The University Of Sheffield.

Analysis of epigenetic changes arising during challenge with *Streptococcus pneumoniae*

By:

Joby Cole

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculties of Medicine & Engineering,
Departments of Infection, Immunity, and Cardiovascular disease;
&
Chemical and Biological Engineering;

March 2018

Acknowledgements

During the course of this PhD I have been fortunate to receive the support and advice of several people.

Firstly, I would like to thank both of my supervisors Professors Mark Dickman and David Dockrell for all of their advice and encouragement throughout the process. During the fellowship application process in particular, David was incredibly generous with his time throughout the iterative course of the grant writing and preparation for the interview.

I would like to acknowledge the help and support I have received from the members of both lab groups and in particular thank Martin Bewley, Hellen Marriot, Paul Collini, Paul Morris and Lucy Morris for their help with the cell culture and infective challenges. I would especially like to thank both Joseph Longworth and Thomas Minshull for the many late hours spent together coaxing the mass spectrometers through yet another batch of samples, and Eleanor Hanson for supply of CHO cells.

I would like to thank my collaborators Dr Mélanie Hamon at The Pasteur Institute for her advice on histone post-translational modifications and Prof Richard Emes and Dr Tom Giles at the University of Nottingham for their help with the bioinformatics analysis.

I would also like to thank the Wellcome Trust for funding my fellowship without whom none of this would have been possible.

I would also like to thank Prof Robert Read for the encouragement to apply for an academic clinical fellowship and the introduction to Prof Dickman which led to this project.

Finally, I would like to thank my parents without whose care and concern I would not find myself in this position, and my parents in-law George and Sheila Kinghorn for their help and support.

Finally, I would like to thank Joanne for being so understanding of all of those late nights, evenings and weekends spent in the lab; and also Nia and Léo whose constructive criticism and feedback at all hours made the whole experience altogether more interesting.

Table of Contents:

List of Tables:

List of Figures:

Abbreviations:

ΔPLY
Isogenic mutant D39-ΔPLY,
1D
One dimensional unfractionated sample,
2D
two dimensional liquid chromatography method,
AA
Amino Acid,
ACN
Acetonitrile,
AGC
Automatic Gain Control,
AMPK
AMP-activated protein kinase,
ANOVA
Analysis of variance,
Aof1
amine oxidase flavin containing 1,
ARHGEF2
Rho guanine nucleotide exchange factor 2,
ASH1
absent, small, or homeotic-1,
AUC
Area Under the Curve,
BCG
Bacillus Calmette-Guérin,
bp
base pair,
BRG1
Brahma-related gene 1,
BSA
Bovine serum albumin,
BWA
Burrows-Wheeler Alignment,
C1QBP
Complement component 1 Q subcomponent-binding protein, mitochondrial,
cagPAI
cytotoxin-assoicated gene A pathogenicity island,
CBA
Colombia blood agar,
CCDC22
coiled-coil domain containing protein 22,
CCL3L1
C-C motif chemokine ligand 3 like 1,
CCL4L1
C-C motif chemokine ligand 4 like 1,
CCR5
chemokine receptor 5,
cfu
Colony forming units,
ChIP
Chromatin Immunoprecipitation,
ChIP-Seq
Chromatin Immunoprecipitation and next generation sequencing,
CHO
Chinese Hamster Ovary,
CID
Collision induced dissociation,
CO_2
Carbon Dioxide,
CT
Cycle threshold,
CV
Coefficient of variation,
CXCL10
C-X-C motif chemokine ligand 10,
CXCL2
chemokine C-X-C motif ligand 2,
CYTOR
Cytoskeleton regulator RNA,
Da
Daltons,
dapi
4',6-diamidino-2-phenylindole,
DDA
Data dependent acquisition,
DDA120
Data dependent Acquisition with MS2 resolution of 120 000 at 200 m/z,
DDA30
Data dependent Acquisition with MS2 resolution of 30 000 at 200 m/z,
DDA60
Data dependent Acquisition with MS2 resolution of 60 000 at 200 m/z,
DDX46
Deadbox 46,
DEFB125
Defensin beta 125,
DGLUCY
D-Glutamate Cyclase 3,
DIA
Data independent acquisition,
DIA30

Data Independent Acquisition, with MS1 resolution of 30 000 at 200m/z and MS2 resolution of 15 000 at 200m/z,

DIA60
Data Independent Acquisition, with MS1 resolution of 60 000 at 200m/z and MS2 resolution of 30 000 at 200m/z,

DIAvw
Data Independent Acquisition variable window,

DNA
Deoxyribonucleic acid,

DNMT
DNA methyltransferase,

DTT
dithiothreitol, 46

DUSP4
Dual specificity phosphatase 4,

DUX4
Double homeobox 4,

EDTA
ethylenediaminetetraacetic acid,

EGR3
Early growth response 3,

EGTA
ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid,

ENCODE
Encyclopaedia of DNA elements,

ESI
Electrospray ionisation,

FASP
Filter aided separation of proteins,

FCS
Fetal calf serum,

FDR
False Discovery Rate,

FDR
False discovery rate,

FITC
Fluorescein isothiocyanate,

FSC
Forward scatter,

GADD45B
Growth arrest and DNA damage inducible beta,

GCRMA
robust multi-array average expression with help of probe sequence,

GJA3
Gap junction protein alpha 3,

GO
Gene Ontology,

GRIN1
glutamate ionotropic receptor N-methyl-D-aspartate type subunit 1,

H_2O
Dihydrogen monoxide,

H_2SO_4
Sulphuric acid,

H3.3K27me2K36me1
Histone H3.3 lysine 27 dimethylation lysine 36 monomethylation,

H3.3K27me2K36me2
Histone H3.3 Lysine 27 dimethylation lysine 36 dimethylation,

H3.3K36me2
Histone H3.3 lysine 36 dimethylation,

H3K14ac
Histone H3 lysine 14 acetylation,

H3K14me3
Histone H3 lysine 14 trimethylation,

H3K27me2K36me2
Histone H3 lysine 27 dimethylation lysine36 dimethylation,

H3K27me3
Histone H3 lysine 27 trimethylation,

H3K27me3K36me1
Histone H3 lysine 27 trimethylation lysine 36 monomethylation,

H3K4me3
Histone 3 Lysine 4 Trimethylation,

H3K9me2
Histone H3 lysine 9 dimethylation,

H3K9me2K14ac
Histone H3 lysine 9 dimethylation Lysine 14 acetylation),

H3K9me3
Histone H3 lysine 9 trimethylation,

H3K9me3K14ac
Histone H3 lysine 9 trimethylation, lysine 14 acetylation,

H3K9MeS10PhosK14Ac
Histone 3 lysine 9 methylation serine 10 phosphorylation and lysine 14 acetylation,

H3S10
Hinstone H3 serine 10,

H3S10phos
Histone H3 Serire 10 phosphorylation,

H4K16ac
Histone 4 lysine 16 acetylation,

H4K20me1
Histon H4 lysine 20 monomethyl,

H4K20me2

Histone H4 lysine 20 dimethylation,

HDAC
Histone deacetylase,

HDACi
Histone Deacetylase Inhibitor,

HDHD5
Haloacid dehalogenase like hydrolase domain containing 5,

HFBA
heptafluorobutyric acid,

HGH1
Human growth hormone 1 homolog,

HPLC
high performance liquid chromatography,

hsp
Heat shock protein,

IAA
Iodoacetic acid,

IFN-γ
Interferon gamma,

IGV
Integrated genomics viewer,

IL-1β
interleukin 1 beta,

IL-18
Interleukin 18,

IL-6
Interleukin 6,

InlB
internalin B, 34

IPD
Invasive Pneumococcal Disease,

K
Lysine,

KAT6A
Lysine acetyltransferase 6A,

KCl
potassium chloride,

Kdm1b
lysine demethylase 1B,

Kdm3a
lysine demethylase 3a,

Kdm3b
lysine demethylase 3b,

KEGG
Kyoto Encyclopaedia of Genes and Genomes,

KEGG
Kyoto encyclopaedia of genes and genomes,

KMT
Lysine methyltransferase,

LC-MS/MS
Liquid chromatography and tandem mass spectrometry,

LFQ
Label Free Quantification,

LINC01667
Long non-protein coding RNA 1667,

Lowess
locally weighted scatterplot smoothing,

LPS
Lipopolysaccharide,

LSD1
Lysine specific demethylase 1,

LSD2
lysine-specific histone demethylase 2,

m/z
Mass to charge ratio,

MACS
Model-based Analysis of ChIP-Seq,

MAPK
Mitogen activated protein kinase,

MDM
Monocyte Derived Macrophages,

MgCl
Magnesium Chloride,

mgf
Mascot Generic Format,

MI
Mock infected,

MIP-1β
Macrophage inflammatory protein 1 β,

MIP-1α/LD78β
macrophage inflammatory protein,

miR155
micro RNA 155 host gene,

mRNA
Messenger ribonucleic acid,

MS
Mass spectrometry,

MS/MS
Tandem mass spectrometry,

MS2
Tandem mass spectrometry spectra,

MUC3A
mucin 3A,

NaCl
Sodium Chloride,

NCE
Normalised Collision Energy,

NGS
Next Generation Sequencing,

NLRP3
nod like receptor pyrin domain containing 3,

NLS
N-Laurosyl-sarcosine,
NO
Nitrogen oxide,
NOD
Nucleotide binding oligomerization domain-containing protein,
NP40
Nonidet P-40,
NPAS3
Neuronal Per-Arnt-Sim domain 3,
NQO1
NAD(P)H dehydrogenase quinone 1,
NR4A1
Nuclear receptor subfamily 4 group A member 1,
NR4A2
Nuclear receptor subfamily 4 group A member 2,
NRF2
Nuclear erythroid factor 2,
NSC
Normalised strand coeficient,
OD
Optical Density,
PBS
Phosphate buffered saline,
PCR2
Ploycomb repressive complex 2,
PGC
Porous Graphitic Column,
PHF8
PHD finger protein 8,
PI
Protease Inhibitor,
PLY
Pneumolysin,
ppm
Parts per million,
PRM
Parallel reaction monitoring,
PSM
Peptide sequence match,
PTM
Post-translational modification,
PTX3
Pentraxin 3,
PVDF
Polyvinylidene fluoride membrane,
QE
Orbitrap Q-Exactive HF,
QE
QExactive HF Orbitrap,
qPCR
quantitative polymerase chain reaction,
Rag1
recombination activating gene 1,
rcf
Relative Centrifugal Force,
RIPK2
Receptor-interacting Serine/threonine-protein kinase 2,
RLE
Relative log expression,
RNA45SN1
45S Pre-Ribosomal N1,
RNASeq
RNA Sequencing,
ROCK1P1
Rho associated coiled-coil containing protein kinase 1 pseudogene 1,
ROS
Reactive oxygen species,
RSC
Relative strand correlation,
S. pneumoniae
Streptococcus pneumoniae,
SCC
Side scatter,
SCID
Severe combined immunodeficiency,
SD
Standard deviation,
SDS
Sodium dodecyl sulphate,
SEM
standard error of the mean,
Sirt1
Sirtuin 1,
Sirt2
Sirtuin2,
SLIT2
Slit homolog 2 protein,
SMARCA
SW/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 2,
SRM
Selective reaction monitoring,
STAB1
stabilin 1,
Suv39h
SET domain of the human homolog of *Drosophila* Su(var)3-9,
SWI/SNF
SWitch/Sucrose non-fermentable,
TBS/tween

Solution of Tris-HCL, sodium Chloride, and tween-20,

TCA
Trichloroacetic acid,

TCEP
Tris(2-carboxyethyl)phosphine hydrochloride,

TEAB
Triethylammonium bicarbonate,

TEMED
Tetramethylethylenediamine,

TFA
Trifluoroacetic acid,

TLR
Toll-like receptor,

TLR4
Toll like receptor 4,

TNF
Tumour Necrosis factor,

TNF-α
Tumour necrosis factor alpha,

ToF
Time of flight,

TRIM24
tripartite motif containing 24,

TRXR1
Thioredoxin reductase,

TSS
Transcription start site,

TTTY23
testis specific transcript Y linked 23,

UBTD1
Ubiquitin domain-containing protein1,

UCK1
Uridine-cytidine kinase 1,

UV
ultraviolet,

VPS13C
vacuolar protein sorting 13C,

XIC
Extracted ion Chromatogram,

Thesis abstract:

Streptococcus pneumoniae is the leading cause of community acquired pneumonia. The pathogenesis of pneumococcal disease is not fully understood. Nasopharyngeal colonisation precedes pneumococcal disease and is influenced by the effectiveness of innate immune responses. As the macrophage is a key component in the innate immune response, the interaction between *S. pneumoniae* and the macrophage will be crucial in controlling susceptibility to pneumococcal disease. It is increasingly recognised that epigenetic mechanisms play key roles in the host pathogen interaction. Mass spectrometry has emerged as a powerful tool for the study of global changes in histone post-translational modifications (PTM) one of the key epigenetic mechanisms.

I hypothesised that *S. pneumoniae* induces immediate changes in the host's epigenome by altering histone post-translational modifications, in turn regulating the innate immune responses by modulating gene expression and may explain some of the variation observed in the susceptibility to invasive pneumococcal disease.

I established a proteomic approach to determine the extent and nature of PTMs in primary human macrophages following challenge with *S. pneumoniae* and the contribution of pneumolysin, a key virulence factor, to changes in histone PTMs following bacterial challenge of MDMs. This illustrated that the relative abundance of several histone PTMs is modified in response to challenge with *S. pneumoniae.*

I demonstrated that the bacterial virulence factor pneumolysin causes significant changes in both gene expression (503 differentially expressed genes) and protein expression in primary MDMs.

Finally, I used RNA-Seq and ChIP-Seq, to attempt to determine the relationship between gene expression levels and PTMs of the associated histones. These analyses demonstrated a number of innate immune response genes to be differentially enriched, demonstrating the role of epigenetic regulation of gene expression associated with the innate immune response upon challenge with *S. pneumoniae.*

Chapter 1: Introduction

1.1 *Streptococcus pneumoniae*:

1.1.1 The public health burden of *Streptococcus pneumoniae*:

Streptococcus pneumoniae (*S. pneumoniae*) is a Gram-positive bacterium that remains a major cause of invasive disease in humans despite the availability of polysaccharide and protein conjugate vaccines (Ladhani et al., 2013). It causes approximately 5000 cases of invasive disease a year in England and Wales (" GOV.UK,"2018.). Invasive pneumococcal disease (IPD) is defined as the isolation of *S. pneumoniae* from otherwise sterile sites such as blood or cerebral spinal fluid. *S. pneumoniae* is the leading cause of community-acquired pneumonia (Holter et al., 2015). Invasive pneumococcal disease accounts for at least 2 million deaths globally (O'Brien et al., 2009) and represents therefore a major burden on health services. The presence of *S. pneumoniae* in the nasopharynx, without causing disease, is termed colonisation. This has been shown to precede invasive disease (Kadioglu et al., 2008a; Ramos-Sevillano et al., 2011). Although there is a role for the adaptive immune response, as vaccination with pneumococcal conjugate vaccine has been shown to reduce IPD through antibody and T cell responses (Bryant et al., 2010; Ladhani et al., 2013), the likelihood of severe disease or of IPD is thought to be governed by the interaction between the innate immune response and *S. pneumoniae*. It is not fully understood how *S. pneumoniae* evades the host's innate immune defences, nor how aberrant host responses contribute to disease severity. Microorganisms have evolved a variety of mechanisms to evade host defences. These include inhibition of phagocytosis through the action of polysaccharide capsule (Hyams et al., 2010), subversion of bacterial killing, inhibition of inflammation, modulation of cell trafficking and survival (Baxt et al., 2013; Fernie-King et al., 2002). Several of these mechanisms rely on the modulation of host gene expression (Jenner and Young, 2005; Nau et al., 2002).

1.1.2 The role of pneumolysin as a virulence factor:

S. pneumoniae has a number of different virulence factors (Kadioglu et al., 2008a). Among these one of the most important is pneumolysin (PLY), a cholesterol-dependent cytolysin (Marriott et al., 2008). It is produced by almost all invasive clinical isolates of *S. pneumoniae* and has been implicated in migration from lung tissue to blood (Hu et al., 2015), in host immune subversion and as a trigger of inflammatory responses. In human cells PLY is a potent activator of gene transcription and potentially interacts with several pattern recognition receptors to mediate innate immune responses. These include toll-like receptor 4 (TLR4) and nod-like receptor pyrin domain containing 3 (NLRP3) and potentially other NLRs, which help coordinate the host response to pneumococcal infection (McNeela et al., 2010; Srivastava et al., 2005; Witzenrath et al., 2011). Although the exact role played by TLR4 in pneumolysin remains controversial. *S. pneumoniae* mutants lacking PLY in murine models display attenuated virulence but cause chronic bacteraemia rather than overwhelming sepsis and death (Benton et al., 1995). This underlines the importance of PLY in activating the host response, which leads to bacterial clearance. It therefore appears that although PLY is required to activate a successful innate immune response, PLY achieves this at the expense of generating a vigorous inflammatory response, which in some individuals is excessive and contributes to disease pathogenesis.

Furthermore, PLY is responsible for the differential expression of 142 genes in an undifferentiated THP-1 cell model of infection. Importantly a number of these genes are involved in host immune responses, thereby reinforcing the role of PLY as a key virulence factor (Rogers et al., 2003).

1.2 The innate immune system:

Invasive disease has been shown to be preceded by colonisation (Bogaert et al., 2004; Gray et al., 1980). The host's innate immune response contributes to the clearance of colonising bacteria and those that reach the lower airway, thereby playing a key role in determining whether *S. pneumoniae* colonisation progresses to pneumonia or IPD. The mammalian immune system has traditionally been described as comprising two different types of immune responses. First, the innate immune response, a general response to pathogens which was initially thought to occur with identical intensity on repeated exposure and is composed chiefly of phagocytic cells, cells that release cytokines and chemokines and natural killer cells. The second, the adaptive immune response, is a more specialised response that develops in response to each new pathogen it encounters, gives rise to a stronger response on repeated exposure and is composed of proliferating B and T Lymphocytes allowing it to develop a memory of antecedent events with great specificity. The innate immune system is responsible for detection and recognition of pathogens, phagocytosis of cellular debris and pathogens, cytokine release and regulation of the inflammatory response. These different processes are achieved by specialised cells involved primarily in phagocytic clearance (neutrophils, macrophages and monocytes), or cells not primarily involved in phagocytic clearance but that can activate host responses following exposure to pathogens (dendritic cells, basophils, mast cells, eosinophils) or destruction of malignant cells (natural killer cells). However, in practice, there is overlap between roles, and the distinction is less clear (Delves and Roitt, 2000). In addition, the innate immune system primes the adaptive response by way of antigen presenting cells and release of cytokines during infections. Therefore, the two systems can be viewed as being complementary to each other.

1.2.1 Macrophages:

Macrophages can arise from bone marrow-derived precursor cells known as monocytes. However, it has been shown in mice, that the ontology of tissue-resident macrophages varies (Ginhoux and Jung, 2014). Some tissue-resident macrophages, such as intestinal, dermal and cardiac macrophages, are continually renewed from circulating monocytes. Others, such as alveolar macrophages, arise prior to birth from primarily foetal liver derived stem cells and monocyte derived cells only contribute to renewal when resident cells are depleted to an extent that replication of resident cells cannot replenish numbers. Microglial cells, conversely, arise primarily from embryonic yolk sac derived precursors. Finally, Langerhans cells rise from a combination of both. These tissue resident macrophages can then be maintained within that tissue during normal homeostatic conditions (Yona et al., 2013). Furthermore, during tissue injury these resident cells release signals that stimulate circulating monocytes to migrate into tissues, where they then differentiate into either dendritic cells or tissue-specific macrophages. Moreover, different macrophage populations vary in their functional activity according to the environmental stimuli they encounter. This has led to the terms M1 and M2 polarisation being used to describe these different roles in mice. The current M1/M2 paradigm is likely to be an oversimplification and describes the extremes of the continuum, as it relies on descriptions of in vitro murine models with distinct stimuli

which don't occur in isolation in nature. Furthermore, it does not take into account the context of the stimulation nor the stage of maturation of the macrophages as it has been shown that stimulation of monocytes gives rise to different phenotypes to mature macrophages being stimulated (Martinez and Gordon 2014).. Finally, it is likely that the different lineages of macrophages such as tissue resident macrophages versus monocyte derived macrophages may have distinct responses. The M1 polarised macrophage is better suited to some aspects of bacterial killing, whereas the M2 macrophage is better suited to tissue repair (Martinez et al., 2008). However, there is overlap between functional states, as M2 macrophages are better suited to anti helminthic responses and some bacteria will cause a mixed response (Benoit et al., 2008). As macrophages are the resident tissue phagocyte, they are key in controlling infection They also have an important role in escalating the inflammatory response by releasing cytokines, and provide linkage to the adaptive immune response by recruiting T lymphocytes and dendritic cells, making them the cornerstone of the innate immune response.

1.3 Epigenetics:

1.3.1 Chromatin remodelling:

Cell function is governed by the dynamic translation of the genetic code into proteins in response to a diverse range of external and endogenous signals. The deoxyribonucleic acid (DNA) contained within the cell's nucleus encodes a series of genes, and it is by varying their expression that the cell is able to adapt to its environment and fulfil its role. The term epigenetics was initially used to describe heritable changes in the transcriptional phenotype of cells without altering DNA sequence (Berger et al., 2009). More recently it is also used to include transient changes in chromatin state that result in gene expression changes (Natoli, 2010). The epigenome is the net result of the interaction between the host genome and its environment. It is recognised as making a critical contribution to the control of gene expression in cells. Epigenetic changes, such as DNA methylation, histones post-translational modifications and miRNA interactions, regulate gene transcription through alterations other than changes in the DNA sequence (Jenuwein and Allis, 2001).

DNA is coiled around a histone octamer (composed of pairs of H2A, H2B, H3 and H4 histones) to form a nucleosome (Kornberg, 1974). In turn, chains of nucleosomes along with linker DNA and Histone H1 form chromatin (Figure 1.1, adapted from (Cole et al., 2014)). These histones are highly conserved throughout eukaryotic cells. The manner in which DNA is wound around histones and the modelling of the chromatin that results, dictates the availability of DNA sequences to the various transcriptional elements which regulate gene transcription. The traditional view, which is probably an oversimplification, states that chromatin can be densely packed, in which case it is termed heterochromatin. In this state, the DNA sequence is so tightly coiled that gene transcription is not possible. Conversely, in euchromatin, the bound DNA is loosely coiled and the coding sequence of genes is available for transcription. Evidence exists in drosophila of 5 distinct chromatin states based on different protein interactions and transcriptional states of the bound DNA sequence (Filion et al., 2010). Therefore, the control over chromatin density and the proteins bound to it, including the post-translational modifications of histones (PTM), are, in part, responsible for the regulation of gene transcription.

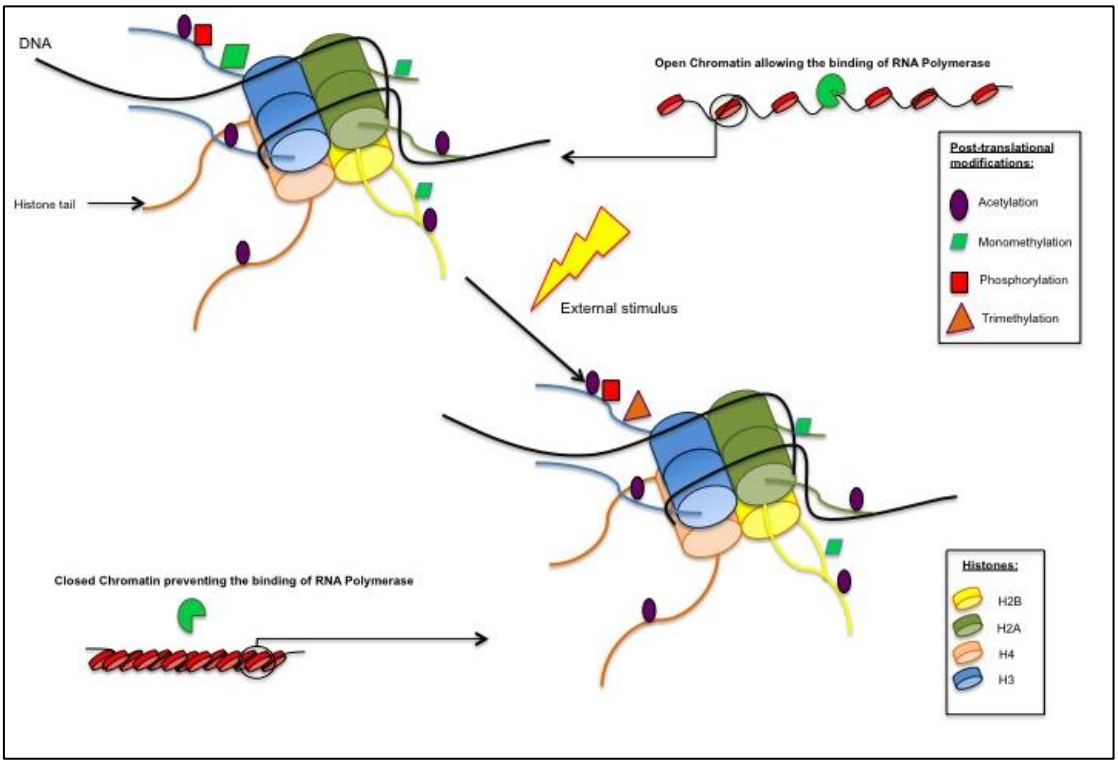


Figure 1.1: Schematic representation of histone post-translational modifications and their influence on chromatin remodelling. The figure shows a nucleosome composed of an octamer of histones with a schematic representation of the post-translational modifications (PTM) on the histone tails. Following an external stimulus the levels of trimethylation on the 4th lysine of Histone H3 (H3K4me3) are modified and along with other PTMs (not represented) this leads to a change in the conformation of the chromatin and subsequently the RNA polymerase is no longer able to bind the DNA. (adapted from (Cole et al., 2014)).

1.3.2 DNA methylation:

DNA methylation refers to the chemical addition of a methyl group to the cytosine base. This process is regulated by DNA methyltransferase (DNMT) 1 for maintenance and DNMT3A and DNMT3B for *de novo* methylation in embryonic tissues. The methylation of cytosine bases occurs principally at the CpG dinucleotides. Studies have shown that between 50 to 90% of these are methylated in mammalian tissues (Ehrlich et al., 1982). DNA methylation associated with the promoter region of genes leads to reduced gene expression, as the promoter region is not available to initiate transcription. Importantly DNA methylation appears to be conserved over time and is heritable (Smith and Meissner, 2013).

1.3.3 Histone post-translational modifications:

Modifications of the N-terminal tails of histones influence the availability of the bound DNA sequence by altering chromatin conformation or interacting with transcription factors (Berger, 2007; Jenuwein and Allis, 2001). These changes in the N-terminal domain of the histone tail, termed PTMs, are caused by chemical alterations, such as phosphorylation, methylation, acetylation, ubiquitination, propionylation and sumoylation. These chemical modifications cause adjoining histone tails to be either attracted to each other or repulsed (Rhee et al., 2014). The PTMs not only regulate the density of the chromatin binding but can also block the binding of transcription factors and RNA polymerase (Gerber and Shilatifard, 2003). There are a number of different

enzymatic processes involved in the control of these PTMs (Jenuwein and Allis, 2001). Acetylation is mediated by histone acetyltransferase. Deacetylation is controlled by histone deacetylase, which in humans include the 18 different enzymes grouped in 4 classes of which the sirtuins (class 3) are an important component.

Certain PTMs have been associated with up-regulation, or down-regulation, of the associated gene expression (Jenuwein and Allis, 2001). While it has been recognised for some time that bacteria and viruses can induce chronic changes to the epigenome influencing cancer susceptibility and the pathogenesis of chronic inflammatory diseases, the consequences of epigenetic changes during acute bacterial infections are not well understood (Esteller, 2007; Paschos and Allday, 2010).

1.3.4 Epigenetic influences on macrophage development:

The processes governing the differentiation of monocytes into macrophages have been extensively studied (Saeed et al., 2014; Wallner et al., 2016). It has been demonstrated that circulating monocyte differentiation into macrophages is associated with the combinatorial histone modification lysine 9 methylation serine 10 phosphorylation and lysine 14 acetylation on histone H3 (H3K9MeS10PhosK14Ac). However, differentiation into dendritic cells is associated with acetylation of lysine 16 on histone H4 (H4K16Ac) (Nicholas et al., 2014). Prior exposure to certain stimuli (β Glucan or lipopolysaccharide (LPS) leads to the differential maturation of monocytes into macrophages and is also associated with specific PTMs (Saeed et al., 2014). Immature cells exposed to Histone deacytelase inhibitors (HDACi), such as Trichostatin A, develop greater acetylation levels on H3 and H4, which is associated with increased production of the pro-inflammatory cytokine tumour necrosis factor α (TNF-α). Whereas dynamic changes in acetylation, which are higher in macrophages at baseline, do not appear to be a significant regulator of TNF-α in mature macrophages, potentially as these acetylation changes are offset by other aspects of chromatin remodelling that occur during maturation (Lee et al., 2003). This implies that maturation of cells influences epigenetic regulation and the extent to which dynamic changes in acetylation regulate gene expression by external stimuli, leading to a better adaption to particular functional roles. Other studies have suggested that maturation-related changes in innate immune responses could be regulated by epigenetic changes in histones as a response to prior stimulation of pattern recognition receptors. Macrophages derived from stem cells exposed to Toll-like receptor (TLR) 2 agonists prior to differentiation or during it, were found to have decreased ROS production (Yáñez et al., 2013) compared to macrophages from unstimulated stem cells. More recently, the differentiation of monocytes into macrophages has been shown to involve the removal of DNA methylation mediated repression of genes implicated in phagocytosis, which are rapidly activated after the deposition of histone PTMs and miRNA changes (Wallner et al., 2016).

1.3.5 Epigenetic regulation of macrophage function:

Macrophages fulfil diverse roles to enable the optimal functioning of the innate immune system. In order to perform these roles correctly, the macrophage is dependent on gene transcription. Genetic variation in the form of single nucleotide polymorphism for instance, contributes to the intensity of these responses influencing the susceptibility to infectious diseases. More recently, epigenetic variation has emerged as an additional point of regulation (Jenuwein and Allis, 2001). This allows a mechanism by which the environment can influence the transcriptional response (Galli et al., 2011). Epigenetic mechanisms have been shown to play a pivotal role in the regulation of macrophage

function in a range of different conditions. Smoking-induced inflammation in alveolar macrophages has been shown to be mediated by histone PTMs (Philibert et al., 2012), and the inflammatory response to particulate matter leading to decreases in DNA methylation by inhibiting DNMT1 activity (Miousse et al., 2014).

1.3.6 Trained immunity:

The epigenetics of macrophages has mainly been studied in the context of macrophage functional phenotype and/or differentiation (Nicholas et al., 2014; Saeed et al., 2014; Satoh et al., 2010; Yáñez et al., 2013). Both DNA methylation and histones PTMs have been shown to play a role in this process. Importantly, epigenetic mechanisms also explain, in part, the emerging field of "trained immunity" (Netea et al., 2011). Although the phenomenon of "trained immunity" has been recognized for a long time, it is only recently that we have been able to explain the mechanisms underpinning macrophage training. Indeed, as early as the 1970s observations were made of increased survival of mice following infection with the parasites *Babesia microti* and *Babesia rodhaini* after prior *Bacillus Calmette-Guérin* (BCG) vaccination (Clark et al., 1976). Increased survival in mice following pre-treatment with low dose *Candida albicans* when challenged subsequently with lethal doses of *C. albicans* or *Staphylococcus aureus* was also observed (Bistoni et al., 1986). There is evidence that macrophages are capable of developing a form of memory of antecedent events leading to a modified response and a survival advantage that appears to be mediated by epigenetic mechanisms. This evidence is also drawn from observations of individuals post BCG vaccination. Macrophages from individuals vaccinated with BCG were found to have increased TNF-α and Interleukin 1β (IL-1β) following repeat exposure to *Mycobacteria tuberculosis* and / or *Candida albicans* (Kleinnijenhuis et al., 2012). Observations such as these have been made in a number of different models deficient in adaptive immunity, such as plants, where it is termed "systemic acquired resistance" (Durrant and Dong, 2004) and in invertebrates (Moret and Siva-Jothy, 2003; Pham et al., 2007). Historically, crossover protection has been reported following vaccination in humans or challenge with one organism, which confer survival benefits against a subsequent infectious challenge. Experiments in healthy volunteers demonstrated that following vaccination with BCG, upon re-challenge with bacteria, *Mycobacteria* or *Candida*, monocytes had increased production of pro-inflammatory cytokines, and this was associated with increased levels of trimethylation of lysine 4 on Histone H3 (H3K4me3) (Kleinnijenhuis et al., 2012). Kleinnijenhuis and colleagues also demonstrated that the mechanism of training monocytes with BCG is independent of TLR2 and TLR4, instead it relies on nucleotide-binding oligomerization domain-containing protein (NOD) 2 receptors. They then established the functional consequences of training macrophages in a lymphocyte-deficient murine model (using severe combined immunodeficiency (SCID) mice). They showed that BCG vaccination led to significant survival and decreased fungal burden in kidneys when mice were subsequently challenged with a lethal dose of *C. albicans*, thereby providing *in vivo* evidence of innate immune memory.

Innate memory has also been seen when pre-treating recombination activating gene (Rag1) deficient mice, which lack functional T- and B- cells and represent another model of severe combined immunodeficiency, with low dose *C. albicans*. Increased survival was observed compared to CCR2-deficient mice when re-infected with a lethal dose of *C. albicans* (Quintin et al., 2012), emphasizing that survival involves training of $CCR2^{+}$ monocytes and not lymphocytes. Genome-wide chromatin immunoprecipitation

and sequencing (ChIP-Seq) for the histone PTM H3K4me3 showed increased levels of this PTM at the promoter regions of genes involved in inflammation, such as TNF-α, interleukin 6 (IL-6) and interleukin 18 (IL-18). This highlights an epigenetic mechanism underlying the basis for training the innate immune system.

1.3.7 Epigenetics and tolerance:

The macrophage response to external stimuli is a careful balancing act between generating a pro-inflammatory response to clear pathogens and the need to subsequently resolve the inflammatory process and generate a tissue healing response to prevent excessive and chronic inflammation leading to tissue injury. The mechanisms facilitating this include epigenetic regulation. The principal models used to study inflammation have been those of endotoxic shock. Acute lipopolysaccharide (LPS) exposure, signals via TLR4, to induce expression of a number of different pro-inflammatory genes. In contrast, repeated exposure of monocytes to LPS leads to a tolerant monocyte, which produces markedly lower pro-inflammatory cytokines, such as TNF-α, than when first exposed. In murine models of sepsis, following caecal ligation and puncture, LPS tolerance increases survival (Wheeler et al., 2008). Repeated exposure to certain stimuli, such as TLR4 agonists, stimulates tolerance, whereas the response to some other stimuli induces an accentuated pro-inflammatory response, consistent with the model of trained immunity (Ifrim et al., 2014). Tolerance helps decrease the risk of death from an exaggerated pro-inflammatory state, as observed in sepsis. Upon repeated LPS challenge, the activation of pro-inflammatory cytokine genes, such as IL-6, is reduced. As it is the same TLR4 pathway that is activated by repeated exposure to LPS, both signal transduction pathways and epigenetic changes regulate this response leading to the tolerance phenotype.

The promoter regions of pro-inflammatory genes in macrophages during the initial response and on re-challenge with LPS have been shown to be associated with varying levels of H3K4me3 and H4 acetylation (Foster et al., 2007). Foster and colleagues demonstrated that on LPS re-challenge, global H4 acetylation levels in the promoter regions of genes involved in LPS tolerance were decreased (Foster et al., 2007). However, genes that continued to be actively transcribed on re-challenge maintained their H4 acetylation level. Foster and colleagues also showed that the TLR4 response to repeated LPS exposure led to a loss of H3K4me3 in the promoter regions of genes that demonstrated LPS-tolerance, while in genes associated with non-tolerance, H3K4me3 was maintained. Furthermore, when pre-treated with Trichostatin A (a histone deacetylase inhibitor) during the initial LPS exposure, the pro-inflammatory gene expression of the macrophage is similar to the naïve macrophage upon re-stimulation, restoring the secretion of IL-6. This was also seen when pre-treating with pargyline (an inhibitor of H3K4 demethylase), confirming that both H4 acetylation and H3K4me3 marks are important components in regulating responses during tolerance (Table 1 adapted from (Cole et al., 2014)). Interestingly, the authors observed that the LPS-induced gene responses that did not demonstrate LPS tolerance, included those involved in antimicrobial responses, ensuring these responses were not down-regulated on re-challenge (Foster et al., 2007).

Table 1.1: Summary of selected histone PTMs discussed above.

External stimuli	Post-translational modification:	Pathway	Outcome	Murine experiment outcome

BCG vaccine	↑ H3K4me3	NOD2	↑ TNF-α, IL-1β, IL-6, IL-18	Increased survival in candidaemia
LPS challenge	↑ H3K4me3 & H4ac	TLR4		
LPS re-challenge	↓ H3K4me3 & H4ac	TLR4	↓ TNF-α and IL-6	Increased survival in caecal ligation and puncture model
ROM A	↑ H3K4me		↓ IL-6	
Listeriolysin	↓ H3S10phos & H4ac		↓ IL-6	

1.3.8 Epigenetic control of macrophage polarisation:

Macrophage polarization is controlled in part by epigenetic regulation. Macrophage polarisation has been described in terms of M1 (or classically-activated) and M2 (or alternatively-activated) although in reality these occur across a spectrum. In particular, during the induction of an M1 phenotype in macrophages, the repressive marks trimethylation of lysine 27 (H3K27me3), trimethylation of lysine 9 (H3K9me3), and trimethylation of lysine 20 (H4K20me3) are removed, which is followed by an increase in the activator marks phosphorylation of serine 10 (H3S10phos), global levels of acetylation on Histone H4 and H3K4me3 (De Santa et al., 2009a; Stender et al., 2012). This, in turn, leads to activation of genes involved in the acute response to external stimuli associated with the M1 phenotype. The epigenetic regulation of inflammatory responses has triggered interest in applying novel bromodomain inhibitors, such as the synthetic compound I-BET, to modulate these responses. The M2 phenotype is associated with the removal of the H3K27me3 repressive mark and is mediated by the H3K27 demethylase Jumonji domain containing-3 (*JMJD3*) (Satoh et al., 2010). Satoh and colleagues suggest that *JMJD3* is involved in both the mediation of M1 and M2 phenotypes, but only appears to play a minor role in the M1 response, enabling fine-tuning (De Santa et al., 2009a; Satoh et al., 2010). In contrast, it plays a more significant role in M2 polarization, allowing transcription of the transcription factor, interferon regulatory factor (Irf) 4, playing a crucial role in M2 polarization (Satoh et al., 2010).

Further studies have assessed global H4 acetylation levels, a PTM associated with enhanced transcription, in polarized macrophages. The results were mixed but showed that some genes, and in particular those whose expression was modified by the cytokines used to polarise the MDMs, had increased levels of H4 acetylation at their promoter regions (Zhang et al., 2011). However, not all genes whose expression varied during polarization showed changes in H4 acetylation levels. This suggests that multiple different PTMs may interact to govern polarisation. Furthermore, the association of DNA methylation and combinatorial PTMs have not been studied to date.

Thus macrophages are subject to epigenetic regulation by multiple mechanisms including histone PTMs, DNA methylation and miRNAs and these operate both during differentiation from monocyte to macrophage, but also in differentiated tissue macrophages. This complex regulation allows fine-tuning of both a strong initial response and protection from the sustained effects of innate immune activation. As illustrated above, epigenetic mechanisms play a role in a variety of processes including immune training, induction of tolerance and macrophage polarisation (Table 2, adapted from (Cole et al., 2014)). Finally, these regulatory mechanisms have also been shown to play a role in a number of different disease processes, such as cancer and coronary

artery disease. This may well contribute to the link between inflammatory responses (and by extension to infection) and the pathogenesis of cancer or atherosclerosis.

Table 1.2: Summary of selected known PTMs and influence on macrophage polarisation.

Phenotype	Function	Decreased PTMs	Increased PTMs
M1 (or "classically activated")	Secrete IL-1β, IL-6, IL-12, NO generation, intracellular bacterial killing	H3K27me3, H3K9me2, H4K20me3	H4ac, H3S10phos, H3K4me3
M2 (or "alternatively activated")	Repression of inflammation, Arginase production, anti-helminthic properties	H3K27me3	

1.4 Macrophages and infections:

The subversion of the host immune system is one of the key components of bacterial pathogenesis. In effect, microorganisms highjack the host gene expression to their benefit. In macrophages, post-translational modifications have been shown to be induced by a number of different bacterial components including LPS, listeriolysin and pneumolysin (Hamon et al., 2007). Experiments using both *Legionella pneumophila* (Rolando et al., 2013), and *Listeria monocytogenes* (Eskandarian et al., 2013) have shown that bacterial interaction with the THP-1 monocytic cell line also induces histone PTM.

In the case of *Legionella* the histone PTMs caused during infection have a direct survival benefit. *Legionella pneumophila* has been shown to use posttranslational modification to modulate the gene expression within the infected cell to aid intracellular replication (Rolando et al., 2013). This results in the pathogen secreting ROM A, a SET domain containing methyltransferase, leading to increased H3K14 methylation and decreased levels of acetylation. These histone PTMs were associated with the promoter regions of genes, such as IL-6. This in turn, shows that *L. pneumophila* modulates host gene expression in particular, in genes relating to innate immunity, to enhance its intracellular survival.

Listeriolysin (LLO), a pore-forming cytolysin similar to PLY, secreted by *Listeria monocytogenes* has been shown to cause dephosphorylation at serine 10 of H3 and decreases in the levels of acetylated H4 in THP-1 cells. This correlated with a change in the transcriptional profile in HeLa cells, which was associated with a decrease in IL-6 and other genes involved in innate immune responses (Hamon et al., 2007). This lends strength to the theory that bacteria use epigenetic modulation in order to limit the inflammatory response and increase their survival.

The consequences of epigenetic changes during acute bacterial infections are not fully understood. They can enable pathogen subversion of the host's immune responses, but may also be used by the host to modify the consequences of these responses, for example to limit the damage induced by sustained inflammation or enhance responses on re-challenge by training.

1.5 Susceptibility to infection:

1.5.1 Genetic susceptibility to infection:

The basis of susceptibility to infection is not fully understood. There is clear evidence that some host genetic factors play a role in host susceptibility to infection, as

evidenced by single gene defects, as is the case in IL-12R β1 receptor deficiency which is associated with invasive mycobacterial infections (Lammas et al., 2000), and by genome-wide studies identifying key polymorphisms (Chapman and Hill, 2012), such as single nucleotide polymorphisms in surfactant proteins and susceptibility to meningococcal infections (Jack et al., 2006). In the case of pneumococcal infection interleukin-1 associated kinase 4 (IRAK4) deficiency or nuclear factor κB modulator protein (NEMO) deficiency (Brouwer et al., 2009) have been associated with disease susceptibility. However, this does not fully explain host susceptibility. It has recently been postulated that some of the factors influencing the development of our immune system are epigenetic (Gómez-Díaz et al., 2012) in nature and, as such, I propose that some of the susceptibility infections and in particular to pneumococcal infection may be epigenetic.

1.5.2 Epigenetic susceptibility to infection:

Bacterial pathogens have evolved a number of different mechanisms to subvert the host's defences. In particular, it has been recognized that they can manipulate the host's own intracellular signalling and alter the host cell's gene expression (Bhavsar et al., 2007). The importance of epigenetic changes occurring as part of the pathogenesis of infectious diseases is becoming increasingly understood. There are a number of examples of pathogens altering gene transcription by modifying the host cell's epigenome, for example *Legionella pneumophilia, Helicobacter pylori* and *Listeria monocytogenes* have been shown to alter histone PTMs, thereby altering gene expression during infections (Ding et al., 2010; Hamon et al., 2007; Rolando et al., 2013).

When gastric epithelial cells were infected with *H. pylori,* a time dependent dephosphorylation of serine 10 of histone H3 (H3S10) was observed as well as a decrease in actetylation levels on lysine 23 (Ding et al., 2010). In the case of dephosphorylation, the investigators found that these changes were associated with the promoter regions of genes involved in inflammation and were associated with induction of IL-8 in gastric epithelial cells. Interestingly, this mechanism is likely to be both organism and cell specific, since in THP-1 cells the dephosphorylation of serine 10 on Histone H3 (H3S10) was associated with the promoter region for the gene coding for IL-6 and was associated with an increase in release of IL-6 (Pathak et al., 2006). The exact mechanism by which *H. pylori* causes alterations in histone PTMs change is not fully understood, although it is apparent that it involves the cytotoxin-associated gene A pathogenicity island (cagPAI), since deletion of this, but not cagA or other factors, induced the dephosphorylation of H3S10 (Ding et al., 2010). These changes were also associated with up-regulation of the oncogene c-Jun and with down-regulation of heat shock protein (hsp) 70, showing they not only contribute to inflammation but also to tumour development. Fehri et al showed *H. pylori* infection resulted in dephosphorylation of H3S10 but also of threonine 3 of H3 (Fehri et al., 2009). These events were also linked to the cagPAI and a functional type 4 secretion system and were absent in a ΔcagL mutant. cagL has been associated with binding to the cell surface. The changes were associated with bacterial induced reduction of cell division cycle 25 phosphatase and a resultant reduction in activation of the H3 vaccinia-related kinase (VRK) I and *H. pylori* induced pre-mitotic arrest.

In the case of *L. pneumophilia,* the bacterium uses a type 4 secretion system effector regulator of methylation A (RomA), a SET domain containing methyltransferase, which trimethylates lysine 14 on histone H3 of the host cell and

reduces histone H3 lysine 14 acetylation (H3K14ac), thus switching off gene transcription (Rolando et al., 2013). Following ChIP-Seq, the authors established that the switch to the repressive histone H3 lysine 14 trimethylation mark (H3K14me3) was associated with the promoter sites of genes involved in immune responses. In particular, cytokines such as TNF-α and IL-6; PRR such as TLR5 which responds to flagellin, and the Nod-like receptor Nacht, LRR and PYD domains containing protein 3 (Nalp 3). This provides support to the theory that pathogens, including bacteria and viruses, can use chromatin remodelling strategies to turn off immune responses.

In the case of *L. monocytogenes*, secretion of the surface protein internalin B (InlB) leads to translocation of a host cell class 3 histone deacetylase (HDAC), sirtuin 2 (Sirt2), to the nucleus where it deacetylates lysine 18 of H3. This is associated with a decrease in expression of genes involved in DNA binding and immune responses (Eskandarian et al., 2013). Moreover, *L. monocytogenes* also secretes listeriolysin O (a pore-forming cytolysin), which causes dephosphorylation of H3S10 and decreases the level of H4 acetylation (Hamon et al., 2007). These responses occurred by a pore-forming independent mechanism. This correlated with a change in the transcriptional profile of HeLa cells, which was associated with a decrease in genes involved in innate immune responses including the neutrophil chemokine C-X-C motif ligand 2 (CXCL2) and dual specificity phosphatase 4 (DUSP4) , a phosphatase involved in regulating mitogen activated protein kinase (MAPK) signalling.

It remains unclear what the consequences of epigenetic changes during acute bacterial infections are. Since PLY plays a pivotal role in modifying multiple aspects of the host response to pneumococci, including microbiocidal activity, generation of pro-inflammatory cytokines and immune cell viability, epigenetic modifications in macrophages in response to PLY have the potential to play a key contribution to the outcome of infection.

There is a growing body of evidence that epigenetic traits can be passed from one generation to the next. In rats, hepatic injury induced with the hepatotoxic tetracarbon CCL_4, in the parent generation leads to a protective effect in the offspring, mediated by modulation in trimethylation levels of H3K27me3 (Zeybel et al., 2012). Moreover, it has been recognised for some time that bacteria and viruses can induce chronic changes to the epigenome inducing susceptibility to cancer and the pathogenesis of chronic inflammatory diseases (Esteller, 2007; Paschos and Allday, 2010). These heritable traits are transmitted without a change in DNA sequence, therefore it is possible that the heritable epigenetic make-up of an individual may determine the response of their off-spring to future infection and this may explain some differences in innate responses between individuals.

Therefore, it is possible that epigenetics could play a number of roles in the host pathogen interaction and determine susceptibility to infectious diseases. First, this could be due to a pathogen released virulence factor that alters histone PTMs and leads to conformational change in the chromatin in turn blocking transcription (as described in Figure 1.1) and therefore immune responses. In addition, it is possible that, as is seen in the response to LPS tolerance, repeated exposure leads to changes in the histone PTMs, again leading to conformational change of the chromatin which impairs gene transcription and leads to a reduced immune response (as is described in Figure 1.1.2).

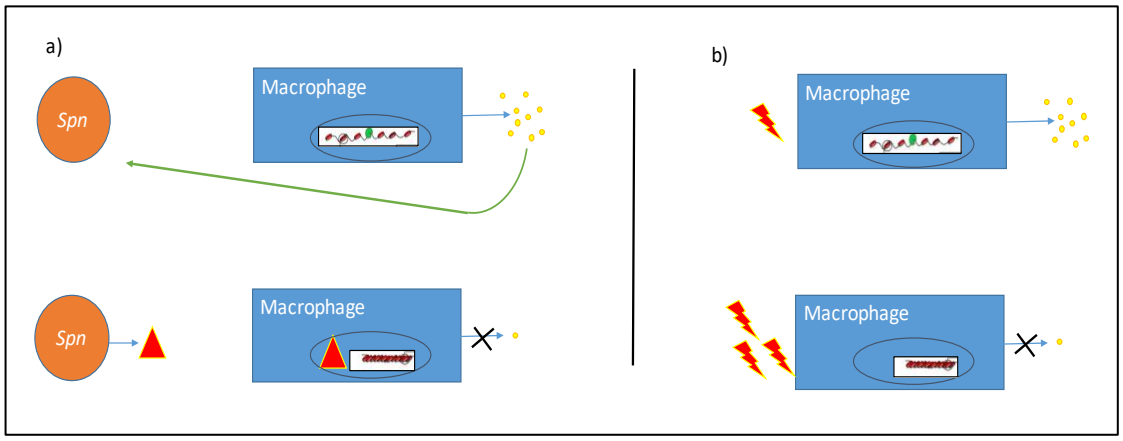


Figure 1.1.2: Schematic representation of the role played by histone PTMs in the host pathogen interaction. Panel a depict the release of pro-inflammatory cytokines following exposure of macrophage to a pathogen. This is then impaired when a virulence factor released by the pathogen leads to transformation from euchromatin in the nucleus to heterochromatin and therefore the inability for RNA polymerase to bind to the DNA and initiate gene transcription leading to a decrease in the innate immune response. Panel b illustrates the phenomenon of tolerance, whereby upon repeated exposure to a stimuli such as LPS, epigenetic changes occur to impair the release of pro-inflammatory cytokines and therefore hamper the immune response.

1.6 Epigenetic manipulation of macrophages and infection:

There is mounting interest in the use of HDACi in the fields of oncology, HIV and inflammatory disorders. Their use has been associated with survival benefits in murine models of sepsis. In lethal LPS challenge models, the use of suberoylanilide hydroxamic acid (a HDACi) was associated with increased survival (Li et al., 2010). However, other groups have reported detrimental effects of HDACi on macrophage killing ability with decreased clearance of bacteria and fungi (Mombelli et al., 2011; Roger et al., 2011). Rodger demonstrated decreased nitric oxide (NO) and reactive oxygen species (ROS) production in macrophages treated with a HDACi, thereby impairing microbial killing of *Escherichia coli* and *Staphylococcus aureus*. Nevertheless, they also observed increased survival following LPS challenges. Tubustatin (a HDAC6 inhibitor) was associated with increases in circulating monocytes numbers and the authors suggested that this may lead to faster bacterial clearance and therefore explain the survival advantage seen when used in the CLP model in mice (T. Zhao et al., 2014b). They also showed Tubustatin reversed the increased acetylation levels seen in sepsis and thereby the organ dysfunction, lymphoid apoptosis and the attenuated stress responses (T. Zhao et al., 2014a). Moreover, recent studies of the use of Cambinol (a Sirtuin 1 & 2 inhibitor) was associated with anti-inflammatory properties and led to attenuation of innate immune responses protecting mice from endotoxic and toxic shock (Lugrin et al., 2013). Therefore, it is likely that HDACi could play a role in the control of inflammatory responses, such as sepsis and other autoimmune conditions, but, as they may impair initial macrophage response to infection, further studies will be required to optimize the timing of such drug therapy. Other strategies are likely to be developed, with the advent of specific small molecules targeted at individual modifications, such as the novel *JMJD3* H3K27 demethylase inhibitors which lead to a modulation in the inflammatory response to LPS (Kruidenier et al., 2012). These conflicting results and possible future developments are summarised in Figure 1.2.

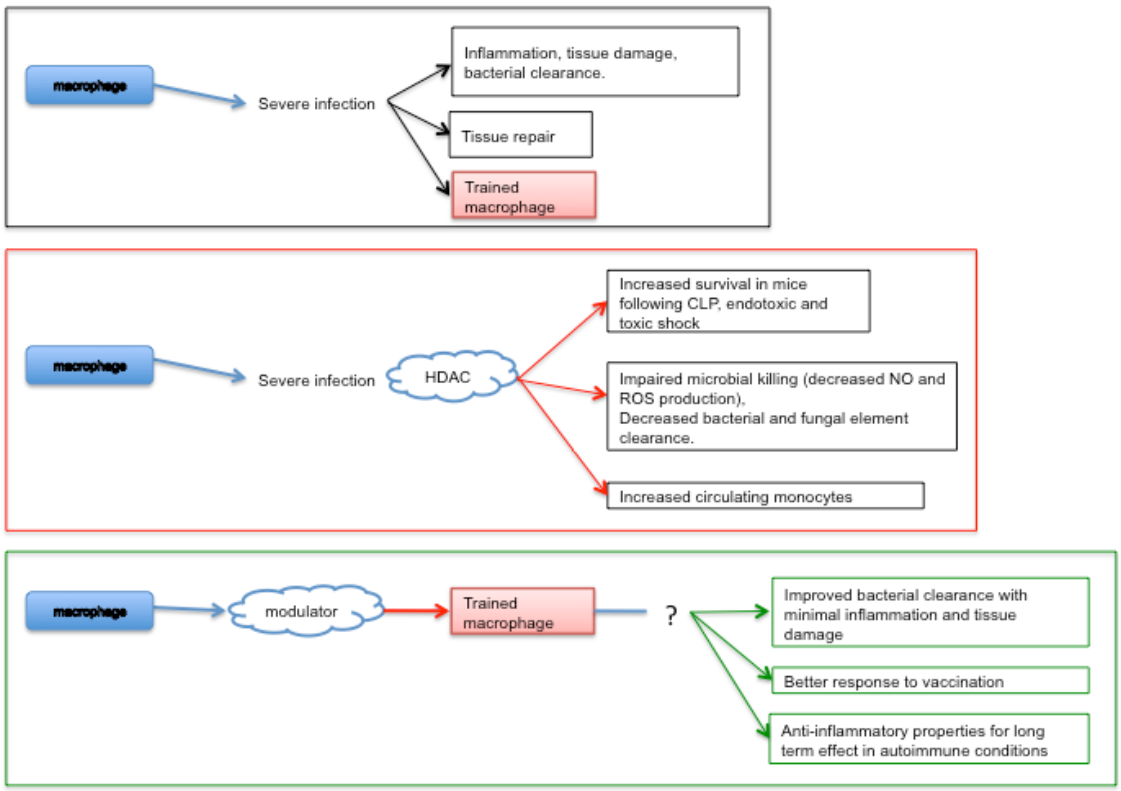


Figure 1.2: Potential epigenetic modulation. This figure summarises the current knowledge of the conflicting outcomes associated with the use of HDACi on macrophage function. The green box suggests what might be possible in the future using novel epigenetic modulators such as CRISPR CAS9 editors to alter PTM and specifically enhance or repress individual gene expression, or by influencing polarisation of the macrophage control inflammation.

Most of the pharmacological interventions to date have involved HDACi, and it is well recognised that these have a range of off-target effects. Furthermore, older studies have used pan HDACi. This may explain the conflicting results associated with their use in animal studies with these agents demonstrating both protection in shock but also impairment of microbial killing. As more specific compounds arise, such as specific bromodomain inhibitors and / or advances in clustered regularly interspaced short palindromic repeats associated protein 9 (CRISPR/CAS9) technology, (see Figure 1.3), this will enable targeted manipulation of macrophage function by inhibition of individual genes within a tissue-specific macrophage. These approaches should enable clinicians to more selectively modulate the inflammatory response or heighten key antibacterial responses, which will be potentially crucial in an era of ever-increasing antimicrobial resistance (Figure 1.2). Furthermore, if histone PTMs are associated with susceptibility to infections, they could be used to screen individuals for susceptibility to infection to enable preventative strategies such as targeted vaccination, or be themselves targets for novel therapies to induce a more favourable epigenetic profile. Indeed, recent advances in the field of clustered regularly interspaced short palindromic repeats (CRISPR) technology has enabled targeted modification of the epigenome in an effort to manipulate gene regulation. This approach uses a programmable CRISPR/CAS9 construct fused with an acetyltransferase (p300), allowing the acetylation of the promoter region of specific genes (Hilton et al., 2015). This ground-breaking development paves the way for reversible editing of the epigenome and thereby a specific approach targeting cell function by influencing gene expression without altering the underlying DNA. This could lead to the targeted resolution of inflammatory

processes or to the manipulation of individual epigenomes to reduce the susceptibility to disease.

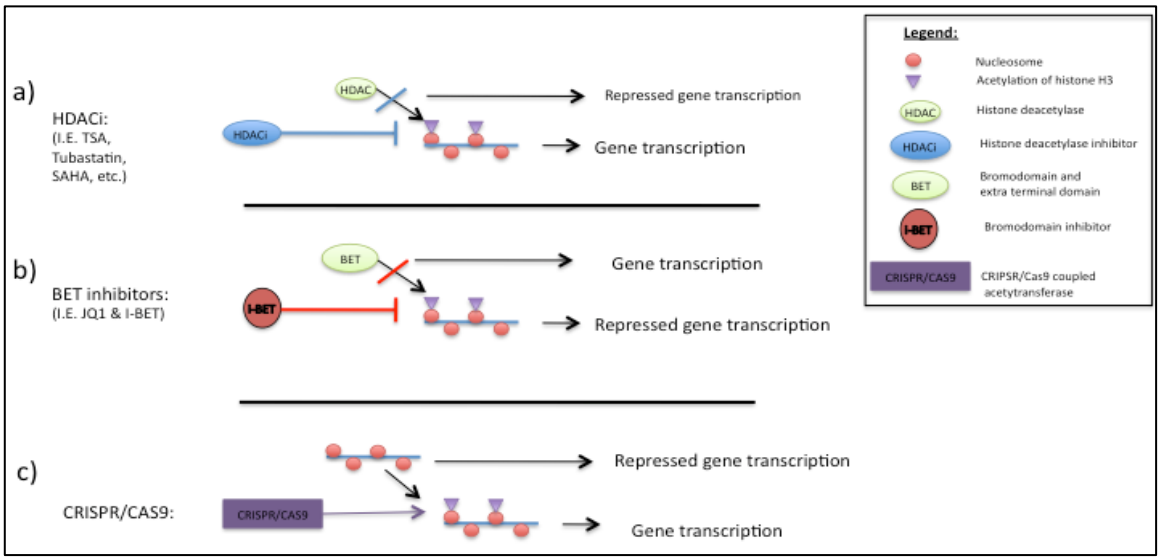


Figure 1.3: Diagram depicting different mode of actions of PTMs modulators.
a) Schematic representation of the mode of action of HDACi. Leading to the maintenance of the acetyl post-translational modification and thereby ongoing gene transcription. b) Schematic representation of Bromodomain inhibitors such as I-BET151 or JQ1, these inhibit the binding of Bromodomains such as BRD4 to the underlying chromatin thereby inhibiting gene transcription. Bromodomains are involved in reading acetylation and are found on a variety enzymes such as methyltransferases, transcriptional co-activators and are involved in recruiting multi molecular complexes that modify chromatin or contribute to gene transcription (Muller et al., 2011).c) Schematic representation of the mode of action of CRISPR/CSAS9 construct, this leads to the acetylation of the promoter region of a specific gene leading to its transcription. (Adapted from (Cole et al., 2016).

1.7 Identification and characterisation of histone PTMs:

As discussed above, there is a growing body of evidence illustrating the importance of epigenetic changes in the host's response to infection. Both *Listeria monocytogenes* and *Legionella pneumophilia* have been shown to alter the host epigenome to provide them with a survival advantage (Hamon et al., 2007; Rolando et al., 2013). Furthermore, the underlying macrophage polarisation and therefore its response to bacterial challenge, is associated with distinct histone PTMs (Ivashkiv, 2013). Of the different aspects of the epigenome, microRNA, DNA methylation and histone PTMs, the later have been most extensively studied in the context of infections. Therefore, I chose to study histone PTMs in the context of host response as they are key to further understand the host-pathogen interaction.

1.7.1 Antibody based approaches:

Several different approaches have been developed for the study of histone PTMs. The commonest of which is the use of antibodies developed against individual histone PTMs. This allows a semi-quantitative detection of a PTMs using western blotting or dot blot. This approach has been successfully used by a number of groups to determine the changes in abundance of histone PTMs following infection (Hamon et al., 2007). However, as with all techniques, there are a number of limitations with respect to antibody based techniques for probing histone PTMs, including problems associated with selectivity, epitope occlusion and cross reactivity (Egelhofer et al., 2011; Fuchs et al., 2011). Furthermore, it is not currently commonplace to establish combinatorial modifications by these approaches and they require prior knowledge of the modification

to be identified, introducing bias, and moreover won't allow the detection of novel PTMs.

1.7.2 Characterising histone PTMs by mass spectrometry:

Mass spectrometry (MS) has emerged as a powerful method to characterize and quantify histone PTMs. It allows unbiased identification and quantification of multiple histone PTMs including combinations in a single analysis; a significant advantage over other antibody based semi-quantitative approaches, such as western blotting. Recently, a plethora of different approaches have been described to perform the study of histone PTMs (Moradian et al., 2014; Young et al., 2010) These include different types of analysis, from intact proteins in top down (Pesavento et al., 2004, 2008), large peptides following Glu-C digestion in middle down (Molden and Garcia, 2014; Phanstiel et al., 2008), and smaller peptides following digestion with trypsin or Arg-C in bottom up approaches (Minshull et al., 2016; Plazas-Mayorca et al., 2009). Each approach has its own merits (summarised in Figure 1.4) (Karch et al., 2016; Moradian et al., 2014; Zheng et al., 2016)

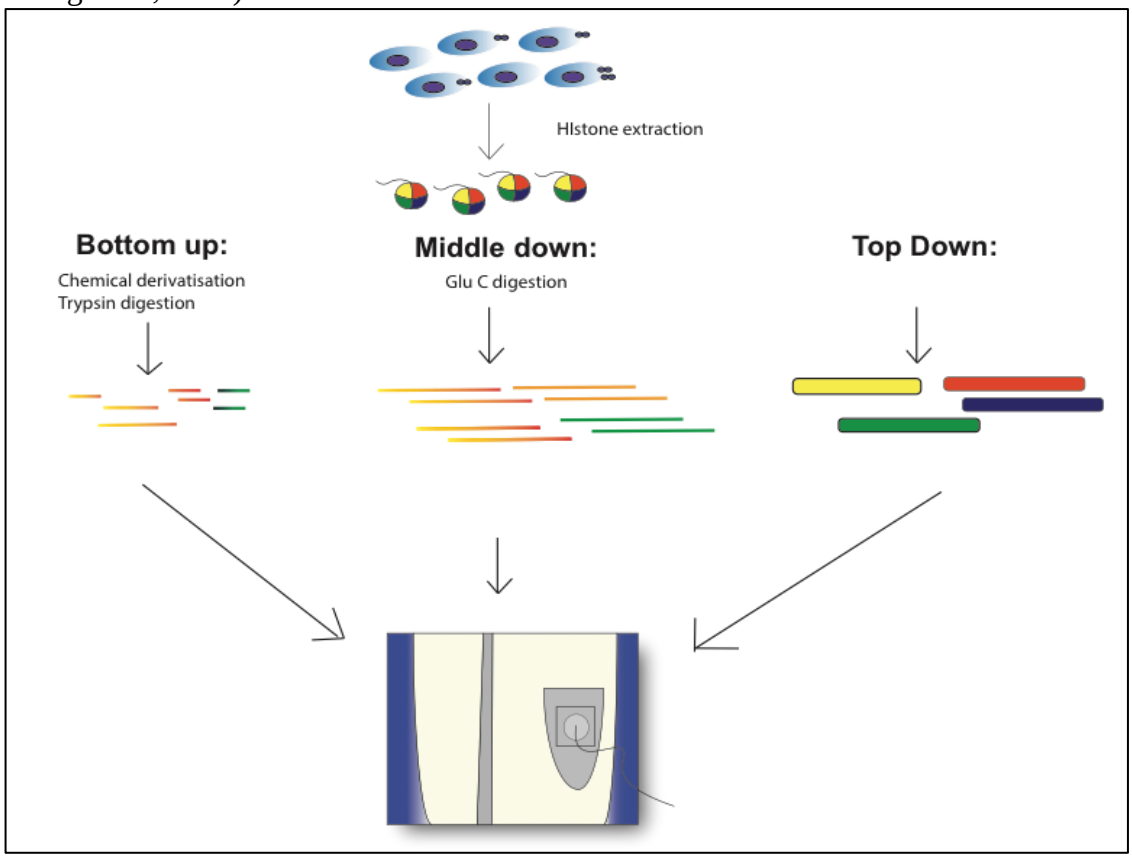


Figure 1.4: Schematic representation of different mass spectrometry approaches for the analysis pf histone PTMs.
Bottom up approaches are preceded by endoproteinase Arg-C or trypsin digestion resulting in multiple small peptides that are then analysed, Middle down approaches are preceded by an endoproteinase Glu-C digestion leading to longer peptides and therefore more combinatorial information but is limited by increased difficulty in analysis, Finally, Top down analysis allows study of intact protein without digestion but is limited by the necessity for complex bioinformatics pipeline set ups.

The top down analysis provides information about combinatorial modifications by analysing the intact protein. However, in light of the extensive modifications of histones and since many of these are isobaric and can co-elute, this results in very

complex spectra where the identification of individual proteoforms proves challenging. Furthermore, certain histone PTMs are not abundant and identifying these proves difficult and requires large amounts of initial protein. In order to maximize coverage, extensive 2 dimensional fractionation is required, limiting its use as a high-throughput approach (Pesavento et al., 2008).

The middle down approach provides a compromise between the other two approaches. It allows for larger peptides than bottom up methods. However, it is hampered by the complexity of analysing tandem mass spectrometry spectra (MS2) to correctly identify the proteoforms. Although it has recently been shown to be equivalent to bottom up approaches for quantification (Sidoli et al., 2015a; Sidoli and Garcia, 2017), it remains rarely used due to the requirement for complex sample separation and bioinformatics analysis.

Bottom up approaches have been the most extensively employed for the study of histone PTMs. These allow improved quantification of individual proteoforms and separation of isobaric peptides by use of reverse phase high performance liquid chromatography (HPLC) and offers a high throughput method (Karch et al., 2016; Minshull et al., 2016; Plazas-Mayorca et al., 2009; Sidoli et al., 2016, 2015b, 2015c).

Given the requirement for complex bioinformatics support to analyse top down and middle down approaches and that the bottom up approach shows comparable results to middle down (Sidoli et al., 2015a), I focused on the development and application of bottom up approaches to analyse histone PTMs.

1.7.3 Data acquisition methods:

1.7.3.1 Data dependent acquisition:

Not only are there different sample preparation approaches to MS, but several different data acquisition strategies have been developed and employed for the analysis of bottom up histone PTMs (Karch et al., 2016; Sidoli et al., 2015b, 2015c, 2016). The commonest approach is data dependent acquisition (DDA) as this doesn't require any prior knowledge of the PTMs to be identified (Plazas-Mayorca et al., 2009). During the analysis, the top N-eluting peptides are selected for fragmentation and analysis. However, the quantification of isobaric co-eluting peptides using this approach proves challenging. In addition, low abundant proteoforms may not be selected for fragmentation and tandem MS (MS/MS) and therefore won't be identified and quantified. In an effort to optimize the number of histone PTMs identified using lower speed instruments such as maXis Time of Flight (ToF) mass spectrometer, off line 2D-HPLC fractionation methods have been developed (Minshull et al., 2016). Although the use of fractionation increases the number of PTMs identified, it does so at the cost of speed as up to 20 hr of MS time were required for each sample.

1.7.3.2 Targeted data acquisition:

In light of this, selective reaction monitoring methods (SRM) and parallel reaction monitoring (PRM) have been developed (Peach et al., 2012; Bourmaud et al., 2016; Peterson et al., 2012). These approaches rely on the establishment of an isolation list for all of the PTMs to target for MS/MS. These are then monitored throughout the HPLC gradient and selected for fragmentation and analysis. These approaches improve the sensitivity, especially for low abundant proteoforms, but are constrained by total cycle time for multiple PTMs as these can elute in different charge states. They are limited by the number of transitions to be monitored throughout the gradient and the

need for prior knowledge of which ones to target. Moreover, they will not allow retrospective analysis for novel PTMs.

1.7.3.3 Data independent acquisition:

In order to overcome some of these limitations data independent acquisition (DIA) methods have been gaining popularity for discovery proteomics and are particularly suited to the study of PTMs (Krautkramer et al., 2015; Sidoli et al., 2015c, 2015b, 2016). A number of different DIA methods have been used to study histone PTMs. One of the first methods developed was SWATH™ (ABSciex), designed for triple ToF instruments. This method was successfully used to identify and quantify histone PTMs (Sidoli et al., 2015b). It involves a series of 85 isolation windows of variable sizes spanning the range of mass to charge ratios (m/z) in which histone PTMs are found (see Appendix table 1.1). Subsequently, Krautkramer et al. used a DIA method with regular 10 m/z isolation windows to identify and quantify the changes in histone PTMs following histone deacetylase inhibitor treatment (Krautkramer et al., 2015). Using this approach enabled a greater reproducibility than conventional DDA with consistently high numbers of proteoforms identified and with lower Coefficient of Variations (CV) in the relative abundance in DIA runs when compared to DDA. Indeed, both the SWATH™ and DIA were able to detect low abundance proteoforms. Previous studies have shown that DIA protocols can be adapted to low resolution ion trap instruments (Sidoli et al., 2015c). In this study from the Garcia group, they demonstrated the adaptability of low resolution DIA using 50 m/z sequential windows. Furthermore, the same group had previously compared both a high resolution to a low resolution instrument for the analysis of Histone PTMs in DDA mode (Karch et al., 2014).

The new mass spectrometry instruments possess increasingly higher speeds making the use of higher resolutions more practical in a high throughput manner. It has been shown that histone PTMs can be studied using low resolution instruments (Karch et al., 2014; Sidoli et al., 2015c). In particular, previous studies have focused on comparing instruments with different resolutions or different data acquisition methods. But the comparison between high and low resolution modes on the same instrument has yet to be done. In this study, I look to optimise and establish the best possible method for the study of histone PTMs.

1.7.4 Chromatin immunoprecipitation and sequencing:

Despite the clear advantages of MS-based techniques to study chromatin modifications, they do not provide any information about the relationship of these protein PTMs with the bound DNA. They are indeed limited to describing changes occurring at a global level and cannot further characterise the changes in histone PTMs occurring at an individual gene level. In order to explore the function of the PTMs and the associated DNA sequence chromatin immunoprecipitation combined with next generation sequencing (ChIP-Seq), techniques have been developed (Barski et al., 2007; Johnson et al., 2007). Briefly, these techniques rely on using an antibody to bind the histone PTMs (or transcription factor) of interest and then the antibody is used to pull down the cross-linked DNA sequence, which is then purified and sequenced. This can then provide information about the position of histone PTMs in relation to gene sequences. ChIP-Seq approaches have allowed large consortium such as NIH Roadmap Epigenomics Consortium, to establish the association of the different patterns of histone PTMs to be associated with certain genetic regions such, as enhancer or promoter

regions, and be associated with activated / repressed gene expression (Roadmap Epigenomics Consortium et al., 2015). However, as with the western blot approaches, these techniques are limited by prior knowledge of the histone PTM, antibody availability and cross reactivity issues.

1.8 Aims and Hypothesis:

S. pneumoniae remains responsible for significant morbidity and mortality, despite the advent of vaccination. The pathogenesis of IPD is not fully understood, although it is clear that IPD is preceded by nasopharyngeal colonisation. As the macrophage is a key component in the innate immune response, it is likely that the host pathogen interaction between *S. pneumoniae* and the macrophage will therefore be crucial in controlling the progression to IPD. It is increasingly recognised that epigenetic mechanisms play key roles in the host pathogen interaction.

I hypothesise that *S. pneumoniae* induces changes in the host's epigenome by altering histone PTMs, which in turn regulates the innate immune responses by modulating gene expression and may explain some of the variation observed in the susceptibility to invasive pneumococcal disease.

As a key modulator of responses, I propose that the bacterial virulence factor pneumolysin may be responsible for some of these epigenetic modifications and in turn influence key effector functions with consequences for the innate immune.

My aims are to:

1) Study the changes in host's epigenome at the level of histone PTMs in response to *S. pneumoniae.*

a) establish a proteomic approach to determine the extent and nature of PTMs in primary human macrophages.

b) determine the contribution of pneumolysin to changes in histone PTMs in following infective challenge of MDMs with *S. pneumoniae* and isogenic pneumolysin deficient mutant.

2) Study the effector consequences of the bacterial virulence factor pneumolysin on the host's immune response.

a) establish the differential gene expression occurring in a pneumolysin dependent manner.

b) establish the changes occurring at the proteomic level in a pneumolysin dependent manner.

3) Use the combined omics approach to determine the relationship between gene expression levels and PTMs of the associated histones.

a) establish the differential transcript expression occurring in response to *S. pneumoniae* challenge using RNA-seq in paired samples.

b) establish the enriched genomic locations following challenge with *S. pneumoniae* using ChIP-seq in paired samples

c) integrate the datasets to infer the role played by histone PTMs in innate immune responses.

Chapter 2: Materials and Methods

Table 2.1: List of consumables and suppliers.

Chemical	Supplier
5 x Gel Loading Dye	New England BioLabs
100 bp DNA ladder	New England BioLabs
2 µm Yellow-Green Fluorescent latex beads	Sigma
4',6-diamidino-2-phenylindole (dapi)	Vectorshield
Acetone	Thermo Fischer Scientific
Acetonitrile (ACN)	Thermo Fischer Scientific
Acrylamide	Thermo Fischer Scientific
Agarose	Sigma-Aldrich
Ammonium bicarbonate	Sigma-Aldrich
Ammonium hydroxide	Sigma-Aldrich
Ammonium sulphate	Sigma-Aldrich
Benzylpenicillin	CrystapenTM, Genus Pharmaceutical
BIX-01294	Sigma-Aldrich
Blue pre-stained standard Broad range protein ladder	New England BioLabs
Bovine serum albumin (BSA)	Sigma-Aldrich
Brain Heart Infusion broth	Oxoid
Bromophenol blue	Sigma-Aldrich
Calf histone standard	Sigma-Aldrich .
CD-CHO media	Cobra Biologics
CHO-S cells	Cobra Biologics
ClarityTM Western ECL	BioRad
Colloidal commassie brilliant blue dye	Sigma-Aldrich
Colombia blood agar (CBA)	Oxoid
Dithiothreitol (DTT)	Sigma-Aldrich
Ethanol	FischerScientific
Ethidium Bromide	Sigma-Aldrich
Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA)	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Fetal calf serum (FCS)	Biosera
Fluorescein isothiocyanate labelled (FITC) goat anti human IgG antibody	Sigma-Aldrich
Ficoll-paque	GE Healthcare Life sciences
Gentamicin	Sanofi
Glacial acetic acid	Thermo Fischer Scientific
Glycerol	Sigma-Aldrich
Glycine	Thermo Fischer Scientific
Heat inactivated low endotoxin newborn calf serum	Gibco, Life technology .

Hepes-KOH	Sigma-Aldrich
Heptafluorobutyric acid (HFBA)	Thermo Fischer Scientific
HPLC grade dihydrogen monoxide	Thermo Fischer Scientific
HT supplement	Thermo Fischer Scientific
Iodoacetic acid (IAA)	Sigma-Aldrich
Isopropanol	Thermo Fischer Scientific
L-Glutamine	Gibco, Life technology
Magnesium chloride (MgCl)	Sigma-Aldrich
Magnetic Protein A	Invitrogen
Magnetic Protein G	Invitrogen
Methanol	VWR Chemicals
Milk	Sigma-Aldrich
N-Laurosyl-sarcosine (NLS)	Sigma-Aldrich
NEBNext® Multiplex Oligos for Illumina (Index primers sets 1-3)	New England BioLabsNew
NEBNext® Ultra™ 2 DNA Library Prep Kit for Illumina®	New England BioLabs
Nonidet P-40 (NP40)	Fissons laboratory reagents
Nuclease free water	Quiagen
Orthophosphoric acid	Thermo Fischer Scientific
Paraformaldehyde	Thermo Fischer Scientific
Phosphate buffered saline (PBS)	Lonza
Polyvinylidene fluoride membrane (PVDF)	Immobilon
Porcine pancreas trypsin	Sigma-Aldrich
Potassium chloride (KCl)	Sigma-Aldrich
Propionic anhydride	Sigma-Aldrich
Proteinase K	Sigma-Aldrich
Puromycin	Gibco
Quantifast SYBR Green	Quiagen
RNAse A	Sigma-Aldrich
Roche Complete EDTA free , Protease inhibitor	Roche
RPMI 1640	Lonza
Saponin	Sigma-Aldrich
Sodium chloride (NaCl)	Thermo Fischer Scientific
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich
Sodium dodecyl sulphate running buffer	National Diagnostics
Sulphuric acid (H_2SO_4)	Thermo Fischer Scientific
Tri Reagent	Sigma-Aldrich ,
Trichloroacetic acid (TCA)	Thermo Fischer Scientific
Triethylammonium bicarbonate buffer (TEAB)	Sigma-Aldrich
Trifluoroacetic acid (TFA)	Thermo Fischer Scientific
Tris–Cl pH 8.0,	Sigma-Aldrich
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)	Sigma-Aldrich
Triton X-100,	ThermoFischer

Tween-20	Sigma-Aldrich
Ultra pure Formaldehyde	Sigma-Aldrich
Ultra-pure Lipopolysacharide (LPS)	Enzo
β-mercaptoethanol	Sigma-Aldrich
Plastics:	
50mL centrifuge tube	Sarstedt
T75 flasks	Corning
24 well plates	Corning
6 well plates	Corning
96 well ELISA plates	Corning
1.5mL microcentrifuge tubes	Sarstetd
2mL lo-bind eppendorfs	Eppendorf
1.5 mL TPX tubes	Diagenode
MicroAmp fast optical 96-well plate	Applied Biosystems

2.1 Growth of bacterial stocks:

Streptococcus pneumoniae serotype 2 strain D39 and the isogenic mutant D39-ΔPLY ((ΔPLY), which has a single amino acid substitution in the pneumolysin (PLY) sequence generating a STOP codon and therefore doesn't express PLY), were kindly obtained from Prof T. Mitchell (University of Brimingham).

All strains were grown in brain heart infusion broth and 20% FCS growth curves were established for all strains using optical density, measured on a spectrophotometer (Jenway 6300) over time, and by plotting the number of colony-forming units (cfu) (by the Miles Misra surface viability count method) plated on Colombia blood agar plates (CBA) at regular time intervals. Stocks were then grown up to mid log phase then frozen at -80°C.

2.1.2 Opsonisation of *Streptococcus pneumoniae*:

Aliquots of bacteria were thawed, washed in PBS twice, the pellet was then re-suspended in RPMI 1640 supplemented with 10% pooled human immune serum from previously vaccinated volunteers with demonstrable antibody levels to serotype 2 pneumococci. The aliquot was incubated for 30 minutes at 37°C in 5% carbon dioxide (CO_2). Then it was centrifuged at 900 relative centrifugal force (rcf) for 3 minutes and the pellet washed in PBS twice prior to re-suspension in RPMI and L-Glutamine and 10% FCS containing low LPS (Dockrell et al., 2001).

2.3 Human monocyte-derived macrophage isolation:

Ethical approval was granted by South Sheffield Regional Ethics committee (07/Q2305/7). Whole blood was obtained from healthy staff and student volunteers at Sheffield Teaching hospitals NHS Foundation Trust following written informed consent. The blood was separated by differential centrifugation using a Ficoll-Paque gradient. Whole blood was layered on Ficoll-Paque in a 2:1 ratio in a 50mL centrifuge tube and centrifuged at 600rcf for 23 minutes. The peripheral blood monocytes in the cloudy layer at the surface were then transferred to a fresh 50 mL tube and washed in

PBS and centrifuged for 10 minutes at 300rcf. The pellet was then washed a second time and then re-suspended in RPMI 1640 medium with 2 mmol/L of L-glutamine and 10% heat inactivated low endotoxin newborn calf serum. The cells were then seeded with or without cover slips at a density of 2.10^6 cells/mL, incubated at 37°C in 5% CO_2. At 24 hours, the cells were washed in PBS and re-suspended in fresh media RPMI 1640 with 2 mmol/L of L-glutamine and 10% heat inactivated FCS. The media was changed every 3-4 days. The monocyte derived macrophages (MDM) were subsequently maintained in RPMI + 10% FCS + 2mmol/L L-glutamine for 13 days.

2.3.1 MDM challenge:

14 day old MDM were washed in PBS, then fresh media (RPMI, 10% FCS and L-Glutamine as above) was replaced. Opsonised D39, ΔPLY bacteria or fresh media (for mock infected) were subsequently added to the cells at a MOI of 10, rested on ice for 1 hour and incubated at 37°C in 5%CO_2 for a further 2 or 3 hours. The MOI was confirmed by measuring the number of cfu following 24hr incubation at 37°C in 5%CO_2 for each aliquot of bacteria. All infections were started at the same time of day (09:00).

2.4 MDM functional assays:

2.4.1 Microscopic assessment of internalisation of *S. pneumoniae*:

Following infection with *S pneumoniae* (with or without pre-stimulation) cells in a 24 well plate seeded on cover slips were washed three times with PBS then 250 μL of 2% paraformaldehyde was added to each well. After at least 30 minutes incubation each well was washed three times in distilled water then the coverslips were incubated with fluorescein isothiocyanate (FITC) labelled goat anti human IgG antibody for 12 minutes. The cover slips were mounted on a glass slide with DAPI and sealed prior to fluorescent microscopy.

2.4.2 *S. pneumoniae* internalisation assay:

At either 3 or 4 hours following infection with *S. pneumoniae,* MDMs in 24 well plates were washed three times in ice cold PBS, then incubated for 30 minutes in media containing RPMI, 10% FCS, 2mmol/L L-glutamine and 40 units/mL of benzylpenicillin and 20 mg/mL gentamicin. The cells were then washed three times in PBS. Cells were then incubated in 250 μL of 2% saponin for 12 minutes at 37°C in 5%CO_2, then 750 μL of PBS was added, followed by vigorous pipetting. The number of internalised viable bacteria were measured by counting the number of cfu on CBA after 24 hours incubation at 37°C in 5%CO_2 contained in these lysates (using the Miles and Misra Method).

2.4.3 Cytokine production measurements:

Supernatants were obtained from cells having been infected at 0, 0.5, 1, 2, 3, and 4 hours, then stored frozen at -80°C prior to analysis.

2.4.4 TNF-α and IL-6 measurements:

Supernatants were analysed as per the manufacturer's guidelines using either Tumour necrosis factor alpha (TNF-α) or interleukin 6 (IL-6) kit (Ready-set-go!™,

eBioscience). Briefly, 96 well ELISA plates were coated with capture antibody (Purified anti-human TNF- α (MAB1) or purified anti-human IL-6) and then left sealed overnight at 4 °C. They were washed four times, in 0.05% Tween-20 and PBS, and the wells blocked in assay diluent at room temperature for 1 hour. The supernatants were then added as were the standards (recombinant human TNF-α or recombinant human IL-6) and incubated for 2 hours at room temperature. The wells were then washed five times and the detection antibody (biotin-conjugated anti-human TNF-α or anti-human IL-6) was added and left to incubate for 1 hour. After washing a further five times avidin-horse radish peroxidase (HRP) was added to each well and left at room temperature for 30 minutes. Plates were then washed a further six times, prior to adding tetramethylbenzidine substrate solution for 15 minutes at room temperature. The reaction was stopped by adding 2M sulphuric acid (H_2SO_4), the plate was then read at 450nm (Multiscan® EX, Thermo Scientific), and the data analysed in GraphPad Prysm version 7.0c. (GraphPad Software).

2.4.5 Flow Cytometry:

Samples were run on a 4 colour FACScalibur™ (BD Biosciences) in the University of Sheffield's core facility using CellQuest Pro (version 6.0). 10 000 gated events were counted and Forward scatter (FSC) and Side scatter (SCC) were used to define populations. Data analysis was performed in FlowJo™ software (version 10.1r5).

2.4.5.1 Flow cytometry confirmation of opsonisation:

Aliquots of 10^6 cfu of serotype 2 D39 *S. pneumoniae* were washed twice in PBS. Then incubated in RPMI with 10% human immune serum for 30 min at 37°C in 5% CO_2. The bacteria were spun down at 900 rcf for 3 min then washed twice in PBS and fixed in 0.5% PFA. The bacteria were then washed once and incubated in 100 µL PBS with 1:100 FITC labelled goat anti human IgG antibody for 20min at room temperature. The bacteria were then washed in PBS and re-suspended in final volume of 300 µL of PBS prior to analysis in the core facility.

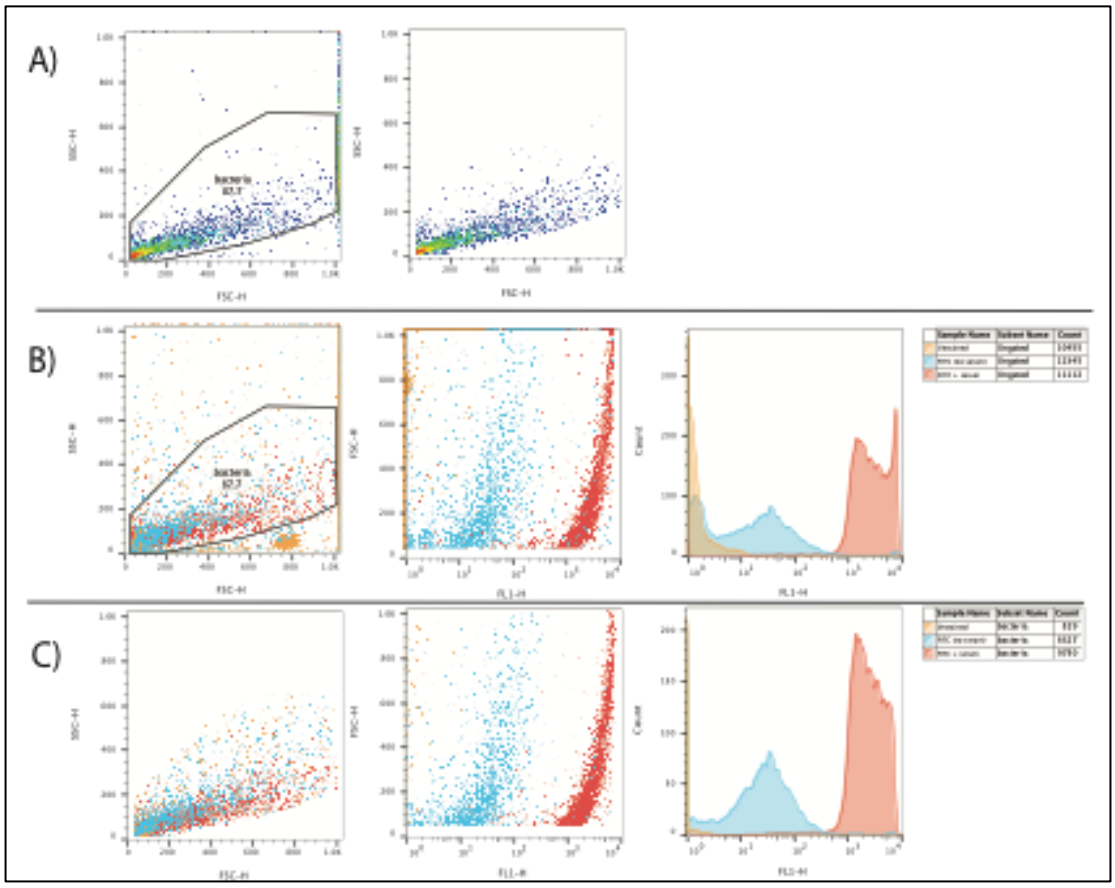


Figure *2*.1: Opsonisation assay.
In order to confirm that the human immune serum was opsonizing the bacteria the following Flow cytometry approach was used. Panels A illustrates the gating strategy used to exclude debris and non-specific FL1 signal. The dot plots and histograms in panels B show the shift in FL1 signal corresponding to the binding of FITC secondary antibody to the bacteria without gating. Panels C how the results of the gated FL1 analysis.

2.5 mRNA analysis:

After 3 hours of infection, cells were washed three times in PBS and lysed in Tri Reagent®, then transferred to 1.5mL microcentrifuge tubes and stored at -80°C.

2.5.1 RNA extraction using Direct-Zol:

Messenger ribonucleic acid (mRNA) extraction was performed following the manufacturer's guidelines for Direct-Zol™ RNA miniPrep (Zymo). Briefly, samples in Tri Reagent® were centrifuged at 12 000 for 1 minute, then the supernatant was transferred to a fresh 1.5 mL tube, 100% ethanol was added in a 1:1 ratio and well mixed. This was then transferred to the Zymo-spin™ column, centrifuged for 1 minute, then washed using Direct-zol™ RNA pre-wash, centrifuged again, then washed in RNA wash buffer and again centrifuged. Finally, the RNA was eluted out in DNase/RNase free water and samples frozen at -80°C.

2.5.2 Microarray mRNA expression analysis:

Using the University of Sheffield's core facility services, Affymetrix chip micro-array (Human Genome U133 plus 2.0 array, Santa Clara, CA) analysis of samples from three individuals was undertaken to further characterise gene expression.

Data analysis was performed in collaboration with Dr Richard Emes at the Advanced Data Analysis Centre (University of Nottingham) using R version 3.4.0. CEL files were read in using Simpleaffy version 2.52.0; background intensity correction, median correction and quantile normalisation was performed using robust multi-array average expression with help of probe sequence (GCRMA), using AffyPLM version 1.52.1. The quality control matrix was generated using Simpleaffy and AffyPLM. Principal component analysis was performed in R. The probes whose intensity were within the lowest 20th centile were removed, using Dplyr version 0.7.1. Differential gene expression was calculated using Limma version 3.32.2.

2.5.3 Pathway analysis:

Pathway analysis was performed from the differentially-expressed gene lists generated. Hypergeometric tests were calculated in R, using *GO.db* version 3.4.1 to search the Gene Ontology (GO) database for molecular function, cellular component, and biological process, using p value cut off of 0.01 and minimum of 3 genes. The Reactome database was searched using ReactomePA version1.20.2 with a q value cut off of 0.01. The Kyoto Encyclopaedia of Genes and Genomes (KEGG) databases was searched using gage version 2.26.1 and visualised using pathview version1.16.1.

2.5.4.1 Next Generation RNA Sequencing:

In addition to microarray mRNA analysis, samples from cells infected with *S.pneumoniae* were also sent for Next Generation Sequencing (NGS). RNA sequencing (RNASeq) via the core facility of the University of Sheffield. Total RNA was checked on a nanodrop and 2100 Bioanalyser RNA nano chip (Agilent) (Figure 2.2). The mRNA was purified using the NEBNext® Poly(A) mRNA Magnetic Isolation module. The sequencing libraries prepared using the NEBNext® Ultra™ 2 Directional RNA kit. Samples indexed using the NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 2) (summarised in Table 2.2). Libraries were checked on a Bioanalyser high sensitivity DNA chip to check average fragment size. (Figure 2.1) Libraries were pooled and run across 2 lanes of an Illumina HiScan™ SQ (Illumina), paired end 2 x100 base pairs (bp).

Table 2.2: RNA-Seq oligo barcode details.

Sample	Condition	Index	Sequence
3/5 D39	*S pneumoniae*	13	AGTCAA
3/5 M	Mock Infected	14	AGTTCC
12/4 D39	*S pneumoniae*	15	ATGTCA
12/4 M	Mock Infected	16	CCGTCC
25/5 D39	*S pneumoniae*	18	GTCCGC
25/5 M	Mock Infected	19	GTGAAA

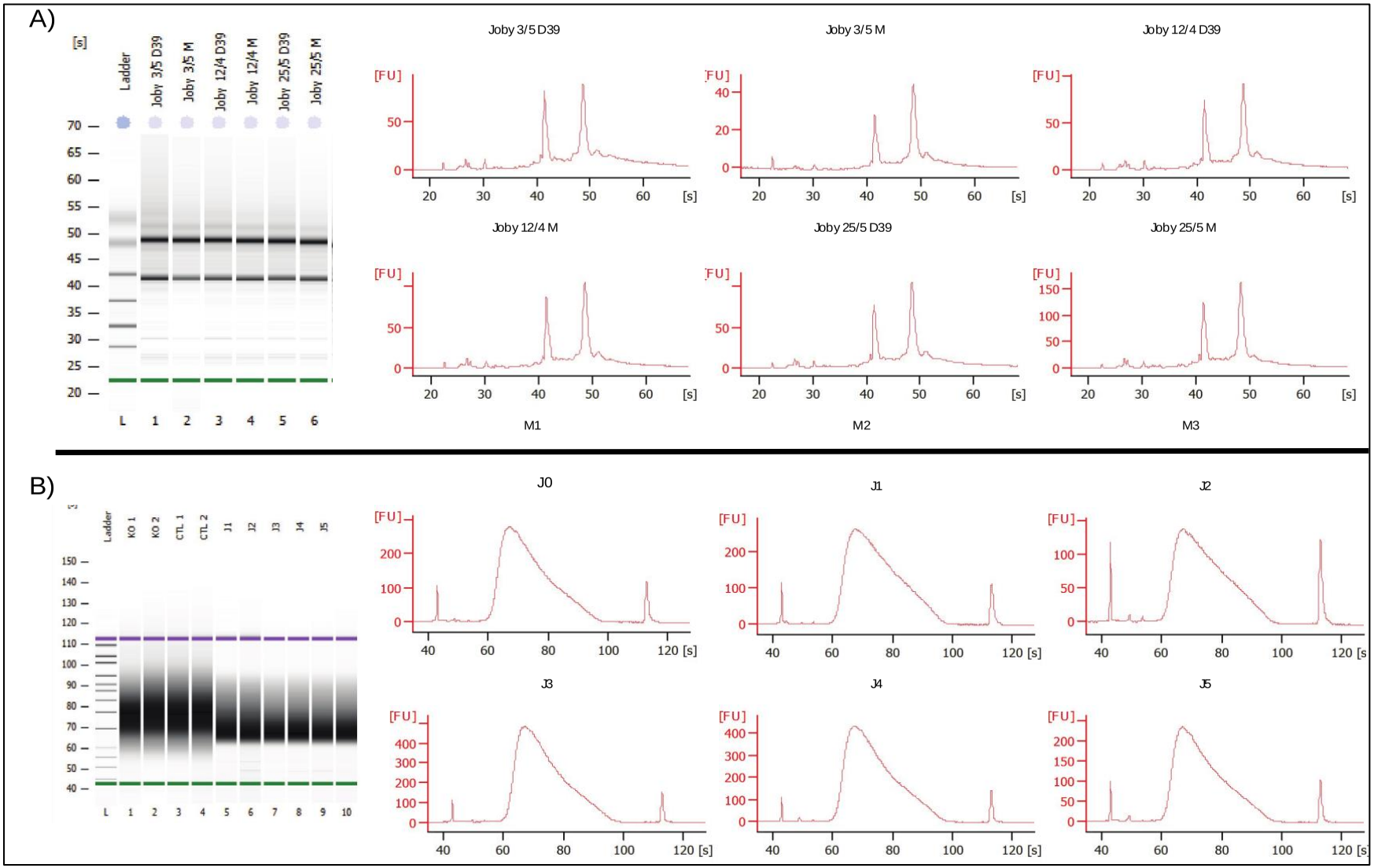


Figure 2.2: Bionanalyser analysis of RNA.

Panel A) This figure illustrates the six RNA samples taken forward for sequencing prior to magnetic bead clean up and library preparation 1 μg of total RNA was used. Panel B) This figure illustrates the results of the library preparation prior to sequencing and confirms good library fragment size.

2.5.4.2 Next Generation RNA-Sequencing data analysis:

RNA-Seq runs were converted to FASTQ files by the core facility. The 6 samples were run as an indexed pool across two lanes, therefore there are 4 FASTQ per sample (2 forward and 2 reverse reads). Quality checks were performed for each sample FASTQ files in FASTQC (version 0.11.5). Data analysis was then performed in GALAXY (version 17.05) following the "new tuxedo" protocol (Pertea et al., 2016), files were concatenated, then alignment was performed using HISAT2 (Galaxy version 2.0.5.2)(Kim et al., 2015) against the HG38 human genome. The transcript assembly and quantification was performed in Stringtie (Galaxy version 1.3.3)(Pertea et al., 2015) using annotations from Hg38 GENCODE v24 downloaded from University of California Santa Cruz (11/08/2017) and NCBIrefseq (genome) downloaded (20/09/2017). Differential transcript expression was then performed in R using Ballgown package (version 2.8.4)(Pertea et al., 2016). In addition, quasi-mapping was performed using Salmon and the UCSC NCBIrefseq transcriptome followed by differential expression analysis carried out using edgeR (Robinson et al., 2010).

2.6 Cell preparation for mass spectrometry:

In order to make the best use of primary cells these were not used for the optimisation of mass spectrometry methods. Instead, a Chinese Hamster Ovary (CHO) cell line was used. CHO-S cells were grown in CD-CHO supplemented with 8mM L-glutamine, 0.012mg/ml puromycin and HT supplement media for either 2 or 4 days then washed in PBS and pelleted into $4x10^6$ cell aliquots before being frozen -20 °C.

For infections and pre-stimulation experiments, MDMs were washed three times in ice cold PBS and scrapped in ice cold PBS with added protease inhibitors ((PI) Roche Complete EDTA free), before being pelleted at 900rcf for 10 minutes then frozen at -80°C prior to preparation for mass spectrometry.

2.6.1 Histone extractions and purification:

All steps were done at 4°C in lo-bind Eppendorf's. Cell pellets were lysed in histone extraction hypotonic lysis solution, left to rotate for 30 minutes, then pelleted at 10 000 rcf for 10 minutes. The supernatant was discarded, the pellet re-suspended in 0.2 M H_2SO_4 and rotated for 4 hours. The tubes were then centrifuged at 16 000 rcf for 10 minutes to remove debris. The supernatant was transferred to a new lo-bind tube and the histones were precipitated out by adding TCA drop wise to a final concentration of 33%. The samples were then left to stand overnight. They were then centrifuged at 16 000 rcf for 10 minutes, the supernatant removed, and the histone smear washed in acetone twice. Finally, the histones were re-suspended in 100 μL of high performance liquid chromatography (HPLC) grade dihydrogen monoxide (H_2O) and transferred to a fresh tube prior to freezing at -80°C.

Table 2.3: Histone extraction hypotonic lysis solution.

Histone extraction hypotonic lysis solution:	
10 mM	Tris–Cl pH 8.0,
1 mM	KCl
1.5 mM	MgCl
1 mM	DTT

1 tablet / 10 mL	Roche complete EDTA free Protease inhibitor tablet

2.6.2 Protein purity and concentration estimation:

5 µl aliquots of purified histones were run on 12% sodium dodecyl sulphate (SDS) polyacrylamide gel. Briefly, 5 µl of H_2O was added to 5 µl of samples or to the calf histone standard. This was then diluted in Loading buffer at a 1 to 1 ratio and denatured at 95°C for 5 minutes. The samples were then loaded on the 0.75mm 12% SDS gel in SDS running buffer along with a 5µL of Blue pre-stained standard broad range ladder. The current was set at 200V for 45 minutes or until the dye had reached the bottom of the gel. The gel was then stained with Commassie stain and agitated at room temperature overnight. The following morning the gel was de-stained by washing in distilled H_2O.

Table 2.4: Western blot protein Loading buffer.

Western blot protein loading buffer	
50 mM	Tris-HCl
0.10%	bromophenol blue,
2%	SDS
100 mM	β-mercaptoethanol
10%	Glycerol

Table 2.5: Commassie stain.

Commassie stain	
1.60%	Orthophosphoric acid
20%	Ethanol
8%	Ammonium sulphate
0.08%	Colloidal Commassie brilliant blue dye

Table 2.6: 12% SDS page gel recipe.

	Stacking gel	Resolving Gel
Distilled Water	3.66 mL	4.38 mL
40% Acrylamide (Fischer Scientific)	750 µL	3 mL
4 x Upper Buffer (0.5MTris pH6.2 0.4% SDS)	1.5 mL	
4 x Lower Buffer (1.5M Tris pH 8 0.4% SDS)		2.5 mL
10 % Ammonium persulfate (Sigma)	75 µL	100µL
Tetramethylethylenediamine (TEMED)	15 µL	20µL

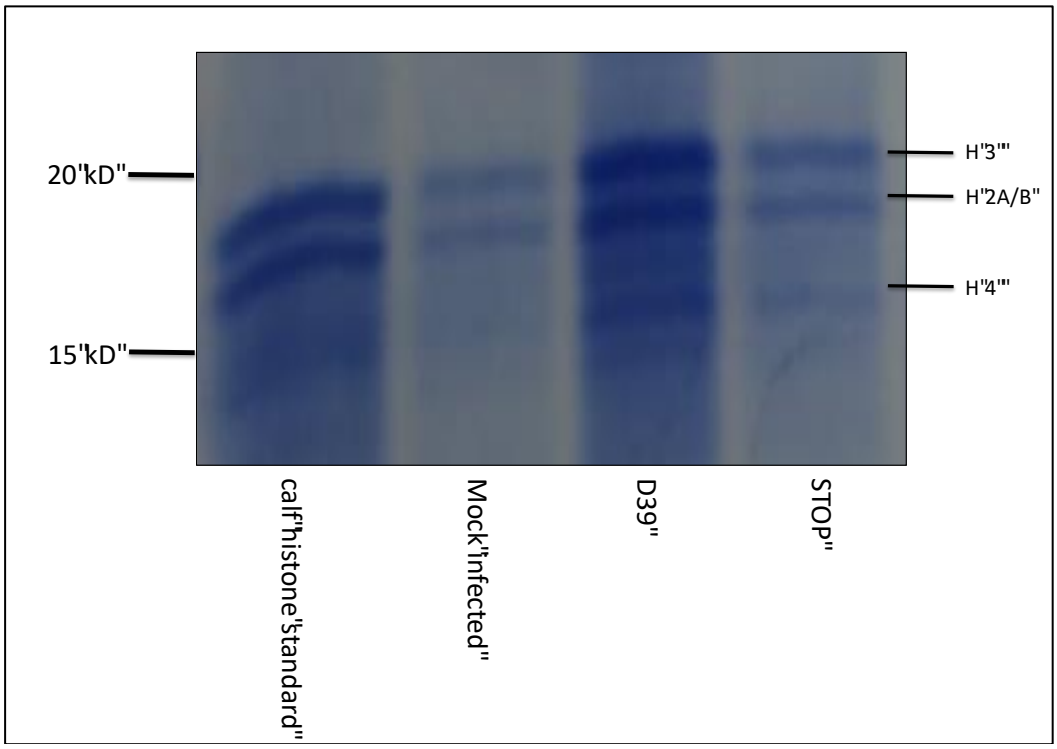


Figure 2.3: Commassie stained 12% SDS PAGE analysis of acid extracted histones. Following acid extraction from three samples, 5 μL aliquots were loaded along with 5 μg of calf histone standard and a protein molecular weight marker as indicated. The histone H3, H2A/B and H4 are highlighted. There is some variability in the quantities of histone isolated as is illustrated by the variability in the intensity of the bands when compared to the calf histone standard.

2.6.3 Chemical derivitization and trypsin digestion:

In order to facilitate the identification of isobaric peptides that ordinarily would coelute, peptides were chemically derivitised to increase their hydrophobicity by the addition of proprionyl groups (Garcia et al., 2007).

10 μL of 100 mM ammonium bicarbonate pH 8 and 4 μl of ammonium hydroxide was added to 10 μg of histone sample. Then 10μL of propionic anhydride in isopropanol (1:3 ratio) was added and ammonium hydroxide used to keep the pH >8.0. The sample was incubated at 37°C for 15 minutes. Then it was dried down in a vacuum centrifuge (Concentrator plus, Eppendorf) and the process repeated. The samples were re-suspended in 40 μL of 100mM ammonium bicarbonate and then tryptically digested overnight. The digestion was stopped by addition of glacial acetic acid and freezing at -80°C for 5 minutes. Finally, the samples were dried down before undergoing a further two rounds of proprionylation.

2.6.4 Fractionation of samples using porous graphitic column:

To aid the identification of co-eluting peptides, samples were fractionated (Minshull and Dickman, 2015) using a porous graphitic column (PGC, Hypercarb™ (50 x 2.1 mm), Thermo Scientific). Samples were re-suspended in 0.1% Trifluoroacetic acid (TFA), and fractionated using an Ultimate 3000® (Dionex, Thermo Scientific) at 30°C with a flow rate of 0.2 mL/min. Fractions were collected from 8 minutes until 32 minutes, using a stepwise gradient as described below. The individually collected

fractions were then dried down in SpeedVac and frozen to -80°C prior to Mass spectrometry analysis. The ultraviolet (UV) detection was set at 214 nm (50Hz) throughout the fractionation process.

Table 2.7: Stepwise gradients for fractionation of proteins on HPLC.

HPLC fractionation of histones:

Time: (min)	Buffer A_1	Buffer B_1
0	100%	0%
25	30%	70%
30	10%	90%
39	95%	5%
45	100%	0%

Buffer A_1 3% ACN 0.1% TFA
Buffer B_1 90% ACN 0.1% TFA

HPLC fractionation of Proteomes:

Time: (min)	Buffer A_1	Buffer B_1
0	98%	2%
110	40%	60%
115	10%	90%
116	98%	2%
125	98%	2%

2.6.5 Hypersep™ Hypercarb tip clean up:

An alternative approach to fractionating samples is to use a Hypersep™ Hypercarb tip to clean the samples of the chemical derivitisation residues. This approach was termed "one dimension" (1D) sample (Minshull and Dickman, 2015). These samples were prepared for mass spectrometry on Orbitrap QE HF, by re-suspending them in 20 µL 0.1% TFA and 3% HPLC grade ACN. Then following the manufacturer's protocol, the Hypersep Hypercarb ™ (ThermoScientific) tips were primed by 3 washes (20 µL each) with elution solution (60% Acetonitrile 0.1% TFA), followed by 5 washes (20 µL each) in cleaning solution (0.1% TFA). The sample was then re-suspended on the tip by pipetting up and down 50 times. The tip was cleaned in 3 x 20 µL of 0.1%TFA, then in a fresh low-bind tube, the samples were eluted using 5 x 20 µL of elution solution 1 (60% ACN 0.1% TFA) and pipetted up and down 10 times, then 5 x 20 µL of elution solution 2 (90% ACN 0.1% TFA). The samples were then dried down in Speedvac and frozen at -80°C prior to mass spectrometry (MS) analysis.

2.6.6 Histone post-translational mass spectrometry: maXis (Brucker):

The samples were re-suspended in 0.1% TFA and analysed using reverse phase high performance liquid capillary chromatography (Ultimate 3000,Thermo Scientific) on a C18 column (150 mm x 75 µm internal diameter, PepMap reversed phase column (Dionex UK)) interfaced with an Ultra-High Resolution Time of Flight mass spectrometer (maXis, Bruker Daltonics), using a captive Electrospray ionisation (ESI), on a stepwise 60 minute gradient (as illustrated below). MS and MS/MS scans were acquired in positive ion mode (m/z 100-1800).

Data was then processed using automatic internal calibration (lock mass 1221.99) and analysed in DataAnalysis v4.1 (Brucker Daltronics). The mascot generic files (mgf) were created in DataAnalysis, the sum peak finder was used with signal to noise ratio of 100 and minimum intensity of 5000. These files were used for Mascot searches performed in Mascot Daemon v2.5.1 (Matrix science).

The following search parameters were used: MS and MS/MS tolerance was set at 0.15 Da, searched against the SwissProt database, using homo sapiens taxonomy; the fixed modifications were set as propionylation of lysines (K) and N terminus; the

variable modifications were defined as: methylpropionylation, dimethylation, trimethylation and acetylation of lysines, phosphorylation of serines and threonines, the digestion enzyme was set as Arg-C with up to 2 missed cleavages, and charge states set at 2+, 3+ and 4+.

After manual verification, quantification was carried out by the integration of smoothed (Gauss method) extracted ion chromatograms and correction factors (to account for the differences in efficiency in ionisation, were applied to the results (Lin et al., 2014) and the process automated in Hist-o-matic (in house software courtesy of T Minshull).

Table 2.8: Stepwise gradients for mass spectrometry.

Gradient for the maXis:		
Time: (min)	**Buffer A_2**	**Buffer B_2**
0	**97%**	**3%**
30	**75%**	**25%**
40	**15%**	**50%**
41	**10%**	**90%**
48	**97%**	**3%**
55	**97%**	**3%**

Buffer A_2: 3% ACN 0.1% FA
Buffer B_2: 80% ACN 0.1% FA

Gradient for the Orbitrap QExactive HF:		
Time: (min)	**Buffer A_2**	**Buffer B_2**
0	**97%**	**3%**
8	**90%**	**8%**
51	**75%**	**25%**
81	**40%**	**60%**
82	**10%**	**90%**
86	**10%**	**90%**
87	**97%**	**3%**
101	**97%**	**3%**

2.6.7.1 Histone posttranslational mass spectrometry: Orbitrap QE HF:

The samples were re-suspended in 0.1% TFA (or 0.05% heptafluorobutyric acid (HFBA), loaded into and run on Ultimate 3000 online capillary liquid chromatography system with PepMap300 C18 trapping column (Thermo Fischer), coupled to Orbitrap Q-Exactive HF ((QE), Thermo Fischer). Peptides were eluted onto a 50 cm x 75 μm Easy-spray PepMap C18 column with a flow rate of 300 nL/min. Peptides were eluted using a stepped gradient (as described in Table 2.8).

Data acquisition was performed in a number of different modes (as described in Table 2.10). Briefly, Data Dependent Acquisition (DDA) was performed in full scan positive mode, scanning 375 to 1500m/z, with an MS1 resolution of 120 000, and Automatic Gain Control (AGC) target of $1x10^6$ and a maximum fill time of 450ms. The top 10 most intense ions from MS1 scan were selected for collision induce dissociation (CID). MS2 resolution was set at either 120 000 (DDA120), 60 000 (DDA 60),) or 30 000 (DDA30) with AGC target of 1x105 and maximum fill time of 450, 220 and 100ms respectively, with isolation window of 2m/z and scan range of 200-2000 m/z, normalised collision energy 27 (NCE).

Data Independent Acquisition (DIA) was performed in three different settings. Firstly, DIA60 (DIA60) had a full scan at a resolution 60 000, AGC target of $3x10^6$, maximum fill time of 55ms, scanning range of 300 to 900 m/z. followed by 10 DIA windows at a resolution of 30 000, AGC target of $1x10^6$, isolation windows of 20m/z and NCE of 26 for DIA60. DIA30 (DIA30) had full scan resolution of 30 000, AGC target of $3x10^6$, maximum fill time 100ms, scanning range of 300 to 900m/z; followed by 10 DIA windows at a resolution of 15 000, AGC target $1x10^6$, isolation windows of 20 m/z, NCE 26. For DIA60 and DIA30 the isolation lists were the same (see Table 2.9).

Finally, DIA variable window (DIAvw) had full scan resolution of 30 000, AGC target of $3x10^6$, maximum fill time 100ms, scanning range of 300 to 900m/z; followed by 85 DIA windows at a resolution of 15 000, AGC target $1x10^6$, maximum fill time 115ms, with an isolation window scheme which varied to resemble SWATH™ (ABSciex), the variable isolation windows are summarised in Table 2.11, and NCE 26.

The DIA method used for the study of MDMs was identical to DIA60 except that the scanning range was adjusted to 300 to 1100 m/z.

Table 2.9: Isolation lists.

DIA60/30 isolation list	DIAvw isolation list		
310.390973	292	580	730
330.400068	359	585	735
350.409163	378	590	740
370.418258	397	595	745
390.427353	416	600	750
410.436448	432.5	605	755
430.445543	446.5	610	760
450.454638	458	615	765
470.463733	465	620	770
490.472828	470	625	775
510.481923	475	630	780
530.491018	480	635	785
550.500113	485	640	790
570.509208	490	645	795
590.518303	495	650	800
610.527398	500	655	805
630.536493	505	660	812
650.545588	510	665	823.5
670.554683	515	670	840
690.563778	522.5	675	859
710.572873	530	680	880.5
730.581968	535	685	904.5
750.591063	540	690	928.5
770.600158	545	695	965
790.609253	550	700	1014.5
810.618348	555	705	
830.627443	560	710	
850.636538	565	715	
870.645633	570	720	
890.654728	575	725	

Table 2.10: Characteristics of data acquisition methods.

		Data-Dependent Acquisition			Data-Independent Acquisition		
		DDA 120	DDA 60	DDA 30	DIA 60	DIA 30	DIA vw
MS1	Resolution	120 000	60 000	30 000	60 000	30 000	30 000
	AGC	1.00E+06	1.00E+06	1.00E+06	3.00E+06	3.00E+06	3.00E+06
	Fill time (ms)	450	450	450	55	100	100
	Scan range (m/z)	375-1500	375-1500	375-1500	300-900*	300-900	300-900
MS2	Resolution	120 000	60 000	30 000	30 000	15 000	15 000
	AGC	1.00E+05	1.00E+05	1.00E+05	1.00E+06	1.00E+06	1.00E+06
	Fill time (ms)	450	220	100	Automatic	Automatic	115
	loop count	10	10	10	10	10	85
	isolation window	2	2	2	20m/z	20m/z	variable
	NCE	27	27	27	26	26	26
	scan range (m/z)				300-900	300-900	300-1100

Table 2.11: DIAvw isolation window size and loop count.

DIA vw isolation list	
isolation window (m/z)	loop count
115	1
19	4
26	2
9	1
5	11
10	1
5	56
9	1
14	1
19	1
24	3
49	1
50	1

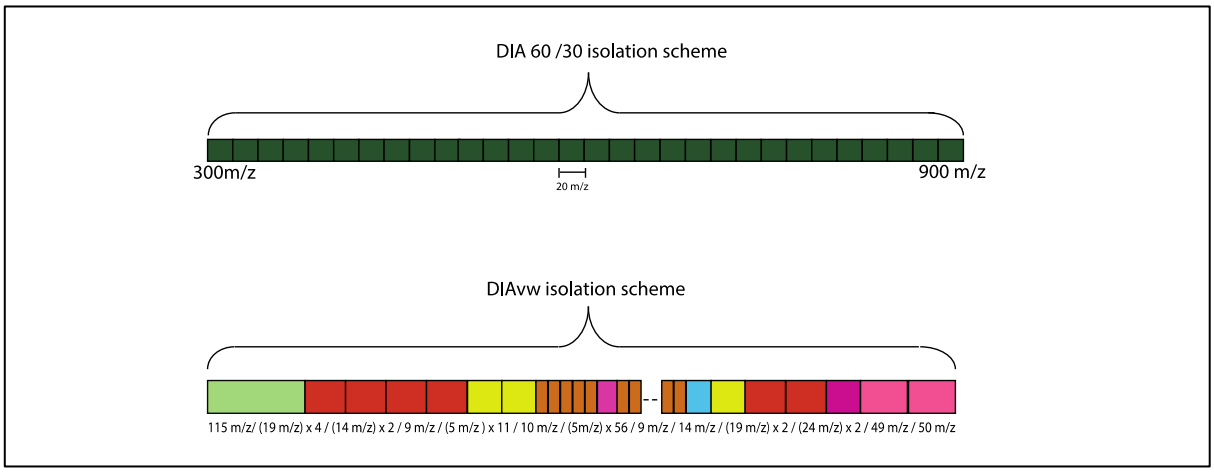


Figure *2*.4: Data independent acquisition schemes.
This figure shows the two isolation window schemes use in panel a) the DIA 60 or DIA30 20m/z isolation windows from 300 to 900m/z. panel b) shows the isolation windows for the DIA variable window method.

2.6.7.2 Analysis of histone PTMs using QExactive HF:

RAW files were converted to mgf using MSConvert (proteowizard) for DDA runs. Searches were performed using Mascot Daemon 2.5.0 (using CHO Uniprot 10029 (downloaded 07/06/2017) for CHO cells searches, or Histone Bespoke database (all of the Human Histone proteins downloaded from Uniprot 31/03/2016). The following search parameters were used: peptide tolerance 10 ppm, MSMS tolerance 0.01 Daltons (Da); digestion enzyme was Arg-C with no missed cleavages; peptide charges of 2, 3 and 4+; fixed modifications (propionyl (K) and propionyl (N-term)) and variable modifications (acetyl (K), methylpropionyl (K), dimethyl (K) and trimethyl (K)), false discovery rate (FDR) was set to less than 2%.

Relative abundance quantification was performed in conjunction with Skyline (Schilling et al., 2012) to extract the area for each peptide proteoform. The relative abundance of histone peptides was determined by the integration of the extracted ion chromatograms to determine the area under the curve (AUC). This value was then used to determine relative abundances for each peptide, by dividing the AUC of one proteoform of a peptide by the sum total of the AUC of all forms of that peptide. For DIA PTM identification was performed in Skyline (using prior knowledge of elution profile, isotope dot Product > 0.90 and <5 parts per million (ppm) (Schilling et al., 2012). Relative abundance of Histone PTMs and identification was also determined using Epiprofile 2.0 (Yuan et al., 2015).

2.6.8 Label-free proteome analysis:

In order to prepare protein samples for shotgun proteomic analysis, 14 day old MDMs were challenged with opsonised *S. pneumoniae* D39 or ΔPLY for either 3 or 6 hours prior to being washed three times in PBS, scrapped in PBS and Roche Complete EDTA free Protease inhibitor and then centrifuged 900 rcf for 10min. The cell pellet was then frozen at -80 °C.

2.6.8.1 Cell lysis:

Cell pellets were re-suspended in FASP Lysis buffer vortexed and then sonicated in water bath sonicator (Fisherbrand) for 10 seconds, then boiled for 5 minutes at 95 °C. The lysates were clarified by centrifuging at 16 000 rcf for 5 minutes.

Table 2.12: FASP lysis buffer.

FASP protein lysis buffer	
4%	SDS
100mM	TEAB
0.1M	TCEP

2.6.8.2 Protein concentration estimation:

In order to estimate the protein concentration in the lysates, a BioRad RC DC™ protein assay (Reducing agent and detergent compatible protein assay, BioRad) was performed as per the manufacturer's instructions. Briefly, 5 µl of DC Reagent S was added to each 250 µl of DC Reagent A. Serial dilutions of bovine serum albumin (BSA) to form the protein standard were performed (from 0.2 mg/ml to 1.5 mg/ml protein). 2.5 µL of lysates were diluted in 22.5 µL of HPLC grade H_2O. Then 125 µl RC Reagent I was added to each tube, vortexed, and incubated for 1 minute at room temperature. Then 125 µl RC Reagent 2 was added into each tube, vortexed then centrifuged (15 000 rcf for 5 minutes). The supernatant was discarded and the tube dried. 127 µl of Reagent A´ was added to the tubes, then vortexed and incubated at room temperature for 5 minutes.

Finally, 1 ml of DC Reagent B was added to each tube and vortexed then incubated at room temperature for 15 minutes, prior to the absorbance being read on a spectrophotometer at 750nm.

2.6.8.3 Filter aided separation of protein:

In order to facilitate the preparation of the protein lysates prior to MS analysis, the samples were processed as per the Filter aided separation protocol (FASP)(Wiśniewski et al., 2009). 100 µg of protein lysates were mixed with 8M Urea dissolved in triethylammonium bicarbonate (TEAB) and added to the filter mounted in low-bind eppendorf. The tubes were centrifuged at 14 000 rcf for 30 minutes, the flow through discarded and further two 200 µL wash performed.

Then 100 µL Iodoacetic acid (92 µg in 8M Urea) was added to each filter to alkylated the proteins. The membranes were then incubated in the dark at room temperature for 20 minutes. Then centrifuged at 14 000 rcf for 30min. The flow-through was discarded. Then the membrane was washed three times with 100 µL of 8M urea, followed by 15 minute centrifugation at 14 000 rcf. The membranes were then washed three times in 100 µL of 100 mM TEAB, followed by 20 minute centrifugation at 14 000 rcf.

The lysates were then subjected to trypsin (2 µg, porcine pancreas) digestion overnight at 37 °C.

The following morning the filters were transferred to fresh low-bind tubes and the digested proteins were eluted in 40 µL of TEAB by centrifuging 14 000 rcf for 10 minutes repeated once, and finally 40 µL of 0.5M sodium chloride was used to elute the final peptides from the filter. The peptides were then frozen at -20 °C.

2.6.8.4 Commassie stain of protein digestion:

In order to confirm that the FASP preparation had indeed resulted in digestion of the proteins, an aliquot was run on a 12%SDS page and then post stained with Commassie (as described 2.6.2).

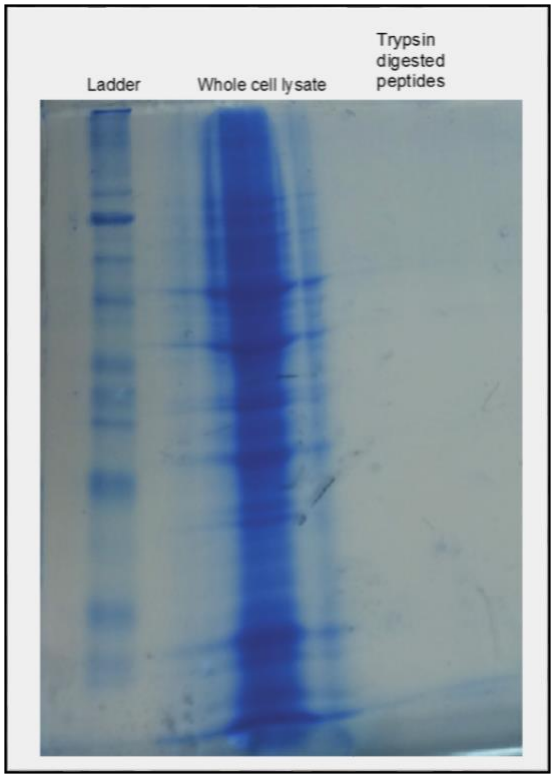


Figure 2.5: Commassie stained protein lysates.
12% SDS gel with broad range protein ladder in lane 1, whole cell lysate in lane 2, then post trypsin digestion and FASP separation in lane 3. There are no discernible proteins suggesting a complete digestion occurred.

2.6.8.4 Hypercarb tip clean up:

In order to desalt the peptides prior to mass spectrometry analysis, they were re-suspended in 100 µL of 0.1%TFA 3% ACN. 1 µL was then further diluted up to 20 µL prior to being bound to the Hypersep™ Hypercarb tip (as described in 2.6.5) The desalted peptides were then dried down using SpeedVac and frozen at -20 °C.

2.6.8.5 Hypercarb fractionation of FASP prepared peptides:

In order to reduce the complexity of the peptide mixture, samples were fractionated using a Hypercarb™ HPLC Column (100 x 2.1 mm, ThermoScientific). Samples were re-suspended in 0.1% TFA and 3% ACN, and fractionated using an Ultimate 3000® (Dionex, Thermo Scientific) at 30°C with a flow rate of 0.2 mL/min. 3 minute fractions were collected from 15 minutes until 90 minutes, using a stepwise gradient as described below (Table 2.9) resulting in 25 fractions. The individually collected fractions were then dried down in SpeedVac and frozen at -80°C prior to Mass spectrometry analysis. The UV detection was set at 214 nm (50Hz) throughout the fractionation process.

Prior to mass spectrometry, the fractions were pooled as described in Table 2.13. The first few fractions and the last were not run as they have previously been shown not to contain any peptides of interest. The pooled fractions were dried and re-suspended in 3% ACN 0.1% TFA.

Table 2.13: Pooling of FASP fractions.

Time fraction collected (min)	Pooled fraction:	Time fraction collected (min)	Pooled fraction:
16-18		51-53	E
19-20		54-56	E
21-23	A	57-59	E
24-25	A	60-62	E
26-28	B	63-65	C
29-31	B	66-68	C
32-34	C	69-71	B
35-37	C	72-74	B
38-40	D	75-77	A
41-43	D	78-80	A
44-46	D	81-83	
48-50	D	84-86	
		87-89	

2.6.8.6 Shot gun proteomic analysis of label-free proteome:

The trypsin-digested pooled fractions were re-suspended in 0.1%TFA and 3% ACN and loaded into and run on Ultimate 3000 online capillary liquid chromatography

system with PepMap300 C18 trapping column (Thermo Fischer), coupled to Orbitrap Q-Exactive HF (Thermo Fischer). Peptides were eluted onto a 50 cm x 75 μm Easy-spray PepMap C18 column with a flow rate of 300 nL/min. Peptides were eluted using a gradient of 3% to 35% over 75 minutes. Data acquisition was performed in full scan positive mode, scanning 375 to 1500m/z, with an MS1 resolution of 120 000, and AGC target of $1x10^6$. The top 10 most intense ions from MS1 scan were selected for CID. MS2 resolution was of 30 000 with AGC target of $1x10^5$ and maximum fill time of 60ms, with isolation window of 2m/z and scan range of 200-2000 m/z, normalised collision energy 27.

2.6.8.7 MaxQuant data analysis:

The raw data from the MS runs were analysed in MaxQuant (version 1.5.6.5). The search settings were as follows: trypsin/P digestion, with up to 2 missed cleavages, fixed modification was Carbamidomethyl (C), variable modifications were oxidation (M) and acetylation (Protein N-term), Label Free Quantification (LFQ) was performed with a minimum neighbours of 3 and average number of neighbour of 6. Peptide tolerance was set at 4.5ppm and minimum peptide length of 7 Amino acid (AA), maximum peptide mass of 4600Da, Protein FDR was set at 0.01.

2.6.8.8 Statistical analysis:

Downstream analysis was performed in R version 3.4.0. The protein identification files were read in. Initial filtering of likely false identifications was performed by excluding results matched to a reverse sequence database entitled "Reverse", then those matching contaminant database search were also removed. Finally, only the identifications with at least 2 unique peptides (peptides unique to one protein thereby reducing the ambiguity) were taken forward to the rest of the analysis. The label-free intensities were then median-corrected for each sample and log2 transformed. Differential protein expression was calculated using Limma version 3.32.2. Principle component analysis was performed to assess the impact of biological variations.

2.7 Chromatin immunoprecipitation:

2.7.1 Cross linking cells for ChIP-Seq:

14 day old MDM following challenge with opsonised D39 or mock infected were washed in PBS at 3 or 4 hours then fixed in 1% ultra pure formaldehyde in 1640 RPMI and 10% FCS and 2mmol L-glutamine for 10 minutes at room temperature with gentle agitation then quenched by the addition of 0.125 M glycine for 5 minutes at room temperature with gentle agitation. The cells were then washed in PBS x 3 and scraped, then pelleted at 900 rcf for 10 minutes prior to freezing at -80°C for storage prior to chromatin immunoprecipitation (ChIP).

2.7.2 Chromatin purification:

Table 2.14: Chromatin immunoprecipitation lysis buffers.

ChIP lysis buffer 1	
50mM	Hepes-KOH pH7.5
140mM	NaCl
1mM	EDTA
10%	Glycerol
0.50%	Nonidet P-40
0.25%	Triton X-100
1 tablet / 10 mL	Roche complete EDTA free Protease inhibitor tablet
ChIP lysis buffer 2	
200mM	NaCl
1mM	EDTA
0.5mM	EGTA
5mM	Tris pH 8
1 tablet / 10 mL	Roche complete EDTA free Protease inhibitor tablet
ChIP lysis buffer 3	
1mM	EDTA
5mM	Tris pH 8
0.5mM	EGTA
0.50%	N-Laurosyl-sarcosine
1 tablet / 10 mL	Roche complete EDTA free Protease inhibitor tablet

The cross-linked cell pellet was re-suspended in 1mL of Lysis buffer 1, rocked at 4°C for 30 minutes, then centrifuged for 5 minutes at 4°C 900 rcf. The pellet was then re-suspended in 1 mL of lysis buffer 2, rocked for 30 minutes at room temperature. Then it was centrifuged for 10 minutes at 4°C 900 rcf. Finally it was re-suspended in 300 μL of lysis buffer 3 and rested on ice for 10 minutes prior to sonication.

2.7.3 Sonication:

Samples were transferred to 1.5mL TPX tubes and sonicated 4 x 10 minutes + 1 x 5 minutes on high, 30 seconds on 30 seconds off, in ice cold water on a Bioruptor® (Diagenode). The samples were then centrifuged at 14000 rcf for 10 minutes at 4°C. The sheared chromatin was then divided into aliquots and frozen at -80°C.

2.7.3.1 Sonication efficiency evaluation:

30 μl of sample was diluted to up to 200 μl with water. Then it was incubated at 95ºC for 20 minutes for a quick de-crosslinking. This was then incubated for 10 minutes at 37ºC with 0.5 mg/ml RNase A and then this was followed by 30 minutes incubation at 50ºC with 0.5 mg/ml Proteinase K.

2.7.3.2 MinElute® PCR purification kit (Quiagen):

After reversal of the crosslinking the DNA was cleaned on Quiagen MinElute® PCR purification kit as per manufacturer's protocol. 1 mL of PB buffer was added to 200 µl of sample and 40 µl 3M sodium acetate pH5.2. This was added to the MinElute column and centrifuged at 17 900 rcf for 1 minute. The flow-through was discarded. The column was washed in 700 µl of PE buffer and centrifuged for 1 minute, the flow-through discarded and then re-centrifuged and the column transferred to a new low-bind tube. 10 µl of EB buffer was added to the centre of the column and centrifuged for 1 minute then a further 10µl of EB buffer was added and the centrifugation repeated.

2.7.3.3 Nanodrop measurement:

The Optical Density (OD) of 2µl of the eluted DNA was measured on a NanoDrop 2000 Spectrophotometer (Thermo Scientific) to estimate the concentration.

2.7.3.4 1% Agarose gels:

Table 2.15: TAE buffer.

TAE buffer	
40mM	Tris pH8.0
20mM	Acetic acid
1mM	EDTA

Sonication evaluation was performed by running the remaining eluted DNA (18µl) on a 1% agarose gel. Briefly, 1g of Agarose was dissolved in TAE buffer by heating. It was then allowed to cool prior to addition of 5µl Ethidium Bromide. The 4µl of 5 x Gel Loading Dye was added to each sample prior to loading on the gel and running for 45 minutes at 80V or until the dye had reached the end of the gel. In addition, a 100bp DNA ladder was also run.

2.7.4 Chromatin immunoprecipitation:

Table 2.16: Chromatin immunoprecipitation wash and elution buffers.

Chip Dilution buffer	
0.01%	SDS,
1.10%	Triton X-100
1.2 mM	EDTA
16,7 mM	Tris-HCl pH 8.0
167mM	NaCl,
1 tablet / 10 mL	Roche complete EDTA free Protease inhibitor tablet
ChIP wash buffer A	
50mM	Tris pH8
150mM	NaCl
1mM	EDTA
0.10%	SDS

1%	NP40
0.50%	Deoxycholate
ChIP wash buffer B	
1%	NP40
50mM	Tris pH8
500mM	NaCl
1mM	EDTA
0.10%	SDS
0.50%	Deoxycholate
ChIP wash buffer C	
50mM	Tris pH8
250mM	Lithium Chloride
1mM	EDTA
1%	NP40
0.50%	Deoxycholate
Elution buffer	
1%	SDS
0.1M	Sodium bicarbonate

50µl of sheared chromatin was used for each pull down. It was diluted up to 300 µl with Chip Dilution buffer. The appropriate amount of histone antibody (see Table 2.17) was added (or not, for beads only pull downs) and incubated with rotation overnight at 4ºC. The next day 30µl of Magnetic Protein A and Protein G beads (50:50 mix, Invitrogen) was added to each immunoprecipitation and rotated in the cold room for 4 hours. The beads were then washed twice with 500 µl of wash buffer A. The samples were rotated for 1 minute, before being put back on magnetic rack between washes. Then they were washed once with 500 µl of wash buffer B. They were then rotated for 1 minute and put back on the magnetic rack.

Finally, the samples were washed once with 500 µl of wash buffer C, rotated for 1 minute and put back on the magnetic rack. Following the washes, the DNA was eluted from the beads in 200 L of Elution buffer, by incubating on a Thermomixer (Eppendorf) for 40 minutes at 65ºC, shaking at 600 rpm.

The 10% input samples were processed from this point by adding Elution buffer up to 200µl.

The samples were then incubated at 37ºC for 15 minutes with 5µl RNAase A (10mg/ml). Then 2 µL of Proteinase K (20 mg/ml) and 20 µl of NaCl (5M) were added and it was incubated overnight on a Thermomixer at 65 ºC shaken at 600rpm. The following day the samples were processed cleaned on MinElute PCR purification kit as described in 2.12.2.2 MinElute® PCR purification kit (Quiagen). This was then frozen at -20 ºC prior to NGS or quantitative polymerase chain reaction (qPCR).

Table 2.17: Antibodies for pull downs.

Antibody	Quantity (µg)	Suplier	Lot number
Total H3	2	Abcam ab1791	GR103804-2
H3K4me1	2.5	Abcam ab8895	GR312093-1
H3K4me3	3	Active Motif 39159	12613005

H3K9me2	4	Abcam ab1220	GR183500-3
H3K27ac	5	Active Motif 39133	3184008
H3K27me3	3	Millipore 07-449	2826067

2.7.5 Quantitative polymerase chain reaction:

2.7.5.1 Preparing primer-mixes for qPCR:

Primers were designed using Primer-BLAST (NCBI) or obtained from the literature (Yoshida et al., 2015a). Settings for design were PCR products of length < 250 bases, not spanning an exon junction, from the Genome for Homo Sapiens. Primers were obtained from Sigma. Working stocks of 100 µM of each oligonucleotide were made in nuclease-free water (Quiagen) following the instruction from Sigma. These were mixed by gentle pipetting and allowed to rest at room temperature for 15 minutes. Then 20 µl of Forward and 20 µL of Reverse primer were mixed with 960 µL of nuclease-free water to make 2 µM stocks aliquoted in 100 µL and frozen -20 °C. Table 2.19 summarises the primers used.

2.7.5.2 Quantitative Polymerase Chain Reaction:

Following PCR clean-up, the 2 µl of Deoxyribonucleic Acid (DNA) from CHIP was diluted in 8 µL nuclease-free water and added to 2.5 µL of primer mix (final concentration of 200 mM per primer) and 12.5 µL of Quantifast SYBR Green mixed by gentle pipetting and then transferred to a MicroAmp fast optical 96-well plate. These then underwent qPCR analysis using a 7500 fast real-time PCR system (Applied Biosystems). In addition to DNA samples, negative controls containing no DNA were also run for each primer mix. The amplification settings were as described in Table 2.18. In addition Melting Curve analysis was performed for each new primer used from 60 to 95 °C.

Table 2.18: qPCR Amplification settings.

Temperature	Time	Cycles
95 °C	5 min	
95 °C	15 s	40
60 °C	60 s	
5 °C	Until end	

Table 2.19: Primer characteristics.

Name	Direction	Sequence	Length	GC %	Melting Tempera-ture °C	Self Complemen-tarity	Self complemen-tarity 3'	Product length
Ornithine decarboxylase antizyme 1 (OAZ-1)	Forward primer	GCCTCAGGGTAGCTGCTAAG	20	59.89	60	7	3	132
	Reverse primer	ATGAAGACATGGTCGGCTCG	20	60.18	55	4	2	
Actin	Forward primer	CACCACCATGTACCCTGGCAT	21	62.42	57.14	4	2	170
	Reverse primer	CAGTGAGGACCCTGGATGTGAC	22	62.31	59.09	4	3	
Glyceraldehyde-3-phosphate dehydrogenase isoform 1 (GAPDH)	Forward primer	GGTCTCCTCTGACTTCAACAGC	22	60.61	54.55	4	2	187
	Reverse primer	AGAGTTGTCAGGGCCCTTTTCT	23	62.04	47.83	8	0	
C-C motif Chemokine 3 precursor (CCL3)	Forward primer	TCCAAGAATAGCATCTTTGAGC	22	56.49	40.91	6	3	103
	Reverse primer	CCCTCTTCACACTCACAGGAG	21	59.73	57.14	3	3	
Signal transducer and activator of transcription 1 (STAT1)	Forward primer	CGCAGGAAAGCGAAACTACC	20	59.56	55	3	0	115
	Reverse primer	GGAACAGCCGCGTCTAATTG	20	59.63	55	6	2	
Unconventional myosin-x (Myo10)	Forward primer	CGCCAACCACACCGTCT	17	59.93	64.71	3	3	85
	Reverse primer	CTGCTTCGGGCTTAGAGTGG	20	60.46	60	3	0	
Antigen peptide Transporter 1 (TAP1)	Forward primer	TGGTAGCAAGCGTGGAAATC	20	58.55	50	4	2	82
	Reverse primer	GAGAAGGGTGTGCGTGATG	19	58.54	57.89	2	1	
Insulin	Forward primer	CAGCTGGAGAACTACTGCAACTAGA	25	62.07	48	7	2	97
	Reverse primer	GCTGGTTCAAGGGCTTTATTCC	22	59.83	50	4	2	

2.7.6.1 Library preparations for next generation sequencing for ChIP-Seq:

The library preparation for NGS was performed in conjunction with the Microarray and Next Generation Sequencing Core Facility of the University of Sheffield. A selection of samples was checked on a 2100 Bioanalyser to confirm the size of the input DNA. The input quantity was checked using the high sensitivity Qubit dsDNA kit. The DNA from the chromatin immunoprecipitations and the 10% input samples were prepared as per the manufacturer's guidelines using NEBNext® Ultra™ 2 DNA Library Prep Kit for Illumina® (New England BioLabs). Samples were indexed using NEBNext® Multiplex Oligos for Illumina (Index primers sets 1-3) (NewEngland Biollabs), adaptor contamination was removed using bead purification, then pooled into two by the core facility and then sent for sequencing to the Edinburgh Genomics facility for sequencing on HiSeq 4000. (Table 2.20 summarises the pooling and indexing strategies).

Table 2.20: ChIP-Seq indexing and pooling strategies.

Sample	Index	Sequence	Pool	Sample	Index	Sequence	Pool
M2527AC	3	TTAGGC	A	D2527M3	1	ATCACG	B
M25K4M3	4	TGACCA	A	M2527M3	2	CGATGT	B
D254M3	5	ACAGTG	A	D327M3	11	GGCTAC	B
M1227AC	15	ATGTCA	A	M25K9	12	CTTGTA	B
M327AC	16	CCGTCC	A	M327M3	13	AGTCAA	B
M25K4M1	18	GTCCGC	A	D25K9	14	AGTTCC	B
D327AC	20	GTGGCC	A	M1227M3	19	GTGAAA	B
D1227AC	22	CGTACG	A	D1227M3	21	GTTTCG	B
D25K4M1	23	GAGTGG	A	M3K9	26	ATGAGC	B
M25%	25	ACTGAT	A	M12K9	31	CACGAT	B
D3K4M1	27	ATTCCT	A	D3K9	34	CATGGC	B
M12%	17	GTAGAG	A	D12K9	42	TAATCG	B
M3%	24	GGTAGC	A				
M3K4M1	28	CAAAAG	A				
M12K4M1	29	CAACTA	A				
D3%	32	CACTCA	A				
D12%	33	CAGGCG	A				
D25%	35	CATTTT	A				
D12K4M1	36	CCAACA	A				
D2527AC	6	GCCAAT	A				
D12K4M3	7	CAGATC	A				
D3K4M3	8	ACTTGA	A				
M3K4M3	9	GATCAG	A				
M12K4M3	10	TAGCTT	A				

2.7.6.2 Next generation sequencing for ChIP-Seq data analysis:

The ChIP-Seq runs were converted to FASTQ files by the core facility. The FASTQ files were aligned to human genome by Dr Tom Giles (University of Nottingham) using the Human genome HG38 assembly downloaded from UCSC. Briefly, the FASTQ files for each ChIP were concatenated. Then the FASTQ file was cleaned using the Illuminaclip function of Trimmomatic followed by trimming of leading and trailing ends and Sliding Window trimming. The reads were then mapped using Burrows-Wheeler Aligner. The SAM file was then sorted using samtools and the files deduplicated using Picard tools. Then, MACS2 was used to perform the peak picking and BEDTools to look for the overlap between the biological replicates. Finally statistical significant enriched genomic locations were calculated using Diffbind in R (Ross-Innes et al., 2012).

2.8 Western blot analysis:

Table 2.21: Laemmli buffer recipe.

Laemmli buffer (2X)	
0.125M	Tris pH6,8
4%	SDS
20%	Glycerol
100mM	DTT
	Bromophenol blue

Following infection and / or pre-stimulation, cells were washed three times in PBS and then lysed in Laemmli buffer. The samples were then drawn through an insulin syringe in order to homogenise them prior to boiling for 5 minutes. The samples and 5µl of Blue Wide Range Protein ladder were then loaded on to a 12% SDS page gel and run at 200V for 45min or until the dye had reached the bottom of the gel.

2.8.1 Semi-dry transfer:

The proteins were then transferred from the resolving gel to a polyvinylidene fluoride (PVDF) membrane. To achieve this, the PVDF membrane was activated in methanol then equilibrated in transfer buffer (47mM Tris-Base, 0.04% SDS) and then placed on top of filter paper imbibed with transfer buffer, the resolving portion of the SDS page gel was then placed on top of the membrane and above this, a further filter paper layer was applied. The gel sandwich was then mounted in the cassette of a Tran-Blot Turbo (BIO-Rad) and run for 15 minutes at 25V.

2.8.2 Chemiluminescence:

Following the transfer, the PVDF membrane was blocked in 5% milk in 0.1M Tris-HCL, 0.16M sodium chloride, 0.05% tween-20 (TBS/tween) by gentle agitation at room temperature. The membrane was then incubated overnight in fresh 5% milk in TBS/tween with the primary antibody (1:2500, except for α-tubulin used at 1:2000), at 5°C while rotating. The antibodies used are listed in Table 2.17 as well as Polyclonal anti-mouse α-tubulin. The following day the membranes were washed three times in TBS/tween, then incubated for 1 hr with rotation at room temperature in 5% milk in TBS/tween with the secondary antibody (HRP-linked goat anti-mouse or goat anti-rabbit (Dako), used at 1:5000).

The membranes were then washed a further three times in TBS/tween prior to 5 minutes incubation with Clarity™ Western ECL and were then imaged on ChemiDoc™ XRS+ system (BioRad). Images were exported as TIFF and densitometry of the bands analysed in Image J (version 1.50g). In order to correct for protein loading differences the ratio of the modified Histone H3 signal to that of α-tubulin was calculated.

2.9 Statistical analysis:

Statistical analysis was performed using either R for the transcriptomic and proteomic analysis, or for all other experiments in Prism version 7.0c (Graphpad). Data is presented as mean and standard deviation (SD) or standard error of the mean (SEM). For all experiments a minimum of 3 biological replicates was used. Comparison between two paired groups employed a paired t-test, for comparison of 3 or more a one-

way Analysis of variance (ANOVA) with Tukey's post-test was performed. For comparison of multiple observations in more than two groups a 2-way ANOVA with Tukey's multiple comparison test was used. Distribution was analysed using Shapiro-Wilk test in R.

Chapter 3: Development and optimization of mass spectrometry workflows for the analysis of histone PTMs

3.1 Introduction:

Mass spectrometry (MS) has emerged as a powerful method to characterise and quantify histone PTMs. It allows unbiased identification and quantification of multiple histone PTMs including combinations in a single run.

There are several different data acquisition strategies that have been developed and employed for the bottom up analysis of histone PTMs. Data dependent acquisition (DDA) is most commonly used and doesn't require any prior knowledge of the PTMs (Plazas-Mayorca et al., 2009). During the analysis, the top N eluting peptides are selected for fragmentation and analysis. However, the quantification of isobaric co-eluting peptides using this approach proves challenging. In addition, low abundant proteoforms may not be selected for MS/MS and therefore won't be identified and quantified.

In light of this, selective reaction monitoring methods (SRM) and parallel reaction monitoring (PRM) have been developed (Bourmaud et al., 2016; Peach et al., 2012; Peterson et al., 2012). These approaches rely on the establishment of an isolation list for all of the different peptide proteoforms to target for tandem mass spectrometry (MS/MS). These are then monitored throughout the high performance liquid chromatography (HPLC) gradient and selected for fragmentation and analysis. Although these approaches improve the sensitivity, they are constrained by total cycle time for multiple PTMs. They are limited by the number of transitions to be monitored throughout the gradient and the need for prior knowledge of which ones to target. Moreover, they will not allow retrospective analysis for novel PTMs.

In order to overcome some of these limitations, data independent acquisition (DIA) methods have been gaining in popularity for discovery proteomics and are particularly suited to the study of PTMs (Krautkramer et al., 2015; Sidoli et al., 2015c, 2015b, 2016). A number of different DIA methods have been used to study histone PTMs (as discussed in Chapter 1). One of the first methods developed was SWATH™ (ABSciex), designed for triple TOF instruments. This method was successfully used to identify and quantify histone PTMs (Sidoli et al., 2015b). It involves a series of 85 isolation windows of variable sizes spanning the range m/z in which histone PTMs are found (see Tables 2.9 and 2.11). Subsequently, Krautkramer *et al.* used a DIA method with regular 10m/z isolation windows to identify and quantify the changes in histone PTMs following histone deacetylase inhibitor treatment (Krautkramer et al., 2015). Using this approach enabled a greater reproducibility than conventional DDA with consistently high numbers of proteoforms identified and with lower coefficient of variations (CV) in the relative abundance in DIA runs when compared to DDA. Indeed, both the SWATH™ and DIA were able to detect low abundance proteoforms. In addition previous studies have shown that DIA protocols can be adapted to low resolution ion trap instruments (Karch et al., 2014; Sidoli et al., 2015c).

The latest mass spectrometry instruments possess increasingly higher speed and resolution. The increased resolution however, comes with increased scanning time. As it has been shown that histone PTMs can be studied using low resolution instruments (Karch et al., 2014; Sidoli et al., 2015c), initial work focused on studying the effects of higher resolution over increased number of scans for the analysis of histone PTMs. In particular, as previous studies have focused on comparing instruments with different

resolutions or different data acquisition methods, in this chapter I sought to compare a range of alternative data dependent acquisition (DDA) methods and data independent acquisition (DIA) methods for the identification and quantification of the histone PTMs.

The purpose of this chapter is to develop and adapt existing methodology to enable the identification and quantification of previously identified histone PTMs.

I hypothesised that a data independent acquisition method would allow identification and quantification of histone PTMs with greatest ease and flexibility.

The aim of this chapter was to:

- Develop and optimise a one dimensional liquid chromatography mass spectrometry (LC-MS/MS) method for the study of histone PTMs on a high resolution QExactive HF Orbitrap mass spectrometer.
- Compare a range of data acquisition methods in conjunction with varying resolution and cycle times to determine which method is best suited for the analysis of histone PTMs in conjunction with downstream bioinformatics.

In addition, I compared these approaches with previously optimised two dimensional liquid chromatography separation (2D) methods in conjunction with ultra-high resolution time of flight mass spectrometry.

In order to minimize the use of primary cells for the optimization of the MS method, I used Chinese hamster ovary (CHO) cells, an important biological system for the production of biopharmaceuticals. Despite their prevalent use in industry, the epigenetics of CHO cells have not been widely studied. In this section of my thesis I characterized and quantified changes in histone PTMs of a CHO-S line between day 2 and day 4 of culture.

3.2 Results and Discussion:

3.2.1 Comparative analysis of histone PTMs on a maXis (UHR TOF) and QExactive HF Orbitrap:

Initially the analysis of PTMs was carried out using protocols developed in collaboration with Dr Minshull (Department of Chemical and Biological Engineering, University of Sheffield) using a novel 2D HPLC method coupled to high resolution time of flight mass spectrometer (maXis, Bruker). However, this method required extensive sample fractionation and lengthy MS analysis time. Therefore, a novel approach was developed in conjunction with a new mass spectrometer in the laboratory, a QExactive HF Orbitrap.

3.2.2 Comparative analysis of 2D vs 1D LC MS for the analysis of histone PTMs:

Previous studies have shown that in order to obtain good coverage of histone PTMs sample fractionation was required when using the MaXis MS (Minshull et al., 2016). However, this is associated with having to run multiple fractions for each sample resulting in up to 20 hours of MS time for each histone analysis. In order to decrease the MS time needed to analyse samples, the results of fractionated samples (2D) were compared to an unfractionated single sample desalted using Hypersep^{TM} Hypercarb tip (1D). Histones were purified from CHO cell pellets grown to day 2 or day 4, by hypotonic lysis and acid extraction (as described in 2.6.1. Histone extraction and purification). The purified histones were subsequently analysed on a 12% SDS page gel to assess purity and estimate concentration when comparing to a known calf histone standard (as illustrated in Figure 2.3). 10 μg of purified histones was then subjected to chemical derivatization and digestion (Benjamin A Garcia et al., 2007) before being either desalted and fractionated offline using reverse phase HPLC with Hypercarb column (Minshull et al., 2016) prior to RP LC MS analysis (2D LC MS) or desalted using Hypersep Hypercarb tips prior to RP LC MS analysis (1D LC MS). Both the tips and the column use the same porous graphitic carbon (PGC) as the stationary phase to achieve the desalting. Analysis was performed on two different mass spectrometers a maXis UHR TOF and the QExactive HF Orbitrap (QE) (Figure 3.1).

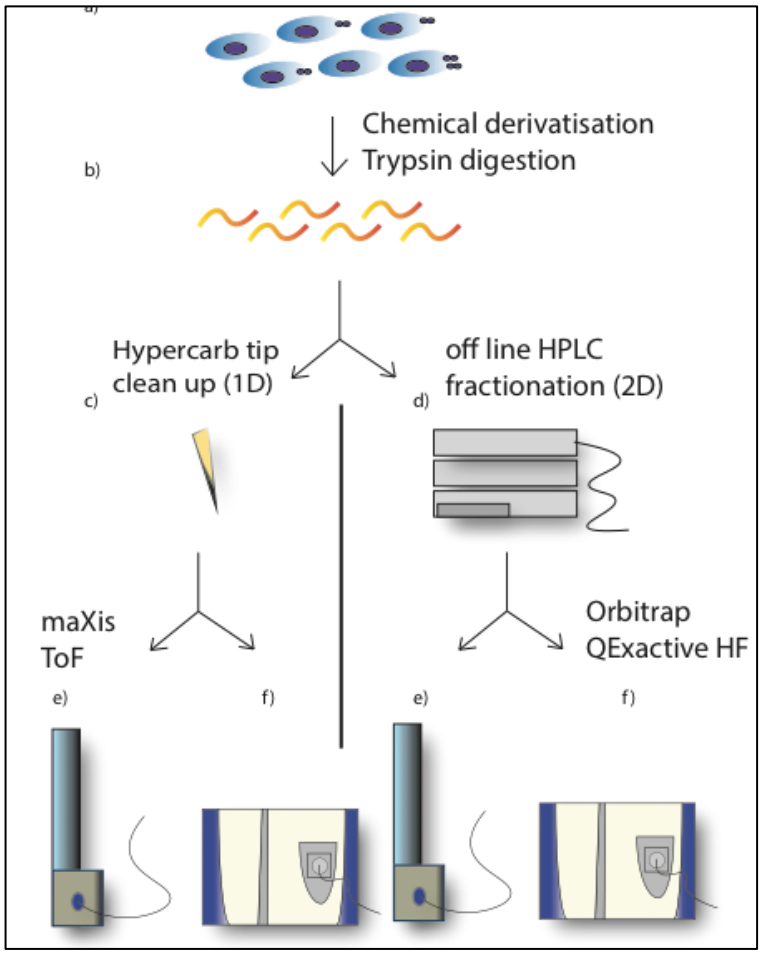


Figure 3.1: Schematic representation of mass spectrometry workflow.
a) CHO cells are collected and undergo hypotonic lysis, b) the purified histones are propionylated, digested and re-propionylated, The peptides are then desalted using either c) Hypersep^{TM} Tip (1D), or d) desalted and fractionated on Hypercarb column on a HPLC (2D). The single sample or fractions are then analysed on either e) the maXis or the f) QE.

For the data acquired using the maXis, the identification of PTMs was carried out using Mascot searches in combination with manual verification of PTMs using Data Analysis (BrukerDaltronics). The relative abundance was calculated using in house software (Hist-O-Matic). For the data acquired on the QE, the analysis was carried out using Epiprofile and / or Skyline for the manual verification of the PTMs, as well as the relative abundance quantification. As the analysis of datasets from each instrument was performed using separate software, it is not ideally suited to a direct comparison between instruments.

However, it is possible to compare the sample preparation methods on each instrument. A summary of the identification and relative quantification of the histone PTMs is shown in Figure 3.2. The same number of histone proteoforms were identified using the two sample preparation methods. Next, the relative abundance of the two sample preparation methods was compared. This showed that using the MaXis, I identified the same number of histone PTMs and their relative abundance was consistent using the two different sample preparation techniques (Figure 3.2).

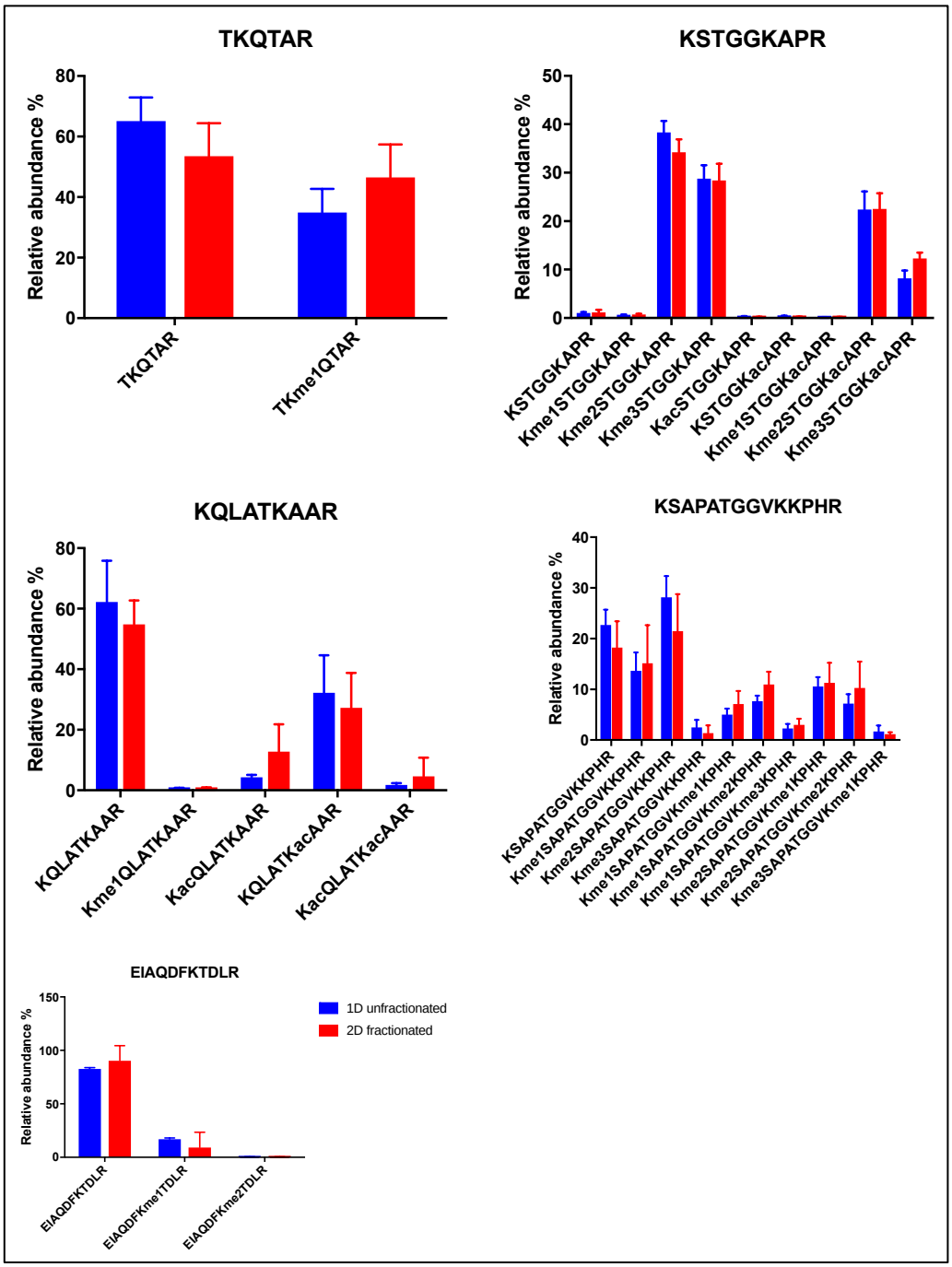


Figure 3.2: Comparison of sample preparation methods on relative abundance of Histone PTMs.

The relative abundance of histone H3 PTMs from day 4 CHO cells was measured using a one hour gradient on the maXis for samples prepared using either 1D LC MS dimension (blue) or 2D LC MS (red). There were no significant differences between the two methods for the five principle peptides of Histone H3 shown here. n=4.

The comparison of the two different sample preparation methods was also carried out using the QExactive HF (Figure 3.3). The results show that similar to the data obtained on the maXis instrument, the two different sample preparations gave equivalent relative abundances across a wide number of peptide proteoforms, and that the 1D LC MS method identified all of the PTMs seen in the 2D LC MS sample preparation method (one of the limitations of 1D preparations previously observed on the MaXis (Minshull et al., 2016)).

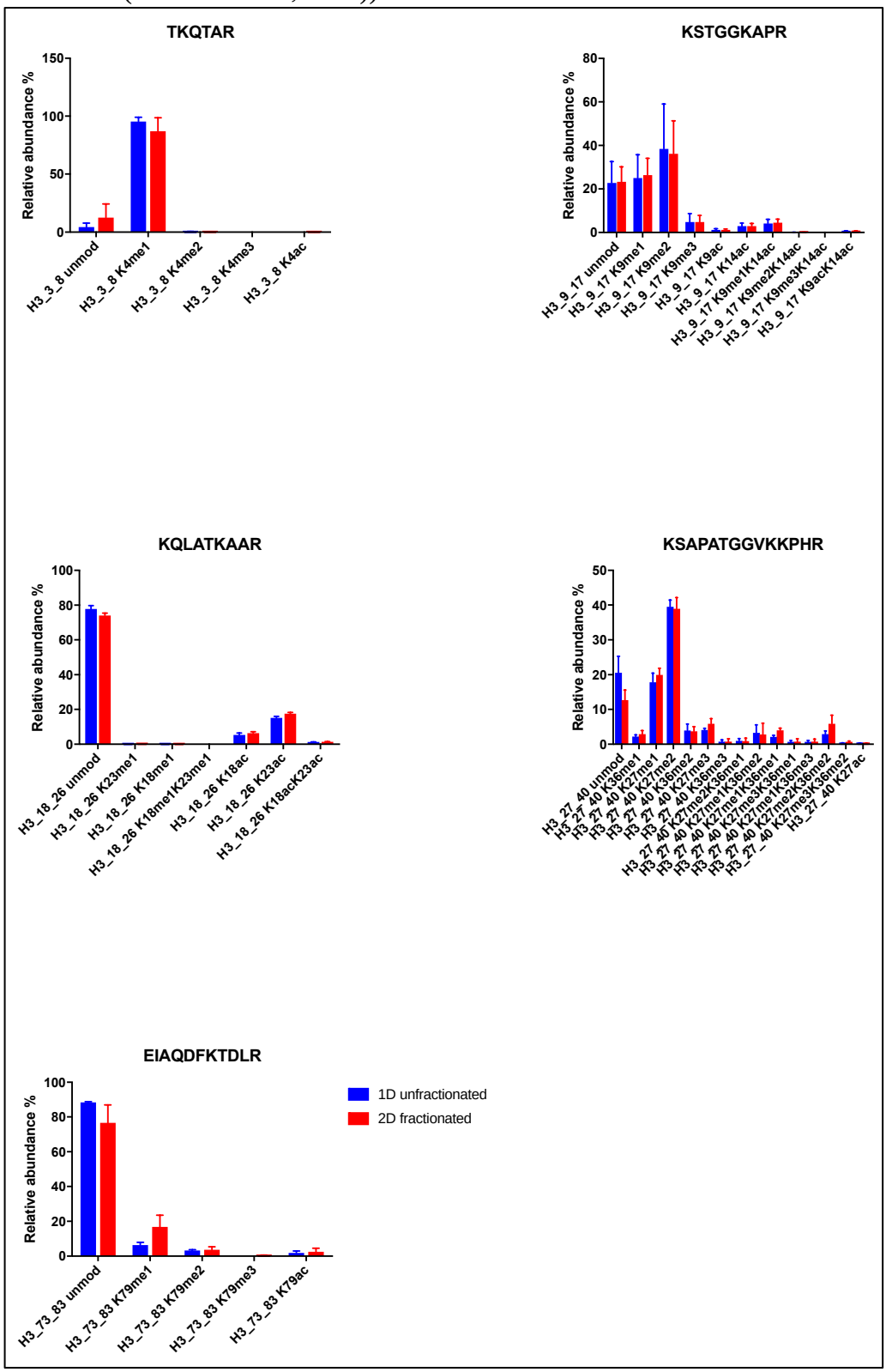


Figure 3.3: Comparison of sample preparation methods on the relative abundance of Histone PTMs using QExactive HF Orbitrap.
The relative abundance of histone H3 PTMs from day 4 CHO cells was measured using a 105min run on the QE in DDA for samples prepared using either 1D LC MS(blue) or 2D LC MS (red). There were no significant differences between the two methods. n=4.

The two different MS instruments had different analytical columns and gradient times which makes the direct comparison between the two instruments difficult. Moreover, during the optimization of the QE methodology the HPLC coupled to the MS methods included a 5 minute wash period prior to the valve change and the MS analysis starts, to ensure the instrument didn't become contaminated. This led to the loss of early eluting proteoforms such as H3 lysine 9 dimethylation (H3K9me2) (Figure 3.4) and makes a detailed comparison in the relative abundances between the two instruments challenging.

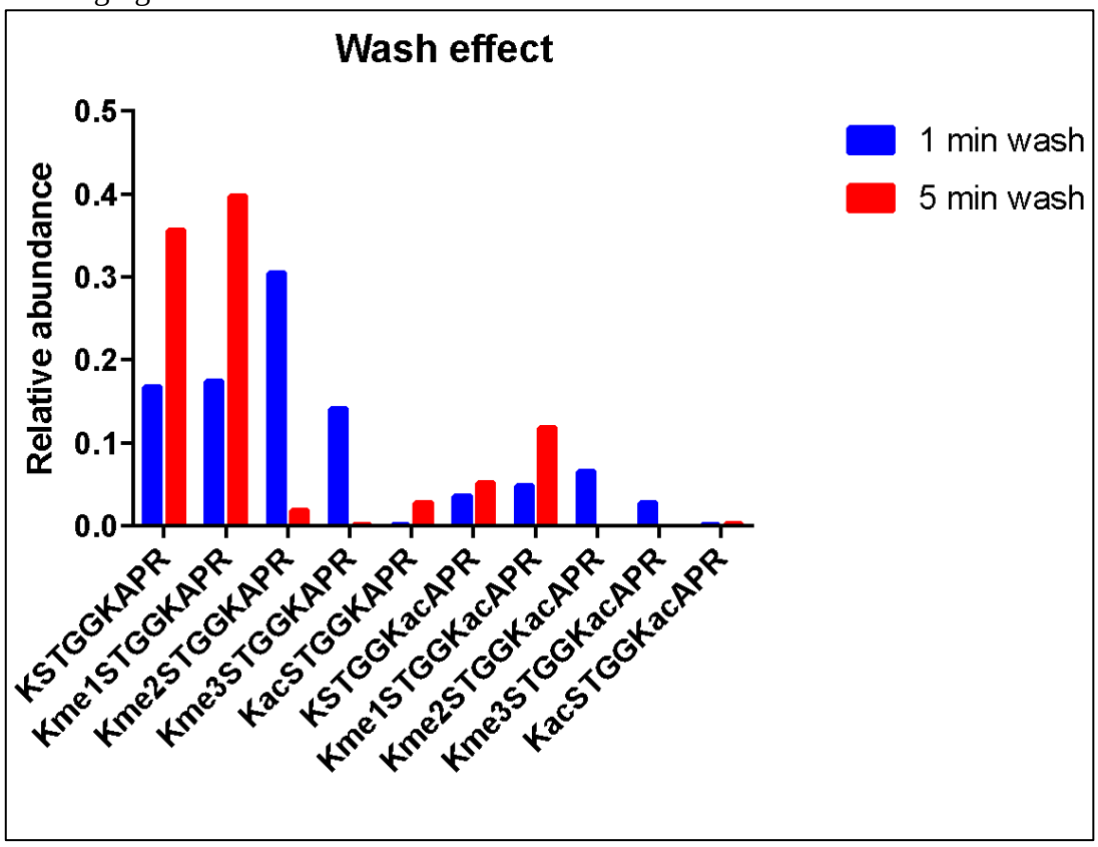


Figure 3.4: Comparison of 1 min and 5 min wash periods prior to MS analysis on the relative abundance of histone PTMs.
The relative abundance of the different acetylation and methylation proteoforms for the KSTGGKAPR peptide are compared. In blue are the abundances obtained following a 1 min wash showing higher abundance of the early eluting proteoforms (dimethylation and trimethylation of lysine 9 on histone H3 (H3K9me2, H3K9me3), and the combinatorial marks histone H3 lysine 9 dimethylation and trimethylation with lysine 14 acetylation (H3K9me2K14ac and H3K9me3K14ac) than was obtained following a 5min wash (in red) from a repeat injection of the same sample. (n=1)

3.2.3 Comparison of number of proteoforms identified on a QExactive HF Orbitrap to that of the maXis (ToF):

As previously established, the number of histone PTMs identified using the maXis in a 1D LC MS approach was inferior to that of the 2D LC MS method (Minshull et al., 2016), although a significant reduction in the amount of MS time was

achieved using the 1D approach. In order to overcome this, the QExactive HF Orbitrap was used in light of its faster speed, reduced scanning time and high resolution. As one of the main limitations of 1D LC MS workflows on the maXis was a reduced number of different PTMs identified when compared to 2D LC MS analysis, a comparison of the total number of PTMs identified between the two instruments would indicate whether the 1D LC MS analysis was a suitable approach on the QExactive HF.

The number of proteoforms identified using the 1D LC MS method on each instrument was compared (Figure 3.5). The results show that a greater number of peptide proteoforms were identified on the QExactive HF compared to the 2D LC MS analysis on the maXis. This highlights the benefits of the newer instrument, the QExactive HF, in identifying more PTMs than the maXis. Furthermore, it is able to achieve this using the 1D LC MS workflow, in effect reducing the MS time by 18 hours per biological sample, as total run time for 1D LC MS on the QE is 105 minutes compared to around 20 hours for a 2D LC MS workflow on the maXis.

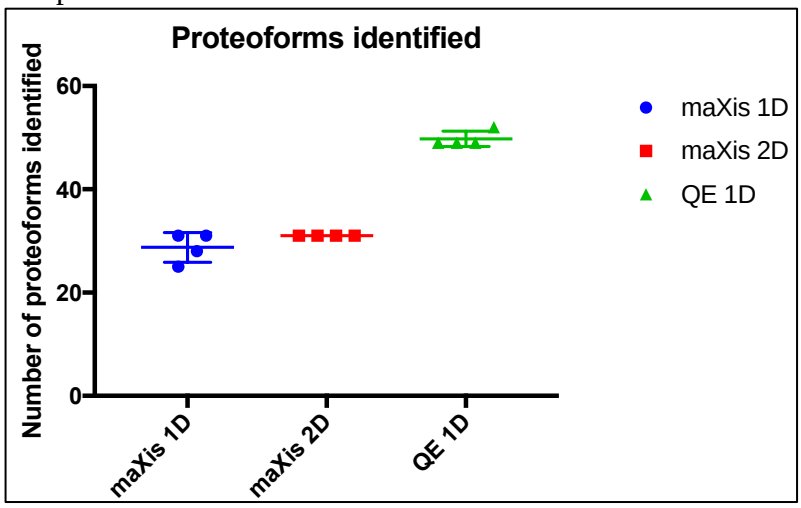


Figure 3.5: Comparison of number of proetoforms identified on each instrument.

The number of proteoforms for Histone H3 identified by manual verification of the spectra following analysis on the MaXis was compared to that of the number of proteoforms identified by Epiprofile and manually verified in Skyline following analysis on the QE from 4 separate replicates.

Given that two different data analysis approaches were used for each instrument a direct comparison in the relative abundances is however, not possible.

3.2.3 Development and optimization of MS workflows on the QE HF for the characterization and quantification of histone PTMs:

Following demonstration of the ability of the QE HF to rapidly identify large numbers of peptide proteoforms and histone PTMs in comparison to previously developed methods on the maXis (UHR TOF) instrument, further studies were performed to develop and optimize MS workflows on the QE HF for the characterization and quantification of histone PTMs.

Initial work focused on studying the impact of MS2 resolution in data dependent acquisition methods (described in Chapter 2) for the characterization and quantification of histone PTMs. Three different DDA methods with MS2 resolutions of 120 000, 60 000 and 30 000 were compared. In addition to DDA methods, I also studied DIA methods and the impact of reducing the resolution with MS2 isolation windows of 20 m/z at a resolution of 30 000 or 15 000. Finally, an emulated SWATH™ protocol with variable window size (DIAvw) was studied as this was one of the first DIA methods

developed for the study of histone PTMs (Sidoli et al., 2015b).

In order to establish which data acquisition approach (DDA or DIA) is best suited to the study of histone PTMs and to determine the influence of different MS2 resolution settings on the identification and quantification of histone PTMs each sample was run in the 6 different data acquisition methods (summarized in Table 2.10).

The baseline characteristics of each data acquisition method were established. The duty cycles for each method were calculated (Figure 3.6 panel a). This illustrates that the higher the resolution used, the slower the instrument scanned, resulting in longer cycle times. This however was kept below 5 seconds for the DDA allowing at least 7 MS1 scans in a 30 second peak which is typical for the elution of the different peptide proteoforms using the online chromatography employed in this study. The DIAvw had the longest duty cycle of 5.1 seconds and the shortest 1.4 seconds with DDA30. As would be expected as the resolution was decreased and the cycle time decreased.

Next the number of MS1 and MS2 scans for each method was examined using RawMeat. The lower resolution methods were associated with increased numbers of MS1 scans which should enable greater accuracy in the label-free quantification (Figure 3.6 panel b). A higher number of MS2 scans was obtained in DIA mode which should enable greater discrimination and quantification of isobaric peptides (Figure 3.6 panel c).

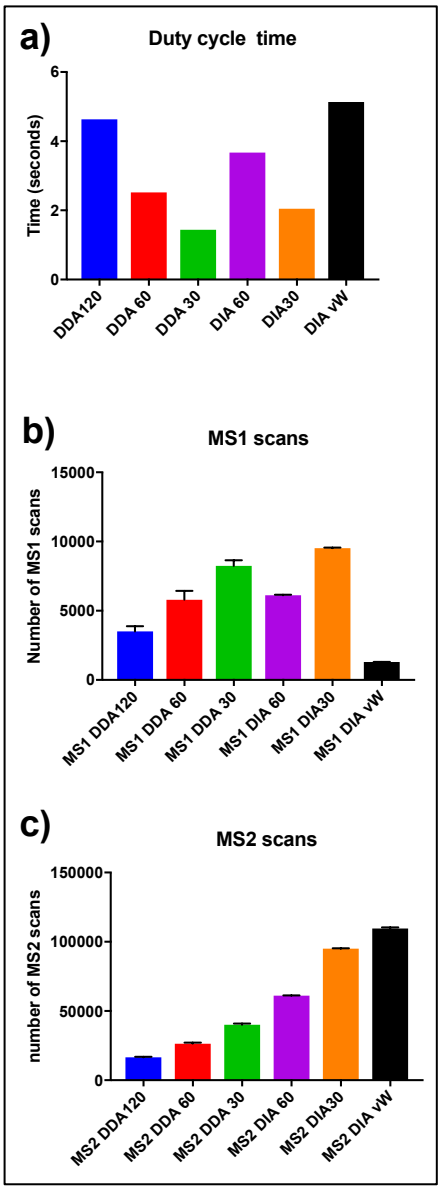


Figure 3.6: Comparison of the effect of different data acquisition methods on the duty cycle time and number of MS1 and MS2 scans.

a) The duty cycle time for each data acquisition method is represented. As the resolution is decreased the duty cycle time decreases. b) This shows the number of MS1 and c) the number of MS2 scans for each data acquisition method. As less time is spent on MS2 scans more MS1 scans are undertaken. (n=6)

3.2.4 Comparison of the impact of resolution on the identification of histone PTMs using data dependent acquisition:

To assess the ability of the three different DDA methods to accurately identify histone PTMs, the RAW files were converted to mgf using MSConvert (proteowizard). Searches were performed using Mascot Daemon 2.5.0 as this has previously been shown to be the most reliable for the identification of histone PTMs (Yuan et al., 2014). The mgfs were searched using CHO proteome (downloaded from Uniprot (downloaded 07/06/2017), using Arg-C digestion, a peptide tolerance 10 ppm, and an MSMS tolerance 0.01 Da, no missed cleavages, Peptide charges of 2, 3 and 4+; fixed modifications (propionyl (K) and propionyl (N-term)) and variable modifications (acetyl (K), methylpropionyl (K), dimethyl (K) and trimethyl (K)), FDRs were set to

less than 2%.

In order to ascertain the impact of the variation in resolution of the DDA methods on the ability to accurately identify histone PTMs, the number of peptide sequence matches (PSMs) and total number of queries in each of the different DDA methods were analysed (see Figure 3.8). The results show that the lower resolution scans were associated with greater number of queries and consequently a greater number of PSMs. However, the conversion rate (the proportion of PSMs to the number of queries) was lower in the DDA30 method.

I next examined the results of Mascot searches in the three different DDA methods, and their ability to correctly identify the PTMs of histone H3. The Mascot search results give a peptide proteoform ion score based on the probability of this being identified by chance. A score of 46 or greater is associated a 5% confidence threshold. (The results are summarized in Figure 3.8 panel a). The results show that as the MS2 resolution decreased, more scans were performed and more peptide proteoforms were identified. To determine the accuracy of these potential identifications the Mascot ion scores associated with each proteoform were examined. The highest Mascot ion score, and therefore the most confident identification, was the same across all three methods despite the increasing ppm error in the lower resolution scans (Figure 3.8 panel b). Indeed, the proportion of proteoforms with Mascot ion scores greater than 46 was higher in the higher resolution scans (77%, 75% and 70% respectively).

Correctly identifying the position of PTMs can be challenging given that histone peptides are heavily modified and the near isobaric nature of acetylation and trimethylation. In order to further disambiguate the position of PTMs, the Mascot Delta score was calculated for each of the proteoforms identified across the different methods. Previous work in the field of Phosphoproteomics has determined that a Mascot Delta score of greater than 17 was associated with accurate location of phosphorylation (Savitski et al., 2011). Mascot Delta scores were calculated by taking the difference between the highest ion score for a given proteoform and the score for the next possible proteoform (Figure 3.7).

The results show that despite identifying a greater number of total proteoforms, the lower resolution scans did not do so with the same degree of confidence. The higher resolution scans had a higher proportion of proteoforms with a Mascot ion score greater than 46. However, the proportion of proteoforms with a Mascot Delta score of greater than 17 was the same with all three data acquisition methods (approximately 20% (Figure 3.8 panel A)).

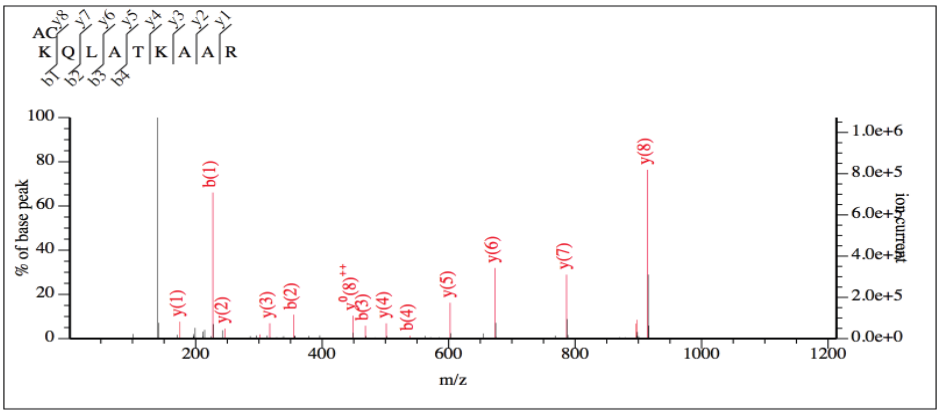
In summary, the results show that the increased number of MS2 scans afforded by the lower resolution DDA method (DDA30) resulted in a higher number of queries, PSMs and a higher number of peptide proteoforms identified with a Mascot Ion score greater than 46, with no difference in the proportion of peptide proteoforms with Delta scores >17. These results indicate that no significant benefit is gained by performing DDA analysis using high resolution MS2 scans on the QE HF for the analysis of histone PTMs.

Mascot Search Results

Peptide View

MS/MS Fragmentation of **KQLATKAAR**
Found in **G3H2T7** in **CHO_Uniprot10029**, Histone H3 OS=Cricetulus griseus GN=I79_004539 PE=3 SV=1

Match to Query 4993: 1139.666878 from(570.840715,2+) intensity(3438939.5000) rtinseconds(4034.4657) index(9423)
Title: JOBY_D2_A1_dda120.12211.12211.2 File:"JOBY_D2_A1_dda120.raw", NativeID:"controllerType=0 controllerNumber=1 scan=12211"



Monoisotopic mass of neutral peptide Mr(calc): 1139.6662
Fixed modifications: Propionyl (K),Propionyl (N-term) (apply to specified residues or termini only)
Variable modifications:
K1 : Acetyl (K)
Ions Score: 54 **Expect:** 7e-005
Matches : 16/80 fragment ions using 30 most intense peaks (help)

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	227.1390	114.0731	210.1125	105.5599			K							9
2	355.1976	178.1024	338.1710	169.5892			Q	914.5418	457.7745	897.5152	449.2613	896.5312	448.7693	8
3	468.2817	234.6445	451.2551	226.1312			L	786.4832	393.7452	769.4567	385.2320	768.4726	384.7400	7
4	539.3188	270.1630	522.2922	261.6498			A	673.3992	337.2032	656.3726	328.6899	655.3886	328.1979	6
5	640.3665	320.6869	623.3399	312.1736	622.3559	311.6816	T	602.3620	301.6847	585.3355	293.1714	584.3515	292.6794	5
6	824.4876	412.7475	807.4611	404.2342	806.4771	403.7422	K	501.3144	251.1608	484.2878	242.6475			4
7	895.5247	448.2660	878.4982	439.7527	877.5142	439.2607	A	317.1932	159.1002	300.1666	150.5870			3
8	966.5619	483.7846	949.5353	475.2713	948.5513	474.7793	A	246.1561	123.5817	229.1295	115.0684			2
9							R	175.1190	88.0631	158.0924	79.5498			1

All matches to this query

Score	Mr(calc)	Delta	Sequence
54.2	1139.6662	0.0006	KQLATKAAR
15.0	1139.6662	0.0006	KKEQKAAR
8.2	1139.6662	0.0006	KKEQKAAR
8.2	1139.6662	0.0006	KKEQKAAR
8.2	1139.6662	0.0006	KKEQKAAR
7.9	1139.6662	0.0006	KAKLQKSR
7.9	1139.6663	0.0006	KIQVGTKGR
7.9	1139.6662	0.0006	KVAAISVAQR
7.5	1139.6663	0.0006	KKQQEKR

Delta Score:
54.2 - 15.0 = 39.2

Figure 3.7: Delta Score calculation.
This figure shows the Mascot search result for the KQLATKAAR peptide with a single acetylation at lysing 18. The Delta score calculation is illustrated. From the highest Mascot Ion score for the query (54.2) the score for the next possible match (KKEQKAAR, 15.0) is subtracted giving a Delta score of 39.2.

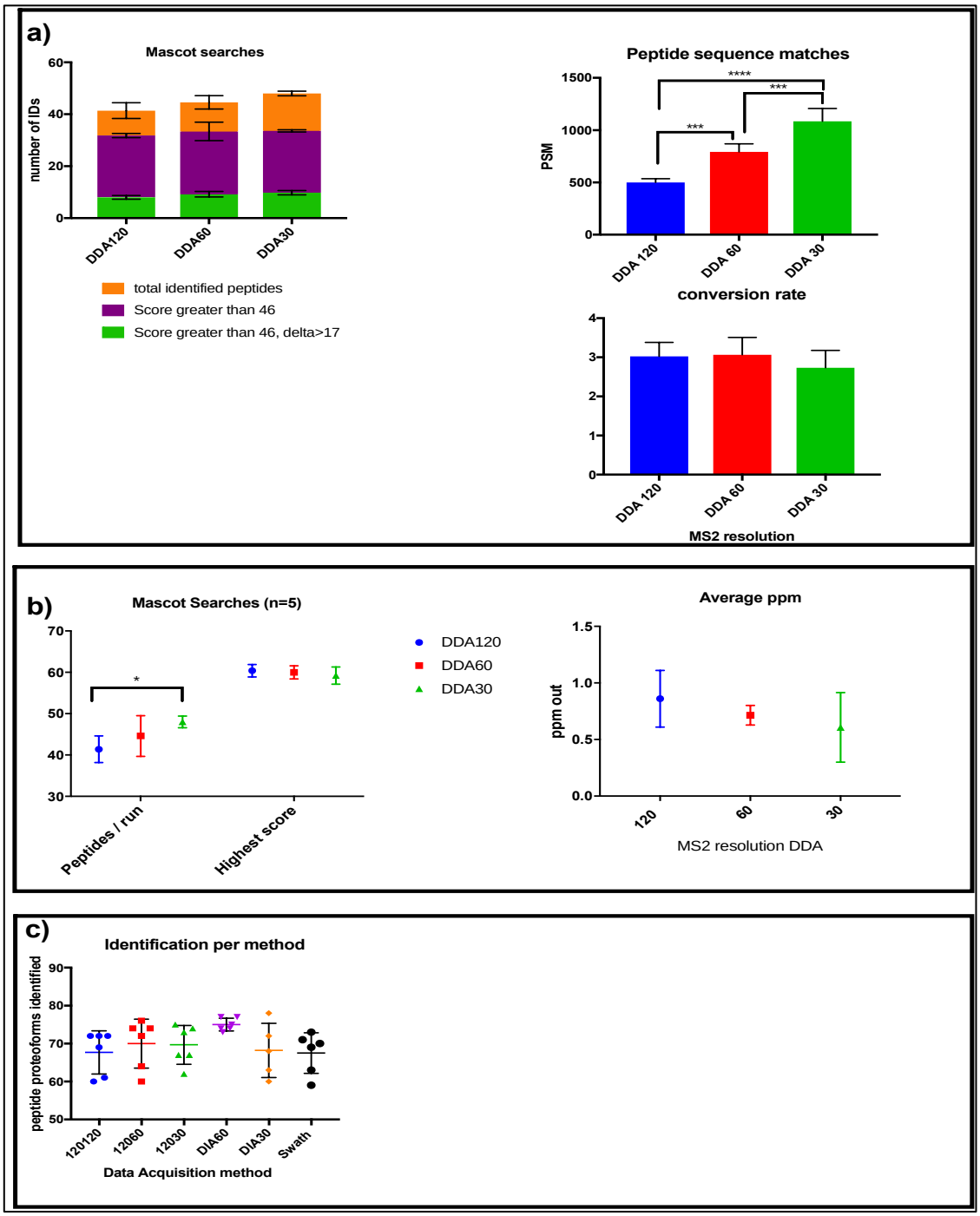


Figure 3.8: Analysis of data dependent acquisition methods.
A) The DDA runs were searched using Mascot Daemon, FDR was set at less than 2% using a reversed sequence library. Peptides from H3.1 and H4 were examined. The proteoforms with Mascot Ion scores >46 and Delta Score >17 are represented. The peptide sequence matches above the identity threshold and the conversion rate are also represented for each method. There were significantly more peptide sequence matches in the lower resolution runs B) This panel illustrates the number of peptide proteoforms identified in each data dependent acquisition methods and the highest Mascot score associated with it and the average ppm error for each correct identification. There were significantly less peptide proteoforms identified in the higher resolution DDA120 than in the DDA30. C) The number of peptide proteoforms identified using Epiprofile 2.0 in each different data acquisition method illustrates that there was slightly higher number of peptide proteoforms identified with the DIA 60 approach. n=6, represented are the mean and standard deviation. (* $p<0.05$, *** $p<0.0005$, **** $p<0.0001$; one way ANOVA with Tukeys multiple comparisons test).

3.2.5 Comparison of data independent and data dependent acquisition methods for the analysis of histone PTMs:

Having examined the ability of the different DDA methods to identify PTMs, further analysis was performed on the QE HF using DIA methods (described in Chapter 2) to enable a comparison between DDA and DIA methods. In order to establish the identification of proteoforms in DIA runs, data analysis was performed using the software Epiprofile, which was specifically developed for the identification and quantification of Histone PTMs and is capable of handling both DDA and DIA data (Yuan et al., 2015).

Briefly, the three DIA methods (described in greater detail in Chapter 2) employed are:

DIA60, a method using sequential 20m/z isolation windows from 300 to 900m/z with MS1 scans every 10 at a resolution of 60 000 at 200m/z and an MS2 resolution set at 30 000 at 200 m/z;

DIA30, a method using sequential 20m/z isolation windows from 300 to 900m/z with MS1 scans every 10 at a resolution of 30 000 at 200 m/z and an MS2 resolution set at 15 000 at 200 m/z;

DIAvw, a method using 85 sequential variable isolation windows to mimic the SWATH^{TM} method.

The total number of peptide proteoforms identified across all of the different acquisition methods in Epiprofile for histones H3 and H4 was compared (Figure 3.8 panel c). The results showed that on average 69 proteoforms were identified in each method (ranging from 60 to 77). 47 proteoforms were identified in all of the runs (68% of average identified) and 90% of all proteoforms were identified in at least 3 out of 6 runs in each method. This showed that DIA60 identified slightly more proteoforms (although this did not reach statistical significance), 75 across all 6 runs, of which 96% were identified in at least 3 runs, than the other methods.

3.2.6 Quantification of histone PTMs:

Having established that all of the DDA and DIA methods were able to consistently identify the same pattern of lysine methylation and acetylation on Histone H3 and H4 and that this is consistent with the number of different acetylation and methylation sites identified in the literature (albeit in different cell lines) (Kulej et al., 2015), their ability to accurately report the relative quantification of histone PTMs was assessed. The relative abundance of each histone proteoform is calculated by measuring the area under the curve (AUC) of the extracted ion chromatogram (XIC) corresponding to the individual PTMs for each peptide, dividing the AUC of one PTM of a peptide by the sum total of the AUC of all PTMs for that peptide (DiMaggio et al., 2009).The quantification was performed in Epiprofile (Yuan et al., 2015). In addition to the use of Epiprofile further validation of the relative quantification was performed in Skyline (MacLean et al., 2010; Schilling et al., 2012), which can analyse data from both DIA and DDA, and determine the relative abundance for each peptide proteoform .

The relative abundance of histone PTMs in CHO cells has been shown to vary over time and in response to evolutionary pressures (Feichtinger et al., 2016). Furthermore, it has been shown that global levels of H3 acetylation decrease over time (Paredes et al., 2013). Therefore, the relative abundances of histone PTMs between day 2 and day 4 of culture, as I expected to find a difference between the two, was analysed across each of the different acquisition methods (Figure 3.9). The results show a wide

range of histone PTMs were identified and quantified (summarised in Figure 3.10). In order to further analyse the quantitative differences obtained across these different methods I focused on a number of peptide proteoforms that were initially identified as having changed in abundance between day 2 and day 4.

Proteoform	Day 2	Day 4	Fold Change	Significant:
KSTGGKAPR			-1.12	**
Kme1STGGKAPR			-0.71	**
Kme2STGGKAPR			0.68	**
Kme3STGGKAPR			0.65	**
KacSTGGKAPR			2.78	
KSTGGKacAPR			-1.91	
Kme1STGGKacAPR			-1.64	
Kme2STGGKacAPR			-6.54	**
Kme3STGGKacAPR			-6.90	
KacSTGGKacAPR			-0.74	
KQLATKAAR			-0.08	**
KQLATKme1AAR			0.21	
Kme1QLATKAAR			-0.28	
Kme1QLATKme1AAR			0.62	
KacQLATKAAR			0.52	
KQLATKacAAR			0.33	**
KacQLATKacAAR			0.81	
KSAPATGGVKKPHR			0.74	**
KSAPATGGVKme1KPHR			-0.54	
Kme1SAPATGGVKKPHR			0.59	**
Kme2SAPATGGVKKPHR			0.31	**
KSAPATGGVKme2KPHR			-0.19	
Kme3SAPATGGVKKPHR			-0.48	**
KSAPATGGVKme3KPHR			0.23	
Kme2SAPATGGVKme1KPHR			-0.63	**
Kme1SAPATGGVKme2KPHR			-0.83	**
Kme1SAPATGGVKme1KPHR			-0.74	**
Kme3SAPATGGVKme1KPHR			-1.41	**
Kme1SAPATGGVKme3KPHR			-1.23	
Kme2SAPATGGVKme2KPHR			-0.98	**
Kme3SAPATGGVKme2KPHR			-1.83	
KacSAPATGGVKKPHR			0.87	
YQKSTELLIR			0.00	
YQKme1STELLIR			1.24	
YQKme2STELLIR			0.31	
YQKacSTELLIR			-0.02	
EIAQDFKTDLR			-0.03	**
EIAQDFKme1TDLR			0.04	
EIAQDFKme2TDLR			-0.11	
EIAQDFKacTDLR			2.06	**
VTIMPKDIQLAR			-0.88	**
VTIMPKacDIQLAR			8.40	**
GKGGKGLGKGGAKR			-0.59	**
GKacGGKGLGKGGAKR			-2.54	
GKGGKacGLGKGGAKR			-2.71	**
GKGGKGLGKacGGAKR			-2.21	**
GKGGKGLGKGGAKacR			0.76	**
GKacGGKacGLGKGGAKR			-2.11	
GKacGGKGLGKacGGAKR			-1.39	
GKacGGKGLGKGGAKacR			0.89	
GKGGKacGLGKacGGAKR			-2.31	
GKGGKacGLGKGGAKacR			1.01	**
GKGGKGLGKacGGAKacR			0.14	
GKacGGKacGLGKacGGAKR			-1.40	
GKacGGKacGLGKGGAKacR			1.78	
GKacGGKGLGKacGGAKacR			0.38	
GKGGKacGLGKacGGAKacR			0.93	
GKacGGKacGLGKacGGAKacR			1.17	

Figure 3.10: Heatmap of all of the different PTMs identified for histone H3 and H4 using DIA 60 for day 2 and day 4 CHO cells.

This figure illustrates the total number of different histone PTMs identified across histone H3 and H4 and the ability of DIA 60 to identify changes in relative abundance between the two days. n=3, ** $p<0.05$ Students t test.

As the overall ion intensity of a proteoform may influence the accuracy of both correct identification and quantification, examples of proteoforms with high, medium and low ion intensities (which was arbitrarily defined as XIC intensities of $>9x10^9$, $>8x10^7$,$>3x10^5$ respectively) were studied. A number of peptide proteoforms covering both a range of different ion intensities (low, medium and high) and a range of different relative abundances, where relative abundance reflects the percentage of the peptide proteoform with respect to the total abundance of all peptide proteoforms for the corresponding peptide on both histone H3 and histone H4 were selected.

3.2.6.1 High intensity PTMs:

Figure 3.9 panel A shows that each method was able to confidently demonstrate the change in relative abundance of the highly abundant peptide GKGGKGLGKGGAKR, from histone H4, between day 2 and day 4.The ability to correctly identify changes in the single acetylated form of KQLATKAAR peptide of histone H3 on lysine 23 was examined next (see Figure 3.9 panel A). In light of the isobaric nature of the acetylation on K18 or K23 and as both forms co-elute, the relative abundance is based on the proportion of diagnostic y and b ions belonging to one form or the other (Yuan et al., 2015). The results show that all of the different methods, except DIA30, reported the change in relative abundance of the peptide proteoform. However, it should be noted that the DIA30 analysis showed the same trend with an increase in K23 acetylation with reciprocal decrease in the unmodified form, but failed to reach statistical significance. For the KQLATKAAR peptide the DIA30 analysis showed the same trend in increasing in K23 acetylation with reciprocal decrease in the unmodified form, but failed to reach significance.

3.2.6.2 Mid Intensity PTMs:

In the mid intensity peptides, such as the dual acetylated peptides GKGGKacGLGKGGAKacR of H4 or KacQLATKacAAR of H3, the difference in relative abundance between day 2 and day 4 samples were observed in all methods, reaching statistical significance except in the DIA30 (Figure 3.9 panel B).

In addition to looking for changes in relative abundances between the two days, I also looked at the relative abundance of a peptide in which I did not expect to see a change in abundance, such as the unmodified peptide YQSTELLIR (Figure 3.9 panel B). All of the DDA, the DIA60 and DIA30 showed the same relative abundance between the two days. Although the relative abundance of the unmodified YQSTELLIR peptide from H3 across all of the methods was consistent, the DIAvw method showed greater variability compared to the other data acquisition methods with a low level of change in between the abundance of the day 2 and day 4 samples.

3.2.6.3 Low intensity PTMs:

Finally, in the lower intensity, proteoforms such as YQKacSTELLIR on histone H3 were analysed (see Figure 3.9 panel C). The results showed that there were no significant differences between the methods. However, increased variability was observed between replicates, for all of the methods, as would be expected for low

intensity, low abundance peptide proteoforms. When the changes in the relative abundance of KacSAPATGGVKKPHR (H3K27ac) between day 2 and day 4 were examined, an increase in acetylation in all three DDA methods and in both DIA60 and DIAvw was observed. The DIA30 was unable to distinguish the increase seen in the relative abundance of H3K27ac (Figure 3.9 panel C).

Overall, all of the DIA methods used in this study were able to identify the same trend in changes in the relative abundance of the more prominent PTMs (Figure 3.9 panel D). The DIA30, that has lower resolution, displayed greater variability in the fold change in abundance between the two days as evidenced by the greater standard deviations in the KacQLATKacAAR peptide. The DIAvw that has the greatest cycle time displayed greater variability in the fold change in abundance between the two days for the YQSTELLIR peptide.

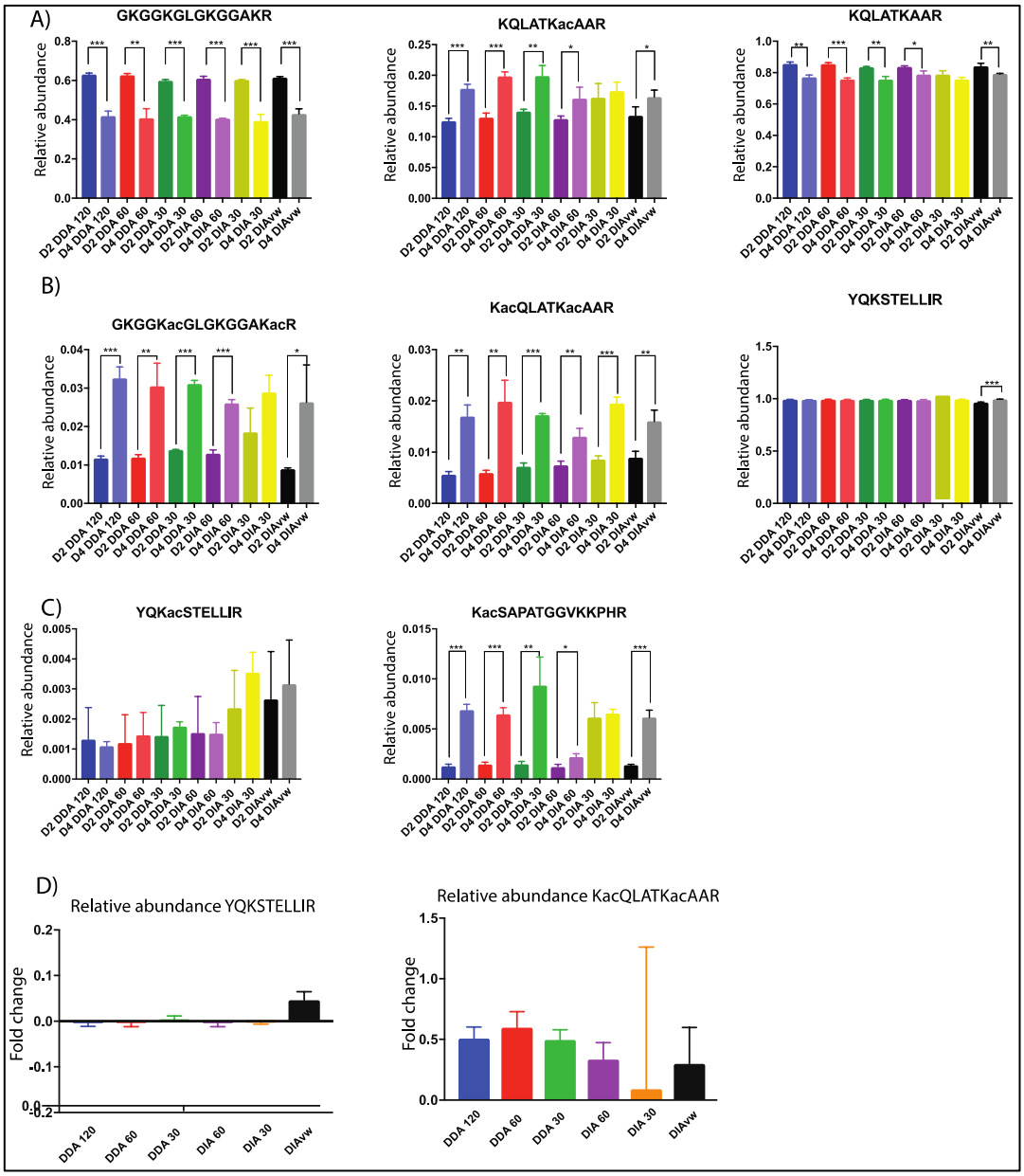


Figure 3.9: Comparison of relative abundance of histone PTMs between data acquisition methods.

The relative abundances of histone post-translational modifications for day 2 and day 4 CHO cells were calculated using Epiprofile 2.0 for the different data acquisition methods. Panel A illustrate three histone

PTMs with high MS1 intensity GKGGKGLGKGGAKR on histone H4, KQLATKacAAR and the unmodified form on histone H3. There is good concordance between the different methods except for the DIA 30, which showed the same trend but didn't reach statistical significance. Panel B illustrates the relative abundance of a three peptide PTMs with middle intensity of MS1 signal. GKGGKacGLGKGGKacR and KacQLATKacAAR representing both lower relative abundance peptides for the respective peptides and YQSTELLIR the unmodified form of which is highly abundant. Panel C illustrates two peptide PTMs of both low relative abundance and low MS1 intensity. Overall there is good concordance between the different data acquisition methods in their ability to distinguish significant difference between the abundance of PTM between day 2 and day4 CHO cells. In addition, with the exception of DIAvw, the peptide PTMs which didn't change between the two days were correctly identified. Panel D illustrates the fold change between the relative abundance of PTM for day 2 and day4 CHO cell for the YQSTELLIR peptide and the KacQLATKacAAR peptide. For all the panels, n=3, bar charts represent the mean and standard deviation, * $p<0.05$, ** $p<0.01$, *** $p<0.001$ Students t test.

3.2.7 Reproducibility of the relative abundance quantification:

In order to establish the reliability of the relative abundance measurements of each method, three technical replicates for day 2 and day 4 were examined and the coefficient of variation (CVs) calculated for each peptide proteoform identified in all replicates (see Figure 3.11). The results show, as expected, that there was greater variability in the proteoforms with the lowest intensities in all data acquisition methods. Three quarters of the CVs were 20% or below for the DIA60 method. The median CV varied from 10% for DIA60 to 15% for DIA30. In comparing all of the proteoforms together there was a trend to smaller CVs with the DIA60 compared to the other methods (Figure 3.11), suggesting this is the most reliable quantification method.

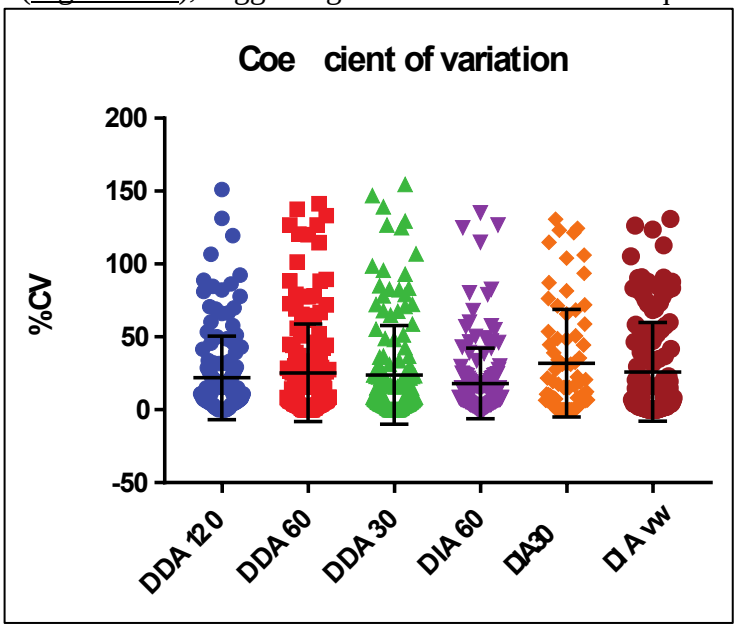


Figure 3.11: Coefficient of variation.
The coefficient of variation was calculated for all of the quantified PTM in each of the different data acquisition methods for both day 2 and day 4 samples (n=6). As expected the PTM with very low relative abundance had greater variation. Overall the DIA 60 had a trend towards lower CVs. n=6 mean and standard deviation represented.

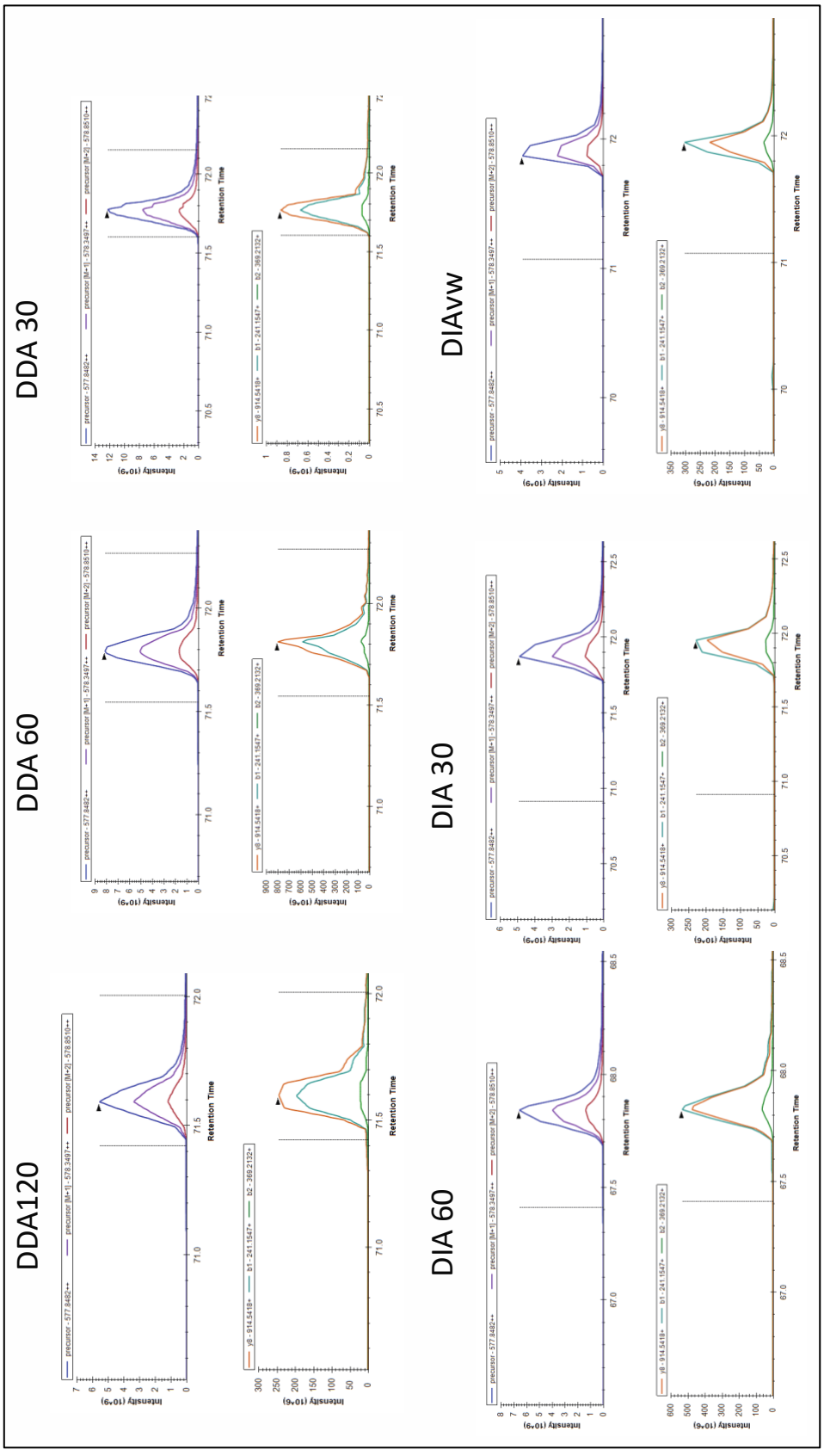


Figure 3.12: Extracted ion chromatograms.
This figure illustrates the extracted ion chromatogram for each of the different data acquisition methods for the unmodified KQLATKAAR peptide of H3. The extracted ion chromatograms are of similar quality across the different data acquisition methods and show an average elution time of 30s.

When comparing the reproducibility in the nanoLC and retention time between

each run, there was excellent reproducibility (CVs <1% between replicates) demonstrating that the variability in correctly quantifying the proteoforms is due to the lower number of MS1 scans. Furthermore, the chromatography for each proteoform was comparable between each data acquisition method (Figure 3.12). Typically, proteoforms are seen eluting over 30 seconds this would equate to between 6 MS1 scans in DDA120 and up to 20 in DDA30 due to the shorter cycle time. Furthermore, the DIA60 would result in 9 MS1 scans whereas DIA30 would have 14, suggesting that the modest decrease in the CVs is due to improved resolution rather than the number of MS1 scans.

3.3 Conclusions:

In this chapter, a number of different data acquisition methods were compared on a QExactive HF Orbitrap mass spectrometer for the identification and quantification of histone PTMs. The 1D LC MS workflow analysed on the QE identified more histone PTMs when compared to the 2D LC MS workflow in conjunction with the maXis, demonstrating the advantages of this instrument for the analysis of histone PTMs. Furthermore, the relative abundances of both the 2D LC MS and 1D LC MS analysis was identical thereby validating this workflow on the QExactive HF Orbitrap. Both data dependent and data independent methods were used to analyse changes in relative abundance of histone PTMs in CHO cells. This approach was able to identify 71 histone proteoforms for histone H3 and H4 and quantified 64 across each of the different acquisition methods.

The advantages of DDA mean that the confidence in correctly identifying and quantifying PTMs can be achieved with lower resolution MS2 scans when coupled with search engines such as Mascot. Indeed, the lower resolution DDA30 method was associated with the greatest number of PSMs ($p<0.001$; Figure 3.8), with equal ability to obtain high ion peptide scores following Mascot searches than the higher resolution scans. However, the advantages of DIA methods over DDA, namely the ability to accurately apportion relative abundances to isobaric co-eluting proteoforms and the fact that they offer greater flexibility to re-search data for novel PTMs, outweigh any disadvantages incurred by the technique. Moreover, the DIA60 was associated with increased reliability in terms of lower CVs for the relative abundance of PTMs as compared to the other data acquisition methods. Therefore, these results demonstrate that this approach is the optimum method from the range of methods studied and will be utilized for all subsequent analysis of histone PTMs on the QExactive HF Orbitrap.

Chapter 4: The impact of pneumolysin on the Epigenomic landscape

4.1 Introduction:

Pneumolysin is one of the key virulence factors of *S. pneumoniae* (Kadioglu et al., 2008b). It is present in the majority of clinical isolates causing IPD (Gray et al., 1980; Hu et al., 2015). In murine models of bacteraemia pneumolysin sufficient mutants are associated with increased lethality compared to pneumolysin deficient mutants, linking the toxin to virulence (Benton et al., 1995). Pneumolysin has been shown to be responsible for the differential expression of multiple genes in undifferentiated THP-1 cells (Rogers et al., 2003) but its impact on gene expression in primary MDMs has not yet been established. Furthermore, the transmission of *S. pneumoniae* between hosts has been linked to the presence of inflammation in the nasopharynx and pneumolysin has been shown to promote inflammation, increase transmission and foster the survival *ex vivo* of the *S. pneumoniae* (Zafar et al., 2017). It has been also been suggested that pneumolysin facilitates blood stream invasion by *S. pneumoniae* (Hu et al., 2015). This highlights the importance of pneumolysin as a key virulence factor of *S. pneumoniae* due to its role in the transmission of *S. pneumoniae* between hosts, in the progression from nasopharyngeal colonisation to IPD, the stimulation of inflammation and pneumolysin's cytotoxic effects (Kadioglu et al., 2008b).

The purpose of this chapter is to establish the pneumolysin-dependent changes in the host's response following challenge with *S. pneumoniae*. As a key bacterial virulence factor, pneumolysin may be responsible for epigenetic modification and these in turn influence key effector functions with consequences for the innate immune response. It is proposed to study pneumolysin-dependent perturbation of the transcriptome, proteome and epigenetic level in primary MDMs. It is planned to study differential gene expression in primary MDMs using microarrays, and differential protein expression using label-free quantitative proteomics. Next, having established the optimal MS method to study histone PTMs (Chapter 3), this method will be used to describe pneumolysin-dependent changes in relative abundance of histone PTMs. Finally, the transcriptomics and proteomics datasets will be integrated in order to tease out greater understanding of the host's response. Chapter 5 will focus on the integration of the histone PTM profile with the associated transcriptomic data sets in order to draw out conclusion regarding the role played by the histone PTMs in the host-pathogen interaction.

In this chapter I set out to test the following hypothesis:

Is pneumolysin responsible for changes in histone PTMs and is this associated with changes in effector functions such as transcriptomic and proteomic profiles?

4.2 Results:

Changes in histone PTMs have been shown to occur as early as 20 minutes following exposure to listeriolysin (LLO) (Hamon et al., 2007). Therefore, initial experiments were performed to determine the earliest time at which MDMs changed their activity in response to bacterial challenge, initially by measuring release of pro-inflammatory cytokines following exposure to *S. pneumoniae* or its pneumolysin deficient mutant (ΔPLY). In order to assess the role played by pneumolysin in the challenge of MDMs I used an isogenic mutant of the D39 parent strain which has a STOP codon inserted at the beginning of the pneumolysin (PLY) gene resulting in no

pneumolysin being produced. This mutant will be referred to throughout the rest of the chapter as ΔPLY.

4.2.1 Model of infective challenge in MDMs:

In order to establish the transcriptional, epigenetic and proteomic changes, MDMs were challenged with opsonised *S. pneumoniae* serotype 2 stain D39 or the isogenic mutant lacking pneumolysin (ΔPLY) or mock infected.

4.2.2 Pro-inflammatory cytokine release during infection:

In order to establish the earliest time point at which the host's response is modified by the bacterial challenge, cytokine production was measured establishing the kinetics of the host cell's functional response. Supernatants were collected at time 0, 0.5, 1, 2, 3, 4, 5, 6 and 7 hours following challenge with both strains. The levels of the pro-inflammatory cytokines TNF-α and IL-6 were measured in the supernatants using ELISA. The results presented in Figure 4.1 show that the release of TNF-α and IL-6 is detectable as early as 3 hours following exposure to bacteria. The data shows that at 4 hours following exposure to bacteria the MDMs have recognised the pathogen and started releasing pro-inflammatory cytokines in response to either strain ($p<0.001$). The pneumolysin-deficient strain is associated with significantly higher amounts of TNF-α released than isogenic parent strain at the 4 hour time point. This may be due to the increased variability seen in response to *S. pneumoniae* and the small number of replicates (n=3).

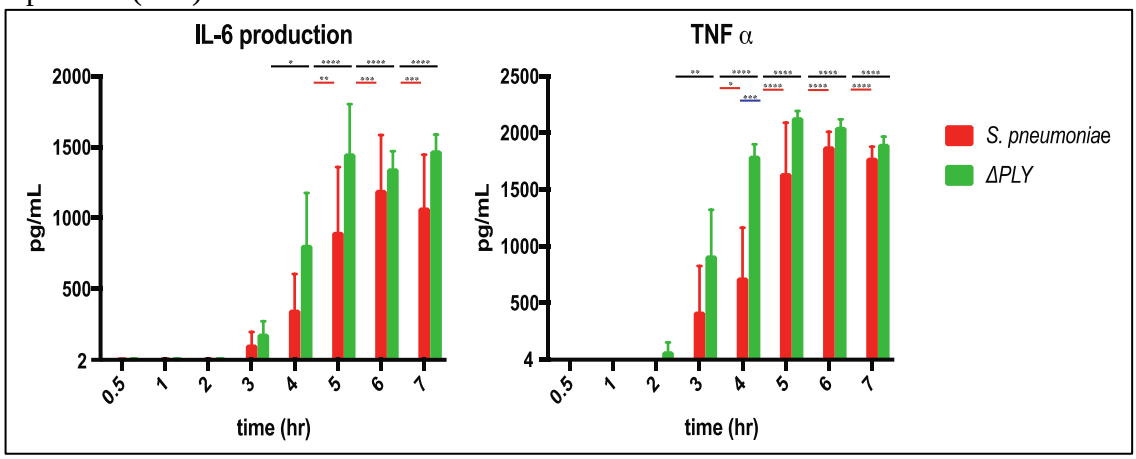


Figure 4.1: Pro-inflammatory cytokine release following challenge with S. pneumoniae with or without pneumolysin.

Alterations in production of TNF-α and IL–6 over time. Mean and Standard Error of the Mean (SEM) of 3 biological replicates run in technical duplicates are shown. There is a significant difference in the rise of TNF-α between the two strains at 4 hr, (black line highlights differences between control and ΔPLY, red line between control and *S pneumoniae*, blue line *S pneumoniae*, and ΔPLY, * $p<0.05$, ** $p<0.01$, **** $p<0.0001$).

4.2.2 Intracellular estimation of viability:

In order to ensure that both strains of *S. pneumoniae* are internalised by MDMs to similar extents and therefore provide comparable numbers of intracellular bacteria to stimulate alterations in the proteome and transcriptome and host epigenome, the number of viable intracellular bacteria at 3 hours in cells challenged with either strain were

measured. The results show that both strains were readily phagocytosed by the MDMs (Figure 4.2) and there is no significant difference in the number of viable bacteria 3 hours following challenge with *S. pneumoniae* or ΔPLY.

Therefore, following 3 hours of exposure, MDMs have ingested bacteria in similar numbers and started to release pro-inflammatory cytokines. These results confirm that 3 hours is a suitable time point to use for the study of early changes in abundance of PTMs of histones. Furthermore, a similar time point has been used to study transcriptomic effects of pneumolysin in undifferentiated THP-1 cells (Rogers et al., 2003) and changes in response to exogenous pneumolysin have been demonstrated to occur as early as 20 min following exposure (Hamon et al., 2007). Therefore, this time point should enable me to elicit the impact of pneumolysin on the epigenome, transcriptome and proteome.

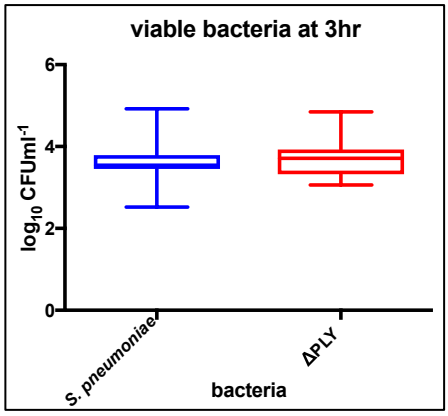


Figure 4.2: Intracellular bacteria following 3 hour challenge with S. pneumoniae or ΔPLY mutant.
14 day old MDMs were challenged with either *S. pneumoniae* or the isogenic pneumolysin negative mutant ΔPLY. At 3 hours the cells were lysed. The lysates were then plated out to count the number of viable intracellular bacteria. The number of bacteria in the wash were subtracted from the lysates and the results expressed as $\log_{10}$ of cfu/mL and represented as box plots with min and max whiskers. (n=8, paired t test p=0.66)

Part A. Studying the effect of *S pneumoniae* infection on the transcriptomic response in MDMs.

4.3 Establishing pneumolysin-dependent differential gene expression in MDMs:

It has been previously shown that in undifferentiated THP-1 cells (a human monocytic cell line) differential gene expression in a pneumolysin-dependent manner was observed (Rogers et al., 2003). These included a number of immune response genes such as macrophage inflammatory protein 1 β (MIP-1β), mannose binding lectin 1, IL-8 and prostaglandin E synthase. To date it is not known whether these pneumolysin-dependent transcriptional differences occurred in primary human macrophages or not. In order to answer this question and establish if any of these transcriptional differences occurred in a pneumolysin-dependent manner, a transcriptome wide analysis was performed following exposure of MDMs to either the parent strain D39 or ΔPLY to quantify changes in mRNA expression.

Three hours following the infection of MDMs with either D39 or ΔPLY, mRNA was extracted and hybridised to the Human Genome U133 Plus 2.0 Array affymetrix Chip for analysis, that is comprised of 54 675 probe sets which according to the manufacturers literature gives almost complete coverage of the human protein coding genes, was carried out by the University of Sheffield's core facility (as described in Chapter 2). Data analysis was performed in R. Briefly, quality control of each sample was estimated using AffyBatch function from the simpleaffy package (version 2.52.0) (quality control plots Appendix figure 4.1.1-3). The probes whose intensity fell within the lowest 20th centile were removed (as these often correspond to noise) leaving 39 344 (of 54 675) probe-sets. Probes-sets were then background corrected using gcRMA (Figure 4.3).

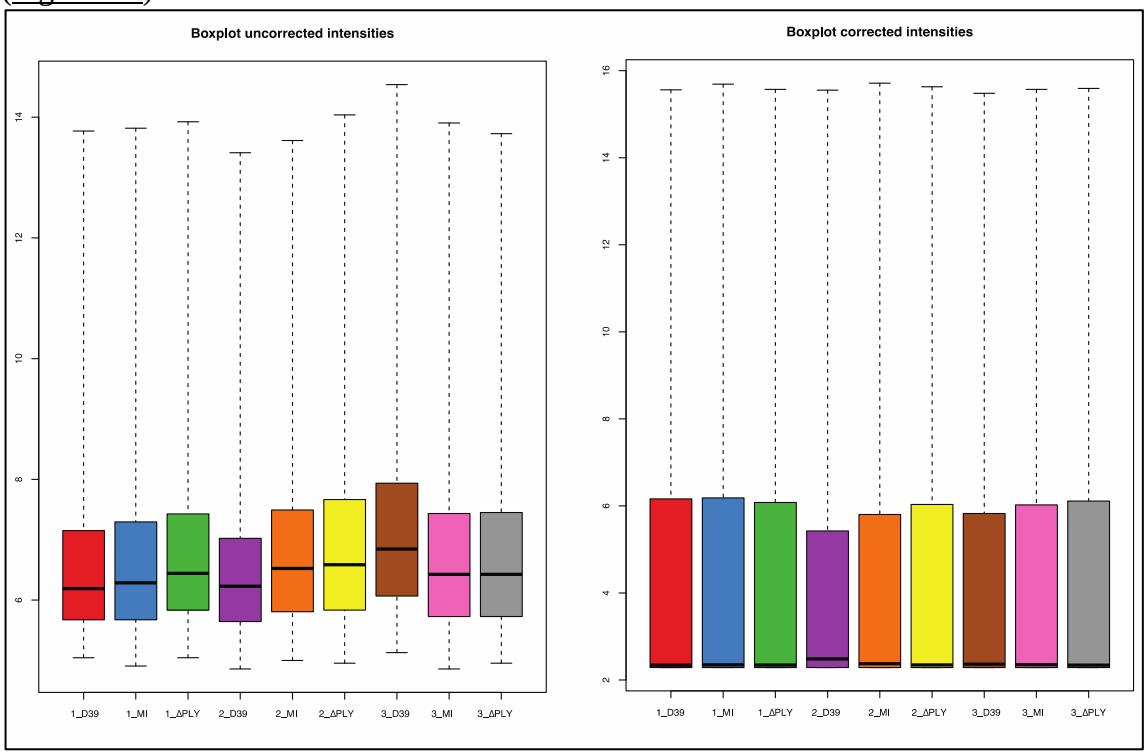


Figure 4.3: Boxplots of microarray samples before and after background correction.

These boxplots represent the intensities of all of the probes on each microarray before and after gcRMA background correction.

Next, I established the statistically significant differences between the probe-sets due to differences between donors (using cut-off of adjusted p value <0.05, from the moderated F statistic test for each probe). This resulted in 34 probe-sets which were excluded from subsequent analysis (Appendix table 4.1).

Then, the differences between the probe sets due to the different conditions (mock infected (MI), *S. pneumoniae*, ΔPLY) were established (using a cut-off of adjusted p value <0.05, from the moderated F statistic test for each probe). This resulted in 1 872 probe-sets whose expression changed significantly as a result of either bacterial strain used in the challenge. "Next, in order to look for the differentially expressed genes between the MI and *S. pneumoniae* challenged cells within those 1 872 probes, adjusted p values (to account for multiple test correction, using FDR) were calculated from the moderated t statistics using an empirical Bayes method in limma. This has become one of the key approaches for the analysis of microarray experiments (Ritchie et al 2015) as it allows detection of small differences between sample, especially if n

number is small, by taking into account all of the expression data from a replicate to estimate the variance for each gene and compute a t statistic which in turn is used to calculate and adjusted p value. Next, the comparison between the MI and ΔPLY challenged cells was performed in the same manner. Briefly, adjusted p values were calculated from the moderated t statistic giving rise to a list of differentially expressed genes. Finally, the comparison between *S. pneumoniae* and ΔPLY challenged cells was performed to determine which probe-sets were differentially expressed in a pneumolysin-dependent and independent manner."

The fold changes for each comparison was also performed and Volcano plots were drawn (Figure 4.5) illustrating the fold change and $\log_{10}$ p value. This resulted in 1 596 probe-sets with a p value <0.05 and 1 243 after multiple test correction (p<0.05 and false discovery rate (FDR <0.05; Appendix table 4.2). Figure 4.4 illustrates the overlap in differentially expressed probe-sets. There are 503 which are differentially expressed in a pneumolysin-dependent manner and 234 in an independent manner.

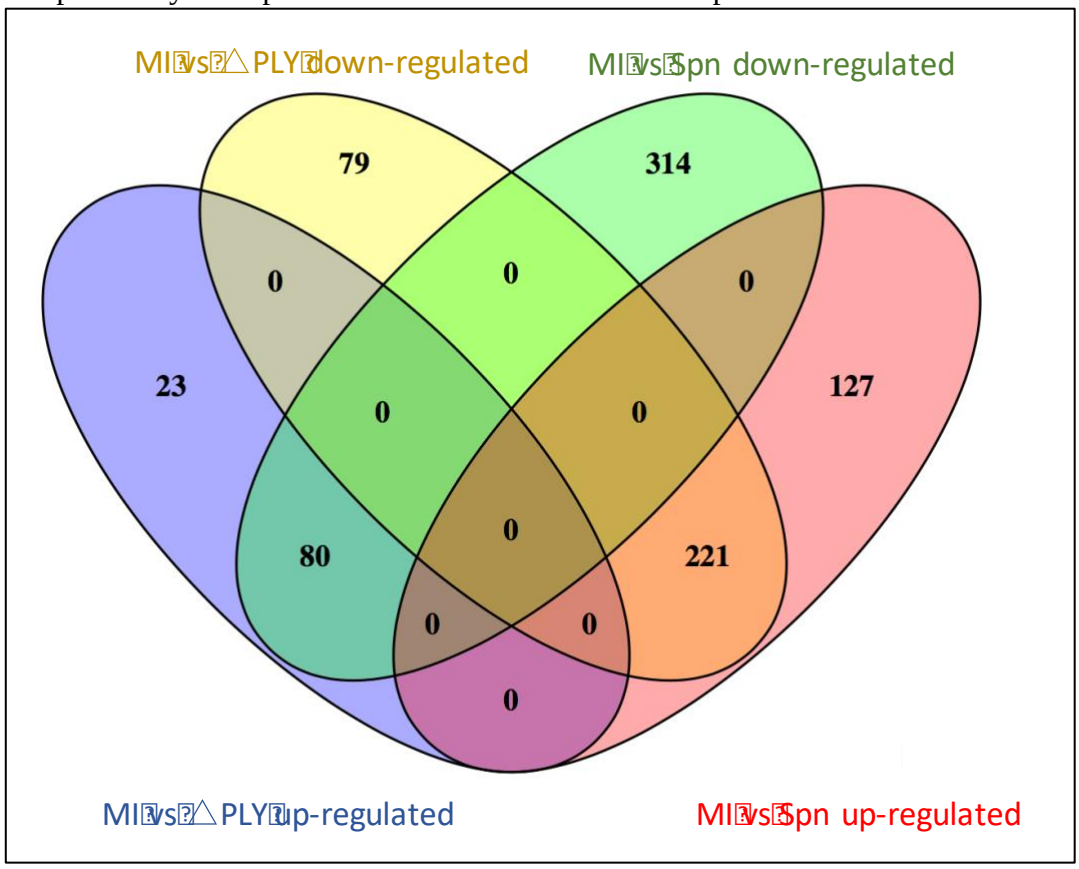


Figure 4.4: Butterfly plot of differentially expressed probe-sets greater than 2 fold change.

The butterfly plot illustrates the differentially expressed probe-sets with an adjusted p value of <0.05 and an absolute fold change greater than 2. There are 742 probes differentially expressed in response to challenge with Spn, of which 348 were up-regulated and 394 were down-regulated. In response to challenge with the mutant strain there were 403 probes with 103 up-regulated and 300 down regulated. There were 221 probes whose expression was up-regulated in response to challenge with Spn and down regulated in the ΔPLY mutant challenge. There were also 80 probes whose expression was increased in the absence of PLY and decreased in the challenge with Spn. n=3, adjusted p value <0.05, fold change >2 or <-2.

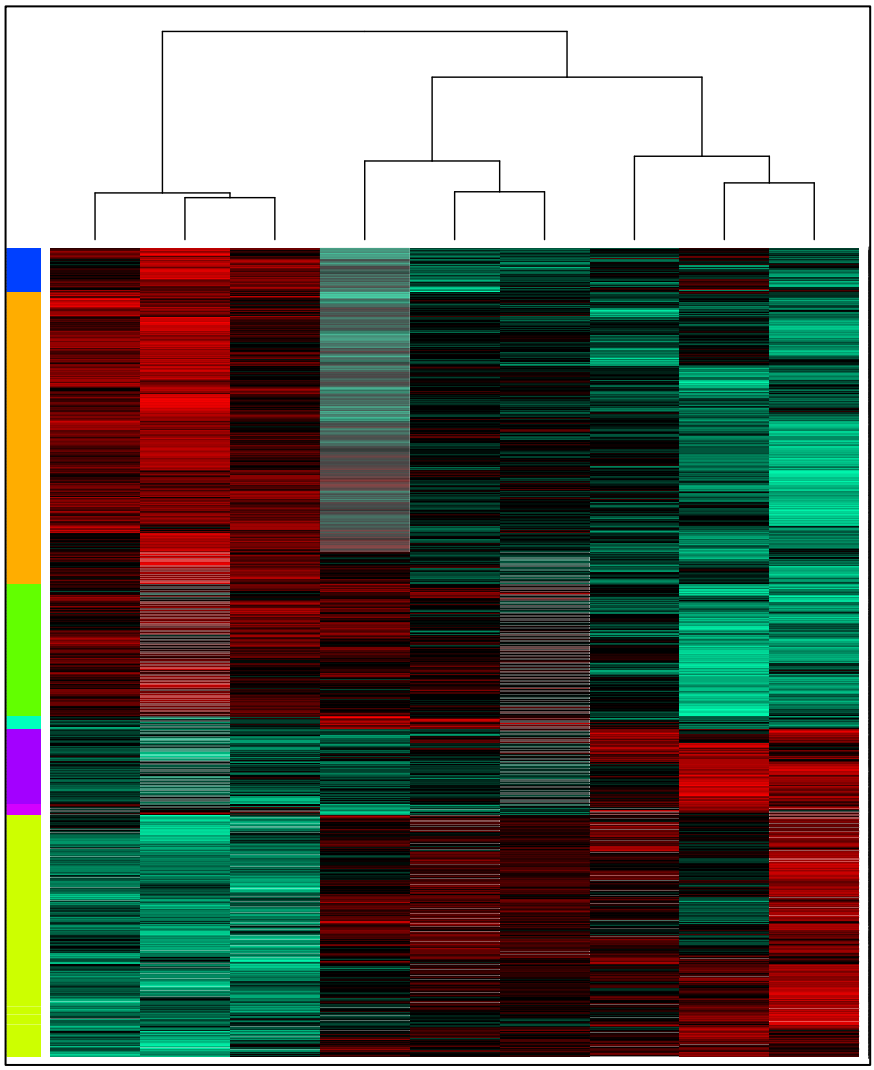


Figure 4.4.1 Heatmap of the 1 872 probes identified by ANOVA as being significantly different.
The first three replicates correspond to the mock infected samples, then the next three are the samples challenged with the mutant ΔPLY, and finally the last three samples are those challenged with *Streptococcus pneumoniae.*

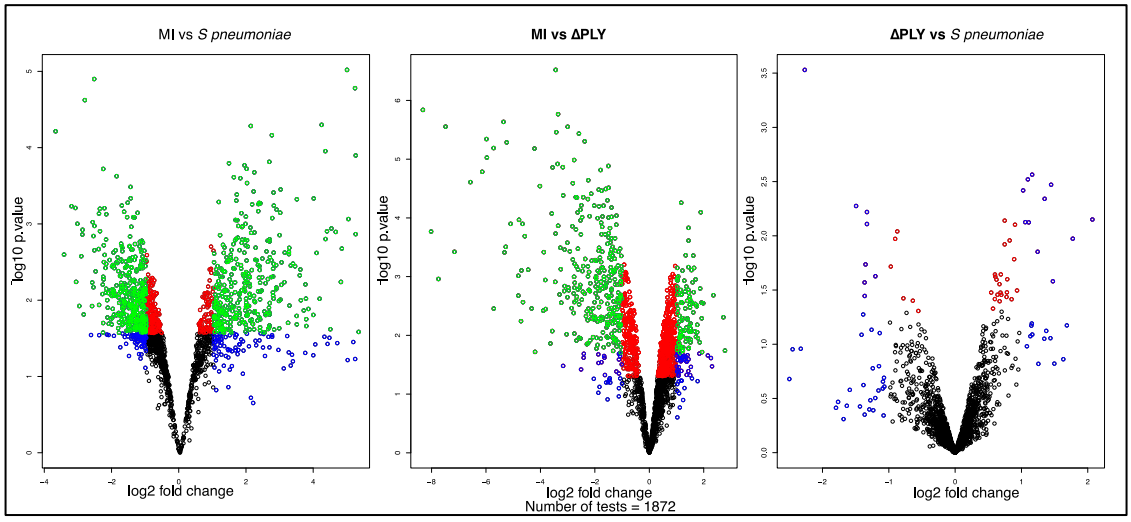


Figure 4.5: Volcano plots of the differentially expressed probe-sets.
These volcano plots illustrate the comparisons between either mock infected (MI) and *S. pneumoniae*, MI and ΔPLY or between ΔPLY and *S. pneumoniae*. The probe-sets in blue have a p value <0.05, in red an adjusted p value <0.05 (following FDR correction) and in green an adjusted p value <0.05 and an absolute fold change greater than 1. There are more probe-sets that are down regulated in response to infection (64% in each case).

These results are consistent with previous research in THP-1 cells that showed 142 genes to be differentially expressed in a PLY dependent manner and 40 to be PLY independent (Rogers et al., 2003). This highlights that pneumolysin elicits a strong host response with more probe-sets being differentially expressed than following challenge with the ΔPLY mutant.

4.3.2 Bioinformatic analysis of differentially expressed genes:

In order to further analyse the transcriptional differences between each condition, the probe-sets were converted to Ensembl identifiers and pathway enrichment analysis was performed on the list of differentially expressed genes (Appendix table 4.2) for both comparisons.

4.3.2.1 Gene ontology:

The list of differentially expressed genes for both comparisons was used to assess which of the gene ontology (GO) pathways are enriched using GoStats package to perform the hypergeometric test for Molecular function, Biological Process and Cellular component.

4.3.2.1.1 Cellular component:

The top ten GO Cellular component terms were similar between both sets of differentially expressed genes (Figure 4.6). There were more terms enriched in the pneumolysin intact strain than the mutant ΔPLY strain (61 vs 35). This is likely to be secondary to the greater number of differentially expressed genes in the *S. pneumoniae* infection.

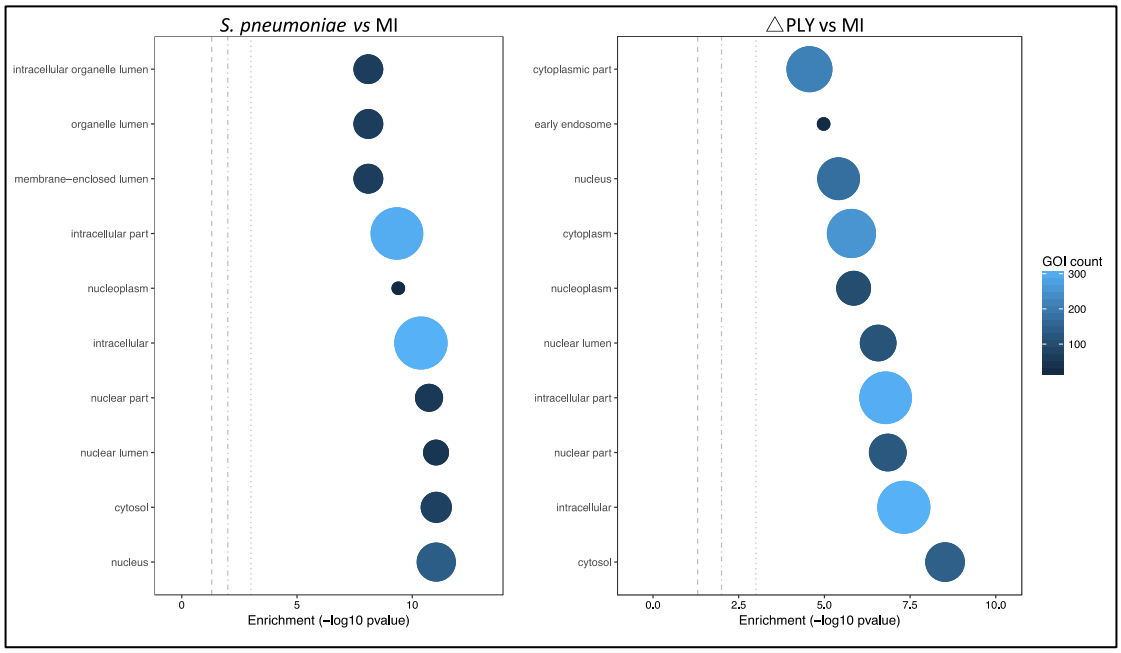


Figure 4.6: Gene Ontology Cellular component term enrichment.
This figure shows the top ten enriched GO cellular component terms in both the pneumolysin mutant and the parent strain analysis. The bubble size and colour correspond to the number of genes that have been identified as belonging to that GO term. The bubbles are plotted along the x axis according to the $-\log_{10}$ p value for the enrichment.

4.3.2.1.2 Gene ontology biological processes:

Next the enrichment for the GO biological process terms were examined. This revealed that the top ten most enriched terms in both cases were predominantly related to cell metabolism (Figure 4.7). However, there were also a number of terms relating to cellular responses to "stress" (17 in *S. pneumoniae* challenge and 15 in the ΔPLY) and in particular to oxidative stress responses. The host's oxidative stress responses have been highlighted as playing a key role in the host-pathogen interaction following infection with *S pneumoniae* in lung epithelial cells (Zahlten et al., 2015). Furthermore, nuclear factor erythroid 2 (NRF2), the master regulator of antioxidant responses, plays a pivotal role in the protection against lung injury (H. Zhao et al., 2017). Indeed, NRF2 knockout mice had higher mortality rates in response to lethal intraperitoneal LPS injections or CLP than there NRF2 positive counterparts (Thimmulappa et al., 2016). Moreover, this was associated with greater inflammatory response in the lungs of NRF2 knockout mice following intra-tracheal LPS instillation. Therefore, the oxidant stress response pathways are pivotal in the response to infections through the regulation of inflammatory responses.

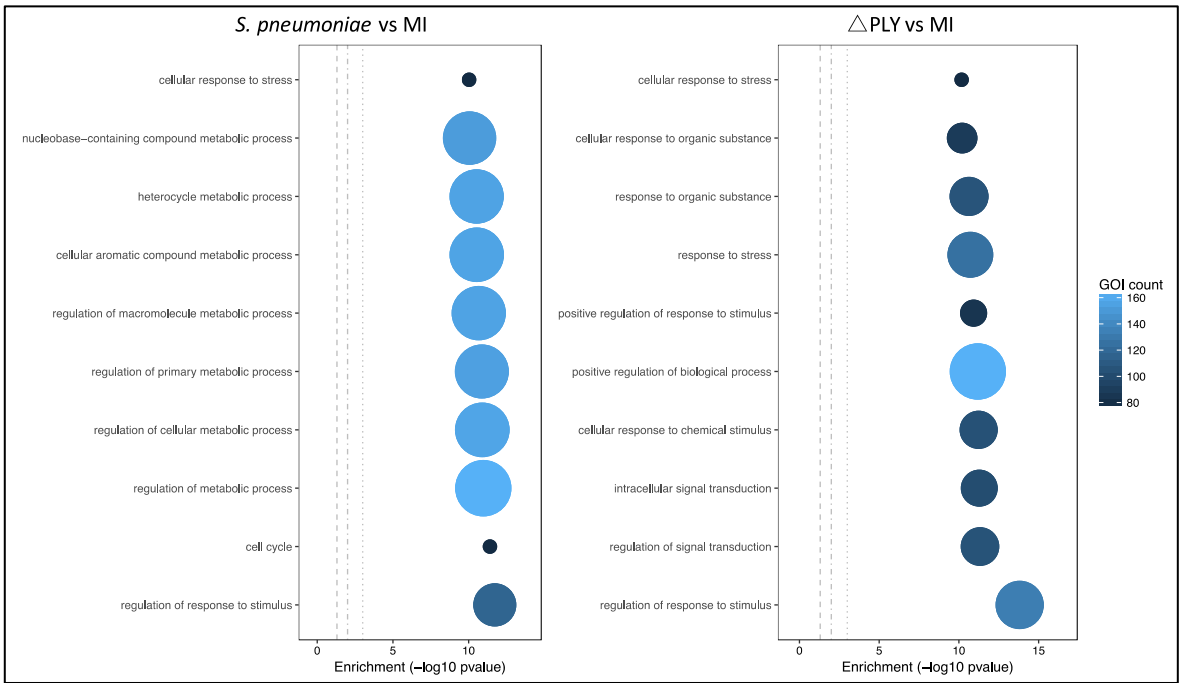


Figure 4.7: Gene ontology biological processes enriched terms.
This figure shows the top ten enriched GO biological processes terms in both the pneumolysin mutant and the parent strain analysis. The bubble size and colour correspond to the number of genes which have mapped to the GO term. The bubbles are plotted along the x axis according to the –log10 p value for the enrichment.

The volcano plots for the differentially expressed genes belonging to the GO term for oxidative stress response were then plotted for each strain (Figure 4.8). These results highlighted that the differentially expressed genes in the pneumolysin-dependent manner were predominantly upregulated whereas the comparison involving the pneumolysin deficient mutant showed these were downregulated. Indeed, the TNF, HMOX1 and PTGS2 genes were strongly up-regulated in a pneumolysin-dependent manner and down regulated in the ΔPLY response.

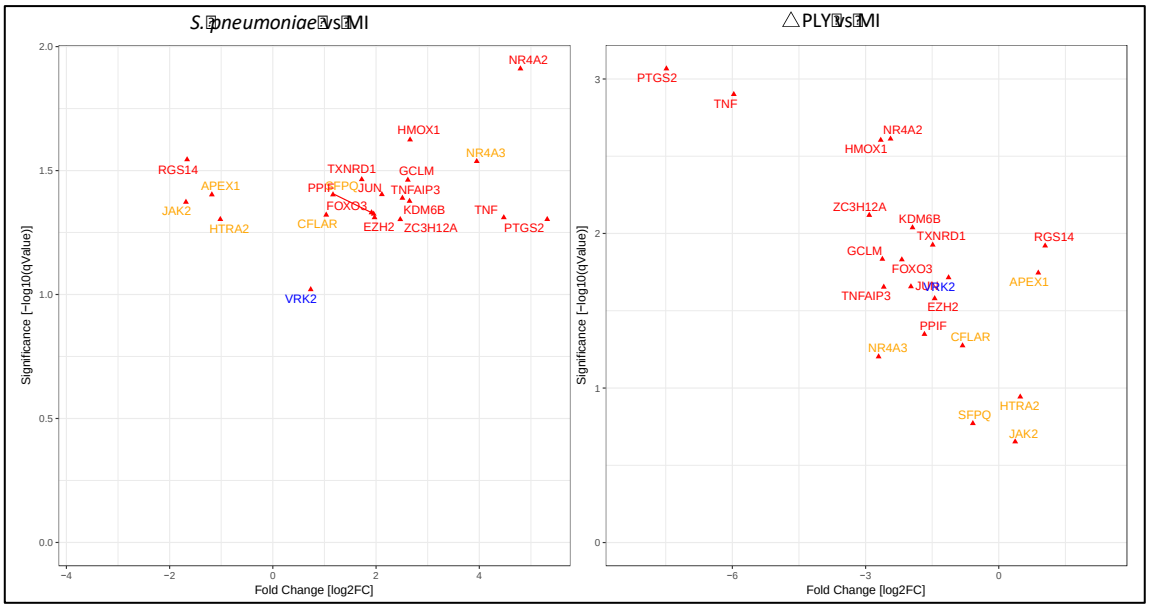


Figure 4.8: Volcano plot of the genes belonging to the oxidative stress response pathways for the responses to infective challenge with either *S. pneumoniae* or ΔPLY.

This figure represents two volcano plots with the changes in the gene expression for the members of GO term for oxidative stress response pathway, compared to MI. The highlighted genes are found within the significantly expressed list (q value <0.05). In red are the gene names that are found in challenge with both strains, in blue the gene that is only found in the response to ΔPLY challenge and in orange the pneumolysin-dependent genes.

4.3.2.1.3 Gene ontology Molecular Function terms:

The GO enriched terms belonging to the molecular functions were also evaluated (see Figure 4.9). This highlighted predominantly "binding" and "nuclear functions" in response to both infections.

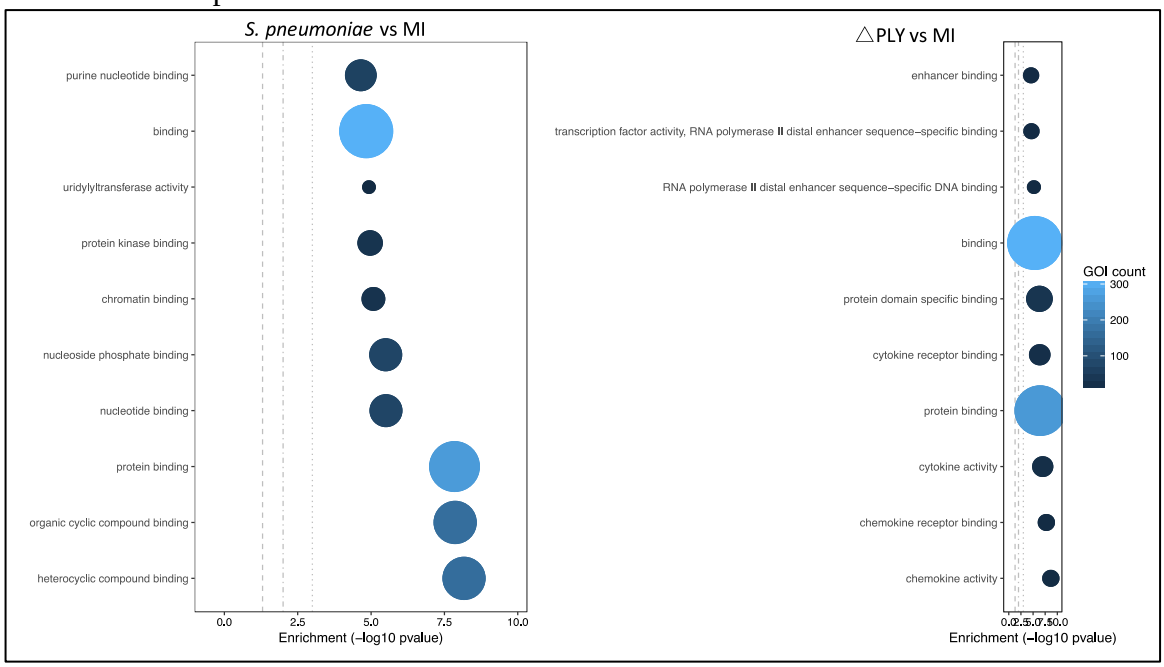


Figure 4.9: Gene Ontology Molecular Functions Enriched terms.

This figure shows the top ten enriched GO biological processes terms in both the pneumolysin mutant and the parent strain analysis. The bubble size and colour correspond to the number of genes that have mapped to the GO term. The bubbles are plotted along the x axis according to the –log10 p value for the enrichment.

4.3.2.2 Kyoto encyclopaedia of genes and genomes enrichment:

The Kyoto encyclopaedia of genes and genomes database (KEGG) was also used to perform pathway enrichment searches for the differentially expressed genes as a result of challenge with either strain (see Table 4.1). In both comparisons the TNF-α pathway and the NF-Kappa B signalling pathway were enriched in response to infective challenges with either *S. pneumoniae* or ΔPLY. In addition, in response to the ΔPLY challenge the NOD-like receptor signalling, MAPK signalling pathway and Toll-like receptor signalling pathway were also significantly enriched.

Overall the transcriptomic analysis has demonstrated a number of innate immune responses to be enriched following challenge with *S. pneumoniae.* Indeed, the TNF-α signalling pathway, the NFKB pathway and the oxidative stress response are differentially expressed. Furthermore, there are more genes that are differentially expressed in a pneumolysin dependent manner. The gene ontology analysis of these

demonstrated that there were several metabolic pathways enriched than in the pneumolysin independent analysis.

Table 4.1: Summary of significantly enriched KEGG pathways relevant to infection.

	MI vs *S. pneumoniae*		MI vs △PLY	
Pathway	p.val	q.val	p.val	q.val
MAPK signalling pathway	2.18E-04	5.20E-02	**1.36E-06**	**3.16E-04**
ErbB signalling pathway	**1.53E-04**	**3.65E-02**	8.82E-03	1.00E+00
Cytokine-cytokine receptor interaction	1.22E-03	2.93E-01	**2.27E-08**	**5.29E-06**
Chemokine signalling pathway	1.85E-02	1.00E+00	**1.68E-08**	**3.92E-06**
NF-kappa B signalling pathway	**1.51E-09**	**3.60E-07**	**6.67E-13**	**1.55E-10**
Apoptosis	**2.80E-08**	**6.68E-06**	**1.76E-08**	**4.10E-06**
Necroptosis	**2.78E-07**	**6.63E-05**	**6.05E-06**	**1.41E-03**
Toll-like receptor signalling pathway	2.66E-03	6.35E-01	**4.40E-08**	**1.03E-05**
NOD-like receptor signalling pathway	**2.12E-06**	**5.06E-04**	**2.10E-07**	**4.90E-05**
TNF signalling pathway	**8.34E-09**	**1.99E-06**	**1.93E-15**	**4.50E-13**
Fluid shear stress and atherosclerosis	**2.05E-04**	**4.90E-02**	**1.50E-07**	**3.49E-05**

4.3.3 XGR enrichment analysis:

In addition to performing pathway enrichment of genes using NIPA (see chapter 2), I also repeated the analysis of the differentially expressed genes using XGR which performs the similar hypergeometric enrichment analysis but can also perform background correction using the expression of monocyte derived macrophage cells to give a more cell type specific analysis.

Initially the canonical pathway analysis was performed for genes whose expression was up-regulated in response to challenge with Spn (adjusted p value <0.05). This revealed 32 over-represented pathways (Table 4.1.1). Importantly this demonstrated that both the TNF and the NFKB signalling pathways were enriched. This mirrors the analysis seen without background correction looking at KEGG and GO biological process enrichment analyses.

Analysis of the 2 fold down-regulated terms did not reveal any enriched canonical pathways.

Table 4.1.1 Canonical pathway analysis of up-regulated differentially expressed probes following challenge of MDMs with *Streptococcus pneumoniae* using XGR.

Term Name	FDR	Genes
AP-1 transcription factor network	0.0004	ATF3, CXCL8, FABP4, FOSL2, JUN, MAFG, NR3C1
Calcineurin-regulated NFAT-dependent transcription in lymphocytes	0.0011	CREM, CXCL8, JUN, PTGS2, TLE4, TNF
Validated transcriptional targets of TAp63 isoforms	0.0011	CDKN1A, JAG1, PERP, PMAIP1
TNF receptor signaling pathway	0.0015	MAP4K5, NSMAF, SQSTM1, TNF, TNFAIP3, TRAF1
Downstream signaling in naive CD8+ T cells	0.0029	B2M, JUN, STAT4, TNF, TNFRSF4, TNFRSF9
Validated targets of C-MYC transcriptional repression	0.0029	CDKN1A, CFLAR, FOXO3, WNT5A
Direct p53 effectors	0.0029	ATF3, CDKN1A, DDIT4, DUSP5, JUN, PERP, PMAIP1, PRDM1
Regulation of nuclear SMAD2/3 signaling	0.0029	ATF3, CDKN1A, FOXO3, JUN, NR3C1, RUNX2, SKIL, TGIF1
ATF-2 transcription factor network	0.0029	ATF3, CSRP2, CXCL8, DUSP5, JUN
Genes related to Wnt-mediated signal transduction	0.0029	B2M, CSNK1A1, JUN, TLE1, WNT5A
Validated transcriptional targets of AP1 family members Fra1 and Fra2	0.0049	CXCL8, FOSL2, HMOX1, JUN
Canonical NF-kappaB pathway	0.0074	TNF, TNFAIP3, UBE2D3
Presenilin action in Notch and Wnt signaling	0.0074	CSNK1A1, JUN, TLE1
Regulation of retinoblastoma protein	0.0084	BRD2, CDKN1A, JUN, RUNX2
Tumor Necrosis Factor Pathway.	0.0084	CFLAR, JUN, TNF, TNFAIP3
Genes encoding secreted soluble factors	0.012	CLCF1, CXCL8, IL15, IL1RN, PDGFB, TNF, WNT5A
TGF-beta receptor signaling	0.013	PPP1R15A, SKIL, XIAP, ZFYVE16
Notch signaling pathway	0.013	CDKN1A, JAG1, RBBP8
Signaling mediated by p38-alpha and p38-beta	0.013	JUN, MEF2A, PTGS2
Calcium signaling in the CD4+ TCR pathway	0.024	CREM, JUN, PTGS2
FoxO family signaling	0.024	CSNK1A1, FOXO3, RALA
Validated transcriptional targets of deltaNp63 isoforms	0.024	FOSL2, PERP, RAB38
Caspase cascade in apoptosis	0.032	LMNA, TNF, XIAP
CD40/CD40L signaling	0.032	JUN, TNFAIP3, TRAF1
Ensemble of genes encoding ECM-associated proteins including ECM-aff	0.032	CLCF1, CXCL8, IL15, IL1RN, PDGFB, TNF, WNT5A
HIF-1-alpha transcription factor network	0.032	HMOX1, JUN, PKM
IL1-mediated signaling events	0.032	IL1RN, JUN, SQSTM1
LPA receptor mediated events	0.032	CXCL8, JUN, LPAR2
Notch-mediated HES/HEY network	0.032	RBBP8, RUNX2, TLE1
Glucocorticoid receptor regulatory network	0.034	CDKN1A, CXCL8, JUN, NR3C1
Regulation of nuclear beta catenin signaling and target gene transcription	0.034	CXCL8, JUN, TLE1, TLE4
HIV-1 Nef: Negative effector of Fas and TNF-alpha	0.045	CFLAR, TNF, TRAF1

Next the analysis was repeated looking at the differentially expressed genes whose expression was down-regulated by greater than 2 fold in response to challenge with the ΔPLY mutant. This revealed that the pathways for TNF, NFKB and CD40 signalling were significantly over-represented (Table 4.1.2). This highlights the importance of these pathways in the response to bacterial challenge. It also suggests that pneumolysin is responsible for the increase in pro-inflammatory signals being released and that the host is able to modulate its response to infection by decreasing these pathways in the context of the less virulent organism.

Table 4.1.2 Canonical pathway analysis of down-regulated differentially expressed probes following challenge of MDMs with ΔPLY using XGR.

Term Name	FDR	Genes
Genes encoding secreted soluble factors	6.10E-12	CCL1, CCL20, CCL3L3, CCL4, CLCF1, CXCL1, CXCL2, CXCL3, CXCL8, IL15, IL1A, IL1B, INHBA, KITLG, TNF, WNT5A
Ensemble of genes encoding extracellular matrix and extracellular matrix-associated proteins	8.00E-10	CCL1, CCL20, CCL3L3, CCL4, CLCF1, CRIM1, CXCL1, CXCL2, CXCL3, CXCL8, IL15, IL1A, IL1B, INHBA, KITLG, TNF, TNFAIP6, WNT5A
Ensemble of genes encoding ECM-associated proteins including ECM-affilaited proteins, ECM regulators and secreted factors	2.20E-09	CCL1, CCL20, CCL3L3, CCL4, CLCF1, CXCL1, CXCL2, CXCL3, CXCL8, IL15, IL1A, IL1B, INHBA, KITLG, TNF, WNT5A
TNF receptor signaling pathway	0.0000013	BIRC3, MAP4K5, NFKB1, NSMAF, SQSTM1, TNF, TNFAIP3, TRAF1
Direct p53 effectors	0.000026	AIFM2, ATF3, CDKN1A, DDIT4, DUSP5, JUN, MCL1, PMAIP1, PRDM1, TNFRSF10A
CD40/CD40L signaling	0.000039	BIRC3, JUN, NFKB1, TDP2, TNFAIP3, TRAF1

Calcineurin-regulated NFAT-dependent transcription in lymphocytes	0.00013	CREM, CXCL8, JUN, PTGS2, TLE4, TNF
Validated transcriptional targets of AP1 family members Fra1 and Fra2	0.00017	CXCL8, FOSL2, HMOX1, JUN, THBD
IL23-mediated signaling events	0.00017	CXCL1, IL1B, NFKB1, STAT4, TNF
Ceramide signaling pathway	0.00017	BIRC3, NFKB1, NSMAF, TNF
Canonical NF-kappaB pathway	0.00017	NFKB1, TNF, TNFAIP3, UBE2D3
AP-1 transcription factor network	0.00023	ATF3, CXCL8, FOSL2, JUN, MAFG, NR3C1
IL1-mediated signaling events	0.00039	IL1A, IL1B, JUN, NFKB1, SQSTM1
Tumor Necrosis Factor Pathway.	0.00039	BIRC3, JUN, NFKB1, TNF, TNFAIP3
Validated targets of C-MYC transcriptional repression	0.00065	CDKN1A, FOXO3, TJP2, WNT5A
ATF-2 transcription factor network	0.0009	ATF3, CSRP2, CXCL8, DUSP5, JUN
IL27-mediated signaling events	0.0034	IL1B, STAT4, TNF
Validated transcriptional targets of TAp63 isoforms	0.0034	CDKN1A, JAG1, PMAIP1
HIF-1-alpha transcription factor network	0.0038	HMOX1, JUN, MCL1, PKM
Glucocorticoid receptor regulatory network	0.0055	CDKN1A, CXCL8, JUN, NFKB1, NR3C1
C-MYB transcription factor network	0.0063	BIRC3, CDKN1A, KITLG, PTGS2
HIV-1 Nef: Negative effector of Fas and TNF-alpha	0.0063	BIRC3, NFKB1, TNF, TRAF1
IL12-mediated signaling events	0.0063	CCL4, IL1B, NFKB1, STAT4
Signaling mediated by p38-alpha and p38-beta	0.0068	JUN, MEF2A, PTGS2
Regulation of nuclear SMAD2/3 signaling	0.011	ATF3, CDKN1A, FOXO3, JUN, NR3C1, RUNX2
Validated transcriptional targets of deltaNp63 isoforms	0.013	FOSL2, IL1A, RAB38
Calcium signaling in the CD4+ TCR pathway	0.013	CREM, JUN, PTGS2
Caspase cascade in apoptosis	0.019	BIRC3, LMNA, TNF
Signaling events mediated by HDAC Class I	0.019	MXD1, NFKB1, TNF
LPA receptor mediated events	0.019	CXCL8, JUN, NFKB1
Regulation of retinoblastoma protein	0.019	CDKN1A, JUN, RUNX2
Fas Signaling Pathway	0.02	CSNK1A1, CXCL8, IL1A, NFKB1
Angiopoietin receptor Tie2-mediated signaling	0.029	CDKN1A, NFKB1, TNF
Genes related to Wnt-mediated signal transduction	0.029	CSNK1A1, JUN, WNT5A
Regulation of Androgen receptor activity	0.04	JUN, NR3C1, REL
Regulation of Telomerase	0.04	JUN, MXD1, NFKB1

Finally, the analysis of the up-regulated genes in response to challenge with ΔPLY revealed 4 canonical pathways to be over-represented. Three of these were cell signalling pathways highlighting the role play by these during challenge with bacteria.

Table 4.1.3 Canonical pathway analysis of up-regulated differentially expressed probes following challenge of MDMs with ΔPLY using XGR.

Term Name	FDR	Genes
B Cell Antigen Receptor	0.0049	BLNK, ITPKB, PIK3CD
Members of the BCR signaling pathway	0.0056	BLNK, NFATC1, PIK3CD
Class I PI3K signaling events	0.0056	BLNK, PIK3CD, PLEKHA2
BCR signaling pathway	0.016	BLNK, CAMK2G, NFATC1

Overall, the analysis of the microarray with background correction using XGR re-iterated the importance of the TNF and NFKB pathways in the response to *Streptococcus pneumoniae* infections. It did not show significant difference between in the enriched pathways compared to analysis with NIPA. This is probably due to the fact that the NIPA analysis uses a similar background correction by taking into account all of the genes identified (including those not differentially expressed) during the analysis as part of the "universe" prior to performing its hypergeometric test.

Part B Studying the effect of *S pneumoniae* infection on the proteome in MDMs

4.4 Label-free proteomic analysis:

Having established that there is a transcriptional difference in the host cell response following challenge with *S. pneumoniae* or ΔPLY, further studies were performed to study the effects of *S. pneumoniae* on the proteome of MDMs.

4.4.1 Three-hour following bacterial challenge:

Three hours following the exposure of MDMs with either *S. pneumoniae* or ΔPLY, cells from three biological replicates were lysed and protein quantification carried out (see Chapter 2). Proteins were then subjected to reduction and alkylation prior to trypsin digestion with the use of FASP (see Chapter 2 for further details). Samples were then desalted using either Hypersep^{TM} tips or using hypercarb column and off line HPLC fractionation. The samples were then analysed using mass spectrometry on a Orbitrap QE HF. The RAW files from each condition (mock infected, challenged with *S. pneumoniae* and ΔPLY) in the 3 biological replicates were then searched using MaxQuant to perform the protein identifications and label-free quantification (LFQ) as described in Chapter 2. Data analysis was performed in R. Briefly, the matches to the reverse and the contaminant database were removed, then only the majority protein identifications with more than two unique peptides were kept (Table 4.2 summarises the data filtering process).

Table 4.2: Summary of protein identifications from MDMs following infection with either *S. pneumoniae* or ΔPLY.

The three biological replicates are represented by A, B and C. The number of majority protein identifications at each stage are reported, showing the removal of reverse, contaminant and less than two unique peptide matches.

	Mock Infected	*S pneumoniae*	ΔPLY	

Donor	A	B	C	A	B	C	A	B	C	average
Majority protein identifications	1318	1468	1435	1347	1385	1322	1410	1422	1418	1392
reverse matches removed	1316	1465	1433	1346	1385	1317	1406	1414	1407	1388
contaminant matches removed	1300	1449	1420	1331	1370	1304	1390	1399	1396	1373
>2 unique peptide matches	1102	1232	1200	1150	1217	1130	1132	1172	1097	1159

4.4.2 Label-free quantification:

Following initial protein identifications from the MDMs infected with either *S. pneumoniae* or ΔPLY, further analysis was performed to identify those proteins differentially expressed. Protein quantification was performed using label-free quantification within the MaxQuant software (Cox et al., 2014) and see Chapter 2)

The LFQ intensity was then normalised by the median of the LFQ intensities for that sample and log2 transformed. (Figure 4.10 illustrates the correction of intensities.)

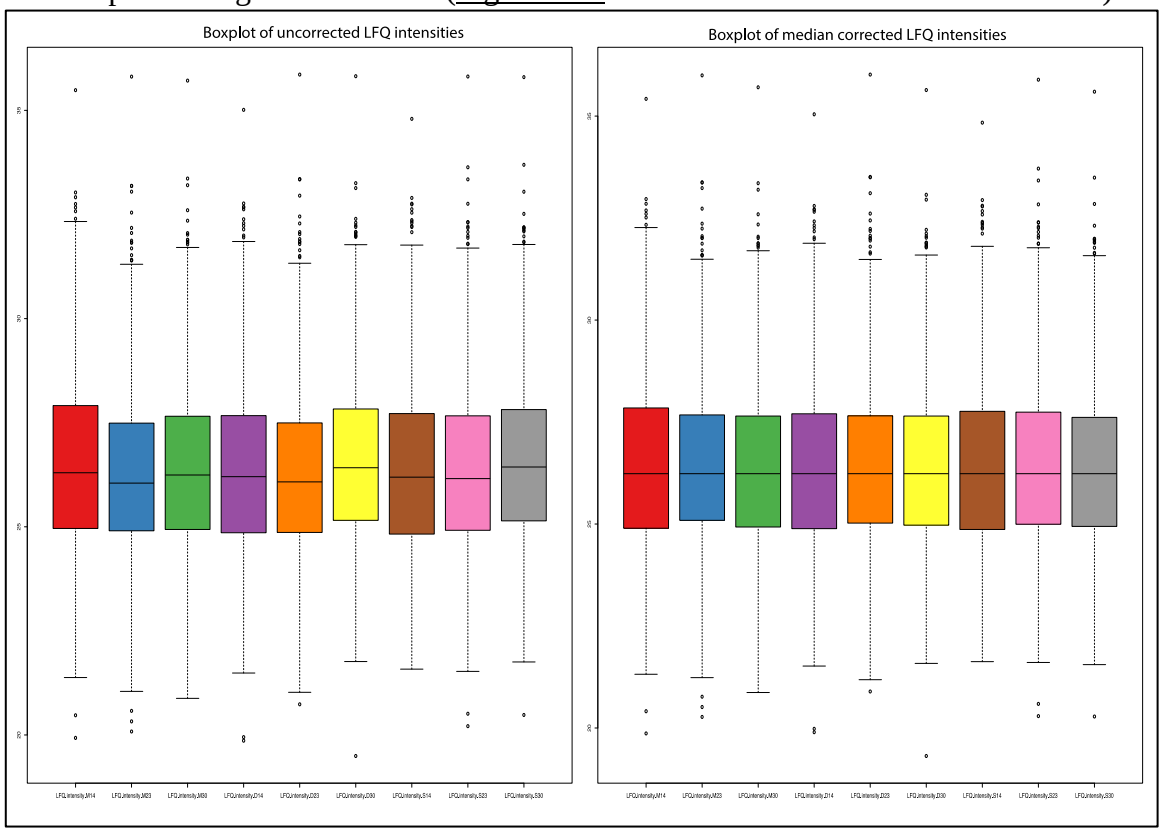


Figure 4.10: Boxplot of LFQ intensities before and after median correction prior to differential analysis.
These boxplots illustrate the log 2 LFQ intensity for the proteins identified before and after the median correction.

Then, a repeated measures ANOVA using Limma package of the LFQ intensities for the remaining majority protein identifications was performed, to enable comparison to the microarray analysis (Goeminne et al., 2016). This showed 5 majority protein identifications with an F statistic value of less than 0.05. The moderated t tests were then calculated for these five proteins and multiple test correction applied (summarised in Table 4.3).

Table 4.3: Differentially expressed proteins at 3 hours following challenge with *S pneumoniae* or ΔPLY.

	MI vs *S pneumoniae*			MI vs PLY				
Majority Protein ID	**log FC**	**p val**	**q val**	**log FC**	**p val**	**q val**	**protein name**	**Gene name**
Q07021; I3L3Q7; I3L3B0	**-1.88**	**0.002**	**0.012**	-1.014	0.133	0.133	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP
P56134; C9JJT5;C9J U26;G3V325	**-1.839**	**0.006**	**0.015**	**-2.801**	**0.010**	**0.017**	ATP synthase subunit f, mitochondrial	**ATP5J2**
V9GYM8; Q92974; Q5VY93	**-0.748**	**0.029**	**0.048**	**-1.089**	**0.005**	**0.011**	Rho guanine nucleotide exchange factor 2	**ARHGEF2**
P12277; G3V4N7	0.945	0.054	0.068	**3.175**	**0.002**	**0.011**	Creatine kinase B-type	CKB
H0Y8C6; O00410	-0.422	0.337	0.337	**-1.941**	**0.029**	**0.036**	Importin-5	**IPO5**

Complement component 1 Q subcomponent-binding protein (C1QBP) is a multifunctional protein involved in inflammation and infection. It has been associated with enabling entry of pathogens such as *Listeria monocytogenes* into cells (Braun et al., 2000). It has also been shown to be involved in complement mediated immune suppression via phosphoinositide-3-kinase (Waggoner et al., 2005). Furthermore, it is a competitive inhibitor of the hyaluronidase of *S. pneumoniae* (Yadav et al., 2009). C1QBP has been demonstrated to localise at the mitochondria, where in mice it has been shown to protect against oxidative stress mediated death (McGEE and Baines, 2011). Conversely, it has also been demonstrated to be pro-apoptotic by allowing calcium influx to the mitochondria, and is inhibited by Mcl-1 (Xiao et al., 2014). In murine sepsis models, C1QBP has led to increased survival associated with decreased release of IL-6 compared to C1QBP deficient mice (Sasaki et al., 2017). Therefore, the decrease in the abundance of C1QBP in response to both bacterial challenges maybe due to the host releasing it to combat the pathogen's hyaluronidase, be seen as a strategy to limit its inhibitory activity on the complement mediated immune response, be associated with the alteration in protection against oxidative stress or reduction in early apoptosis.

Rho guanine nucleotide exchange factor 2 (ARHGEF2) is thought to fulfil numerous roles including innate immune responses. It has been implicated in the detection of intracellular microbial components alongside NOD1 in *Shigella flexneri* invasion (Fukazawa et al., 2008). It has also been shown to regulate the microbial sensing via the NOD2 pathway (Zhao et al., 2012). Thus, the decrease in abundance may be secondary to the detection of intracellular *S. pneumoniae*.

Importin-5 is involved in the transport of ribosomal proteins into the nucleus (Jäkel and Gürlich, 1998). The increase in creatine kinase B abundance may reflect the increased energy demands placed on the host cell by infection. Surprisingly the ATP synthase subunit f, which is involved in the generation of ATP and therefore energy generation, is decreased in response to exposure to both strains.

In light of the paucity of differentially expressed proteins a principle component analysis was performed for the proteins identified in all of the samples (Figure 4.11). The results show that the differences between donors is greater than the effect of the bacterial challenge at this early time point.

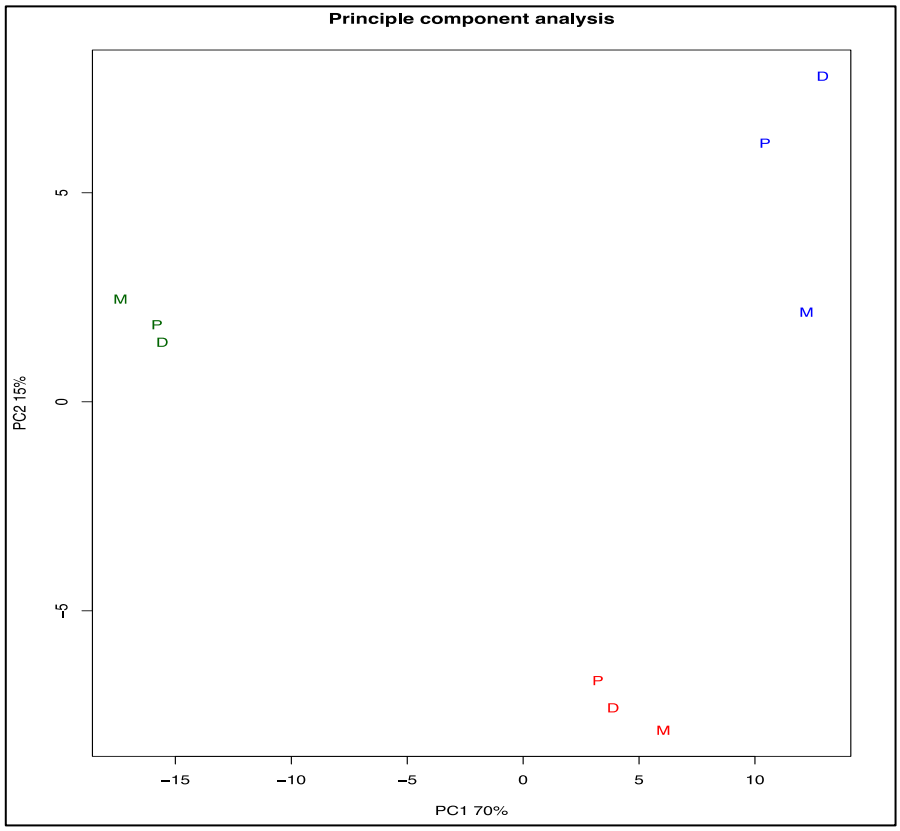


Figure 4.11: Principle component analysis of the proteins identified in all of the runs.

Principle component analysis of the majority protein identification seen in all of the 12 samples. The three donors are coloured red, green and blue respectively. The mock infected samples are represented by the letter M, the *S pneumoniae* challenged samples are represented by the letter D, and the ΔPLY by the letter P. This illustrates that the majority of the difference between the samples is due to differences between donors rather than experimental condition. (n=3)

One possible explanation for the paucity of differentially expressed proteins following bacterial challenge may be that the changes in the cell's proteome are not large enough to be detected by a shotgun proteomic approach at such an early time point. In light of this, the infections were repeated and extended to six hours to see if the later time point gave rise to larger differences in the global proteome. The cells were then collected and the label-free proteome characterised as for the 3 hour time point (described in Chapter 4.4.1; Appendix figure 4.2 shows the LFQ intensities following normalisation, Appendix table 4.2 summarises the data filtering process).

The filtered and median corrected LFQ intensities were then used to detect statistical differences by ANOVA followed by calculating the moderated t tests between the mock infected and either *S. pneumoniae* or ΔPLY. This identified 32 majority protein identifications that were differentially expressed between the three conditions (summarised in Appendix table 4.3). The results show that 16 proteins are differentially expressed in response to infections with *S pneumoniae,* and 22 in response to ΔPLY. Figure 4.12 illustrates the overlap between each bacterial challenge. The results show that 8 proteins were differentially expressed in a pneumolysin-dependent manner and 9 were common to bacterial strains. Given the small number of differentially expressed proteins, pathway enrichment analysis is not appropriate.

Although the absolute number of differentially expressed proteins is small, this is in keeping with a number previously published proteomic studies. In a study of THP-

1 cells infected with *Mycobaterium tuberculosis* only 61 proteins were found to be differentially regulated (P. Li et al., 2017). In a study of alveolar macrophages infected with porcine reproductive and respiratory syndrome virus only 95 proteins were differentially expressed (Qu et al., 2017). Furthermore, this study has used primary cells and the difference in-between donors is such that small differences as a result of treatment may be masked (as illustrated in Figure 4.11).

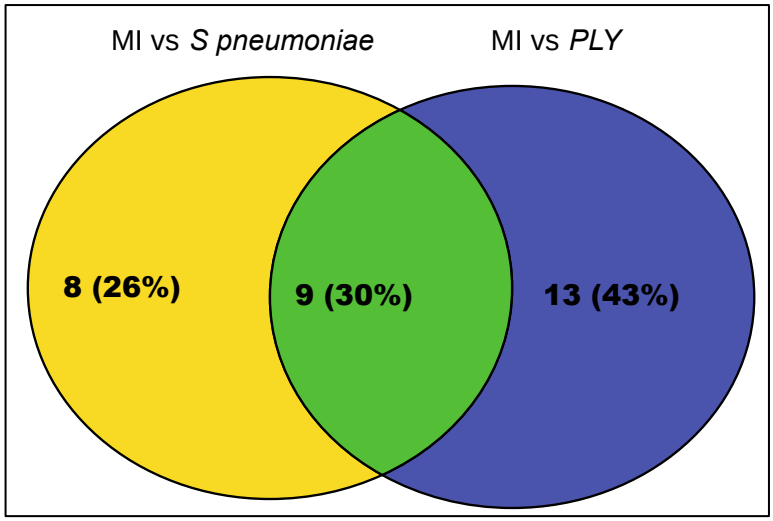


Figure 4.12: Venn diagram of differentially expressed proteins after 6 hour infection. This Venn diagram illustrates the 32 proteins that were identified as differentially regulated following 6 hour challenge with wither *S pneumoniae* or ΔPLY.

Of the pneumolysin-dependent differentially expressed proteins, the serine / threonine protein kinase 1, (also known as oxidative stress response protein 1, OXSR1) has been shown to regulate responses secondary to environmental stress (Chen et al., 2004). This is interesting, as it complements the transcriptomic data where the oxidative stress responses were over-represented in the pathway analysis.

In addition, the probable global transcription activator SNF2L2 / Brahma protein encoded by switch/sucrose non-fermentable (SWI/SNF) related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 2 / 4(SMARCA 2 / 4) genes and the transcription activator Brahma-related gene 1 (BRG1) protein encoded by SMARCA4 are associated with regulation of gene transcription by chromatin remodelling (Wilson and Roberts, 2011). The ubiquitin domain-containing protein 1 (UBTD 1) is associated with the cellular senescence via a positive feedback loop with tumour protein (TP53; p53)(Zhang et al., 2015). The ATP dependent RNA helicase DDX46 (DDX46) is associated with pre-mRNA splicing (Will et al., 2002). The probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase encoded for by the NOP2 gene has been shown to be associated with the assembly of the large subunit of the ribosome and may play a role in cell cycle and proliferation (Sloan et al., 2013). The coiled-coil domain containing protein 22 (CCDC22) has been shown to be involved in the NFKB signalling and its depletion leads to blockade of signalling (Starokadomskyy et al., 2013). The 6-phosphogluconolactonase protein is part of the pentose phosphate pathway (Collard et al., 1999).

Taken together, the pneumolysin-dependent differentially expressed proteins are predominantly involved in the regulation of gene transcription possibly by chromatin remodelling, modulation of the TP53 and NFKB pathways as well as interference with the assembly of the ribosome.

4.5.1 Comparison of proteomic analysis of the 3 hour and 6 hour infective challenge time points:

The overlap between the differentially expressed proteins identified from the 3 hour and 6 hour infective challenges was established (Figure 4.13). The results showed that of the total majority proteins identified in the mass spectrometry analysis, 38% were common to both analyses. Of the proteins identified as having changes in their abundance by the ANOVA analysis, 4 of the 5 at 3 hours and 16 of the 32 at 6 hours were identified in both searches. Of these 20 proteins, LFQ intensities were measured in both searches in 8 cases. The fold changes in the LFQ intensities at each time point are compared for both analyses (see Figure 4.14).

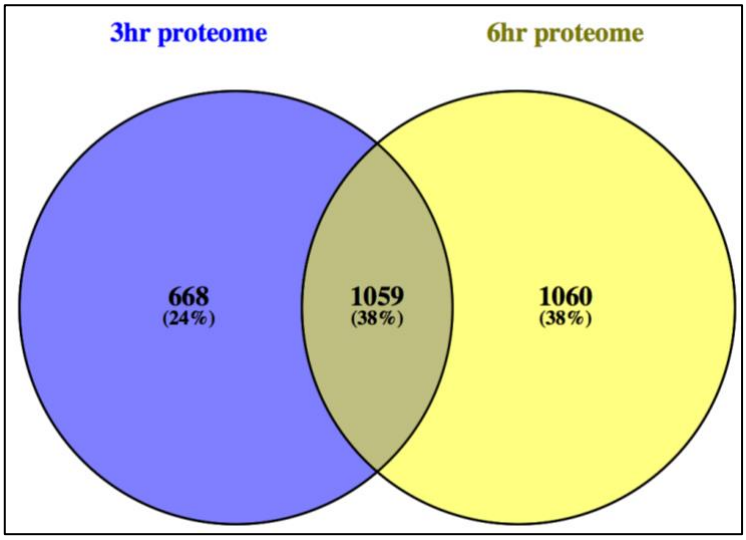


Figure 4.13: Overlap of the proteins identified after 3 and 6 hour bacterial challenges.
Venn diagram illustrating the overlap in the majority protein IDs from both the 3 hour and the 6 hour proteomes.

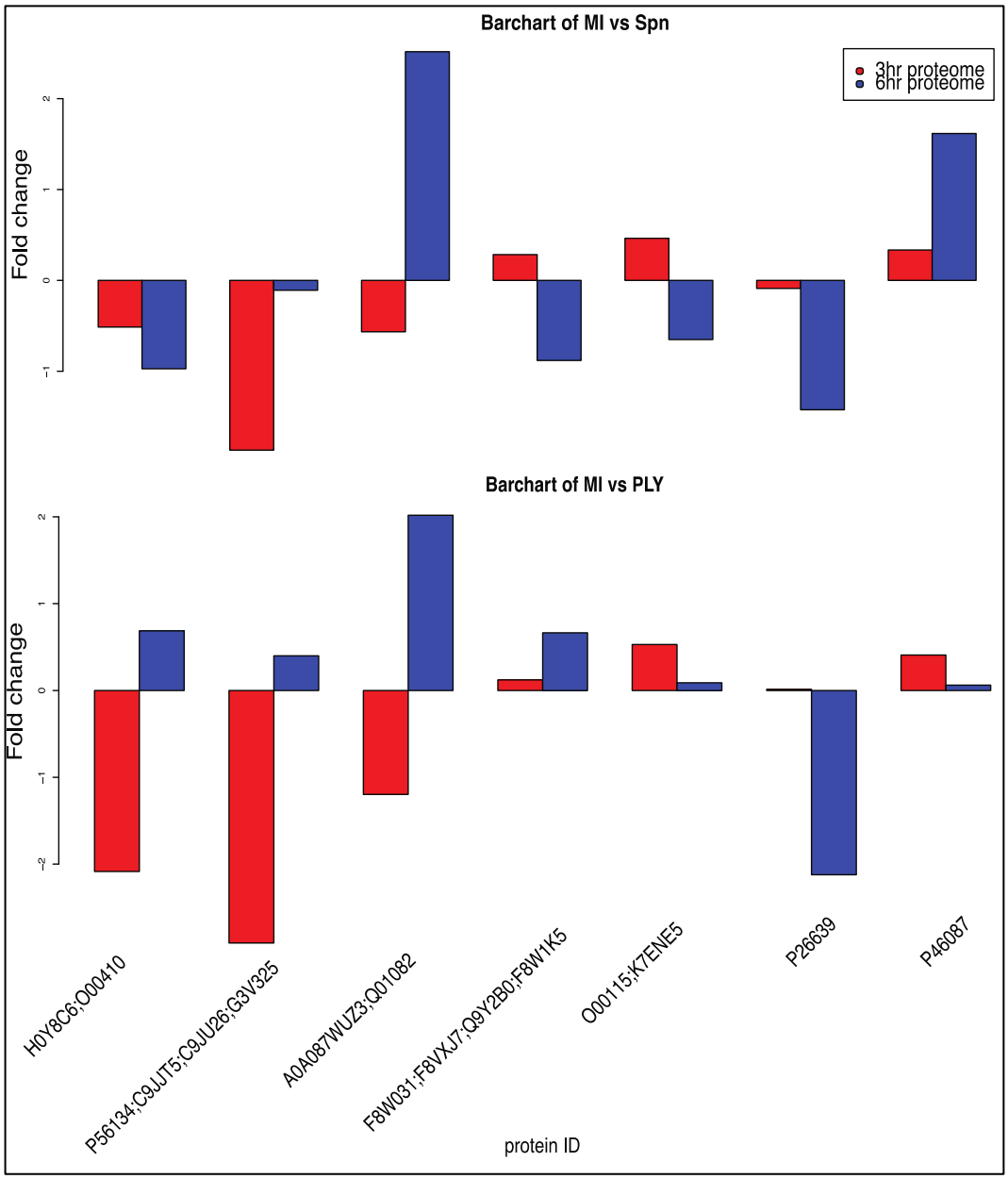


Figure 4.14: Fold change in abundance of proteins after 3 hours and 6 hours bacterial challenges.

This figure illustrates the changes in the abundances of the proteins identified as differentially expressed by ANOVA between MI and *S pneumoniae* challenged cells (top) and the ΔPLY challenged cells (bottom). The red bar shows the fold change at 3 hours and the blue bar that of the 6 hour analysis. Protein ID are the Uniprot identifiers or the majority protein identifications from the MaxQuant analysis.

For 2 of the 8 proteins quantified in both searches the fold changes were in the same direction at both time points. Threonine-tRNA ligase was identified in both searches and in both comparisons as having decreased expression in response to infective challenges. Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase**)** was also found to have an increase in abundance in response to infection at both time points in both comparisons.

4.5.2 Integration of transcriptomic and proteomic analyses:

The central dogma of biology proposed by Crick in 1958 and widely accepted since, dictates that DNA is translated into mRNA, which is in turn translated into proteins (Crick, 1958). However, large number of studies have demonstrated that there is poor correlation between transcriptome and proteome data sets (Chen et al., 2002; Ghazalpour et al., 2011; Pascal et al., 2008; Stare et al., 2017). Indeed several mechanisms control the conversion of mRNA to protein that are not limited to transcriptional efficiency (abundance of mRNA) but also translational efficiency including post-translational modifications of the mRNA, half-life of mRNA transcripts and increased protein degradation which may explain the lack of correlation seen between the two datasets in some studies (Stare et al., 2017). More recently, the correlation between the two datasets has been improved by the use of RNA-to-protein ratios demonstrating that it is theoretically possible to infer the protein abundance from the transcript level, although this does remain controversial and requires a spike in PRM method to establish the absolute protein abundance for each cell line studied to obtain the RNA-to-protein ratios (Edfors et al., 2016).

Having established the changes in the transcriptome and the proteome of MDMs following challenge with *S. pneumoniae* or ΔPLY, next the overlap between the two was examined. The 5 proteins that were identified at the 3 hour time point as being differentially regulated were mapped using *Ensembl* identifiers to the microarray data (Figure 4.15). This shows that although there appears to be reasonable correlation between the fold changes in mRNA transcripts and the fold changes in abundance of proteins, this is not statistically significant ($p>0.05$). This may be in part due to the microarray and proteomes coming from different donors or to the small number of comparisons being made (n=5) as I limited the comparison to the proteins that were differentially regulated.

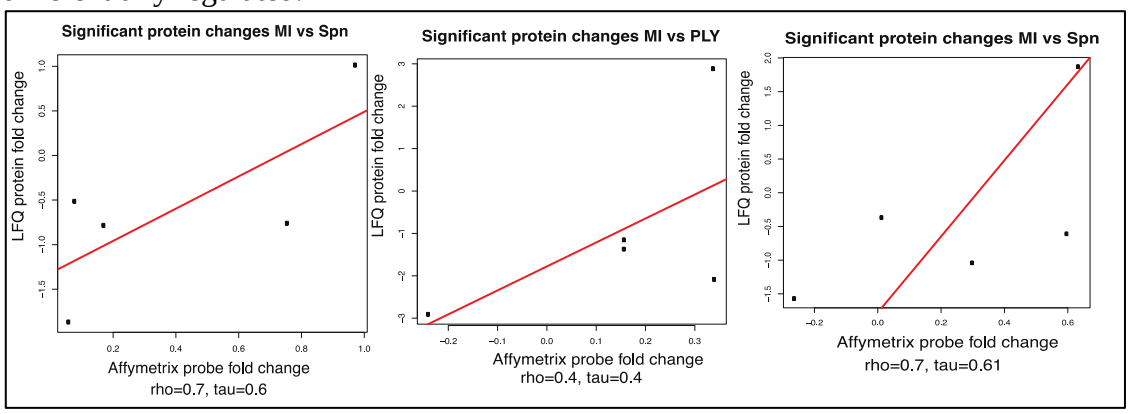


Figure 4.15: Correlation between proteomic data and microarray data obtained after 3 hour infective challenge.

The scatterplots illustrate the correlation between the fold changes in the microarray probe intensities and the LFQ intensities between the MI and *S. pneumoniae* challenged cells (top), MI and PLY (middle) and *S. pneumoniae* and PLY (bottom). n=5, Pearson's correlation and p values were calculated for each comparison. The regression line is plotted in red.

In addition, the comparison between the microarray and the 6 hour proteome was carried out (see Figure 4.16). This shows poor correlation between the two data sets which maybe down to the small number of comparisons being made (n=32), to the different time points (3 hours for the microarray and 6 hours for the proteome) or due the fact that the two data sets are from different donors.

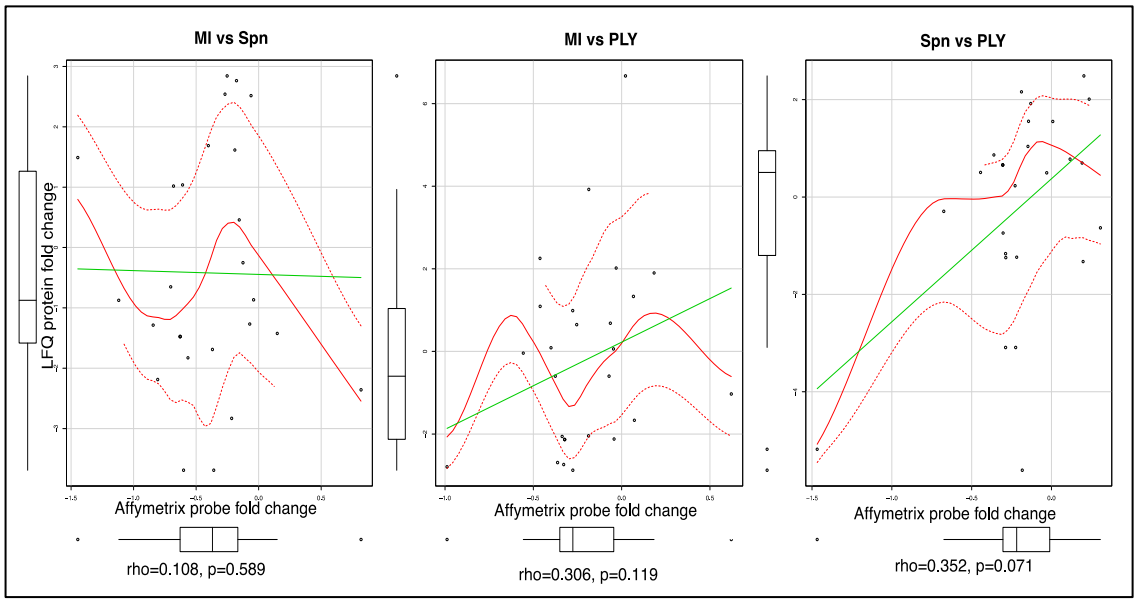


Figure 4.16: Correlation between proteomic data following a 6 hour infective challenge and microarray data following a 3 hour infective challenge.

The scatterplots illustrate the correlation between the fold changes in the microarray probe intensities at 3 hours and the LFQ intensities at 6 hours between the MI and *S. pneumoniae* challenged cells (top), MI and PLY (middle) and *S. pneumoniae* and PLY (bottom). n=32, Spearman rho and p values calculated for each comparison. The regression lines are plotted in green and the locally weighted scatterplot smoothing (Lowess) fits to the data are plotted in red.

However, the results show that the proteins that were found to be significantly differentially expressed at the 6 hour time point (by ANOVA), the corresponding Affymetrix probes for two of them were also differentially expressed at 3 hours (Table 4.4). Neither or these proteins were identified in the 3 hour proteomic analysis.

The vacuolar protein sorting 13 C protein (VPS13C) is associated with mitochondrial function. Loss of VPS13C is associated with mitochondrial fragmentation, and loss of membrane potential as well as an increase in maximal respiration(Lesage et al., 2016). Haloacid dehalogenase like hydrolase domain containing 5 (HDHD5) protein is not well described in the literature, it is associated with cat eye syndrome in humans (Footz et al., 2001). Cat eye syndrome is a rare genetic disease caused by tri or tetrasomy of the short arm of chromosome 22. It is associated with a number defects affecting different organs, the commonest of which is a coloboma (up to 50%).

Table 4.4: Differentially expressed proteins with corresponding changes in microarray values.

	Microarray	**log fold change**			**p value**			
Name	**probe ID**	***Spn* vs MI**	**△PLY vs MI**	***Spn* vs △PLY**	***Spn* vs MI**	**△PLY vs MI**	***Spn* vs △PLY**	**F.value**
VPS13C	218396_at	**-0.565**	**-0.337**	-0.228	**0.012**	0.094	0.235	0.036
HDHD5	218592_s_at	**-1.119**	**-0.989**	-0.130	**0.008**	**0.015**	0.701	0.016
	Proteome							
VPS13C	Q709C8	**-1.828**	**-2.059**	0.231	**0.002**	**0.0002**	0.674	0.0005
HDHD5	Q9BXW7	**-0.874**	**-2.794**	1.920	0.229	**0.0002**	0.001	0.0003

Part C Studying the effect of *S pneumoniae* infection on Histone PTMs in MDMs.

4.6 Histone post-translational modifications:

Having established that there is pneumolysin-dependent effect on the transcriptome, and to a lesser extent on the proteome of MDMs, in the bacterial challenge model, subsequent studies were performed to study the effect on the PTMs landscape to provide further insight at the epigenetic level. As described in Chapter 2 following challenge with either *S pneumoniae* or ΔPLY, cells were collected and the histones purified, chemically derivitised, trypsin digested, desalted and analysed by mass spectrometry. Data analysis was performed in Epiprofile with manual verification using Skyline (as described in Chapters 2 and Chapter 3).

4.6.2 Studying the effect of *S. pneumoniae* infection on Histone PTMs:

Histone samples were prepared as described in Chapter 2 and similarly to the sample preparation for analysis on the maXis (Chapter 4.5.1) acid extracted histones were propionylated twice, trypsin digested and re-propionylated. Then, they were desalted using Hypersep™ tips and re-suspended in 0.05% HFBA and 3% ACN prior to analysis on the QExactive HF Orbitrap. Histone PTMs were identified using Epiprofile (Yuan et al., 2015) and manually verified using Skyline (Schilling et al., 2012). The quantification was performed in Epiprofile for all of the peptides except KSAPATGGVKKPHR and KSAPSTGGVKKPHR which were quantified manually using Skyline as Epiprofile failed to correctly assign all of the PTMs in these peptides. A summary of the relative quantification of the histone PTMs on histone H3 is shown in (Figure 4.18)

Table 4.5: Summary of the significant changes in relative abundance of histone PTMs identified following bacterial challenge.

Histone PTM	*S. pneumoniae*	△PLY
H3K4me1	**Up**	**Down**
H3K9me2	**Down**	**Up**
H3K23ac	**Up**	**Up**
H3K27me2	Down	Up
H3K27me2K36me2		Up
H3K27me3K36me1		Down
H3.3K27me2K36me1	Up	Up
H3.3K27me2K36me2	Down	Down
H3.3K36me2	Up	Down
H3K79me2	**Down**	**Up**

The results show that in response to challenge with *S. pneumoniae* there is an increase in H3K4me1 relative to the ΔPLY, the decrease in H3K9me2 is maintained and interestingly in response to ΔPLY the relative abundance of H3K9me2 is increased compared to MI (Figure 4.18 panel A). There is an increase in the level of H3K23ac in

response to both bacterial challenges, although the response to *S. pneumoniae* appears to be greatest (Figure 4.18 panel B). The results show a small decrease in H3K27me2 following challenge *S. pneumoniae* when compared to ΔPLY. In addition, there seemed to be a relative increase in the level of the combinatorial mark K27 and K36 dimethylation (H3K27me2K36me2) and a reciprocal drop K27 trimethylation and K36 monomethylation (H3K27me3K36me1) in response to ΔPLY challenge compared to MI.

The results also show that on the peptide KSAPSTGGVKKPHR of histone H3.3 there is an increase in the relative abundance of K36me2. In addition, there was an increase in the combinatorial PTMs K27 dimethylation K36 monomethylation (H3.3K27me2K36me1), with a reciprocal drop in the dimethylation of lysines 27 and 36 (H3.3K27me2K36me2). Analysis of the EIAQDFKTDLR peptide shows that although the trend to an increase in H3K79me1 in response to *S. pneumoniae* challenge was observed it does not reach significance in this dataset. However, a decrease in the level of the H3K79me2 was observed.

There were no changes in the relative abundance of the histone PTMs on the YQSTELLIR and the VTIMPKDIQLAR peptides (Appendix figure 4.3).

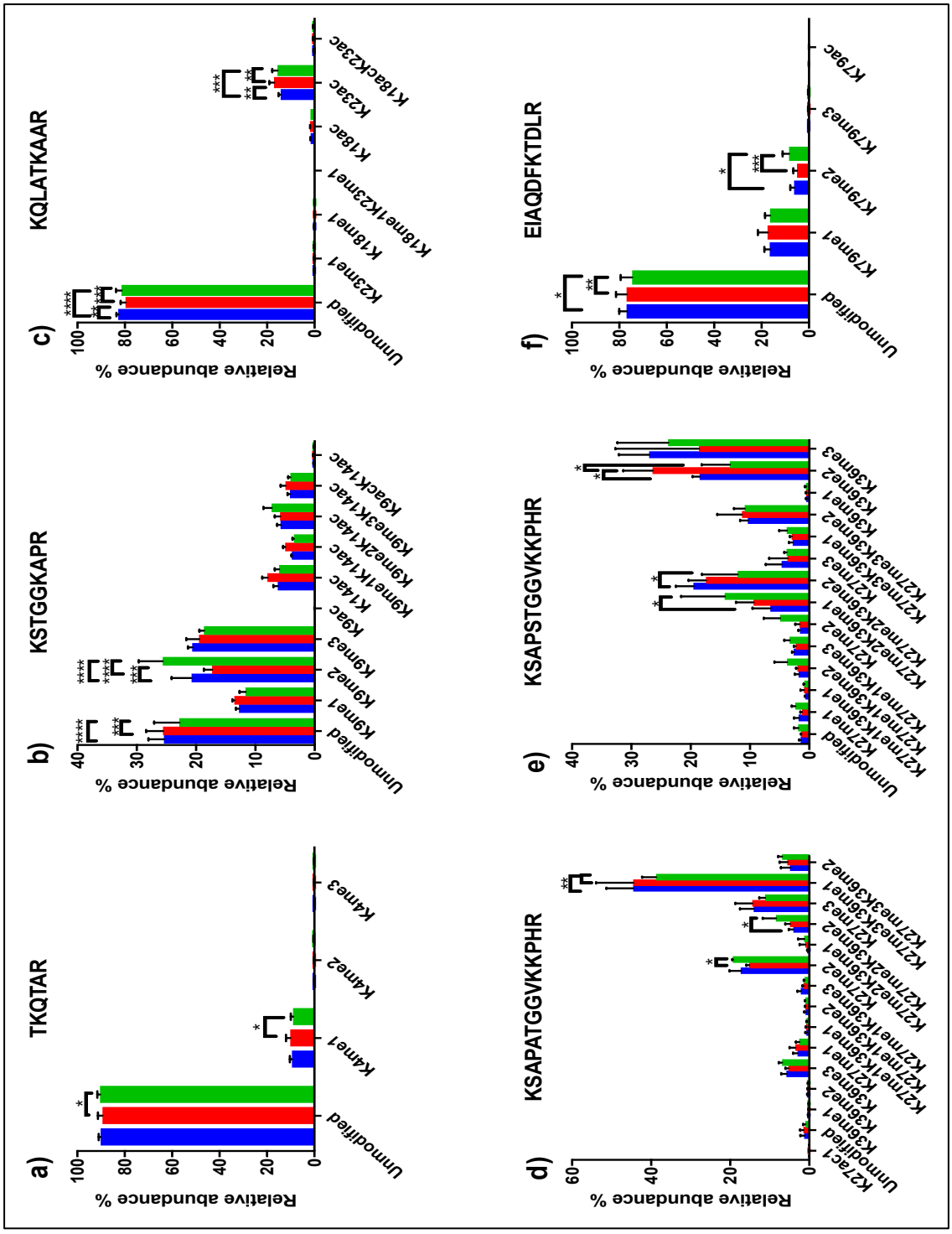


Figure 4.18: Changes in the relative abundance of the PTMs on histone H3 in a pneumolysin-dependent manner on the QE.

These bar charts summarise the relative abundance in the histone PTMs identified on histone H3. Panel a) on the TKQTAR peptide there is a relative increase in the level of H3K4me1 following challenge with *S. pneumoniae* (Green) in comparison to the ΔPLY (red). panel b) on the KSTGGKAPR peptide there is a drop in the level of H3K9me2 in a pneumolysin-dependent manner compared to both mock infected (MI; blue) and ΔPLY. panel c) on the KQLATKAAR there is a relative increase in the level of H3K23ac in response to both bacterial challenges. Panel d) on the KSAPATGGVKKPHR peptide; there was relatively less abundant H3K27me2 following challenge with *S. pneumoniae* than with ΔPLY. In addition there is an increase in the level of H3K27me2K36me2 and a reciprocal drop in H3K27me3K36me1 in response to ΔPLY challenge compared to MI. Panel e) on the EIAQDFKTDLR peptide there is an increase in the

relative abundance of H3K79me2 following challenge with ΔPLY. Finally, panel f) on the KSAPSTGGVKKPHR peptide of H3.3 there is an increase in H3.3K36me2 and in H3.3K27me2K36me1 with a reciprocal drop in H3.3K27me2K36me2. (n=3. One way ANOVA, * p<0.05, ** p<0.01, *** p<0.001).

Having identified changes in relative abundance on histone H3, I then examined the changes in relative abundance for the GKGGKGLGKGGAKR and KVLR peptides of histone H4 (Figure 4.19). The results showed a significant increase in the abundance of acetylation on K16 (H4K16ac). In addition, there was also a decrease in the dimethylated form of K20 (H4K20me2) and a reciprocal increase in the monomethylated form (H4K20me1) in a pneumolysin-dependent manner.

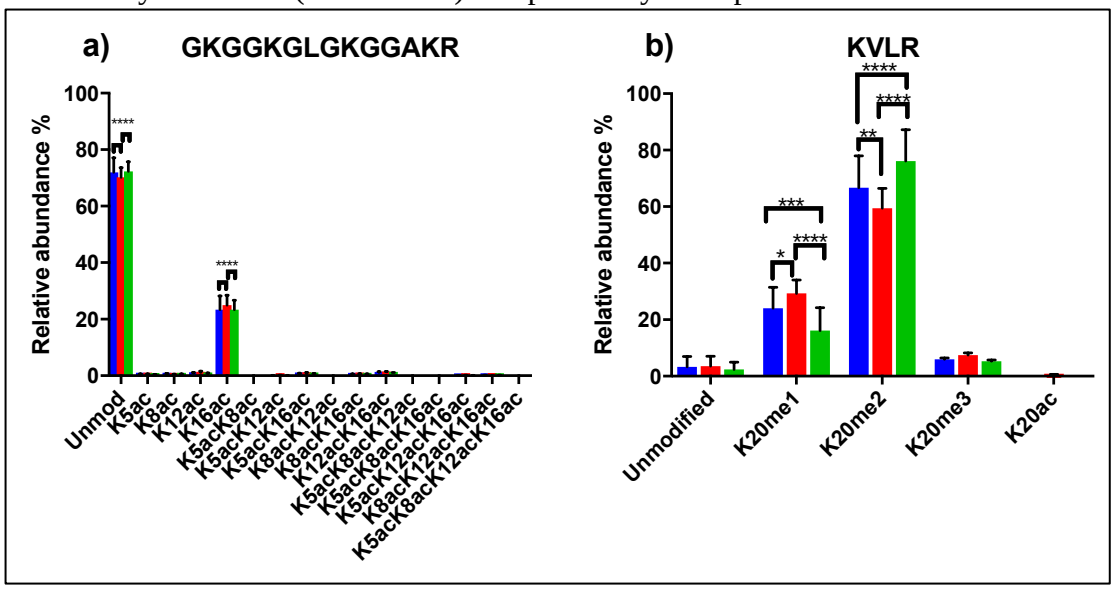


Figure 4.19: Relative quantification of histone H4 PTMs upon challenge with *S. pneumoniae*.

These bar charts summarise the relative abundance in the histone PTMs identified on histone H4. a) Relative abundance of GKGGKGLGKGGAKR peptide proteoforms reveals a relative increase in the level of H4K16ac following challenge with *S. pneumoniae* (green) in comparison to the ΔPLY (red). b) Relative abundance of KVLR peptide proteoforms reveals there is a decrease in the level of H4K20me2 in a pneumolysin-dependent manner compared to both MI (blue) and ΔPLY and a reciprocal increase in the H4K20me1. Conversely for the cells challenged with ΔPLY there was a rise in H4K20me2 and a fall in the level of H4K20me1. (n=3. One way ANOVA, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001).

A summary of the changes in the relative abundance of PTMs in response to infection with *S. pneumoniae* in a pneumolysin-dependent manner is summarised in Table 4.5. The following sections will discuss the roles of these epigenetics marks.

4.7.1 H3K4me1:

The relative abundance of H3K4me1 levels was shown to increase in response to infective challenge with *S. pneumoniae* when compared to the challenge with ΔPLY (see Figure 4.18). H3K4me1 is predominantly associated with active or poised enhancers and also found, to a lesser extent, on promoters (Barski et al., 2007; Heintzman et al., 2007). However, a clear function for this mark has yet to be established. H3K4me1 has been found with K9 acetylation at the transcription start sites of genes (Liang et al., 2004). A careful balance occurs between the trimethylated and monomethylated form at promoter regions, as when H3K4me1 is found on its own it is

associated with gene repression but when it is in minority compared to H3K4me3 then it is seen at the site of actively transcribed genes (Cheng et al., 2014). It has been shown to delineate boundaries of promoters in conjunction with the H3K4me3 mark that are associated with active gene transcription (Cheng et al., 2014). Furthermore, the concomitant presence of H3K27ac at enhancer sites is associated with gene activation (Okabe et al., 2017). Although the exact contribution of H3K4me1 to active gene transcription is less clear, as gene transcription was only marginally impaired when H3K4me1 was removed from enhancers (Dorighi et al., 2017).

In terminally differentiated cells such as macrophages, the enhancer landscape is remodelled by TLR4 stimulation and is associated with the formation of de novo enhancers marked by H3K4me1 (Kaikkonen et al., 2013). The deposition of H3K4me1 at enhancer sites is associated with faster and stronger response following re-stimulation with LPS, which may represent a memory like mechanism for these cells (Ostuni et al., 2013). Indeed the H3K4me1 PTM is maintained after the removal of H3K27ac (Saeed et al., 2014). Infection with Epstein Barr virus was associated with a loss of repressive marks and an increased in H3K4me1 and H3K27ac at enhancer sites highlighting the important role played by H3K4me1 in the host response to infection (Okabe et al., 2017) Furthermore, H3K4me1 was shown to be deposited in the enhancer region of host defence genes in macrophages infected with *Mycobacterium tuberculosis* (Bouttier et al., 2016). H3K4me1 at enhancers has been shown to be associated with different transcriptional states depending on which PTMs it coexists with, suggesting that there are multiple layers of control over gene transcription at these sites (Soldi et al., 2017). The presence of both H3K4me1 and H3K27ac at enhancers has been associated with active transcription (Pradeepa, 2017), whereas the presence of both H3K4me1 and H3K36me2 was found at intronic and extragenic regions, the co-localisation of H3K4me1 with either H3K79me2 or H3K36me3 was mainly found at introns (Soldi et al., 2017). The authors went on to show that the combination of H3K4me1/H3K27ac/K79me2 and H3K4me1/H3K27ac/K36me3 at enhancers was associated with distinct aspects of the macrophage inflammatory response. The H3K4me1/H3K27ac/H3K79me2 enhancers were found to be associated with the GO term “positive response to JUN kinase activity” term, whereas the H3K4me1/H3K27ac/K36me3 was associated with the “response to lipopolysaccharide” term. The authors conclude that the distinct enhancer motif control different aspects of the inflammatory response.

H3K4me1 is removed by the demethylase lysine specific demethylase 1 (LSD1)(Shi et al., 2004). It appears to not have any activity on trimethylated lysine residues. It is also responsible for demethylation of lysine 9 residues on Histone H3 (Metzger et al., 2005). The exact mechanism of the demethylation is unclear but appears to be a flavin dependent mono amine oxidative process (Forneris et al., 2005). Set9 is the methyltransferase responsible for methylation of H3K4 (Nishioka et al., 2002). It was found that methylation of H3K4 was associated with inhibition of methylation of H3K9 by SET domain of the human homolog of *Drosophila* Su(var)3-9 (Suv39h1). As the methylation of H3K9 is associated with the formation of heterochromatin and gene repression, the authors suggest that H3K4methylation is therefore associated with the maintenance of active transcription (Nishioka et al., 2002).

Finally, the increase in H3K4me1 in response to challenge with *S. pneumoniae* infection in MDMs is consistent with the literature suggesting that after TLR4 stimulation (Kaikkonen et al., 2013), LPS challenge (Ostuni et al., 2013), or infection with EBV (Okabe et al., 2017) the levels of H3K4me1 increase. Given that the majority of the literature supports a role for H3K4me1 at enhancers in active or poised genes, it

is likely that the increase observed in my system relates to the recruitment of genes to the immune response.

4.7.2 H3K9me2:

The relative abundance of H3K9me2 levels was seen to decrease in response to challenge with *S. pneumoniae*, but increase in response to challenge with ΔPLY (see Figure 4.18).

H3K9me2 is primarily associated with gene repression by orchestrating heterochromatin formation and covers large genomic domains (Kouzarides, 2002; Lienert et al., 2011; Tachibana et al., 2002). It is thought to play a critical role in the differentiation process of pluripotent cells into differentiated cells (Wen et al., 2009). Although some controversy persists, it has been suggested that dimethylation of H3K9 occurs as part of the differentiation process of embryonic stem cells into tissue specific lineages (Wen et al., 2009). Following differentiation, up to 30% of the genome was found to be associated with regions of H3K9me2. However, more recently, the increase in global amounts of H3K9me2 upon differentiation has been challenged (Lienert et al., 2011). The authors propose that local variations in H3K9me2 abundance rather than large changes in global levels play a role in the stabilisation of transcriptional response of differentiated cells compared to their pluripotent stem cell ancestors. They show that the presence of H3K9me2 is exclusive to the regions of active gene transcription. Regulation of the differentiation of monocytes into macrophages has been shown to involve the modulation of H3K9me2 levels (Hoeksema and de Winther, 2016).

H3K9me2 has also been demonstrated to be key in regulating inflammatory responses in terminally differentiated cells such as macrophages. The presence of H3K9me2 at the promoter sites for genes involved in interferon and antiviral responses was associated with decreased gene transcription (Fang et al., 2012). The response to LPS challenge in macrophages was associated with removal of H3K9me2 and increased gene transcription (Saccani and Natoli, 2002). Furthermore, H3K9me2 has been shown to be removed from the promoter regions of previously transcriptionally silent genes involved in the inflammatory response following LPS stimulation, which are mediated via NFKB dependent pathways (Essen et al., 2010). Indeed, in macrophages, the stress response transcription factor ATF7 has been show to recruit G9a leading to dimethylation of H3K9, which was reversed by stimulation with LPS (Yoshida et al., 2015a).

Dimethylation of H3K9 is mediated by the methyltransferase G9a (Gazzar et al., 2008; Scheer and Zaph, 2017; Tachibana et al., 2002). The removal of H3K9me2 can involve a number of different demethylases such as LSD1 (Metzger et al., 2005), amine oxidase flavin containing 1 (Aof1 also known as lysine-specific histone demethylase 2 (LSD2) or lysine demethylase 1B (Kdm1b)(Essen et al., 2010), Kdm3a/Kdm3b . (Ebata et al., 2017) or PHD finger protein 8 (PHF8)(Zhu et al., 2010).

Finally, the decrease observed in the relative abundance of H3K9me2 following bacterial challenge may represent the removal of repressive pressure on certain genes to allow activation of distinct transcriptomic modules as part of the host's immune response to infection.

4.7.3 H3K23ac:

Following infective challenge with either *S. pneumoniae* or ΔPLY in MDMs there is a rise in the relative abundance of H3K23ac compared to the mock infected cells (see Figure 4.18).

The acetylation of lysine 23 has not been extensively described in the literature. In *Arabidopsis* the presence of H3K23ac is associated with active gene transcription (Lu et al., 2015; L. Wang et al., 2017). In Drosophila H3K23ac was found at the promoter sites of actively transcribed genes (Bodai et al., 2012). In humans, the acetylation of H3K23 has been found to be recognised by tripartite motif containing 24 (TRIM24), which reads the combinatorial modification H3K4 unmodified and H3K23ac (Tsai et al., 2010). This in turn binds the oestrogen receptor and leads to increased oestrogen dependent gene expression. High levels of TRIM24 and H3K23ac in breast cancer tissue was associated with poor survival(Ma et al., 2016). Moreover, the lysine acetyltransferase 6A (KAT6A) has been shown to acetylate H3K23 recruiting TRIM24 and play a role in the formation of gliomas (Lv et al., 2017). Therefore, given the increase in relative abundance of H3K23ac following infective challenge in my model, it is likely that this occurs at the site of genes that are actively transcribed and therefore may constitute part of the host's response to infection.

4.7.4 KSAPATGGVKKPHR peptide:

Following challenge with *S. pneumoniae*, there was a relative decrease in the abundance of H3K27me2 when compared to the ΔPLY challenged cells.

H3K27me2 has been shown in embryonic stem cells to be one of the most abundant PTM, where it was found primarily in intergenic regions and appears to play a role in preventing aberrant acetylation of lysine 27 (Ferrari et al., 2014). This suggests that it is a primarily repressive mark. H3K27 is methylated by the polycomb repressive complex 2 (PCR2) to both dimethylated and trimethylated states (Conway et al., 2015; Juan et al., 2016).

Furthermore, the authors showed that the two modifications were associated with functionally distinct regions of their embryonic stem cells (Juan et al., 2016). The different methylation states of lysine 27 are associated with distinct transcriptional profiles, the monomethylated form, in combination with H3K36me3, is seen in regions of active gene transcription, in contrast the dimethylated form is shown to play a role in gene repression (Ferrari et al., 2014), as is the trimethylated form (Barski et al., 2007). The H3K27me2 mark has been associated with repression of metabolic genes, and the H3K27me3 of developmental genes (Juan et al., 2016). In CD4+ T cells both the H3K27me2 and H3K27me3 mark were associated with gene repression (Wang et al., 2008). Both the dimethylated and the trimethylated form of lysine 27 are demethylated by JMJD3 (Burchfield et al., 2015; Chen et al., 2012; De Santa et al., 2009a). Indeed, JMJD3 was shown to remove the H3K27me3 in macrophages stimulated with LPS (De Santa et al., 2009b). JMJD3 has been shown to be essential for the differentiation and polarisation of macrophages (Satoh et al., 2010). Highlighting the critical role played by this modification in the host response.

Given that this mark is decreased following challenge with *S. pneumoniae* compared to the challenge with ΔPLY, it is possible that this predominantly repressive mark is being removed to allow the activation of genes related to metabolism as part of the cell's response to infective challenge.

In the cells challenged with ΔPLY, there was also an increase in the relative abundance of H3K27me2K36me2 and decrease in the H3K27me3K36me1 PTMs (see Figure 4.18). This highlights the advantages of mass spectrometry to identify combinatorial marks over antibody based approaches which would not have identified them.

The literature is rather sparse when examining combinatorial modifications. Indeed, as they have primarily been identified by MS, functional studies describing the roles played by combinatorial modifications are lacking. The dimethylation of H3K27 and H3K36 was found to be increased in diabetic rats (W. Wang et al., 2017). One of the early studies of histone PTMs across species has suggested that the combinatorial modifications have arisen by the co-evolution of two or more enzymes that maintain the different PTMs across eukaryotic species, the authors suggest that as the species become more complex (from eukaryotes to mammals), the complexity and combinatorial nature of PTMs becomes more prominent, indeed they demonstrate that several combinatorial modifications are lacking in yeast (Garcia et al., 2007). Furthermore, they suggest that the more complex an organism is, the more repressive marks it possesses. Another study suggests that lysine methylation can act as a buffer for the maintenance of epigenetic marks during cell division as the levels acquired by the daughter cell are not as abundant as the parental one (Xu et al., 2011).

H3K36 methylation has been shown to be associated with euchromatin and therefore with active gene transcription (Miao and Natarajan, 2005; Vakoc et al., 2006). H3K36me2 has been shown to antagonise PRC2 and therefore the methylation of lysine 27 (Ballaré et al., 2012; Yuan et al., 2011). The authors suggest that this controls the progression of repressive chromatin. It has also been shown to be deposited after double stranded DNA breaks, as part of the repair mechanism (Fnu et al., 2011). Both H3K36 mono and di-methylation are methylated by the absent, small, or homeotic-1 (ASH1)) subfamily of SET domain proteins (Tanaka et al., 2007).

Given that methylation of K27 is associated with gene repression and conversely the methylation of K36 is associated with active gene transcription the role of the combinatorial modifications is likely to be complex, and may be involved in the fine tuning of these events.

4.7.5 KSAPSTGGVKKPHR peptide:

On the KSAPSTGGVKKPHR peptide of H3.3 there is an increase in H3.3K36me2 and in H3.3K27me2K36me1 with a reciprocal drop in H3.3K27me2K36me2 in response to challenge with *S. pneumoniae* and similar changes in H3.3K27me2K36me1 and H3.3K27me2K36me2 in response to challenge with ΔPLY (see Figure 4.18).

Histone H3.3 is a variant of histone H3.1/2 that differs by 4 amino acids at positions 87, 89, 90 and 96 (Szenker et al., 2011). H3.3 has been found to be enriched at regions of active transcription in Hela cells (Xu et al., 2010) suggesting that it plays a key role in gene transcription control. Furthermore, H3.3 was found to be deposited in exchange for H3.1 at sites of active transcription (Schwartz and Ahmad, 2005). It has also been found to be associated with the enhancer regions of actively transcribed genes (Chen et al., 2013). H3.3 has been shown to accumulate with age in murine tissues and correlates with altered patterns of H3 PTMs (Tvardovskiy et al., 2017). The exact role of H3.3 is still to be more fully elucidated and the function of PTMs on H3.3 are likely to be similar to its counterparts on H3.1 but have yet to be extensively determined.

4.7.6 H3K79 methylation:

Following challenge with the ΔPLY there is an increase in the relative abundance of H3K79me2 on the EIAQDFKTDLR with a reciprocal drop in the unmodified form of the peptide (see Figure 4.18).

The precise role played by H3K79me1 is yet to be fully understood. It has been found to localise at both promoter regions of active transcription sites and intergenic silent transcription (Steger et al., 2008). Indeed, the authors show that with successive levels of methylation (dimethylated and trimethylated) the marks were closer to the promoter regions and co-localised with H3K4me2/me3. They suggest therefore that the monomethylated form might act as a template for long range enhancers and a poised mark at promoters whereas the di- and trimethylated form act as an activator at the promoters. It has been proposed that H3K79me1 might play a role in the localisation of long range enhancers. Furthermore, it has been shown that H3K79 methylation is associated with H3K4me1 at promoter sites and correlates with H3K9ac suggesting an interplay between the two PTMs (Zhang et al., 2004). However, the observation that gene transcription is maintained at the same rate following depletion of H3K79 methylation raises the possibility that its function may be distinct from transcription regulation and instead play a role in genome stability. This contrasts with other studies that have shown H3K79me1 level to be highly predictive of gene transcription (Karlić et al., 2010). H3K79me1 has been associated with active gene transcription whereas the trimethylated form, H3K79me3 is associated with repressed transcription (Barrand and Collas, 2010; Barski et al., 2007). The function of the dimethylated lysine 79 is yet to be fully elucidated, as H3K79me2 has been linked to active (Okada et al., 2005) transcription in some studies but not others (Barski et al., 2007). In cardiac myocytes H3K79me2 has been shown to control different transcriptional networks at different stages of the cells differentiation (Cattaneo et al., 2016). Moreover, H3K79me2 has been shown to play a role in the control of gene transcription during the haemopoetic differentiation of acute myeloid leukaemia by enhancing DNA accessibility (Ye et al., 2015). Thus the exact role of the H3K79 methylation remains to be fully explained.

4.7.7 Histone H4:

In response to the bacterial challenge with *S. pneumoniae*, there is a rise in the relative abundance of H4K16ac. In murine embryonic stem cells acetylation of histone H4K16 has been associated with active gene transcription as it was found not only in the transcription start site of genes but also in active enhancers accompanied by H3K4me1 (Taylor et al., 2013). However, in yeast cells H4K16ac was not associated with active transcription (Kurdistani et al., 2004). Moreover, in human embryonic kidney 293 cells H4K16ac only had a modest effect on transcriptional regulation, suggesting that as with many other PTMs the role of H4K16ac may vary from cell type to cell type (Horikoshi et al., 2013). The exact role for H4K16ac in human cells is therefore yet to be fully determined. It has been shown to be associated with double stranded DNA repair mechanisms by maintaining chromatin in an accessible open structure and thus enabling detection of DNA damage (Sharma et al., 2010; Zhong et al., 2017). In human CD4+ T cells the combination of H4K16ac and with other acetylated PTMs was associated with moderate to high rates of gene transcription (Wang et al., 2008). H4K16ac has been demonstrated to be removed by sirtuin 1 (Sirt1)(Vaquero et al., 2004). Hypoacetylation of H4K16 was associated with the demethylation of H3K79me1 and with recruitment of Histone H1 leading to repression of gene transcription.

H4K20 methylation has been shown to be critical for embryonic development (Oda et al., 2009). The authors demonstrated that following depletion of PR-Set7/Set8/KMT5a, the lysine methyltransferase (KMT) responsible for the formation of H4K20me1 there was a loss of both H4K20me1 and H4K20me2/3 as well

as defects in the condensation of the chromatin highlighting its importance for normal embryogenesis. Indeed depletion of H4K20me1 was associated with mitotic defects (Julien and Herr, 2004). The authors suggest that the progression from monomethylation to the dimethylated form of H4K20 may cause irreversible changes in higher order chromatin structure. H4K20me1 has been shown to associate with regions of silent gene transcription and co-localises with H4K9me1 (Sims et al., 2006). H4K20me2 has also been associated with a role in double stranded DNA repair (Pellegrino et al., 2017). *Mycobacterium tuberculosis* has been shown to induce the Set8 methyltransferase via the induction of the microRNA 30e-3p, which causes H4K20 methylation and this in turn leads to up-regulation of NAD(P)H dehydrogenase quinone 1 (NQO1) and thioredoxin reductase (TRXR1) which are associated with the negative regulation of inflammation and apoptosis respectively, thus forming part of the pathogens immune evasion strategy (Singh et al., 2017).

4.8 Conclusions:

In this chapter the effect of pneumolysin on MDMs has been examined at the transcriptional, proteomic and epigenome level. Following 3 hour exposure to *S. pneumoniae* or ΔPLY, I demonstrated that MDMs have ingested either bacteria in similar numbers and started to release pro-inflammatory cytokines. Three hours following the infection of MDMs with either D39 or ΔPLY, mRNA was extracted and affymetrix Chip analysis performed to determine pneumolysin-dependent differentially expressed genes.

The microarray analysis showed that 503 probes are differentially expressed in pneumolysin-dependent manner. A number of probes relate to genes involved in the immune response such as TNF, and HMOX1. Pathway analysis of the differentially expressed probes highlighted the importance of metabolic pathways in response to infection, and in particular the role for oxidative stress responses. The oxidative stress response has been shown to be up-regulated following *S pneumoniae* infections in lung epithelial cells (Zahlten et al., 2015). As this response is also demonstrated in MDMs this provides further evidence for the important role played by oxidative stress responses following infections with *S pneumoniae.* Given the role played by NRF2 as a master regulator of oxidative stress responses in the lung this may provide a further therapeutic avenue by modulating NRF2 levels during infection. Furthermore, the KEGG pathway analysis for both of the infective challenges highlighted NFKB, NOD-like receptor and TNF signalling as enriched pathways emphasising their importance in the host's response to infective challenge.

The label-free quantitative proteomic analysis revealed only small numbers of significantly differentially expressed proteins. 5 proteins were differentially expressed after 3 hours of infective challenge and 8 proteins were differentially expressed in a pneumolysin-dependent manner following a 6 hour challenge. This data set also showed an oxidative stress response protein being differentially expressed further reinforcing

the pivotal role played by oxidative stress responses in the context of infections. One of the reasons for the relatively small number of differentially expressed proteins may be the large differences between the biological replicates. Furthermore, several studies have shown only modest correlation between transcriptomic and proteomic datasets (Chen et al., 2002; Ghazalpour et al., 2011; Pascal et al., 2008; Stare et al., 2017), which might further explain the small number of differentially expressed proteins. Moreover, the translation from altered transcriptional response to altered proteome may be longer than the 6 hour time point studied, and it might be of benefit to extend the time point further. Indeed, in the study of proteomic response to *M. tuberculosis* and pathogenic porcine reproductive and respiratory syndrome virus both authors used a 24 hour time point (H. Li et al., 2017; Qu et al., 2017). It is possible that the differentially expressed proteins are of relatively low abundance and therefore not readily detected using a shotgun proteomics approach, as the more abundant structural proteins such as cytoskeletal proteins were readily identified but not differentially expressed.

Finally, the global histone PTMs landscape study has shown that there are several changes in the relative abundance of histone PTMs in response to *S pneumoniae.* Using the data independent method optimised in Chapter 3, I was able to identify and quantify 84 different peptide proteoforms. Of these, there were 5 whose relative abundance changed in a pneumolysin dependent manner. This is to my knowledge the first time that this approach has been used to study the changes in relative abundance of histone PTMs in primary MDMs in response to *S. pneumoniae.* The advantages of mass spectrometry were emphasised by the identification of combinatorial marks on both H3 and H3.3 that would not have been possible using antibody based approaches. Interestingly several of these modifications are associated with gene transcription. The exact function of PTMs is yet to be fully described. It is likely to vary between cell types, furthermore individual PTMs won't act in isolation but rather be part of a combinatorial code of fine tuning, in response environmental stressors, such as response to bacterial infection and at later time points to DNA damage. Nevertheless, in light of the changes observed in response to challenge with *S. pneumoniae* namely the rise in H3K4me1, H3K23ac, H4K16ac and given the decrease in H3K9me2, H3K79me2 it is possible that these changes represent the removal of repressive marks and the increase in marks associated with active gene transcription thereby allowing the host response to occur. This highlights the important role played by PTMs in the response to infection and therefore offer novel possible immunomodulating therapeutic avenues through the use of emerging classes of drugs that target the enzymes that regulate these histone PTMs.

Chapter 5: The role of histone PTMs in the host's response to infection.

5.1 Introduction:

There are few studies of the impact of live bacteria on histone PTMs in primary cells and the functional consequences of these are not fully understood. Following infection of cell lines with *L. monocytogenes* and or *L. pneumophilia* changes in histone PTMs have been demonstrated and in the case of *L. pneumophilia* these were associated with changes in the immune responses (Eskandarian et al., 2013; Rolando et al., 2013; Rolando and Buchrieser, 2014).

Following previous studies in Chapter 4 that identified a number of histone PTMs changing abundance upon challenge with *S. pneumoniae*, further studies were performed to provide additional mechanistic insight into the epigenomics associated with *S. pneumoniae*. In an attempt to infer a role for the changes observed in the relative abundance of histone PTMs during challenge with *S. pneumoniae.* Having established that histone PTMs were modulated during early infection, next I sought to establish if these could account for changes in the transcriptome. In order to establish this, I performed RNA-seq to establish the transcriptional activity of these cells at this early time point. Then I performed ChIP-seq for key histone PTMs to characterise the genomic location of these and their relationship to the differentially expressed genes. I obtained a further set of paired RNA, cross linked chromatin and histone samples from 9 individuals. As there were only very few significant changes in the relative abundance of PTMs in response to challenge with ΔPLY and in order to keep the number of samples to sequence and the amount of MS time down only the mock infected and the cells challenged with *S. pneumoniae* as described in Chapter 2 were taken forward for further analysis. Briefly, cells from nine different donors were either mock infected or challenged with *S. pneumoniae,* then at 3 hours following challenge cells were either taken for histone analysis, or RNA was collected in TriReagent™ or cross linked in formaldehyde for ChIP-seq. This resulted in paired samples for which RNA-seq, ChIP-seq and histone PTM analysis was possible.

In this chapter I hypothesised that modifications in histone PTMs arising in response to bacterial challenge occur at specific genomic location and are in turn associated with changes in gene transcription.

5.1 Analysis of the relative abundance of histone PTMs following challenge with *S. pneumoniae*:

Following mass spectrometry analysis of the relative abundance of histone PTMs following challenge with *S. pneumoniae* and ΔPLY in monocyte derived macrophages (in Chapter 4), I identified a number of changes in the relative abundance of histone PTMs in a pneumolysin dependent manner. In this Chapter, I obtained paired samples of RNA, cross-linked chromatin and histones from 9 different individuals with and without in vitro challenge to *S. pneumoniae*. The sample preparations were as described in Chapter 3.

Figures 5.1, 5.2 and 5.3 illustrates the changes in relative abundance seen following 3 hour challenge with *S. pneumoniae* analysed using Orbitrap QExactive HF. This enabled the identification of several PTMs (as described in Chapter 4 and summarised in Appendix table 5.1). The changes in the relative abundance of histone

PTMs in H3 are consistent with those seen in the comparison of challenge to both *S. pneumoniae* described in Chapter 4.

The results show a number of histone PTMs are altered in abundance in response to challenge with *S. pneumoniae* including an increase in H3K4me1 and a decrease in H3K9me2 (which is maintained compared to MI) (Figure 5.1). In addition, the results show that there is also an increase in the level of H3K23ac, a decrease H3K27me2 and a reciprocal increase in the abundance of H3K27me3 in response to challenge with *S. pneumoniae*. However, despite increasing the number of samples analysed, although a similar trend was observed, the changes in the combinatorial modifications described in Chapter 4 (in response to challenge with ΔPLY) did not reach statistical significance in this subset of samples. There was no significant change in the relative abundance of PTMs on the YQSTELLIR, EIAQFKTDLR and VTIMPKDIQLAR peptides of histone H3.

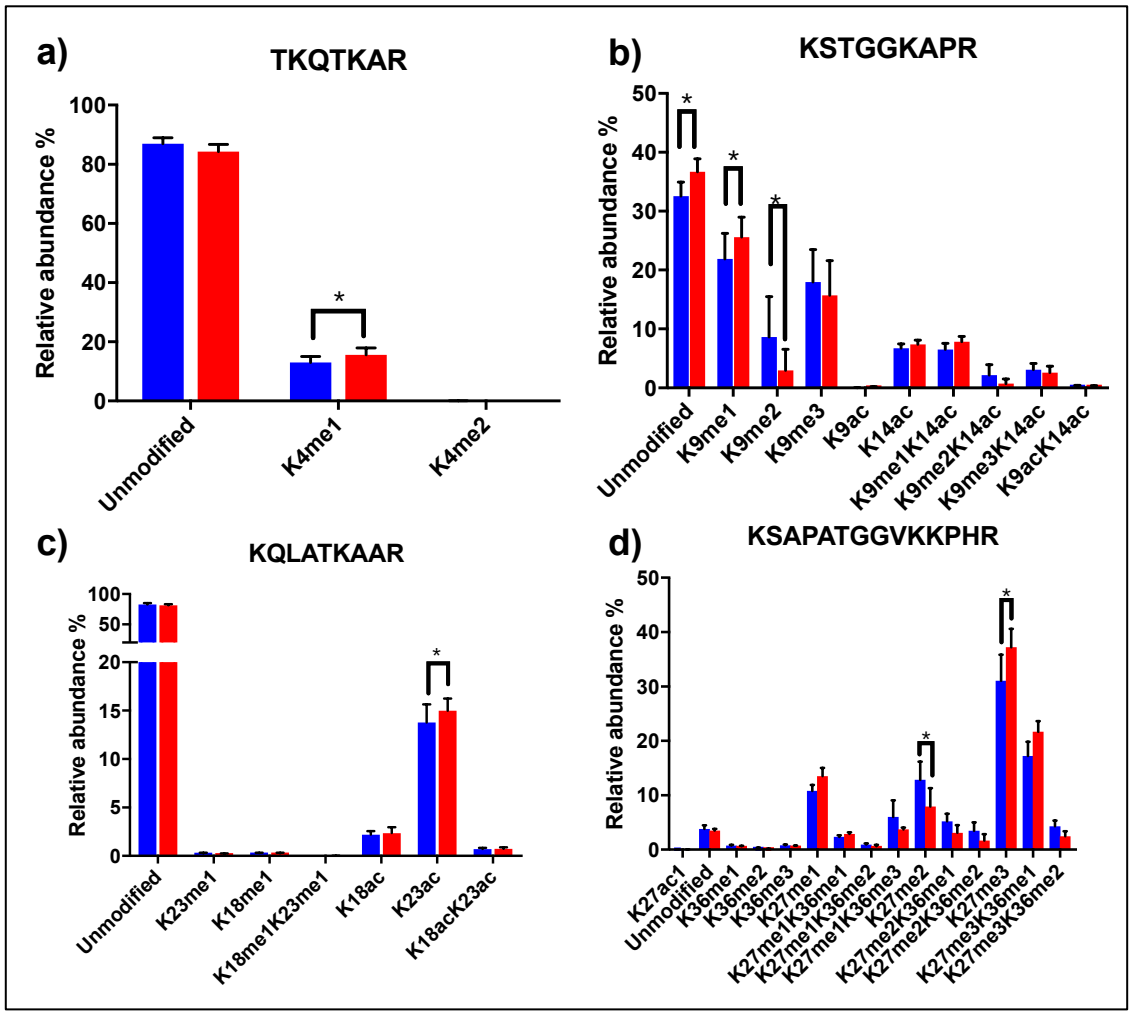


Figure 5.1: Changes in relative abundance of histone PTMs on histone H3 following 3 hour challenge with *S. pneumoniae* analysed on Orbitrap QExactive HF.

This figure illustrates the changes in the relative abundance of the PTMs observed for the different peptides of Histone H3 in blue are the abundances of the mock infected samples and in red those of the samples challenged with *S. pneumoniae*. Panel a) represents the decrease in H3K4me of the TQTAR peptide (n=7, t tests, * $p<0.05$), Panel b) highlights the decrease in H3K9me2 and reciprocal increase in H3K9me1 and the unmodified form of the KSTGGKAPR peptide (n=8, t test * $p<0.05$), Panel c) illustrates the increase in H3K23ac of the KQLATKAAR peptide (n=8, t tests, * $p<0.05$), Panel d) represents the increase in H3K27me3 and reciprocal fall in the abundance of H3K27me2 (n=8, t tests, * $p<0.05$).

Analysis of the PTMs on histone H3.3 revealed that on the identified KSAPSTGGVKKPHR peptide, the relative abundance of lysine 36 dimethylation (H3.3K36me2) was significantly increased (Figure 5.2).

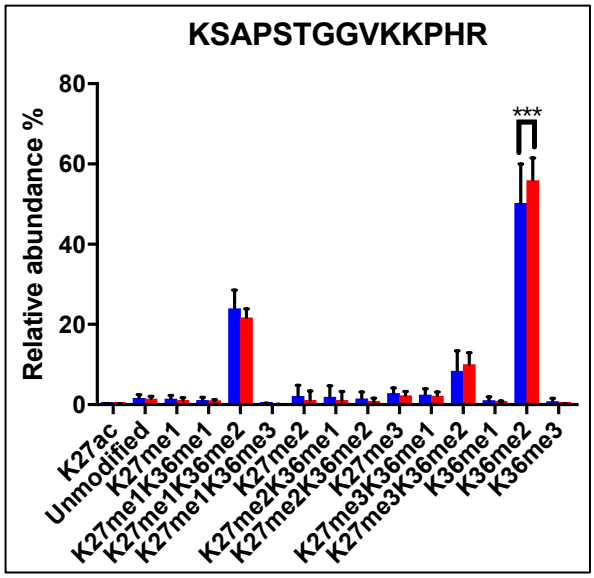


Figure 5.2: Changes in relative abundance of histone PTMs on histone H3.3 following 3 hour challenge with *S. pneumoniae* analysed on Orbitrap QExactive HF.

This figure illustrates the relative abundance of all of the PTMs observed for the KSAPSTGGVKKPHR peptide of Histone H3.3. n=8, t tests, *** p <0.001, q<0.001, in blue are the abundances of the mock infected samples and in red those of the samples challenged with *S. pneumoniae.*

On histone H4 (Figure 5.3) there was an increase in the monomethylation of lysine 20 (H4K20me1) and a reciprocal decrease in the dimethylated form (H4K20me2). There was also a trend towards increase in the abundance of H4K16ac, however this failed to reach significance in this subset of samples.

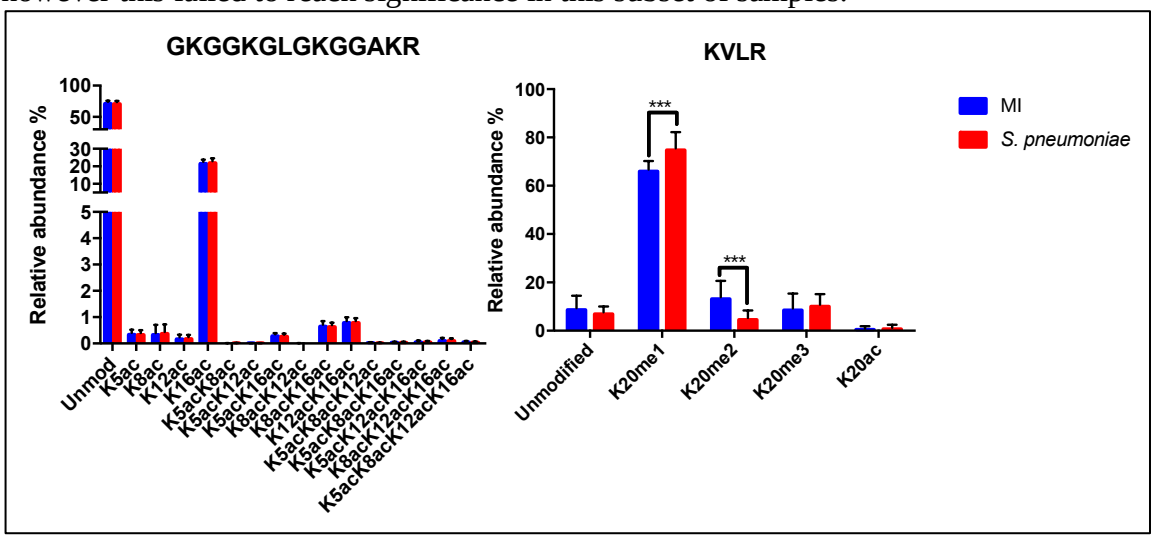


Figure 5.3: Changes in relative abundance of histone PTMs on histone H4 following 3 hour challenge with *S. pneumoniae* analysed on Orbitrap QExactive HF.

This figure illustrates the relative abundance of all of the PTMs observed for the different peptides of Histone H3. n=8, t tests, *** p <0.001, in blue are the abundances of the mock infected samples and in red those of the samples challenged with *S. pneumoniae.*

This data therefore confirms that the relative abundance of several different histone PTMs change in response to challenge with *S. pneumoniae* in two separate experiments. As discussed in Chapter IV, several of these marks have been associated with the control of gene expression and thus may play a role in host's response to infection.

5.2 Transcriptomic analysis of MDMs challenged with *S pneumoniae*:

Having established the changes occurring at a global histone PTMs level, next the changes in gene transcription were established in paired samples. RNA-seq was performed as described in Chapter 2. Briefly, following 3 hour challenge of MDMs with *S. pneumoniae* or mock infected, RNA was extracted from samples lysed in Tri Reagent using Direct-Zol™ extraction kits as per the manufacturer's protocol. The extracted RNA was then transferred to the core facility who prepared the libraries and sequenced the samples on the HiScan™ SQ (as described in Chapter 2).

Quality control of the RNA-seq runs was performed by FASTQC (as described in chapeter 2). This demonstrated that overall the RNA-seq raw data was of good quality. Read depth was estimated based on number reads made as between 2-3X (read depth= (fragment length x number of reads) / genome length). Overall this conforms with the number of reads required by ENCODE for an RNA-seq experiment to compare abundances of transcripts between two conditions. In addition, the mean coverage depth was measured using QualiMap and was found to be 22. This takes into account the number of reads per genomic region and therefore represents a better estimate of depth of sequencing as it isn't influenced by the regions of the genome not identified.

Initially, the samples were analysed following the "new Tuxedo protocol" (Pertea et al., 2016), which uses HISAT2 (Kim et al., 2015) to align the samples to the reference HG38 genome, then Stringtie (Pertea et al., 2015) was used to generate the transcript count tables and finally differential expression was assessed using Ballgown in R. This identified 197,782 transcripts of which when performing differential expression only 199 had a p value of less than 0.001. Unfortunately, none of the transcripts identified were differentially regulated after multiple test correction is taken into consideration (FDR <0.05) (Appendix table 5.2 shows the top 10 transcripts).

Therefore, in order to assess differential expression at gene level rather than look for novel transcripts, I used the quasi-mapping to the reference transcriptome approach from Salmon (Patro et al., 2017) to generate transcript level count tables and assessed differential expression using EdgeR (Robinson et al., 2010).

This approach identified 155 010 transcripts (76% of the 203 903 known transcripts in Ensembl). Of these, when performing differential expression analysis, 347 had a p value of less than 0.05, and 47 were considered to be differentially expressed, as the transcripts had a q value of less than 0.05 (FDR) (Table 5.1).

In order to establish the total number of genes identified at least once by the RNA-seq the Salmon transcript count data was converted to gene level count data using txtimport. This revealed 24 104 genes identified in at least one sample (61% of human genes of the 39 297 in the Ensmbl biomart database used to convert transcript to genes).

Table 5.1: Differentially expressed transcripts.

Transcripts identified following quasi-mapping to the RefSeq transcriptome and differential expression assessed in EdgeR, all upregulated following infective challenge (n=3):

RefSeq ID	Gene name	q value	Fold change Spn to MI

NR_033926.1	actin gamma 1 pseudogene 20	0.01961	-3.79
NM_021006.5	C-C motif chemokine ligand 3 like 1 variant 1	0.00193	-9.16
NR_111964.1	C-C motif chemokine ligand 3 like 1, variant 2	0.00232	-8.98
NR_111969.1	C-C motif chemokine ligand 4 like 1 variant CCL4Ldelta2	0.03280	-7.51
NM_001565.3	C-X-C motif chemokine ligand 10	0.00000	-15.59
NM_001199880.1	early growth response 3, transcript variant 2	0.01521	-3.83
NM_001199881.1	early growth response 3, transcript variant 3	0.01521	-3.83
NM_004430.2	early growth response 3, transcript variant 1	0.01421	-3.77
XM_005273425.3	early growth response 3 (EGR3), transcript variant X1	0.01521	-3.83
XM_005273426.3	early growth response 3 (EGR3), transcript variant X3	0.01521	-3.84
XM_011544429.2	early growth response 3 (EGR3), transcript variant X2	0.01557	-3.82
NM_015675.3	growth arrest and DNA damage inducible beta	0.00011	-3.97
XM_017026822.1	Nuclear growth arrest and DNA damage inducible beta (GADD45B), transcript variant X1	0.00010	-3.98
NM_000619.2	Interferon gamma	0.00000	-207.52
XR_937617.2	uncharacterized LOC105372754	0.00012	-6.65
XR_001754988.1	uncharacterized LOC107985489	0.00732	-5.17
XR_001756620.1	uncharacterized LOC107987437	0.01521	-7.26
NR_001458.3	MIR155 host gene	0.00752	-3.51
NM_001202233.1	nuclear receptor subfamily 4 group A member 1, transcript varint 3	0.00193	-2.99
NM_001202234.1	nuclear receptor subfamily 4 group A member 1, transcript varint 4	0.00247	-2.90
NM_002135.4	nuclear receptor subfamily 4 group A member 1, transcript varint 1	0.00193	-2.99
NM_173157.2	nuclear receptor subfamily 4 group A member 1, transcript variant 2	0.00171	-3.03
XM_005268822.3	nuclear receptor subfamily 4 group A member 1 (NR4A1), transcript variant X1	0.00230	-2.95
XM_005268824.3	nuclear receptor subfamily 4 group A member 1 (NR4A1), transcript variant X4	0.00211	-2.98
XM_006719363.1	nuclear receptor subfamily 4 group A member 1 (NR4A1), transcript variant X3	0.00193	-2.98
XM_006719364.3	nuclear receptor subfamily 4 group A member 1 (NR4A1), transcript variant X5	0.00232	-2.96
XM_017019247.1	nuclear receptor subfamily 4 group A member 1 (NR4A1), transcript variant X2	0.00193	-3.00
NM_006186.3	nuclear receptor subfamily 4 group A member 2	0.00193	-2.81
XM_005246621.3	uclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X1	0.00211	-2.84
XM_005246622.3	Homo sapiens nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X6	0.00247	-2.91
XM_006712553.3	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X3	0.00300	-2.76
XM_011511246.1	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X5	0.00247	-2.76
XM_017004219.1	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X2	0.00247	-2.83
XM_017004220.1	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X4	0.00353	-2.73
XR_001738751.1	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X8	0.00294	-2.71
XR_001738752.1	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X9	0.00391	-2.79
XR_427087.3	nuclear receptor subfamily 4 group A member 2 (NR4A2), transcript variant X7	0.00284	-2.74
NM_002852.3	Pentraxin 3	0.00353	-4.94
NR_002716.3	U2 small nuclear 1	0.00006	-7.51

NR_004394.1	**U6 small nuclear 1**	**0.00000**	**-13.62**
NR_125730.1	**U6 small nuclear 2**	**0.02262**	**-3.11**
NR_104084.1	**U6 small nuclear 7**	**0.00000**	**-12.20**
NR_104088.1	**U6 small nuclear 8**	**0.00000**	**-12.20**
NR_104080.1	**U6 small nuclear 9**	**0.02262**	**-3.11**
NM_000594.3	**Tumour necrosis factor**	**0.00193**	**-13.23**
NM_001178096.1	**Coagulation factor III, tissue factor (F3) transcript variant 2**	**0.01943**	**-2.56**
NM_001993.4	**Coagulation factor III, tissue factor (F3) transcript variant 1**	**0.02060**	**-2.54**

The RNA-seq analysis revealed 47 transcripts that were differentially expressed after challenge with *S. pneumoniae* ($q<0.05$). These relate to 18 different genomic locations. Given the low number of genes, pathway analysis was not performed.

The C-C motif chemokine ligand 3 like 1(CCL3L1) gene encodes the chemotactic cytokine known as macrophage inflammatory protein (MIP-1α/LD78β). It attracts both lymphocytes and monocytes as part of the inflammatory response. It has been shown to potently bind to chemokine receptor 5 (CCR5) but also receptors 1 and 3 (Struyf et al., 2001). It therefore plays a key role in the host response to the macrophage tropic Human immunodeficiency virus-1 infection (Proost et al., 2000). Importantly it is upregulated during pneumococcal infection and associated with macrophage recruitment (Fillion et al., 2001).

C-C motif chemokine ligand 4 like 1(CCL4L1) gene encodes the chemotactic cytokine know as MIP-1β. It was first identified following endotoxin stimulation of macrophages (Wolpe et al., 1988). Like MIP-1α, it binds to CCR5 cell surface receptor and acts as a chemoattractant and pro-inflammatory cytokine. Both CCL3L1 and CCL4L1 can competitively bind to CCR5 thereby impairing HIV-1 entry into cells (Shao et al., 2007). It is also upregulated by pneumococcal infection and is linked to macrophage recruitment (Fillion et al., 2001).

The C-X-C motif chemokine ligand 10 (CXCL10) is a C-X-C chemokine that is secreted in response to inflammatory stimuli, it acts as a chemoattractant promoting inflammation and binds to the CXCR3 receptor (Liu et al., 2011). It has been shown to play a crucial role in the pathogenesis of a number of different infections such as *Legionella pneumophilia.* In murine models CXCL10 has been shown to be one of the most highly up regulated genes following an intranasal *S. pneumoniae* infection, highlighting the importance of these chemokines in the host response to infection. More recently, it has been shown to regulate the production of pro-inflammatory cytokines in monocytes by binding to the CXCR3 receptor and initiating downstream p38 MAPK signalling (Q. Zhao et al., 2017).

Early growth response 3 (EGR3) is a member of the EGR transcription factor family whose members are involved in the regulation of growth and differentiation (Tourtellotte and Milbrandt, 1998). It was initially seen to be induced by growth factors (Patwardhan et al., 1991). EGR3 deficient mice were found to develop sensory ataxia as it was shown to be critical for muscle spindle development (Tourtellotte and Milbrandt, 1998). More recently, it has been shown to play an important role in the regulation of transcription of inflammatory genes in T cell activation by interacting with NFKB (Safford et al., 2005; Wieland et al., 2005). Furthermore, deficiency in EGR-2 and EGR-3 in B and T cells results in a lethal autoimmune syndrome (Li et al., 2012). Moreover, adhesion of *S. pneumoniae* to pharyngeal epithelial cells has been shown to induce the upregulation of EGR-3 amongst other transcription factors (Bootsma et al.,

2007). EGR-3 has been found to be up-regulated following infection of human monocytes with different gram positive bacteria such as *Staphylococcus aureus, S. pneumoniae* and *L. monocytogenes* (Tchatalbachev et al., 2010). This highlights the importance of EGR-3 in the regulation of the inflammatory response.

Growth arrest and DNA damage inducible beta (GADD45B) is involved in the response to DNA damage in particular in response to environmental stresses (Jiang et al., 2016) and mediates apoptosis via p38/JNK pathway (Takekawa and Saito, 1998). GADD45B has been shown to play an important role in the regulation of the inflammatory response in mice (Salerno et al., 2009). The authors showed that GADD45B deficient mice were more susceptible to septic shock from low dose LPS challenge. It has a role as a stress sensor in mice. Indeed, in an intraperitoneal sepsis model, GADD45B deficient mice recruited less macrophages to the peritoneum. Furthermore, the bone marrow derived macrophages from GADD45B deficient mice produced less ROS in response to LPS, highlighting its importance in the host's innate immune response (Salerno et al., 2012). It has been shown to be suppressed in a number of cancers (Ou et al., 2015; Zerbini et al., 2004).

Interferon gamma (IFN-γ) is a key component of both the adaptive and innate immune responses (Schroder et al., 2004). It has been shown to be secreted by monocyte derived macrophages in response to IL-12 and IL-18 co-stimulation (Darwich et al., 2009). IFN-γ has been demonstrated to play a number of different immunomodulatory roles in particular activating macrophages and leading to T cell differentiation (Gessani and Belardelli, 1998) but also priming for the production of antimicrobial components such as NO or ROS; therefore demonstrating its pivotal role in the host-pathogen response. Furthermore, IFN-γ has been shown to be essential to NO production in response to pneumolysin stimulation in murine macrophages (Braun et al., 1999).

microRNA 155 host gene (miR155) is a small non protein coding RNA (Elton et al., 2013). It plays an important role in post translational control of gene expression by interacting directly with messenger RNA for example (Guo et al., 2010). Macrophages were found to up-regulate the production of miR155 in response to TLR, TNF-α or IFN-γ stimulation (O'Connell et al., 2007). Additionally, the production of pro-inflammatory cytokines was found to be reduced in miR155 deficient mice (Jablonski et al., 2016). It has also been shown to be critical for clearance of *S. pneumoniae* from the nasopharynx of mice (Verschoor et al., 2014). miR-155 has been implicated in a number of different cancers and it has been postulated that it may represent a link between aberrant inflammatory processes and pro-oncogenic state (Elton et al., 2013).

Nuclear receptor subfamily 4 group A member 1 (NR4A1, also known as Nur77) is a member of the steroid thyroid hormone receptor family (Ryseck et al., 1989). It has been shown to play a number of metabolic roles in skeletal muscle cells and hepatocytes (Zhan et al., 2012). It has been demonstrated to bind to liver kinase B1 and lead to attenuation of AMP-activated protein kinase (AMPK) thereby modulating glucose metabolism (Zhan et al., 2012). Moreover, in T cells NR4A1 has been shown to have a pro-apoptotic role by interacting with Bcl-2 (Lin et al., 2004). Furthermore, the regulation of activated T cells has been shown to play a role in bacterial clearance in the lung (Marriott et al., 2012). Nuclear receptor subfamily 4 group A member 2 (NR4A2) also known as Nurr1, along with NR4A1 have been demonstrated in macrophages to play important roles in the regulation of inflammation as they are induced by a number of different pro-inflammatory ligands such as LPS (Pei et al., 2005). More recently it has been demonstrated that NR4A1, and NR4A2 to a lesser extent, induce a number of inflammatory cytokines such TNF-α via modulation of the NFKB pathway (Pei et al.,

2006), highlighting the important role played by these nuclear receptors in the host-pathogen interaction. NR4A2 has also been implicated in the polarisation of macrophages towards an alternative or M2 phenotype (Mahajan et al., 2015, p. 2) suggesting that it may play a role in controlling inflammation and resolution of the inflammatory response.

Pentraxin 3 (PTX3) is a secreted protein that plays a critical role in the innate immune response (Bottazzi et al., 2010). It is involved in resistance to a number of different infections including bacterial (Soares et al., 2006), influenza A virus (Reading et al., 2008) and fungal infections (Garlanda et al., 2002). It modulates the production of NO and TNF-α during the inflammatory response (Perea et al., 2017; Soares et al., 2006) and therefore plays a key role in the resolution of inflammation in macrophages (Shiraki et al., 2016) and the response to *S. pneumoniae* (Cole et al., 2014; Kadioglu et al., 2008b).

The transcripts for both U2 and U6 components of the spliceosome were also upregulated by the bacterial challenge which would be expected given the increase in gene transcription and therefore the need for post-translational modification of pre-mRNA (Will and Lührmann, 2011). It has been suggested that during infection of macrophage alternative splicing events can occur to confer increased survival for intracellular pathogens such as *M. tuberculosis* (Kalam et al., 2017).

Coagulation factor III, also known as tissue factor is a cell surface glycoprotein that playa a critical role in the coagulation cascade (Nemerson, 1988) and as such is an important component of the immune response.

Finally, it is reassuring to see that despite the low number of differentially expressed transcripts, the majority of them relate to the host's immune response with upregulation of pre-dominantly pro-inflammatory cytokines such as TNF-α, CCL4 and CCL3. The overall low numbers of differentially transcription may reflect the early time point, and as I would expect only a subset of cells will have ingested *S. pneumoniae*, it is possible that the overall effect is masked by non-reactive bystander cells.

Moreover, the overall overlap in the RNA-seq and microarray analyses shows that the 64% of the differentially expressed genes from the microarray analysis were also identified in the RNA-seq analysis.

5.3 Chromatin precipitation and sequencing of key histone PTMs:

In order to gain further insight into the role played by the changes in global relative abundance of the histone PTMs observed by MS, ChIP-Seq was employed. Given the potential role played by H3K4me1, H3K27me3 and H3K9me2 as part of the regulation of gene transcription these marks were selected for further analysis. Despite the use of HFBA in the loading buffer the overall abundance of H3K4me3 and H3K27ac marks remains low and I was not able to distinguish any significant differences in their relative abundance after infective challenge. However, given the pivotal role played of both of these marks in the regulation of transcriptional responses (Creyghton et al., 2010; Jenuwein and Allis, 2001), and as they have been widely studied by ChIP-Seq and are difficult to quantitate by mass spectrometry, they too were taken forward for further analysis.

5.3.1 Optimisation of ChIP-Seq workflow: Sonication efficiency:

The success of ChIP-Seq relies on several factors; adequate fragmentation of the chromatin to be able to sequence it efficiently, and antibody specificity are two key components (Landt et al., 2012).

Therefore, the efficiency of the sonication used to shear the crosslinked DNA fragments was first validated by assessing the effect of sonication prior to agarose gel electrophoresis analysis (after reversal of the crosslinking as described in Chapter 2 on an agarose gel (Figure 5.4).

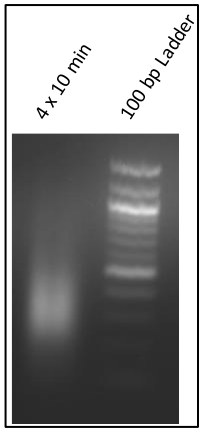


Figure 5.4: Validation of sonication efficiency.
This 2% agarose gel demonstrates the efficiency of the chromatin sheering after 4 x 10min cycles, 30s on / off on high output. 4 x 10^6 cells were lysed and the chromatin sheered 10% of the sample was ran on the 2 % gel after having cross link reversal, RNAase A and Proteinase K treatment, then PCR clean up kit.

The results show that using 4 x 10 minute cycles of 30 seconds on / off on the high output setting, resulted in efficient shearing of the DNA to the correct size range of 200bp. In order to further validate the sonication efficiency, prior to pull down each sample was also analysed on a Bioanalyser prior to sequencing (Figure 5.5). the results show that the majority of the sheared DNA was around 200bp in length.

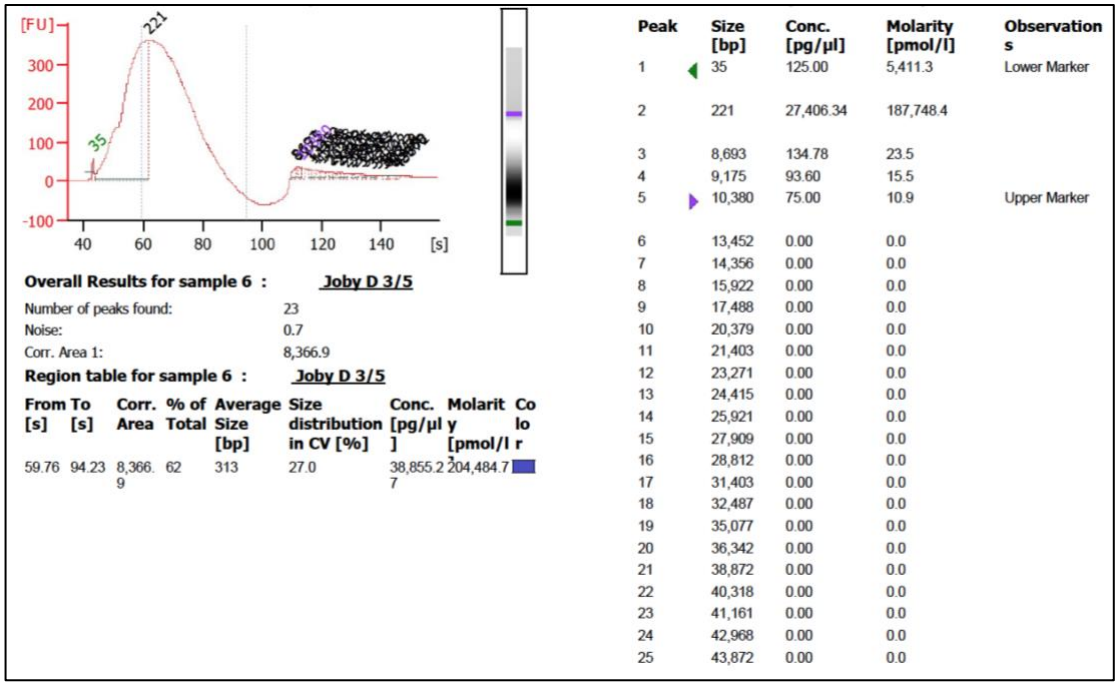


Figure 5.5: Analysis of the DNA fragments after sonication using the Bioanalyser.

This figure shows the bioanalyser analysis results and illustrates the length of the DNA fragments after sonication demonstrating that the majority of DNA fragments are around 221 bp.

5.3.2 Antibody validation:

ChIP-Seq data is only as good as the antibody used to perform the pull downs (Landt et al., 2012). Therefore, only antibodies previously used by the Encode consortia were used (Table 2.17). Furthermore, western blots were carried out for each antibody to ensure only a single band was detected (Figure 5.6).

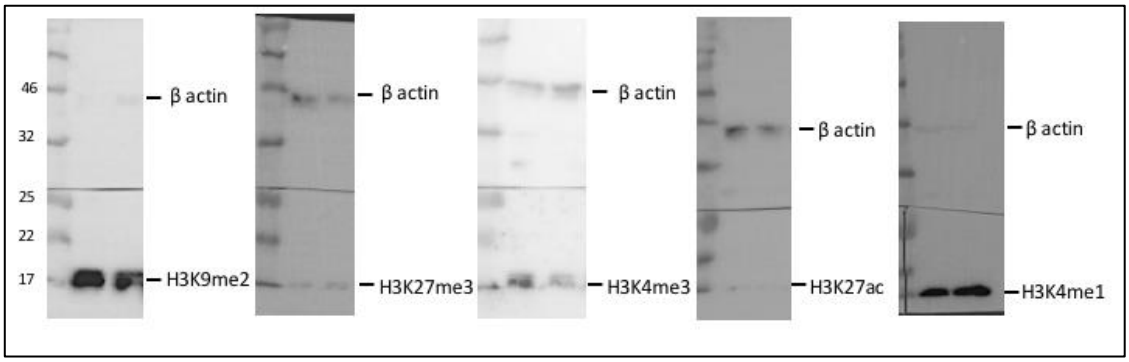


Figure 5.6: Western blot validation of the ChIP grade antibodies used for the ChIP.

5 μL of MDMs lysed in Laemli buffer were run 12% SDS page gel. After semidry transfer, the membranes were incubated overnight with each antibody as described in Chapter 2. This shows a clear single band at the level expected for histone H3.

5.4 ChIP-qPCR:

To confirm that the chromatin immunoprecipitations had worked and were able to enrich DNA compared to non-specific pull downs (beads only), I carried out a series of pull downs followed by qPCR. The primers chosen were either house keeping genes (GAPDH, Actin, OAZ1) or genes known to be targets of H3K9me2 for example Insulin (Mutskov et al., 2007), MYOD10, STAT1 and CCL3 (Yoshida et al., 2015). The initial pull downs were carried out comparing total H3 antibody to beads only and then looking for enrichment in the house keeping genes (Figure 5.7).

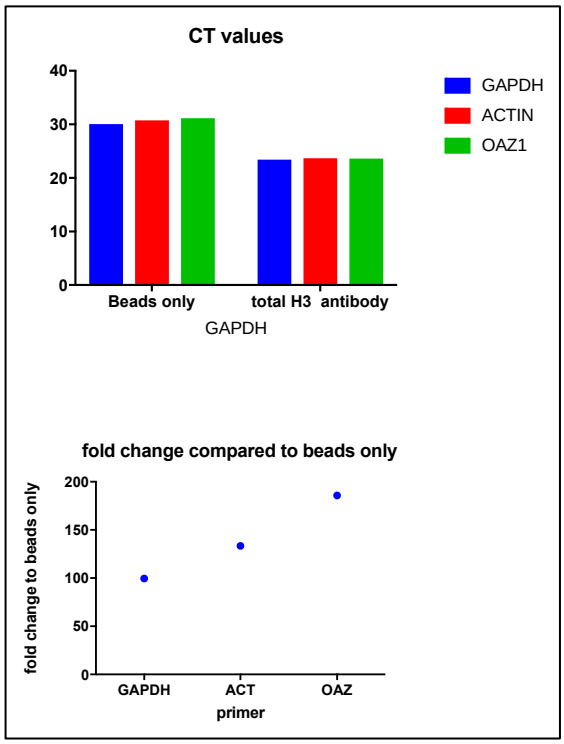


Figure 5.7: Validation of ChIP using qPCR.
ChIP was performed using total H3 antibody to enrich DNA fragments in conjunction with qPCR to detect a number of house keeping genes. This figure demonstrates successful enrichment of the pulldown DNA compared to the non-specific binding seen in the beads only sample as the cycle threshold (CT) values are significantly lower in the total H3 pull down. This in turn corresponds to significant enrichment compared to non-specific binding.

Next, the individual antibodies were assessed to see if enrichment was detectable compared to non-specific binding (Figure 5.8).

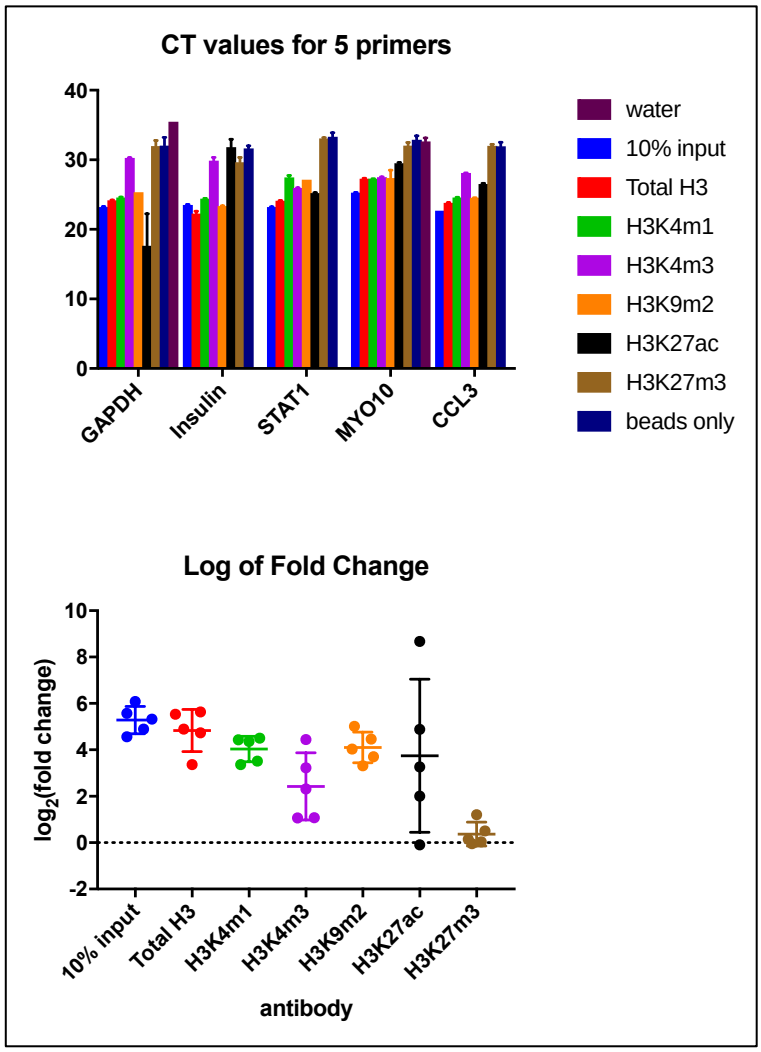


Figure 5.8: ChIP-qPCR validation of the antibodies.
This figure demonstrates that successful enrichment was possible for all of the antibodies. Unfortunately, there was some non-specific signal see in the water only qPCR for the GAPDH and MYOD10 primers although this had a greater CT value than the non-specific binding seen with beads only. As H3K27me3 is thought to be predominantly associated with repressed gene expression the lack of enrichment with all but the primer for insulin maybe due to these genes being expressed in our cells. Furthermore, the primers were chosen to look for known targets of H3K9me2.

Finally, the samples taken forward for ChIP-Seq also had an aliquot run by qPCR to ensure that they had been successfully enriched (Figure 5.9).

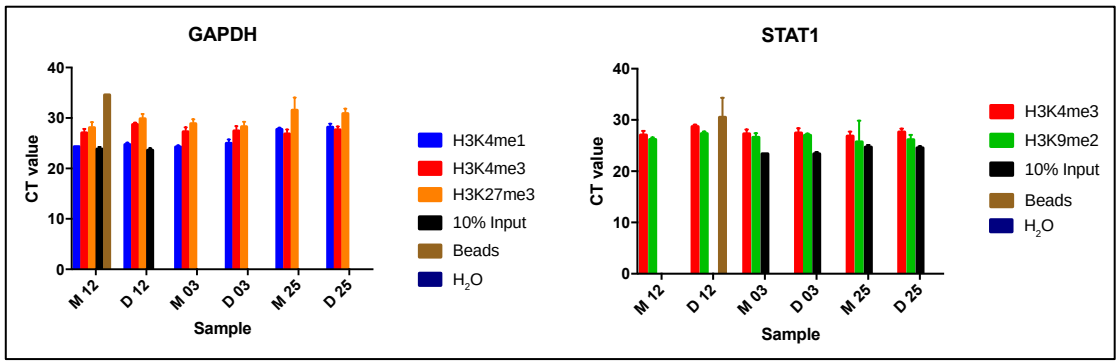


Figure 5.9: ChIP-qPCR QC prior to NGS.
These bar charts illustrate the CT values for the qPCR following the 6 different ChIP of 6 samples prior to library preparation and sequencing. In an effort to constrain costs only 2 "beads only" samples were run. The left hand panel shows the results of pull downs for H3K4me1, H3K4me3 and H3K27me3 and 10% input samples (for the first two samples) using GAPDH primers. The right had panel shows those resulting from the H3K4me3, H3K9me2 pull downs and 10% input samples for the last four samples using primers for STAT1. With each qPCR pull down a "bead only" pull down was included to assess non-specific binding, and a negative control (water).

5.4 ChIP-Seq analysis:

Following optimisation of the ChIP procedure, chromatin immunoprecipitations for H3K4me1, H3K4me3, H3K9me2, H3K27ac and H3K27me3 were performed for 3 biological replicates of either mock infected or cells challenged with *S. pneumoniae* (as described in Chapter 2). The precipitated DNA was then indexed and library preparation performed by the core facility (SITraN, University of Sheffield). A selection of the pull downs were assessed after the libraries were prepared using the Bioanalyser to ensure appropriate size selection had occurred (Appendix figure 5.1) prior to high throughput DNA sequencing using a HiSeq 4000 (Illumina) performed by Edinburgh Genomics (University of Edinburgh). The samples were pooled and run as described in Chapter 2. The ChIP-Seq of histone PTMs has identified a number of PTMs to have distinct profiles, some are considered to form narrow peaks (such as H3K27ac), others form broader peaks (such as H3K9me2) (Landt et al., 2012). In order to achieve the required depth of sequencing suggested by Encyclopaedia of DNA elements (ENCODE), the broad marks (H3K9me2 and H3K27me3) were run across two lanes. The remaining 3 narrow ones (H3K4me1, H3K4me3, H3K27ac) and the input controls were run on another two lanes. The raw data was then processed in collaboration with the Advanced Data Analysis Centre of the University of Nottingham.

5.4.1 ChIP-Seq analysis and quality controls:

Briefly, the FASTQ files for each lane were assessed using FastQC (version 0.11.7, https://www.bioinformatics.babraham.ac.uk/projects/download.html#fastqc). This showed that the median quality scores were greater than 25 (Appendix figure 5.2). The reads from both lanes were then concatenated, and trimmed using Trimomatic (Bolger et al., 2014), alignments were made using Burrows-Wheeler Alignment tool (BWA) (Li and Durbin, 2009), samples were deduplicated using Picard-tools (https://github.com/broadinstitute/picard). Peak calling was made using both narrow and broad peak modes of Model-based Analysis of ChIP-Seq (MACS2) (Zhang et al., 2008), then the peaks were assessed using phantompeakqualtools (https://code.google.com/archive/p/phantompeakqualtools/source/default/source). This showed that the majority of normalised strand coefficients (NSC) were greater than 1.05 and the relative strand correlation (RSC) was greater than 0.8 (Appendix figure 5.3) suggesting that the ChIP had been successful (Landt et al., 2012). Finally, ChipQC (Carroll et al., 2014) in R was used to characterise the peak picking (summarised in Appendix figure 5.4).

5.4.2 ChIP-Seq analysis and genomic distribution of peaks:

To further characterise each pull down, I also used ChIPseeker (Yu et al., 2015) to annotate and visualise the peaks. This revealed that the H3K27ac and H3K4me3 pull downs had sequences that were predominantly located at or near the transcription start

site (TSS) as demonstrated in Figure 5.10 (and Appendix figures 5.5-6), whereas H3K27me3 and H3K9me2 had predominantly distal intergenic occupancy.

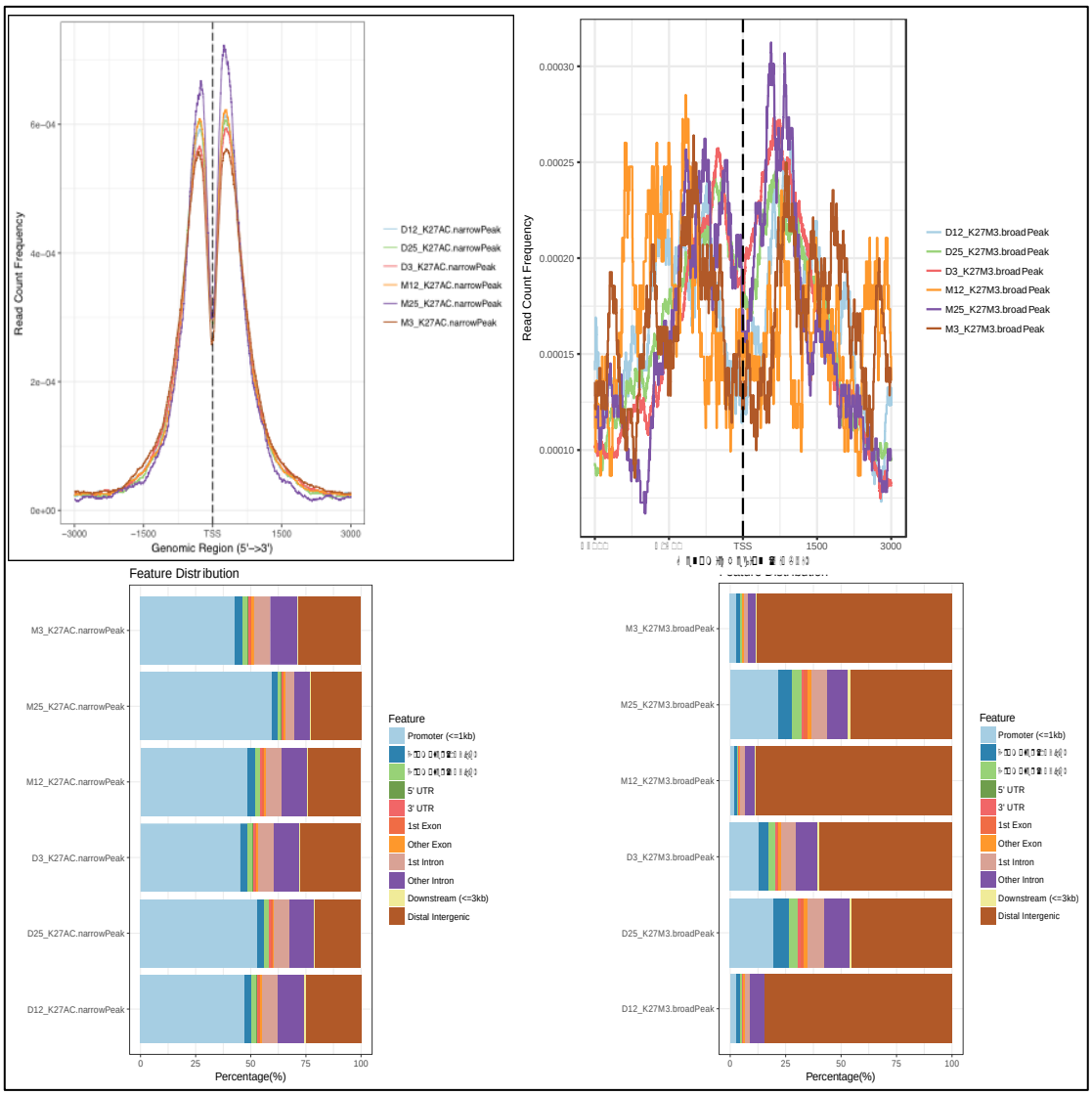


Figure 5.10: Distribution of reads relative to TSS of genes.
Using ChIPseeker, the reads were mapped to genomic locations. This figure illustrates that the majority of the peaks identified following H3K27ac (left side) ChIP-Seq were within 1kb of the TSS. On the other hand (right side) the results from the H3K27me3 ChIP-Seq illustrates that the majority of the peaks are distal intergenic.

5.4.3 ChIP-Seq differential analysis and visualisation:

Differential enrichment for both the peaks picked in narrow and in broad mode was performed using the R package Diffbind (Ross-Innes et al., 2012). This generated a number of BED files containing the differentially bound regions. In addition, using Bedtools (Quinlan and Hall, 2010) a consensus file for each of the pull downs for both the mock infected and the *S. pneumoniae* challenge samples were created. The data was visualised in Integrated Genomics Viewer (IGV) as illustrated in Figure 5.11 (Robinson et al., 2011).

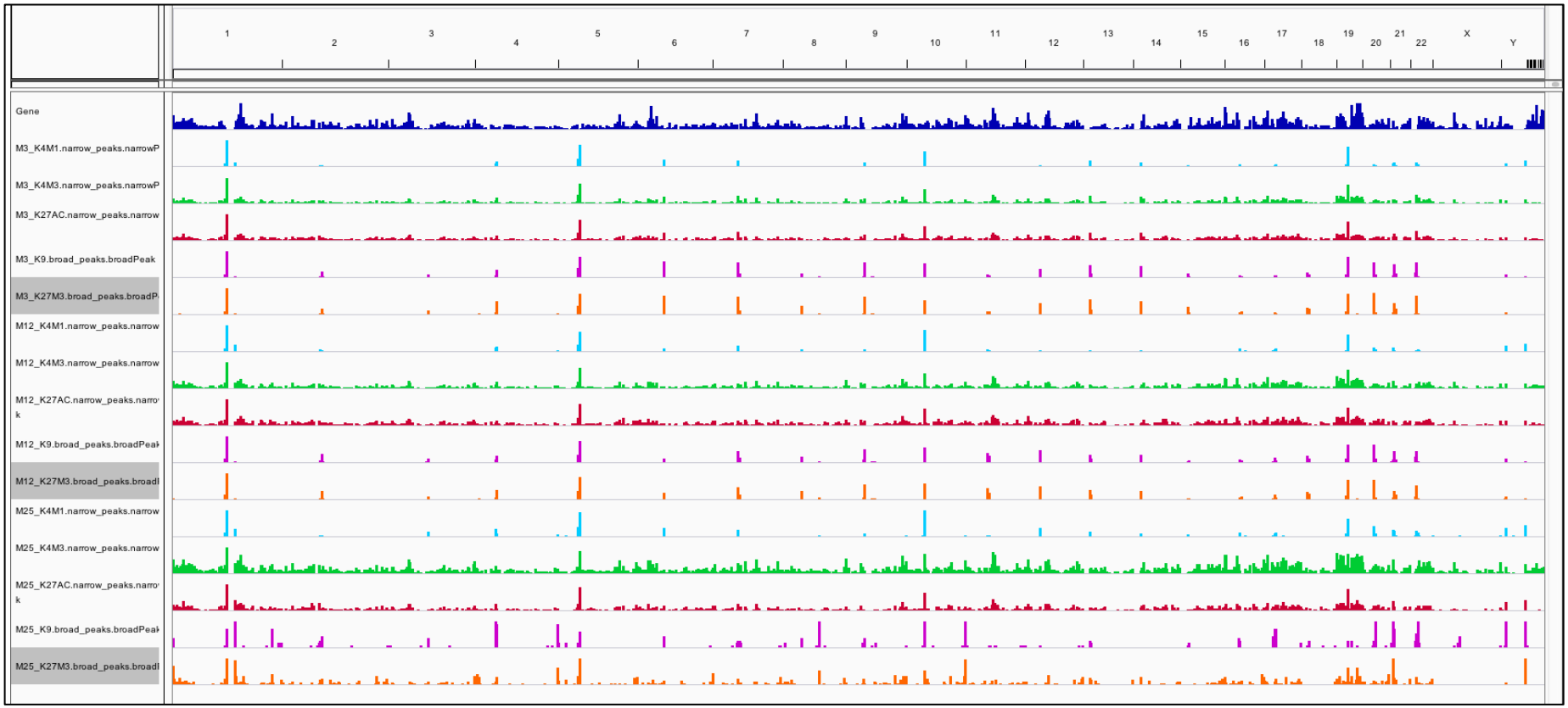


Figure 5.11: IGV genome wide view of the ChIP-Seq results for the mock infected samples.
This genome browser view illustrates the peaks found following the ChIP-Seq for each mark for the mock infected (MI) samples. This highlights the greater number of peaks identified with H3K4me3 (green track) and the H3K27ac (red track) when compared to the pull downs for H3K9me2, H3K27me3 and H3K4me1.

5.4 ChIP-Seq results: Analysis of differentially enriched genomic locations

The analysis of the ChIP-Seq showed that overall the data was of good quality (Appendix figures 5.1-4). The differentially enriched peaks between the two conditions (mock infected and challenged with *S. pneumoniae*) however, only identified a few differentially enriched regions after multiple test correction (FDR<5%).

ChIP-seq suffers from a number of limitations in particular in relation to sensitivity. The overall ability to correctly identify changes in relative abundance of histone PTMs at a given genomic location is governed a number of different factors. Firstly, the quality of the antibody used and its ability to bind to the histone PTM, not be displaced by neighbouring PTMs such as is the case in epitope occlusion and or cross reactivity with other PTMs. It is also possible that the conformation of the chromatin may not allow binding of the antibody to a PTM. Furthermore, the ability to distinguish small changes occurring at a specific location may be below the detection limit of these novel next generation sequencing methods. Moreover, there is a lack of consensus on the interpretation of multiple replicates within the field. Additionally, a global decrease in a PTM occurring across the entire genome may not identify distinct sites of change by ChIP-seq. Finally, it is likely that several layers of redundancy exist within the histone PTM landscape and as a result it is probable that multiple PTMs act in consort to perform their role, leading to dangers of over interpretation if individual PTMs are studied in isolation. Taken together, these factors dictate that some caution be used in the interpretation of the ChIP-seq data and further validation studies using ChIP-qPCR for example are warranted before making firm conclusions.

5.5 H3K27ac:

The ChIP-Seq analysis identified between 20 744 and 32 713 narrow peaks identified in 5 of the 6 samples with one outlier at 8 758.

There was one differentially enriched region for the narrow peak results for H3K27ac pull downs (Table 5.2 and Figure 5.12). This corresponded to chromosome 21 position 8 219 749 to 8 220 258, this is within the intron for NR_038958, also known as LOC100507412 and is a long non-coding RNA. It also happens to occur within a CpG island. The analysis of broad peak enrichment identified a region 10kb upstream of the Gap junction protein alpha 3 (GJA3 gene). Defects in this gene are associated with congenital cataract formation (Yao et al., 2017). It also identified a region within the intron for Corf159, also known as D-Glutamate cyclase (DGLUCY). The enzyme encoded for by this gene is responsible for the metabolism of D-glutamate (Ariyoshi et al., 2017).

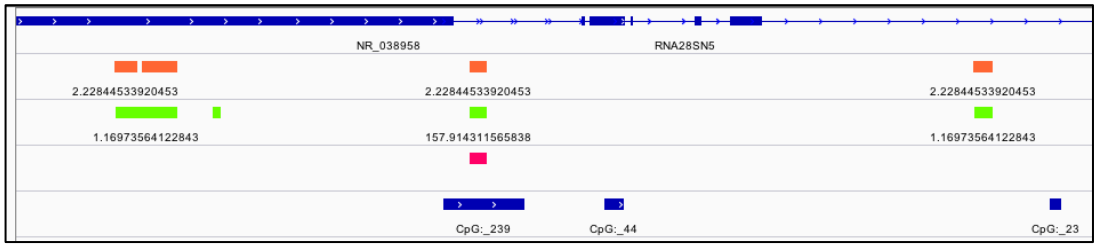


Figure 5.12: IGV visualisation of enriched region for H3K27ac pull downs.
This IGV figure illustrates the region on chromosome 21 that is differentially populated following bacterial challenge with an increase in H3K27ac within a CpG island in a long non-coding RNA. The top track in blue is the transcriptome information, in orange is the consensus BED file for the mock infected and in green for the *S. pneumoniae* challenged. Finally, the bottom track shows the location of the CpG islands. This approach was used to characterise the relationship with coding regions of the genome for all of the statistically differentially enriched regions.

As one of the samples had very few peaks, it is possible that this biological replicate skewed the results, therefore the differential analysis was repeated excluding this biological replicate (Table 5.2 orange rows). This resulted in 5 differentially enriched regions being identified when looking at the broad peak picking results. The first region corresponds to an intron for the long non-coding RNA (LOC101929709 also known as NR_125822). This site is also situated within 10kb of the start site for Receptor-interacting Serine/threonine-protein kinase 2 (RIPK2). RIPK2 is a Serine/threonine/kinase that is essential in the regulation of the innate immune response due to its involvement in the NOD / NFKB signalling pathway (Tigno-Aranjuez et al., 2010; Zhao et al., 2012). The first intron for stabilin 1 (STAB1) was also found to be differentially enriched. This gene is encodes the receptor protein stabilin 1. High levels of stabilin 1 have been associated with down-regulation of pro-inflammatory genes (Palani et al., 2016). The ChIP-Seq analysis suggests that there is less H3K27ac, a mark commonly associated with active gene transcription found at the start of this gene. The transcriptomic analysis shows that there is slightly less transcript (-0.327 fold change) in the bacterial challenge samples than the mock infected although this does not reach statistical significance ($p>0.05$). Finally, the region 60kb downstream for the pseudogene Ankyrin domain 26 pseudogene 1 was also differentially enriched. Ankyrin domain 26 protein is found within the inner cellular membranes of multiple tissues and is thought cooperate with signalling proteins (Raciti et al., 2011). Mutations in Ankyrin domain 26 gene have been associated with acute myeloid leukaemia (Marconi et al., 2017) and thrombocytopenia 2 (Bluteau et al., 2014).

Table 5.2: List of differentially enriched regions for H3K27ac pull downs.

PTM	n	Chr	Start	End	Fold	FDR	Gene	Gene body	Distance
H3K27ac Broad	**2**	**chr8**	**89727950**	**89730853**	**3.57**	**5.32E-05**	**NR_125822 / upstream of RIPK2**	**Intron**	
H3K27ac Broad	**2**	**chr3**	**52494678**	**52501213**	**-1.64**	**0.0199**	**STAB1**	**Intron**	
H3K27ac Broad	***2***	***chr13***	***20174031***	***20175733***	***3.47***	***0.000828***	***Upstream GJA3 / dowstream GJB2***	No	10kb / 10 kb
H3K27ac Broad	2	chr16	46398609	46401505	7.96	0.0379	Downstream from ANKRD26P1	no	60kb
H3K27ac Broad	2	chr5	49601597	49602776	7.85	0.0414	centromere	no	
H3K27ac narrow	2	chr16	46398519	46401602	8.03	0.0107	Downstream from ANKRD26P1	no	60kb
H3K27ac narrow	2	chr5	49601485	49602902	8.24	0.0101	centromere	no	

H3K27ac Broad	**3**	**chr8**	**89727950**	**89730853**	**3.17**	**6.33E-05**	**NR_125822 / upstream of RIPK2**	**Intron**	
H3K27ac Broad	**3**	**chr14**	**91220428**	**91223032**	**4.74**	**0.00985**	**C14orf159**	**Intron**	
H3K27ac Broad	**3**	**chr21**	**8219800**	**8220686**	**5.07**	**0.00985**	**NR_038958**	**Intron**	
H3K27ac Broad	***3***	***chr13***	***20174031***	***20175733***	***4.03***	***3.22E-05***	***Upstream GJA3 / dowstream GJB2***	No	10kb / 10 kb
H3K27ac narrow	**3**	**chr21**	**8219748**	**8220258**	**6.08**	**0.0019**	**NR_038958**	**Intron**	

5.6 H3K9me2:

There were between 4 604 and 13 210 broad peaks identified in 5 of the 6 samples with one outlier at 539.

Despite identifying 9 355 regions in both the mock infected and *S. pneumoniae* challenged samples following ChIP, none of these were differentially enriched after FDR correction.

5.7 H3K27me3:

There were between 12 653 and 30 103 broad peaks identified in 5 of the 6 samples with one outlier with only 1 851 peaks. There are 18 regions with differential H3K27me3 peaks following challenge with *S. pneumoniae* (Table 5.3).

There were two differentially enriched regions following analysis with all three samples; this is a region on chromosome 2 that is situated within 30kb of the start of the cytoskeleton regulator RNA gene (CYTOR) and another region on chromosome Y, 1mb away from the testis specific transcript Y linked 23 (TTTY23).

The differential analysis after excluding the outlying biological replicate (orange rows of Table 5.3) resulted in more regions of interest. One region was on chromosome 22, involving one of the exons for RNA, 45S Pre-Ribosomal N1 gene (RNA45SN1). This region encodes the 45S ribosomal RNA, an important constituent of the ribosome (Yu and Lemos, 2016). On chromosome 2, the intron for the SLIT homolog 2 protein coding gene (SLIT2) was found to be differentially enriched. This protein is found to play a number of different roles and has recently been implicated in the control of inflammation caused by LPS in endothelial cells (Zhao et al., 2014) and has been found to regulate monocyte adhesion to endothelial walls, hence playing a part in the regulation of inflammation (Mukovozov et al., 2015). On chromosome 9 the intron for the glutamate ionotropic receptor N-methyl-D-aspartate type subunit 1 (GRIN1) was differentially enriched. This gene encodes a portion of the N-methyl-D-aspartate receptor which plays a role in memory and learning. Mutations in GRIN1 have been associated with encephalopathies (Zehavi et al., 2017). A further region was on chromosome 12, located at the non-coding RNA transcript NR_120467, a gene of unknown significance.

On chromosome 14, the region upstream of neuronal Per-Arnt-Sim domain 3 (NPAS3), which has been implicated in the development of schizophrenia, is also differentially enriched (Macintyre et al., 2010). The region 10kb upstream of uridine-cytidine kinase 1 (UCK1) was also enriched. UCK1 phosphorylates uridine and cytidine (Meinsma and van Kuilenburg, 2016) and has been associated with neuroblastomas (van Kuilenburg and Meinsma, 2016). Finally, the region 6kb upstream of Human growth hormone 1 homolog (HGH1) gene was also found to be differentially enriched.

Table 5.3: Differentially enriched regions following H3K27me3 diffbind analysis.

PTM	n	Chr	Start	End	Fold	FDR	Gene	Gene body	distance
H3K27me3 Broad	3	chrY	11322543	11322787	3.22	0.0438	downstream TTY23 /	no	1mb
H3K27me3 Broad	***3***	***chr2***	***87423893***	***87425699***	2.47	0.0284	Upstream CYTOR	no	30kb
H3K27me3 Broad	**2**	**chr22_KI270733v1_random**	**133811**	**135280**	7.08	1.31E-05	RNA45SN1	intron	
H3K27me3 Broad	2	chr1_KI270712v1_random	58428	59525	5.96	0.00116			
H3K27me3 Broad	**2**	**chr4**	**20254532**	**20255349**	3.23	0.00557	SLIT2	exon	
H3K27me3 Broad	**2**	**chr9**	**137160768**	**137161470**	3.91	0.0166	GRIN1	exon	
H3K27me3 Broad	**2**	**chr12**	**132330028**	**132330311**	3.59	0.0244	NR_120467	exon	
H3K27me3 Broad	2	chrUn_KI270467v1	2051	3828	-0.73	0.0252			
H3K27me3 Broad	***2***	***chr14***	***32932478***	***32934557***	1.92	0.0152	Upstream NPAS3	no	10kb
H3K27me3 Broad	***2***	***chr9***	***131545680***	***131546557***	2.2	0.0336	Upstream UCK1	no	10kb
H3K27me3 Broad	***2***	***chr2***	***89610914***	***89613026***	2.71	0.0265	Downstream NR_136329	no	15kb
H3K27me3 Broad	2	chrY	11332961	11334300	1.91	1.31E-05	Downstream GYG2P1	no	1mb
H3K27me3 Broad	2	chr21	12965934	12966712	-2.25	0.0356	Upstream ANKRD30BP2	no	1mb
H3K27me3 Broad	2	chr2	91477391	91479074	3.02	0.00557	downstream LOC654342	no	200kb
H3K27me3 Broad	***2***	***chr2***	***87423893***	***87425699***	2.81	0.000791	Upstream CYTOR	no	30kb
H3K27me3 Broad	***2***	***chr2***	***87418396***	***87419860***	3.45	0.0146	Upstream CYTOR	no	33kb
H3K27me3 Broad	2	chr20	26602470	26604119	-3.51	0.0366	centromere / Upstream MIR663AHG	no	400kb
H3K27me3 Broad	2	chr7	56370534	56373033	3.4	0.0326	Downstream LOC650226	no	40kb
H3K27me3 Broad	***2***	***chr8***	***144125644***	***144126008***	1.94	0.0487	Upstream HGH1	no	6kb
H3K27me3 Broad	2	chr20	29326303	29327076	2.63	0.0156	centromere	No	
H3K27me3 narrow	2	chr14_GL000225v1_random	59742	59886	3.84	0.0318			

5.7 H3K4me1:

The number of narrow peaks identified following the pull downs with H3K4me1 ranged from 1 313 to 3 169 across the 6 samples.

The differential binding analysis for the narrow peaks following H3K4me1 pull down with the outlier excluded resulted in 75 significantly different regions (Table 5.4).

The analysis of the broad peaks resulted in 70 peaks (Appendix table 5.3) of which the only difference of note was the presence of enrichment in the exon / transcription start site of long non-protein coding RNA 1667 (LINC01667).

The results also identified enriched regions common to both analyses, including the second exon for mucin 3A (MUC3A) gene. MUC3A encodes a transmembrane mucin (Williams et al., 1999). It has been suggested that mucin in macrophages may play an anti-inflammatory role (Kato et al., 2016). Another is within the Rho associated coiled-coil containing protein kinase 1 pseudogene 1 (ROCK1P1) pseudogene body. A further region was within one of the introns for Double homeobox 4 (DUX4), which was also significantly different between the two conditions. Deficiencies in DUX4 are associated with Facioscapulohumeral muscular dystrophy (Dmitriev et al., 2013). Finally, the region coding for NR_038958, which is of unknown function, was also differentially enriched between the two conditions.

When comparing the differential analysis with the outlier included, the only significant addition was an enriched region of chromosome 20 situated 30kb upstream of the defensin beta 125 (DEFB125) gene, which can act as antimicrobial peptides.

Table 5.4: Differentially enriched regions following H3K4me1 diffbind analysis.

Chromosome	Start	End	Fold	FDR	Gene		distance
chr18	**110245**	**110763**	**7.43**	**2.29E-05**	**ROCK1P1 /MIR8078**	**exon**	
chr18	**109729**	**110029**	**5.89**	**0.00783**	**ROCK1P1 /MIR8078**	**exon**	
chr18	**107882**	**109395**	**2.24**	**0.0111**	**ROCK1P1 /MIR8078**	**exon**	
chr7	**100957785**	**100958930**	**6.1**	**0.00298**	**MUC3A**	**exon**	
chr21	**8219768**	**8220288**	**5.7**	**0.014**	**NR_038958**	**intron**	
chr4	**190177603**	**190178966**	**3.98**	**0.035**	**DUX4**	**intron**	
chr16	46389641	46391270	9.56	1.88E-11	Downstream ANKRD26P1	no	100kb
chr16	46388615	46389363	7.91	3.74E-06	Downstream ANKRD26P1	no	100kb
chr16	46394275	46395099	8.2	0.000344	Downstream ANKRD26P1	no	100kb
chr21	10692370	10693086	4.39	0.048	Downstream TPTE	no	100kb
chr3	***75668894***	***75669629***	***7.82***	***7.13E-06***	***Downstream FRG2C / Upstream LINC00960***	no	10kb / 12kb
chr8	***144125633***	***144126054***	***4.89***	***0.0398***	***Downstream WDR97 / Upstream HGH1***	no	10kb / 12kb
chr20	28897867	28898808	4.72	0.0466	Upstream FRG1DP	no	180kb
chrY	11332971	11334272	7.76	7.13E-06	Downstream GYG2P1	no	1mb
chrY	11329704	11330373	5.7	0.00783	Downstream GYG2P1	no	1mb
chrY	11322566	11322761	4.77	0.0243	Downstream GYG2P1	no	1mb
chrY	11318613	11319054	5.12	0.0267	Downstream GYG2P1	no	1mb
chrY	10775971	10776671	-4.63	0.0492	Downstream TTTY23	no	1mb
chr1	143214220	143215408	6.54	0.00182	Downstream LOC645166	no	200kb
chr1	143184638	143185470	6.19	0.0023	Downstream LOC645166	no	200kb
chr1	143263376	143264803	7.12	0.0033	Downstream LOC645166	no	200kb
chr10	42070441	42071049	-4.16	0.0318	Downstream LOC441666	no	200kb
chr10	41903968	41904399	4.53	0.0413	Downstream LOC441666	no	200kb
chr10	41896173	41896560	4.7	0.0467	Downstream LOC441666	no	200kb
chr10	41879280	41879654	3.58	0.0492	Downstream LOC441666	no	200kb
chr16	34582695	34583823	8.82	9.56E-10	Upstream LINC00273	no	200kb
chr16	34594857	34595785	6.89	0.00125	Upstream LINC00273	no	200kb
chr16	34586491	34587021	4.59	0.00222	Upstream LINC00273	no	200kb
chr16	34571640	34572115	6.24	0.00278	Upstream LINC00273	no	200kb
chr16	34592580	34593661	6.38	0.014	Upstream LINC00273	no	200kb
chr16	34573906	34574451	6.05	0.0222	Upstream LINC00273	no	200kb
chr16	34582134	34582473	5.79	0.0313	Upstream LINC00273	no	200kb
chr16	46380706	46381059	5.62	0.0396	Upstream LINC00273	no	200kb
chr2	89836279	89836822	6.95	0.000408	Upstream NR_136329	no	200kb
chrY	56734619	56734919	5.93	0.00387	Upstream SPRY3	no	250kb
chr10	41883097	41884055	4.75	0.00275	Downstream LOC441666	no	300kb
chr4	***49091787***	***49092425***	***6.43***	***0.00126***	***Downstream CWH43***	no	30kb
chr4	***49110648***	***49110846***	***3.58***	***0.0493***	***Downstream CWH43***	no	30kb
chr22	12693437	12694023	6.01	0.00325	Downstream NR_110761	no	600kb

chr20	***31185032***	***31187393***	***4.19***	***0.0225***	***Upstream DEFB115***	no	70kb
chr20	***47894718***	***47895496***	***5.37***	***0.0484***	***Upstream LINC01522***	no	80kb
chr1_KI270709v1_random	6596	7783	5.42	0.048			
chr14_GL000225v1_random	131547	131970	7.16	0.000532			
chr14_GL000225v1_random	6675	6980	6.86	0.000542			
chr14_GL000225v1_random	81845	82339	6.17	0.0023			
chr14_GL000225v1_random	125653	126052	5.81	0.0103			
chr14_GL000225v1_random	25423	25829	5.46	0.0188			
chr14_GL000225v1_random	45947	46206	5.28	0.0279			
chr14_GL000225v1_random	90884	91265	5.01	0.0313			
chr14_GL000225v1_random	40435	41042	4.93	0.0396			
chr14_GL000225v1_random	85565	86015	3.33	0.0413			
chr14_GL000225v1_random	17007	17269	4.86	0.0413			
chr14_KI270723v1_random	35218	35437	5.03	0.0313			
chr15	17080543	17081024	5.45	0.014			
chr17_KI270729v1_random	2325	3648	4.07	1.30E-05			
chr17_KI270729v1_random	21051	21569	5.49	0.0162			
chr22_KI270733v1_random	165118	165364	5.22	0.0222			
chr22_KI270736v1_random	179479	181786	4.07	0.00126			
chrUn_GL000214v1	125307	125518	4.68	0.0476			
chrUn_GL000216v2	149357	149842	6.63	0.00125			
chrUn_GL000216v2	160613	161004	6.52	0.00126			
chrUn_GL000216v2	158635	159714	4.27	0.0275			
chrUn_GL000220v1	144801	145571	-5.53	0.0132			
chrUn_GL000224v1	3509	4577	6.88	2.78E-05			
chrUn_GL000224v1	944	1901	6.1	0.00278			
chrUn_KI270438v1	103888	105259	4.21	1.65E-12			
chrUn_KI270438v1	111252	112440	2.25	0.000511			
chrUn_KI270438v1	109174	110591	2.04	0.0267			
chrUn_KI270465v1	802	1397	4.75	0.048			
chrUn_KI270519v1	137592	138002	4.13	0.0497			
chrUn_KI270589v1	41729	42716	5.27	0.0218			
chrUn_KI270744v1	119664	121040	7.05	0.0029			
chrUn_KI270744v1	109241	109721	4.45	0.0417			
chrUn_KI270751v1	149528	150682	5.87	0.0192			
chrUn_KI270754v1	18129	18342	5.68	0.00974			

5.8 H3K4me3:

This identified over 19 000 peaks in each sample. However, the diffbind analysis did not discern any significantly differentially enriched regions directly relating to nearby protein coding regions (Appendix table 5.4).

5.9 Integration of ChIP-Seq and RNA-Seq data:

In order to further investigate the relationship between histone PTMs and host transcriptional response, the results of the ChIP-Seq were used to assess the coding regions for the significantly differentially expressed transcripts identified in Table 5.1.

Table 5.5: Analysis of histone motif surrounding differentially expressed genes.

Gene name	H3K27ac	H3K4me3	H3K4me1	H3K27me3	H3K9me2
actin gamma 1 pseudogene 20	TSS	TSS			
C-C motif chemokine ligand 3 like 1 variant 1			30kb down-stream		
C-C motif chemokine ligand 3 like 1, variant 2	TSS	TSS			
C-C motif chemokine ligand 4 like 1 variant CCL4Ldelta2					
C-X-C motif chemokine ligand 10					
early growth response 3, transcript variant 2		TSS		*Spn* only	
growth arrest and DNA damage inducible beta	TSS	TSS			
Nuclear growth arrest and DNA damage inducible beta (GADD45B), transcript variant X1	TSS	TSS			
Interferon gamma					
MIR155 host gene	TSS	TSS			
nuclear receptor subfamily 4 group A member 1,	TSS	TSS			
nuclear receptor subfamily 4 group A member 2	TSS	TSS		*Spn* only	
Pentraxin 3		TSS			
U2 small nuclear 1	TSS	TSS			
U6 small nuclear					
Tumour necrosis factor					
Coagulation factor III, tissue factor (F3)	TSS	TSS			

Having found 17 genes whose expression was found to be significantly differentially expressed and were all up-regulated, I examined these genomic locations in the ChIP-Seq data sets (Table 5.5). This revealed that for the most part (8 out of 17) the transcription start sties (TSS) were predominantly associated with H3K27ac. Furthermore, H3K4me3 was found in the TSS of 11 out of the 17 genes. As both of these marks are associated with active gene transcription and that these genes were up-regulated this was to be expected. Unfortunately, there was no discernible difference between the mock infected samples and those challenged with *S. pneumoniae* (as illustrated in Figure 5.13). In the case of EGR3 and NR4A2 there was also association of the genomic locations with H3K27me3. As this mark is predominantly associated with gene repression this may represent a negative feedback mechanism.

Figure 5.13: IGV view of the NR4A2 gene and the ChIP-Seq profiles for H3K27ac and H3K4me3 and H3K27me3.

This figure illustrates the peaks identified following the ChIP-Seq for H3K27ac, H3K27me3 and H3K4me3 around the NR4A2 gene. The majority of the peaks are identified within 1kb of the start of the coding sequence for this gene.

5.10 Conclusions:

The aim of this chapter was to characterise the changes occurring in MDMs following challenge with *S. pneumoniae* at an epigenomic (histone PTMs) and transcriptomic level and attempt to integrate the two data sets by using ChIP-Seq in order to infer a functional role for the changes observed. This chapter has further highlighted the ability of MS to assess global changes in histone PTMs. Three hours following the challenge of MDMs with *S. pneumoniae* there were a number of changes in the relative abundance of histone PTMs. These changes were similar to those seen following infective challenge in MDMs with *S. pneumoniae* or ΔPLY (in Chapter 4), namely the increase in the relative abundance of H3K4me1, H3K23ac, H3K27me3 and H4K20me1, and a decrease in H3K9me2, H3K27me2 and H4K20me2. As a number of these marks have been associated with the regulation of gene transcription, it is likely that they play a critical role in the host's response to infection.

The RNA-Seq analysis showed that at this very early time point there are 47 transcripts whose expression is significantly modified. Importantly, these include several genes involved in the innate immune response including TNF, INF-γ, CCL3L1, CCL4L1, CXCL10, and so on. This indicates that the although there are very few transcripts whose expression is significantly modified after challenge with *S. pneumoniae*, the vast majority of those that are, are involved in the immune response. Furthermore, I demonstrated significant changes in a number of histone PTMs at the global level after challenge with *S. pneumoniae.* These epigenetic modifications may potentially play a role in the regulation of the transcriptional response.

Finally, the ChIP-Seq experiments identified several thousand peaks in the majority of the pull downs across the genome. The data was overall good quality and it highlighted a number of different genomic regions that were differentially enriched following challenge with *S. pneumoniae.* This confirms that changes not only occur at a global level, as seen by MS, but also impact at the level of individual genes. Interestingly, the H3K27ac pull downs identified an increase in the abundance of H3K27ac upstream of the RIPK2. As this has been associated with the innate immune response this maybe illustrative of epigenetic control of the innate immune response. Unfortunately, the transcript for RIPK2 was not identified in the paired RNA-Seq, limiting the potential to draw any further conclusions.

One of the limitations of this study has been the low number of differentially enriched genomic locations in the ChIP-Seq with the H3K9me2 broad mark. This may be due to technical limitations, as it has been demonstrated that detecting global decreases in marks can be difficult without the use of spike-in standards to normalise against (Orlando et al., 2014). Unfortunately, despite identifying a number of interesting genes related to immune responses, they were not associated with differentially bound histone PTMs in the ChIP-Seq analysis. Finally, it is possible that the early time point has restricted both the number of differentially expressed transcripts and the differentially enriched genomic locations following the ChIP-Seq and repeating the bacterial challenge with a longer duration may prove more fruitful.

Chapter 6: Conclusions and future work

6.1 Conclusions:

S. pneumoniae is the principle cause of community acquired pneumonia and as such poses a significant public health burden (Ladhani et al., 2013). There are up to six thousand cases of invasive pneumococcal disease (IPD) in the UK per year ("GOV.UK,"2018.).As colonisation of the nasopharynx has been shown to precede invasive pneumococcal disease, it is therefore the interaction between the host's innate immune system and the pathogen that will dictate whether IPD occurs. Once bacteria translocate to the lower airway one of the key components of the innate immune response is the alveolar macrophage, while once bacteria spread to the blood splenic macrophages are critical for clearance and therefore it is this interaction between the macrophage and *S. pneumoniae* that I chose to study (Dockrell, 2012). Histone PTMs have been shown to be modulated by pathogens to hijack the host's epigenetic machinery to its own advantage. Thus, in this thesis I set out to study the changes in global histone PTM profiles following challenge with *S. pneumoniae*

Previous studies have suggested that many histone PTM changes occur very rapidly upon exposure to other bacteria (Hamon et al., 2007). Therefore, I sought the earliest time point at which host response was modulated in an attempt to characterise the first changes seen. Following, analysis of cytokine release I established that in my model TNF and IL-6 were released and were detectable following a 3 hour challenge with *S. pneumoniae.* Although, analysis of this time point allowed the identification of several histone PTM changes, the transcriptomic and ChIP-Seq analysis may have been limited by the dynamic range of these early changes.

Initial work focused on the optimisation of mass spectrometry based methods for identification and quantification of histone PTMs. Reductions in the amount of MS and data analysis time was achieved by adapting data independent methods. Using the optimised data independent acquisition methods, I was able to identify and quantify 83 histone proteoforms across histones H3 and H4, and demonstrate significant changes in the abundance of 10 different histone PTMs between MDMs mock infected and those challenged with *S. pneumoniae*. I focussed on lysine methylation and acetylation on histones H3 and H4 as these are the most abundant modifications and have been implicated in the pathogenesis of many disease processes, including infection. Indeed, several pathogens have been found to hijack the host's epigenome to promote its survival. One of the limitations of this approach is that it was restricted to acetylation and methylation modifications and did not examine other post-translational. However, as the data was acquired in a data independent manner it is possible to re-analyse the current data, targeting alternative histone PTMs if required.

Many studies *in vivo* have demonstrated the importance of pneumolysin as a virulence factor (Benton et al., 1995). Initially, in Chapter 4, I focussed on the study of pneumolysin, a key virulence factor for *S. pneumoniae* (Kadioglu et al., 2008b). First, I aimed to establish the importance of pneumolysin on the host's response by studying transcriptomic changes by microarray. Further studies were subsequently aimed at identifying potential proteins that play a critical role in the host's response to pneumolysin. High throughput label-free quantitative proteomics in conjunction with microarray analysis was performed to identify differentially expressed proteins and genes following infection with *S. pneumoniae* or a pneumolysin deficient mutant. Finally, I established the changes in abundance of histone PTMs in a pneumolysin

dependent manner. To conclude I attempted to integrate these findings to demonstrate that there are specific changes in the host's cell response to infection in a pneumolysin dependent manner.

The transcriptomic analysis revealed 503 pneumolysin dependent differentially expressed probes. Further pathway analysis demonstrated that several of these were involved in inflammatory responses. In particular, oxidative stress responses were highlighted as being up regulated during infection.

Quantitative proteomic analysis identified differentially expressed proteins at 3 and 6 hours in a pneumolysin dependent manner following challenge with *S. pneumoniae*. Of note, one of these differentially abundant proteins includes oxidative stress response protein 1, suggesting that stress responses play a key role in the host pathogen interaction at both the transcriptome and proteome level.

The study of pneumolysin dependent changes in relative abundance of histone PTMs revealed 5 statistically significant histone PTMs changes on histone H3. This suggests that the host's response to challenge with *S. pneumoniae* includes modification of global abundance of histone PTMs, and extends on this by linking the changes to a specific virulence factor, supporting the hypothesis that pneumolysin may be responsible for alterations in gene transcription by modulating the epigenome. However, despite establishing pneumolysin specific changes in abundance of histone PTMs, the exact role for these remains unclear. Further work using alternative approaches including ChIP-Seq would potentially provide further insight proved too costly to undertake for both the mutant and the parent strain for all of the modifications. Moreover, a number interesting changes occurred in combinatorial modifications such as H3K27me2K36me2, highlighting the advantages of MS over antibody based approaches to study histone PTMs; it will prove challenging to illicit the role for these combinatorial marks by ChIP-Seq as this is likely to require sequential pull downs and therefore require large amounts of starting chromatin to perform ChIP-ChIP-Seq.

To provide further insight into the effects of *S. pneumoniae* on the epigenome an integrated multi-omic approach was used. Initial work focussed on analysing changes in global histone PTMs using mass spectrometry. The results identified a number of histone PTMs that are altered in abundance following the challenge of MDMs with *S. pneumoniae.* These include increases in H3K4me1, H3K23ac, H3K27me3 and H4K20me1, and a decrease in H3K9me2, H3K27me2 and H4K20me2, similar to the effects seen with pneumolysin. Consistent with previous observations in response to other pathogens (Hamon et al., 2007; Rolando and Buchrieser, 2014), these results demonstrate alterations at the epigenetic level following infection with *S. pneumonia,* highlighting potential important novel histone PTMs associated with this host-pathogen interaction.

Next, I sought to ascertain the role played by the changes seen in histone PTMs abundance following challenge with *S. pneumonia*. In order to achieve this, I used RNA-Seq to establish the transcriptomic changes occurring in response to infection. The RNA-Seq analysis identified several genes involved in the inflammatory response namely, TNF, INF-γ, CCL3L1, CCL4L1, CXCL10 that are differentially expressed and all of which were up-regulated, at this early time point, 3 hours following challenge with *S. pneumoniae.* Reassuringly, a number of these differentially expressed transcripts were also identified in the microarray analysis in Chapter 4 including TNF and NR4A2. These results are consistent with previous studies that have highlighted the role of TNF in the host response to *S. pneumoniae* infections (Jeong et al., 2015; Kirby et al., 2005).

Having demonstrated that upon challenge with *S. pneumoniae* alterations at both the epigenetic and gene level were observed, further analysis was performed using

ChIP-Seq in an effort to link alterations observed at the epigenetic level (histone PTMs) with perturbations observed at the transcriptional level. The analysis of the H3K9me2 ChIP-Seq, a mark associated with gene repression and that of the H3K4me3 ChIP-Seq, a mark associated with active gene transcription when located in the promoter regions did not demonstrate any differentially enriched locations. For the other marks examined, H3K4me1, H3K27me3 and H3K27ac, the ChIP-Seq analysis revealed a number of differentially enriched regions suggesting that these are indeed modulated during the host pathogen interaction.

H3K27ac has been associated with active gene transcription when located in the transcription start site. H3K4me1 has been associated with active or poised enhancer and promoter regions of genes and is associated with active gene transcription. The initial analysis revealed that H3K27ac was found in close proximity to transcription start sites. There were 5 differentially enriched genomic locations following pull down with H3K27ac, two of which involved the gene body of STAB1 and RIPK2. Both of which have been shown to play a role in the inflammatory response. The analysis of the H3K4me1 ChIP-Seq demonstrated 70 differentially enriched genomic locations. This included the MUC3A gene which has been associated with anti-inflammatory properties in macrophages. The analysis of H3K27me3, a mark associated with gene repression, revealed 18 differentially enriched genomic locations. This included the coding region for SLIT2 which has been associated with inflammatory processes.

This confirms that changes not only occur at a global level, as seen by MS, but also impact at the level of individual genes. Interestingly, the H3K27ac pull downs identified an increase in the abundance of H3K27ac upstream of the RIPK2. As this has been associated with the innate immune response this potentially important finding is illustrative of epigenetic control of the innate immune response. Unfortunately, the transcript for RIPK2 was not identified in the paired RNA-seq.

One possible explanation for the paucity of findings following ChIP-Seq is that the global effect demonstrated by MS is not distinguishable by this method due to the relative small changes in abundance occurring across large sections of the genome. Newer, approaches relying on spike-in methods to normalise the signal from each ChIP may be more sensitive to these global changes and represent a useful approach. Indeed, the ability to distinguish changes in the relative abundance of histone PTMs for broad marks has been demonstrated using Drosophilia chromatin spike in (Orlando et al., 2014). Furthermore, the early time point studied (3 hours after exposure to bacteria) may also be contributing to the limited number of statistically significant findings.

6.2 Future work:

In order to further characterise the role played by pneumolysin in the pathogenesis of pneumococcal infections a number of different approaches could be used. It would be of interest to use exogenous pneumolysin to establish if the same changes in histone PTMs occur as are seen following infection with *S. pneumoniae*. Furthermore, different pneumolysin mutants exist that either lack pore forming ability or have different degrees of haemolytic activity. These mutants have been implicated in the pathogenesis of pneumococcal diseases and therefore could be used to establish which aspect of pneumolysin's function is essential for the formation of these histone PTMs. Moreover, having established that *S. pneumoniae* causes changes in histone post translational modifications, future work could be aimed at studying the different serotypes of *S. pneumoniae.* These studies would provide important insight and determine if changes in histone PTMs are specific to capsular serotype, or just governed

by the activity of pneumolysin, or related to other widely expressed virulence factors with more variable expression such as pili (Kadioglu et al., 2008b; Mitchell and Mitchell, 2010). In order to further establish the role of the pneumolysin dependent changes in the abundance of histone PTMs, further work should be performed using ChIP-qPCR to establish the changes in histone PTMs in the promoter regions of key genes involved in the response to infection (such as TNF).

The work presented in this thesis has been focused on the use of monocyte derived macrophages as a model for tissue resident macrophages. However, it is well established that alveolar macrophages arise from two distinct mechanisms. Firstly embryonic derived from a foetal liver linage and secondly from circulating peripheral monocytes (Gordon and Taylor, 2005). In order to further validate the importance of the role played by modulation of histone PTMs in the host's response to infection, future work could include a murine challenge. Such studies could examine the global changes in histone PTMs from alveolar macrophages following intra-nasal challenge. This model mimics pneumonia more closely than *in vitro* models and therefore is a closer representation of the clinical pathology. It will allow the confirmation that the epigenetic changes are indeed observed in tissue resident macrophages as well as other distal sites such as bone marrow or spleen. Moreover, it has been suggested that epigenetic changes are potentially inheritable and therefore future work proposed would also include breeding these mice to examine subsequent generations' histone PTM profile and carry out a phenotypic characterisation of the animals to look for susceptibility or resistance to pneumococcal challenge in the offspring. Furthermore, if the changes in the relative abundance of histone PTMs is also observed in the murine alveolar macrophages it would be important to establish a functional consequence of these changes by measuring cytokine responses, bacterial clearance and attempt to modify these modifications by using epigenetic modulators such as HDACi.

Initial work performed in Chapters 4 and 5 isolated all the MDMs following *S. pneumoniae* infection. However, a caveat associated with this approach is that only a subset of MDMs will have ingested the *S. pneumoniae*. This could have hampered the ability to identify differentially expressed genes and proteins and downstream ChIP analysis. Therefore, to overcome this caveat, it is proposed that future work could focus on using a fluorescent tagged bacteria and cell sorting the MDMs following challenge would guarantee that the cell had ingested at least one pathogen. This may reveal a greater number of differentially expressed proteins, transcripts and histone PTMs as the "noise" from the uninfected cells would be abrogated. However, the number of cells required would prove challenging to perform in primary monocyte derived macrophages and it may require a cell line such as THP-1 cells to obtain the necessary cells following infection.

Although changes in histones PTMs have been shown to occur very rapidly after bacterial challenge (Hamon et al., 2007), and I was able to demonstrate changes in the relative abundance of histone PTMs in MDMs and cytokine release following bacterial challenge at 3 hours, it is possible by studying a later time point that more pronounced changes might be apparent at the transcriptional and proteomic level, and possibly lead to better identification of differentially enriched regions by ChIP-Seq.

As alluded to in Chapter 1, histone PTMs are one of several different epigenetic mechanisms. It would therefore be interesting to ascertain the level of changes occurring in the DNA methylome, microRNA and histone PTMs and try to establish any of the interactions between the three as of these DNA methylation may represent a mechanism for the maintenance of longer term epigenetic "memory" as seen in trained immunity.

Although there is conclusive evidence that the relative abundance of histone PTMs is modified during the bacterial challenge it remains unclear whether this is a pathogen specific phenomenon or a generic host response to stress. It would therefore be of value to repeat the profiling of histone PTMs looking at the interaction with other pathogens such as using *Staphylococcus aureus* another important Gram positive pathogen, Gram negative bacteria and stress induced responses to other microbial components such as LPS, lipopeptides and peptidoglycan challenge.

Finally, in an attempt to further establish a role for these different histone PTMs one approach further work should focus on, is targeting the enzymes responsible for the histone PTMs that were perturbed during the challenge with *S. pneumoniae.* Two possible methods would be either the use of epigenetic inhibitors targeting key enzymes such as the G9a/GLP methyltransferase responsible for the methylation of H3K9me2, using BIX01294. Alternatively inhibiting the demethylases responsible for the decrease in H3K9me2 by using daminozide (a selective inhibitor of lysine demethylase 2/7), IOX-1 (histone demethylase JMJD inhibitor) or GSK Lysine Specific demethylase 1 (LSD1) inhibitor. Similar approaches could be used for the other histone PTMs. The methyltransferases responsible for H3K4me1 are Set9, MLL3 and MLL4. However, specific methyltransferase inhibitors for these enzymes are not very well developed, demethylation of H3K4me1 is mediated by LSD1 and can be inhibited by GSK LSD1 inhibitor. H3K23ac could be modulated by HDAC inhibitors however, one of the principle drawbacks of inhibitor based approaches is that they rarely are very specific and will therefore have a number of different off-target effects. An alternative method would be to use a CRISPR-Cas9 approach to either inhibit the enzymes responsible for the methylation / acetylation by cleaving the genes encoding them prior to infection, or using a CRISPR-Cas9 null construct coupled to an acetyltransferase (Hilton et al., 2015) or a methyltransferase (Xiong et al., 2017) to specifically acetylate the neighbouring histones or methylate the neighbouring DNA of a target gene and subsequently measure the response to infection.

6.3 Concluding remarks:

In conclusion, this thesis describes the development and optimisation of a mass spectrometry based method for the study of histone PTMs. The mass spectrometry approach is a powerful approach for the identification and quantification of histone PTMs, demonstrating a number of advantages over antibody based approaches. I utilised the developed methods to analyse epigenetic changes in monocyte derived macrophages in response to challenge with *S. pneumoniae.* In addition, high throughput transcriptomic and proteomic methods were also used to study changes in transcriptional response and protein expression at an early time to challenge with *S. pneumoniae*. The results identified a number of histone PTMs were indeed modulated as part of the response to pneumococcal challenge, supporting the hypothesis that epigenetics (histone PTMs) play a role in the regulation of the innate immune response. Furthermore, the transcriptomic and ChIP-Seq analyses both showed a number of innate immune response genes to be differentially enriched following infection confirming that the epigenome is modified in response to infective challenge with *S. pneumoniae* and is associated with alterations in the host's transcriptome and proteome. Together, these results, demonstrate the role of epigenetic regulation of gene expression associated with the innate immune response upon challenge with *S. pneumoniae.*

Chapter 7: References

Ariyoshi, M., Katane, M., Hamase, K., Miyoshi, Y., Nakane, M., Hoshino, A., Okawa, Y., Mita, Y., Kaimoto, S., Uchihashi, M., Fukai, K., Ono, K., Tateishi, S., Hato, D., Yamanaka, R., Honda, S., Fushimura, Y., Iwai-Kanai, E., Ishihara, N., Mita, M., Homma, H., Matoba, S., 2017. D-Glutamate is metabolized in the heart mitochondria. Sci. Rep. 7, 43911. https://doi.org/10.1038/srep43911

Ballaré, C., Lange, M., Lapinaite, A., Martin, G.M., Morey, L., Pascual, G., Liefke, R., Simon, B., Shi, Y., Gozani, O., Carlomagno, T., Benitah, S.A., Croce, L.D., 2012. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nat. Struct. Mol. Biol. 19, nsmb.2434. https://doi.org/10.1038/nsmb.2434

Barrand, S., Collas, P., 2010. Chromatin states of core pluripotency-associated genes in pluripotent, multipotent and differentiated cells. Biochem. Biophys. Res. Commun. 391, 762–767. https://doi.org/10.1016/j.bbrc.2009.11.134

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., Zhao, K., 2007. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129, 823–837. https://doi.org/10.1016/j.cell.2007.05.009

Baxt, L.A., Garza-Mayers, A.C., Goldberg, M.B., 2013. Bacterial Subversion of Host Innate Immune Pathways. Science 340, 697–701. https://doi.org/10.1126/science.1235771

Benoit, M., Desnues, B., Mege, J.-L., 2008. Macrophage Polarization in Bacterial Infections. J. Immunol. 181, 3733–3739. https://doi.org/10.4049/jimmunol.181.6.3733

Benton, K.A., Everson, M.P., Briles, D.E., 1995. A pneumolysin-negative mutant of Streptococcus pneumoniae causes chronic bacteremia rather than acute sepsis in mice. Infect. Immun. 63, 448–455.

Berger, S.L., 2007. The complex language of chromatin regulation during transcription. Nature 447, 407–412. https://doi.org/10.1038/nature05915

Berger, S.L., Kouzarides, T., Shiekhattar, R., Shilatifard, A., 2009. An operational definition of epigenetics. Genes Dev. 23, 781–783. https://doi.org/10.1101/gad.1787609

Bhavsar, A.P., Guttman, J.A., Finlay, B.B., 2007. Manipulation of host-cell pathways by bacterial pathogens. Nature 449, 827–834. https://doi.org/10.1038/nature06247

Bistoni, F., Vecchiarelli, A., Cenci, E., Puccetti, P., Marconi, P., Cassone, A., 1986. Evidence for macrophage-mediated protection against lethal Candida albicans infection. Infect. Immun. 51, 668–674.

Bluteau, D., Balduini, A., Balayn, N., Currao, M., Nurden, P., Deswarte, C., Leverger, G., Noris, P., Perrotta, S., Solary, E., Vainchenker, W., Debili, N., Favier, R., Raslova, H., 2014. Thrombocytopenia-associated mutations in the *ANKRD26* regulatory region induce MAPK hyperactivation. J. Clin. Invest. 124, 580–591. https://doi.org/10.1172/JCI71861

Bodai, L., Zsindely, N., Gáspár, R., Kristó, I., Komonyi, O., Boros, I.M., 2012. Ecdysone Induced Gene Expression Is Associated with Acetylation of

Histone H3 Lysine 23 in Drosophila melanogaster. PLOS ONE 7, e40565. https://doi.org/10.1371/journal.pone.0040565
Bogaert, D., de Groot, R., Hermans, P., 2004. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. Lancet Infect. Dis. 4, 144–154. https://doi.org/10.1016/S1473-3099(04)00938-7
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma. Oxf. Engl. 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170
Bootsma, H.J., Egmont-Petersen, M., Hermans, P.W.M., 2007. Analysis of the In Vitro Transcriptional Response of Human Pharyngeal Epithelial Cells to Adherent Streptococcus pneumoniae: Evidence for a Distinct Response to Encapsulated Strains. Infect. Immun. 75, 5489–5499. https://doi.org/10.1128/IAI.01823-06
Bottazzi, B., Doni, A., Garlanda, C., Mantovani, A., 2010. An Integrated View of Humoral Innate Immunity: Pentraxins as a Paradigm. Annu. Rev. Immunol. 28, 157–183. https://doi.org/10.1146/annurev-immunol-030409-101305
Bourmaud, A., Gallien, S., Domon, B., 2016. Parallel reaction monitoring using quadrupole-orbitrap mass spectrometer: Principle and applications. PROTEOMICS n/a-n/a. https://doi.org/10.1002/pmic.201500543
Bouttier, M., Laperriere, D., Memari, B., Mangiapane, J., Fiore, A., Mitchell, E., Verway, M., Behr, M.A., Sladek, R., Barreiro, L.B., Mader, S., White, J.H., 2016. Alu repeats as transcriptional regulatory platforms in macrophage responses to M. tuberculosis infection. Nucleic Acids Res. 44, 10571–10587. https://doi.org/10.1093/nar/gkw782
Braun, J.S., Novak, R., Gao, G., Murray, P.J., Shenep, J.L., 1999. Pneumolysin, a Protein Toxin of Streptococcus pneumoniae, Induces Nitric Oxide Production from Macrophages. Infect. Immun. 67, 3750–3756.
Braun, L., Ghebrehiwet, B., Cossart, P., 2000. gC1q-R/p32, a C1q-binding protein, is a receptor for the InlB invasion protein of Listeria monocytogenes. EMBO J. 19, 1458–1466. https://doi.org/10.1093/emboj/19.7.1458
Brouwer, M.C., de Gans, J., Heckenberg, S.G., Zwinderman, A.H., van der Poll, T., van de Beek, D., 2009. Host genetic susceptibility to pneumococcal and meningococcal disease: a systematic review and meta-analysis. Lancet Infect. Dis. 9, 31–44. https://doi.org/10.1016/S1473-3099(08)70261-5
Bryant, K.A., Block, S.L., Baker, S.A., Gruber, W.C., Scott, D.A., PCV13 Infant Study Group, 2010. Safety and immunogenicity of a 13-valent pneumococcal conjugate vaccine. Pediatrics 125, 866–875. https://doi.org/10.1542/peds.2009-1405
Burchfield, J.S., Li, Q., Wang, H.Y., Wang, R.-F., 2015. JMJD3 as an epigenetic regulator in development and disease. Int. J. Biochem. Cell Biol. 67, 148–157. https://doi.org/10.1016/j.biocel.2015.07.006
Carroll, T.S., Liang, Z., Salama, R., Stark, R., de Santiago, I., 2014. Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. Front. Genet. 5. https://doi.org/10.3389/fgene.2014.00075
Cattaneo, P., Kunderfranco, P., Greco, C., Guffanti, A., Stirparo, G.G., Rusconi, F., Rizzi, R., Di Pasquale, E., Locatelli, S.L., Latronico, M.V.G., Bearzi, C., Papait, R., Condorelli, G., 2016. DOT1L-mediated H3K79me2

modification critically regulates gene expression during cardiomyocyte differentiation. Cell Death Differ. 23, 555–564. https://doi.org/10.1038/cdd.2014.199
Chapman, S.J., Hill, A.V.S., 2012. Human genetic susceptibility to infectious disease. Nat. Rev. Genet. 13, 175–188. https://doi.org/10.1038/nrg3114
Chen, G., Gharib, T.G., Huang, C.-C., Taylor, J.M.G., Misek, D.E., Kardia, S.L.R., Giordano, T.J., Iannettoni, M.D., Orringer, M.B., Hanash, S.M., Beer, D.G., 2002. Discordant Protein and mRNA Expression in Lung Adenocarcinomas. Mol. Cell. Proteomics 1, 304–313. https://doi.org/10.1074/mcp.M200008-MCP200
Chen, P., Zhao, J., Wang, Y., Wang, M., Long, H., Liang, D., Huang, L., Wen, Z., Li, W., Li, X., Feng, H., Zhao, H., Zhu, P., Li, M., Wang, Q., Li, G., 2013. H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. Genes Dev. 27, 2109–2124. https://doi.org/10.1101/gad.222174.113
Chen, S., Ma, J., Wu, F., Xiong, L., Ma, H., Xu, W., Lv, R., Li, X., Villen, J., Gygi, S.P., Liu, X.S., Shi, Y., 2012. The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. Genes Dev. 26, 1364–1375. https://doi.org/10.1101/gad.186056.111
Chen, W., Yazicioglu, M., Cobb, M.H., 2004. Characterization of OSR1, a Member of the Mammalian Ste20p/Germinal Center Kinase Subfamily. J. Biol. Chem. 279, 11129–11136. https://doi.org/10.1074/jbc.M313562200
Cheng, J., Blum, R., Bowman, C., Hu, D., Shilatifard, A., Shen, S., Dynlacht, B.D., 2014. A Role for H3K4 Monomethylation in Gene Repression and Partitioning of Chromatin Readers. Mol. Cell 53, 979–992. https://doi.org/10.1016/j.molcel.2014.02.032
Clark, I.A., Allison, A.C., Cox, F.E., 1976. Protection of mice against Babesia, and Plasmodium with BCG. Nature 259, 309–311. https://doi.org/10.1038/259309a0
Cole, J., Aberdein, J., Jubrail, J., Dockrell, D.H., 2014. The Role of Macrophages in the Innate Immune Response to Streptococcus pneumoniae and Staphylococcus aureus: Mechanisms and Contrasts. Adv. Microb. Physiol. 65, 125–202. https://doi.org/10.1016/bs.ampbs.2014.08.004
Cole, J., Morris, P., Dickman, M.J., Dockrell, D.H., 2016. The therapeutic potential of epigenetic manipulation during infectious diseases. Pharmacol. Ther. 167, 85–99. https://doi.org/10.1016/j.pharmthera.2016.07.013
Collard, F., Collet, J.-F., Gerin, I., Veiga-da-Cunha, M., Van Schaftingen, E., 1999. Identification of the cDNA encoding human 6-phosphogluconolactonase, the enzyme catalyzing the second step of the pentose phosphate pathway. FEBS Lett. 459, 223–226. https://doi.org/10.1016/S0014-5793(99)01247-8
Conway, E., Healy, E., Bracken, A.P., 2015. PRC2 mediated H3K27 methylations in cellular identity and cancer. Curr. Opin. Cell Biol., Differentiation and disease 37, 42–48. https://doi.org/10.1016/j.ceb.2015.10.003
Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., Mann, M., 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics 13, 2513–2526.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., Boyer, L.A., Young, R.A., Jaenisch, R., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. 107, 21931–21936. https://doi.org/10.1073/pnas.1016071107

Crick, F.H., 1958. On protein synthesis. Symp. Soc. Exp. Biol. 12, 138–163.

Darwich, L., Coma, G., Peña, R., Bellido, R., Blanco, E.J.J., Este, J.A., Borras, F.E., Clotet, B., Ruiz, L., Rosell, A., Andreo, F., Parkhouse, R.M.E., Bofill, M., 2009. Secretion of interferon-γ by human macrophages demonstrated at the single-cell level after costimulation with interleukin (IL)-12 plus IL-18. Immunology 126, 386–393. https://doi.org/10.1111/j.1365-2567.2008.02905.x

De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., Burgold, T., Austenaa, L., Bucci, G., Caganova, M., Notarbartolo, S., Casola, S., Testa, G., Sung, W.-K., Wei, C.-L., Natoli, G., 2009a. Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. EMBO J. 28, 3341–3352. https://doi.org/10.1038/emboj.2009.271

De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., Burgold, T., Austenaa, L., Bucci, G., Caganova, M., Notarbartolo, S., Casola, S., Testa, G., Sung, W.-K., Wei, C.-L., Natoli, G., 2009b. Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. EMBO J. 28, 3341–3352. https://doi.org/10.1038/emboj.2009.271

Delves, P.J., Roitt, I.M., 2000. The Immune System. N. Engl. J. Med. 343, 37–49. https://doi.org/10.1056/NEJM200007063430107

DiMaggio, P.A., Young, N.L., Baliban, R.C., Garcia, B.A., Floudas, C.A., 2009. A Mixed Integer Linear Optimization Framework for the Identification and Quantification of Targeted Post-translational Modifications of Highly Modified Proteins Using Multiplexed Electron Transfer Dissociation Tandem Mass Spectrometry. Mol. Cell. Proteomics 8, 2527–2543. https://doi.org/10.1074/mcp.M900144-MCP200

Ding, S.-Z., Fischer, W., Kaparakis-Liaskos, M., Liechti, G., Merrell, D.S., Grant, P.A., Ferrero, R.L., Crowe, S.E., Haas, R., Hatakeyama, M., Goldberg, J.B., 2010. Helicobacter pylori-Induced Histone Modification, Associated Gene Expression in Gastric Epithelial Cells, and Its Implication in Pathogenesis. PLoS ONE 5, e9875. https://doi.org/10.1371/journal.pone.0009875

Dmitriev, P., Stankevicins, L., Ansseau, E., Petrov, A., Barat, A., Dessen, P., Robert, T., Turki, A., Lazar, V., Labourer, E., Belayew, A., Carnac, G., Laoudj-Chenivesse, D., Lipinski, M., Vassetzky, Y.S., 2013. Defective regulation of microRNA target genes in myoblasts from facioscapulohumeral dystrophy patients. J. Biol. Chem. 288, 34989–35002. https://doi.org/10.1074/jbc.M113.504522

Dockrell, D.H., 2012. Pneumococcal Pneumonia Mechanisms of Infection and Resolution. CHEST J. 142, 482. https://doi.org/10.1378/chest.12-0210

Dockrell, D.H., Lee, M., Lynch, D.H., Read, R.C., 2001. Immune-Mediated Phagocytosis and Killing of Streptococcus pneumoniae Are Associated with Direct and Bystander Macrophage Apoptosis. J. Infect. Dis. 184, 713–722. https://doi.org/10.1086/323084

Dorighi, K.M., Swigut, T., Henriques, T., Bhanu, N.V., Scruggs, B.S., Nady, N., Still, C.D., Garcia, B.A., Adelman, K., Wysocka, J., 2017. Mll3 and Mll4

Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. Mol. Cell 66, 568–576.e4. https://doi.org/10.1016/j.molcel.2017.04.018
Durrant, W.E., Dong, X., 2004. Systemic Acquired Resistance. Annu. Rev. Phytopathol. 42, 185–209. https://doi.org/10.1146/annurev.phyto.42.040803.140421
Ebata, K.T., Mesh, K., Liu, S., Bilenky, M., Fekete, A., Acker, M.G., Hirst, M., Garcia, B.A., Ramalho-Santos, M., 2017. Vitamin C induces specific demethylation of H3K9me2 in mouse embryonic stem cells via Kdm3a/b. Epigenetics Chromatin 10, 36. https://doi.org/10.1186/s13072-017-0143-3
Edfors, F., Danielsson, F., Hallström, B.M., Käll, L., Lundberg, E., Pontén, F., Forsström, B., Uhlén, M., 2016. Gene-specific correlation of RNA and protein levels in human cells and tissues. Mol. Syst. Biol. 12, 883. https://doi.org/10.15252/msb.20167144
Egelhofer, T. a, Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A. a, Cheung, M.-S., Day, D.S., Gadel, S., Gorchakov, A. a, Gu, T., Kharchenko, P.V., Kuan, S., Latorre, I., Linder-Basso, D., Luu, Y., Ngo, Q., Perry, M., Rechtsteiner, A., Riddle, N.C., Schwartz, Y.B., Shanower, G. a, Vielle, A., Ahringer, J., Elgin, S.C.R., Kuroda, M.I., Pirrotta, V., Ren, B., Strome, S., Park, P.J., Karpen, G.H., Hawkins, R.D., Lieb, J.D., 2011. An assessment of histone-modification antibody quality. Nat. Struct. Mol. Biol. 18, 91–3. https://doi.org/10.1038/nsmb.1972
Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., Gehrke, C., 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res. 10, 2709–2721.
Elton, T.S., Selemon, H., Elton, S.M., Parinandi, N.L., 2013. Regulation of the MIR155 host gene in physiological and pathological processes. Gene 532, 1–12. https://doi.org/10.1016/j.gene.2012.12.009
Eskandarian, H.A., Impens, F., Nahori, M.-A., Soubigou, G., Coppée, J.-Y., Cossart, P., Hamon, M.A., 2013. A role for SIRT2-dependent histone H3K18 deacetylation in bacterial infection. Science 341, 1238858. https://doi.org/10.1126/science.1238858
Essen, D. van, Zhu, Y., Saccani, S., 2010. A Feed-Forward Circuit Controlling Inducible NF-κB Target Gene Activation by Promoter Histone Demethylation. Mol. Cell 39, 750–760. https://doi.org/10.1016/j.molcel.2010.08.010
Esteller, M., 2007. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat. Rev. Genet. 8, 286–298. https://doi.org/10.1038/nrg2005
Fang, T.C., Schaefer, U., Mecklenbrauker, I., Stienen, A., Dewell, S., Chen, M.S., Rioja, I., Parravicini, V., Prinjha, R.K., Chandwani, R., MacDonald, M.R., Lee, K., Rice, C.M., Tarakhovsky, A., 2012. Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response. J. Exp. Med. 209, 661–669. https://doi.org/10.1084/jem.20112343
Fehri, L.F., Rechner, C., Janßen, S., Mak, T.N., Holland, C., Bartfeld, S., Brüggemann, H., Meyer, T.F., 2009. Helicobacter pylori-induced modification of the histone H3 phosphorylation status in gastric epithelial

cells reflects its impact on cell cycle regulation. Epigenetics 4, 577–586. https://doi.org/10.4161/epi.4.8.10217
Feichtinger, J., Hernández, I., Fischer, C., Hanscho, M., Auer, N., Hackl, M., Jadhav, V., Baumann, M., Krempl, P.M., Schmidl, C., Farlik, M., Schuster, M., Merkel, A., Sommer, A., Heath, S., Rico, D., Bock, C., Thallinger, G.G., Borth, N., 2016. Comprehensive genome and epigenome characterization of CHO cells in response to evolutionary pressures and over time. Biotechnol. Bioeng. 113, 2241–2253. https://doi.org/10.1002/bit.25990
Fernie-King, B., Seilly, D., Davies, A., Lachmann, P., 2002. Subversion of the innate immune response by micro-organisms. Ann. Rheum. Dis. 61, ii8-ii12. https://doi.org/10.1136/ard.61.suppl_2.ii8
Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T., Pasini, D., 2014. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. Mol. Cell 53, 49–62. https://doi.org/10.1016/j.molcel.2013.10.030
Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro Genebra de Jesus, I., Kerkhoven, R.M., Bussemaker, H.J., van Steensel, B., 2010. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell 143, 212–224. https://doi.org/10.1016/j.cell.2010.09.009
Fillion, I., Ouellet, N., Simard, M., Bergeron, Y., Sato, S., Bergeron, M.G., 2001. Role of Chemokines and Formyl Peptides in Pneumococcal Pneumonia-Induced Monocyte/Macrophage Recruitment. J. Immunol. 166, 7353–7361. https://doi.org/10.4049/jimmunol.166.12.7353
Fnu, S., Williamson, E.A., De Haro, L.P., Brenneman, M., Wray, J., Shaheen, M., Radhakrishnan, K., Lee, S.-H., Nickoloff, J.A., Hromas, R., 2011. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. Proc. Natl. Acad. Sci. U. S. A. 108, 540–545. https://doi.org/10.1073/pnas.1013571108
Footz, T.K., Brinkman-Mills, P., Banting, G.S., Maier, S.A., Riazi, M.A., Bridgland, L., Hu, S., Birren, B., Minoshima, S., Shimizu, N., Pan, H., Nguyen, T., Fang, F., Fu, Y., Ray, L., Wu, H., Shaull, S., Phan, S., Yao, Z., Chen, F., Huan, A., Hu, P., Wang, Q., Loh, P., Qi, S., Roe, B.A., McDermid, H.E., 2001. Analysis of the cat eye syndrome critical region in humans and the region of conserved synteny in mice: a search for candidate genes at or near the human chromosome 22 pericentromere. Genome Res. 11, 1053–1070. https://doi.org/10.1101/gr.154901
Forneris, F., Binda, C., Vanoni, M.A., Mattevi, A., Battaglioli, E., 2005. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. FEBS Lett. 579, 2203–2207. https://doi.org/10.1016/j.febslet.2005.03.015
Foster, S.L., Hargreaves, D.C., Medzhitov, R., 2007. Gene-specific control of inflammation by TLR-induced chromatin modifications. Nature 447, 972–978. https://doi.org/10.1038/nature05836
Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L., Strahl, B.D., 2011. Influence of combinatorial histone modifications on antibody and effector protein recognition. Curr. Biol. CB 21, 53–8. https://doi.org/10.1016/j.cub.2010.11.058
Fukazawa, A., Alonso, C., Kurachi, K., Gupta, S., Lesser, C.F., McCormick, B.A., Reinecker, H.-C., 2008. GEF-H1 Mediated Control of NOD1 Dependent

NF-κB Activation by Shigella Effectors. PLOS Pathog. 4, e1000228. https://doi.org/10.1371/journal.ppat.1000228
Galli, S.J., Borregaard, N., Wynn, T.A., 2011. Phenotypic and functional plasticity of cells of innate immunity: macrophages, mast cells and neutrophils. Nat. Immunol. 12, 1035–1044. https://doi.org/10.1038/ni.2109
Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B.D., Allis, C.D., Hunt, D.F., 2007. Organismal Differences in Post-translational Modifications in Histones H3 and H4. J. Biol. Chem. 282, 7641–7655. https://doi.org/10.1074/jbc.M607900200
Garcia, B.A., Mollah, S., Ueberheide, B.M., Busby, S.A., Muratore, T.L., Shabanowitz, J., Hunt, D.F., 2007. Chemical derivatization of histones for facilitated analysis by mass spectrometry. Nat. Protoc. 2, 933–938. https://doi.org/10.1038/nprot.2007.106
Garlanda, C., Hirsch, E., Bozza, S., Salustri, A., Acetis, M.D., Nota, R., Maccagno, A., Riva, F., Bottazzi, B., Peri, G., Doni, A., Vago, L., Botto, M., Santis, R.D., Carminati, P., Siracusa, G., Altruda, F., Vecchi, A., Romani, L., Mantovani, A., 2002. Non-redundant role of the long pentraxin PTX3 in anti-fungal innate immune response. Nature 420, 182. https://doi.org/10.1038/nature01195
Gazzar, M.E., Yoza, B.K., Chen, X., Hu, J., Hawkins, G.A., McCall, C.E., 2008. G9a and HP1 Couple Histone and DNA Methylation to TNFα Transcription Silencing during Endotoxin Tolerance. J. Biol. Chem. 283, 32198–32208. https://doi.org/10.1074/jbc.M803446200
Gerber, M., Shilatifard, A., 2003. Transcriptional Elongation by RNA Polymerase II and Histone Methylation. J. Biol. Chem. 278, 26303–26306. https://doi.org/10.1074/jbc.R300014200
Gessani, S., Belardelli, F., 1998. IFN-γ Expression in Macrophages and Its Possible Biological Significance. Cytokine Growth Factor Rev. 9, 117–123. https://doi.org/10.1016/S1359-6101(98)00007-0
Ghazalpour, A., Bennett, B., Petyuk, V.A., Orozco, L., Hagopian, R., Mungrue, I.N., Farber, C.R., Sinsheimer, J., Kang, H.M., Furlotte, N., Park, C.C., Wen, P.-Z., Brewer, H., Weitz, K., Camp, D.G., Pan, C., Yordanova, R., Neuhaus, I., Tilford, C., Siemers, N., Gargalovic, P., Eskin, E., Kirchgessner, T., Smith, D.J., Smith, R.D., Lusis, A.J., 2011. Comparative Analysis of Proteome and Transcriptome Variation in Mouse. PLoS Genet. 7. https://doi.org/10.1371/journal.pgen.1001393
Ginhoux, F., Jung, S., 2014. Monocytes and macrophages: developmental pathways and tissue homeostasis. Nat. Rev. Immunol. 14, 392–404. https://doi.org/10.1038/nri3671
Goeminne, L.J.E., Gevaert, K., Clement, L., 2016. Peptide-level Robust Ridge Regression Improves Estimation, Sensitivity, and Specificity in Data-dependent Quantitative Label-free Shotgun Proteomics. Mol. Cell. Proteomics 15, 657–668. https://doi.org/10.1074/mcp.M115.055897
Gómez-Díaz, E., Jordà, M., Peinado, M.A., Rivero, A., 2012. Epigenetics of Host–Pathogen Interactions: The Road Ahead and the Road Behind. PLOS Pathog. 8, e1003007. https://doi.org/10.1371/journal.ppat.1003007
Gordon, S., Taylor, P.R., 2005. Monocyte and macrophage heterogeneity. Nat. Rev. Immunol. 5, 953. https://doi.org/10.1038/nri1733

Gray, B.M., Converse, G.M., Dillon, H.C., 1980. Epidemiologic Studies of Streptococcus pneumoniae in Infants: Acquisition, Carriage, and Infection during the First 24 Months of Life. J. Infect. Dis. 142, 923–933. https://doi.org/10.1093/infdis/142.6.923

Guo, H., Ingolia, N.T., Weissman, J.S., Bartel, D.P., 2010. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466, 835. https://doi.org/10.1038/nature09267

Hamon, M.A., Batsché, E., Régnault, B., Tham, T.N., Seveau, S., Muchardt, C., Cossart, P., 2007. Histone modifications induced by a family of bacterial toxins. Proc. Natl. Acad. Sci. U. S. A. 104, 13467–13472. https://doi.org/10.1073/pnas.0702729104

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Calcar, S.V., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E., Ren, B., 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. 39, ng1966. https://doi.org/10.1038/ng1966

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., Gersbach, C.A., 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. 33, 510–517. https://doi.org/10.1038/nbt.3199

Hoeksema, M.A., de Winther, M.P.J., 2016. Epigenetic Regulation of Monocyte and Macrophage Function. Antioxid. Redox Signal. 25, 758–774. https://doi.org/10.1089/ars.2016.6695

Holter, J.C., Müller, F., Bjørang, O., Samdal, H.H., Marthinsen, J.B., Jenum, P.A., Ueland, T., Frøland, S.S., Aukrust, P., Husebye, E., Heggelund, L., 2015. Etiology of community-acquired pneumonia and diagnostic yields of microbiological methods: a 3-year prospective study in Norway. BMC Infect. Dis. 15, 64. https://doi.org/10.1186/s12879-015-0803-5

Horikoshi, N., Kumar, P., Sharma, G.G., Chen, M., Hunt, C.R., Westover, K., Chowdhury, S., Pandita, T.K., 2013. Genome-wide distribution of histone H4 Lysine 16 acetylation sites and their relationship to gene expression. Genome Integr. 4, 3. https://doi.org/10.1186/2041-9414-4-3

Hu, D.-K., Liu, Y., Li, X.-Y., Qu, Y., 2015. In vitro expression of Streptococcus pneumoniae ply gene in human monocytes and pneumocytes. Eur. J. Med. Res. 20. https://doi.org/10.1186/s40001-015-0142-4

Hyams, C., Camberlein, E., Cohen, J.M., Bax, K., Brown, J.S., 2010. The Streptococcus pneumoniae Capsule Inhibits Complement Activity and Neutrophil Phagocytosis by Multiple Mechanisms. Infect. Immun. 78, 704–715. https://doi.org/10.1128/IAI.00881-09

Ifrim, D.C., Quintin, J., Joosten, L.A.B., Jacobs, C., Jansen, T., Jacobs, L., Gow, N.A.R., Williams, D.L., Meer, J.W.M. van der, Netea, M.G., 2014. Trained immunity or tolerance: opposing functional programs induced in human monocytes after engagement of various pattern recognition receptors. Clin. Vaccine Immunol. CVI.00688-13. https://doi.org/10.1128/CVI.00688-13

Ivashkiv, L.B., 2013. Epigenetic regulation of macrophage polarization and function. Trends Immunol. 34, 216–223. https://doi.org/10.1016/j.it.2012.11.001

Jablonski, K.A., Gaudet, A.D., Amici, S.A., Popovich, P.G., Guerau-de-Arellano, M., 2016. Control of the Inflammatory Macrophage Transcriptional

Signature by miR-155. PLOS ONE 11, e0159724. https://doi.org/10.1371/journal.pone.0159724
Jack, D.L., Cole, J., Naylor, S.C., Borrow, R., Kaczmarski, E.B., Klein, N.J., Read, R.C., 2006. Genetic polymorphism of the binding domain of surfactant protein-A2 increases susceptibility to meningococcal disease. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. 43, 1426–1433. https://doi.org/10.1086/508775
Jäkel, S., Gürlich, D., 1998. Importin β, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. EMBO J. 17, 4491–4502. https://doi.org/10.1093/emboj/17.15.4491
Jenner, R.G., Young, R.A., 2005. Insights into host responses against pathogens from transcriptional profiling. Nat. Rev. Microbiol. 3, 281–294. https://doi.org/10.1038/nrmicro1126
Jenuwein, T., Allis, C.D., 2001. Translating the histone code. Science 293, 1074–1080. https://doi.org/10.1126/science.1063127
Jeong, D.-G., Seo, J.-H., Heo, S.-H., Choi, Y.-K., Jeong, E.-S., 2015. Tumor necrosis factor-alpha deficiency impairs host defense against Streptococcus pneumoniae. Lab. Anim. Res. 31, 78–85. https://doi.org/10.5625/lar.2015.31.2.78
Jiang, X., An, Z., Lu, C., Chen, Y., Du, E., Qi, S., Yang, K., Zhang, Z., Xu, Y., 2016. The protective role of Nrf2-Gadd45b against antimony-induced oxidative stress and apoptosis in HEK293 cells. Toxicol. Lett. 256, 11–18. https://doi.org/10.1016/j.toxlet.2016.05.016
Johnson, D.S., Mortazavi, A., Myers, R.M., Wold, B., 2007. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. Science 316, 1497–1502. https://doi.org/10.1126/science.1141319
Juan, A.H., Wang, S., Ko, K.D., Zare, H., Tsai, P.-F., Feng, X., Vivanco, K.O., Ascoli, A.M., Gutierrez-Cruz, G., Krebs, J., Sidoli, S., Knight, A.L., Pedersen, R.A., Garcia, B.A., Casellas, R., Zou, J., Sartorelli, V., 2016. Roles of H3K27me2 and H3K27me3 Examined During Fate Specification of Embryonic Stem Cells. Cell Rep. 17, 1369–1382. https://doi.org/10.1016/j.celrep.2016.09.087
Julien, E., Herr, W., 2004. A Switch in Mitotic Histone H4 Lysine 20 Methylation Status Is Linked to M Phase Defects upon Loss of HCF-1. Mol. Cell 14, 713–725. https://doi.org/10.1016/j.molcel.2004.06.008
Kadioglu, A., Weiser, J.N., Paton, J.C., Andrew, P.W., 2008a. The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. Nat. Rev. Microbiol. 6, 288–301. https://doi.org/10.1038/nrmicro1871
Kadioglu, A., Weiser, J.N., Paton, J.C., Andrew, P.W., 2008b. The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. Nat. Rev. Microbiol. 6, 288–301. https://doi.org/10.1038/nrmicro1871
Kaikkonen, M.U., Spann, N., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., Glass, C.K., 2013. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol. Cell 51, 310–325. https://doi.org/10.1016/j.molcel.2013.07.010
Kalam, H., Fontana, M.F., Kumar, D., 2017. Alternate splicing of transcripts shape macrophage response to Mycobacterium tuberculosis infection.

PLOS Pathog. 13, e1006236. https://doi.org/10.1371/journal.ppat.1006236
Karch, K.R., Sidoli, S., Garcia, B.A., 2016. Identification and Quantification of Histone PTMs Using High-Resolution Mass Spectrometry, in: Methods in Enzymology. Elsevier, pp. 3–29. https://doi.org/10.1016/bs.mie.2015.12.007
Karch, K.R., Zee, B.M., Garcia, B.A., 2014. High resolution is not a strict requirement for characterization and quantification of histone post-translational modifications. J. Proteome Res. 13, 6152–6159. https://doi.org/10.1021/pr500902f
Karlić, R., Chung, H.-R., Lasserre, J., Vlahoviček, K., Vingron, M., 2010. Histone modification levels are predictive for gene expression. Proc. Natl. Acad. Sci. 107, 2926–2931. https://doi.org/10.1073/pnas.0909344107
Kato, K., Uchino, R., Lillehoj, E.P., Knox, K., Lin, Y., Kim, K.C., 2016. Membrane-Tethered MUC1 Mucin Counter-Regulates the Phagocytic Activity of Macrophages. Am. J. Respir. Cell Mol. Biol. 54, 515–523. https://doi.org/10.1165/rcmb.2015-0177OC
Kim, D., Langmead, B., Salzberg, S.L., 2015. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360. https://doi.org/10.1038/nmeth.3317
Kirby, A.C., Raynes, J.G., Kaye, P.M., 2005. The role played by tumor necrosis factor during localized and systemic infection with Streptococcus pneumoniae. J. Infect. Dis. 191, 1538–1547.
Kleinnijenhuis, J., Quintin, J., Preijers, F., Joosten, L.A.B., Ifrim, D.C., Saeed, S., Jacobs, C., van Loenhout, J., de Jong, D., Stunnenberg, H.G., Xavier, R.J., van der Meer, J.W.M., van Crevel, R., Netea, M.G., 2012. Bacille Calmette-Guerin induces NOD2-dependent nonspecific protection from reinfection via epigenetic reprogramming of monocytes. Proc. Natl. Acad. Sci. U. S. A. 109, 17537–17542. https://doi.org/10.1073/pnas.1202870109
Kornberg, R.D., 1974. Chromatin structure: a repeating unit of histones and DNA. Science 184, 868–871.
Kouzarides, T., 2002. Histone methylation in transcriptional control. Curr. Opin. Genet. Amp Dev. 12, 198–209.
Krautkramer, K.A., Reiter, L., Denu, J.M., Dowell, J.A., 2015. Quantification of SAHA-Dependent Changes in Histone Modifications Using Data-Independent Acquisition Mass Spectrometry. J. Proteome Res. 14, 3252–3262. https://doi.org/10.1021/acs.jproteome.5b00245
Kruidenier, L., Chung, C., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H., Eberhard, D., Hutchinson, S., Jones, E., Katso, R., Leveridge, M., Mander, P.K., Mosley, J., Ramirez-Molina, C., Rowland, P., Schofield, C.J., Sheppard, R.J., Smith, J.E., Swales, C., Tanner, R., Thomas, P., Tumber, A., Drewes, G., Oppermann, U., Patel, D.J., Lee, K., Wilson, D.M., 2012. A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. Nature 488, 404–408. https://doi.org/10.1038/nature11262
Kulej, K., Avgousti, D.C., Weitzman, M.D., Garcia, B.A., 2015. Characterization of histone post-translational modifications during virus infection using mass spectrometry-based proteomics. Methods 90, 8–20. https://doi.org/10.1016/j.ymeth.2015.06.008

Kurdistani, S.K., Tavazoie, S., Grunstein, M., 2004. Mapping Global Histone Acetylation Patterns to Gene Expression. Cell 117, 721–733. https://doi.org/10.1016/j.cell.2004.05.023

Ladhani, S.N., Slack, M.P.E., Andrews, N.J., Waight, P.A., Borrow, R., Miller, E., 2013. Invasive Pneumococcal Disease after Routine Pneumococcal Conjugate Vaccination in Children, England and Wales. Emerg. Infect. Dis. 19, 61–68. https://doi.org/10.3201/eid1901.120741

Lammas, D.A., Casanova, J.-L., Kumararatne, D.S., 2000. Clinical consequences of defects in the IL-12-dependent interferon-gamma (IFN-γ) pathway. Clin. Exp. Immunol. 121, 417–425. https://doi.org/10.1046/j.1365-2249.2000.01284.x

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C., Fisher-Aylor, K.I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A.J., Hoffman, M.M., Iyer, V.R., Jung, Y.L., Karmakar, S., Kellis, M., Kharchenko, P.V., Li, Q., Liu, T., Liu, X.S., Ma, L., Milosavljevic, A., Myers, R.M., Park, P.J., Pazin, M.J., Perry, M.D., Raha, D., Reddy, T.E., Rozowsky, J., Shoresh, N., Sidow, A., Slattery, M., Stamatoyannopoulos, J.A., Tolstorukov, M.Y., White, K.P., Xi, S., Farnham, P.J., Lieb, J.D., Wold, B.J., Snyder, M., 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. 22, 1813–1831. https://doi.org/10.1101/gr.136184.111

Lee, J.Y., Kim, N.A., Sanford, A., Sullivan, K.E., 2003. Histone acetylation and chromatin conformation are regulated separately at the TNF-α promoter in monocytes and macrophages. J. Leukoc. Biol. 73, 862–871. https://doi.org/10.1189/jlb.1202618

Lesage, S., Drouet, V., Majounie, E., Deramecourt, V., Jacoupy, M., Nicolas, A., Cormier-Dequaire, F., Hassoun, S.M., Pujol, C., Ciura, S., Erpapazoglou, Z., Usenko, T., Maurage, C.A., Sahbatou, M., Liebau, S., Ding, J., Bilgic, B., Emre, M., Erginel-Unaltuna, N., Guven, G., Tison, F., Tranchant, C., Vidailhet, M., Corvol, J.C., Krack, P., Leutenegger, A.L., Nalls, M.A., Hernandez, D.G., Heutink, P., Gibbs, J.R., Hardy, J., Wood, N.W., Gasser, T., Durr, A., Deleuze, J.F., Tazir, M., Destée, A., Lohmann, E., Kabashi, E., Singleton, A., Corti, O., Brice, A., 2016. Loss of VPS13C Function in Autosomal-Recessive Parkinsonism Causes Mitochondrial Dysfunction and Increases PINK1/Parkin-Dependent Mitophagy., Loss of VPS13C Function in Autosomal-Recessive Parkinsonism Causes Mitochondrial Dysfunction and Increases PINK1/Parkin-Dependent Mitophagy. Am. J. Hum. Genet. Am. J. Hum. Genet. 98, 98, 500, 500–513. https://doi.org/10.1016/j.ajhg.2016.01.014, 10.1016/j.ajhg.2016.01.014

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma. Oxf. Engl. 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Wei, S., Fang, Y., Li, M., Li, X., Li, Z., Zhang, J., Zhu, G., Li, C., Bi, L., Zhang, G., Wang, D., Zhang, X.-E., 2017. Quantitative proteomic analysis of host responses triggered by Mycobacterium tuberculosis infection in human macrophage cells. Acta Biochim. Biophys. Sin. 49, 835–844. https://doi.org/10.1093/abbs/gmx080

Li, P., Wang, R., Dong, W., Hu, L., Zong, B., Zhang, Y., Wang, X., Guo, A., Zhang, A., Xiang, Y., Chen, H., Tan, C., 2017. Comparative Proteomics Analysis

of Human Macrophages Infected with Virulent Mycobacterium bovis. Front. Cell. Infect. Microbiol. 7. https://doi.org/10.3389/fcimb.2017.00065
Li, S., Miao, T., Sebastian, M., Bhullar, P., Ghaffari, E., Liu, M., Symonds, A.L.J., Wang, P., 2012. The Transcription Factors Egr2 and Egr3 Are Essential for the Control of Inflammation and Antigen-Induced Proliferation of B and T Cells. Immunity 37, 685–696. https://doi.org/10.1016/j.immuni.2012.08.001
Li, Y., Liu, B., Fukudome, E.Y., Kochanek, A.R., Finkelstein, R.A., Chong, W., Jin, G., Lu, J., deMoya, M.A., Velmahos, G.C., Alam, H.B., 2010. Surviving lethal septic shock without fluid resuscitation in a rodent model. Surgery 148, 246–254. https://doi.org/10.1016/j.surg.2010.05.003
Liang, G., Lin, J.C.Y., Wei, V., Yoo, C., Cheng, J.C., Nguyen, C.T., Weisenberger, D.J., Egger, G., Takai, D., Gonzales, F.A., Jones, P.A., 2004. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proc. Natl. Acad. Sci. U. S. A. 101, 7357–7362. https://doi.org/10.1073/pnas.0401866101
Lienert, F., Mohn, F., Tiwari, V.K., Baubec, T., Roloff, T.C., Gaidatzis, D., Stadler, M.B., Schübeler, D., 2011. Genomic Prevalence of Heterochromatic H3K9me2 and Transcription Do Not Discriminate Pluripotent from Terminally Differentiated Cells. PLoS Genet. 7. https://doi.org/10.1371/journal.pgen.1002090
Lin, B., Kolluri, S.K., Lin, F., Liu, W., Han, Y.-H., Cao, X., Dawson, M.I., Reed, J.C., Zhang, X., 2004. Conversion of Bcl-2 from Protector to Killer by Interaction with Nuclear Orphan Receptor Nur77/TR3. Cell 116, 527–540. https://doi.org/10.1016/S0092-8674(04)00162-X
Lin, S., Wein, S., Gonzales-Cope, M., Otte, G.L., Yuan, Z.-F., Afjehi-Sadat, L., Maile, T., Berger, S.L., Rush, J., Lill, J.R., Arnott, D., Garcia, B.A., 2014. Stable-isotope-labeled histone peptide library for histone post-translational modification and variant quantification by mass spectrometry. Mol. Cell. Proteomics MCP 13, 2450–2466. https://doi.org/10.1074/mcp.O113.036459
Liu, M., Guo, S., Hibbert, J.M., Jain, V., Singh, N., Wilson, N.O., Stiles, J.K., 2011. CXCL10/IP-10 in infectious diseases pathogenesis and potential therapeutic implications. Cytokine Growth Factor Rev. 22, 121–130. https://doi.org/10.1016/j.cytogfr.2011.06.001
Lu, L., Chen, X., Sanders, D., Qian, S., Zhong, X., 2015. High-resolution mapping of H4K16 and H3K23 acetylation reveals conserved and unique distribution patterns in Arabidopsis and rice. Epigenetics 10, 1044–1053. https://doi.org/10.1080/15592294.2015.1104446
Lugrin, J., Ciarlo, E., Santos, A., Grandmaison, G., dos Santos, I., Le Roy, D., Roger, T., 2013. The sirtuin inhibitor cambinol impairs MAPK signaling, inhibits inflammatory and innate immune responses and protects from septic shock. Biochim. Biophys. Acta 1833, 1498–1510. https://doi.org/10.1016/j.bbamcr.2013.03.004
Lv, D., Jia, F., Hou, Y., Sang, Y., Alvarez, A.A., Zhang, W., Gao, W.-Q., Hu, B., Cheng, S.-Y., Ge, J., Li, Y., Feng, H., 2017. Histone Acetyltransferase KAT6A Upregulates PI3K/AKT Signaling through TRIM24 Binding. Cancer Res. 77, 6190–6201. https://doi.org/10.1158/0008-5472.CAN-17-1388
Ma, L., Yuan, L., An, J., Barton, M.C., Zhang, Q., Liu, Z., 2016. Histone H3 lysine 23 acetylation is associated with oncogene TRIM24 expression and a poor

prognosis in breast cancer. Tumor Biol. 37, 14803–14812. https://doi.org/10.1007/s13277-016-5344-z
Macintyre, G., Alford, T., Xiong, L., Rouleau, G.A., Tibbo, P.G., Cox, D.W., 2010. Association of NPAS3 exonic variation with schizophrenia. Schizophr. Res. 120, 143–149. https://doi.org/10.1016/j.schres.2010.04.002
MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., MacCoss, M.J., 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26, 966–968. https://doi.org/10.1093/bioinformatics/btq054
Mahajan, S., Saini, A., Chandra, V., Nanduri, R., Kalra, R., Bhagyaraj, E., Khatri, N., Gupta, P., 2015. Nuclear Receptor Nr4a2 Promotes Alternative Polarization of Macrophages and Confers Protection in Sepsis♦. J. Biol. Chem. 290, 18304–18314. https://doi.org/10.1074/jbc.M115.638064
Marconi, C., Canobbio, I., Bozzi, V., Pippucci, T., Simonetti, G., Melazzini, F., Angori, S., Martinelli, G., Saglio, G., Torti, M., Pastan, I., Seri, M., Pecci, A., 2017. 5'UTR point substitutions and N-terminal truncating mutations of ANKRD26 in acute myeloid leukemia. J. Hematol. Oncol.J Hematol Oncol 10, 18. https://doi.org/10.1186/s13045-016-0382-y
Marriott, H.M., Daigneault, M., Thompson, A.A.R., Walmsley, S.R., Gill, S.K., Witcher, D.R., Wroblewski, V.J., Hellewell, P.G., Whyte, M.K.B., Dockrell, D.H., 2012. A decoy receptor 3 analogue reduces localised defects in phagocyte function in pneumococcal pneumonia. Thorax 67, 985–992. https://doi.org/10.1136/thoraxjnl-2012-201591
Marriott, H.M., Mitchell, T.J., Dockrell, D.H., 2008. Pneumolysin: a double-edged sword during the host-pathogen interaction. Curr. Mol. Med. 8, 497–509.
Martinez, F.O., Sica, A., Mantovani, A., Locati, M., 2008. Macrophage activation and polarization. Front. Biosci. J. Virtual Libr. 13, 453–461.
McGEE, A.M., Baines, C.P., 2011. Complement 1q-binding protein inhibits the mitochondrial permeability transition pore and protects against oxidative stress-induced death. Biochem. J. 433, 119–125. https://doi.org/10.1042/BJ20101431
McNeela, E.A., Burke, Á., Neill, D.R., Baxter, C., Fernandes, V.E., Ferreira, D., Smeaton, S., El-Rachkidy, R., McLoughlin, R.M., Mori, A., Moran, B., Fitzgerald, K.A., Tschopp, J., Pétrilli, V., Andrew, P.W., Kadioglu, A., Lavelle, E.C., 2010. Pneumolysin Activates the NLRP3 Inflammasome and Promotes Proinflammatory Cytokines Independently of TLR4. PLoS Pathog 6, e1001191. https://doi.org/10.1371/journal.ppat.1001191
Meinsma, R., van Kuilenburg, A.B.P., 2016. Purification, activity, and expression levels of two uridine-cytidine kinase isoforms in neuroblastoma cell lines. Nucleosides Nucleotides Nucleic Acids 35, 613–618. https://doi.org/10.1080/15257770.2015.1124998
Metzger, E., Wissmann, M., Yin, N., Müller, J.M., Schneider, R., Peters, A.H.F.M., Günther, T., Buettner, R., Schüle, R., 2005. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437, 436–439. https://doi.org/10.1038/nature04020
Miao, F., Natarajan, R., 2005. Mapping global histone methylation patterns in the coding regions of human genes. Mol. Cell. Biol. 25, 4650–4661. https://doi.org/10.1128/MCB.25.11.4650-4661.2005

Minshull, T.C., Cole, J., Dockrell, D.H., Read, R.C., Dickman, M.J., 2016. Analysis of histone post translational modifications in primary monocyte derived macrophages using reverse phase×reverse phase chromatography in conjunction with porous graphitic carbon stationary phase. J. Chromatogr. A 1453, 43–53. https://doi.org/10.1016/j.chroma.2016.05.025

Minshull, T.C., Dickman, M.J., 2015. Mass spectrometry analysis of histone post translational modifications. Drug Discov. Today Dis. Models. https://doi.org/10.1016/j.ddmod.2015.03.002

Miousse, I.R., Chalbot, M.-C.G., Aykin-Burns, N., Wang, X., Basnakian, A., Kavouras, I.G., Koturbash, I., 2014. Epigenetic alterations induced by ambient particulate matter in mouse macrophages. Environ. Mol. Mutagen. n/a–n/a. https://doi.org/10.1002/em.21855

Mitchell, A.M., Mitchell, T.J., 2010. Streptococcus pneumoniae: virulence factors and variation. Clin. Microbiol. Infect. 16, 411–418. https://doi.org/10.1111/j.1469-0691.2010.03183.x

Molden, R.C., Garcia, B.A., 2014. Middle-Down and Top-Down Mass Spectrometric Analysis of Co-occurring Histone Modifications. Curr. Protoc. Protein Sci. 77, 23.7.1-28. https://doi.org/10.1002/0471140864.ps2307s77

Mombelli, M., Lugrin, J., Rubino, I., Chanson, A.-L., Giddey, M., Calandra, T., Roger, T., 2011. Histone Deacetylase Inhibitors Impair Antibacterial Defenses of Macrophages. J. Infect. Dis. 204, 1367–1374. https://doi.org/10.1093/infdis/jir553

Moradian, A., Kalli, A., Sweredoski, M.J., Hess, S., 2014. The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications. PROTEOMICS 14, 489–497. https://doi.org/10.1002/pmic.201300256

Moret, Y., Siva-Jothy, M.T., 2003. Adaptive innate immunity? Responsive-mode prophylaxis in the mealworm beetle, Tenebrio molitor. Proc. R. Soc. Lond. B Biol. Sci. 270, 2475–2480. https://doi.org/10.1098/rspb.2003.2511

Mukovozov, I., Huang, Y.-W., Zhang, Q., Liu, G.Y., Siu, A., Sokolskyy, Y., Patel, S., Hyduk, S.J., Kutryk, M.J.B., Cybulsky, M.I., Robinson, L.A., 2015. The Neurorepellent Slit2 Inhibits Postadhesion Stabilization of Monocytes Tethered to Vascular Endothelial Cells. J. Immunol. 195, 3334–3344. https://doi.org/10.4049/jimmunol.1500640

Muller, S., Filippakopoulos, P., Knapp, S., 2011. Bromodomains as therapeutic targets. Expert Rev. Mol. Med. 13. https://doi.org/10.1017/S1462399411001992

Mutskov, V., Raaka, B.M., Felsenfeld, G., Gershengorn, M.C., 2007. The human insulin gene displays transcriptionally active epigenetic marks in islet-derived mesenchymal precursor cells in the absence of insulin expression. Stem Cells Dayt. Ohio 25, 3223–3233. https://doi.org/10.1634/stemcells.2007-0325

Natoli, G., 2010. Maintaining cell identity through global control of genomic organization. Immunity 33, 12–24. https://doi.org/10.1016/j.immuni.2010.07.006

Nau, G.J., Richmond, J.F.L., Schlesinger, A., Jennings, E.G., Lander, E.S., Young, R.A., 2002. Human macrophage activation programs induced by

bacterial pathogens. Proc. Natl. Acad. Sci. 99, 1503–1508. https://doi.org/10.1073/pnas.022649799
Nemerson, Y., 1988. Tissue factor and hemostasis [published erratum appears in Blood 1988 Apr;71(4):1178]. Blood 71, 1–8.
Netea, M.G., Quintin, J., van der Meer, J.W.M., 2011. Trained Immunity: A Memory for Innate Host Defense. Cell Host Microbe 9, 355–361. https://doi.org/10.1016/j.chom.2011.04.006
Nicholas, D., Tang, H., Zhang, Q., Rudra, J., Xu, F., Langridge, W., Zhang, K., 2014. Quantitative Proteomics Reveals a Role for Epigenetic Reprogramming During Human Monocyte Differentiation. Mol. Cell. Proteomics MCP. https://doi.org/10.1074/mcp.M113.035089
Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P., Reinberg, D., 2002. Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes Dev. 16, 479–489. https://doi.org/10.1101/gad.967202
O'Brien, K.L., Wolfson, L.J., Watt, J.P., Henkle, E., Deloria-Knoll, M., McCall, N., Lee, E., Mulholland, K., Levine, O.S., Cherian, T., Hib and Pneumococcal Global Burden of Disease Study Team, 2009. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet 374, 893–902. https://doi.org/10.1016/S0140-6736(09)61204-6
O'Connell, R.M., Taganov, K.D., Boldin, M.P., Cheng, G., Baltimore, D., 2007. MicroRNA-155 is induced during the macrophage inflammatory response. Proc. Natl. Acad. Sci. 104, 1604–1609. https://doi.org/10.1073/pnas.0610731104
Oda, H., Okamoto, I., Murphy, N., Chu, J., Price, S.M., Shen, M.M., Torres-Padilla, M.E., Heard, E., Reinberg, D., 2009. Monomethylation of Histone H4-Lysine 20 Is Involved in Chromosome Structure and Stability and Is Essential for Mouse Development. Mol. Cell. Biol. 29, 2278–2295. https://doi.org/10.1128/MCB.01768-08
Okabe, A., Funata, S., Matsusaka, K., Namba, H., Fukuyo, M., Rahmutulla, B., Oshima, M., Iwama, A., Fukayama, M., Kaneda, A., 2017. Regulation of tumour related genes by dynamic epigenetic alteration at enhancer regions in gastric epithelial cells infected by Epstein-Barr virus. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-08370-7
Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V.M., Su, L., Xu, G., Zhang, Y., 2005. hDOT1L Links Histone Methylation to Leukemogenesis. Cell 121, 167–178. https://doi.org/10.1016/j.cell.2005.02.020
Orlando, D.A., Chen, M.W., Brown, V.E., Solanki, S., Choi, Y.J., Olson, E.R., Fritz, C.C., Bradner, J.E., Guenther, M.G., 2014. Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. Cell Rep. 9, 1163–1170. https://doi.org/10.1016/j.celrep.2014.10.018
Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., Natoli, G., 2013. Latent Enhancers Activated by Stimulation in Differentiated Cells. Cell 152, 157–171. https://doi.org/10.1016/j.cell.2012.12.018
Ou, D.-L., Shyue, S.-K., Lin, L.-I., Feng, Z.-R., Liou, J.-Y., Fan, H.-H., Lee, B.-S., Hsu, C., Cheng, A.-L., 2015. Growth arrest DNA damage-inducible gene

45 gamma expression as a prognostic and predictive biomarker in hepatocellular carcinoma. Oncotarget 6, 27953–27965.
Palani, S., Elima, K., Ekholm, E., Jalkanen, S., Salmi, M., 2016. Monocyte Stabilin-1 Suppresses the Activation of Th1 Lymphocytes. J. Immunol. 196, 115–123. https://doi.org/10.4049/jimmunol.1500257
Paredes, V., Park, J.S., Jeong, Y., Yoon, J., Baek, K., 2013. Unstable expression of recombinant antibody during long-term culture of CHO cells is accompanied by histone H3 hypoacetylation. Biotechnol. Lett. 35, 987–993. https://doi.org/10.1007/s10529-013-1168-8
Pascal, L.E., True, L.D., Campbell, D.S., Deutsch, E.W., Risk, M., Coleman, I.M., Eichner, L.J., Nelson, P.S., Liu, A.Y., 2008. Correlation of mRNA and protein levels: Cell type-specific gene expression of cluster designation antigens in the prostate. BMC Genomics 9, 246. https://doi.org/10.1186/1471-2164-9-246
Paschos, K., Allday, M.J., 2010. Epigenetic reprogramming of host genes in viral and microbial pathogenesis. Trends Microbiol. 18, 439–447. https://doi.org/10.1016/j.tim.2010.07.003
Pathak, S.K., Basu, S., Bhattacharyya, A., Pathak, S., Banerjee, A., Basu, J., Kundu, M., 2006. TLR4-Dependent NF-κB Activation and Mitogen- and Stress-Activated Protein Kinase 1-Triggered Phosphorylation Events Are Central to Helicobacter pylori Peptidyl Prolyl cis-, trans-Isomerase (HP0175)-Mediated Induction of IL-6 Release from Macrophages. J. Immunol. 177, 7950–7958. https://doi.org/10.4049/jimmunol.177.11.7950
Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417. https://doi.org/10.1038/nmeth.4197
Patwardhan, S., Gashler, A., Siegel, M.G., Chang, L.C., Joseph, L.J., Shows, T.B., Le, M.B., Sukhatme, V.P., 1991. EGR3, a novel member of the Egr family of genes encoding immediate-early transcription factors. Oncogene 6, 917–928.
Peach, S.E., Rudomin, E.L., Udeshi, N.D., Carr, S.A., Jaffe, J.D., 2012. Quantitative Assessment of Chromatin Immunoprecipitation Grade Antibodies Directed against Histone Modifications Reveals Patterns of Co-occurring Marks on Histone Protein Molecules. Mol. Cell. Proteomics 11, 128–137. https://doi.org/10.1074/mcp.M111.015941
Pei, L., Castrillo, A., Chen, M., Hoffmann, A., Tontonoz, P., 2005. Induction of NR4A orphan nuclear receptor expression in macrophages in response to inflammatory stimuli. J. Biol. Chem. 280, 29256–29262. https://doi.org/10.1074/jbc.M502606200
Pei, L., Castrillo, A., Tontonoz, P., 2006. Regulation of Macrophage Inflammatory Gene Expression by the Orphan Nuclear Receptor Nur77. Mol. Endocrinol. 20, 786–794. https://doi.org/10.1210/me.2005-0331
Pellegrino, S., Michelena, J., Teloni, F., Imhof, R., Altmeyer, M., 2017. Replication-Coupled Dilution of H4K20me2 Guides 53BP1 to Pre-replicative Chromatin. Cell Rep. 19, 1819–1831. https://doi.org/10.1016/j.celrep.2017.05.016
Perea, L., Coll, M., Sanjurjo, L., Blaya, D., Taghdouini, A.E., Rodrigo-Torres, D., Altamirano, J., Graupera, I., Aguilar-Bravo, B., Llopis, M., Vallverdú, J., Caballeria, J., van Grunsven, L.A., Sarrias, M.-R., Ginès, P., Sancho-Bru, P., 2017. Pentraxin-3 modulates lipopolysaccharide-induced inflammatory

response and attenuates liver injury. Hepatology 66, 953–968. https://doi.org/10.1002/hep.29215
Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., Salzberg, S.L., 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. 11, 1650. https://doi.org/10.1038/nprot.2016.095
Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T., Salzberg, S.L., 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290–295. https://doi.org/10.1038/nbt.3122
Pesavento, J.J., Bullock, C.R., LeDuc, R.D., Mizzen, C. a, Kelleher, N.L., 2008. Combinatorial modification of human histone H4 quantitated by two-dimensional liquid chromatography coupled with top down mass spectrometry. J. Biol. Chem. 283, 14927–37. https://doi.org/10.1074/jbc.M709796200
Pesavento, J.J., Kim, Y.-B., Taylor, G.K., Kelleher, N.L., 2004. Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. J. Am. Chem. Soc. 126, 3386–7. https://doi.org/10.1021/ja039748i
Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., Coon, J.J., 2012. Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. Mol. Cell. Proteomics MCP 11, 1475–1488. https://doi.org/10.1074/mcp.O112.020131
Pham, L.N., Dionne, M.S., Shirasu-Hiza, M., Schneider, D.S., 2007. A Specific Primed Immune Response in Drosophila Is Dependent on Phagocytes. PLoS Pathog 3, e26. https://doi.org/10.1371/journal.ppat.0030026
Phanstiel, D., Brumbaugh, J., Berggren, W.T., Conard, K., Feng, X., Levenstein, M.E., McAlister, G.C., Thomson, J.A., Coon, J.J., 2008. Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. Proc. Natl. Acad. Sci. 105, 4093–4098. https://doi.org/10.1073/pnas.0710515105
Philibert, R.A., Sears, R.A., Powers, L.S., Nash, E., Bair, T., Gerke, A.K., Hassan, I., Thomas, C.P., Gross, T.J., Monick, M.M., 2012. Coordinated DNA methylation and gene expression changes in smoker alveolar macrophages: specific effects on VEGF receptor 1 expression. J. Leukoc. Biol. 92, 621–631. https://doi.org/10.1189/jlb.1211632
Plazas-Mayorca, M.D., Zee, B.M., Young, N.L., Fingerman, I.M., LeRoy, G., Briggs, S.D., Garcia, B.A., 2009. One-Pot Shotgun Quantitative Mass Spectrometry Characterization of Histones. J. Proteome Res. 8, 5367–5374. https://doi.org/10.1021/pr900777e
Pneumococcal disease: cases caused by strains covered by Prevenar 13 vaccine - GOV.UK [WWW Document], n.d. URL https://www.gov.uk/government/publications/pneumococcal-disease-cases-caused-by-strains-covered-by-prevenar-13-vaccine (accessed 3.3.18).
Pneumococcal disease infections caused by serotypes not in Prevenar 13 vaccine - GOV.UK [WWW Document], n.d. URL https://www.gov.uk/government/publications/pneumococcal-disease-caused-by-strains-not-covered-by-prevenar-13-vaccine/pneumococcal-

disease-infections-caused-by-serotypes-not-in-prevenar-13-vaccine (accessed 3.3.18).
Pradeepa, M.M., 2017. Causal role of histone acetylations in enhancer function. Transcription 8, 40–47. https://doi.org/10.1080/21541264.2016.1253529
Proost, P., Menten, P., Struyf, S., Schutyser, E., Meester, I.D., Damme, J.V., 2000. Cleavage by CD26/dipeptidyl peptidase IV converts the chemokine LD78β into a most efficient monocyte attractant and CCR1 agonist. Blood 96, 1674–1680.
Qu, Z., Gao, F., Li, L., Zhang, Y., Jiang, Y., Yu, L., Zhou, Y., Zheng, H., Tong, W., Li, G., Tong, G., 2017. Label-Free Quantitative Proteomic Analysis of Differentially Expressed Membrane Proteins of Pulmonary Alveolar Macrophages Infected with Highly Pathogenic Porcine Reproductive and Respiratory Syndrome Virus and Its Attenuated Strain. PROTEOMICS 17, n/a-n/a. https://doi.org/10.1002/pmic.201700101
Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033
Quintin, J., Saeed, S., Martens, J.H.A., Giamarellos-Bourboulis, E.J., Ifrim, D.C., Logie, C., Jacobs, L., Jansen, T., Kullberg, B.-J., Wijmenga, C., Joosten, L.A.B., Xavier, R.J., van der Meer, J.W.M., Stunnenberg, H.G., Netea, M.G., 2012. Candida albicans infection affords protection against reinfection via functional reprogramming of monocytes. Cell Host Microbe 12, 223–232. https://doi.org/10.1016/j.chom.2012.06.006
Raciti, G.A., Bera, T.K., Gavrilova, O., Pastan, I., 2011. Partial inactivation of Ankrd26 causes diabetes with enhanced insulin responsiveness of adipose tissue in mice. Diabetologia 54, 2911–2922. https://doi.org/10.1007/s00125-011-2263-9
Ramos-Sevillano, E., Moscoso, M., García, P., García, E., Yuste, J., 2011. Nasopharyngeal colonization and invasive disease are enhanced by the cell wall hydrolases LytB and LytC of Streptococcus pneumoniae. PloS One 6, e23626. https://doi.org/10.1371/journal.pone.0023626
Reading, P.C., Bozza, S., Gilbertson, B., Tate, M., Moretti, S., Job, E.R., Crouch, E.C., Brooks, A.G., Brown, L.E., Bottazzi, B., Romani, L., Mantovani, A., 2008. Antiviral activity of the long chain pentraxin PTX3 against influenza viruses. J. Immunol. Baltim. Md 1950 180, 3391–3398.
Rhee, H.S., Bataille, A.R., Zhang, L., Pugh, B.F., 2014. Subnucleosomal Structures and Nucleosome Asymmetry across a Genome. Cell 159, 1377–1388. https://doi.org/10.1016/j.cell.2014.10.054
Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., Amin, V., Whitaker, J.W., Schultz, M.D., Ward, L.D., Sarkar, A., Quon, G., Sandstrom, R.S., Eaton, M.L., Wu, Y.-C., Pfenning, A.R., Wang, X., Claussnitzer, M., Liu, Y., Coarfa, C., Harris, R.A., Shoresh, N., Epstein, C.B., Gjoneska, E., Leung, D., Xie, W., Hawkins, R.D., Lister, R., Hong, C., Gascard, P., Mungall, A.J., Moore, R., Chuah, E., Tam, A., Canfield, T.K., Hansen, R.S., Kaul, R., Sabo, P.J., Bansal, M.S., Carles, A., Dixon, J.R., Farh, K.-H., Feizi, S., Karlic, R., Kim, A.-R., Kulkarni, A., Li, D., Lowdon, R., Elliott, G., Mercer, T.R., Neph, S.J., Onuchic, V., Polak, P., Rajagopal, N., Ray, P., Sallari, R.C., Siebenthall, K.T., Sinnott-Armstrong, N.A., Stevens, M., Thurman, R.E., Wu, J., Zhang, B., Zhou, X.,

Beaudet, A.E., Boyer, L.A., De Jager, P.L., Farnham, P.J., Fisher, S.J., Haussler, D., Jones, S.J.M., Li, W., Marra, M.A., McManus, M.T., Sunyaev, S., Thomson, J.A., Tlsty, T.D., Tsai, L.-H., Wang, W., Waterland, R.A., Zhang, M.Q., Chadwick, L.H., Bernstein, B.E., Costello, J.F., Ecker, J.R., Hirst, M., Meissner, A., Milosavljevic, A., Ren, B., Stamatoyannopoulos, J.A., Wang, T., Kellis, M., 2015. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330. https://doi.org/10.1038/nature14248

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26. https://doi.org/10.1038/nbt.1754

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Roger, T., Lugrin, J., Roy, D.L., Goy, G., Mombelli, M., Koessler, T., Ding, X.C., Chanson, A.-L., Reymond, M.K., Miconnet, I., Schrenzel, J., François, P., Calandra, T., 2011. Histone deacetylase inhibitors impair innate immune responses to Toll-like receptor agonists and to infection. Blood 117, 1205–1217. https://doi.org/10.1182/blood-2010-05-284711

Rogers, P.D., Thornton, J., Barker, K.S., McDaniel, D.O., Sacks, G.S., Swiatlo, E., McDaniel, L.S., 2003. Pneumolysin-Dependent and -Independent Gene Expression Identified by cDNA Microarray Analysis of THP-1 Human Mononuclear Cells Stimulated by Streptococcus pneumoniae. Infect. Immun. 71, 2087–2094. https://doi.org/10.1128/IAI.71.4.2087-2094.2003

Rolando, M., Buchrieser, C., 2014. Legionella pneumophila type IV effectors hijack the transcription and translation machinery of the host cell. Trends Cell Biol. 24, 771–778. https://doi.org/10.1016/j.tcb.2014.06.002

Rolando, M., Sanulli, S., Rusniok, C., Gomez-Valero, L., Bertholet, C., Sahr, T., Margueron, R., Buchrieser, C., 2013. Legionella pneumophila Effector RomA Uniquely Modifies Host Chromatin to Repress Gene Expression and Promote Intracellular Bacterial Replication. Cell Host Microbe 13, 395–405. https://doi.org/10.1016/j.chom.2013.03.004

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., Ali, S., Chin, S.-F., Palmieri, C., Caldas, C., Carroll, J.S., 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481, 389. https://doi.org/10.1038/nature10730

Ryseck, R.P., Macdonald-Bravo, H., Mattéi, M.G., Ruppert, S., Bravo, R., 1989. Structure, mapping and expression of a growth factor inducible gene encoding a putative nuclear hormonal binding receptor. EMBO J. 8, 3327–3335.

Saccani, S., Natoli, G., 2002. Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. Genes Dev. 16, 2219–2224. https://doi.org/10.1101/gad.232502

Saeed, S., Quintin, J., Kerstens, H.H.D., Rao, N.A., Aghajanirefah, A., Matarese, F., Cheng, S.-C., Ratter, J., Berentsen, K., van der Ent, M.A., Sharifi, N., Janssen-Megens, E.M., Ter Huurne, M., Mandoli, A., van Schaik, T., Ng, A., Burden, F., Downes, K., Frontini, M., Kumar, V., Giamarellos-Bourboulis, E.J., Ouwehand, W.H., van der Meer, J.W.M., Joosten, L.A.B.,

Wijmenga, C., Martens, J.H.A., Xavier, R.J., Logie, C., Netea, M.G., Stunnenberg, H.G., 2014. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science 345, 1251086. https://doi.org/10.1126/science.1251086

Safford, M., Collins, S., Lutz, M.A., Allen, A., Huang, C.-T., Kowalski, J., Blackford, A., Horton, M.R., Drake, C., Schwartz, R.H., Powell, J.D., 2005. Egr-2 and Egr-3 are negative regulators of T cell activation. Nat. Immunol. 6, 472. https://doi.org/10.1038/ni1193

Salerno, D.M., Hoffman, B., Liebermann, D.A., 2009. Gadd45 Null Mice and Neutrophils/Macrophages Derived From the Bone Marrow of These Mice Exhibit Impaired Innate-Immune/Inflammatory Reponses. Blood 114, 1350–1350.

Salerno, D.M., Tront, J.S., Hoffman, B., Liebermann, D.A., 2012. Gadd45a and Gadd45b modulate innate immune functions of granulocytes and macrophages by differential regulation of p38 and JNK signaling. J. Cell. Physiol. 227, 3613–3620. https://doi.org/10.1002/jcp.24067

Sasaki, K., Gotoh, K., Miake, S., Setoyama, D., Yagi, M., Igami, K., Uchiumi, T., Kang, D., 2017. p32 is Required for Appropriate Interleukin-6 Production Upon LPS Stimulation and Protects Mice from Endotoxin Shock. EBioMedicine 20, 161–172. https://doi.org/10.1016/j.ebiom.2017.05.018

Satoh, T., Takeuchi, O., Vandenbon, A., Yasuda, K., Tanaka, Y., Kumagai, Y., Miyake, T., Matsushita, K., Okazaki, T., Saitoh, T., Honma, K., Matsuyama, T., Yui, K., Tsujimura, T., Standley, D.M., Nakanishi, K., Nakai, K., Akira, S., 2010. The Jmjd3-Irf4 axis regulates M2 macrophage polarization and host responses against helminth infection. Nat. Immunol. 11, 936–944. https://doi.org/10.1038/ni.1920

Savitski, M.M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., Kuster, B., 2011. Confident phosphorylation site localization using the Mascot Delta Score. Mol. Cell. Proteomics MCP 10, M110.003830. https://doi.org/10.1074/mcp.M110.003830

Scheer, S., Zaph, C., 2017. The Lysine Methyltransferase G9a in Immune Cell Differentiation and Function. Front. Immunol. 8. https://doi.org/10.3389/fimmu.2017.00429

Schilling, B., Rardin, M.J., MacLean, B.X., Zawadzka, A.M., Frewen, B.E., Cusack, M.P., Sorensen, D.J., Bereman, M.S., Jing, E., Wu, C.C., Verdin, E., Kahn, C.R., MacCoss, M.J., Gibson, B.W., 2012. Platform-independent and Label-free Quantitation of Proteomic Data Using MS1 Extracted Ion Chromatograms in Skyline. Mol. Cell. Proteomics MCP 11, 202–214. https://doi.org/10.1074/mcp.M112.017707

Schroder, K., Hertzog, P.J., Ravasi, T., Hume, D.A., 2004. Interferon-γ: an overview of signals, mechanisms and functions. J. Leukoc. Biol. 75, 163–189. https://doi.org/10.1189/jlb.0603252

Schwartz, B.E., Ahmad, K., 2005. Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev. 19, 804–814. https://doi.org/10.1101/gad.1259805

Shao, W., Tang, J., Song, W., Wang, C., Li, Y., Wilson, C.M., Kaslow, R.A., 2007. *CCL3L1* and *CCL4L1*: variable gene copy number in adolescents with and without human immunodeficiency virus type 1 (HIV-1) infection. Genes Immun. 8, 224. https://doi.org/10.1038/sj.gene.6364378

Sharma, G.G., So, S., Gupta, A., Kumar, R., Cayrou, C., Avvakumov, N., Bhadra, U., Pandita, R.K., Porteus, M.H., Chen, D.J., Cote, J., Pandita, T.K., 2010. MOF and Histone H4 Acetylation at Lysine 16 Are Critical for DNA Damage Response and Double-Strand Break Repair. Mol. Cell. Biol. 30, 3582–3595. https://doi.org/10.1128/MCB.01476-09

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., Shi, Y., 2004. Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. Cell 119, 941–953. https://doi.org/10.1016/j.cell.2004.12.012

Shiraki, A., Kotooka, N., Komoda, H., Hirase, T., Oyama, J., Node, K., 2016. Pentraxin-3 regulates the inflammatory activity of macrophages. Biochem. Biophys. Rep. 5, 290–295. https://doi.org/10.1016/j.bbrep.2016.01.009

Sidoli, S., Fujiwara, R., Garcia, B.A., 2016. Multiplexed data independent acquisition (MSX-DIA) applied by high resolution mass spectrometry improves quantification quality for the analysis of histone peptides. PROTEOMICS 16, 2095–2105. https://doi.org/10.1002/pmic.201500527

Sidoli, S., Garcia, B.A., 2017. Middle-down proteomics: a still unexploited resource for chromatin biology. Expert Rev. Proteomics 14, 617–626. https://doi.org/10.1080/14789450.2017.1345632

Sidoli, S., Lin, S., Karch, K.R., Garcia, B.A., 2015a. Bottom-up and middle-down proteomics have comparable accuracies in defining histone post-translational modification relative abundance and stoichiometry. Anal. Chem. 87, 3129–3133. https://doi.org/10.1021/acs.analchem.5b00072

Sidoli, S., Lin, S., Xiong, L., Bhanu, N.V., Karch, K.R., Johansen, E., Hunter, C., Mollah, S., Garcia, B.A., 2015b. Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH) Analysis for Characterization and Quantification of Histone Post-translational Modifications. Mol. Cell. Proteomics MCP 14, 2420–2428. https://doi.org/10.1074/mcp.O114.046102

Sidoli, S., Lu, C., Coradin, M., Wang, X., Karch, K.R., Ruminowicz, C., Garcia, B.A., 2017. Metabolic labeling in middle-down proteomics allows for investigation of the dynamics of the histone code. Epigenetics Chromatin 10, 34. https://doi.org/10.1186/s13072-017-0139-z

Sidoli, S., Simithy, J., Karch, K.R., Kulej, K., Garcia, B.A., 2015c. Low Resolution Data-Independent Acquisition in an LTQ-Orbitrap Allows for Simplified and Fully Untargeted Analysis of Histone Modifications. Anal. Chem. 87, 11448–11454. https://doi.org/10.1021/acs.analchem.5b03009

Sims, J.K., Houston, S.I., Magazinnik, T., Rice, J.C., 2006. A Trans-tail Histone Code Defined by Monomethylated H4 Lys-20 and H3 Lys-9 Demarcates Distinct Regions of Silent Chromatin. J. Biol. Chem. 281, 12760–12766. https://doi.org/10.1074/jbc.M513462200

Singh, V., Prakhar, P., Rajmani, R.S., Mahadik, K., Borbora, S.M., Balaji, K.N., 2017. Histone Methyltransferase SET8 Epigenetically Reprograms Host Immune Responses to Assist Mycobacterial Survival. J. Infect. Dis. 216, 477–488. https://doi.org/10.1093/infdis/jix322

Sloan, K.E., Bohnsack, M.T., Watkins, N.J., 2013. The 5S RNP Couples p53 Homeostasis to Ribosome Biogenesis and Nucleolar Stress. Cell Rep. 5, 237–247. https://doi.org/10.1016/j.celrep.2013.08.049

Smith, Z.D., Meissner, A., 2013. DNA methylation: roles in mammalian development. Nat. Rev. Genet. 14, 204–220. https://doi.org/10.1038/nrg3354

Soares, A.C., Souza, D.G., Pinho, V., Vieira, A.T., Nicoli, J.R., Cunha, F.Q., Mantovani, A., Reis, L.F.L., Dias, A.A.M., Teixeira, M.M., 2006. Dual function of the long pentraxin PTX3 in resistance against pulmonary infection with Klebsiella pneumoniae in transgenic mice. Microbes Infect. 8, 1321–1329. https://doi.org/10.1016/j.micinf.2005.12.017

Soldi, M., Mari, T., Nicosia, L., Musiani, D., Sigismondo, G., Cuomo, A., Pavesi, G., Bonaldi, T., 2017. Chromatin proteomics reveals novel combinatorial histone modification signatures that mark distinct subpopulations of macrophage enhancers. Nucleic Acids Res. 45, 12195–12213. https://doi.org/10.1093/nar/gkx821

Soldi, M., Mari, T., Nicosia, L., Musiani, D., Sigismondo, G., Cuomo, A., Pavesi, G., Bonaldi, T., n.d. Chromatin proteomics reveals novel combinatorial histone modification signatures that mark distinct subpopulations of macrophage enhancers. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx821

Srivastava, A., Henneke, P., Visintin, A., Morse, S.C., Martin, V., Watkins, C., Paton, J.C., Wessels, M.R., Golenbock, D.T., Malley, R., 2005. The Apoptotic Response to Pneumolysin Is Toll-Like Receptor 4 Dependent and Protects against Pneumococcal Disease. Infect. Immun. 73, 6479–6487. https://doi.org/10.1128/IAI.73.10.6479-6487.2005

Stare, T., Stare, K., Weckwerth, W., Wienkoop, S., Gruden, K., 2017. Comparison between Proteome and Transcriptome Response in Potato (Solanum tuberosum L.) Leaves Following Potato Virus Y (PVY) Infection. Proteomes 5, 14. https://doi.org/10.3390/proteomes5030014

Starokadomskyy, P., Gluck, N., Li, H., Chen, B., Wallis, M., Maine, G.N., Mao, X., Zaidi, I.W., Hein, M.Y., McDonald, F.J., Lenzner, S., Zecha, A., Ropers, H.-H., Kuss, A.W., McGaughran, J., Gecz, J., Burstein, E., 2013. CCDC22 deficiency in humans blunts activation of proinflammatory NF-κB signaling. J. Clin. Invest. 123, 2244–2256. https://doi.org/10.1172/JCI66466

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.-E., Chen, J., Lazar, M.A., Blobel, G.A., Vakoc, C.R., 2008. DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. Mol. Cell. Biol. 28, 2825–2839. https://doi.org/10.1128/MCB.02076-07

Stender, J.D., Pascual, G., Liu, W., Kaikkonen, M.U., Do, K., Spann, N.J., Boutros, M., Perrimon, N., Rosenfeld, M.G., Glass, C.K., 2012. Control of Proinflammatory Gene Programs by Regulated Trimethylation and Demethylation of Histone H4K20. Mol. Cell 48, 28–38. https://doi.org/10.1016/j.molcel.2012.07.020

Struyf, S., Menten, P., Lenaerts, J.-P., Put, W., D'Haese, A., De Clercq, E., Schols, D., Proost, P., Van Damme, J., 2001. Diverging binding capacities of natural LD78β isoforms of macrophage inflammatory protein-1α to the CC chemokine receptors 1, 3 and 5 affect their anti-HIV-1 activity and chemotactic potencies for neutrophils and eosinophils. Eur. J. Immunol. 31, 2170–2178. https://doi.org/10.1002/1521-4141(200107)31:7<2170::AID-IMMU2170>3.0.CO;2-D

Szenker, E., Ray-Gallet, D., Almouzni, G., 2011. The double face of the histone variant H3.3. Cell Res. 21, 421–434. https://doi.org/10.1038/cr.2011.14
Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., Shinkai, Y., 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. 16, 1779–1791. https://doi.org/10.1101/gad.989402
Takekawa, M., Saito, H., 1998. A Family of Stress-Inducible GADD45-like Proteins Mediate Activation of the Stress-Responsive MTK1/MEKK4 MAPKKK. Cell 95, 521–530. https://doi.org/10.1016/S0092-8674(00)81619-0
Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D., Kitajima, S., 2007. Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene 397, 161–168. https://doi.org/10.1016/j.gene.2007.04.027
Taylor, G.C.A., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M.M., Bickmore, W.A., 2013. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. Genome Res. 23, 2053–2065. https://doi.org/10.1101/gr.155028.113
Tchatalbachev, S., Ghai, R., Hossain, H., Chakraborty, T., 2010. Gram-positive pathogenic bacteria induce a common early response in human monocytes. BMC Microbiol. 10, 275. https://doi.org/10.1186/1471-2180-10-275
Thimmulappa, R.K., Lee, H., Rangasamy, T., Reddy, S.P., Yamamoto, M., Kensler, T.W., Biswal, S., 2016. Nrf2 is a critical regulator of the innate immune response and survival during experimental sepsis. J. Clin. Invest. 116, 984–995. https://doi.org/10.1172/JCI25790
Tigno-Aranjuez, J.T., Asara, J.M., Abbott, D.W., 2010. Inhibition of RIP2's tyrosine kinase activity limits NOD2-driven cytokine responses. Genes Dev. 24, 2666–2677. https://doi.org/10.1101/gad.1964410
Tourtellotte, W.G., Milbrandt, J., 1998. Sensory ataxia and muscle spindle agenesis in mice lacking the transcription factor Egr3. Nat. Genet. 20, 87. https://doi.org/10.1038/1757
Tsai, W.-W., Wang, Z., Yiu, T.T., Akdemir, K.C., Xia, W., Winter, S., Tsai, C.-Y., Shi, X., Schwarzer, D., Plunkett, W., Aronow, B., Gozani, O., Fischle, W., Hung, M.-C., Patel, D.J., Barton, M.C., 2010. TRIM24 links a non-canonical histone signature to breast cancer. Nature 468, 927–932. https://doi.org/10.1038/nature09542
Tvardovskiy, A., Schwämmle, V., Kempf, S.J., Rogowska-Wrzesinska, A., Jensen, O.N., 2017. Accumulation of histone variant H3.3 with age is associated with profound changes in the histone methylation landscape. Nucleic Acids Res. 45, 9272–9289. https://doi.org/10.1093/nar/gkx696
Vakoc, C.R., Sachdeva, M.M., Wang, H., Blobel, G.A., 2006. Profile of Histone Lysine Methylation across Transcribed Mammalian Chromatin. Mol. Cell. Biol. 26, 9185–9195. https://doi.org/10.1128/MCB.01529-06
van Kuilenburg, A.B.P., Meinsma, R., 2016. The pivotal role of uridine-cytidine kinases in pyrimidine metabolism and activation of cytotoxic nucleoside analogues in neuroblastoma. Biochim. Biophys. Acta 1862, 1504–1512. https://doi.org/10.1016/j.bbadis.2016.05.012
Vaquero, A., Scher, M., Lee, D., Erdjument-Bromage, H., Tempst, P., Reinberg, D., 2004. Human SirT1 Interacts with Histone H1 and Promotes Formation

of Facultative Heterochromatin. Mol. Cell 16, 93–105. https://doi.org/10.1016/j.molcel.2004.08.031

Verschoor, C.P., Dorrington, M.G., Novakowski, K.E., Kaiser, J., Radford, K., Nair, P., Anipindi, V., Kaushic, C., Surette, M.G., Bowdish, D.M.E., 2014. MicroRNA-155 Is Required for Clearance of Streptococcus pneumoniae from the Nasopharynx. Infect. Immun. 82, 4824–4833. https://doi.org/10.1128/IAI.02251-14

Waggoner, S.N., Cruise, M.W., Kassel, R., Hahn, Y.S., 2005. gC1q Receptor Ligation Selectively Down-Regulates Human IL-12 Production through Activation of the Phosphoinositide 3-Kinase Pathway. J. Immunol. 175, 4706–4714. https://doi.org/10.4049/jimmunol.175.7.4706

Wallner, S., Schröder, C., Leitão, E., Berulava, T., Haak, C., Beißer, D., Rahmann, S., Richter, A.S., Manke, T., Bönisch, U., Arrigoni, L., Fröhler, S., Klironomos, F., Chen, W., Rajewsky, N., Müller, F., Ebert, P., Lengauer, T., Barann, M., Rosenstiel, P., Gasparoni, G., Nordström, K., Walter, J., Brors, B., Zipprich, G., Felder, B., Klein-Hitpass, L., Attenberger, C., Schmitz, G., Horsthemke, B., 2016. Epigenetic dynamics of monocyte-to-macrophage differentiation. Epigenetics Chromatin 9. https://doi.org/10.1186/s13072-016-0079-z

Wang, L., Zhang, F., Rode, S., Chin, K.K., Ko, E.E., Kim, J., Iyer, V.R., Qiao, H., 2017. Ethylene induces combinatorial effects of histone H3 acetylation in gene expression in Arabidopsis. BMC Genomics 18, 538. https://doi.org/10.1186/s12864-017-3929-6

Wang, W., Sidoli, S., Zhang, W., Wang, Q., Wang, L., Jensen, O.N., Guo, L., Zhao, X., Zheng, L., 2017. Abnormal levels of histone methylation in the retinas of diabetic rats are reversed by minocycline treatment. Sci. Rep. 7. https://doi.org/10.1038/srep45103

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., Zhao, K., 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. 40, 897–903. https://doi.org/10.1038/ng.154

Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., Feinberg, A.P., 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat. Genet. 41, ng.297. https://doi.org/10.1038/ng.297

Wheeler, D.S., Lahni, P.M., Denenberg, A.G., Poynter, S.E., Wong, H.R., Cook, J.A., Zingarelli, B., 2008. Induction of endotoxin tolerance enhances bacterial clearance and survival in murine polymicrobial sepsis. Shock Augusta Ga 30, 267–273. https://doi.org/10.1097/shk.0b013e318162c190

Wieland, G.D., Nehmann, N., Müller, D., Eibel, H., Siebenlist, U., Sühnel, J., Zipfel, P.F., Skerka, C., 2005. Early growth response proteins EGR-4 and EGR-3 interact with immune inflammatory mediators NF-κB p50 and p65. J. Cell Sci. 118, 3203–3212. https://doi.org/10.1242/jcs.02445

Will, C.L., Lührmann, R., 2011. Spliceosome Structure and Function. Cold Spring Harb. Perspect. Biol. 3. https://doi.org/10.1101/cshperspect.a003707

Will, C.L., Urlaub, H., Achsel, T., Gentzel, M., Wilm, M., Lührmann, R., 2002. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. EMBO J. 21, 4978–4988. https://doi.org/10.1093/emboj/cdf480

Williams, S.J., Munster, D.J., Quin, R.J., Gotley, D.C., McGuckin, M.A., 1999. The MUC3 gene encodes a transmembrane mucin and is alternatively spliced. Biochem. Biophys. Res. Commun. 261, 83–89. https://doi.org/10.1006/bbrc.1999.1001

Wilson, B.G., Roberts, C.W.M., 2011. SWI/SNF nucleosome remodellers and cancer. Nat. Rev. Cancer 11, 481–492. https://doi.org/10.1038/nrc3068

Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., 2009. Universal sample preparation method for proteome analysis. Nat. Methods 6, 359–362. https://doi.org/10.1038/nmeth.1322

Witzenrath, M., Pache, F., Lorenz, D., Koppe, U., Gutbier, B., Tabeling, C., Reppe, K., Meixenberger, K., Dorhoi, A., Ma, J., Holmes, A., Trendelenburg, G., Heimesaat, M.M., Bereswill, S., van der Linden, M., Tschopp, J., Mitchell, T.J., Suttorp, N., Opitz, B., 2011. The NLRP3 inflammasome is differentially activated by pneumolysin variants and contributes to host defense in pneumococcal pneumonia. J. Immunol. Baltim. Md 1950 187, 434–440. https://doi.org/10.4049/jimmunol.1003143

Wolpe, S.D., Davatelis, G., Sherry, B., Beutler, B., Hesse, D.G., Nguyen, H.T., Moldawer, L.L., Nathan, C.F., Lowry, S.F., Cerami, A., 1988. Macrophages secrete a novel heparin-binding protein with inflammatory and neutrophil chemokinetic properties. J. Exp. Med. 167, 570–581. https://doi.org/10.1084/jem.167.2.570

Xiao, K., Wang, Y., Chang, Z., Lao, Y., Chang, D.C., 2014. p32, a novel binding partner of Mcl-1, positively regulates mitochondrial Ca2+ uptake and apoptosis. Biochem. Biophys. Res. Commun. 451, 322–328. https://doi.org/10.1016/j.bbrc.2014.07.122

Xiong, T., Meister, G.E., Workman, R.E., Kato, N.C., Spellberg, M.J., Turker, F., Timp, W., Ostermeier, M., Novina, C.D., 2017. Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-06757-0

Xu, M., Long, C., Chen, X., Huang, C., Chen, S., Zhu, B., 2010. Partitioning of Histone H3-H4 Tetramers During DNA Replication–Dependent Chromatin Assembly. Science 328, 94–98. https://doi.org/10.1126/science.1178994

Xu, M., Wang, W., Chen, S., Zhu, B., 2011. A model for mitotic inheritance of histone lysine methylation. EMBO Rep. 13, 60–67. https://doi.org/10.1038/embor.2011.206

Yadav, G., Prasad, R.L.A., Jha, B.K., Rai, V., Bhakuni, V., Datta, K., 2009. Evidence for Inhibitory Interaction of Hyaluronan-binding Protein 1 (HABP1/p32/gC1qR) with Streptococcus pneumoniae Hyaluronidase. J. Biol. Chem. 284, 3897–3905. https://doi.org/10.1074/jbc.M804246200

Yáñez, A., Hassanzadeh-Kiabi, N., Ng, M.Y., Megías, J., Subramanian, A., Liu, G.Y., Underhill, D.M., Gil, M.L., Goodridge, H.S., 2013. Detection of a TLR2 agonist by hematopoietic stem and progenitor cells impacts the function of the macrophages they produce. Eur. J. Immunol. 43, 2114–2125. https://doi.org/10.1002/eji.201343403

Yao, Y., Zheng, X., Ge, X., Xiu, Y., Zhang, L., Fang, W., Zhao, J., Gu, F., Zhu, Y., 2017. Identification of a novel GJA3 mutation in a large Chinese family with congenital cataract using targeted exome sequencing. PloS One 12, e0184440. https://doi.org/10.1371/journal.pone.0184440

Ye, M., Zhang, H., Yang, H., Koche, R., Staber, P.B., Cusan, M., Levantini, E., Welner, R.S., Bach, C.S., Zhang, J., Krivtsov, A.V., Armstrong, S.A.,

Tenen, D.G., 2015. Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia. Cell Stem Cell 17, 611–623. https://doi.org/10.1016/j.stem.2015.08.011

Yona, S., Kim, K.-W., Wolf, Y., Mildner, A., Varol, D., Breker, M., Strauss-Ayali, D., Viukov, S., Guilliams, M., Misharin, A., Hume, D.A., Perlman, H., Malissen, B., Zelzer, E., Jung, S., 2013. Fate Mapping Reveals Origins and Dynamics of Monocytes and Tissue Macrophages under Homeostasis. Immunity 38, 79–91. https://doi.org/10.1016/j.immuni.2012.12.001

Yoshida, K., Maekawa, T., Zhu, Y., Renard-Guillet, C., Chatton, B., Inoue, K., Uchiyama, T., Ishibashi, K., Yamada, T., Ohno, N., Shirahige, K., Okada-Hatakeyama, M., Ishii, S., 2015a. The transcription factor ATF7 mediates lipopolysaccharide-induced epigenetic changes in macrophages involved in innate immunological memory. Nat. Immunol. advance online publication. https://doi.org/10.1038/ni.3257

Yoshida, K., Maekawa, T., Zhu, Y., Renard-Guillet, C., Chatton, B., Inoue, K., Uchiyama, T., Ishibashi, K., Yamada, T., Ohno, N., Shirahige, K., Okada-Hatakeyama, M., Ishii, S., 2015b. The transcription factor ATF7 mediates lipopolysaccharide-induced epigenetic changes in macrophages involved in innate immunological memory. Nat. Immunol. 16, 1034–1043. https://doi.org/10.1038/ni.3257

Young, N.L., DiMaggio, P.A., Garcia, B.A., 2010. The significance, development and progress of high-throughput combinatorial histone code analysis. Cell. Mol. Life Sci. 67, 3983–4000. https://doi.org/10.1007/s00018-010-0475-7

Yu, G., Wang, L.-G., He, Q.-Y., 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 31, 2382–2383. https://doi.org/10.1093/bioinformatics/btv145

Yu, S., Lemos, B., 2016. A Portrait of Ribosomal DNA Contacts with Hi-C Reveals 5S and 45S rDNA Anchoring Points in the Folded Human Genome. Genome Biol. Evol. 8, 3545–3558. https://doi.org/10.1093/gbe/evw257

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., Zhu, B., 2011. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J. Biol. Chem. 286, 7983–7989. https://doi.org/10.1074/jbc.M110.194027

Yuan, Z.-F., Lin, S., Molden, R.C., Cao, X.-J., Bhanu, N.V., Wang, X., Sidoli, S., Liu, S., Garcia, B.A., 2015. EpiProfile Quantifies Histone Peptides With Modifications by Extracting Retention Time and Intensity in High-resolution Mass Spectra. Mol. Cell. Proteomics MCP 14, 1696–1707. https://doi.org/10.1074/mcp.M114.046011

Yuan, Z.-F., Lin, S., Molden, R.C., Garcia, B.A., 2014. Evaluation of Proteomic Search Engines for the Analysis of Histone Modifications. J. Proteome Res. 13, 4470–4478. https://doi.org/10.1021/pr5008015

Zafar, M.A., Wang, Y., Hamaguchi, S., Weiser, J.N., 2017. Host-to-Host Transmission of Streptococcus pneumoniae Is Driven by Its Inflammatory Toxin, Pneumolysin. Cell Host Microbe 21, 73–83. https://doi.org/10.1016/j.chom.2016.12.005

Zahlten, J., Kim, Y.-J., Doehn, J.-M., Pribyl, T., Hocke, A.C., García, P., Hammerschmidt, S., Suttorp, N., Hippenstiel, S., Hübner, R.-H., 2015. Streptococcus pneumoniae–Induced Oxidative Stress in Lung Epithelial Cells Depends on Pneumococcal Autolysis and Is Reversible by

Resveratrol. J. Infect. Dis. 211, 1822–1830. https://doi.org/10.1093/infdis/jiu806
Zehavi, Y., Mandel, H., Zehavi, A., Rashid, M.A., Straussberg, R., Jabur, B., Shaag, A., Elpeleg, O., Spiegel, R., 2017. De novo GRIN1 mutations: An emerging cause of severe early infantile encephalopathy. Eur. J. Med. Genet. 60, 317–320. https://doi.org/10.1016/j.ejmg.2017.04.001
Zerbini, L.F., Wang, Y., Czibere, A., Correa, R.G., Cho, J.-Y., Ijiri, K., Wei, W., Joseph, M., Gu, X., Grall, F., Goldring, M.B., Zhou, J.-R., Libermann, T.A., 2004. NF-κB-mediated repression of growth arrest- and DNA-damage-inducible proteins 45α and γ is essential for cancer cell survival. Proc. Natl. Acad. Sci. U. S. A. 101, 13618–13623. https://doi.org/10.1073/pnas.0402069101
Zeybel, M., Hardy, T., Wong, Y.K., Mathers, J.C., Fox, C.R., Gackowska, A., Oakley, F., Burt, A.D., Wilson, C.L., Anstee, Q.M., Barter, M.J., Masson, S., Elsharkawy, A.M., Mann, D.A., Mann, J., 2012. Multigenerational Epigenetic Adaptation of the Hepatic Wound-Healing Response. Nat. Med. 18, 1369–1377. https://doi.org/10.1038/nm.2893
Zhan, Y., Chen, Y., Zhang, Q., Zhuang, J., Tian, M., Chen, H., Zhang, L., Zhang, H., He, J., Wang, W., Wu, R., Wang, Y., Shi, C., Yang, K., Li, A., Xin, Y., Li, T.Y., Yang, J.Y., Zheng, Z., Yu, C., Lin, S.-C., Chang, C., Huang, P., Lin, T., Wu, Q., 2012. The orphan nuclear receptor Nur77 regulates LKB1 localization and activates AMPK. Nat. Chem. Biol. 8, 897. https://doi.org/10.1038/nchembio.1069
Zhang, K., Siino, J.S., Jones, P.R., Yau, P.M., Bradbury, E.M., 2004. A mass spectrometric “Western blot” to evaluate the correlations between histone methylation and histone acetylation. PROTEOMICS 4, 3765–3775. https://doi.org/10.1002/pmic.200400819
Zhang, X.-W., Wang, X.-F., Ni, S.-J., Qin, W., Zhao, L.-Q., Hua, R.-X., Lu, Y.-W., Li, J., Dimri, G.P., Guo, W.-J., 2015. UBTD1 induces cellular senescence through an UBTD1–Mdm2/p53 positive feedback loop. J. Pathol. 235, 656–667. https://doi.org/10.1002/path.4478
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137. https://doi.org/10.1186/gb-2008-9-9-r137
Zhang, Z., Song, L., Maurer, K., Bagashev, A., Sullivan, K., 2011. Monocyte polarization: the relationship of genome-wide changes in H4 acetylation with polarization. Genes Immun. 12, 445–456. https://doi.org/10.1038/gene.2011.17
Zhao, H., Anand, A.R., Ganju, R.K., 2014. Slit2-Robo4 pathway modulates LPS-induced endothelial inflammation and its expression is dysregulated during endotoxemia. J. Immunol. Baltim. Md 1950 192, 385–393. https://doi.org/10.4049/jimmunol.1302021
Zhao, H., Eguchi, S., Alam, A., Ma, D., 2017. The role of nuclear factor-erythroid 2 related factor 2 (Nrf-2) in the protection against lung injury. Am. J. Physiol. - Lung Cell. Mol. Physiol. 312, L155–L162. https://doi.org/10.1152/ajplung.00449.2016
Zhao, Q., Kim, T., Pang, J., Sun, W., Yang, X., Wang, J., Song, Y., Zhang, H., Sun, H., Rangan, V., Deshpande, S., Tang, H., Cvijic, M.E., Westhouse, R., Olah, T., Xie, J., Struthers, M., Salter-Cid, L., 2017. A novel function of

CXCL10 in mediating monocyte production of proinflammatory cytokines. J. Leukoc. Biol. 102, 1271–1280. https://doi.org/10.1189/jlb.5A0717-302

Zhao, T., Li, Y., Bronson, R.T., Liu, B., Velmahos, G.C., Alam, H.B., 2014a. Selective histone deacetylase-6 inhibition attenuates stress responses and prevents immune organ atrophy in a lethal septic model. Surgery 156, 235–242. https://doi.org/10.1016/j.surg.2014.03.033

Zhao, T., Li, Y., Liu, B., Halaweish, I., Mazitschek, R., Alam, H.B., 2014b. Selective inhibition of histone deacetylase 6 alters the composition of circulating blood cells in a lethal septic model. J. Surg. Res. 190, 647–654. https://doi.org/10.1016/j.jss.2014.01.056

Zhao, Y., Alonso, C., Ballester, I., Song, J.H., Chang, S.Y., Guleng, B., Arihiro, S., Murray, P.J., Xavier, R., Kobayashi, K.S., Reinecker, H.-C., 2012. Control of NOD2 and Rip2-dependent innate immune activation by GEF-H1. Inflamm. Bowel Dis. 18, 603–612. https://doi.org/10.1002/ibd.21851

Zheng, Y., Huang, X., Kelleher, N.L., 2016. Epiproteomics: quantitative analysis of histone marks and codes by mass spectrometry. Curr. Opin. Chem. Biol., Chemical genetics and epigenetics * Molecular imaging 33, 142–150. https://doi.org/10.1016/j.cbpa.2016.06.007

Zhong, J., Ji, L., Chen, H., Li, X., Zhang, J. 'an, Wang, X., Wu, W., Xu, Y., Huang, F., Cai, W., Sun, Z.S., 2017. Acetylation of hMOF Modulates H4K16ac to Regulate DNA Repair Genes in Response to Oxidative Stress. Int. J. Biol. Sci. 13, 923–934. https://doi.org/10.7150/ijbs.17260

Zhu, Z., Wang, Y., Li, X., Wang, Y., Xu, L., Wang, X., Sun, T., Dong, X., Chen, L., Mao, H., Yu, Y., Li, J., Chen, P.A., Chen, C.D., 2010. PHF8 is a histone H3K9me2 demethylase regulating rRNA synthesis. Cell Res. 20, 794–801. https://doi.org/10.1038/cr.2010.75

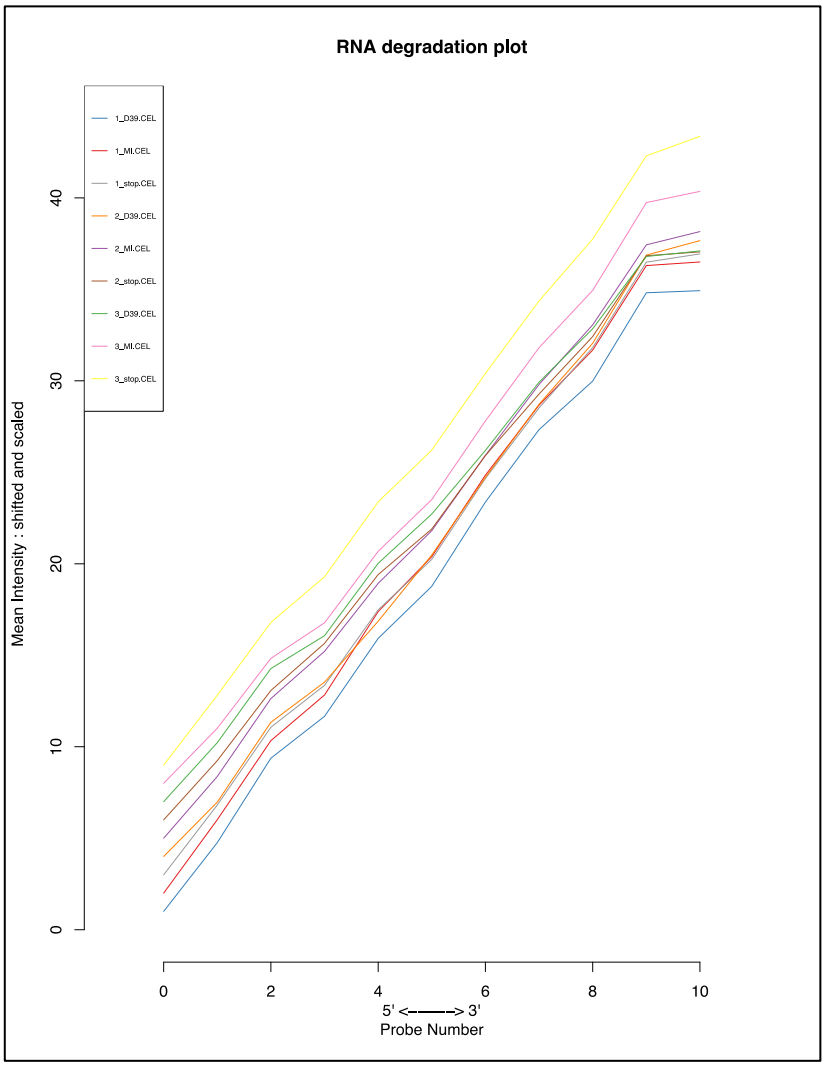
Chapter 8: Appendices

Appendix table 1.1: Data independent isolation list.

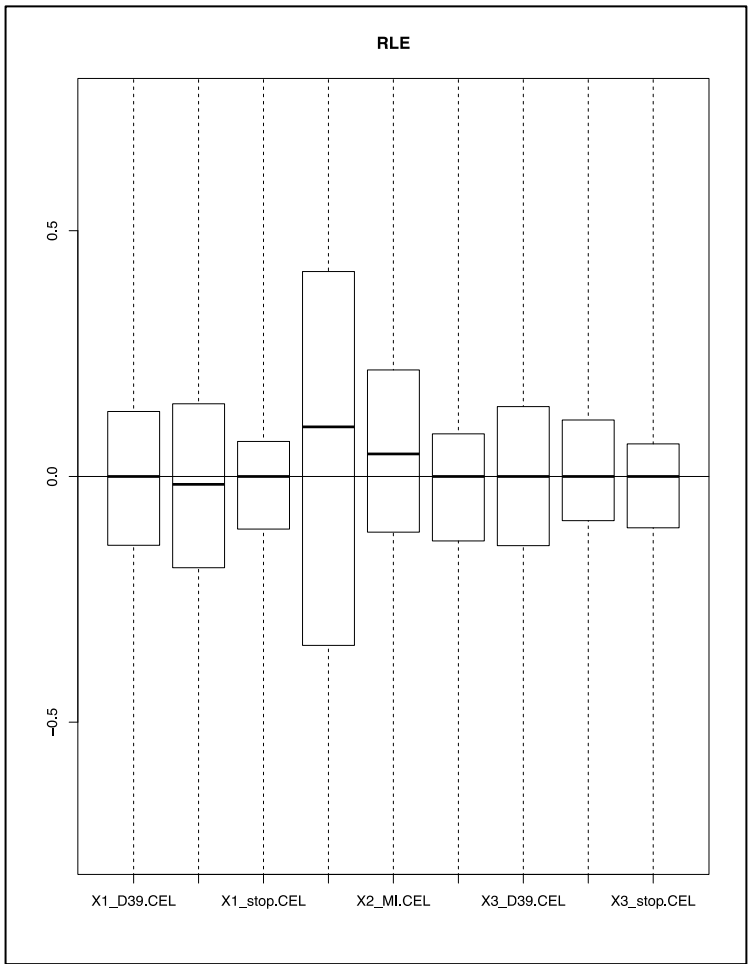
DIA60/30 isolation list	DIAvw isolation list		
310.390973	292	580	730
330.400068	359	585	735
350.409163	378	590	740
370.418258	397	595	745
390.427353	416	600	750
410.436448	432.5	605	755
430.445543	446.5	610	760
450.454638	458	615	765
470.463733	465	620	770
490.472828	470	625	775
510.481923	475	630	780
530.491018	480	635	785
550.500113	485	640	790
570.509208	490	645	795
590.518303	495	650	800
610.527398	500	655	805
630.536493	505	660	812
650.545588	510	665	823.5
670.554683	515	670	840
690.563778	522.5	675	859
710.572873	530	680	880.5
730.581968	535	685	904.5
750.591063	540	690	928.5
770.600158	545	695	965
790.609253	550	700	1014.5
810.618348	555	705	
830.627443	560	710	
850.636538	565	715	
870.645633	570	720	
890.654728	575	725	

Appendix table 2.1: Contaminants removed prior to analysis.

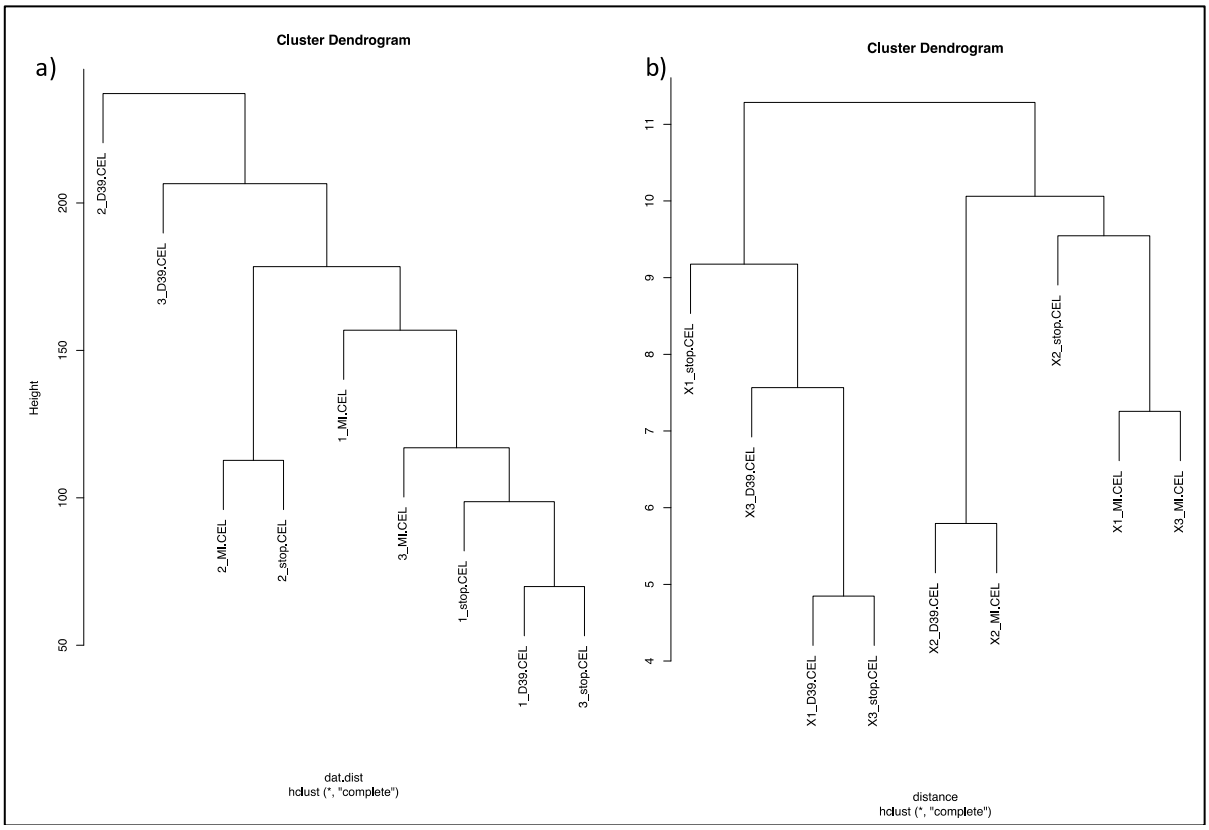
Contaminant	**Sequence**
solexa-forward	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
truseq-forward-contam	AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC
truseq-reverse-contam	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTA
nextera-forward-read-contam	CTGTCTCTTATACACATCTCCGAGCCCACGAGAC
multiplexing-forward	GATCGGAAGAGCACACGTCT
nextera-reverse-read-contam	CTGTCTCTTATACACATCTGACGCTGCCGACGA
solexa-reverse	AGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG



Appendix figure 4.1.1: Microarray quality control: RNA degradation plot.



Appendix figure 4.1.2: Microarray quality control: Relative log expression (RLE).



Appendix figure 4.1.3: Microarray quality control: Cluster dendrogram before and after normalisation.

Appendix table 4.1: Probes excluded from downstream analysis.

Probe ID	pt2...pt1	pt3...pt1	pt2...pt3	AveExpr	P.Value	adj.P.Val
215492_x_at	1.03	0.00	1.03	3.52	1.6E-07	0.004
220558_x_at	1.53	0.14	1.39	4.41	2.3E-07	0.004
209728_at	-3.71	0.12	-3.83	11.25	7.6E-07	0.010
205001_s_at	-5.10	0.63	-5.73	6.67	1.8E-06	0.017
221291_at	2.57	-0.07	2.65	3.44	2.5E-06	0.020
230846_at	2.70	0.63	2.06	5.68	3.9E-06	0.026
219759_at	6.31	0.45	5.86	7.09	5.6E-06	0.032
201909_at	-7.77	0.22	-7.99	8.94	6.4E-06	0.032
223062_s_at	4.21	0.80	3.42	7.95	8.0E-06	0.033
1558199_at	3.64	-0.31	3.94	4.90	8.4E-06	0.033
201069_at	-4.77	0.36	-5.13	7.01	1.0E-05	0.034
239824_s_at	0.70	-0.59	1.29	8.52	1.1E-05	0.034
212532_s_at	-0.81	-0.15	-0.67	7.95	1.2E-05	0.034
201196_s_at	-0.94	0.39	-1.33	9.09	1.2E-05	0.034
230360_at	5.94	0.01	5.93	4.45	1.3E-05	0.034
227462_at	5.03	0.47	4.56	8.48	1.7E-05	0.040
203475_at	7.15	0.35	6.80	6.54	1.8E-05	0.040
223423_at	-3.29	-0.39	-2.90	6.70	1.9E-05	0.040

229862_x_at	0.41	-0.01	0.42	5.38	2.2E-05	0.040
204916_at	2.63	0.01	2.62	6.78	2.4E-05	0.040
226836_at	0.78	-0.12	0.90	3.91	2.5E-05	0.040
1553551_s_at	0.91	-0.04	0.95	13.56	2.6E-05	0.040
228160_at	-2.73	-0.33	-2.39	5.60	2.6E-05	0.040
223640_at	1.52	-0.33	1.85	11.31	2.9E-05	0.040
209147_s_at	1.50	-0.11	1.61	5.87	2.9E-05	0.040
202777_at	-0.65	-0.12	-0.53	9.98	3.0E-05	0.040
204531_s_at	1.31	0.03	1.28	5.48	3.0E-05	0.040
214701_s_at	1.67	-0.02	1.70	3.04	3.0E-05	0.040
212062_at	3.23	-0.24	3.47	5.83	3.0E-05	0.040
203433_at	1.12	-0.16	1.28	7.83	3.0E-05	0.040
204443_at	-3.05	-0.18	-2.87	4.65	3.3E-05	0.042
215333_x_at	0.78	0.00	0.78	2.48	3.4E-05	0.042
228353_x_at	2.19	0.68	1.50	6.70	3.6E-05	0.042
38487_at	-6.00	-1.10	-4.90	8.35	4.2E-05	0.049

Appendix table 4.2: Differentially expressed probes.

Pneumolysin dependent (503)		Common (506)		Pneumolysin independent (234)
226331_at	218502_s_at	236140_at	216248_s_at	205476_at
222756_s_at	227354_at	1552553_a_at	243366_s_at	232797_at
222805_at	206723_s_at	228188_at	204622_x_at	204470_at
233255_s_at	229383_at	222173_s_at	1552315_at	237252_at
221861_at	1554539_a_at	234979_at	204621_s_at	207850_at
219349_s_at	235306_at	209340_at	225827_at	209098_s_at
214228_x_at	227533_at	223446_s_at	220005_at	1553449_at
212234_at	244519_at	207840_at	244804_at	240231_at
204552_at	219574_at	235094_at	242329_at	239876_at
227446_s_at	211676_s_at	218617_at	209959_at	223218_s_at
211715_s_at	223620_at	212747_at	203665_at	228258_at
226016_at	222146_s_at	202266_at	219210_s_at	221756_at
201447_at	244008_at	219166_at	212086_x_at	204103_at
222201_s_at	219581_at	220731_s_at	202284_s_at	204182_s_at
224978_s_at	204669_s_at	207038_at	228825_at	219540_at
206278_at	207978_s_at	202643_s_at	209684_at	226893_at
203253_s_at	202727_s_at	204440_at	203297_s_at	212724_at
32723_at	1557432_at	213524_s_at	209446_s_at	205114_s_at
218862_at	223470_at	225539_at	1556385_at	227140_at
242062_at	236802_at	214059_at	203810_at	210511_s_at
224185_at	220987_s_at	201464_x_at	244341_at	210118_s_at
212704_at	37170_at	1554334_a_at	228846_at	228496_s_at
225632_s_at	230252_at	201881_s_at	217996_at	213327_s_at

AFFX-M27830_M_at	203765_at	205599_at	232744_x_at	230526_at
227678_at	226479_at	203473_at	219239_s_at	203887_s_at
242943_at	228075_x_at	203241_at	243179_at	243463_s_at
228189_at	239901_at	202114_at	231199_at	39402_at
226080_at	228702_at	201845_s_at	226650_at	223217_s_at
219885_at	204861_s_at	1556744_a_at	1555847_a_at	243338_at
208875_s_at	213056_at	226507_at	226041_at	208811_s_at
236381_s_at	236782_at	203785_s_at	203811_s_at	243664_at
227112_at	203554_x_at	230380_at	236583_at	213353_at
203690_at	226529_at	203246_s_at	229367_s_at	240108_at
225154_at	203980_at	208760_at	219243_at	211538_s_at
222716_s_at	227062_at	217998_at	221213_s_at	226906_s_at
224788_at	228922_at	202014_at	236404_at	203888_at
213189_at	204015_s_at	219016_at	223394_at	227200_at
225022_at	242125_at	241472_at	38290_at	226574_at
223741_s_at	206074_s_at	202644_s_at	205193_at	209774_x_at
225455_at	201964_at	1565544_at	225171_at	215743_at
209705_at	210785_s_at	223013_at	232431_at	205067_at
232436_at	204526_s_at	218472_s_at	230298_at	201369_s_at
223493_at	207826_s_at	226541_at	203879_at	214056_at
228594_at	242919_at	203344_s_at	1555860_x_at	219165_at
226068_at	225767_at	224699_s_at	205990_s_at	223289_s_at
228968_at	36920_at	243046_at	228964_at	226534_at
235529_x_at	235385_at	224864_at	201844_s_at	221757_at
228670_at	223961_s_at	219911_s_at	226423_at	208296_x_at
226030_at	228220_at	226219_at	210786_s_at	226630_at
229983_at	202983_at	227182_at	238738_at	212365_at
239804_at	219317_at	202329_at	203298_s_at	206173_x_at
222586_s_at	224622_at	229899_s_at	201846_s_at	235407_at
229333_at	219925_at	206332_s_at	1554015_a_at	217371_s_at
208708_x_at	1560145_at	229574_at	228066_at	219161_s_at
230332_at	204112_s_at	225567_at	209457_at	229795_at
1554057_at	205885_s_at	223412_at	1552316_a_at	202085_at
218218_at	203753_at	204159_at	204236_at	202385_s_at
222392_x_at	227749_at	232012_at	239331_at	1555841_at
225497_at	214467_at	202113_s_at	243296_at	221688_s_at
213271_s_at	244462_at	244840_x_at	208707_at	201782_s_at
229560_at	229431_at	204181_s_at	235242_at	223553_s_at
229317_at	230099_at	218696_at	204472_at	210044_s_at
41329_at	205251_at	1555274_a_at	207630_s_at	215224_at
202224_at	1552634_a_at	219677_at	208093_s_at	214100_x_at
213218_at	230329_s_at	228799_at	236401_at	239001_at

204014_at	242143_at	232369_at	201266_at	1553987_at
220631_at	214157_at	213281_at	217999_s_at	207616_s_at
212798_s_at	226659_at	225040_s_at	209200_at	1568954_s_at
228284_at	222663_at	227991_x_at	202887_s_at	207533_at
213352_at	244650_at	225283_at	204180_s_at	232543_x_at
212665_at	201751_at	220934_s_at	209681_at	229355_at
222762_x_at	203048_s_at	1556007_s_at	235765_at	226952_at
224581_s_at	227344_at	205449_at	36711_at	225443_at
225910_at	235470_at	201448_at	240154_at	203099_s_at
213016_at	212723_at	201986_at	213700_s_at	213872_at
223219_s_at	227220_at	41387_r_at	227402_s_at	212902_at
230178_s_at	208860_s_at	202779_s_at	214061_at	225136_at
213572_s_at	225290_at	230941_at	1562056_at	224250_s_at
228190_at	227184_at	208686_s_at	243496_at	226853_at
222849_s_at	218377_s_at	235423_at	204094_s_at	31845_at
219861_at	206028_s_at	219966_x_at	226542_at	208648_at
209004_s_at	229891_x_at	219008_at	205407_at	207438_s_at
208290_s_at	228588_s_at	226820_at	217997_at	225702_at
228817_at	227295_at	213425_at	242268_at	205126_at
225840_at	225954_s_at	55872_at	221044_s_at	209239_at
218179_s_at	1555910_at	203179_at	232311_at	206026_s_at
219467_at	234295_at	1553693_s_at	229762_at	226189_at
225246_at	235071_at	213923_at	228869_at	220346_at
222369_at	225992_at	227313_at	223980_s_at	227066_at
204912_at	223178_s_at	219458_s_at	215498_s_at	223485_at
228041_at	208869_s_at	230748_at	226743_at	219397_at
229018_at	228324_at	205992_s_at	203128_at	228328_at
235206_at	241908_at	232330_at	214508_x_at	202460_s_at
216202_s_at	210962_s_at	226762_at	242403_at	229398_at
218352_at	203596_s_at	241824_at	230511_at	200733_s_at
223243_s_at	218842_at	234986_at	215933_s_at	210190_at
59644_at	226855_at	241769_at	219694_at	228590_at
228837_at	222443_s_at	202147_s_at	219777_at	202622_s_at
226031_at	201585_s_at	204178_s_at	213677_s_at	223288_at
203411_s_at	202530_at	235716_at	231234_at	1552277_a_at
1569003_at	209995_s_at	225955_at	216834_at	239835_at
1552733_at	235890_at	205419_at	219315_s_at	244826_at
211763_s_at	212905_at	244511_at	233952_s_at	222395_s_at
1555950_a_at	1558511_s_at	209928_s_at	240859_at	221064_s_at
226644_at	219169_s_at	213134_x_at	227385_at	237444_at
218323_at	232014_at	208966_x_at	226275_at	202551_s_at
202281_at	212387_at	202911_at	202859_x_at	202414_at

225435_at	218700_s_at	223206_s_at	242853_at	241371_at
211458_s_at	226157_at	218300_at	219481_at	204512_at
228772_at	1555920_at	204286_s_at	222450_at	207181_s_at
212331_at	230091_at	232835_at	223275_at	225267_at
225824_at	207571_x_at	201878_at	239486_at	206025_s_at
228392_at	225163_at	218310_at	226180_at	205038_at
238041_at	229670_at	202334_s_at	222605_at	203743_s_at
203376_at	208644_at	225755_at	219625_s_at	225061_at
230836_at	221965_at	201431_s_at	40255_at	227630_at
235698_at	244397_at	219475_at	203925_at	223144_s_at
218370_s_at	210719_s_at	225864_at	223028_s_at	214911_s_at
212557_at	203404_at	212089_at	216061_x_at	232652_x_at
204928_s_at	226221_at	209782_s_at	227626_at	217813_s_at
218449_at	213294_at	205220_at	229373_at	205122_at
228760_at	232134_at	231183_s_at	225120_at	217985_s_at
204787_at	218432_at	242058_at	222449_at	219386_s_at
219826_at	213891_s_at	228812_at	225842_at	201146_at
238013_at	219174_at	227008_at	225432_s_at	206976_s_at
210028_s_at	206861_s_at	209357_at	209967_s_at	211297_s_at
229394_s_at	204957_at	232287_at	1554014_at	221058_s_at
201453_x_at	220183_s_at	235670_at	232052_at	223451_s_at
228454_at	229123_at	223059_s_at	242558_at	238337_s_at
224973_at	225001_at	213797_at	244859_at	235338_s_at
209115_at	213626_at	1554171_at	203482_at	218168_s_at
212382_at	224793_s_at	239300_at	202988_s_at	221797_at
212923_s_at	232210_at	225890_at	212799_at	226673_at
202190_at	227418_at	214486_x_at	228700_at	228280_at
238963_at	209434_s_at	218880_at	225262_at	229058_at
211063_s_at	218131_s_at	212470_at	37028_at	236953_s_at
225843_at	221677_s_at	225921_at	205053_at	225950_at
235422_at	244312_at	209539_at	221428_s_at	203721_s_at
209658_at	235159_at	218252_at	232024_at	1568680_s_at
211786_at	221813_at	242688_at	212757_s_at	203490_at
244788_at	206126_at	AFFX-HUMRGE/M10098_M_at	204970_s_at	229870_at
225528_at	230032_at	241418_at	201925_s_at	1561167_at
223599_at	213480_at	229742_at	210029_at	222665_at
209790_s_at	214290_s_at	1559584_a_at	211081_s_at	212991_at
224492_s_at	225666_at	227255_at	222792_s_at	214368_at
218431_at	228465_at	204720_s_at	222529_at	212644_s_at
1566901_at	234989_at	203321_s_at	201195_s_at	224639_at
221825_at	218927_s_at	223331_s_at	227395_at	223988_x_at

228095_at	201449_at	228077_at	225557_at	231897_at
238812_at	217591_at	50374_at	206118_at	213229_at
219947_at	213351_s_at	232356_at	223027_at	202693_s_at
204646_at	226969_at	209099_x_at	207606_s_at	223157_at
202439_s_at	218259_at	244087_at	209477_at	212018_s_at
64418_at	223070_at	244415_at	200731_s_at	220935_s_at
221073_s_at	202970_at	221652_s_at	219412_at	218543_s_at
224900_at	208883_at	209269_s_at	203428_s_at	203839_s_at
212659_s_at	229584_at	201368_at	212528_at	228961_at
216243_s_at	228661_s_at	208706_s_at	217875_s_at	227878_s_at
213015_at	230141_at	225350_s_at	232231_at	239891_x_at
225484_at	221806_s_at	211704_s_at	242103_at	33148_at
209199_s_at	225856_at	227570_at	208078_s_at	226017_at
201034_at	210541_s_at	217922_at	224739_at	227178_at
224130_s_at	212911_at	207980_s_at	204521_at	219774_at
242439_s_at	1553957_at	1559249_at	215203_at	227580_s_at
203263_s_at	212553_at	227621_at	219500_at	224461_s_at
227450_at	208785_s_at	223528_s_at	229908_s_at	229632_s_at
205192_at	229041_s_at	201020_at	1554638_at	1554173_at
205504_at	236610_at	208912_s_at	242920_at	226089_at
210826_x_at	235417_at	201489_at	231297_at	239154_at
242990_at	221877_at	203073_at	215499_at	218859_s_at
224175_s_at	228963_at	209447_at	201466_s_at	202301_s_at
223405_at	1555833_a_at	39248_at	222613_at	241393_at
212388_at	235926_at	216997_x_at	201465_s_at	210776_x_at
209889_at	213237_at	236685_at	1554132_a_at	212625_at
226609_at	213045_at	225212_at	224740_at	202386_s_at
217792_at	222371_at	207540_s_at	231431_s_at	232413_at
214435_x_at	227157_at	204157_s_at	202459_s_at	219322_s_at
218315_s_at	230779_at	214755_at	209411_s_at	219340_s_at
1558345_a_at	203319_s_at	205548_s_at	208392_x_at	220587_s_at
211563_s_at	213227_at	213625_at	1554149_at	213286_at
218570_at	227346_at	214684_at	218158_s_at	211105_s_at
218113_at	211661_x_at	235864_at	204794_at	220326_s_at
220329_s_at	226093_at	216268_s_at	1552318_at	226727_at
218938_at	225338_at	201626_at	211506_s_at	221752_at
202020_s_at	204817_at	204131_s_at	206036_s_at	214441_at
225749_at	224467_s_at	212748_at	226034_at	224618_at
227007_at	209761_s_at	203620_s_at	1559060_a_at	1562612_at
223580_at	236288_at	242625_at	225095_at	204382_at
226222_at	203420_at	1552846_s_at	212476_at	219684_at
212407_at	221223_x_at	209567_at	218530_at	201858_s_at

229298_at	238558_at	221940_at	204285_s_at	210538_s_at
212867_at	228630_at	218177_at	229566_at	208021_s_at
228523_at	202417_at	243579_at	1555858_at	221492_s_at
212527_at	226434_at	229821_at	225266_at	226874_at
227897_at	212188_at	226831_at	238880_at	225623_at
238057_at	205884_at	217886_at	212943_at	201279_s_at
223335_at	218452_at	213593_s_at	239629_at	228549_at
227210_at	212192_at	203542_s_at	1560071_a_at	242467_at
204871_at	226683_at	1555167_s_at	227571_at	218679_s_at
37577_at	222691_at	243303_at	204689_at	235692_at
227252_at	227559_at	219378_at	217738_at	203723_at
218152_at	226339_at	202365_at	214857_at	228385_at
204146_at	235728_at	201080_at	209771_x_at	219417_s_at
226867_at	213188_s_at	235005_at	212419_at	232405_at
224730_at	230100_x_at	222148_s_at	240383_at	224670_at
224558_s_at	225039_at	210793_s_at	229055_at	209433_s_at
236375_at	1566557_at	203823_at	209882_at	1555960_at
213598_at	212374_at	235339_at	226338_at	212429_s_at
205642_at	212386_at	238624_at	225252_at	219910_at
218230_at	218000_s_at	226949_at	208328_s_at	214946_x_at
238004_at	223466_x_at	204752_x_at	214696_at	212041_at
224643_at	203481_at	1569906_s_at	226875_at	202084_s_at
221925_s_at	220992_s_at	203358_s_at	212672_at	202664_at
224500_s_at	209431_s_at	225848_at	203741_s_at	223978_s_at
241391_at	225558_at	242121_at	208944_at	218076_s_at
228927_at	209721_s_at	228122_at	209378_s_at	235181_at
228793_at	211913_s_at	216218_s_at	218491_s_at	215299_x_at
206303_s_at	228084_at	229695_at	217739_s_at	218244_at
224682_at	229393_at	205811_at	207655_s_at	206769_at
228217_s_at	228544_s_at	219147_s_at	204683_at	220097_s_at
213396_s_at	222130_s_at	1554251_at	202776_at	201563_at
203348_s_at	203508_at	220035_at	210027_s_at	237104_at
AFFX-M27830_5_at	212693_at	201753_s_at	207559_s_at	1554768_a_at
223590_at	213064_at	218592_s_at	223797_at	1552618_at
212740_at	225860_at	222989_s_at	224692_at	1569139_s_at
211115_x_at	225918_at	1556064_at	210428_s_at	218833_at
235739_at	39891_at	232148_at	218178_s_at	201349_at
221666_s_at	218280_x_at	1554980_a_at	219734_at	204151_x_at
211316_x_at	222538_s_at	207113_s_at	219582_at	201528_at
228595_at	201280_s_at	212441_at	207667_s_at	213554_s_at
242725_at	204226_at	231907_at	240013_at	225751_at
230261_at	209704_at	238273_at	219448_at	224217_s_at

203089_s_at	227726_at	204032_at	226652_at	219131_at
218396_at	219487_at	232406_at	213309_at	221081_s_at
201384_s_at	229362_at	209694_at	236554_x_at	
218957_s_at	225769_at	222406_s_at	232504_at	
243993_at	227224_at	207030_s_at	223145_s_at	
218400_at	204788_s_at	212663_at	204923_at	
225076_s_at	222811_at	213670_x_at	223711_s_at	
222203_s_at	218827_s_at	218130_at	203846_at	
233656_s_at	205842_s_at	208998_at	242734_x_at	
243037_at	219081_at	240499_at	222955_s_at	
225313_at	203525_s_at	209770_at	36030_at	
219130_at	218797_s_at	1569362_at	223445_at	
203086_at	218870_at	203553_s_at	200815_s_at	
235464_at	211922_s_at	231548_at	218304_s_at	
212510_at	225762_x_at	204748_at	220560_at	
203351_s_at	219045_at	203640_at	204336_s_at	
224060_s_at	208925_at	226254_s_at	235222_x_at	
223320_s_at	210113_s_at	228940_at	203302_at	
209085_x_at	202951_at	225645_at	205771_s_at	

Appendix table 4.0: KEGG pathway enrichment analysis:

	MI vs *S pneumoniae*			**MI vs PLY**	
Pathway	**p.val**	**FDR q.val**	**Pathway**	**p.val**	**FDR q.val**
hsa04668 TNF signaling pathway	**3.0E-08**	**6.5E-06**	**hsa04668 TNF signaling pathway**	**3.2E-12**	**5.9E-10**
hsa05202 Transcriptional misregulation in cancer	**3.0E-07**	**6.5E-05**	**hsa05202 Transcriptional misregulation in cancer**	**5.5E-09**	**1.0E-06**
hsa04210 Apoptosis	**4.3E-06**	**9.5E-04**	**hsa04064 NF-kappa B signaling pathway**	**8.9E-09**	**1.6E-06**
hsa04064 NF-kappa B signaling pathway	**3.9E-05**	**8.6E-03**	**hsa05134 Legionellosis**	**6.6E-08**	**1.2E-05**
hsa04380 Osteoclast differentiation	**4.6E-05**	**1.0E-02**	**hsa04210 Apoptosis**	**9.0E-08**	**1.7E-05**
hsa05210 Colorectal cancer	**2.1E-04**	**4.6E-02**	**hsa04621 NOD-like receptor signaling pathway**	**1.3E-07**	**2.4E-05**
hsa05142 Chagas disease (American trypanosomiasis)	4.1E-04	9.0E-02	**hsa04010 MAPK signaling pathway**	**6.9E-07**	**1.3E-04**
hsa04010 MAPK signaling pathway	4.4E-04	9.7E-02	**hsa05132 Salmonella infection**	**3.3E-06**	**6.0E-04**
hsa04664 Fc epsilon RI signaling pathway	4.8E-04	1.1E-01	**hsa05323 Rheumatoid arthritis**	**4.8E-06**	**8.8E-04**
hsa05166 HTLV-I infection	5.0E-04	1.1E-01	**hsa04060 Cytokine-cytokine receptor interaction**	**5.5E-06**	**1.0E-03**
hsa04662 B cell receptor signaling pathway	6.4E-04	1.4E-01	**hsa05164 Influenza A**	**3.8E-05**	**7.0E-03**
hsa05164 Influenza A	7.5E-04	1.6E-01	**hsa04932 Non-alcoholic fatty liver disease (NAFLD)**	**5.5E-05**	**1.0E-02**
hsa04068 FoxO signaling pathway	8.3E-04	1.8E-01	**hsa05166 HTLV-I infection**	**7.6E-05**	**1.4E-02**
hsa04120 Ubiquitin mediated proteolysis	9.9E-04	2.2E-01	**hsa05133 Pertussis**	**9.4E-05**	**1.7E-02**
hsa05203 Viral carcinogenesis	1.1E-03	2.4E-01	**hsa04380 Osteoclast differentiation**	**1.0E-04**	**1.9E-02**

hsa05145 Toxoplasmosis	1.1E-03	2.4E-01	**hsa04620 Toll-like receptor signaling pathway**	**1.2E-04**	**2.3E-02**
hsa04070 Phosphatidylinositol signaling system	1.2E-03	2.6E-01	**hsa05200 Pathways in cancer**	**1.6E-04**	**3.1E-02**
hsa04015 Rap1 signaling pathway	1.5E-03	3.3E-01	hsa05321 Inflammatory bowel disease (IBD)	2.9E-04	5.4E-02
hsa03410 Base excision repair	1.6E-03	3.5E-01	hsa04062 Chemokine signaling pathway	3.2E-04	5.9E-02
hsa05161 Hepatitis B	1.7E-03	3.6E-01	hsa05120 Epithelial cell signaling in Helicobacter pylori infection	3.8E-04	7.0E-02
hsa04621 NOD-like receptor signaling pathway	1.8E-03	4.0E-01	hsa04640 Hematopoietic cell lineage	4.3E-04	7.9E-02
hsa04620 Toll-like receptor signaling pathway	1.9E-03	4.2E-01	hsa05140 Leishmaniasis	5.5E-04	1.0E-01
hsa05014 Amyotrophic lateral sclerosis (ALS)	2.1E-03	4.6E-01	hsa05203 Viral carcinogenesis	6.5E-04	1.2E-01
hsa00562 Inositol phosphate metabolism	2.6E-03	5.6E-01	hsa05161 Hepatitis B	1.1E-03	2.0E-01
hsa05160 Hepatitis C	2.7E-03	5.9E-01	hsa05143 African trypanosomiasis	1.4E-03	2.6E-01
hsa05169 Epstein-Barr virus infection	3.0E-03	6.5E-01	hsa05210 Colorectal cancer	1.7E-03	3.2E-01
hsa04666 Fc gamma R-mediated phagocytosis	3.1E-03	6.8E-01	hsa05169 Epstein-Barr virus infection	2.4E-03	4.4E-01
hsa04140 Regulation of autophagy	3.8E-03	8.2E-01	hsa05211 Renal cell carcinoma	2.5E-03	4.6E-01
hsa04144 Endocytosis	3.8E-03	8.3E-01	hsa05332 Graft-versus-host disease	2.5E-03	4.7E-01
hsa03018 RNA degradation	4.1E-03	8.9E-01	hsa05160 Hepatitis C	2.7E-03	5.0E-01
hsa05200 Pathways in cancer	4.4E-03	9.6E-01	hsa04115 p53 signaling pathway	2.8E-03	5.2E-01
hsa00010 Glycolysis / Gluconeogenesis	5.3E-01	1.0E+00	hsa05146 Amoebiasis	2.8E-03	5.2E-01
hsa00030 Pentose phosphate pathway	1.9E-01	1.0E+00	hsa04068 FoxO signaling pathway	2.8E-03	5.2E-01
hsa00040 Pentose and glucuronate interconversions	2.2E-01	1.0E+00	hsa04940 Type I diabetes mellitus	3.0E-03	5.6E-01
hsa00051 Fructose and mannose metabolism	2.1E-01	1.0E+00	hsa05162 Measles	3.1E-03	5.7E-01
hsa00071 Fatty acid degradation	3.2E-01	1.0E+00	hsa05142 Chagas disease (American trypanosomiasis)	3.4E-03	6.4E-01
hsa00072 Synthesis and degradation of ketone bodies	2.7E-02	1.0E+00	hsa04930 Type II diabetes mellitus	4.5E-03	8.3E-01
hsa00230 Purine metabolism	9.1E-02	1.0E+00	hsa05168 Herpes simplex infection	4.6E-03	8.6E-01
hsa00240 Pyrimidine metabolism	2.1E-02	1.0E+00	hsa05144 Malaria	4.8E-03	8.9E-01
hsa00250 Alanine, aspartate and glutamate metabolism	6.3E-02	1.0E+00	hsa00010 Glycolysis / Gluconeogenesis	2.4E-01	1.0E+00
hsa00270 Cysteine and methionine metabolism	6.3E-03	1.0E+00	hsa00052 Galactose metabolism	6.9E-02	1.0E+00
hsa00280 Valine, leucine and isoleucine degradation	3.6E-01	1.0E+00	hsa00230 Purine metabolism	3.7E-02	1.0E+00
hsa00310 Lysine degradation	4.6E-01	1.0E+00	hsa00240 Pyrimidine metabolism	1.6E-02	1.0E+00
hsa00340 Histidine metabolism	2.4E-02	1.0E+00	hsa00270 Cysteine and methionine metabolism	2.5E-02	1.0E+00
hsa00380 Tryptophan metabolism	8.7E-02	1.0E+00	hsa00310 Lysine degradation	5.0E-02	1.0E+00
hsa00410 beta-Alanine metabolism	1.9E-01	1.0E+00	hsa00380 Tryptophan metabolism	1.1E-01	1.0E+00
hsa00450 Selenocompound metabolism	9.3E-03	1.0E+00	hsa00450 Selenocompound metabolism	2.3E-02	1.0E+00
hsa00480 Glutathione metabolism	4.2E-01	1.0E+00	hsa00480 Glutathione metabolism	1.7E-01	1.0E+00
hsa00510 N-Glycan biosynthesis	3.7E-01	1.0E+00	hsa00520 Amino sugar and nucleotide sugar metabolism	1.4E-01	1.0E+00
hsa00512 Mucin type O-Glycan biosynthesis	1.9E-01	1.0E+00	hsa00561 Glycerolipid metabolism	2.0E-01	1.0E+00

hsa00520 Amino sugar and nucleotide sugar metabolism	1.3E-01	1.0E+00	hsa00562 Inositol phosphate metabolism	2.6E-01	1.0E+00
hsa00561 Glycerolipid metabolism	4.6E-01	1.0E+00	hsa00564 Glycerophospholipid metabolism	1.5E-01	1.0E+00
hsa00563 Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1.4E-01	1.0E+00	hsa00565 Ether lipid metabolism	2.5E-02	1.0E+00
hsa00564 Glycerophospholipid metabolism	2.4E-01	1.0E+00	hsa00590 Arachidonic acid metabolism	2.1E-01	1.0E+00
hsa00565 Ether lipid metabolism	1.1E-01	1.0E+00	hsa00600 Sphingolipid metabolism	1.4E-01	1.0E+00
hsa00590 Arachidonic acid metabolism	4.9E-01	1.0E+00	hsa00620 Pyruvate metabolism	1.0E-01	1.0E+00
hsa00591 Linoleic acid metabolism	1.8E-01	1.0E+00	hsa00670 One carbon pool by folate	3.1E-02	1.0E+00
hsa00592 alpha-Linolenic acid metabolism	1.4E-01	1.0E+00	hsa00760 Nicotinate and nicotinamide metabolism	6.2E-02	1.0E+00
hsa00600 Sphingolipid metabolism	3.5E-01	1.0E+00	hsa00860 Porphyrin and chlorophyll metabolism	1.2E-01	1.0E+00
hsa00604 Glycosphingolipid biosynthesis - ganglio series	5.8E-02	1.0E+00	hsa01100 Metabolic pathways	5.8E-01	1.0E+00
hsa00620 Pyruvate metabolism	2.7E-01	1.0E+00	hsa01200 Carbon metabolism	4.7E-01	1.0E+00
hsa00630 Glyoxylate and dicarboxylate metabolism	3.6E-02	1.0E+00	hsa01230 Biosynthesis of amino acids	2.8E-01	1.0E+00
hsa00650 Butanoate metabolism	1.7E-01	1.0E+00	hsa03008 Ribosome biogenesis in eukaryotes	1.7E-02	1.0E+00
hsa00670 One carbon pool by folate	9.6E-02	1.0E+00	hsa03013 RNA transport	9.2E-02	1.0E+00
hsa00760 Nicotinate and nicotinamide metabolism	4.0E-02	1.0E+00	hsa03015 mRNA surveillance pathway	3.6E-01	1.0E+00
hsa00770 Pantothenate and CoA biosynthesis	8.0E-02	1.0E+00	hsa03022 Basal transcription factors	1.3E-01	1.0E+00
hsa00860 Porphyrin and chlorophyll metabolism	2.4E-02	1.0E+00	hsa03030 DNA replication	1.4E-02	1.0E+00
hsa00900 Terpenoid backbone biosynthesis	1.1E-01	1.0E+00	hsa03320 PPAR signaling pathway	2.7E-01	1.0E+00
hsa00910 Nitrogen metabolism	7.2E-02	1.0E+00	hsa03410 Base excision repair	1.1E-02	1.0E+00
hsa00983 Drug metabolism - other enzymes	1.2E-01	1.0E+00	hsa03420 Nucleotide excision repair	2.8E-02	1.0E+00
hsa01100 Metabolic pathways	2.5E-01	1.0E+00	hsa03430 Mismatch repair	4.1E-02	1.0E+00
hsa01200 Carbon metabolism	3.5E-01	1.0E+00	hsa03440 Homologous recombination	1.1E-01	1.0E+00
hsa01212 Fatty acid metabolism	3.6E-01	1.0E+00	hsa04012 ErbB signaling pathway	7.9E-03	1.0E+00
hsa01230 Biosynthesis of amino acids	4.7E-02	1.0E+00	hsa04014 Ras signaling pathway	1.6E-02	1.0E+00
hsa03008 Ribosome biogenesis in eukaryotes	6.8E-03	1.0E+00	hsa04015 Rap1 signaling pathway	1.0E-02	1.0E+00
hsa03013 RNA transport	2.9E-01	1.0E+00	hsa04020 Calcium signaling pathway	2.5E-01	1.0E+00
hsa03015 mRNA surveillance pathway	2.2E-01	1.0E+00	hsa04022 cGMP-PKG signaling pathway	3.1E-02	1.0E+00
hsa03030 DNA replication	1.4E-02	1.0E+00	hsa04024 cAMP signaling pathway	2.3E-02	1.0E+00
hsa03040 Spliceosome	4.7E-01	1.0E+00	hsa04066 HIF-1 signaling pathway	5.7E-02	1.0E+00
hsa03320 PPAR signaling pathway	2.9E-01	1.0E+00	hsa04070 Phosphatidylinositol signaling system	1.6E-01	1.0E+00
hsa03420 Nucleotide excision repair	3.4E-02	1.0E+00	hsa04071 Sphingolipid signaling pathway	2.7E-02	1.0E+00
hsa03430 Mismatch repair	2.2E-02	1.0E+00	hsa04080 Neuroactive ligand-receptor interaction	5.5E-01	1.0E+00
hsa03440 Homologous recombination	9.2E-02	1.0E+00	hsa04110 Cell cycle	9.6E-02	1.0E+00
hsa03460 Fanconi anemia pathway	5.6E-02	1.0E+00	hsa04120 Ubiquitin mediated proteolysis	1.3E-02	1.0E+00

hsa04012 ErbB signaling pathway	8.4E-03	1.0E+00	hsa04130 SNARE interactions in vesicular transport	8.2E-02	1.0E+00
hsa04014 Ras signaling pathway	8.1E-02	1.0E+00	hsa04140 Regulation of autophagy	1.1E-01	1.0E+00
hsa04020 Calcium signaling pathway	2.0E-01	1.0E+00	hsa04141 Protein processing in endoplasmic reticulum	8.4E-02	1.0E+00
hsa04022 cGMP-PKG signaling pathway	1.3E-02	1.0E+00	hsa04142 Lysosome	5.1E-01	1.0E+00
hsa04024 cAMP signaling pathway	6.8E-03	1.0E+00	hsa04144 Endocytosis	1.6E-01	1.0E+00
hsa04060 Cytokine-cytokine receptor interaction	1.1E-02	1.0E+00	hsa04150 mTOR signaling pathway	6.0E-03	1.0E+00
hsa04062 Chemokine signaling pathway	1.2E-01	1.0E+00	hsa04151 PI3K-Akt signaling pathway	5.1E-02	1.0E+00
hsa04066 HIF-1 signaling pathway	1.9E-02	1.0E+00	hsa04152 AMPK signaling pathway	9.6E-02	1.0E+00
hsa04071 Sphingolipid signaling pathway	9.6E-02	1.0E+00	hsa04261 Adrenergic signaling in cardiomyocytes	1.6E-01	1.0E+00
hsa04080 Neuroactive ligand-receptor interaction	7.4E-01	1.0E+00	hsa04310 Wnt signaling pathway	5.1E-02	1.0E+00
hsa04110 Cell cycle	5.5E-03	1.0E+00	hsa04330 Notch signaling pathway	1.4E-01	1.0E+00
hsa04114 Oocyte meiosis	4.1E-01	1.0E+00	hsa04350 TGF-beta signaling pathway	3.0E-02	1.0E+00
hsa04115 p53 signaling pathway	9.4E-03	1.0E+00	hsa04360 Axon guidance	1.0E-01	1.0E+00
hsa04130 SNARE interactions in vesicular transport	5.9E-02	1.0E+00	hsa04370 VEGF signaling pathway	5.4E-02	1.0E+00
hsa04141 Protein processing in endoplasmic reticulum	7.2E-02	1.0E+00	hsa04390 Hippo signaling pathway	6.4E-01	1.0E+00
hsa04142 Lysosome	6.3E-01	1.0E+00	hsa04510 Focal adhesion	2.4E-02	1.0E+00
hsa04145 Phagosome	5.8E-01	1.0E+00	hsa04512 ECM-receptor interaction	3.2E-01	1.0E+00
hsa04146 Peroxisome	3.7E-01	1.0E+00	hsa04514 Cell adhesion molecules (CAMs)	6.0E-01	1.0E+00
hsa04150 mTOR signaling pathway	5.1E-02	1.0E+00	hsa04520 Adherens junction	2.8E-01	1.0E+00
hsa04151 PI3K-Akt signaling pathway	3.7E-02	1.0E+00	hsa04540 Gap junction	3.5E-01	1.0E+00
hsa04152 AMPK signaling pathway	2.3E-01	1.0E+00	hsa04550 Signaling pathways regulating pluripotency of stem cells	4.9E-02	1.0E+00
hsa04261 Adrenergic signaling in cardiomyocytes	1.7E-02	1.0E+00	hsa04610 Complement and coagulation cascades	9.9E-02	1.0E+00
hsa04270 Vascular smooth muscle contraction	6.2E-01	1.0E+00	hsa04611 Platelet activation	2.4E-01	1.0E+00
hsa04310 Wnt signaling pathway	8.3E-02	1.0E+00	hsa04612 Antigen processing and presentation	9.3E-02	1.0E+00
hsa04330 Notch signaling pathway	3.6E-01	1.0E+00	hsa04614 Renin-angiotensin system	4.1E-02	1.0E+00
hsa04350 TGF-beta signaling pathway	7.0E-02	1.0E+00	hsa04622 RIG-I-like receptor signaling pathway	1.7E-02	1.0E+00
hsa04360 Axon guidance	9.5E-02	1.0E+00	hsa04623 Cytosolic DNA-sensing pathway	6.0E-02	1.0E+00
hsa04370 VEGF signaling pathway	7.6E-02	1.0E+00	hsa04630 Jak-STAT signaling pathway	7.1E-02	1.0E+00
hsa04390 Hippo signaling pathway	3.8E-01	1.0E+00	hsa04650 Natural killer cell mediated cytotoxicity	4.1E-02	1.0E+00
hsa04510 Focal adhesion	1.7E-01	1.0E+00	hsa04660 T cell receptor signaling pathway	1.6E-02	1.0E+00
hsa04512 ECM-receptor interaction	3.6E-01	1.0E+00	hsa04662 B cell receptor signaling pathway	1.9E-02	1.0E+00
hsa04514 Cell adhesion molecules (CAMs)	3.3E-01	1.0E+00	hsa04664 Fc epsilon RI signaling pathway	1.7E-02	1.0E+00
hsa04520 Adherens junction	3.1E-01	1.0E+00	hsa04666 Fc gamma R-mediated phagocytosis	1.4E-01	1.0E+00
hsa04540 Gap junction	4.1E-01	1.0E+00	hsa04670 Leukocyte transendothelial migration	4.9E-01	1.0E+00
hsa04550 Signaling pathways regulating pluripotency of stem cells	1.3E-02	1.0E+00	hsa04672 Intestinal immune network for IgA production	1.5E-01	1.0E+00

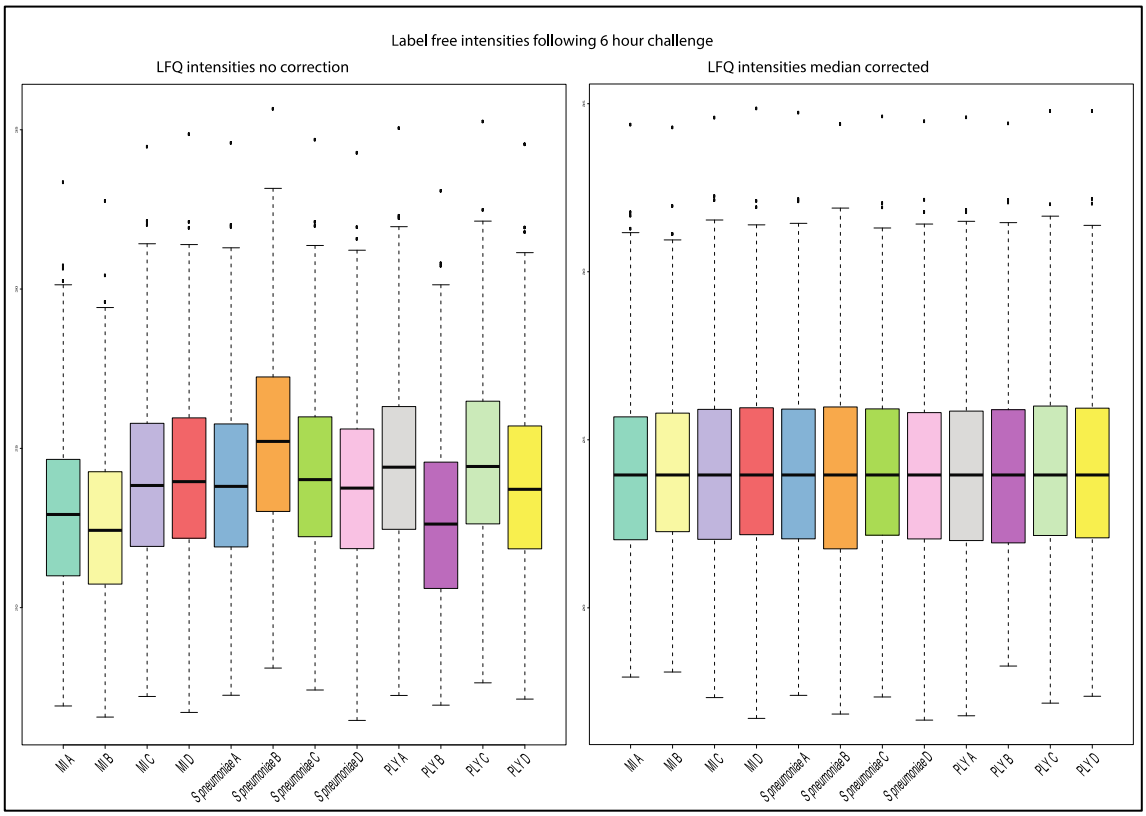
hsa04610 Complement and coagulation cascades	3.4E-01	1.0E+00	hsa04710 Circadian rhythm	6.9E-02	1.0E+00
hsa04611 Platelet activation	4.2E-02	1.0E+00	hsa04713 Circadian entrainment	3.9E-01	1.0E+00
hsa04612 Antigen processing and presentation	1.4E-01	1.0E+00	hsa04722 Neurotrophin signaling pathway	9.0E-02	1.0E+00
hsa04614 Renin-angiotensin system	1.2E-01	1.0E+00	hsa04723 Retrograde endocannabinoid signaling	1.7E-01	1.0E+00
hsa04622 RIG-I-like receptor signaling pathway	1.1E-01	1.0E+00	hsa04725 Cholinergic synapse	2.0E-01	1.0E+00
hsa04630 Jak-STAT signaling pathway	2.4E-02	1.0E+00	hsa04726 Serotonergic synapse	4.7E-01	1.0E+00
hsa04640 Hematopoietic cell lineage	1.1E-01	1.0E+00	hsa04727 GABAergic synapse	3.5E-01	1.0E+00
hsa04650 Natural killer cell mediated cytotoxicity	9.4E-03	1.0E+00	hsa04728 Dopaminergic synapse	5.4E-01	1.0E+00
hsa04660 T cell receptor signaling pathway	6.4E-03	1.0E+00	hsa04750 Inflammatory mediator regulation of TRP channels	5.0E-02	1.0E+00
hsa04670 Leukocyte transendothelial migration	2.0E-01	1.0E+00	hsa04810 Regulation of actin cytoskeleton	8.3E-02	1.0E+00
hsa04672 Intestinal immune network for IgA production	1.4E-01	1.0E+00	hsa04910 Insulin signaling pathway	5.8E-01	1.0E+00
hsa04710 Circadian rhythm	4.7E-02	1.0E+00	hsa04911 Insulin secretion	3.3E-01	1.0E+00
hsa04713 Circadian entrainment	1.1E-01	1.0E+00	hsa04912 GnRH signaling pathway	1.4E-01	1.0E+00
hsa04720 Long-term potentiation	5.3E-01	1.0E+00	hsa04913 Ovarian steroidogenesis	1.5E-01	1.0E+00
hsa04722 Neurotrophin signaling pathway	9.9E-02	1.0E+00	hsa04914 Progesterone-mediated oocyte maturation	4.0E-01	1.0E+00
hsa04723 Retrograde endocannabinoid signaling	2.7E-01	1.0E+00	hsa04915 Estrogen signaling pathway	5.2E-02	1.0E+00
hsa04724 Glutamatergic synapse	3.5E-01	1.0E+00	hsa04916 Melanogenesis	1.7E-01	1.0E+00
hsa04725 Cholinergic synapse	2.7E-02	1.0E+00	hsa04917 Prolactin signaling pathway	1.8E-02	1.0E+00
hsa04727 GABAergic synapse	2.0E-01	1.0E+00	hsa04918 Thyroid hormone synthesis	2.8E-01	1.0E+00
hsa04728 Dopaminergic synapse	2.6E-01	1.0E+00	hsa04919 Thyroid hormone signaling pathway	2.3E-01	1.0E+00
hsa04740 Olfactory transduction	1.0E+00	1.0E+00	hsa04920 Adipocytokine signaling pathway	7.5E-02	1.0E+00
hsa04750 Inflammatory mediator regulation of TRP channels	1.5E-02	1.0E+00	hsa04921 Oxytocin signaling pathway	2.5E-02	1.0E+00
hsa04810 Regulation of actin cytoskeleton	1.2E-01	1.0E+00	hsa04922 Glucagon signaling pathway	1.5E-02	1.0E+00
hsa04910 Insulin signaling pathway	5.0E-01	1.0E+00	hsa04950 Maturity onset diabetes of the young	5.1E-02	1.0E+00
hsa04911 Insulin secretion	7.3E-02	1.0E+00	hsa04960 Aldosterone-regulated sodium reabsorption	1.0E-01	1.0E+00
hsa04912 GnRH signaling pathway	3.4E-02	1.0E+00	hsa04962 Vasopressin-regulated water reabsorption	2.3E-02	1.0E+00
hsa04913 Ovarian steroidogenesis	3.8E-01	1.0E+00	hsa04971 Gastric acid secretion	2.8E-01	1.0E+00
hsa04914 Progesterone-mediated oocyte maturation	4.4E-02	1.0E+00	hsa04973 Carbohydrate digestion and absorption	1.4E-01	1.0E+00
hsa04915 Estrogen signaling pathway	1.6E-02	1.0E+00	hsa04977 Vitamin digestion and absorption	4.4E-02	1.0E+00
hsa04916 Melanogenesis	1.3E-01	1.0E+00	hsa04978 Mineral absorption	1.6E-01	1.0E+00
hsa04917 Prolactin signaling pathway	4.1E-02	1.0E+00	hsa05010 Alzheimer's disease	9.2E-02	1.0E+00
hsa04918 Thyroid hormone synthesis	1.3E-01	1.0E+00	hsa05012 Parkinson's disease	5.9E-01	1.0E+00
hsa04919 Thyroid hormone signaling pathway	2.0E-01	1.0E+00	hsa05014 Amyotrophic lateral sclerosis (ALS)	3.4E-02	1.0E+00
hsa04920 Adipocytokine signaling pathway	1.1E-01	1.0E+00	hsa05016 Huntington's disease	5.1E-01	1.0E+00
hsa04921 Oxytocin signaling pathway	2.5E-02	1.0E+00	hsa05020 Prion diseases	1.3E-02	1.0E+00

hsa04922 Glucagon signaling pathway	5.8E-03	1.0E+00	hsa05030 Cocaine addiction	3.1E-02	1.0E+00
hsa04930 Type II diabetes mellitus	8.3E-03	1.0E+00	hsa05031 Amphetamine addiction	7.0E-02	1.0E+00
hsa04932 Non-alcoholic fatty liver disease (NAFLD)	1.8E-02	1.0E+00	hsa05034 Alcoholism	7.2E-01	1.0E+00
hsa04940 Type I diabetes mellitus	1.0E-01	1.0E+00	hsa05100 Bacterial invasion of epithelial cells	3.0E-01	1.0E+00
hsa04950 Maturity onset diabetes of the young	1.5E-01	1.0E+00	hsa05131 Shigellosis	2.3E-01	1.0E+00
hsa04960 Aldosterone-regulated sodium reabsorption	8.2E-02	1.0E+00	hsa05145 Toxoplasmosis	6.3E-03	1.0E+00
hsa04962 Vasopressin-regulated water reabsorption	2.8E-02	1.0E+00	hsa05150 Staphylococcus aureus infection	1.8E-01	1.0E+00
hsa04970 Salivary secretion	6.9E-01	1.0E+00	hsa05152 Tuberculosis	2.4E-01	1.0E+00
hsa04971 Gastric acid secretion	1.3E-01	1.0E+00	hsa05204 Chemical carcinogenesis	3.2E-01	1.0E+00
hsa04972 Pancreatic secretion	2.4E-01	1.0E+00	hsa05205 Proteoglycans in cancer	6.9E-02	1.0E+00
hsa04973 Carbohydrate digestion and absorption	3.2E-02	1.0E+00	hsa05206 MicroRNAs in cancer	2.5E-02	1.0E+00
hsa04975 Fat digestion and absorption	2.9E-01	1.0E+00	hsa05212 Pancreatic cancer	1.4E-02	1.0E+00
hsa04976 Bile secretion	5.6E-01	1.0E+00	hsa05213 Endometrial cancer	3.6E-02	1.0E+00
hsa04977 Vitamin digestion and absorption	1.3E-01	1.0E+00	hsa05214 Glioma	1.4E-02	1.0E+00
hsa04978 Mineral absorption	4.0E-01	1.0E+00	hsa05215 Prostate cancer	8.2E-03	1.0E+00
hsa05010 Alzheimer's disease	6.6E-01	1.0E+00	hsa05217 Basal cell carcinoma	1.8E-01	1.0E+00
hsa05012 Parkinson's disease	7.2E-01	1.0E+00	hsa05218 Melanoma	1.7E-02	1.0E+00
hsa05016 Huntington's disease	7.5E-01	1.0E+00	hsa05219 Bladder cancer	1.9E-02	1.0E+00
hsa05030 Cocaine addiction	9.0E-03	1.0E+00	hsa05220 Chronic myeloid leukemia	8.2E-02	1.0E+00
hsa05031 Amphetamine addiction	8.8E-03	1.0E+00	hsa05221 Acute myeloid leukemia	1.9E-01	1.0E+00
hsa05032 Morphine addiction	6.9E-01	1.0E+00	hsa05222 Small cell lung cancer	7.2E-03	1.0E+00
hsa05033 Nicotine addiction	2.8E-01	1.0E+00	hsa05223 Non-small cell lung cancer	4.4E-02	1.0E+00
hsa05034 Alcoholism	3.3E-01	1.0E+00	hsa05230 Central carbon metabolism in cancer	1.4E-02	1.0E+00
hsa05100 Bacterial invasion of epithelial cells	3.4E-01	1.0E+00	hsa05231 Choline metabolism in cancer	5.3E-02	1.0E+00
hsa05120 Epithelial cell signaling in Helicobacter pylori infection	8.8E-03	1.0E+00	hsa05310 Asthma	6.9E-02	1.0E+00
hsa05131 Shigellosis	2.4E-01	1.0E+00	hsa05322 Systemic lupus erythematosus	5.6E-01	1.0E+00
hsa05132 Salmonella infection	2.6E-02	1.0E+00	hsa05330 Allograft rejection	9.9E-02	1.0E+00
hsa05133 Pertussis	4.9E-02	1.0E+00	hsa05410 Hypertrophic cardiomyopathy (HCM)	6.2E-03	1.0E+00
hsa05134 Legionellosis	1.4E-02	1.0E+00	hsa05412 Arrhythmogenic right ventricular cardiomyopathy (ARVC)	2.0E-02	1.0E+00
hsa05140 Leishmaniasis	1.2E-02	1.0E+00	hsa05414 Dilated cardiomyopathy	8.6E-03	1.0E+00
hsa05143 African trypanosomiasis	1.3E-02	1.0E+00	hsa05416 Viral myocarditis	2.0E-01	1.0E+00
hsa05144 Malaria	1.4E-01	1.0E+00			
hsa05146 Amoebiasis	4.8E-02	1.0E+00			
hsa05150 Staphylococcus aureus infection	1.8E-01	1.0E+00			
hsa05152 Tuberculosis	2.0E-02	1.0E+00			
hsa05162 Measles	2.7E-02	1.0E+00			

hsa05168 Herpes simplex infection	2.5E-02	1.0E+00			
hsa05205 Proteoglycans in cancer	4.5E-02	1.0E+00			
hsa05206 MicroRNAs in cancer	2.6E-01	1.0E+00			
hsa05211 Renal cell carcinoma	8.2E-03	1.0E+00			
hsa05212 Pancreatic cancer	7.6E-03	1.0E+00			
hsa05213 Endometrial cancer	4.7E-02	1.0E+00			
hsa05214 Glioma	7.6E-03	1.0E+00			
hsa05215 Prostate cancer	8.9E-03	1.0E+00			
hsa05217 Basal cell carcinoma	1.7E-01	1.0E+00			
hsa05218 Melanoma	3.9E-02	1.0E+00			
hsa05219 Bladder cancer	9.2E-02	1.0E+00			
hsa05220 Chronic myeloid leukemia	4.3E-02	1.0E+00			
hsa05221 Acute myeloid leukemia	1.9E-01	1.0E+00			
hsa05222 Small cell lung cancer	7.6E-02	1.0E+00			
hsa05223 Non-small cell lung cancer	1.8E-01	1.0E+00			
hsa05230 Central carbon metabolism in cancer	3.1E-02	1.0E+00			
hsa05231 Choline metabolism in cancer	1.3E-01	1.0E+00			
hsa05310 Asthma	1.9E-01	1.0E+00			
hsa05320 Autoimmune thyroid disease	4.1E-01	1.0E+00			
hsa05321 Inflammatory bowel disease (IBD)	2.8E-02	1.0E+00			
hsa05322 Systemic lupus erythematosus	4.7E-01	1.0E+00			
hsa05323 Rheumatoid arthritis	8.8E-02	1.0E+00			
hsa05330 Allograft rejection	7.7E-02	1.0E+00			
hsa05332 Graft-versus-host disease	9.2E-02	1.0E+00			
hsa05340 Primary immunodeficiency	7.2E-02	1.0E+00			
hsa05410 Hypertrophic cardiomyopathy (HCM)	6.7E-02	1.0E+00			
hsa05412 Arrhythmogenic right ventricular cardiomyopathy (ARVC)	1.3E-01	1.0E+00			
hsa05414 Dilated cardiomyopathy	3.1E-02	1.0E+00			
hsa05416 Viral myocarditis	4.6E-01	1.0E+00			

Appendix table 4.3: 6hr LFQ data filtering.

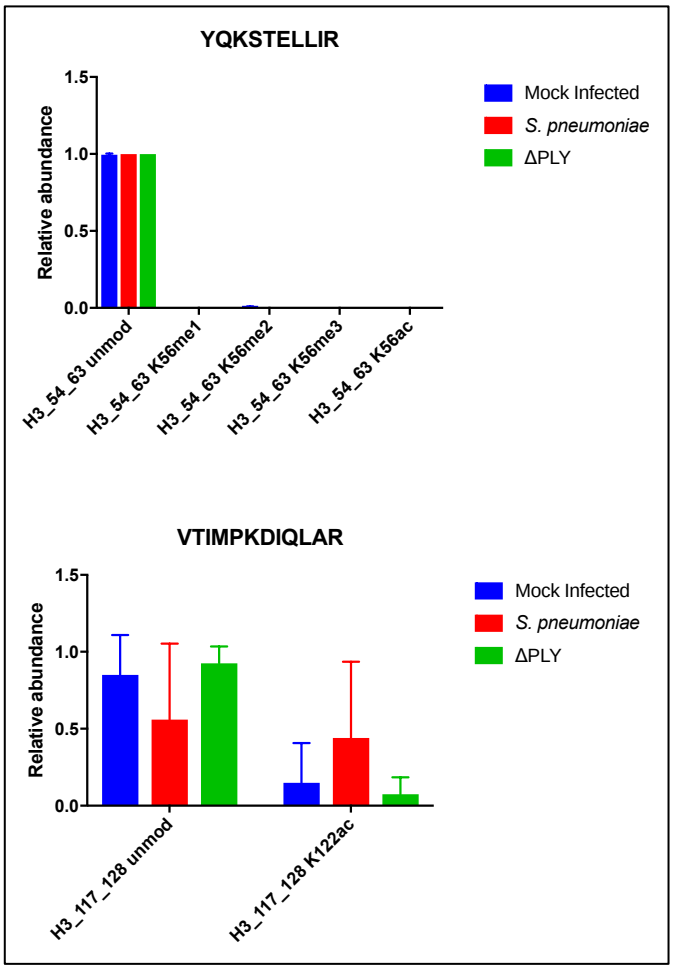
	Mock Infected				*S pneumoniae*				△PLY				
Donor	Donor A	Donor B	Donor C	donor D	Donor A	Donor B	Donor C	donor D	Donor A	Donor B	Donor C	donor D	average
Majority protein identificati	2261	1656	2283	2445	2419	1823	2312	2349	2510	1883	2558	2314	2194
reverse matches removed	2244	1642	2264	2426	2397	1813	2295	2331	2491	1872	2534	2295	2210
contaminant matches remo	2224	1626	2248	2405	2377	1799	2277	2312	2472	1853	2514	2275	2192
>2 unique peptide matches	1836	1420	1856	1962	1933	1538	1869	1891	1970	1602	2006	1870	1808



Appendix figure 4.2: Boxplots of six hour LFQ intensity before and after median correction.

Appendix table 4.4: Differentially expressed proteins identified following ANOVA.

	MI vs *S pneumoniae*			MI vs PLY				
Majority Protien ID	logFC	P.Value	adj.P.Val	logFC	P.Value	adj.P.Val	protein name	Gene name
Q08426	1.49E+00	5.25E-02	8.34E-02	6.67E+00	1.29E-08	4.14E-07	Peroxisomal bifunctional enzyme	EHHADH
H0Y8C3;Q9NZJ7	-8.67E-01	8.03E-02	1.08E-01	-2.88E+00	2.34E-05	3.75E-04	Mitochondrial carrier homolog 1	MTCH1
Q14139	-1.69E+00	5.47E-02	8.34E-02	3.92E+00	7.24E-05	7.72E-04	Ubiquitin conjugation factor E4 A	UBE4A
P36871	-2.19E+00	5.77E-03	1.42E-02	-2.69E+00	1.02E-04	8.14E-04	Phosphoglucomutase-1 (	PGM1
Q709C8	-1.83E+00	2.00E-03	6.42E-03	-2.06E+00	1.78E-04	1.04E-03	Vacuolar protein sorting-associated protein 13C	VPS13C
Q92508	2.29E-01	7.61E-01	7.61E-01	-2.36E+00	1.96E-04	1.04E-03	Piezo-type mechanosensitive ion channel component 1	PIEZO1
O95865;A0A140T971;Q5SSV3;Q5SRR8	-1.48E+00	3.31E-02	5.88E-02	-2.13E+00	2.95E-04	1.09E-03	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2	DDAH2
O14974;A0A0U1RQZ1;H0YIM2;H0YIS3	1.02E+00	8.13E-02	1.08E-01	2.25E+00	3.21E-04	1.09E-03	Protein phosphatase 1 regulatory subunit 12A	PPP1R12A
Q9BXW7	-8.74E-01	2.29E-01	2.62E-01	-2.79E+00	2.39E-04	1.09E-03	Haloacid dehalogenase-like hydrolase domain-containing 5	HDHD5
H0Y8X4;O43598	-2.52E-01	7.29E-01	7.52E-01	-2.74E+00	3.39E-04	1.09E-03	2'-deoxynucleoside 5'-phosphate N-hydrolase 1	DNPH1
Q16401;F2Z3J2	2.76E+00	1.92E-06	6.14E-05	1.90E+00	6.73E-04	1.66E-03	26S proteasome non-ATPase regulatory subunit 5	PSMD5
Q92616	-7.39E-01	1.55E-01	1.91E-01	-1.79E+00	6.15E-04	1.66E-03	eIF-2-alpha kinase activator GCN1	GCN1
Q495G5;Q8IVH4;D6RIS5	7.06E-01	1.79E-01	2.13E-01	1.94E+00	6.61E-04	1.66E-03	Methylmalonic aciduria type A protein, mitochondrial	MMAA
Q08380;K7EJY8;K7ESM3;K7EKQ5;K7EP36	-2.36E+00	1.57E-04	1.26E-03	-1.94E+00	1.30E-03	2.96E-03	Galectin-3-binding protein	LGALS3BP
E9PF10;O75694	3.86E-01	5.89E-01	6.29E-01	2.01E+00	1.53E-03	3.27E-03	Nuclear pore complex protein Nup155	NUP155
A0A087WUZ3;Q01082	2.52E+00	1.73E-03	6.42E-03	2.02E+00	2.02E-03	4.05E-03	Spectrin beta chain, non-erythrocytic 1	SPTBN1
Q5U5X0	1.66E+00	4.79E-03	1.28E-02	1.47E+00	5.25E-03	9.57E-03	Complex III assembly factor LYRM7	LYRM7
P26639	-1.42E+00	4.11E-02	6.93E-02	-2.12E+00	5.38E-03	9.57E-03	Threonine--tRNA ligase, cytoplasmic	TARS
M0R300;M0R0P8;Q13459	-1.27E+00	2.17E-02	4.09E-02	-2.05E+00	6.71E-03	1.13E-02	Unconventional myosin-Ixb	MYOB9B
P62244;I3L3P7	4.58E-01	2.43E-01	2.68E-01	1.09E+00	7.51E-03	1.20E-02	40S ribosomal protein S15a	RPS15A
P23219;A0A087X296	-2.83E+00	1.96E-03	6.42E-03	-1.67E+00	1.70E-02	2.60E-02	Prostaglandin G/H synthase 1	PTGS1
Q9Y5X2	1.63E+00	1.18E-02	2.47E-02	-1.22E+00	2.31E-02	3.36E-02	Sorting nexin-8	SNX8
O95747;C9JIG9	-2.36E+00	2.39E-05	3.82E-04	-1.03E+00	4.02E-02	5.59E-02	Serine/threonine-protein kinase OSR1	OXSR1
O60826	2.54E+00	4.21E-03	1.23E-02	9.88E-01	1.50E-01	1.92E-01	Coiled-coil domain-containing protein 22	CCDC22
F8W031;F8VXJ7;Q9Y2B0;F8W1K5	-8.82E-01	5.89E-02	8.56E-02	6.65E-01	1.50E-01	1.92E-01	Protein canopy homolog 2	CNPY2
A0A0C4DG89;Q7L014;D6RJA6	1.69E+00	1.83E-03	6.42E-03	6.49E-01	1.74E-01	2.15E-01	Probable ATP-dependent RNA helicase DDX46	DDX46
Q9HAC8;Q8WUN7	2.84E+00	3.87E-04	2.48E-03	6.83E-01	2.49E-01	2.95E-01	Ubiquitin domain-containing protein 1 / 2	UBTD2
O00115;K7ENE5	-6.51E-01	9.23E-02	1.18E-01	8.86E-02	8.15E-01	9.00E-01	Deoxyribonuclease-2-alpha	DNASE2
P46087	1.62E+00	1.39E-03	6.42E-03	6.09E-02	9.07E-01	9.32E-01	Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	NOP2
R4GMN1;Q8NHP6	-1.29E+00	6.85E-03	1.57E-02	-4.13E-02	9.32E-01	9.32E-01	Motile sperm domain-containing protein 2	MOSPD2
O95336;M0R261	-1.18E+00	1.23E-02	2.47E-02	-6.15E-02	8.98E-01	9.32E-01	6-phosphogluconolactonase	PGLS
P51531;A0A0U1RQZ9;F6VDE0;Q9HBD4;A0A0A0MT49;P51532;A0A0U1RRD6;A0A0U1RRF8;A0A0A0MSS5;F6UH26;A0A0U1RQU0;A0A0U1RR09;F6XE55;A0A0U1RQW7;A0A0U1RRN2;B1ALG1;A0A0U1RR83;A0A0U1RQX3;F6T8Q0;B1ALG2;A0A0U1RR26;A0A0U1RRG6;B4DNT1;B1ALF6	-3.69E+00	3.61E-05	3.85E-04	-5.99E-01	4.37E-01	4.99E-01	Probable global transcription activator SNF2L2 / Transcription activator BRG1	SMARCA2 / SMARCA 4



Appendix figure 4.3: Relative abundance of the histone PTMs on YQSTELLIR and VTIMPKDIQLAR peptide.

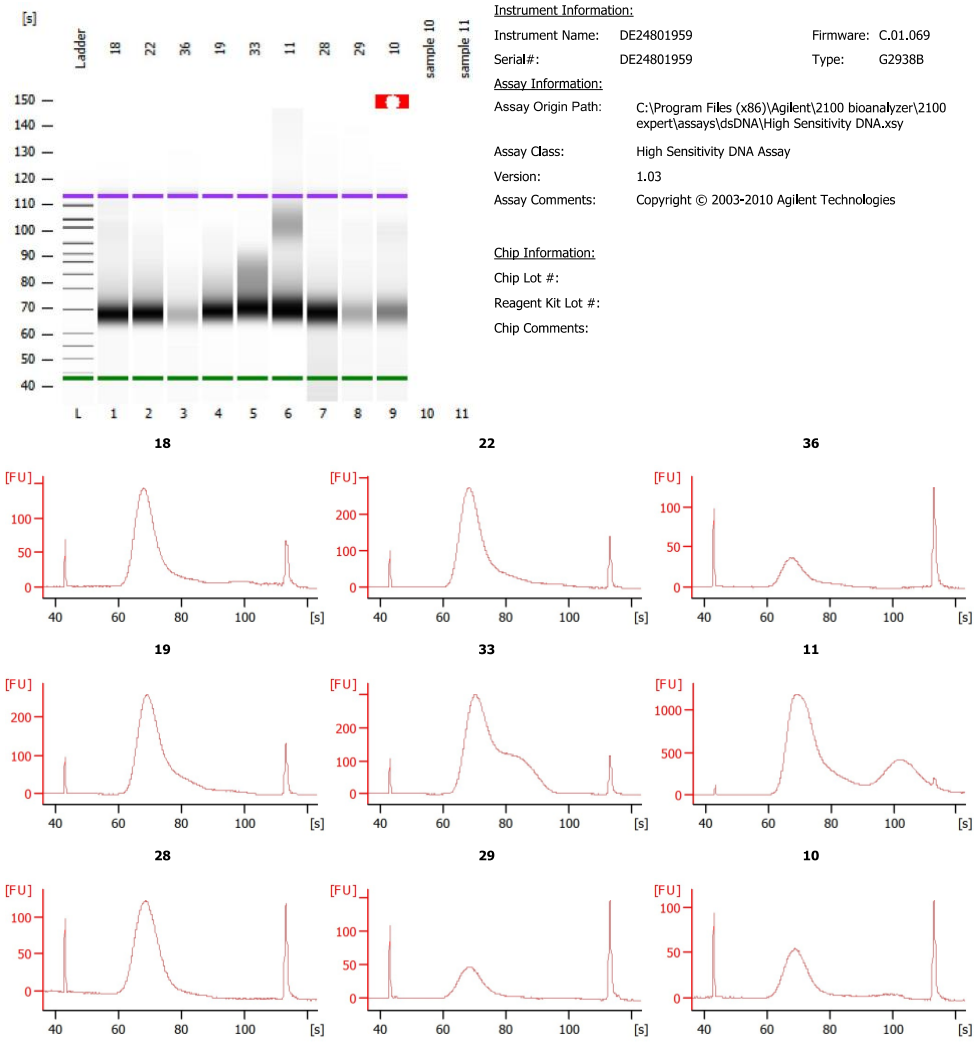
This figure illustrates the relative abundance in the histone PTMs identified on both peptides. There was no significant change observed between the infective challenges and the MI samples. n=3, on way ANOVA.

Appendix table 5.1 Summary of the histone PTMs for H3 and H4 identified:

	maXis	QE□0.1%□TFA	QE□0.05%□HFBA
H3_3_8 unmod			
H3_3_8 K4me1			
H3_3_8 K4me2			
H3_3_8 K4me3			
H3_9_17 unmod			
H3_9_17 K9me1			
H3_9_17 K9me2			
H3_9_17 K9me3			
H3_9_17 K9ac			
H3_9_17 K14ac			
H3_9_17 K9me1K14ac			
H3_9_17 K9me2K14ac			
H3_9_17 K9me3K14ac			
H3_9_17 K9acK14ac			
H3_18_26 unmod			
H3_18_26 K23me1			
H3_18_26 K18me1			
H3_18_26 K18me1K23me1			
H3_18_26 K18ac			
H3_18_26 K23ac			
H3_18_26 K18acK23ac			
H3K27ac1K36me0			
H3K27me0K36me0			
H3K27me0K36me1			
H3K27me0K36me2			
H3K27me0K36me3			
H3K27me1K36me0			
H3K27me1K36me1			
H3K27me1K36me2			
H3K27me1K36me3			
H3K27me2K36me0			
H3K27me2K36me1			
H3K27me2K36me2			
H3K27me3K36me0			
H3K27me3K36me1			
H3K27me3K36me2			
H3.3K27ac			
H3.3K27me0K36me0			
H3.3K27me1			
H3.3K27me1K36me1			
H3.3K27me1K36me2			
H3.3K27me1K36me3			
H3.3K27me2			
H3.3K27me2K36me1			
H3.3K27me2K36me2			
H3.3K27me3			
H3.3K27me3K36me1			
H3.3K27me3K36me2			
H3.3K36me1			
H3.3K36me2			
H3.3K36me3			
H3_54_63 unmod			
H3_54_63 K56me1			
H3_54_63 K56me2			
H3_54_63 K56ac			
H3_73_83 unmod			
H3_73_83 K79me1			
H3_73_83 K79me2			
H3_73_83 K79me3			
H3_73_83 K79ac			
H3_117_128 unmod			
H3_117_128 K122ac			
H4_4_17 unmod			
H4_4_17 K5ac			
H4_4_17 K8ac			
H4_4_17 K12ac			
H4_4_17 K16ac			
H4_4_17 K5acK8ac			
H4_4_17 K5acK12ac			
H4_4_17 K5acK16ac			
H4_4_17 K8acK12ac			
H4_4_17 K8acK16ac			
H4_4_17 K12acK16ac			
H4_4_17 K5acK8acK12ac			
H4_4_17 K5acK8acK16ac			
H4_4_17 K5acK12acK16ac			
H4_4_17 K8acK12acK16ac			
H4_4_17 K5acK8acK12acK16ac			
H4_20_23 unmod			
H4_20_23 K20me1			
H4_20_23 K20me2			
H4_20_23 K20me3			
H4_20_23 K20ac			

Appendix table 5.2: Top ten transcripts identified using "new Tuxedo protocol".

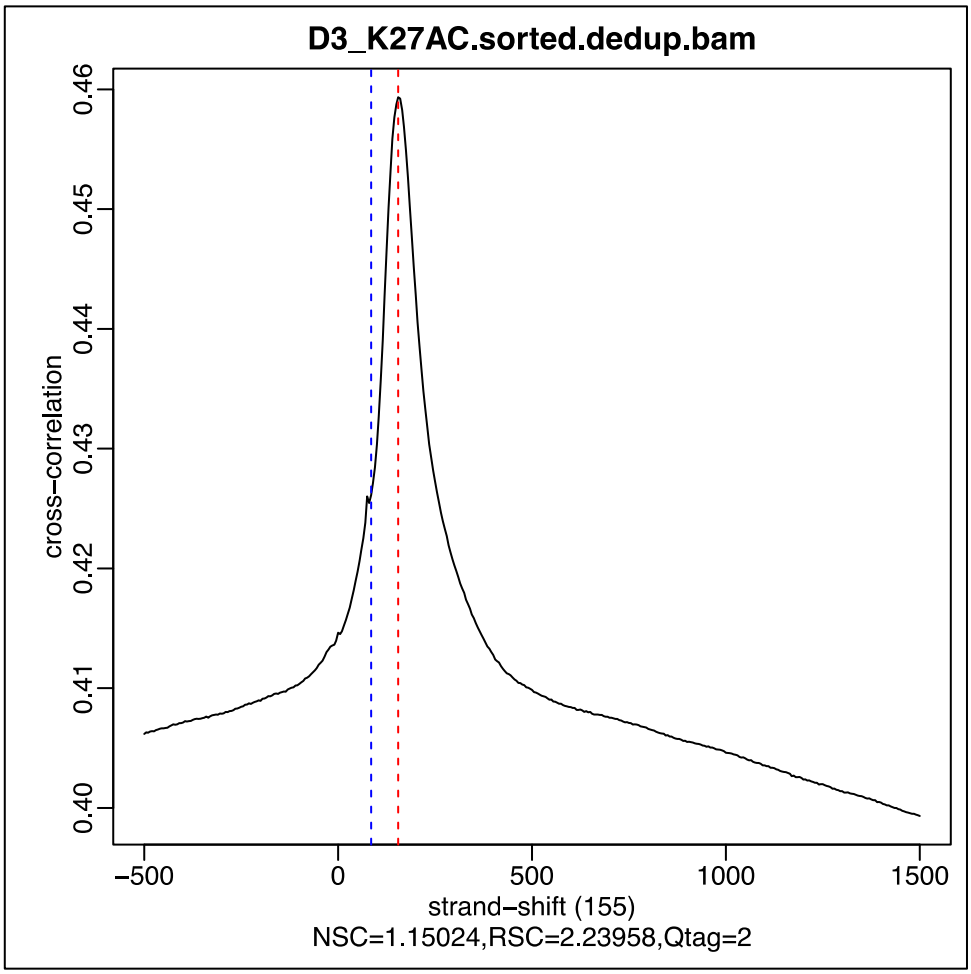
Transcript_Name	Log_FC	p value	q value	chr
uc059mdu.1	0.553223688	2.65E-06	0.1838	chr15
uc062xni.1	0.549128506	4.49E-06	0.1838	chr4
uc021yro.2	0.302302898	5.56E-06	0.1838	chr6
uc058tfo.1	2.887789605	6.20E-06	0.1838	chr12
uc001isy.4	-1.57550692	7.63E-06	0.1838	chr10
uc061zwi.1	-0.73943678	9.66E-06	0.1939	chr21
uc062jir.1	-0.31967312	1.64E-05	0.2328	chr3
uc058nws.1	0.28993914	1.78E-05	0.2328	chr12
uc011lhj.3	-0.80503139	1.85E-05	0.2328	chr8
uc060moz.1	-0.72478913	2.06E-05	0.2328	chr17_KI270862v1_alt



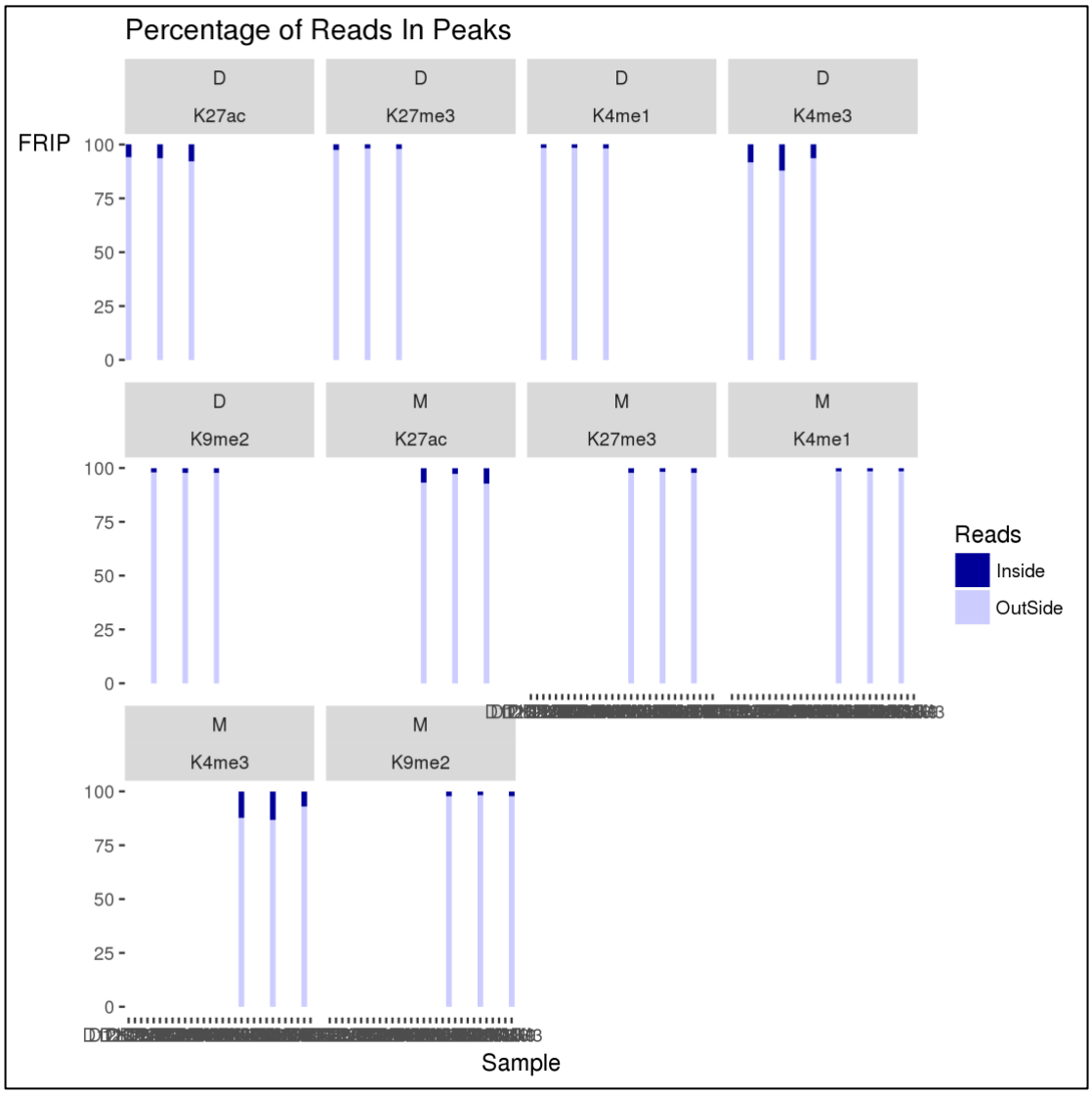
Appendix figure 5.1: Bioanalyser trace of ChIP following library preparation and size selection.

This figure illustrates the post library preparation and magnetic bead size selection and PCR clean up prior to sequencing for a selection of the pull downs.

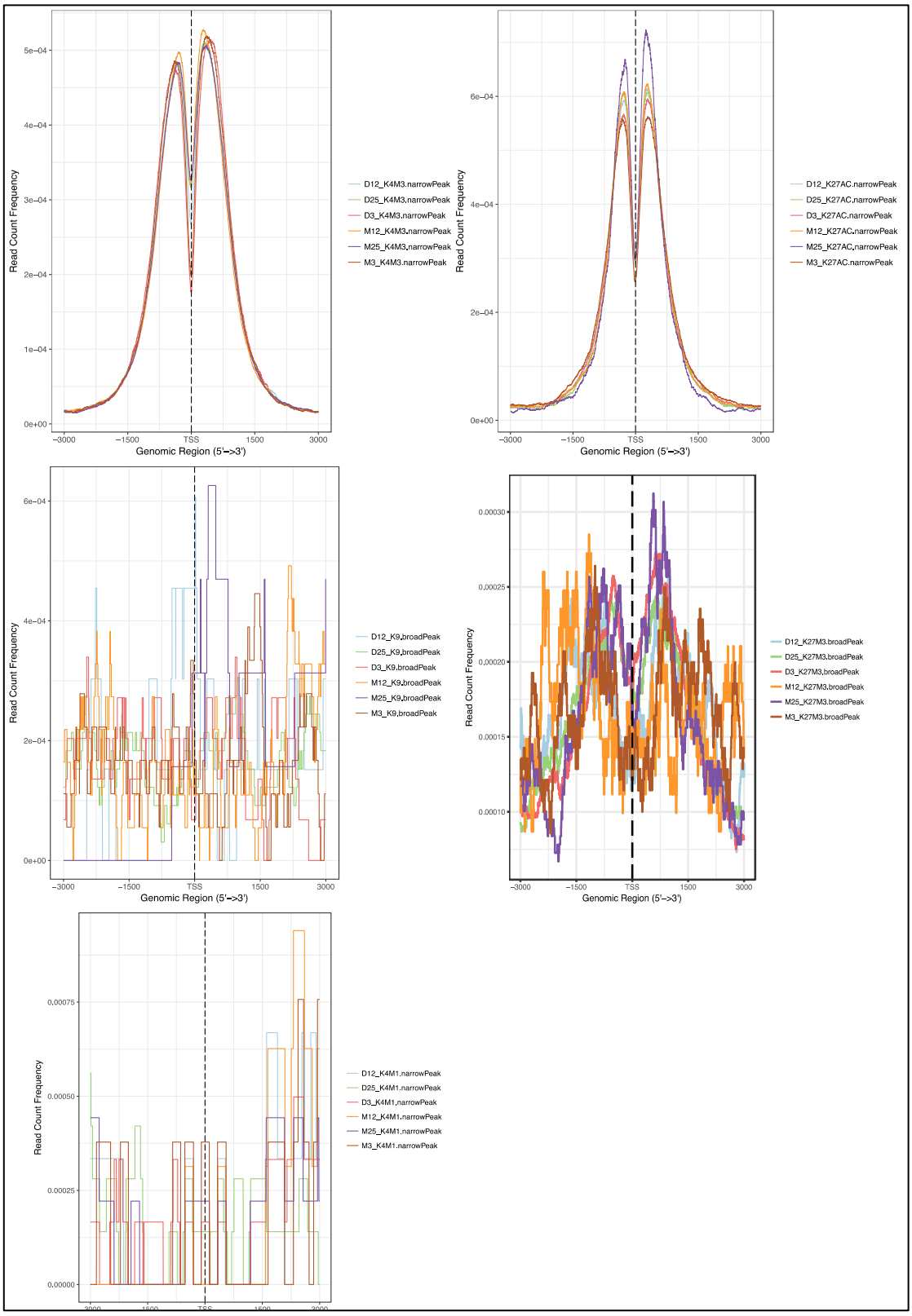
Appendix figure 5.2: FASTQC assessment of ChIP-Seq raw data.
This figure illustrates the results from the FASTQC analysis of the raw sequencing data obtained following ChIP-Seq. The mean Phred scores are greater than 25 and therefore of good quality.



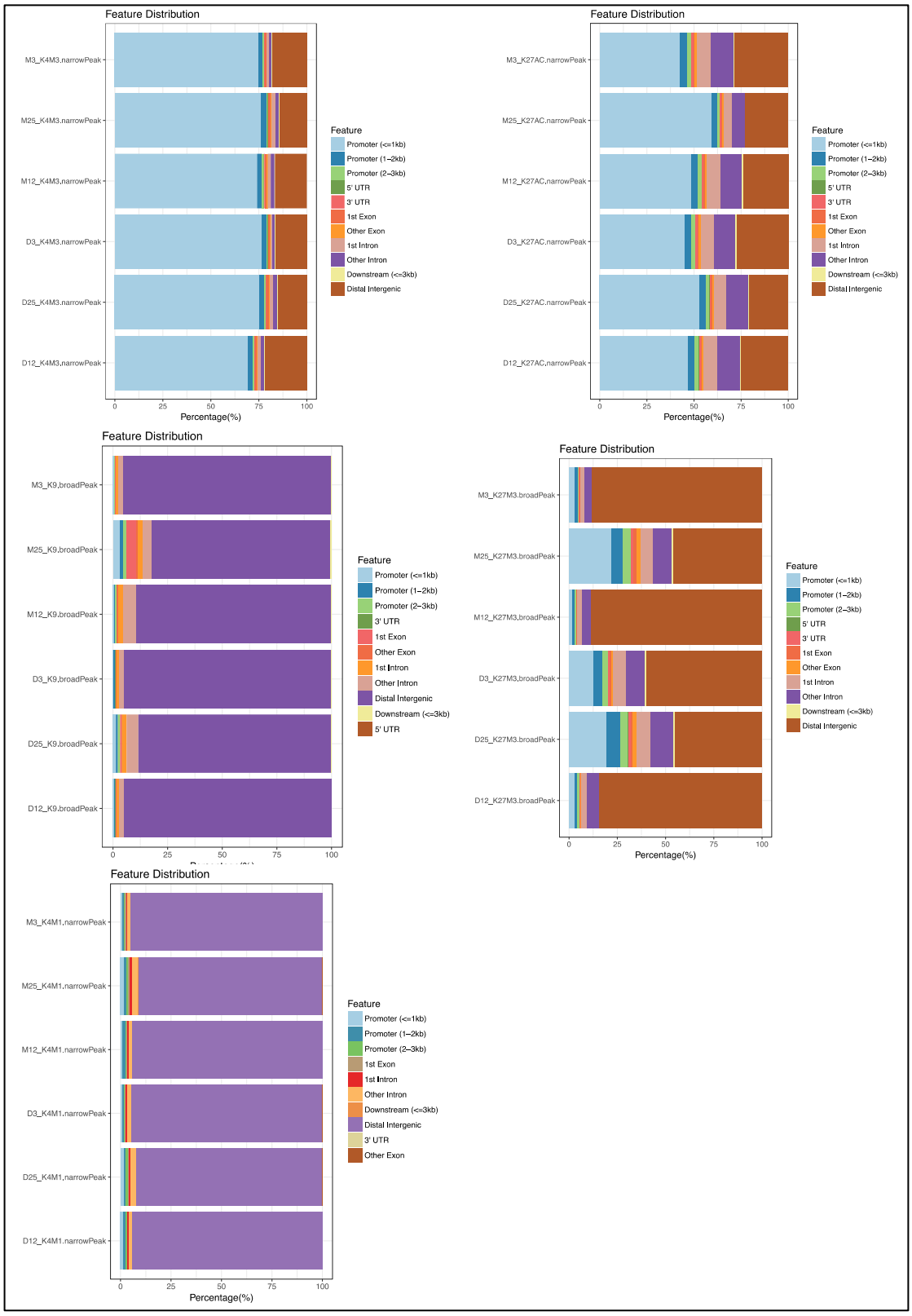
Appendix figure 5.3: Phantom peak tools assessment of the peak picking.
This figure illustrates the assessment of the peaks found after running MACS2 for the H3K27ac pull down in sample D3. Both the NSC and the RSC show good results.



Appendix figure 5.4: ChIPQC assessment of the peak picking.
This figure illustrates the abundance of reads found to be inside peaks after running MACS2 in narrow mode for each of the different ChIPs.



Appendix figure 5.5: Average profile of ChIP peaks binding to TSS regions.
This figure summarises the average profile of the ChIP peaks to the TSS regions for each pull down picked with MACS2 in either narrow mode (H3K4me1, H3K4me3 and H3K27ac) or broad mode (H3K9me2 and H3K27me3).



Appendix figure 5.6: Genomic annotation barplot of ChIP peaks.
This figure summarises the genomic annotations derived from ChIPseeker for each pull down picked with MACS2 in either narrow mode (H3K4me1, H3K4me3 and H3K27ac) or broad mode (H3K9me2 and H3K27me3).

Appendix table 5.3: Broad peak diffbind analysis for H3K4me1 pull downs.

Chr	Start	End	Fold	FDR	Gene		distance
chr18	**107169**	**110747**	**3.63**	**0.000222**	**ROCK1P1 /MIR8078**	**exon**	
chr7	**100957788**	**100958905**	**6.11**	**0.00129**	**MUC3A**	**exon**	
chr21	**9820916**	**9821846**	**4.16**	**0.0459**	**LINC01667**	**exon / TSS**	
chr21	**8219787**	**8220666**	**5.71**	**0.00575**	**NR_038958**	**intron**	
chr16	46389646	46391265	9.56	3.59E-08	Downstream ANKRD26P1	no	100kb
chr16	46394297	46395082	8.17	0.000222	Downstream ANKRD26P1	no	100kb
chr16	46386360	46389343	8.88	0.000236	Downstream ANKRD26P1	no	100kb
chr21	10692370	10693086	4.39	0.028	Downstream TPTE	no	100kb
chr21	***8420049***	***8420586***	***5.04***	***0.0461***	***Downstream NR_038958 / Upstream MIR6724-4***	no	10kb / 12kb
chr3	***75668894***	***75669628***	***7.82***	***1.70E-06***	***Downstream FRG2C / Upstream LINC00960***	no	10kb / 12kb
chr20	31062978	31064143	5.85	0.0025	Upstream DEFB115	no	120kb
chr20	28897883	28898806	4.74	0.0237	Upstream FRG1DP	no	180kb
chrY	11332971	11334272	7.76	1.70E-06	Downstream GYG2P1	no	1mb
chrY	11318134	11319045	5.48	0.00691	Downstream GYG2P1	no	1mb
chrY	26670402	26671899	2.33	0.0483	Downstream TTTY3B	no	1mb
chr1	143214220	143215839	6.91	0.000101	Downstream LOC645166	no	200kb
chr1	143251611	143256873	7.09	0.000146	Downstream LOC645166	no	200kb
chr1	143263626	143264795	7.01	0.00167	Downstream LOC645166	no	200kb
chr1	143184642	143186236	4.24	0.00735	Downstream LOC645166	no	200kb
chr16	34581396	34583815	8.95	0.000236	Upstream LINC00273	no	200kb
chr16	34594857	34595848	6.94	0.000269	Upstream LINC00273	no	200kb
chr16	34586505	34587005	4.59	0.000724	Upstream LINC00273	no	200kb
chr16	34592590	34593646	6.36	0.00645	Upstream LINC00273	no	200kb
chr16	34573910	34574451	6.05	0.0109	Upstream LINC00273	no	200kb
chr16	34575638	34576696	3.29	0.0412	Upstream LINC00273	no	200kb
chr2	89836283	89837201	6.81	0.000205	Upstream NR_136329	no	200kb
chrY	56734065	56734910	5.96	0.00167	Upstream SPRY3	no	250kb
chr10	41883103	41884055	4.75	0.000657	Downstream LOC441666	no	300kb
chr10	41879280	41879636	3.52	0.0239	Downstream LOC441666	no	300kb
chr15	17080543	17081670	4.16	0.011	Updtream CHEK2P2	no	3mb
chr4	49632669	49635611	6.43	0.00575	Downstream CWH43	no	400kb
chr22	12693437	12694023	6.01	0.00148	Downstream NR_110761	no	600kb
chr20	***31185346***	***31187381***	***6.63***	***0.000289***	***Upstream DEFB115***	no	70kb
chrUn_KI270438v1	103891	105254	4.18	1.41E-11			
chr1_KI270709v1_random	6608	7767	5.4	0.0322			
chr14_GL000225v1_random	5434	6978	7.61	4.53E-06			
chr14_GL000225v1_random	64968	66357	7.32	8.40E-06			

chr14_GL000225v1_random	130222	131970	7.38	1.45E-05			
chr14_GL000225v1_random	67256	68342	5.04	0.00063			
chr14_GL000225v1_random	24991	25825	6.29	0.00715			
chr14_GL000225v1_random	125125	126036	6.31	0.0017			
chr14_GL000225v1_random	44328	44605	5.43	0.0073			
chr14_GL000225v1_random	45973	46204	5.3	0.0146			
chr14_GL000225v1_random	90884	91263	5.03	0.0163			
chr14_GL000225v1_random	136994	138232	5.05	0.0171			
chr14_GL000225v1_random	85568	86617	3.71	0.0184			
chr14_GL000225v1_random	17000	17268	4.84	0.0237			
chr14_GL000225v1_random	56981	57477	4.79	0.0239			
chr14_GL000225v1_random	49794	51004	2.58	0.0438			
chr14_KI270723v1_random	35218	37122	6.41	0.00076			
chr14_KI270724v1_random	481	1334	5	0.0171			
chr17_KI270729v1_random	2325	3633	4	0.0057			
chr17_KI270729v1_random	21060	21556	5.53	0.0073			
chr22_KI270733v1_random	165119	166362	5.47	0.0073			
chr22_KI270733v1_random	162219	163702	5.12	0.0438			
chr22_KI270736v1_random	179486	181786	4.13	0.00022			
chrUn_GL000214v1	117073	118932	6.18	0.0008			
chrUn_GL000216v2	160603	161004	6.53	0.00029			
chrUn_GL000216v2	149357	149840	6.63	0.00029			
chrUn_GL000216v2	14889	16143	5.4	0.0080			
chrUn_GL000216v2	158642	159698	4.26	0.0101			
chrUn_GL000220v1	143882	145560	-5.61	0.0057			
chrUn_GL000224v1	3520	5750	7.89	1.50E-06			
chrUn_GL000224v1	505	2748	5.19	0.00022			
chrUn_GL000224v1	16760	19239	5.83	0.0025			
chrUn_KI270438v1	109177	112425	2.14	0.0077			
chrUn_KI270442v1	92121	92712	2.56	0.0327			
chrUn_KI270589v1	41617	42623	5.73	0.0035			
chrUn_KI270744v1	119601	121462	7.08	0.0014			
chrUn_KI270744v1	108139	109571	5.12	0.0015			

Appendix table 5.4: Diffbind analysis for H3K4me3 peaks.

ab	n	Chr	Start	End	Fold	FDR	Gene	Gene body	distance
H3K4me3 Broad	2	chrUn_KI270744v1	119089	121463	7.24	3.20E-05			
H3K4me3 Broad	2	chrUn_GL000214v1	116810	118662	6.13	0.0159			
H3K4me3 Broad	2	chr20	31064959	31065244	6.1	0.0332	Upstream DEFB115	no	200bp
H3K4me3 Broad	2	chr4	49140532	49140889	3.34	0.0432	centromere / downstream CWH43	no	60kb
H3K4me3 Broad	3	chrUn_KI270744v1	119089	121463	6.45	0.00218			
H3K4me3 Broad	3	chrUn_GL000214v1	131146	131864	5.2	0.00691			
H3K4me3 Broad	3	chr20	31064959	31065244	5.07	0.00594	Upstream DEFB115	no	200bp
H3K4me3 narrow	2	chrUn_KI270744v1	119598	120332	6.41	0.00223			
H3K4me3 narrow	2	chrUn_GL000216v2	166609	166932	6.3	0.00366			
H3K4me3 narrow	2	chr14_GL000225v1_random	74454	74902	-4.58	0.0355			
H3K4me3 narrow	2	chr2	89836279	89836820	7.39	1.42E-05	Upstream NR_136329	no	200bp
H3K4me3 narrow	2	chr20	31060276	31060908	7.14	0.00223	Upstream DEFB115	no	200bp
H3K4me3 narrow	2	chr20	31064954	31065293	5.14	0.0174	Upstream DEFB115	no	200bp
H3K4me3 narrow	2	chr1	143184636	143185468	2.08	0.0331	Downstream LOC1645166	no	200bp
H3K4me3 narrow	2	chrY	56706930	56708146	6.37	0.00284	Upstream SPRY3	no	250kb
H3K4me3 narrow	2	chr3	93470363	93470799	12.68	0.00642	centromere / downstream PROS1	no	400kb
H3K4me3 narrow	2	chr4	49120883	49121208	3.33	0.0475	centromere / downstream CWH43	no	60kb
H3K4me3 narrow	3	chrUn_KI270744v1	119598	120332	5.64	0.0444			
H3K4me3 narrow	3	chr20	31064954	31065293	4.62	0.0131	Upstream DEFB115	no	200bp
H3K4me3 narrow	3	chrY	56706930	56708146	5.91	0.000446	Upstream SPRY3	no	250kb