonstrated that some mycobacterial species like *M. bovis*, can activate all three complement pathways, however, only 2 have been shown to be activated by *M. tuberculosis*, the classical, and alternate pathways [158,160].

In this study, the genes associated with the complement system pathway were significantly upregulated, including those activating the classical and alternate pathways complement system, after infection with the WT strain only. The complement subcomponent subunit C1q was significantly upregulated (Figure 3.8.4.1.1.2) in the classical pathway of the complement system enriched by the WT SDEGs. In addition to the C1q protein, both the C1 receptors, C1r and C1s, were significantly upregulated. This in turn elicited the up-regulation of the C4a and C4b proteins of the classical pathway of the Complement system (Figure 3.8.4.1.1.2). The complement C4 components (C4a and C4b) play a crucial role in the complement system by activating the downstream paths of the classical pathway [161].

The alternate pathway was also positively enriched in the current study, by the up-regulation of C3b peptidases (Figure 3.8.4.1.1.2) and C3 receptors, C3ar1 (Figure 3.8.4.1.1.2). As an essential factor in the complement system pathways, the complement C3 component is cleaved into 2 parts, C3a and C3b, to eventually form the terminal complement complex C5b-C9 or Membrane attack complex (MAC), which can infiltrate target cell membranes, forming pores that successively lyse the target cells [160,162]. These 2 pathways allow for the complement to execute its functions via stimulation of chemokine secretion [158]. The absence of this pathway in the host response to the  $\Delta mtp$ -mutant infection illustrates that the complement system pathway was activated due to the presence of MTP in the bacterium, hence further highlighting its role.

### 4.3.2 *The effect of MTP on host-pathogen interactions*

The pathogen's phenotype and the immune status of the host play a role in determining the specific host-pathogen interactions that follow after infection. Recognition of bacilli by the host cells is the start of the complex and dynamic host-pathogen interactions, that determine the outcome of the infection by pathogenic mycobacteria [158]. This PRR-dependent entry of the bacilli into host cells, which is key to the determination of the fate of *M. tuberculosis* upon infection has already been demonstrated to be affected by the MTP.

This, together with the further analysis of the host-pathogen interaction genes (Appendix 10) in the present study, showed that MTP affects the regulation and expression of these genes associated with host-pathogen interaction (Figure 3.8.4.2.1 and 3.8.4.1.1). This suggested that MTP plays a role in the specific host-pathogen interaction in infected Balb/C mice. Further analysis of the top 5 pathways (Dendritic Cell Maturation pathway, the triggering receptor expressed on myeloid cells 1 (TREM1) Signalling pathway, Role of NFAT in Regulation of the Immune Response pathway, Nuclear factor kappa-light-chain-enhancer of activated B cells (NFκB) signalling pathway and iCOS-iCOSL Signalling in T Helper Cells pathway) enriched by the host-pathogen interaction gene set suggested that MTP is an important antigenic protein, whose presence on the *M. tuberculosis* cell wall results in a specific interaction of the bacilli with the host, which leads to a specific response in the host, as shown below.

#### 4.3.2.1 *TREM1 Signalling pathway*

The TREM1 molecule is part of the Immunoglobulin (Ig) family of cell surface receptors, and is associated with the activation of pro-inflammatory immune responses [163]. It is associated with the cellular immune response pathways and cytokine signalling pathways, and its main functions include cell-to-cell signalling and interaction; Haematological system development and function and immune cell trafficking [163]. This was the most highly enriched pathway associated with genes involved in the host-pathogen interactions after WT infection in this study. The following molecules up-regulated during the WT infection are associated in the TREM1 signalling pathway: transmembrane receptors Cd40, Cd82, Cd83 and Cd86 molecules, Fc fragment of IgG receptor IIb (FCGR2B), Tlr1 and Tlr2 as well as the class II, major histocompatibility complex (MHC), trans-activator (CIITA) and NLR family, CARD domain containing 5 (Nlrc5), which are transcription regulators (Figure 3.8.4.2.1.2A). The stimulation of TREM1 observed in the WT infection by these transmembrane receptors activates the transcription regulators such as NF-κB. These, in turn, trigger the secretion of chemokines and cytokines like the monocyte chemotactic protein 1 (MCP-1), and the pro-inflammatory response (Figure 3.8.4.2.1.2D). This study also contributes to the knowledge on TREM1 as currently, its natural ligand is unknown.

#### 4.3.2.2 DC response

DCs are the major antigen presenting cells (APCs) that are associated with both the innate and adaptive immune systems [89]. Thus, their role during *M. tuberculosis* infection in the initiation of an immune response is expected [55]. DC responses are influenced by the specific infecting *M. tuberculosis* strain [55]. The initial events following *M. tuberculosis* infection also stimulate DCs and their response [55]. The functional capability of DCs is controlled and measured by their maturation state [164]. DC maturation results in the presentation of pathogen antigens and an upregulation of chemokine receptors, as well as activation of CD4<sup>+</sup> T cells.

In the current study, the DC Maturation pathway was enriched only during WT infection, whereas the SDEGs from the  $\Delta mtp$ -mutant infection did not enrich this pathway. The DC Maturation pathway was the most activated pathway by the genes unique to WT infection, whereas the genes unique to the  $\Delta mtp$ -mutant infection negatively enriched this pathway (Figure 3.8.2.2.1). One of its top functions include Cell-To-Cell signalling and Interaction [164]. The host-pathogen interaction gene set also showed this pathway as highly enriched (2<sup>nd</sup> most activated as seen in Figure 3.8.4.2.1.1), involving the most number of genes (Figure 3.8.2.8A). Antigens are captured by receptors on the plasma membrane of naïve DCs which thereafter undergo maturation.

In this study, the Fc fragment of Immunoglobulin G receptor Ia (Fcgr1), Fcgr3, Fcgr2b, Fcgr4 and Tlr2 were the receptors that were upregulated. This led to the upregulation of co-stimulatory molecules (CD40, CD80, CD83, CD86 and CD1d) during the maturation pathway (Figure 3.8.4.2.1.2B). This maturation is triggered by several factors like microbial antigens, other host immune cells and cytokines [164]. However, the absence of the DC maturation pathway in the host response to the  $\Delta mtp$ -mutant infection in our study suggests that the deletion of MTP has a deleterious effect on the factors that induce the maturation of DCs. Thus, the early interactions of *M. tuberculosis* that are specifically facilitated by MTP are important, and influence the initiation of a DC maturation response. Previously, the maturation of DCs *in vivo* was found to be mediated via TLR2 [157], and this study showed the upregulation of both TLR and DC maturation pathway, corroborating the involvement of these 2 pathways with each other. Inhibiting these process, as seen in the  $\Delta mtp$ -mutant infection leads to an unsuccessful immune response.

#### 4.3.2.3 Role of NFAT in Regulation of the Immune Response, NF-KB Signalling pathway and *iCOS-iCOSL* Signalling in T Helper Cells

The enrichment of these 3 pathways is similar to the 2 most highly enriched pathways, the TREM1 signalling and the DC Maturation Pathways, and affect the signalling and function of APCs (Figure

3.8.4.2.1.2C and Figure 3.8.4.2.1.2E). They also upregulated transmembrane molecules that interact with the *M. tuberculosis* antigens and elicit an immune response, i.e. the increased expression of CD80 and CD86 on APCs (Figure 3.8.4.2.1.2C), leading to the induction of signalling pathways that result in the secretion of NF- $\kappa$ B- and NFAT-inducible cytokines (Figures 3.8.4.2.1.2C-D), promoting the adaptive immune response [165,166]. The NF $\kappa$ B transcriptional factor regulates the TLR pathway which was positively enriched in this study as previously mentioned. Interestingly, NF $\kappa$ B was found to be upregulated after infection of epithelial cells with WT strain in a study conducted by Mvubu *et al* (2015), which also showed the positive enrichment of the TLR pathway [132]. Taken together, the absence of all these pathways in the host response to the  $\Delta mtp$ -mutant infection strongly implies that MTP plays a major role in the interactions that occur between the *M. tuberculosis* and its host.

### **4.3.3 MTP plays a role in the host immune response**

It is well known that pulmonary *M. tuberculosis* infection elicits intense innate and adaptive immune responses. The study of the host transcriptome after infection with *M. tuberculosis* has immensely improved the understanding of immunological mechanisms and cellular pathways of *M. tuberculosis* infection [6]. In this study, RNA was sequenced at an early time point, (Day14) which was reported to be the start of early adaptive immune responses [139]. Therefore, it was expected that the genes regulated in this study, would include those of the innate response, and some of the early adaptive immune response.

Overall, the GO, canonical pathway, network and upstream regulator analysis suggested that MTP has a significant impact on biological functions and pathways that are important for host immunity during *M. tuberculosis* infection, e.g. the complement system pathway was found not to be enriched during infection with the  $\Delta mtp$ -mutant, implying that MTP plays a role in the innate immune host defence against *M. tuberculosis*. These findings were further corroborated by our extended analysis of the SDEGs associated with host-pathogen interaction. The most highly activated pathway here, TREM1 signalling pathway is associated with various parts of both the innate and adaptive immune system [163]. The DC maturation pathway is involved in cellular immune response, cytokine signalling and pathogen-influenced signalling [164].

#### **4.3.3.1 Phagosome function pathway**

Pathogens reside within membrane-bound vacuoles inside macrophages, and these mature into phagosomes that assist in containing the infection. The Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway (Figure 3.8.5.2.3) is important in the maturation of the phagosomal compartment, and the killing mechanisms of the phagosome [47,48]. Another pathway that was enriched by the genes associated with host-pathogen interaction was the phagosome formation pathway.

Phagosome formation occurs when specific receptors on a phagocyte surface recognize ligands on the bacterial cell surface leading to the engulfment and containment of the bacilli into a vacuole identified as a phagosome, during *M. tuberculosis* infection [167]. The genes (C-type lectin domain family 7 member A (CLE7A), Fc fragment of IgG receptor I (FCGR1), FCGR 3, FCGR 2b, FCGR 4, TLR1 and TLR2) associated with this pathway were all down regulated during infection with the  $\Delta mtp$ -mutant strain (Figure 3.8.5.2.4), suggesting that MTP plays a role in the containment of infecting bacilli by the host via this pathway, and therefore it is involved in a specific immune response elicited by the host. Furthermore, the Production of Nitric Oxide and Reactive Oxygen Species in Macrophages pathway was enriched only in the WT infection. This correlates with the upregulation of genes involved in the phagosome formation pathway. *M. tuberculosis* growth inhibition that is dependent of Nitric oxide has been found to be linked to TLR activation in murine macrophages [157], and in this current study, the TLR signalling pathway was found to be enriched as well, corroborating with previous studies. This strengthens the hypothesis that MTP is involved in the anti-microbial activities elicited by the host upon *M. tuberculosis* infection.

#### 4.3.3.2 Antigen Presentation

Essential to the advancement of the innate immunity, and development of the adaptive immunity, is the presentation of processed antigens to T cells by APCs. APCs were observed in a few of the pathways that were enriched by the WT infection model. These included the Dendritic Cell (DC) Maturation pathway (Figure 3.8.4.2.1.2B), the Role of NFAT in Regulation of the Immune Response pathway (Figure 3.8.4.2.1.2C) and the iCOS-iCOSL Signalling in T Helper Cells pathway (Figure 3.8.4.2.1.2E). DCs are known to be the main APC and important initiators of immune responses that present antigens in a major histocompatibility complex (MHC) class I and II molecule -specific context [35]. During DC maturation, expression of MHC Class I and II molecules are upregulated, enhancing their capability to present antigens [164]. In this study, a negative enrichment was observed in the DC maturation pathway during infection with the  $\Delta mtp$ -mutant strain. This was followed by a negative enrichment in the antigen presentation pathway, which resulted from the downregulation of the genes involved (Figure 3.8.2.8A). The MHC complexes in the  $\Delta mtp$ -mutant infection were seen to be downregulated (Figure 3.8.5.1B). In contrast, peptide complexes, cytokines such as IFN- $\gamma$  and transcription regulators like NLRC5, that increase the innate and adaptive immune responses of the host [164] were upregulated in the WT infection (Figure 3.8.5.1A)

#### 4.3.3.3 Upstream regulators: Transcription factors

The activation of upstream regulators such as the transcription factors STAT1 and NF $\kappa$ B (Figure 3.8.5.1.1) in the WT infection resulted in a greater immune response elicited by the host, compared to the  $\Delta mtp$ -mutant infection. NF $\kappa$ B transcription factors play a significant role in the expression of pro-inflammatory genes including cytokines, chemokines, and adhesion molecules, thus contributing

significantly to the immune response [165]. The NF $\kappa$ B signalling pathway (Figure 3.8.4.2.1.2D) was positively enriched during WT infection, and negatively enriched in the  $\Delta mtp$ -mutant infection (Figure 3.8.2.2.1). STAT1 regulates cytokine production, and as observed in this study, the WT infection greatly upregulated STAT1, which resulted in transcription of a larger number of regulatory molecules than those regulated by STAT1 in the  $\Delta mtp$ -mutant infection (Figure 3.8.5.1.3).

The pathway involved in the multi-subunit nuclear factor of activated T cells (NFAT) transcription factor family in Regulation of the Immune Response was also downregulated during the  $\Delta mtp$ -mutant infection (Figure 3.8.4.2.1.1). This TF is also important in the expression of cytokines, and is therefore, important in immune response. These pathways all impact on the ability of the host to present the antigen to immune cells, as can be seen in the Antigen presentation pathway. As expected, this pathway was lowly enriched in the  $\Delta mtp$ -mutant infection, compared to the WT infection. These novel findings suggest that MTP plays a significant role in the presentation of *M. tuberculosis* antigens to immune cells by the host.

#### 4.3.3.3.1 The *M. tuberculosis* DosR transcription factor

Activation of the host immune system by host transcription factors also regulates bacterial gene expression in response to host immunity. For example, the bacterial transcription factor DosR controls expression of almost 50 genes in *M. tuberculosis* [144]. The regulation of this TF by bacilli infecting mice lungs was previously found to be dependent on the host immune system activation, specifically, the presence of NOS2 and IFNG [144]. In the present study, expression of these 2 was dependent on the presence of MTP, similar to the activation of the host immune system, hence suggesting that MTP plays a role on the regulation of host as well as bacterial TFs. However, this requires further interrogation.

#### 4.3.3.4 Cytokines and chemokines

Bacterial pili have also been reported to induce host immune response via cytokine/chemokine production [16]. MTP was reported not to have a significant effect on the cytokine production during infection of epithelial cells with *M. tuberculosis* [12,18]. In the present study, the difference in the cytokine genes regulated by the WT infection compared to those by the  $\Delta mtp$ -mutant infection, suggested that MTP plays a role in cytokine and chemokine production during *M. tuberculosis* infection in mice lungs. Infection with the WT strain of *M. tuberculosis* induced the expression of various cytokine genes, in contrast to the  $\Delta mtp$ -mutant infection (Figure 3.8.5.1.1A). The most regulated cytokines were Interferon Gamma (IFNG) and tumour necrosis factor (TNF), which were both highly activated during infection with WT, and repressed during  $\Delta mtp$ -mutant infection. These are major pro-inflammatory cytokines produced during *M. tuberculosis* infection, to control anti-mycobacterial

mechanisms of the host that could damage the host itself. They include the reactive oxygen intermediates (ROI) and reactive nitrogen intermediates (RNI) within phago-lysosomes [158]. In the current study, the 2 cytokines upregulated the expression of nitric oxide synthase 2 (NOS2) (Figure 3.8.5.2.2). Increasing RNI production facilitates the killing of *M. tuberculosis* intracellularly [158]. The production of TNF- $\alpha$  has been shown to be dependent on the expression of TLR2 [157] and hence our study also correlates with these previous findings as TNF was downregulated in the  $\Delta mtp$ -mutant infection, after a downregulation of TLR2 was observed. Furthermore, inflammatory response measured by the *in vivo* TNF response was directly affected by the presence of pili in *S. pneumoniae* [141]. Here, the authors found that compared with their respective wild-type strains, the pili deficient strains of *S. pneumoniae* caused a significantly lower TNF and IL-6 response [141]. Another study showed the activation of the IL-8 cytokine after recognition of curli pili of *Salmonella enterica* serotype Typhimurium by TLR2, showing evidence of the MTP induction of the immune system, supported by our current study [154].

The pathway most highly enriched by genes associated with host-pathogen interactions, TREM1 signalling pathway, activates expression and secretion of cytokines and chemokines, resulting in the expression of pro-inflammatory cytokines such as TNF, IL-1 $\beta$ , IL-6 and IL-18 [163]. The maturation of DCs is triggered by a number of cytokines such as IL-12, IL-15, IL-6, TNF and type I IFNs [164], and other cells of the innate immune system e.g. natural killer (NK) cells and other lymphocytes. These findings are in agreement with a previous study that demonstrated that curli pili have in *E. coli* can activate pro-inflammatory cytokines such as IL-6, IL-8, and TNF [168]. The *in vitro* study by Ramsugit *et al* (2015) showed contradictory findings on the ability of MTP to trigger a cytokine response. In their study, the epithelial cell cytokine production of TNF, MCP-1, IFN, the granulocyte colony-stimulating factor (G-CSF), IL-1, IL-4, IL-6 and IL-8 was tested. Only the production of TNF was significantly reduced after 48hrs of infection with the  $\Delta mtp$ -mutant strain. IL-4, IL-8, and IFN were also lowly induced, but not significantly less than the WT strain [18]. This difference could be a result of the different infection times used, as well as the infection models.

This study showed that the cytokine MIF (Macrophage Migration Inhibitory Factor) Regulation of Innate Immunity pathway was activated only in the WT infection, that is, in the presence of MTP (Figure 3.8.5.2.2). This pathway is an essential mediator of the innate immune system [169]. Furthermore, Crosstalk between Dendritic cells and Natural Killer Cells pathway was also positively enriched only during WT infection, and negatively enriched in the  $\Delta mtp$ -mutant infection (Figure 3.8.5.2.1). This pathway is associated with many cytokines and molecules that modulate the immune system [170], like IL-12, TNF and IFNs. These were upregulated in the WT infection, supporting the hypothesis that MTP plays a role in modulating the host immune system. Our current results support findings from other studies in that infection with the WT strain of *M. tuberculosis* initiated a cascade of

immune responses and inflammatory signals from the host. Some of cytokines have been previously implicated in studies of the host immune response to *M. tuberculosis* infection, for example, IFN, TNF, IL-4 receptor [171,172]. Furthermore, in a gene-expression profile study, these genes have been found to be preferentially upregulated in patients with active TB disease [171–173].

#### **4.4 MTP is associated with immune related networks and functions**

Network 3 lacked the calcifediol molecule in the  $\Delta mtp$ -mutant infection, and this resulted in its association with Drug Metabolism, Lipid Metabolism, Small Molecule Biochemistry whilst the WT infection was associated with Cancer, Organismal Injury and Abnormalities, Dermatological Diseases and Conditions. Similarly, the atypical chemokine receptor 1 (ACKR1) molecule was present in WT infection Network 4 (Figure 3.8.3.1.1A), but absent in the  $\Delta mtp$ -mutant, resulting in a difference in the network structure (Figure 3.8.3.1.1B), and the functions this network was associated with in the 2 infections. Network 4 in the WT infection was associated with Cellular Movement, Hematological System Development and Function, and Immune Cell Trafficking, whilst in the  $\Delta mtp$ -mutant it was associated in Cellular Compromise, Cell Signaling, Molecular Transport. The ACKR molecule is a receptor found on the plasma and endosome membrane, and is known to be involved in immune response activities such as stimulus response and signaling. It is also involved in the control of chemokine levels and localization [174]. In our study, the absence of MTP hindered the transcription of this molecule, resulting in the network 4 lacking the Immune Cell Trafficking function, implying that MTP plays a role in immune response.

From the unique genes, the 2 most statistically significant networks showed a similar result, where the WT associated networks were associated with immunologically related functions such as Immunological Disease, and antimicrobial inflammatory response, whereas the top  $\Delta mtp$ -mutant networks were not related to these functions (Table 3.8.3.2.1). This difference in the functions associated with networks of the shared and unique SDEGs demonstrate the association of MTP with immune related networks and functions. *M. tuberculosis* is known to elicit immune related networks that play major roles in the host response, such as natural killer (NK) cell activation and immune cell antibacterial activity [134].

Overall, the gene ontology, canonical pathway and network analysis in this study suggests that MTP has a significant impact on the biological functions, and pathways that are essential for host immunity during *M. tuberculosis* infection. Furthermore, these outcomes were supported by our comprehensive analysis, using z- scores of upstream regulators among the SDEGs that were affected by MTP. Taken together, the activation of a greater immune response to WT infection compared to  $\Delta mtp$ -mutant infection may explain the significantly higher increase in the size of lungs and spleen of WT compared to  $\Delta mtp$ -mutant infected mice. Pili have been shown to a play role in the response of mice to other

bacteriak21 infections. For example, during *S. pneumonia*, pili were found to be potent enhancers of the inflammatory response of the host [105,157], as seen in the current study.

#### **4.5 Conclusions**

The study findings indicate that while the removal of the gene encoding MTP does not completely prevent the growth of the bacilli, the lack of MTP expression does cause a growth deficit in the *in vitro* growth of *M. tuberculosis* under laboratory culture conditions. Thus, MTP does play a role in the growth of *M. tuberculosis in vitro*. MTP was demonstrated to play a significant role in the specific host-pathogen interactions that follow *M. tuberculosis* infection, resulting in host immune responses essential to the hosts' defence by triggering the innate immune response and inflammatory response. Taken together, transcriptome analyses of lung tissue infected with the MTP-deficient strain of the F15/LAM4/KZN family of *M. tuberculosis* has shown MTP to be a strong immunogen. Further studies using the *mtp*-complemented strain would allow for the exact determination of the direct relationship between MTP and the host response. This will be addressed post-Master's using RT-PCR. These findings provide further supporting evidence to previous studies that suggested that MTP is a good therapeutic and vaccine candidate.

#### **4.6 Limitations of study**

Due to financial constraints, RNA-Seq was performed in duplicate and on lung homogenates from only a single time interval. The RNA-Seq data were not validated by RT-qPCR due to time constraints. However, a previous study in our research group demonstrated good correlation between these 2 techniques. Several factors may have impacted on the variability observed in the bacterial lung implantation at 24hr, such as the difference in the initial inoculum used for aerosolization of the WT,  $\Delta mtp$ -mutant and complemented strains, as observed in the CFU/ml results of the inoculum. This was difficult to control, since care had to be exercised not to shear the MTP by vigorous methods to produce single cells. Despite the same inoculum being used for each strain, the 5 replicate mice displayed variability in the implantation at 24 hr, which suggests that the aerosolization technique by the equipment, the GlasCol, may not have been reliable. The use of 100  $\mu$ L of the organ homogenate for plating out for the CFUs may not have been optimal for determining the CFU/mL compared to a larger volume.

In addition, only half of the organ was homogenised for CFUs, as the other half was stored down for RNA extraction. Furthermore, the mice used were not pathogen free and a specific pathogen-free facility was not available to harbour the mice for the *in vivo* experiments. This may have impacted on the gene expression and implantation results. However, uninfected controls were used to normalize results against these background values.

It is recommended that the RNA be treated for genomic DNA that could contaminate downstream experiments, however, this step is optional. During our study, DNase treatment was causing RNA degradation, thus it was decided against. The RNA preparation method used in this study fractionates the RNA away from the DNA (Acid phenol in TRIzol). Furthermore, the RNeasy kit used to purify the RNA has a specific column that elutes RNA alone, thus the RNA was not treated for genomic DNA before sequencing. MOPS gels (Supplementary Figure A2 in Appendix) were used to check the quality of RNA prior to RNA Seq, and DNA contamination was not observed at this point. MOPS gels will reveal bright bands of DNA at the top of the gel as DNA sizes are large. Furthermore, the RNA samples had good A260/280 and high 260/230 ratios Nanodrop ratios, and RNA integrity numbers (RIN) between 6.90 and 9.80 (Table 3.5.1). Moreover, DNA contamination within the samples was removed during the mRNA isolation step.

While complementation did not restore growth *in vivo* in this study, in our previous *in vitro* studies the complementation was found to restore biofilm production as well as adhesion and invasion of *M. tuberculosis* on macrophages and epithelial cells [11,18,111]. The inability of the *mtp* complementation to restore the growth deficit observed in the *mtp* deficient strain suggests that the deletion of the *mtp* gene might have concurrently resulted in other changes in the genome, which the complementation did not restore. Such changes can be detected by whole genome sequencing which was not performed in this study, and has potential to reveal more evidence to the effect of the *mtp* gene if explored in the future.

Survival kinetics were not performed in this study, or prior to determine the virulence of the  $\Delta mtp$ -mutant strain. These studies are vital in determining the effect of MTP on virulence of *M. tuberculosis*. Furthermore, the infecting bacterial transcriptome response is key to ascertaining the overall effect of MTP, and therefore, future work should include elucidating the role of MTP in the regulation of the pathogen's transcriptome. The development of improved bioinformatics tools for use with deep-sequencing RNA technologies in future studies is essential for the analysis of the host and the pathogen transcriptomes simultaneously during infection.

#### **4.7 Recommendations for future work**

The effectiveness of any transcriptome profiling technique is analytically subject to the use of RNA samples that accurately reflect the actual ratios of the infected hosts' RNA from their tissues, thus RT-PCR should be performed on randomly selected genes to validate the RNA-Seq data. Since RNA-Seq was studied at only a single time point in this study, several identified genes/pathways of interest, such as NOS2 and IFNG can be interrogated further by RT-PCR to ascertain at which time point transcription occurs during infection. The organs had been stored for this purpose.

In addition, the certain pathways of interest can be further interrogated, such as the DC maturation pathway, where the maturation of DCs can be measured *in vitro* after infection with 2 strains. Other phenotypic assays can be carried out to explore pathways at different time intervals using the expression of molecules for example, the measurement of NO and ROS in macrophages infected with the 2 strains can be monitored at different time points. Furthermore, multiplex cytokine analysis using the BioPlex System can be performed on stored serum samples that had been collected at the different time points to confirm the gene expression findings.

## References

- [1] Glickman MS, Jacobs WR. Microbial Pathogenesis of *Mycobacterium tuberculosis*: Dawn of a Discipline. *Cell* 2001;104:477–85. doi:10.1016/j.rmed.2006.08.006.
- [2] WHO. Global Tuberculosis report 2015:1–204. doi:10.1007/s13398-014-0173-7.2.
- [3] WHO. Global Tuberculosis Report 2012 2012.
- [4] Gordon S V, Eiglmeier K, Garnier T, Brosch R, Parkhill J, Barrell B, et al. Genomics of *Mycobacterium bovis*. *Tuberculosis (Edinb)* 2001;81:157–63. doi:10.1054/tube.2000.0269.
- [5] Moodley P, Shah NS, Tayob N, Connolly C, Zetola N, Gandhi N, et al. Spread of extensively drug-resistant tuberculosis in KwaZulu-Natal province, South Africa. *PLoS One* 2011;6:e17513. doi:10.1371/journal.pone.0017513.
- [6] Kang DD, Lin Y, Moreno J-R, Randall TD, Khader SA. Profiling Early Lung Immune Responses in the Mouse Model of Tuberculosis. *PLoS One* 2011;6:e16161. doi:10.1371/journal.pone.0016161.
- [7] Klemm P, Schembri M a. Bacterial adhesins: function and structure. *Int J Med Microbiol* 2000;290:27–35. doi:10.1016/S1438-4221(00)80102-2.
- [8] Bermudez EL and, Goodman J. *Mycobacterium tuberculosis* invades and replicates within type II alveolar cells . *Infect Immun* 1996;64.
- [9] Ashiru OT, Pillay M, Sturm a W. Adhesion to and invasion of pulmonary epithelial cells by the F15/LAM4/KZN and Beijing strains of *Mycobacterium tuberculosis*. *J Med Microbiol* 2010;59:528–33. doi:10.1099/jmm.0.016006-0.
- [10] Alteri CJ, Xicohténcatl-Cortes J, Hess S, Caballero-Olín G, Girón J a, Friedman RL. *Mycobacterium tuberculosis* produces pili during human infection. *Proc Natl Acad Sci U S A* 2007;104:5145–50. doi:10.1073/pnas.0602304104.
- [11] Ramsugit S, Guma S, Pillay B, Jain P, Larsen MH, Danaviah S, et al. Pili contribute to biofilm formation in vitro in *Mycobacterium tuberculosis*. *Antonie Van Leeuwenhoek* 2013. doi:10.1007/s10482-013-9981-6.
- [12] Ramsugit S. The role of *Mycobacterium tuberculosis* pili ( MTP ) as an adhesin 2014:1–140.
- [13] Alteri C. Novel Pili of *Mycobacterium tuberculosis*. The University of Arizona., 2005.
- [14] Ramsugit S, Pillay M. Pili of *Mycobacterium tuberculosis*: current knowledge and future prospects. *Arch Microbiol* 2015. doi:10.1007/s00203-015-1117-0.
- [15] Ramsugit S, Pillay M. *Mycobacterium tuberculosis* Pili promote adhesion to and invasion of THP-1 macrophages. *Jpn J Infect Dis* 2014;67:476–8.
- [16] Kagnoff MF, Eckmann L. Epithelial cells as sensors for microbial infection. *J Clin Invest* 1997;100:6–10. doi:10.1172/JCI119522.
- [17] Rock RB, Hu S, Gekker G, Sheng WS, May B, Kapur V, et al. *Mycobacterium tuberculosis* – Induced Cytokine and Chemokine Expression by Human Microglia and Astrocytes : Effects of Dexamethasone 2005;192.

- [18] Ramsugit S, Pillay B, Pillay M. Evaluation of the role of *Mycobacterium tuberculosis* pili (MTP) as an adhesin, invasin, and cytokine inducer of epithelial cells. *Braz J Infect Dis* 2015. doi:10.1016/j.bjid.2015.11.002.
- [19] Mehrotra J, Bishai WR. Regulation of virulence genes in *Mycobacterium tuberculosis* 2001;182:171–82.
- [20] Albuquerque P, Paes HC, Tavares AH, Fernandes L, Bocca AL, Silva-pereira I, et al. Transcriptomics in Health and Disease 2014:265–87. doi:10.1007/978-3-319-11985-4.
- [21] Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol* 2008;26:1135–45.
- [22] Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 2008;9:387–402. doi:10.1146/annurev.genom.9.081307.164359.
- [23] Cook GM, Berney M, Gebhard S, Heinemann M, Robert A, Danilchanka O, et al. Physiology of Mycobacteria. *Adv Microb Physiol* 2009;55:81–319. doi:10.1016/S0065-2911(09)05502-7.Physiology.
- [24] Sakamoto K. The pathology of *Mycobacterium tuberculosis* infection. *Vet Pathol* 2012;49:423–39. doi:10.1177/0300985811429313.
- [25] World Health Organization. WHO | Stop TB Partnership. World Health Organization; 2006.
- [26] World Health Organization. Global TB factsheet 2005. World Health Organization; 2005.
- [27] Riska PF, Jacobs WR, Bloom BR, McKittrick J, Chan J. Specific identification of *Mycobacterium tuberculosis* with the luciferase reporter mycobacteriophage: use of p-nitro-alpha-acetyl-amino-beta-hydroxy propiophenone. *J Clin Microbiol* 1997;35:3225–31.
- [28] Lawn SD, Wood R, Wilkinson RJ. Changing concepts of “latent tuberculosis infection” in patients living with HIV infection. *Clin Dev Immunol* 2011;2011.
- [29] Health D of. The national HIV and syphilis prevalence survey, South Africa 2007. Pretoria: 2008.
- [30] Abdool Karim SS, Churchyard GJ, Karim QA, Lawn SD. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *Lancet (London, England)* 2009;374:921–33. doi:10.1016/S0140-6736(09)60916-8.
- [31] Ioerger TR, Koo S, No E-G, Chen X, Larsen MH, Jacobs WR, et al. Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa. *PLoS One* 2009;4:e7778. doi:10.1371/journal.pone.0007778.
- [32] Velayati AA, Masjedi MR, Farnia P, Tabarsi P, Ghanavi J, ZiaZarifi AH, et al. Emergence of new forms of totally drug-resistant tuberculosis bacilli: Super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran. *Chest* 2009;136:420–5.
- [33] Migliori GB, De Iaco G, Besozzi G, Centis R, Cirillo DM. First tuberculosis cases in Italy resistant to all tested drugs. *Euro Surveill Bull Eur Sur Les Mal Transm = Eur Commun Dis Bull* 2007;12.
- [34] Klopper M, Warren RM, Hayes C, van Pittius NCG, Streicher EM, M??ller B, et al. Emergence

- and spread of extensively and totally drug-resistant tuberculosis, South Africa. *Emerg Infect Dis* 2013;19:449–55.
- [35] Udwardia ZF, Amale RA, Ajbani KK, Rodrigues C. Totally drug-resistant tuberculosis in India. *Clin Infect Dis* 2012;54:579–81. doi:10.1093/cid/cir889.
- [36] Udwardia ZF. Totally drug-resistant tuberculosis in India: Who let the djinn out? *Respirology* 2012;17:741–2.
- [37] Statistics South Africa. Statistical release Mortality and causes of death in South Africa , 2011 Findings from death notification. 2014.
- [38] Coscolla M, Gagneux S. Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Semin Immunol* 2014;26:431–44. doi:10.1016/j.smim.2014.09.012.
- [39] Ford C, Yusim K, Ioerger T, Feng S, Chase M, Greene M, et al. *Mycobacterium tuberculosis*--heterogeneity revealed through whole genome sequencing. *Tuberculosis (Edinb)* 2012;92:194–201. doi:10.1016/j.tube.2011.11.003.
- [40] Rose G, Cortes T, Comas I, Coscolla M, Gagneux S, Young DB. Mapping of genotype-phenotype diversity among clinical isolates of *Mycobacterium tuberculosis* by sequence-based transcriptional profiling. *Genome Biol Evol* 2013;5:1849–62. doi:10.1093/gbe/evt138.
- [41] Victor TC, de Haas PEW, Jordaan AM, van der Spuy GD, Richardson M, van Soolingen D, et al. Molecular characteristics and global spread of *Mycobacterium tuberculosis* with a western cape F11 genotype. *J Clin Microbiol* 2004;42:769–72.
- [42] Gandhi NR, Brust JCM, Moodley P, Weissman D, Heo M, Ning Y, et al. Minimal diversity of drug-resistant *Mycobacterium tuberculosis* strains, South Africa. *Emerg Infect Dis* 2014;20:426–33. doi:10.3201/eid2003.131083.
- [43] Gandhi NR, Moll A, Sturm AW, Pawinski R, Govender T, Lalloo U, et al. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet (London, England)* 2006;368:1575–80. doi:10.1016/S0140-6736(06)69573-1.
- [44] Smith KLJ, Saini D, Bardarov S, Larsen M, Frothingham R, Gandhi NR, et al. Reduced virulence of an extensively drug-resistant outbreak strain of *Mycobacterium tuberculosis* in a murine model. *PLoS One* 2014;9:e94953. doi:10.1371/journal.pone.0094953.
- [45] Pillay M, Sturm a W. Evolution of the extensively drug-resistant F15/LAM4/KZN strain of *Mycobacterium tuberculosis* in KwaZulu-Natal, South Africa. *Clin Infect Dis* 2007;45:1409–14. doi:10.1086/522987.
- [46] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998;393:537–44. doi:10.1038/31159.
- [47] Ruiru Shi and Isamu Sugawara. Pathophysiology of Tuberculosis. In: Mahboub B, editor. *Tuberc. - Curr. Issues Diagnosis Manag., InTech*; 2013, p. 127–40. doi:10.5772/56396.

- [48] Dannenberg, Jr. AM. Pathogenesis of Human Pulmonary Tuberculosis: Insights from the Rabbit Model. Washington, DC: American Society of Microbiology; 2006. doi:10.1128/9781555815684.
- [49] Kumar A, Farhana A, Guidry L, Saini V, Hondalus M, Steyn AJC. Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert Rev Mol Med* 2011;13:1–25.
- [50] Subbian S, Bandyopadhyay N, Tsenova L, O’Brien P, Khetani V, Kushner NL, et al. Early innate immunity determines outcome of *Mycobacterium tuberculosis* pulmonary infection in rabbits. *Cell Commun Signal* 2013;11:60. doi:10.1186/1478-811X-11-60.
- [51] Krishnan N. A transcriptomic approach for studying the activation of dendritic cells in response to mycobacterial infections. University College London, 2007.
- [52] Gengenbacher M, Kaufmann SHE. *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiol Rev* 2012;36:514–32. doi:10.1111/j.1574-6976.2012.00331.x.
- [53] Barry CE, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, et al. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol* 2009;7:845–55. doi:10.1038/nrmicro2236.
- [54] Brighenti S, Lerm M. How *Mycobacterium tuberculosis* Manipulates Innate and Adaptive Immunity – New Views of an Old Topic. *Underst Tuberc – Anal Orig Mycobacterium Tuberc Pathog* 2012:208–34. doi:urn:nbn:se:liu:diva-75444 LA - English.
- [55] O’Garra A, Redford PS, McNab FW, Bloom CI, Wilkinson RJ, Berry MPR. The immune response in tuberculosis. *Annu Rev Immunol* 2013;31:475–527. doi:10.1146/annurev-immunol-032712-095939.
- [56] Smith I. *Mycobacterium tuberculosis* Pathogenesis and Molecular Determinants of Virulence. *Clin Microbiol Rev* 2003;16:463–96. doi:10.1128/CMR.16.3.463.
- [57] Lappin MB, Campbell JD. The Th1-Th2 classification of cellular immune responses: concepts, current thinking and applications in haematological malignancy. *Blood Rev* 2000;14:228–39. doi:10.1054/blre.2000.0136.
- [58] SCHLUGER NW, ROM WN. The Host Immune Response to Tuberculosis. *Am J Respir Crit Care Med* 1998;157:679–91. doi:10.1164/ajrccm.157.3.9708002.
- [59] Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol* 2010;11:373–84.
- [60] Amaral EP, Lasunskaja EB, D’Imp??rio-Lima MR. Innate immunity in tuberculosis: How the sensing of mycobacteria and tissue damage modulates macrophage death. *Microbes Infect* 2016;18:11–20. doi:10.1016/j.micinf.2015.09.005.
- [61] Bermudez LE, Sangari FJ, Kolonoski P, Petrofsky M, Goodman J. The Efficiency of the Translocation of *Mycobacterium tuberculosis* across a Bilayer of Epithelial and Endothelial Cells as a Model of the Alveolar Wall Is a Consequence of Transport within Mononuclear Phagocytes and Invasion of Alveolar Epithelial Cells. *Infect Immun* 2002;70:140–6.

- doi:10.1128/IAI.70.1.140.
- [62] Glynn JR, Whiteley J, Bifani PJ, Kremer K, van Soolingen D. Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg Infect Dis* 2002;8:843–9. doi:10.3201/eid0805.020002.
- [63] Lin Y, Zhang M, Barnes PF. Chemokine production by a human alveolar epithelial cell line in response to *Mycobacterium tuberculosis*. *Infect Immun* 1998;66:1121–6.
- [64] Sharma M, Sharma S, Roy S, Varma S, Bose M. Pulmonary epithelial cells are a source of interferon- $\gamma$  in response to *Mycobacterium tuberculosis* infection. *Immunol Cell Biol* 2007;85:229. doi:10.1038/sj.icb.7100037.
- [65] Hall-Stoodley L, Watts G, Crowther JE, Balagopal A, Torrelles JB, Robison-Cox J, et al. *Mycobacterium tuberculosis* binding to human surfactant proteins A and D, fibronectin, and small airway epithelial cells under shear conditions. *Infect Immun* 2006;74:3587–96. doi:10.1128/IAI.01644-05.
- [66] Pasula R, Downing JF, Wright JR, Kachel DL, Davis TE, Martin WJ. Surfactant Protein A (SP-A) Mediates Attachment of *Mycobacterium tuberculosis* to Murine Alveolar Macrophages. *Am J Respir Cell Mol Biol* 1997;17:209–17. doi:10.1165/ajrcmb.17.2.2469.
- [67] Dubnau E, Smith I. *Mycobacterium tuberculosis* gene expression in macrophages 2003;5:629–37.
- [68] Chatterjee D, Khoo KH. Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. *Glycobiology* 1998;8:113–20.
- [69] Nathan C, Shiloh MU. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci U S A* 2000;97:8841–8.
- [70] Welin A. Survival strategies of *Mycobacterium tuberculosis* inside the human macrophage. 2011.
- [71] Armstrong JA, Hart PD. Response of cultured macrophages to *Mycobacterium tuberculosis*, with observations on fusion of lysosomes with phagosomes. *J Exp Med* 1971;134:713–40.
- [72] Govender VS, Ramsugit S, Pillay M. *Mycobacterium tuberculosis* adhesins: Potential biomarkers as anti-tuberculosis therapeutic and diagnostic targets. *Microbiol (United Kingdom)* 2014;160:1821–31. doi:10.1099/MIC.0.082206-0.
- [73] Henderson RA, Watkins SC, Flynn JL. Activation of human dendritic cells following infection with *Mycobacterium tuberculosis*. *J Immunol* 1997;159:635–43.
- [74] Giacomini E, Iona E, Ferroni L, Miettinen M, Fattorini L, Orefici G, et al. Infection of Human Macrophages and Dendritic Cells with *Mycobacterium tuberculosis* Induces a Differential Cytokine Gene Expression That Modulates T Cell Response. *J Immunol* 2001;166:7033–41. doi:10.4049/jimmunol.166.12.7033.
- [75] Tascon RE, Soares CS, Ragno S, Stavropoulos E, Hirst EM, Colston MJ. *Mycobacterium tuberculosis*-activated dendritic cells induce protective immunity in mice. *Immunology*

- 2000;99:473–80.
- [76] Nakayama M. Antigen Presentation by MHC-Dressed Cells. *Front Immunol* 2014;5:672. doi:10.3389/fimmu.2014.00672.
- [77] Crevel R Van. Innate immunity to *Mycobacterium tuberculosis*. *Clin Microbiol ...* 2002;15:294–309. doi:10.1128/CMR.15.2.294.
- [78] Tailleux L, Schwartz O, Herrmann J-L, Pivert E, Jackson M, Amara A, et al. DC-SIGN is the major *Mycobacterium tuberculosis* receptor on human dendritic cells. *J Exp Med* 2003;197:121–7.
- [79] Kleinnijenhuis J, Oosting M, Joosten LAB, Netea MG, Crevel R Van. Innate Immune Recognition of *Mycobacterium tuberculosis* 2011;2011. doi:10.1155/2011/405310.
- [80] Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol* 2001;2:675–80. doi:10.1038/90609.
- [81] Li Y, Wang Y, Liu X. The role of airway epithelial cells in response to mycobacteria infection. *Clin Dev Immunol* 2012;2012.
- [82] Lerner TR, Borel S, Gutierrez MG. The innate immune response in human tuberculosis. *Cell Microbiol* 2015;17:1277–85. doi:10.1111/cmi.12480.
- [83] Hossain MM, Norazmi M-N. Pattern recognition receptors and cytokines in *Mycobacterium tuberculosis* infection--the double-edged sword? *Biomed Res Int* 2013;2013:179174. doi:10.1155/2013/179174.
- [84] Schlesinger LS, Bellinger-Kawahara CG, Payne NR, Horwitz MA. Phagocytosis of *Mycobacterium tuberculosis* is mediated by human monocyte complement receptors and complement component C3. *J Immunol* 1990;144:2771–80.
- [85] Kang PB, Azad AK, Torrelles JB, Kaufman TM, Beharka A, Tibesar E, et al. The human macrophage mannose receptor directs *Mycobacterium tuberculosis* lipoarabinomannan-mediated phagosome biogenesis. *J Exp Med* 2005;202:987–99. doi:10.1084/jem.20051239.
- [86] Mortaz E, Adcock IM, Tabarsi P, Masjedi MR, Mansouri D, Velayati AA, et al. Interaction of Pattern Recognition Receptors with *Mycobacterium tuberculosis*. *J Clin Immunol* 2015;35:1–10. doi:10.1007/s10875-014-0103-7.
- [87] Sasindran SJ, Torrelles JB. *Mycobacterium tuberculosis* infection and inflammation : what is beneficial for the host and for the bacterium ? 2011;2:1–16. doi:10.3389/fmicb.2011.00002.
- [88] Aderem A, Ulevitch RJ. Toll-like receptors in the induction of the innate immune response. *Nature* 2000;406:782–7.
- [89] Flores ODC. Innate and adaptive immune responses in the lungs. Contribution to protection against mycobacterial infections. 2011.
- [90] Harding C V, Boom WH. Regulation of antigen presentation by *Mycobacterium tuberculosis*: a role for Toll-like receptors. *Nat Rev Microbiol* 2010;8:296–307. doi:10.1038/nrmicro2321.
- [91] Kawai T, Akira S. The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int Immunol*

- 2009;21:317–37. doi:10.1093/intimm/dxp017.
- [92] Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell* 2010;140:805–20. doi:10.1016/j.cell.2010.01.022.
- [93] Takeda K. Toll-like receptors in innate immunity. *Int Immunol* 2004;17:1–14. doi:10.1093/intimm/dxh186.
- [94] Franchi L, Warner N, Viani K, Nuñez G. Function of Nod-like receptors in microbial recognition and host defense. *Immunol Rev* 2009;227:106–28. doi:10.1111/j.1600-065X.2008.00734.x.
- [95] Ferwerda G, Girardin SE, Kullberg B-J, Le Bourhis L, de Jong DJ, Langenberg DML, et al. NOD2 and Toll-Like Receptors Are Nonredundant Recognition Systems of *Mycobacterium tuberculosis*. *PLoS Pathog* 2005;1:e34. doi:10.1371/journal.ppat.0010034.
- [96] Brooks MN, Rajaram MVS, Azad AK, Amer AO, Valdivia-Arenas MA, Park J-H, et al. NOD2 controls the nature of the inflammatory response and subsequent fate of *Mycobacterium tuberculosis* and *M. bovis* BCG in human macrophages. *Cell Microbiol* 2011;13:402–18. doi:10.1111/j.1462-5822.2010.01544.x.
- [97] da Silva Neto BR, de Fátima da Silva J, Mendes-Giannini MJS, Lenzi HL, de Almeida Soares CM, Pereira M. The malate synthase of *Paracoccidioides brasiliensis* is a linked surface protein that behaves as an anchorless adhesin. *BMC Microbiol* 2009;9:272. doi:10.1186/1471-2180-9-272.
- [98] Pizarro-Cerdá J, Cossart P. Bacterial Adhesion and Entry into Host Cells. *Cell* 2006;124:715–27. doi:10.1016/j.cell.2006.02.012.
- [99] Kumar S, Puniya BL, Parween S, Nahar P, Ramachandran S. Identification of Novel Adhesins of *M. tuberculosis* H37Rv Using Integrated Approach of Multiple Computational Algorithms and Experimental Analysis. *PLoS One* 2013;8:e69790. doi:10.1371/journal.pone.0069790.
- [100] Pethe K, Alonso S, Biet F, Delogu G, Brennan MJ, Locht C, et al. The heparin-binding haemagglutinin of *M. tuberculosis* is required for extrapulmonary dissemination. *Nature* 2001;412:190–4. doi:10.1038/35084083.
- [101] Diaz-Silvestre H, Espinosa-Cueto P, Sanchez-Gonzalez A, Esparza-Ceron MA, Pereira-Suarez AL, Bernal-Fernandez G, et al. The 19-kDa antigen of *Mycobacterium tuberculosis* is a major adhesin that binds the mannose receptor of THP-1 monocytic cells and promotes phagocytosis of mycobacteria. *Microb Pathog* 2005;39:97–107. doi:10.1016/j.micpath.2005.06.002.
- [102] Kinkhikar AG, Vargas D, Li H, Mahaffey SB, Hinds L, Belisle JT, et al. *Mycobacterium tuberculosis* malate synthase is a laminin-binding adhesin. *Mol Microbiol* 2006;60:999–1013. doi:10.1111/j.1365-2958.2006.05151.x.
- [103] Ramsugit S, Pillay M. Identification of *Mycobacterium tuberculosis* adherence-mediating components: a review of key methods to confirm adhesin function. *Iran J Basic Med Sci* 2016;19:579–84.
- [104] Jonas K, Tomenius H, Kader A, Normark S, Römling U, Belova LM, et al. Roles of curli,

- cellulose and BapA in *Salmonella* biofilm morphology studied by atomic force microscopy. *BMC Microbiol* 2007;7:70. doi:10.1186/1471-2180-7-70.
- [105] Proft T, Baker EN. Pili in Gram-negative and Gram-positive bacteria - Structure, assembly and their role in disease. *Cell Mol Life Sci* 2009;66:613–35. doi:10.1007/s00018-008-8477-4.
- [106] Danne C, Dramsi S. Pili of Gram-positive bacteria: roles in host colonization. *Res Microbiol* 2012;163:645–58. doi:10.1016/j.resmic.2012.10.012.
- [107] Rock RB, Hu S, Gekker G, Sheng WS, May B, Kapur V, et al. *Mycobacterium tuberculosis*-induced cytokine and chemokine expression by human microglia and astrocytes: effects of dexamethasone. *J Infect Dis* 2005;192:2054–8. doi:10.1086/498165.
- [108] Zhang M, Lin Y, Iyer D V, Gong J, Abrams JS, Barnes PF. T-cell cytokine responses in human infection with *Mycobacterium tuberculosis*. *Infect Immun* 1995;63:3231–4.
- [109] Torres M, Herrera T, Villareal H, Rich EA, Sada E. Cytokine profiles for peripheral blood lymphocytes from patients with active pulmonary tuberculosis and healthy household contacts in response to the 30-kilodalton antigen of *Mycobacterium tuberculosis*. *Infect Immun* 1998;66:176–80.
- [110] Hammar M, Arnqvist A, Bian Z, Olsén A, Normark S. Expression of two *csg* operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. *Mol Microbiol* 1995;18:661–70.
- [111] Naidoo N, Ramsugit S, Pillay M, Parsons LM, Somoskövi A, Gutierrez C, et al. *Mycobacterium tuberculosis* pili (MTP), a putative biomarker for a tuberculosis diagnostic test. *Tuberculosis* 2014;94:338–45. doi:10.1016/j.tube.2014.03.004.
- [112] Cooper BAM, Dalton DK, Stewart T a, Griffin JP, Russell DG, Orme IM. Disseminated Tuberculosis in Interferon  $\gamma$  Gene-disrupted Mice. *J Exp Med* 1993;178:2243–7. doi:10.1084/jem.178.6.2243.
- [113] Ashiru OT, Pillay M, Sturm AW. *Mycobacterium tuberculosis* isolates grown under oxygen deprivation invade pulmonary epithelial cells. *Anaerobe* 2012;18:471–4.
- [114] Mehrotra J, Bishai WR. Regulation of virulence genes in *Mycobacterium tuberculosis*. *Int J Med Microbiol* 2001;291:171–82. doi:10.1078/1438-4221-00113.
- [115] Ordway DJ, Orme IM. Animal Models of Mycobacteria Infection. *Curr Protoc Immunol.*, 2011, p. CHAPTER: Unit–19.5. doi:10.1002/0471142735.im1905s30.Animal.
- [116] Nalpas NC, Park SDE, Magee DA, Taraktsoglou M, Browne JA, Conlon KM, et al. Whole-transcriptome, high-throughput RNA sequence analysis of the bovine macrophage response to *Mycobacterium bovis* infection in vitro. *BMC Genomics* 2013;14:1471.
- [117] Skvortsov T a, Ignatov D V, Majorov KB, Apt a S, Azhikina TL. *Mycobacterium tuberculosis* Transcriptome Profiling in Mice with Genetically Different Susceptibility to Tuberculosis. *Acta Naturae* 2013;5:62–9.
- [118] McLoughlin KE, Nalpas NC, Rue-Albrecht K, Browne JA, Magee DA, Killick KE, et al. RNA-

- seq Transcriptional Profiling of Peripheral Blood Leukocytes from Cattle Infected with *Mycobacterium bovis*. *Front Immunol* 2014;5:396. doi:10.3389/fimmu.2014.00396.
- [119] Rachman H, Strong M, Ulrichs T, Grode L, Schuchhardt J, Mollenkopf H, et al. Unique transcriptome signature of *Mycobacterium tuberculosis* in pulmonary tuberculosis. *Infect Immun* 2006;74:1233–42. doi:10.1128/IAI.74.2.1233-1242.2006.
- [120] Jacobsen M, Repsilber D, Gutschmidt A, Neher A, Feldmann K, Mollenkopf HJ, et al. Candidate biomarkers for discrimination between infection and disease caused by *Mycobacterium tuberculosis*. *J Mol Med* 2007;85:613–21. doi:10.1007/s00109-007-0157-6.
- [121] Gonzalez-Juarrero M, Kingry LC, Ordway DJ, Henao-Tamayo M, Harton M, Basaraba RJ, et al. Immune response to *Mycobacterium tuberculosis* and identification of molecular markers of disease. *Am J Respir Cell Mol Biol* 2009;40:398–409. doi:10.1165/rcmb.2008-0248OC.
- [122] Mollenkopf H-J, Hahnke K, Kaufmann SHE. Transcriptional responses in mouse lungs induced by vaccination with *Mycobacterium bovis* BCG and infection with *Mycobacterium tuberculosis*. *Microbes Infect* 2006;8:136–44. doi:10.1016/j.micinf.2005.06.015.
- [123] Mandlik A, Swierczynski A, Das A, Ton-That H. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol* 2008;16:33–40. doi:10.1016/j.tim.2007.10.010.
- [124] Larsen MH, Biermann K, Jacobs WR. Laboratory Maintenance of *Mycobacterium tuberculosis* 2007:1–5. doi:10.1002/9780471729259.mc10a01s6.
- [125] Naidoo CC, Pillay M. Increased in vitro fitness of multi- and extensively drug-resistant F15/LAM4/KZN strains of *Mycobacterium tuberculosis*. *Clin Microbiol Infect* 2013. doi:10.1111/1469-0691.12415.
- [126] Ordway DJ, Orme IM, Oberhelman RA, Soto-castellares G, Gilman RH, Caviedes L, et al. NIH Public Access 2011;10:612–20. doi:10.1016/S1473-3099(10)70141-9.Diagnostic.
- [127] Leggett RM, Ramirez-Gonzalez RH, Clavijo BJ, Waite D, Davey RP. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. *Front Genet* 2013;4:288. doi:10.3389/fgene.2013.00288.
- [128] Nishimura M, Tanaka S, Ihara F, Muroi Y, Yamagishi J, Furuoka H, et al. Transcriptome and histopathological changes in mouse brain infected with *Neospora caninum*. *Sci Rep* 2015;5:7936. doi:10.1038/srep07936.
- [129] Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012;7:562–78. doi:10.1038/nprot.2012.016.
- [130] Trapnell C, Williams B a, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 2010;28:511–5. doi:10.1038/nbt.1621.
- [131] Chen F, Zhang C, Jia X, Wang S, Wang J, Chen Y, et al. Transcriptome Profiles of Human Lung

- Epithelial Cells A549 Interacting with *Aspergillus fumigatus* by RNA-Seq. *PLoS One* 2015;10:e0135720. doi:10.1371/journal.pone.0135720.
- [132] Mvubu NE, Pillay B, Gamielidien J, Bishai W, Pillay M. Canonical pathways, networks and transcriptional factor regulation by clinical strains of *Mycobacterium tuberculosis* in pulmonary alveolar epithelial cells. *Tuberculosis* 2016;97:73–85. doi:10.1016/j.tube.2015.12.002.
- [133] Lipner EM, Garcia BJ, Strong M. Network Analysis of Human Genes Influencing Susceptibility to Mycobacterial Infections. *PLoS One* 2016;11:e0146585. doi:10.1371/journal.pone.0146585.
- [134] Subbian S, O'Brien P, Kushner NL, Yang G, Tsenova L, Peixoto B, et al. Molecular immunologic correlates of spontaneous latency in a rabbit model of pulmonary tuberculosis. *Cell Commun Signal* 2013;11:16. doi:10.1186/1478-811X-11-16.
- [135] Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009;25:1105–11. doi:10.1093/bioinformatics/btp120.
- [136] Zhang J. Transcriptome Analysis Reveals Novel Entry Mechanisms and a Central Role of SRC in Host Defense during High Multiplicity Mycobacterial Infection. *PLoS One* 2013;8:e65128. doi:10.1371/journal.pone.0065128.
- [137] Beisiegel M, Mollenkopf H, Hahnke K, Koch M, Dietrich I, Reece ST, et al. Combination of host susceptibility and *Mycobacterium tuberculosis* virulence define gene expression profile in the host 2009;3369–84. doi:10.1002/eji.200939615.
- [138] Velayati AA, Farnia P, Masjedi MR. Pili in totally drug resistant *Mycobacterium tuberculosis* (TDR-TB). *Int J Mycobacteriology* 2012;1:57–8. doi:10.1016/j.ijmyco.2012.04.002.
- [139] Shepelkova G, Pommerenke C, Alberts R, Geffers R, Evstifeev V, Apt A, et al. Analysis of the lung transcriptome in *Mycobacterium tuberculosis*-infected mice reveals major differences in immune response pathways between TB-susceptible and resistant hosts. *Tuberculosis (Edinb)* 2013;93:263–9. doi:10.1016/j.tube.2012.11.007.
- [140] Mueller-Ortiz SLL, Sepulveda E, Olsen MRR, Jagannath C, Wanger ARR, Norris SJJ. Decreased infectivity despite unaltered C3 binding by a Delta hbhA mutant of *Mycobacterium tuberculosis*. *Infect Immun* 2002;70:6751. doi:10.1128/IAI.70.12.6751.
- [141] Barocchi MA, Ries J, Zogaj X, Hemsley C, Albiger B, Kanth A, et al. A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci U S A* 2006;103:2857–62. doi:10.1073/pnas.0511017103.
- [142] Azhikina T, Skvortsov T. *Mycobacterium tuberculosis* Transcriptome In Vivo Studies – A Key to Understand the Pathogen Adaptation Mechanism. *In Vivo (Brooklyn)* 1998.
- [143] Mvubu NE, Pillay B, Gamielidien J, Bishai W, Pillay M. Canonical pathways, networks and transcriptional factor regulation by clinical strains of *Mycobacterium tuberculosis* in pulmonary alveolar epithelial cells. *Tuberculosis* 2015;97:73–85. doi:10.1016/j.tube.2015.12.002.
- [144] Schnappinger D, Schoolnik GK, Ehrhart S. Expression profiling of host pathogen interactions: how *Mycobacterium tuberculosis* and the macrophage adapt to one another. *Microbes Infect*

- 2006;8:1132–40. doi:10.1016/j.micinf.2005.10.027.
- [145] Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol* 2016;17:13. doi:10.1186/s13059-016-0881-8.
- [146] Bah SY, Dickinson P, Forster T, Kampmann B, Ghazal P. Immune oxysterols: Role in mycobacterial infection and inflammation. *J Steroid Biochem Mol Biol* 2016. doi:10.1016/j.jsbmb.2016.04.015.
- [147] Rivas-Santiago B, Schwander SK, Sarabia C, Diamond G, Klein-Patel ME, Hernandez-Pando R, et al. Human {beta}-defensin 2 is expressed and associated with *Mycobacterium tuberculosis* during infection of human alveolar epithelial cells. *Infect Immun* 2005;73:4505–11. doi:10.1128/IAI.73.8.4505-4511.2005.
- [148] Rivas-Santiago B, Hernandez-Pando R, Carranza C, Juarez E, Contreras JL, Aguilar-Leon D, et al. Expression of Cathelicidin LL-37 during *Mycobacterium tuberculosis* Infection in Human Alveolar Macrophages, Monocytes, Neutrophils, and Epithelial Cells. *Infect Immun* 2008;76:935–41. doi:10.1128/IAI.01218-07.
- [149] Tomlinson GS, Bell LCK, Walker NF, Tsang J, Brown JS, Breen R, et al. HIV-1 infection of macrophages dysregulates innate immune responses to *Mycobacterium tuberculosis* by inhibition of interleukin-10. *J Infect Dis* 2014;209:1055–65. doi:10.1093/infdis/jit621.
- [150] Lee S-W, Wu LS-H, Huang G-M, Huang K-Y, Lee T-Y, Weng JT-Y, et al. Gene expression profiling identifies candidate biomarkers for active and latent tuberculosis. *BMC Bioinformatics* 2016;17:S3. doi:10.1186/s12859-015-0848-x.
- [151] Cliff JM, Cho J-E, Lee J-S, Ronacher K, King EC, van Helden P, et al. Excessive Cytolytic Responses Predict Tuberculosis Relapse After Apparently Successful Treatment. *J Infect Dis* 2016;213:485–95. doi:10.1093/infdis/jiv447.
- [152] Ingenuity Pathway Analysis. Canonical Pathway: Agranulocyte Adhesion and Diapedesis | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3A8tsnh>.
- [153] Mogensen TH. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin Microbiol Rev* 2009;22:240–73, Table of Contents. doi:10.1128/CMR.00046-08.
- [154] Tükel C, Raffatellu M, Humphries AD, Wilson RP, Andrews-Polymenis HL, Gull T, et al. CsgA is a pathogen-associated molecular pattern of *Salmonella enterica* serotype Typhimurium that is recognized by Toll-like receptor 2. *Mol Microbiol* 2005;58:289–304. doi:10.1111/j.1365-2958.2005.04825.x.
- [155] Ingenuity Pathway Analysis. Canonical Pathway: Toll-like Receptor Signaling | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3Aachz>.
- [156] Pan F, Zhao Y, Zhu S, Sun C, Lei L, Feng X, et al. Different transcriptional profiles of RAW264.7 infected with *Mycobacterium tuberculosis* H37Rv and BCG identified via deep sequencing. *PLoS One* 2012;7:e51988. doi:10.1371/journal.pone.0051988.

- [157] Drennan MB, Nicolle D, Quesniaux VJF, Jacobs M, Allie N, Mpagi J, et al. Toll-like receptor 2-deficient mice succumb to *Mycobacterium tuberculosis* infection. *Am J Pathol* 2004;164:49–57. doi:10.1016/S0002-9440(10)63095-7.
- [158] Gupta A, Kaul A, Tsolaki AG, Kishore U, Bhakta S. *Mycobacterium tuberculosis*: immune evasion, latency and reactivation. *Immunobiology* 2012;217:363–74. doi:10.1016/j.imbio.2011.07.008.
- [159] Lintner KE, Wu YL, Yang Y, Spencer CH, Hauptmann G, Hebert LA, et al. Early Components of the Complement Classical Activation Pathway in Human Systemic Autoimmune Diseases. *Front Immunol* 2016;7:36. doi:10.3389/fimmu.2016.00036.
- [160] Manivannan S, Rao VN, Ramanathan VD. Interaction between mycobacteria and macrophages through the complement system. *Int J Pharma Bio Sci* n.d.;5:B129–39.
- [161] Boshra H, Gelman AE, Sunyer JO. Structural and Functional Characterization of Complement C4 and C1s-Like Molecules in Teleost Fish: Insights into the Evolution of Classical and Alternative Pathways. *J Immunol* 2004;173:349–59. doi:10.4049/jimmunol.173.1.349.
- [162] Ingenuity Pathway Analysis. Canonical Pathway: Complement System | IPA. FFFF.svnRevisionFFFF 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3Axy6l> (accessed August 9, 2016).
- [163] Ingenuity Pathway Analysis. Canonical Pathway: TREM1 Signaling | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3A3j6wq>.
- [164] Ingenuity Pathway Analysis. Canonical Pathway: Dendritic Cell Maturation | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3A3ttx1>.
- [165] Ingenuity Pathway Analysis. Canonical Pathway: NF-κB Signaling | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3Acj4>.
- [166] Ingenuity Pathway Analysis. Canonical Pathway: Role of NFAT in Regulation of the Immune Response | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3A3m04r>.
- [167] Ramachandra L, Smialek JL, Shank SS, Convery M, Boom WH, Harding C V. Phagosomal processing of *Mycobacterium tuberculosis* antigen 85B is modulated independently of mycobacterial viability and phagosome maturation. *Infect Immun* 2005;73:1097–105. doi:10.1128/IAI.73.2.1097-1105.2005.
- [168] Bian Z, Brauner A, Li Y, Normark S. Expression of and cytokine activation by Escherichia coli curli fibers in human sepsis. *J Infect Dis* 2000;181:602–12. doi:10.1086/315233.
- [169] Calandra T, Froidevaux C, Martin C, Roger T. Macrophage migration inhibitory factor and host innate immune defenses against bacterial sepsis. *J Infect Dis* 2003;187 Suppl:S385-90. doi:10.1086/374752.
- [170] Ingenuity Pathway Analysis. Canonical Pathway: Natural Killer Cell Signaling | IPA 2016. <https://reports.ingenuity.com/rs/report/cpathway?id=ING%3A1bk4k>.
- [171] Mistry R, Cliff JM, Clayton CL, Beyers N, Mohamed YS, Wilson PA, et al. Gene-Expression

- Patterns in Whole Blood Identify Subjects at Risk for Recurrent Tuberculosis. *J Infect Dis* 2007;195:357–65. doi:10.1086/510397.
- [172] Kaufmann SHE. How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol* 2001;1:20–30. doi:10.1038/35095558.
- [173] Cliff JM, Andrade INJ, Mistry R, Clayton CL, Lennon MG, Lewis AP, et al. Differential Gene Expression Identifies Novel Markers of CD4+ and CD8+ T Cell Activation Following Stimulation by *Mycobacterium tuberculosis*. *J Immunol* 2004;173:485–93. doi:10.4049/jimmunol.173.1.485.
- [174] Bonecchi R, Graham GJ. Atypical Chemokine Receptors and Their Roles in the Resolution of the Inflammatory Response. *Front Immunol* 2016;7:224. doi:10.3389/fimmu.2016.00224.

## Appendices

### Appendix 1: Media

#### A. Supplemented Middlebrook 7H9 broth (+ 10% OADC + 0.5% glycerol + 0.05% Tween 80)

4.71g Middlebrook 7H9 powder

100 ml OADC

10 ml of 50% (w/v) glycerol

2.5 ml 20% Tween-80.

900ml distilled water

4.71g of Middlebrook 7H9 powder was dissolved in approximately 900mL of distilled water and autoclaved at 121°C for 15 minutes. 10ml of 50% (w/v) glycerol, 2.5ml of 20% Tween-80 and 100ml of OADC were added after cooling and media was stored at 4°C.

#### B. Supplemented Middlebrook 7H11 Agar (+ 10% OADC + 0.5% glycerol)

21g Middlebrook 7H11 powder

900 ml distilled water

100 ml OADC

10 ml of 50% (w/v) glycerol

21g of Middlebrook 7H11 powder was dissolved in approximately 900mL of distilled water and autoclaved at 121°C for 15 minutes. 10ml of 50% (w/v) glycerol and 100ml of OADC were added after cooling the media\*, and 12.5mL was aliquoted into 65mm petri dishes, and stored at 4°C.

\*For the selective media the following antibiotics were added at this point:

200.000 units/L of Polymixin B

20 mg/L of Amphotericin B

100 mg/L of Carbenicilin

20 mg/L of Trimethoprim

#### C. 20% Tween-80

20 mL Tween-80 into

80 mL distilled water

Tween-80 was dissolved in water and sterilized by filtration through 0.22-µm membrane.

#### D. 50% (w/v) Glycerol

50g glycerol

100ml distilled water

50 g of glycerol was dissolved in approximately 80ml of autoclaved distilled water and the volume was brought up to 100ml. The solution was sterilized by filtration through a 0.22 $\mu$ m membrane into a sterile container.

E. Phosphate buffered saline (PBS) (+ 0.05% tween)

10PBS tablets (Oxoid)

1000ml distilled water

Ten PBS tablets were dissolved in 1000ml autoclaved distilled water. The PBS was autoclaved at 121 $^{\circ}$ C for 15 minutes. 2.5 ml 20% Tween 80 was added, and thereafter decanted into 20ml aliquots and refrigerated at 4 $^{\circ}$ C until use.

F. Diethylpyrocarbonate (DEPC) treated water

1ml 0.1% DEPC

1 L Distilled water

1ml of 0.1% DEPC was added to 1L of distilled water and left at room temperature for overnight. The water was autoclaved at 121 $^{\circ}$ C for 15 minutes and allowed to cool before use.

H. 0.5 M EDTA

9.305 g EDTA

40 ml DEPC water

EDTA was dissolved in DEPC water, and brought to 50 ml volume with DEPC water. Thereafter, the solution was autoclaved.

I. 5 X MOPS Buffer, 1 L

41.86 g MOPS

4.115 g sodium acetate

800 ml DEPC water

10 ml 0.5 M Ethylenediaminetetraacetic acid (EDTA)

800ml of DEPC treated water was added to MOPS, EDTA and sodium acetate, and the volume was made up to 1 L with DEPC water. 200 ml aliquots were made and the solution was autoclaved for 15 min at 121  $^{\circ}$ C.

J. 37 % formaldehyde

1.85 g paraformaldehyde

3.5 mL H<sub>2</sub>O

90  $\mu$ L 1 N NaOH.

Water was added to paraformaldehyde and heated in a boiling water bath. NaOH was added and the mixture was agitated for ~1 min. Thereafter, it was cooled until running water and was sterilized by filtration through a 0.22µm membrane

## Appendix 2: MOPS 1% agarose gel

0.5g agarose

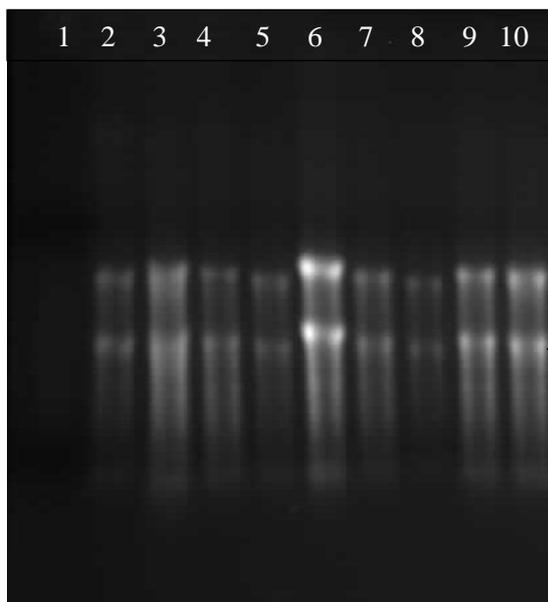
10ml 10x MOPS buffer

36 ml DEPC treated water

9 ml Formaldehyde (37%)

5 µl Gel red

1. 0.5 g of agarose was heated in microwave oven 36 ml DEPC water until dissolved.
2. 5 ml of 10x MOPS running buffer and 9 ml 37% formaldehyde was added in a fume hood after cooling a 'hand-warm' temperature.
3. Gel red was added to the agarose and gently swirled to mix
4. The gel was poured into a casting tray, avoiding air bubbles, and left to set for 30 minutes at room temperature.
5. To 6 µl of each RNA sample, 2 µl of loading dye was added, and the mixture was heated at 70°C for 10 minutes and cooled on ice for 2 minutes.
6. The samples were loaded into gel and ran at 90V for 1 hours in 1xMOPS buffer.
7. Bands were visualized.



Lanes:

1. Molecular Weight Marker
2. M3 D14  $\Delta mtp$ -mutant
3. M3 D21  $\Delta mtp$ -mutant
4. M3 D14 WT
5. M4 D14  $\Delta mtp$ -mutant
6. M4 D21  $\Delta mtp$ -mutant
7. M4 D14 WT
8. M5 D14  $\Delta mtp$ -mutant
9. M5 D21  $\Delta mtp$ -mutant
10. M5 D14 WT

**Figure A2: MOPS RNA gel image of RNA isolated from BALB/c mice infected lungs.** Each time point had 5 independent biological replicates, and 2 biological replicates with the highest RIN numbers were selected for sequencing. RNA from the lungs of mice challenged with the WT and the  $\Delta mtp$ -mutant for 14 and 21 days was extracted. M3, M4, M5 represent Mouse replicate 3, 4 and 5 respectively. The intensity of the bands corresponds to the concentration of the RNA sample. Only RNA from Day 14 was sequenced.

### **Appendix 3: RNA-Seq data analysis pipeline**

#### **Setting up the machine/server and installing software**

Create a directory called bin in the Desktop where you will unzip all your software.

>Install Java in the Linux computer/server (from application centre)

>Install FastsX tool kit (download the zip file and unzip it within the command line)

#### **Commands used:**

>Install FastQC within the command line using: **sudo apt-get install fastqc**

>Download zipped files of bowtie2, tophat, cufflinks and unzip them from a command line

>Install Tophat: **sudo apt-get install tophat,**

    samtools: **sudo apt-get install samtools**

    cufflinks: **sudo apt-get install cufflinks**

#### **Concatenating files:**

1. create a new file/directory: **mkdir newfile**

2. joined Read 1 and Read 2 inside this file, and named them mtp3\_R1.gz and mtp3\_R2.gz respectively.

```
$ cat part1.fastq.gz part2.fastq.gz >merged_file.fastq.gz
```

e.g. \$ cat SKuv13-1\_S1\_L004\_R2\_001.fastq.gz SKuv13-2\_S1\_L004\_R2\_001.fastq.gz

```
>WT13_R2.fastq.gz
```

NB: Keep the data files in a zipped format, as most of the NGS commands will take zipped formats. If you have to unzip for something specifically, unzip the file and keep the original zipped, use, then delete it.

#### **Quality analysis**

Run FastQC using the following command: **fastqc ./sample\_name.fastq**

Asses the quality scores of each read

### Uploading files into the server:

**\$ scp -r path\_for\_file\_to\_upload username@server\_address:/home/georgina**

e.g. \$ scp -r /home/ub2/Desktop/RNA-Seq\_Data/RNA/ georgina@146.230.128.25:/home/georgina

### Downloading files from the server:

**\$ scp -r username@server\_address:file\_to\_be\_copied/ where-to-copy**

e.g.\$ scp -r georgina@146.230.128.25:/home/georgina/Mycobacterium\_tuberculosis\_H37RV\_tar.gz  
/home/ub2/Documents/Georgina\_RNA-Seq\_Data

### Downloading files i.e. reference genome from the web to the server:

1. log into the server and use the command: **\$ wget [OPTION]... [URL]...**

e.g. For Mouse genome: wget ftp://igenome:G3nom3s4u@ussd-ftp.illumina.com/Mus\_musculus/UCSC/mm10/Mus\_musculus\_UCSC\_mm10.tar.gz

### Unzipping the ref genome:

**\$ tar -zxvf yourdownload.tar.gz**

e.g. \$ tar -zxvf Mus\_musculus\_UCSC\_mm10.tar.gz

NB: this reference genome contains all the important files required for the analysis, and these are already standardized e.g. the Bowtie2index is there already

### Running Tophat2:

>Put the path for tophat when running on the laptop

>In this example, brain is 1 variable, and adrenal is another. Run tophat using the following command:

**\$ tophat (-p 4) -o output\_sampleA\_dir -G genes.gtf --transcriptome-index cds  
Homo\_sapiens/UCSC/hg38/Sequence/Bowtie2Index/genome sampleA\_1.fq.gz sampleA\_2.fq.gz**

(this is for paired end (two files from each end) but for a single end sequencing, use each file separately)

e.g. \$ tophat2 -o tophat2\_output\_brain -G  
/home/georgina/Homo\_sapiens/UCSC/hg38/Annotation/Genes/genes.gtf --transcriptome-  
index=known/mrna /home/georgina/Homo\_sapiens/UCSC/hg38/Sequence/Bowtie2Index/genome  
brain\_1.fastq brain\_2.fastq

NB -index=known/mrna creates a transcriptome index and puts it in a directory called 'known', and the index files will have the suffix 'mrna'.

NB when running Tophat for subsequent samples, i.e. adrenal, remove the -G and the .gtf file because the transcriptome index has already been made:

```
e.g. $ tophat2 -o tophat2_output_adrenal --transcriptome-index=known/mrna
/home/georgina/Homo_sapiens/UCSC/hg38/Sequence/Bowtie2Index/genome adrenal_1.fastq.fastq
adrenal_2.fastq
```

Just make sure you use the same transcriptome index, and put the right paths.

### **Running Cuffdiff:**

put the path for cuffdiff when running on the laptop

```
$ cuffdiff -o outputname -L variable1,variable2 -b .fa_file .gtf_file accepted_hits.bam_variable1
accepted_hits.bam_variable2
```

```
e.g. $ cuffdiff -o brain_vs_adrenal -L brain,adrenal -b
/home/georgina/Homo_sapiens/UCSC/hg38/Sequence/WholeGenomeFasta/genome.fa
/home/georgina/Homo_sapiens/UCSC/hg38/Annotation/Genes/genes.gtf
tophat2_output_brain/accepted_hits.bam tophat2_output_adrenal/accepted_hits.bam
```

### **Queueing jobs on the server**

Use the command: **sbatch run\_scriptname-to-be-run**

You use 2 different files:

1. a .sh file containing the command you want to run. This file will specify the paths where the jobs will be run from, and where the output will go e.g. tophat\_mtp1.sh and cuff\_UN\_WT.sh
  2. a .sh file containing the script to run the 1st file e.g. run\_tophat\_mtp1.sh and run\_cuff\_UN\_WT.sh
- >>these files run together and must correspond

### **Editing the script file:**

1. Change -p value to the correct one for that job i.e. for the small dataset change it from -p 140 to -p 20. Note: -p value has a minimum of 20
2. Change the file names within the script i.e. input and output file names, and give them their right, specific paths. Change the output file name within the script and give them their right, specific paths i.e. the output file should be specific for the job being run so that there is no confusion when handling the outputs later. Especially with Tophat jobs because each sample will have its own Tophat run.
3. Make sure the command is right according to the job you are running, e.g. if running Tophat, check with the Tophat command under the 'Running Tophat2' section.
4. Make sure you have all the required arguments in your command, e.g. if running Tophat for the first time, you will need to have the "-G \_\_\_\_ .gtf" argument.
5. Double check the paths of all your commands i.e. your .gtf, .fa, Index paths.
6. Change the name of the script file and save it to be specific to which job its running.

### **Editing the run file:**

1. Change the name of the script file within the run file. Do not change anything in the run file except the name of the script file.

>>the nodes value corresponds to the -p value. For every 20 of the -p, there is 1 of the nodes. So for -p 140, nodes=7

>also change the name of the run file to fit what job its running.

Upload the new scripts into the server where it will run from and give the edited script files permission using: **\$ chmod u+x filename**

### **Submitting to the queue:**

Run the new edited run script: **\$ sbatch new-run-file**

You will get an output file called "slurm\_\_\_\_.out" once your job starts running and it will show you the progress of your job: **\$ more slurm\_\_\_\_.out**

>always open this file to make sure your job is running and that the script has permission before you leave it

### **Running MeV on Linux:**

navigate to the MeV\_4\_8 folders in the terminal: **/home/Documents/MeV\_4\_8**

make tmev.sh executable: **chmod u+x tmev.sh**

run MeV: **./tmev.sh**

MeV can hereafter be run by double-clicking the tmev.sh file.

### **Downstream Analysis**

1. Filter gene\_exp.diff files for all conditions for significant genes, and genes with a log1.5 fold change up/down regulated.

2. Draw Venn diagrams for interested conditions

3. Save the genes of interest to be used for heat maps in a .txt file. Get the log1.5fold changes for each gene using the following script:

```
$ grep -w -f file_input_with_genes_of_interest.txt A549_vs_Beijing/gene_exp.diff | awk -F "\t" '$10 >= 2 || $10 <= -2 {print $10}'> Output.2fold_change.txt
```

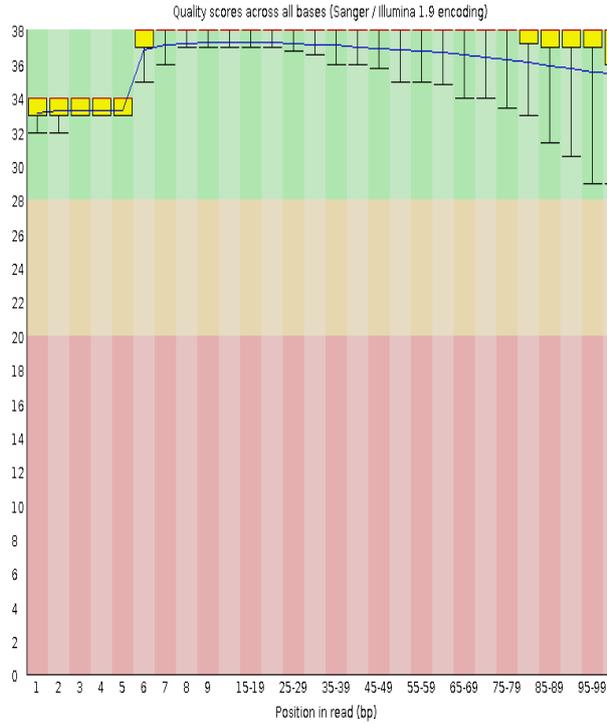
```
eg1: $ grep -w -f UIvsWT_UIvsmtmp_input.txt cuffdiff_trial_UI_vs_WT/gene_exp.diff | awk -F "\t" '$10 >= 1.5 || $10 <= -1.5 {print $10}'> UIvsWT_UIvsmtmp_output.txt
```

and

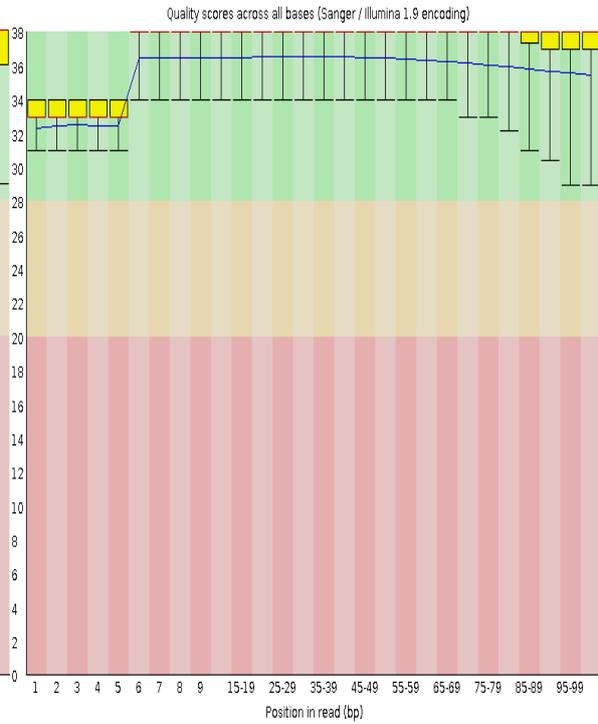
```
eg2: $ grep -w -f UIvsWT_UIvsmtmp_input.txt cuffdiff_trial_UI_vs_mtp/gene_exp.diff | awk -F "\t" '$10
>= 1.5 || $10 <= -1.5 {print $10}'> UIvsWT_UIvsmtmp_output_b.txt
```

## Appendix 4: Read QC

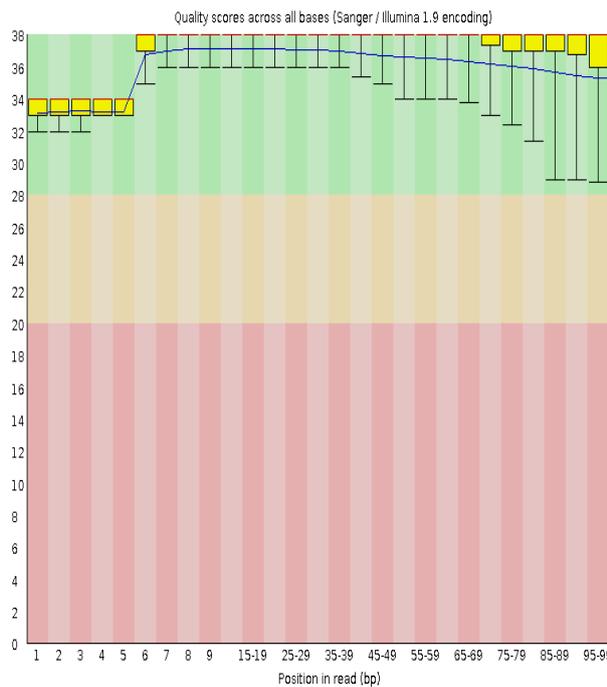
**1i.)**



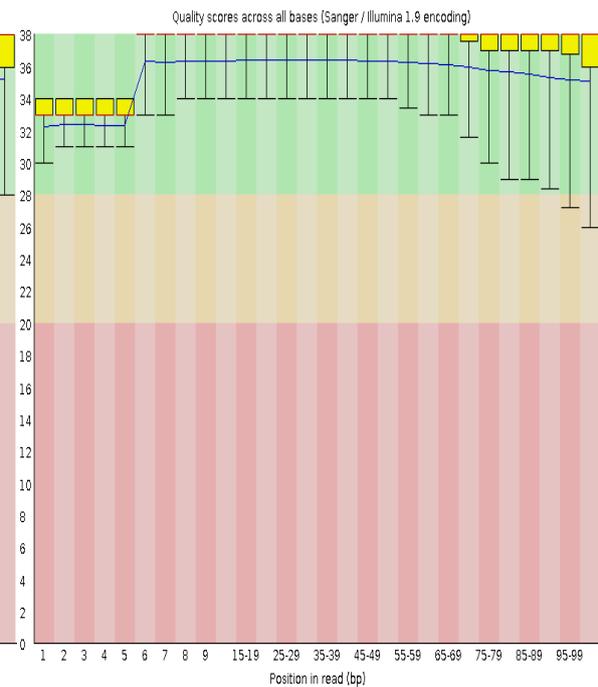
**ii.)**



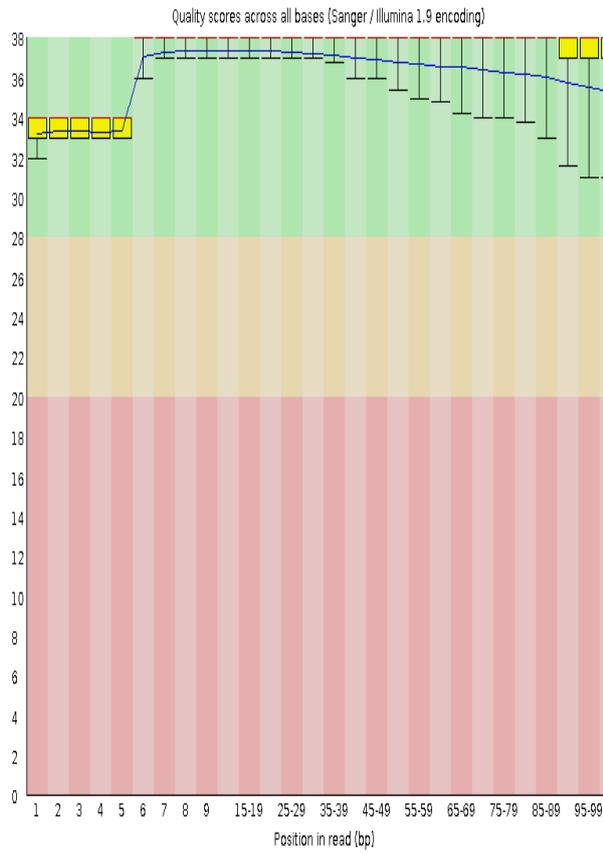
**2i)**



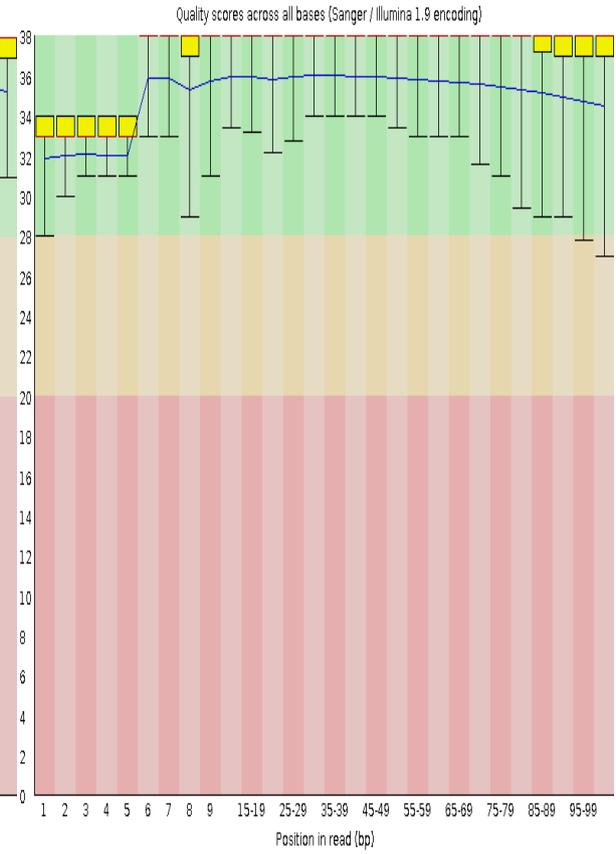
**ii)**



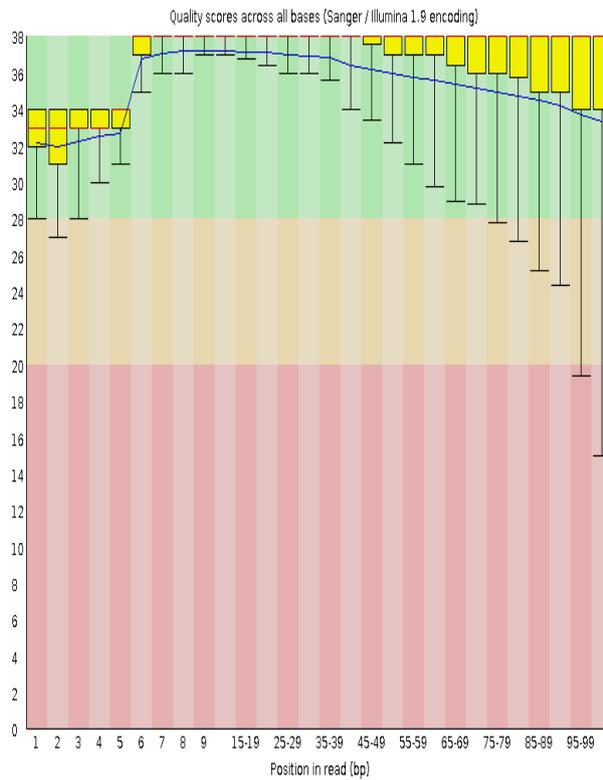
**3i.)**



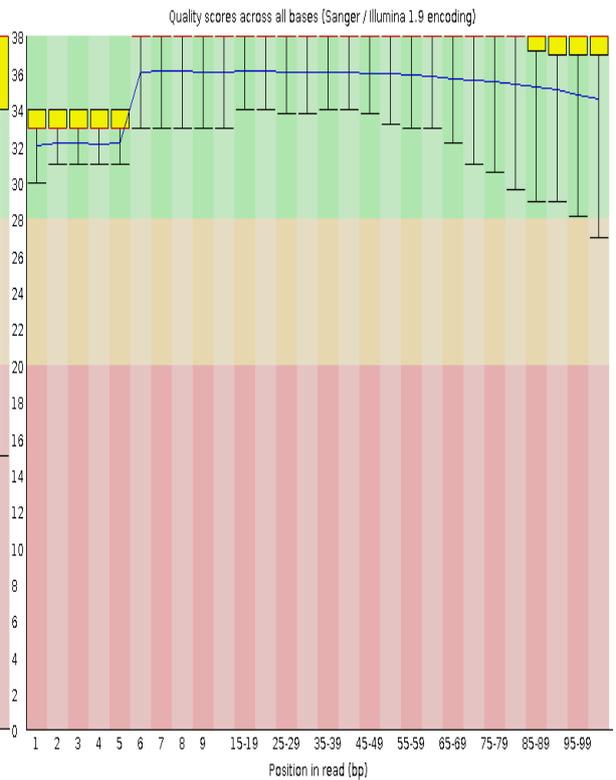
**ii.)**



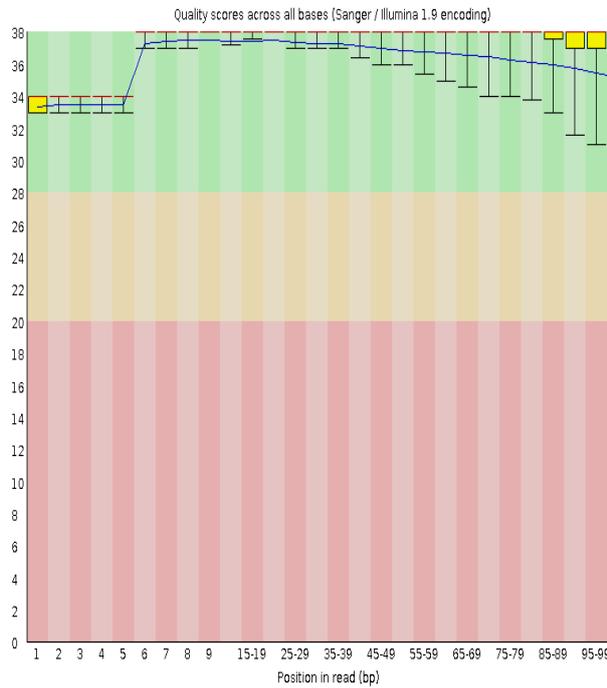
**4i.)**



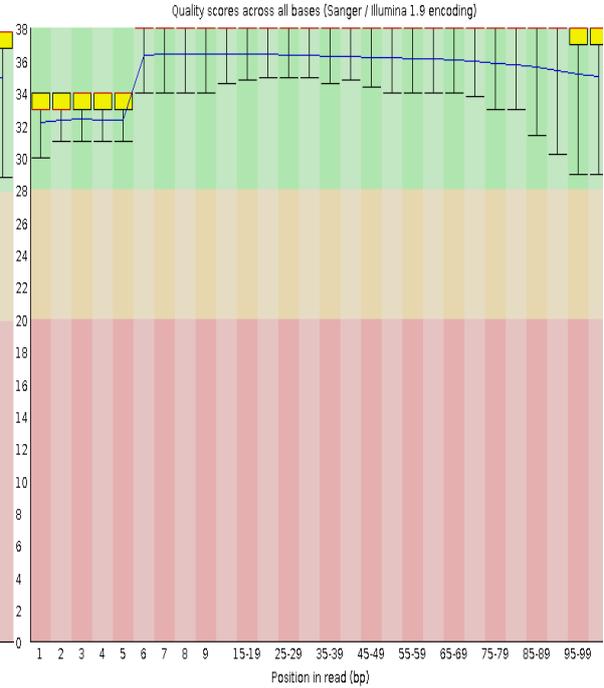
**ii.)**



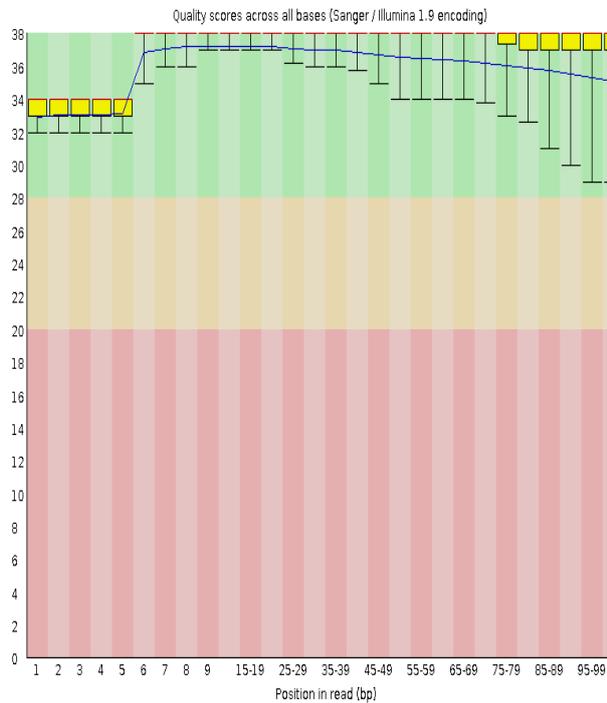
5i.)



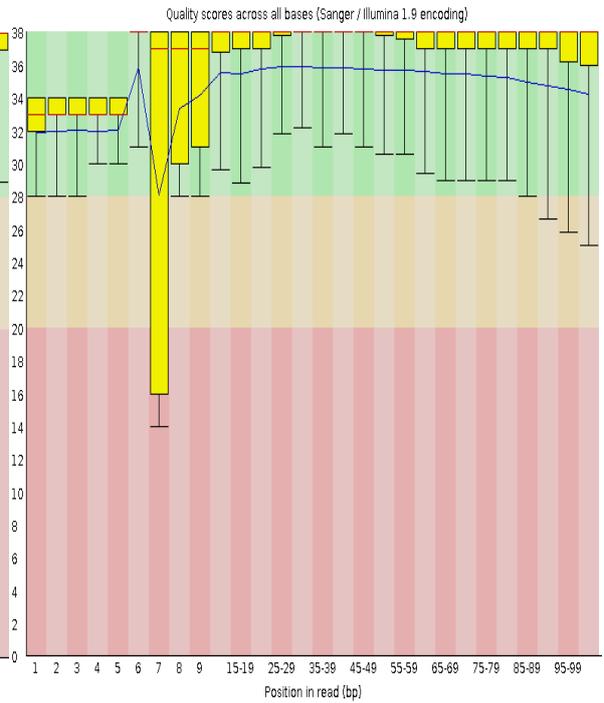
ii.)



6i.)



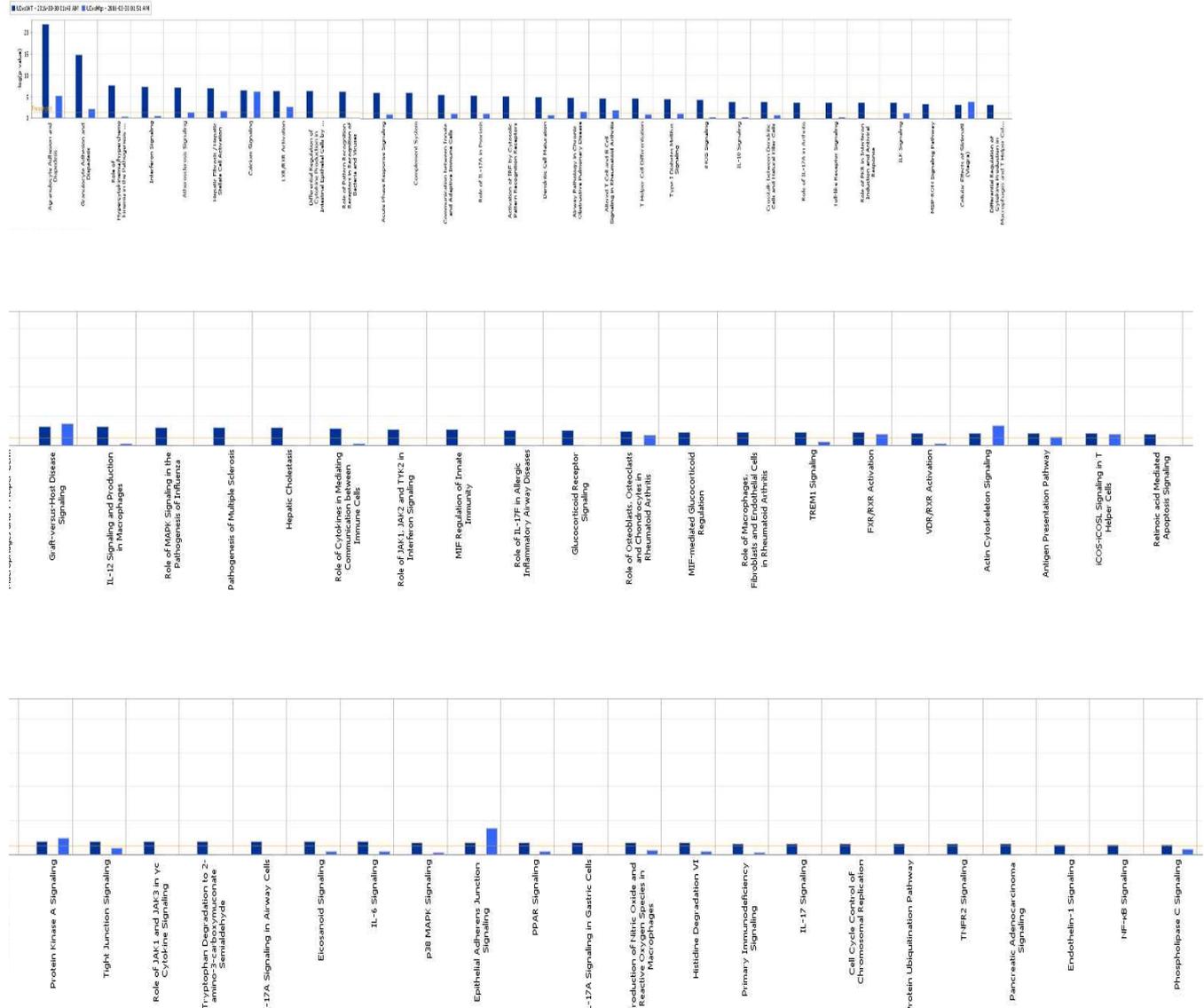
ii.)

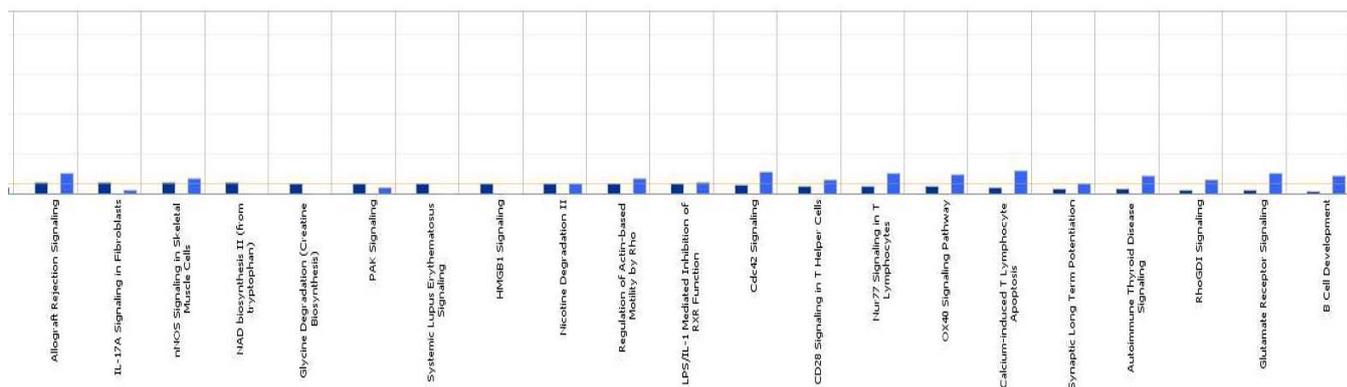


**Figure A4: FastQC output of the Illumina reads.** Forward (i), and reverse (ii) reads from the 6 libraries were generated from the RNA obtained from the 1 and 2) uninfected, 3 and 4) WT and 5 and 6)  $\Delta mtp$ -mutant infected lungs. The mean quality score across the sequence length is represented by the

blue line. The green, orange and pink colours represent areas of good, acceptable and bad quality scores respectively.

### Appendix 5: IPA bar graphs





**Figure A5: IPA canonical pathways enriched from all genes, in response to the WT and  $\Delta mtp$ -mutant infection relative to the uninfected.** Dark blue represents WT and light blue represents  $\Delta mtp$ -mutant (p-value < 0.05).

## Appendix 6: IPA functions

**Table A6: Functions associated with pathways enriched by SDEGs from the WT infection.**

WT infection		
Category	Function Annotation	p-value
Inflammatory Response	inflammation of joint	8.87E-29
Inflammatory Response	inflammation of organ	2.43E-20
Inflammatory Response	inflammation of body region	1.17E-18
Inflammatory Response	inflammation of body cavity	5.48E-17
Inflammatory Response	inflammation of respiratory system component	4.05E-16
Inflammatory Response	inflammation of intestine	3.77E-15
Inflammatory Response	inflammation of large intestine	7.60E-15
Inflammatory Response	inflammation of lung	9.36E-15
Inflammatory Response	inflammation of central nervous system	1.98E-14
Inflammatory Response	inflammation of liver	1.68E-13
Inflammatory Response	inflammation of secretory structure	6.03E-08
Inflammatory Response	inflammation of pancreas	8.04E-08
Inflammatory Response	inflammatory response	2.18E-26
Inflammatory Response	antimicrobial response	1.03E-24
Inflammatory Response	activation of leukocytes	7.62E-24
Inflammatory Response	activation of phagocytes	2.27E-20
Inflammatory Response	activation of myeloid cells	4.66E-20
Inflammatory Response	activation of antigen presenting cells	7.84E-19

Inflammatory Response	activation of macrophages	5.48E-18
Inflammatory Response	activation of mononuclear leukocytes	5.00E-13
Inflammatory Response	activation of lymphocytes	2.04E-12
Inflammatory Response	activation of neutrophils	2.46E-12
Inflammatory Response	activation of T lymphocytes	1.00E-11
Inflammatory Response	activation of microglia	3.16E-10
Inflammatory Response	quantity of phagocytes	1.25E-22
Inflammatory Response	quantity of dendritic cells	9.67E-13
Inflammatory Response	quantity of neutrophils	4.06E-12
Inflammatory Response	quantity of macrophages	1.06E-08
Inflammatory Response	quantity of monocytes	1.45E-07
Inflammatory Response	rheumatoid arthritis	4.88E-21
Inflammatory Response	cell movement of phagocytes	5.32E-20
Inflammatory Response	cell movement of dendritic cells	1.89E-13
Inflammatory Response	cell movement of neutrophils	6.08E-13
Inflammatory Response	cell movement of macrophages	2.69E-11
Inflammatory Response	cell movement of monocytes	1.19E-08
Inflammatory Response	cell movement of bone marrow-derived macrophages	4.18E-07
Inflammatory Response	immune response of leukocytes	4.61E-19
Inflammatory Response	immune response of cells	2.90E-17
Inflammatory Response	immune response of T lymphocytes	1.62E-14
Inflammatory Response	immune response of antigen presenting cells	3.09E-11
Inflammatory Response	immune response of phagocytes	1.53E-09
Inflammatory Response	immune response of helper T lymphocytes	1.24E-07
Inflammatory Response	antiviral response	1.03E-18
Inflammatory Response	experimentally-induced arthritis	3.46E-18
Inflammatory Response	accumulation of leukocytes	6.75E-18
Inflammatory Response	accumulation of mononuclear leukocytes	2.62E-16
Inflammatory Response	accumulation of lymphocytes	3.34E-16
Inflammatory Response	accumulation of T lymphocytes	3.95E-13
Inflammatory Response	accumulation of granulocytes	6.74E-11
Inflammatory Response	accumulation of myeloid cells	1.08E-09
Inflammatory Response	accumulation of neutrophils	4.69E-09
Inflammatory Response	accumulation of phagocytes	1.08E-08
Inflammatory Response	recruitment of neutrophils	1.86E-17
Inflammatory Response	recruitment of phagocytes	3.42E-17

Inflammatory Response	recruitment of macrophages	3.97E-09
Inflammatory Response	recruitment of dendritic cells	5.46E-09
Inflammatory Response	recruitment of monocytes	1.74E-08
Inflammatory Response	migration of phagocytes	1.09E-15
Inflammatory Response	migration of dendritic cells	9.63E-14
Inflammatory Response	migration of macrophages	4.33E-10
Inflammatory Response	colitis	2.93E-14
Inflammatory Response	Encephalitis	6.10E-13
Inflammatory Response	Nephritis	2.18E-12
Inflammatory Response	chemotaxis of leukocytes	2.22E-12
Inflammatory Response	chemotaxis of phagocytes	1.08E-11
Inflammatory Response	chemotaxis of myeloid cells	4.26E-11
Inflammatory Response	chemotaxis of T lymphocytes	1.22E-10
Inflammatory Response	chemotaxis of mononuclear leukocytes	2.63E-10
Inflammatory Response	chemotaxis of lymphocytes	1.05E-09
Inflammatory Response	chemotaxis of neutrophils	1.40E-08
Inflammatory Response	chemotaxis of antigen presenting cells	2.14E-08
Inflammatory Response	chemotaxis of granulocytes	3.05E-08
Inflammatory Response	chemotaxis of dendritic cells	4.61E-08
Inflammatory Response	chemotaxis of monocytes	1.28E-07
Inflammatory Response	experimental autoimmune encephalomyelitis	2.82E-12
Inflammatory Response	innate immune response	3.22E-12
Inflammatory Response	infiltration by neutrophils	8.98E-12
Inflammatory Response	cellular infiltration of phagocytes	3.56E-09
Inflammatory Response	cellular infiltration by macrophages	7.56E-08
Inflammatory Response	glomerulonephritis	1.24E-11
Inflammatory Response	response of phagocytes	3.82E-11
Inflammatory Response	polyarthritis	4.31E-11
Inflammatory Response	chemoattraction of leukocytes	5.26E-11
Inflammatory Response	chemoattraction of phagocytes	6.49E-11
Inflammatory Response	chemoattraction of myeloid cells	7.25E-08
Inflammatory Response	chemoattraction of antigen presenting cells	5.67E-07
Inflammatory Response	mobilization of phagocytes	7.57E-11
Inflammatory Response	mobilization of neutrophils	7.84E-08
Inflammatory Response	cell-mediated response	3.05E-10
Inflammatory Response	cell-mediated response of T lymphocytes	2.24E-08
Inflammatory Response	influx of phagocytes	3.59E-10

Inflammatory Response	influx of neutrophils	1.13E-09
Inflammatory Response	antibacterial response	9.03E-09
Inflammatory Response	fever	9.89E-09
Inflammatory Response	insulinitis	3.08E-08
Inflammatory Response	antigen presentation	4.09E-08
Inflammatory Response	rejection	4.91E-08
Inflammatory Response	phagocytosis of leukocytes	6.83E-08
Inflammatory Response	phagocytosis of phagocytes	2.00E-07
Inflammatory Response	cellulitis	6.88E-08
Inflammatory Response	ulcerative colitis	7.37E-08
Inflammatory Response	Dermatitis	1.46E-07
Inflammatory Response	TH1 immune response	2.85E-07
Inflammatory Response	antibody response	4.17E-07
Inflammatory Response	adjuvant arthritis	5.32E-07
Cellular Movement	leukocyte migration	1.11E-28
Cellular Movement	T cell migration	6.12E-19
Cellular Movement	migration of antigen presenting cells	1.51E-16
Cellular Movement	Lymphocyte migration	1.72E-16
Cellular Movement	migration of mononuclear leukocytes	3.77E-16
Cellular Movement	migration of phagocytes	1.09E-15
Cellular Movement	migration of cells	2.30E-15
Cellular Movement	migration of dendritic cells	9.63E-14
Cellular Movement	migration of macrophages	4.33E-10
Cellular Movement	NK cell migration	4.30E-09
Cellular Movement	migration of myeloid cells	1.77E-08
Cellular Movement	migration of granulocytes	5.10E-07
Cellular Movement	recruitment of cells	1.84E-24
Cellular Movement	recruitment of leukocytes	1.37E-23
Cellular Movement	recruitment of myeloid cells	1.86E-18
Cellular Movement	recruitment of neutrophils	1.86E-17
Cellular Movement	recruitment of granulocytes	2.27E-17
Cellular Movement	recruitment of phagocytes	3.42E-17
Cellular Movement	recruitment of antigen presenting cells	3.67E-13
Cellular Movement	recruitment of mononuclear leukocytes	1.34E-10
Cellular Movement	recruitment of lymphocytes	1.61E-10
Cellular Movement	recruitment of T lymphocytes	1.90E-10
Cellular Movement	recruitment of macrophages	3.97E-09

Cellular Movement	recruitment of dendritic cells	5.46E-09
Cellular Movement	recruitment of inflammatory leukocytes	1.06E-08
Cellular Movement	recruitment of monocytes	1.74E-08
Cellular Movement	cell movement of leukocytes	2.79E-24
Cellular Movement	cell movement of myeloid cells	2.25E-20
Cellular Movement	cell movement of phagocytes	5.32E-20
Cellular Movement	cell movement of T lymphocytes	1.02E-19
Cellular Movement	cell movement of mononuclear leukocytes	6.01E-17
Cellular Movement	cell movement of lymphocytes	1.21E-16
Cellular Movement	cell movement	3.66E-16
Cellular Movement	cell movement of granulocytes	4.98E-16
Cellular Movement	cell movement of antigen presenting cells	1.64E-15
Cellular Movement	cell movement of dendritic cells	1.89E-13
Cellular Movement	cell movement of neutrophils	6.08E-13
Cellular Movement	cell movement of macrophages	2.69E-11
Cellular Movement	cell movement of hematopoietic progenitor cells	6.17E-10
Cellular Movement	cell movement of eosinophils	8.65E-10
Cellular Movement	cell movement of natural killer cells	1.81E-09
Cellular Movement	cell movement of monocytes	1.19E-08
Cellular Movement	cell movement of memory T lymphocytes	1.93E-08
Cellular Movement	cell movement of helper T lymphocytes	5.67E-08
Cellular Movement	cell movement of naive lymphocytes	2.79E-07
Cellular Movement	cell movement of bone marrow-derived macrophages	4.18E-07
Cellular Movement	cell movement of leukocyte cell lines	5.13E-07
Cellular Movement	cellular infiltration	7.66E-21
Cellular Movement	cellular infiltration by leukocytes	1.04E-20
Cellular Movement	cellular infiltration by mononuclear leukocytes	7.72E-17
Cellular Movement	cellular infiltration by lymphocytes	9.45E-15
Cellular Movement	cellular infiltration by granulocytes	1.62E-14
Cellular Movement	infiltration by T lymphocytes	2.92E-12
Cellular Movement	infiltration by neutrophils	8.98E-12
Cellular Movement	cellular infiltration of phagocytes	3.56E-09
Cellular Movement	cellular infiltration by macrophages	7.56E-08
Cellular Movement	mobilization of leukocytes	5.09E-16
Cellular Movement	mobilization of cells	5.83E-13
Cellular Movement	mobilization of myeloid cells	1.14E-12

Cellular Movement	mobilization of phagocytes	7.57E-11
Cellular Movement	mobilization of granulocytes	7.73E-08
Cellular Movement	mobilization of neutrophils	7.84E-08
Cellular Movement	attraction of leukocytes	5.43E-13
Cellular Movement	attraction of phagocytes	2.08E-12
Cellular Movement	attraction of lymphocytes	5.34E-11
Cellular Movement	attraction of mononuclear leukocytes	6.49E-11
Cellular Movement	attraction of cells	1.87E-10
Cellular Movement	attraction of T lymphocytes	5.40E-10
Cellular Movement	attraction of antigen presenting cells	7.60E-10
Cellular Movement	attraction of myeloid cells	3.08E-08
Cellular Movement	attraction of neutrophils	5.55E-08
Cellular Movement	chemotaxis of cells	9.75E-13
Cellular Movement	chemotaxis of leukocytes	2.22E-12
Cellular Movement	chemotaxis of phagocytes	1.08E-11
Cellular Movement	chemotaxis of myeloid cells	4.26E-11
Cellular Movement	chemotaxis of T lymphocytes	1.22E-10
Cellular Movement	chemotaxis of mononuclear leukocytes	2.63E-10
Cellular Movement	chemotaxis of lymphocytes	1.05E-09
Cellular Movement	chemotaxis of neutrophils	1.40E-08
Cellular Movement	chemotaxis of antigen presenting cells	2.14E-08
Cellular Movement	chemotaxis of granulocytes	3.05E-08
Cellular Movement	chemotaxis of dendritic cells	4.61E-08
Cellular Movement	chemotaxis of monocytes	1.28E-07
Cellular Movement	chemotaxis of lymphatic system cells	5.67E-07
Cellular Movement	trafficking of leukocytes	2.70E-12
Cellular Movement	trafficking of mononuclear leukocytes	7.50E-12
Cellular Movement	trafficking of lymphocytes	1.05E-10
Cellular Movement	trafficking of T lymphocytes	1.76E-10
Cellular Movement	homing of cells	4.07E-12
Cellular Movement	influx of leukocytes	3.98E-11
Cellular Movement	influx of cells	5.03E-11
Cellular Movement	influx of phagocytes	3.59E-10
Cellular Movement	influx of neutrophils	1.13E-09
Cellular Movement	chemoattraction of leukocytes	5.26E-11
Cellular Movement	chemoattraction of phagocytes	6.49E-11
Cellular Movement	chemoattraction	1.00E-08

Cellular Movement	chemoattraction of myeloid cells	7.25E-08
Cellular Movement	chemoattraction of antigen presenting cells	5.67E-07
Cellular Movement	transmigration of cells	1.36E-09
Cellular Movement	transmigration of leukocytes	1.96E-09
Immune Cell Trafficking	leukocyte migration	1.11E-28
Immune Cell Trafficking	T cell migration	6.12E-19
Immune Cell Trafficking	migration of antigen presenting cells	1.51E-16
Immune Cell Trafficking	Lymphocyte migration	1.72E-16
Immune Cell Trafficking	migration of mononuclear leukocytes	3.77E-16
Immune Cell Trafficking	migration of phagocytes	1.09E-15
Immune Cell Trafficking	migration of dendritic cells	9.63E-14
Immune Cell Trafficking	migration of macrophages	4.33E-10
Immune Cell Trafficking	NK cell migration	4.30E-09
Immune Cell Trafficking	migration of myeloid cells	1.77E-08
Immune Cell Trafficking	migration of granulocytes	5.10E-07
Immune Cell Trafficking	cell movement of leukocytes	2.79E-24
Immune Cell Trafficking	cell movement of myeloid cells	2.25E-20
Immune Cell Trafficking	cell movement of phagocytes	5.32E-20
Immune Cell Trafficking	cell movement of T lymphocytes	1.02E-19
Immune Cell Trafficking	cell movement of mononuclear leukocytes	6.01E-17
Immune Cell Trafficking	cell movement of lymphocytes	1.21E-16
Immune Cell Trafficking	cell movement of granulocytes	4.98E-16
Immune Cell Trafficking	cell movement of antigen presenting cells	1.64E-15
Immune Cell Trafficking	cell movement of dendritic cells	1.89E-13
Immune Cell Trafficking	cell movement of neutrophils	6.08E-13
Immune Cell Trafficking	cell movement of macrophages	2.69E-11
Immune Cell Trafficking	cell movement of eosinophils	8.65E-10
Immune Cell Trafficking	cell movement of natural killer cells	1.81E-09
Immune Cell Trafficking	cell movement of monocytes	1.19E-08
Immune Cell Trafficking	cell movement of memory T lymphocytes	1.93E-08
Immune Cell Trafficking	cell movement of helper T lymphocytes	5.67E-08
Immune Cell Trafficking	cell movement of naive lymphocytes	2.79E-07
Immune Cell Trafficking	cell movement of bone marrow-derived macrophages	4.18E-07
Immune Cell Trafficking	cell movement of leukocyte cell lines	5.13E-07
Immune Cell Trafficking	activation of leukocytes	7.62E-24
Immune Cell Trafficking	activation of phagocytes	2.27E-20

Immune Cell Trafficking	activation of myeloid cells	4.66E-20
Immune Cell Trafficking	activation of antigen presenting cells	7.84E-19
Immune Cell Trafficking	activation of macrophages	5.48E-18
Immune Cell Trafficking	activation of mononuclear leukocytes	5.00E-13
Immune Cell Trafficking	activation of lymphocytes	2.04E-12
Immune Cell Trafficking	activation of neutrophils	2.46E-12
Immune Cell Trafficking	activation of T lymphocytes	1.00E-11
Immune Cell Trafficking	activation of microglia	3.16E-10
Immune Cell Trafficking	recruitment of leukocytes	1.37E-23
Immune Cell Trafficking	recruitment of myeloid cells	1.86E-18
Immune Cell Trafficking	recruitment of neutrophils	1.86E-17
Immune Cell Trafficking	recruitment of granulocytes	2.27E-17
Immune Cell Trafficking	recruitment of phagocytes	3.42E-17
Immune Cell Trafficking	recruitment of antigen presenting cells	3.67E-13
Immune Cell Trafficking	recruitment of mononuclear leukocytes	1.34E-10
Immune Cell Trafficking	recruitment of lymphocytes	1.61E-10
Immune Cell Trafficking	recruitment of T lymphocytes	1.90E-10
Immune Cell Trafficking	recruitment of macrophages	3.97E-09
Immune Cell Trafficking	recruitment of dendritic cells	5.46E-09
Immune Cell Trafficking	recruitment of inflammatory leukocytes	1.06E-08
Immune Cell Trafficking	recruitment of monocytes	1.74E-08
Immune Cell Trafficking	cellular infiltration by leukocytes	1.04E-20
Immune Cell Trafficking	cellular infiltration by mononuclear leukocytes	7.72E-17
Immune Cell Trafficking	cellular infiltration by lymphocytes	9.45E-15
Immune Cell Trafficking	cellular infiltration by granulocytes	1.62E-14
Immune Cell Trafficking	infiltration by T lymphocytes	2.92E-12
Immune Cell Trafficking	infiltration by neutrophils	8.98E-12
Immune Cell Trafficking	cellular infiltration of phagocytes	3.56E-09
Immune Cell Trafficking	cellular infiltration by macrophages	7.56E-08
Immune Cell Trafficking	accumulation of leukocytes	6.75E-18
Immune Cell Trafficking	accumulation of mononuclear leukocytes	2.62E-16
Immune Cell Trafficking	accumulation of lymphocytes	3.34E-16
Immune Cell Trafficking	accumulation of T lymphocytes	3.95E-13
Immune Cell Trafficking	accumulation of granulocytes	6.74E-11
Immune Cell Trafficking	accumulation of myeloid cells	1.08E-09
Immune Cell Trafficking	accumulation of neutrophils	4.69E-09
Immune Cell Trafficking	accumulation of phagocytes	1.08E-08

Immune Cell Trafficking	mobilization of leukocytes	5.09E-16
Immune Cell Trafficking	mobilization of myeloid cells	1.14E-12
Immune Cell Trafficking	mobilization of phagocytes	7.57E-11
Immune Cell Trafficking	mobilization of granulocytes	7.73E-08
Immune Cell Trafficking	mobilization of neutrophils	7.84E-08
Immune Cell Trafficking	adhesion of immune cells	7.91E-15
Immune Cell Trafficking	attraction of leukocytes	5.43E-13
Immune Cell Trafficking	attraction of phagocytes	2.08E-12
Immune Cell Trafficking	attraction of lymphocytes	5.34E-11
Immune Cell Trafficking	attraction of mononuclear leukocytes	6.49E-11
Immune Cell Trafficking	attraction of T lymphocytes	5.40E-10
Immune Cell Trafficking	attraction of antigen presenting cells	7.60E-10
Immune Cell Trafficking	attraction of myeloid cells	3.08E-08
Immune Cell Trafficking	attraction of neutrophils	5.55E-08
Immune Cell Trafficking	chemotaxis of leukocytes	2.22E-12
Immune Cell Trafficking	chemotaxis of phagocytes	1.08E-11
Immune Cell Trafficking	chemotaxis of myeloid cells	4.26E-11
Immune Cell Trafficking	chemotaxis of T lymphocytes	1.22E-10
Immune Cell Trafficking	chemotaxis of mononuclear leukocytes	2.63E-10
Immune Cell Trafficking	chemotaxis of lymphocytes	1.05E-09
Immune Cell Trafficking	chemotaxis of neutrophils	1.40E-08
Immune Cell Trafficking	chemotaxis of antigen presenting cells	2.14E-08
Immune Cell Trafficking	chemotaxis of granulocytes	3.05E-08
Immune Cell Trafficking	chemotaxis of dendritic cells	4.61E-08
Immune Cell Trafficking	chemotaxis of monocytes	1.28E-07
Immune Cell Trafficking	chemotaxis of lymphatic system cells	5.67E-07
Immune Cell Trafficking	trafficking of leukocytes	2.70E-12
Immune Cell Trafficking	trafficking of mononuclear leukocytes	7.50E-12
Immune Cell Trafficking	trafficking of lymphocytes	1.05E-10
Immune Cell Trafficking	trafficking of T lymphocytes	1.76E-10
Immune Cell Trafficking	influx of leukocytes	3.98E-11
Immune Cell Trafficking	influx of phagocytes	3.59E-10
Immune Cell Trafficking	influx of neutrophils	1.13E-09
Immune Cell Trafficking	chemoattraction of leukocytes	5.26E-11
Immune Cell Trafficking	chemoattraction of phagocytes	6.49E-11
Immune Cell Trafficking	chemoattraction of myeloid cells	7.25E-08
Immune Cell Trafficking	chemoattraction of antigen presenting cells	5.67E-07

Immune Cell Trafficking	transmigration of leukocytes	1.96E-09
Cell-To-Cell Signaling and Interaction	recruitment of cells	1.84E-24
Cell-To-Cell Signaling and Interaction	recruitment of leukocytes	1.37E-23
Cell-To-Cell Signaling and Interaction	recruitment of myeloid cells	1.86E-18
Cell-To-Cell Signaling and Interaction	recruitment of neutrophils	1.86E-17
Cell-To-Cell Signaling and Interaction	recruitment of granulocytes	2.27E-17
Cell-To-Cell Signaling and Interaction	recruitment of phagocytes	3.42E-17
Cell-To-Cell Signaling and Interaction	recruitment of antigen presenting cells	3.67E-13
Cell-To-Cell Signaling and Interaction	recruitment of mononuclear leukocytes	1.34E-10
Cell-To-Cell Signaling and Interaction	recruitment of lymphocytes	1.61E-10
Cell-To-Cell Signaling and Interaction	recruitment of T lymphocytes	1.90E-10
Cell-To-Cell Signaling and Interaction	recruitment of macrophages	3.97E-09
Cell-To-Cell Signaling and Interaction	recruitment of dendritic cells	5.46E-09
Cell-To-Cell Signaling and Interaction	recruitment of inflammatory leukocytes	1.06E-08
Cell-To-Cell Signaling and Interaction	recruitment of monocytes	1.74E-08
Cell-To-Cell Signaling and Interaction	activation of leukocytes	7.62E-24
Cell-To-Cell Signaling and Interaction	activation of cells	5.43E-21
Cell-To-Cell Signaling and Interaction	activation of phagocytes	2.27E-20
Cell-To-Cell Signaling and Interaction	activation of myeloid cells	4.66E-20
Cell-To-Cell Signaling and Interaction	activation of antigen presenting cells	7.84E-19
Cell-To-Cell Signaling and Interaction	activation of macrophages	5.48E-18
Cell-To-Cell Signaling and Interaction	activation of mononuclear leukocytes	5.00E-13
Cell-To-Cell Signaling and Interaction	activation of lymphocytes	2.04E-12
Cell-To-Cell Signaling and Interaction	activation of neutrophils	2.46E-12
Cell-To-Cell Signaling and Interaction	activation of T lymphocytes	1.00E-11
Cell-To-Cell Signaling and Interaction	activation of neuroglia	2.32E-11
Cell-To-Cell Signaling and Interaction	activation of microglia	3.16E-10
Cell-To-Cell Signaling and Interaction	response of mononuclear leukocytes	4.02E-19
Cell-To-Cell Signaling and Interaction	response of lymphocytes	1.34E-17
Cell-To-Cell Signaling and Interaction	T cell response	4.85E-15
Cell-To-Cell Signaling and Interaction	response of phagocytes	3.82E-11
Cell-To-Cell Signaling and Interaction	response of helper T lymphocytes	5.26E-11
Cell-To-Cell Signaling and Interaction	response of myeloid cells	4.17E-10
Cell-To-Cell Signaling and Interaction	response of granulocytes	1.58E-07
Cell-To-Cell Signaling and Interaction	immune response of leukocytes	4.61E-19
Cell-To-Cell Signaling and Interaction	immune response of T lymphocytes	1.62E-14
Cell-To-Cell Signaling and Interaction	immune response of antigen presenting cells	3.09E-11

Cell-To-Cell Signaling and Interaction	immune response of phagocytes	1.53E-09
Cell-To-Cell Signaling and Interaction	immune response of helper T lymphocytes	1.24E-07
Cell-To-Cell Signaling and Interaction	adhesion of blood cells	5.34E-16
Cell-To-Cell Signaling and Interaction	adhesion of immune cells	7.91E-15
Cell-To-Cell Signaling and Interaction	adhesion of endothelial cells	4.39E-07
Cell-To-Cell Signaling and Interaction	stimulation of cells	1.53E-14
Cell-To-Cell Signaling and Interaction	stimulation of mononuclear leukocytes	5.67E-12
Cell-To-Cell Signaling and Interaction	stimulation of lymphocytes	8.39E-11
Cell-To-Cell Signaling and Interaction	stimulation of leukocytes	1.00E-10
Cell-To-Cell Signaling and Interaction	stimulation of T lymphocytes	3.60E-10
Cell-To-Cell Signaling and Interaction	attraction of leukocytes	5.43E-13
Cell-To-Cell Signaling and Interaction	attraction of phagocytes	2.08E-12
Cell-To-Cell Signaling and Interaction	attraction of lymphocytes	5.34E-11
Cell-To-Cell Signaling and Interaction	attraction of mononuclear leukocytes	6.49E-11
Cell-To-Cell Signaling and Interaction	attraction of cells	1.87E-10
Cell-To-Cell Signaling and Interaction	attraction of T lymphocytes	5.40E-10
Cell-To-Cell Signaling and Interaction	attraction of antigen presenting cells	7.60E-10
Cell-To-Cell Signaling and Interaction	attraction of myeloid cells	3.08E-08
Cell-To-Cell Signaling and Interaction	attraction of neutrophils	5.55E-08
Cell-To-Cell Signaling and Interaction	chemoattraction of leukocytes	5.26E-11
Cell-To-Cell Signaling and Interaction	chemoattraction of phagocytes	6.49E-11
Cell-To-Cell Signaling and Interaction	chemoattraction	1.00E-08
Cell-To-Cell Signaling and Interaction	chemoattraction of myeloid cells	7.25E-08
Cell-To-Cell Signaling and Interaction	chemoattraction of antigen presenting cells	5.67E-07
Cell-To-Cell Signaling and Interaction	binding of mononuclear leukocytes	1.46E-10
Cell-To-Cell Signaling and Interaction	binding of leukocytes	1.87E-10
Cell-To-Cell Signaling and Interaction	binding of blood cells	2.25E-10
Cell-To-Cell Signaling and Interaction	binding of cells	2.33E-09
Cell-To-Cell Signaling and Interaction	binding of lymphocytes	6.56E-08
Cell-To-Cell Signaling and Interaction	binding of vascular smooth muscle cells	2.61E-07
Cell-To-Cell Signaling and Interaction	cell-mediated response of T lymphocytes	2.24E-08
Cell-To-Cell Signaling and Interaction	interaction of cells	2.66E-08
Cell-To-Cell Signaling and Interaction	interaction of mononuclear leukocytes	1.58E-07
Cell-To-Cell Signaling and Interaction	induction of cells	4.09E-08
Cell-To-Cell Signaling and Interaction	induction of hematopoietic cell lines	6.88E-08
Cell-To-Cell Signaling and Interaction	phagocytosis of leukocytes	6.83E-08
Cell-To-Cell Signaling and Interaction	phagocytosis of phagocytes	2.00E-07

Cell-To-Cell Signaling and Interaction	respiratory burst of cells	2.36E-07
Cell-To-Cell Signaling and Interaction	suppression of blood cells	3.79E-07
Humoral Immune Response	production of antibody	3.83E-20
Humoral Immune Response	quantity of immunoglobulin	4.49E-20
Humoral Immune Response	quantity of IgG	5.57E-16
Humoral Immune Response	quantity of IgG1	3.23E-10
Humoral Immune Response	quantity of IgA	9.84E-10
Humoral Immune Response	quantity of IgG2a	3.13E-09
Humoral Immune Response	quantity of IgG3	1.92E-08
Humoral Immune Response	quantity of plasma cells	5.77E-08
Humoral Immune Response	morphology of germinal center	1.07E-07
Cell-mediated Immune Response	cell movement of T lymphocytes	1.02E-19
Cell-mediated Immune Response	cell movement of memory T lymphocytes	1.93E-08
Cell-mediated Immune Response	cell movement of helper T lymphocytes	5.67E-08
Cell-mediated Immune Response	T cell migration	6.12E-19
Cell-mediated Immune Response	NK cell migration	4.30E-09
Cell-mediated Immune Response	infiltration by T lymphocytes	2.92E-12
Cell-mediated Immune Response	T cell homeostasis	6.39E-11
Cell-mediated Immune Response	chemotaxis of T lymphocytes	1.22E-10
Cell-mediated Immune Response	T cell development	1.22E-10
Cell-mediated Immune Response	development of helper T lymphocytes	1.02E-07
Cell-mediated Immune Response	differentiation of helper T lymphocytes	1.74E-10
Cell-mediated Immune Response	differentiation of T lymphocytes	9.00E-09
Cell-mediated Immune Response	differentiation of Th2 cells	2.38E-07
Cell-mediated Immune Response	trafficking of T lymphocytes	1.76E-10
Cell-mediated Immune Response	recruitment of T lymphocytes	1.90E-10
Hematopoiesis	differentiation of leukocytes	7.66E-13
Hematopoiesis	differentiation of helper T lymphocytes	1.74E-10
Hematopoiesis	differentiation of mononuclear leukocytes	2.41E-10
Hematopoiesis	differentiation of lymphocytes	9.23E-10
Hematopoiesis	differentiation of T lymphocytes	9.00E-09
Hematopoiesis	differentiation of Th2 cells	2.38E-07
Hematopoiesis	development of blood cells	2.22E-12
Hematopoiesis	development of leukocytes	1.75E-11
Hematopoiesis	T cell development	1.22E-10
Hematopoiesis	development of lymphocytes	5.54E-10
Hematopoiesis	development of helper T lymphocytes	1.02E-07

Hematopoiesis	development of hematopoietic system	4.76E-07
Hematopoiesis	proliferation of hematopoietic cells	2.57E-07
Hematopoiesis	proliferation of hematopoietic progenitor cells	4.13E-07
Hematopoiesis	maturation of dendritic cells	3.56E-07
Hypersensitivity Response	cell movement of eosinophils	8.65E-10
Antigen Presentation	antigen presentation	4.09E-08
Cell Signaling	mobilization of Ca <sup>2+</sup>	1.64E-07
Cell Signaling	flux of Ca <sup>2+</sup>	3.61E-07
Cell Signaling	replication of viral replicon	3.79E-07
Cell Signaling	induction of nitric oxide	4.71E-07

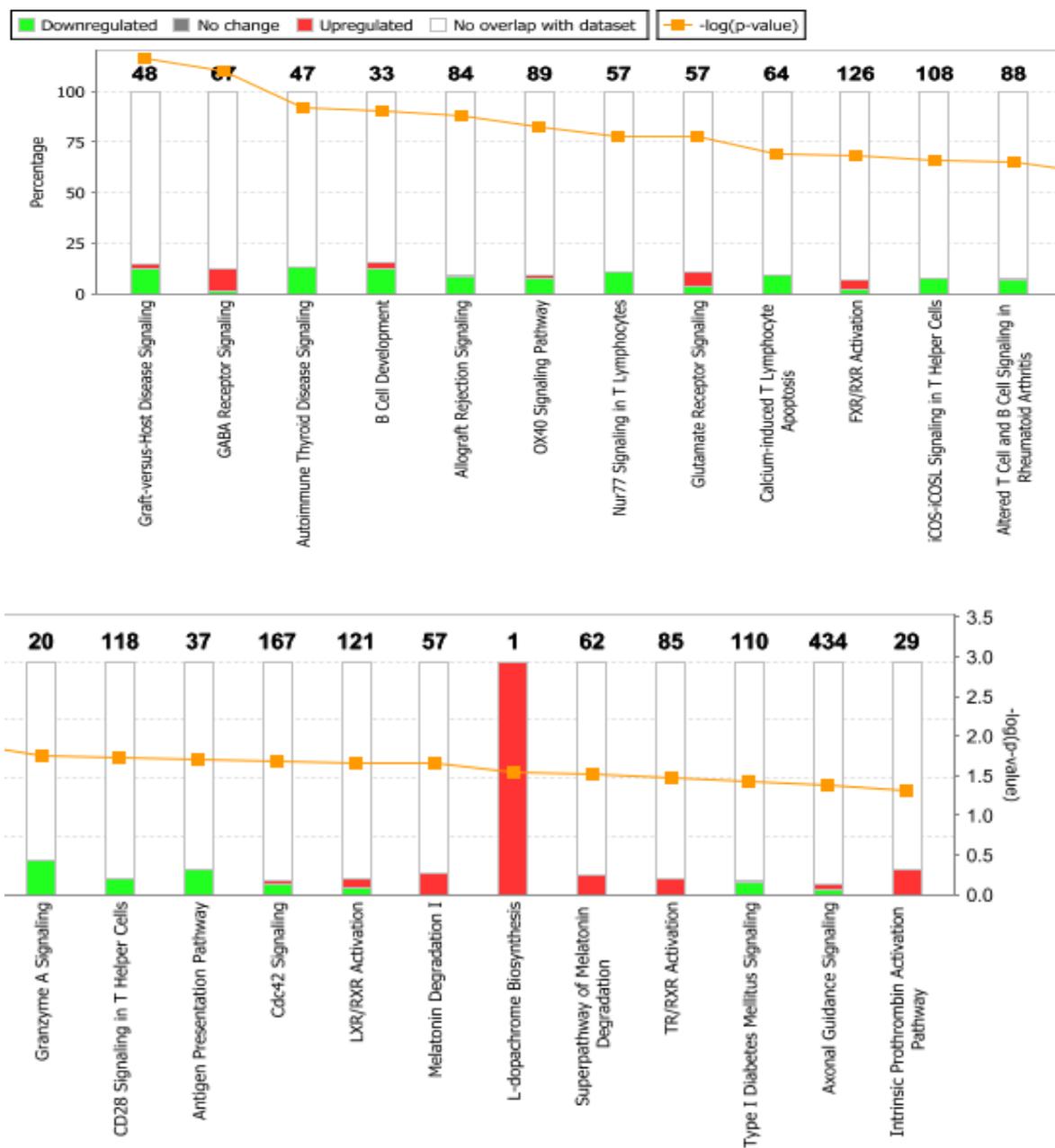
## Appendix 7: IPA functions

**Table A7: Functions associated with pathways enriched by SDEGs from the  $\Delta mtp$ -mutant infection.**

$\Delta mtp$ -mutant		
Category	Function Annotation	p-value
Cell-To-Cell Signaling and Interaction	signal transduction	3.20E-20
Cell-To-Cell Signaling and Interaction	communication of cells	1.55E-19
Cell-To-Cell Signaling and Interaction	communication	2.78E-19
Cell-To-Cell Signaling and Interaction	neurotransmission	1.24E-05
Cell-To-Cell Signaling and Interaction	synaptic transmission	1.51E-05
Cell-To-Cell Signaling and Interaction	GABA-mediated receptor currents	5.08E-04
Cell-To-Cell Signaling and Interaction	density of excitatory synapses	1.67E-03
Cell-To-Cell Signaling and Interaction	density of synapse	1.83E-02
Cell-To-Cell Signaling and Interaction	binding of lung cell lines	2.36E-03
Cell-To-Cell Signaling and Interaction	binding of blood platelets	3.07E-03
Cell-To-Cell Signaling and Interaction	developmental process of synapse	3.04E-03
Cell-To-Cell Signaling and Interaction	release of neurotransmitter	4.07E-03
Cell-To-Cell Signaling and Interaction	degeneration of synapse	7.56E-03
Cell-To-Cell Signaling and Interaction	reorganization of focal adhesions	7.56E-03
Cell-To-Cell Signaling and Interaction	detachment of tumor cell lines	8.61E-03
Cell-To-Cell Signaling and Interaction	assembly of intercellular junctions	1.13E-02
Cell-To-Cell Signaling and Interaction	adhesion of blood platelets	1.44E-02
Cell-To-Cell Signaling and Interaction	adhesion of phagocytes	1.72E-02
Cell-To-Cell Signaling and Interaction	propagation of action potential	2.00E-02
Cell-To-Cell Signaling and Interaction	action potential of cells	2.24E-02
Cell-To-Cell Signaling and Interaction	action potential of neurons	2.62E-02

Cell-To-Cell Signaling and Interaction	action potential of interneurons	2.83E-02
Cell-To-Cell Signaling and Interaction	AMPA mediated synaptic current	2.52E-02
Cell-To-Cell Signaling and Interaction	long-term potentiation of mossy fibers	2.52E-02
Cell-To-Cell Signaling and Interaction	phagocytosis of kidney cell lines	2.52E-02
Cell-To-Cell Signaling and Interaction	inhibitory postsynaptic potential	2.60E-02
Cell-To-Cell Signaling and Interaction	abnormal quantity of norepinephrine	2.62E-02
Cell-To-Cell Signaling and Interaction	secretion of neurotransmitter	2.76E-02
Inflammatory Response	binding of blood platelets	3.07E-03
Inflammatory Response	ulcerative dermatitis	4.62E-03
Inflammatory Response	adhesion of blood platelets	1.44E-02
Inflammatory Response	adhesion of phagocytes	1.72E-02
Inflammatory Response	inflammation of joint	1.53E-02
Inflammatory Response	function of mast cells	1.68E-02
Inflammatory Response	phagocytosis of kidney cell lines	2.52E-02
Cellular Movement	innervation of outer hair cells	3.14E-03
Cellular Movement	homing of Th17 cells	4.62E-03
Cellular Movement	distribution of neurons	1.11E-02
Cell-mediated Immune Response	homing of Th17 cells	4.62E-03
Cell-mediated Immune Response	proliferation of thymocytes	1.47E-02
Cell-mediated Immune Response	formation of thymocytes	2.09E-02
Immune Cell Trafficking	homing of Th17 cells	4.62E-03
Immune Cell Trafficking	adhesion of phagocytes	1.72E-02
Cell Signaling	metabolism of cyclic AMP	8.61E-03
Cell Signaling	catabolism of cyclic AMP	1.14E-02
Cell Signaling	inhibition of cyclic AMP	1.53E-02
Cell Signaling	cell surface receptor linked signal transduction	2.58E-02
Antimicrobial Response	clearance of adenoviridae	1.53E-02
Hypersensitivity Response	function of mast cells	1.68E-02
Humoral Immune Response	lack of B lymphocytes	2.60E-02

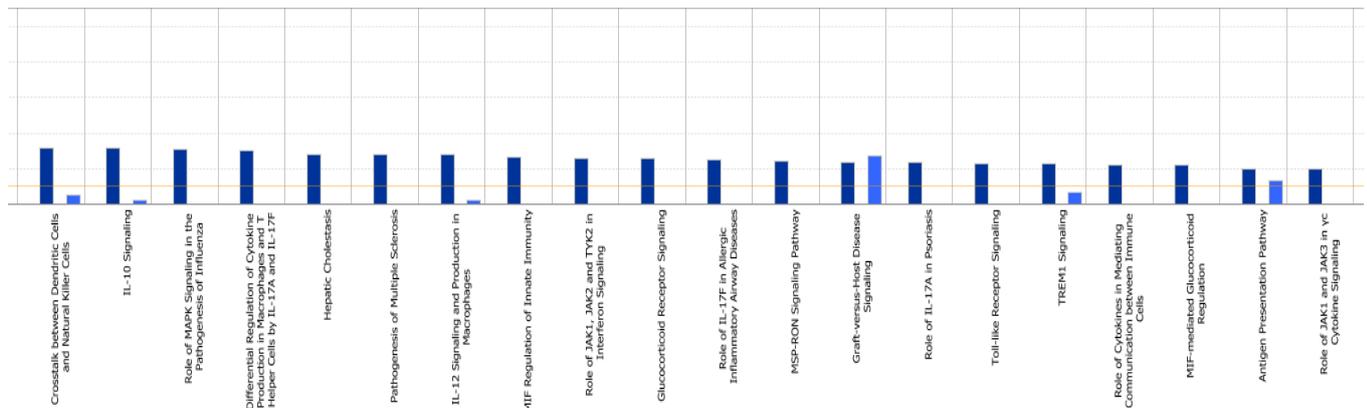
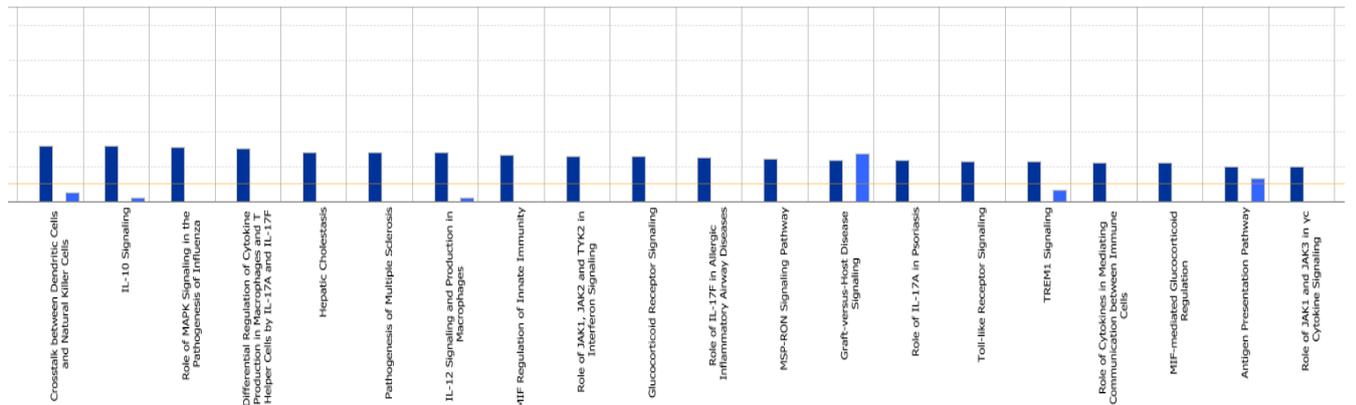
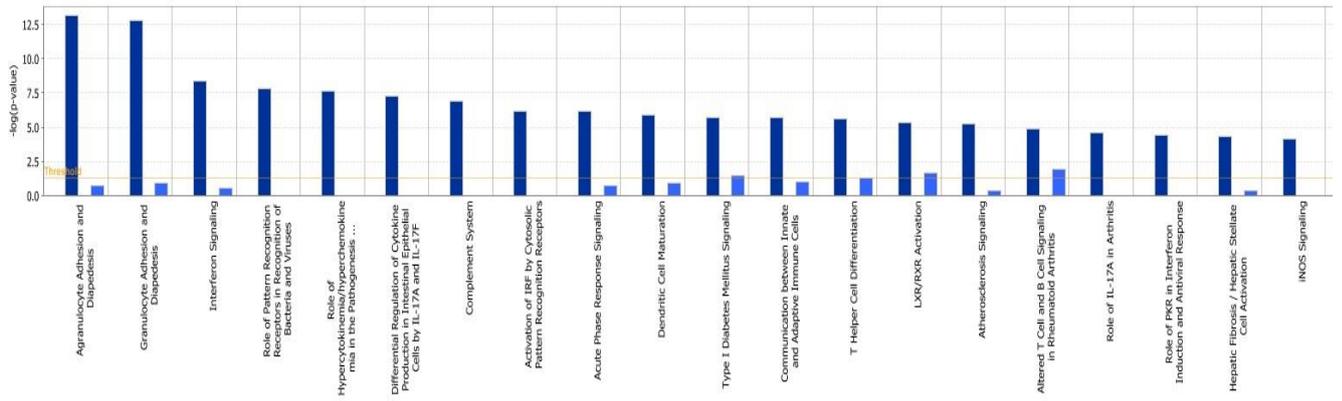
## Appendix 8: IPA bar graphs

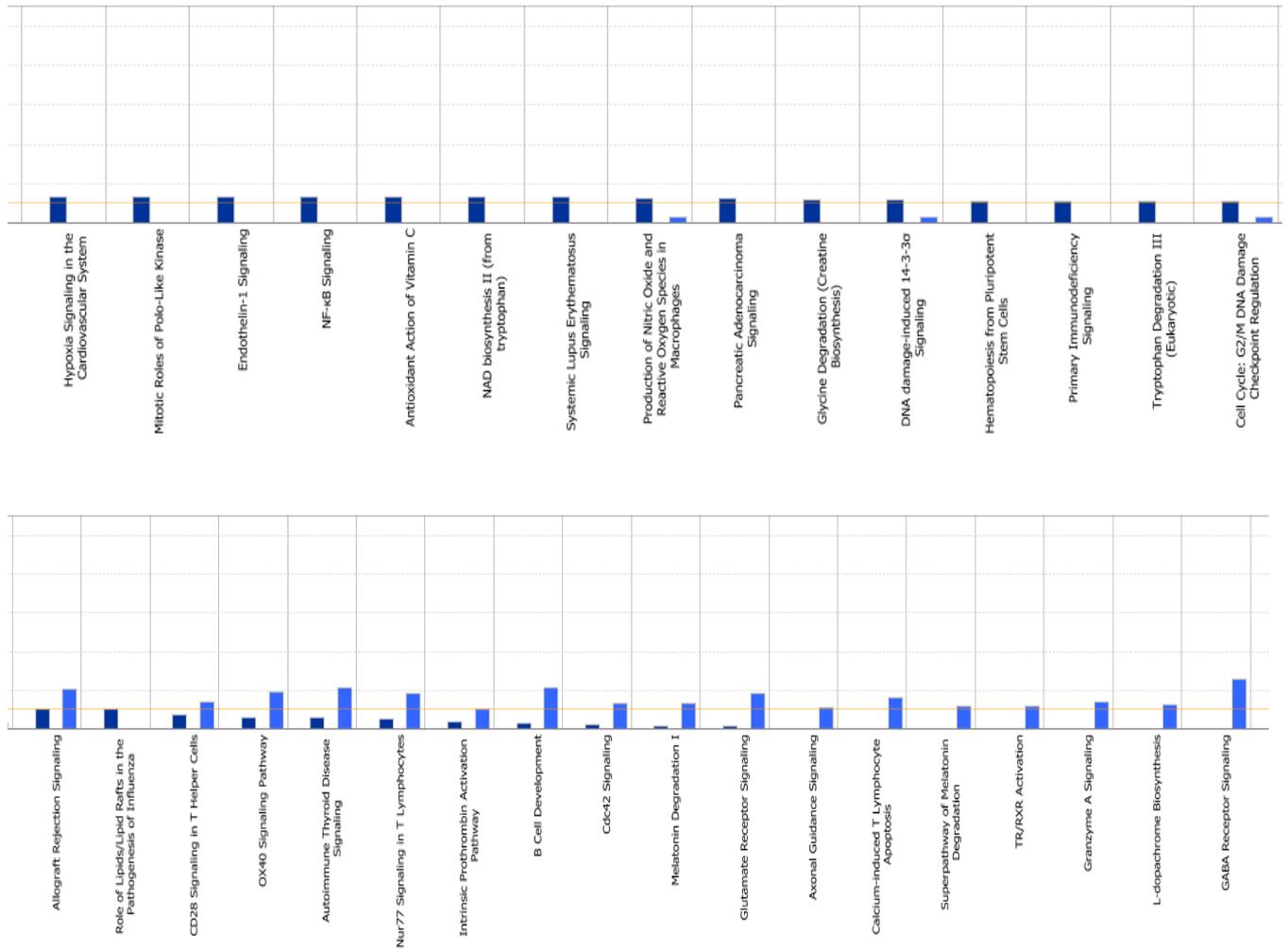


**Figure A8: Canonical pathways enriched from the unique  $\Delta mtp$ -mutant infection genes as shown in IPA.** Red represents the genes that were up-regulated and green represents genes that were down regulated in the pathway.

## Appendix 9: IPA bar graphs

■ UlvsWT\_WT - 2016-03-31 01:20 AM ■ Ulvsmtp\_mtp (Unique 930 genes) - 2016-03-31 01:34 AM





**Figure A9.** Comparison analysis of the canonical pathways enriched from the unique genes as shown in IPA. Dark blue represents WT and light blue represents  $\Delta mtp$ -mutant ( $p$ -value < 0.05).

#### Appendix 10: Selected genes associated with host-pathogen interactions

**Table A10:** SDEGs associated with host-pathogen interactions.

Gene	UI vs WT			UI vs Mtp		
	log2(fold_change)	q_value	significant	log2(fold_change)	q_value	significant
<i>Tlr1</i>	1,05213	0,015531	Yes	-0,07803	0,925691	No
<i>Tlr2</i>	1,01732	0,020848	Yes	-0,37415	0,539432	No
<i>C3ar1</i>	1,51365	0,001934	Yes	0,690664	0,2375	No
<b><i>Itga2b</i></b>	<b>-1,06062</b>	<b>0,029518</b>	<b>Yes</b>	<b>-3,0907</b>	<b>0,003917</b>	<b>Yes</b>
<i>Itga8</i>	-1,32416	0,00628	Yes	-0,54906	0,435652	No
<i>Itgam</i>	1,14451	0,01183	Yes	-0,89914	0,105395	No
<i>C4b</i>	1,58554	0,001934	Yes	0,06405	0,925664	No
<b><i>Igf1</i></b>	<b>1,3504</b>	<b>0,003475</b>	<b>Yes</b>	<b>1,18514</b>	<b>0,030456</b>	<b>Yes</b>
<b><i>Igfbp2</i></b>	<b>-1,4404</b>	<b>0,001934</b>	<b>Yes</b>	<b>-1,42984</b>	<b>0,013355</b>	<b>Yes</b>

<i>Igfbp3</i>	-1,69101	0,001934	Yes	-0,64387	0,265486	No
<i>Igfn1</i>	-1,2349	0,043156	Yes	-0,0225	1	No
<i>Igip</i>	1,07011	0,030972	Yes	0,735156	0,325532	No
<i>Igtp</i>	2,39848	0,001934	Yes	0,23047	0,735453	No
<i>Ciita</i>	1,40541	0,001934	Yes	-0,39929	0,50068	No
<i>Tap1</i>	1,78121	0,001934	Yes	-0,40511	0,47241	No
<i>Tap2</i>	1,42327	0,001934	Yes	-0,63186	0,238256	No
<i>Fcgr1</i>	1,39909	0,001934	Yes	-0,58635	0,441838	No
<i>Fcgr2b</i>	0,907643	0,033775	Yes	-1,03643	0,07617	No
<i>Fcgr3</i>	0,900109	0,031517	Yes	-0,89503	0,107038	No
<i>Fcgr4</i>	2,73725	0,001934	Yes	-0,38016	0,705959	No
<b><i>Fcgrt</i></b>	<b>-0,97318</b>	<b>0,022433</b>	<b>Yes</b>	<b>-1,18799</b>	<b>0,039061</b>	<b>Yes</b>
<i>Clec12a</i>	1,2766	0,001934	Yes	-0,27309	0,70591	No
<i>Clec14a</i>	-1,42536	0,007545	Yes	-0,66403	0,319351	No
<i>Clec4d</i>	1,10329	0,044756	Yes	-1,02023	0,267896	No
<i>Clec4e</i>	2,23019	0,001934	Yes	0,04585	0,954277	No
<i>Clec4n</i>	1,62217	0,001934	Yes	-0,88743	0,210745	No
<i>Clec5a</i>	1,32626	0,035129	Yes	0,642922	0,509838	No
<i>Clec7a</i>	1,14075	0,008653	Yes	-0,51007	0,384983	No
<i>Fpr2</i>	1,06106	0,017405	Yes	-0,3086	0,646696	No
<i>Cd14</i>	1,53342	0,001934	Yes	-0,75403	0,234637	No
<i>Cd163</i>	1,03602	0,037021	Yes	0,862111	0,202033	No
<i>Cd163l1</i>	1,37858	0,042631	Yes	-0,48037	0,632877	No
<i>Cd1d1</i>	1,75068	0,001934	Yes	1,17743	0,160803	No
<i>Cd209a</i>	-1,2476	0,031517	Yes	-0,50808	0,526761	No
<i>Cd274</i>	2,9262	0,001934	Yes	0,812163	0,136719	No
<i>Cd28</i>	1,06285	0,007545	Yes	0,292245	0,665075	No
<i>Cd300lg</i>	-1,59795	0,001934	Yes	-1,16734	0,051337	No
<i>Cd3e</i>	1,38095	0,00628	Yes	0,214637	0,822679	No
<i>Cd3g</i>	1,50286	0,001934	Yes	0,158937	0,821701	No
<i>Cd4</i>	1,17937	0,014525	Yes	-0,137	0,87704	No
<i>Cd40</i>	1,2271	0,003475	Yes	-0,39882	0,570013	No
<i>Cd40lg</i>	2,14313	0,009667	Yes	1,52679	0,156823	No
<i>Cd53</i>	1,43728	0,001934	Yes	0,100706	0,883455	No
<i>Cd63</i>	1,29953	0,001934	Yes	0,196227	0,780553	No
<i>Cd69</i>	1,37153	0,001934	Yes	0,071733	0,939162	No
<i>Cd72</i>	1,66875	0,033775	Yes	-0,90293	0,486789	No
<i>Cd80</i>	1,39102	0,048456	Yes	0,174791	0,88681	No
<i>Cd83</i>	0,900068	0,037021	Yes	-0,73941	0,230279	No
<i>Cd86</i>	1,7798	0,001934	Yes	0,661663	0,272344	No
<b><i>Cd9</i></b>	<b>-1,10082</b>	<b>0,012779</b>	<b>Yes</b>	<b>-1,49473</b>	<b>0,00667</b>	<b>Yes</b>
<i>Nlrc5</i>	1,83406	0,001934	Yes	-0,12355	0,852611	No