Epigenetic Regulation of the Myeloid Cell Lineage

Laurynas Pliuškys

Pembroke College

Nuffield Department of Orthopaedics,

Rheumatology and Musculoskeletal Sciences

University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

Michaelmas Term, 2014

Epigenetic Regulation of the Myeloid Cell Lineage

Laurynas Pliuškys

Pembroke College

Nuffield Department of Orthopaedics,
Rheumatology and Musculoskeletal Sciences

Abstract

The myeloid cell lineage is a fundamental element of the immune system and it can give rise to a diverse set of terminally differentiated cells, such as macrophages or osteoclasts among many others. Mutations or misregulation of gene expression may lead to severe clinical conditions, such as arthritis, osteoporosis or cancers. Epigenetics, the regulation of gene expression and chromatin remodelling, is implicated in cell differentiation, function and disease, and hence it is a promising new area to explore in order to explain underlying cellular mechanisms.

Firstly, human macrophage subtypes were studied. Chemokine (C-C motif) ligand (CCL) 1 and mannose receptor were validated to be granulocyte macrophage (GM) colony stimulating factor (CSF) induced macrophage markers, while CCL2 was specifically expressed in macrophage CSF (MCSF) macrophage population. By utilising publicly available high-throughput sequencing data, new biomarkers dehydrogenase/reductase (SDR family) member 2 and CCL26 were discovered to be MCSF-macrophage specific while guanylate binding protein 5 and apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A were highly up-regulated in GMCSF cells.

Secondly, a range of gene knock-down techniques for the myeloid cell lineage were optimised and established. Lentiviral short-hairpin RNA (shRNA) delivery methods were shown to induce an undesirable pro-inflammatory response in macrophages. Furthermore, the frequently utilised cytomegalovirus promoter for gene expression was shown to be completely silenced in macrophage populations. Locked nucleic acids were selected as a suitable alternative to shRNA knock-down and by employing this new tool it was shown that a histone demethylase lysine (K)-specific demethylase (KDM) 6B is fundamental for macrophage differentiation.

Finally, a small molecule GSK-J4, a potent inhibitor of histone demethylases KDM6A, KDM6B and KDM5B specific for H3K27me3 and H3K4me3, respectively, was used to dissect epigenetic signalling in osteoclasts and multiple myeloma. In osteoclasts it was shown to act mainly by inhibiting transcriptional changes required for osteoclastogenesis when MCSF-macrophages are stimulated with Receptor Activator Of Nuclear Factor Kappa-B Ligand (RANKL), as indicated by the differential increase in H3K27me3 marks, leading to inhibition of c-Jun and potentially abolition of transcription factor AP-1, required for the transcriptional initiation of nuclear factor of activated T-cells 1 (NFATc1). In multiple myeloma cells, GSK-J4 causes a dramatic increase in expression, further supported by the build-up of global H3K4me3 marks, which results in the upregulation of the unfolded protein response pathway. In both cell systems, there is an early upregulation of metallothionein genes, which in multiple myeloma was shown to increase potentially due to rapid influx of zinc ions within the first 30 minutes, and as such may cause

induction of apoptosis in multiple myeloma and may inhibit differentiation of osteoclasts.

Contents

Acknowledgements

This doctorate work would not have been possible without the support and guidance of many people. Firstly, I would like to thank my supervisor Professor Udo Oppermann whose expertise, advice, trust in my capabilities and numerous hours discussing my work helped me successfully complete my thesis work. I am also indebted to my joint supervisors Professor Chris Ponting from the Department of Physiology, Anatomy and Genetics at the University of Oxford, as well as Dr Nicola R Harker and Dr Rabinder Prinjha from GlaxoSmithKline, for their time discussing my progress each term and offering their network of colleagues for further help with my work. I am also grateful to our collaborators Dr Laurens Kruidenier and Dr Palwinder Mander from GlaxoSmithKline for turning our collaboration on macrophage biomarkers into a fruitful project with a potential future publication.

I appreciate the long hours invested by Dr Marco Antonio Mendoza Parra, Dr Shankara Narayanan Pattabhiraman and Dr Mannu Kamalraj Walia in troubleshooting the chromatin immunoprecipitation protocol, made possible by Professor Hinrich Gronemeyer at the Institut Génétique Biologie Moléculaire Cellulaire in Strasbourg, France, who hosted me on multiple occasions. I thank Dr Giuseppe Gallone and Mr Harry Clifford for their time spent sharing their experience on processing of high-throughput sequencing data with me. I am particularly indebted to Mr Marco Licciardello for the effort instructing me on lentiviral transfection optimisation and the subsequent quality controls, made possible by the kind invitation from Dr Stefan Kubicek to visit his research group

based at the Research Center for Molecular Medicine of the Austrian Academy of Sciences in Vienna, Austria. Finally, I thank Dr Edyta Marcon and Professor Jack Greenblatt from the University of Toronto for hosting me in their laboratory and teaching me CRISPR-Cas9 gene knock-down methodology.

I am grateful for the everyday support, friendliness, wit and useful discussions to members of our research group, in no particular order, Dr James Dunford, Kelly Rooke, Dr Edward Hookway, Dr Anneke Kramm, Dr Fernando Martinez Estrada, Dr Clarence Yapp, Dr Adam Cribbs, Dr Peter Cain, Dr Ka Hing Che, Graham Wells, Na Wu, Dr Martin Philpott, Emma Jane Goode as well as the researchers at the Botnar Research Centre Dr Seint Lwin, Dr Srinivasa Rao and Zuzana Kalivodova.

This doctorate would have never been possible without the incredible work done by the team at the Doctoral Training Centre, namely Professor David Gavaghan, Professor Charlotte Deane, Professor Elspeth Garman, Lisa Bligh, Samantha Miles, Francesca Summers, and many others. I am also grateful to the Engineering and Physical Sciences Research Council as well as GlaxoSmithKline for fully funding my student fees and supporting me through a generous scholarship.

The attentive friendship by Ross Purple, Matthew Gibbs, Dr Oscar Cordero Llana, Gabriel Schenk, Richard Reid, Milda Petkevičienė, Gabrielė Subačiūtė, Mantas Kalvelis, Damien Merlo, Dr Hanif Javanmard, Jason Andrew Kaminski, Jonah Macan and many more who are not named here, has given me the strength to cope in tough times and made the life in Oxford and other places during my research visits exciting and memorable.

Finally, I bow to my parents Erika Kulikaitė-Gintalienė and Dainius Gintalas for their love and care, for their limitless effort and patience raising me as a happy, motivated and curious person, for their encouragement, confidence and trust in my life decisions and for their financial support during my studies.

Attributions

The locked nucleic acids (LNAs) were designed by Dr Morten Lindow, synthesised by Dr Yann Tessier and the collaboration was efficiently managed by Susanna Obad of Santaris Pharma A/S in Copenhagen, Denmark.

The ideas for validating macrophage biomarkers Activin A and chemokine (C-C motif) ligand 2 were initially explored and shared with me by Dr Palwinder Mander and Dr Laurens Kruidenier from GlaxoSmithKline.

The blood from anonymous volunteers was used in accordance with the terms of the informed consent form and collected ethically and painlessly by Kelly Rooke from our research group.

The training for confocal and high-throughput microscopy was kindly provided by Dr Clarence Yapp.

The extensive collaboration with Dr John Christianson and Federica Lari from the Ludwig Institute of Cancer Research at the University of Oxford, Dr Michelle Reed and Professor Ulrich Günther from the Wellcome Trust Center for Biomolecular NMR at the University of Birmingham, as well as Dr Charlotte Pawlyn from the Institute for Cancer Research London helped validate the findings from ChIP-Seq and RNA-Seq experiments in multiple myeloma JJN3 cell line treated with GSK-J4.

Associated Publications

Ka Hing Che*, **Laurynas Pliuskys***, Clarence Yapp, Ed Hookway, Fernando O Martinez, James Dunford, Kelly Rooke, Takeshi Kashima, Adam Cribbs, Laurens Kruidenier, Rab Prinjha, Susanna Obad, Morten Lindow, Henrik Oerum, Hinrich Gronemeyer, Kalle Soderstrom, Nick Athanasou, Udo Oppermann. Histone demethylases control RANKL programs required for formation and function of human osteoclasts. *Manuscript in preparation.*

* Equally contributing authors

Edward Hookway*, Federica Lari*, Peter Cain*, **Laurynas Pliuskys***, Charlotte Pawlyn*, Michelle Reed*, Deborah Brotherton*, Faith Davies, Anthony Tumber, KaHing Che, Ralph Mazitschek, Cansu Kucukkose, Duygu Koca, Ferah Gulacti, Fernando Martinez, Cyrille C. Thinnes, Radoslav Novak, Christopher J Schofield, Daniel Korr, Cora Mund, Dave Wilson, Rab Prinjha, Claire M. Edwards, Giuseppe Gallone, Chris P. Ponting, Paul E. Brennan, Mark Bunnage, Alison O'Mahony, Hinrich Gronemeyer, Ulrich L Günther, Morten Lindow, Henrik Ørum, Susanna Obad, Gareth Morgan, Nick Athanasou, John C Christianson, Udo Oppermann. Activation of the Integrated Stress Response by a Jumonji Histone Demethylase Inhibitor Induces Pro-Apoptotic Metabolic Reprogramming in Multiple Myeloma. *Manuscript in preparation*

* Equally contributing authors

Chapter 1

Introduction

In terms of numbers, blood cells comprise about a third of all human cells and play a number of vital functions in higher organisms. While it can be disputed whether blood is an organ or a tissue it is arguably fundamental to the function of human body by transporting nutrients and metabolic waste, vigilantly preventing infection and repairing damaged tissues. The huge diversity of functions carried out by different cell types, mostly arising from the same progenitor cells, exist because of complex regulatory mechanisms of gene expression that determine the fate and specialisation of cells. While many mechanisms are redundant, errors in the cellular machinery may lead to various disorders, diseases and malignancies. Throughout this doctorate thesis there is a lot of emphasis on the immunity aspect of blood macrophages, bone regeneration, and the malignancy of multiple myeloma, and how all these phenomena are interlinked and regulated on an epigenetic level. Importantly, significant amount of time has been spent establishing methods and techniques to study the aforementioned processes in primary human cells.

1. The myeloid cell lineage

Myeloid cells are part of the haematopotietic cells. Haematopoietic cells arise from a common progenitor, the multipotent hematopoietic stem cell (HSC) that resides

in the bone marrow. While regenerating themselves, HSCs also produce an estimated 1 trillion cells on a daily basis in the average human that will differentiate into more specialised cells [1]. The important question is how do HSCs "decide" whether to regenerate or to differentiate into one of the numerous haematopoietic cell types.

Early studies were lacking specific markers for HSCs and downstream cell types, so the formation of colonies of differentiated cells were used to derive a model of HSC renewal versus downstream specialisation [1]. A clear cumulative gamma distribution, rather than a Poisson distribution, was observed when the cumulative number of colony forming units (CFUs) was plotted against the number of CFUs within those colonies – i.e. they are the second generation HSCs within those colonies that retained their ability to regenerate and can generate new colonies if the experiment was to be repeated [2]. Based on this observation, a gamma distribution of probabilities for the cells to retain HSC properties and to regenerate was assumed. As such, a mathematical model, based on a set of experimentally determined parameters, was designed and tested using the Monte Carlo method. The results of the simulation highly correlated with the observation of CFUs and so a model of stochastic variation of HSC differentiation was born [2].

However, now there are markers established for most haematopoietic cell types, including their intermediate states. This allowed more accurate cell type definition, which also led to the conclusion that haematopoietic cells differentiate in a binary step-wise fashion [3]. At each subsequent downstream step there are only two cell

fate choices defined by the availability of extracellular receptors, the surrounding cytokine environment and active nuclear transcription programmes.

The first binary step for a HSC is to "decide" whether to start differentiating or to remain a HSC and regenerate. HSCs were originally defined by a set of the following markers: expression of Sca-1 (Spinocerebellar Ataxia 1) and Mac-1 (Macrophage-1 antigen, consisting of CD11b and CD18), absence of CD4 (cluster of differentiation 4) and Lin (lineage markers, a mix of CD3, CD14, CD19 and CD56) and low expression of Thy-1 (Thy-1 cell surface antigen), also briefly denoted as Thy-1^{low}Sca-1^{+}Lin^{-}Mac-1^{+}CD4^{-} [4]. Subsequently, by taking into consideration a few more markers, a later study revealed the complexities of HSCs: long-term, short-term and multipotent progenitor [5]. Flk-2 was found to be a marker defining the longevity of HSCs – long term HSCs were all negative for Flk-2, while HSCs expressing Flk-2 after injecting into mice stopped regenerating and giving rise to new haematopoietic cell lineages after 8-12 weeks. Finally, HSCs lacking Thy-1 appear to be completely functionally more mature which would indicate they are downstream from short-term HSCs as they express Flk-2 (Figure 1).

Once HSCs have become a multipotent progenitor they are to develop into either a common myeloid progenitor (CMP) or a common lymphoid progenitor (CLP) [6, 7]. The CLPs further differentiate into B cells, T cells and natural killer cells, however, these cell types

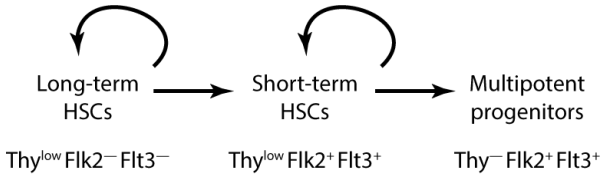


Figure 1 The change in marks of HSCs progressing downstream the differentiation path. The appearance of Flk2 and Flt3 signify stem cells ready to differentiate rather than regenerate.

are outside the scope of this thesis. The CMPs once again have a binary choice – they either turn into granulocyte-macrophage or erythrocyte-megakaryocyte progenitors (GMP and EMP, respectively) [6]. Even though these binary switches may appear somewhat artificial, each cell population has been described with a set of biomarkers. For example, the CMPs in a backward reductionist approach were described as $FcγR^{lo}CD34^{+}$ (low Fcγ receptor-II/III, positive for CD34) [6]. Finally, GMPs then have a further binary choice of differentiating into granulocytes or macrophages.

While the surface biomarkers used to define various cell populations are descriptive, sometimes they lack mechanistic reasoning. Early in development, progenitor cells of the myeloid cell lineage start expressing receptors for GMCSF (granulocyte-macrophage colony-stimulating factor), MCSF (macrophage colony-stimulating factor) and GCSF (granulocyte colony-stimulating factor) [8]. However, while signalling of these cytokines has a significant effect on proliferation and maturation of cells, they are not essential for lineage commitment, as has been show in mice studies. Only MCSFR (MCSF receptor) deletion was lethal to the hematopoietic lineage commitment, while GCSFR (GCSF receptor) and GMCSFR (GMCSF receptor) deletion did not eliminate the lineage [9, 10]. This finding suggests there may be other important factors involved that determine the differentiation path.

Since lineage commitment requires vast changes in gene expression, one group of targets to look at is transcription factors (TFs). SPI1 (spleen focus forming virus (SFFV) proviral integration oncogene, also known as PU.1) was identified as a vital transcription factor for the commitment to the whole hematopoietic lineage, since homozygous deletion of the gene results in the loss of myeloid and lymphoid cells

[11]. While SPI1 deletion was embryonic lethal in mice in the latter study, a different group showed that SPI1 knock-out mice are born but die shortly of septicaemia and their lifespan can only be extended for a short time by administering high doses of antibiotics. Again, those mice were shown to be devoid of B cells, T cells, neutrophils and macrophages [12]. On the other hand, both studies indicate the presence of fully functional erythroid and megakaryocyte populations which may hint at SPI1 involvement in later differentiation stages of the affected populations rather than initial lineage commitment.

C/EBPα (CCAAT/enhancer binding protein alpha) is another transcription factor and it is expressed exclusively in myeloid cells. Mice with a homozygous knock-out were shown to lack granulocytes and their downstream cells, neutrophils and eosinophils, but other myeloid cell types were unaffected [13]. Even more surprisingly, while the promoters of GMCSFR, MCSFR and GCSFR genes contain a C/EBPα binding site, only GCSFR was absent in C/EBPα double knock-out animals. This result indicates a redundancy in the activation of monocytic gene expression programmes.

ICSBP (interferon consensus sequence binding protein 1) is yet another transcription factor, however this one is predominantly expressed in macrophages, B and induced T cells, especially upon induction with IFN-γ [14]. Bone marrow cells that have the ICSBP gene knocked-out have significantly reduced formation of macrophages in response to GMCSF or MCSF but have increased numbers of granulocytes upon stimulation with GMCSF and GCSF [15]. ICSBP-negative macrophages are unable to produce nitric oxide or reactive oxygen species required for the killing of

pathogens [15]. Finally, myeloid cells that lack ICSBP show impaired apoptosis which may explain the defects in the myeloid cell lineage. The same authors even propose ICSBP may act as a master switch between macrophage and granulocyte commitment (Figure 2).

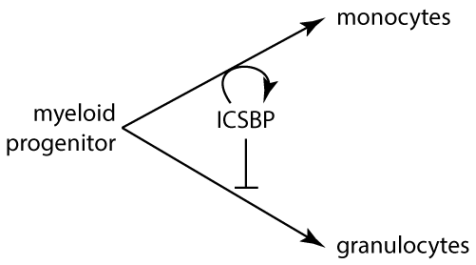


Figure 2 The proposed role of ICSBP acting as a master switch for myeloid progenitor cell fate commitment.

2. Macrophages and their subtypes

Monocytes are the precursors of macrophages which circulate in the blood stream and have a well-defined cell surface marker called CD14 (cluster of differentiation 14) which allows isolation of an apparently pure cell population as a starting point [16]. Monocytes constitute about 10% of all blood leukocytes and due to the large number of TLRs (toll-like receptors) on the cells surface they are able to respond rapidly to foreign antigens [17]. Once a stimulus is identified by monocytic cells, they can migrate from the blood stream to the tissue containing the chemo-attractant. Monocytes can then be differentiated into pro- or anti-inflammatory macrophages, or in other terms M1 and M2 macrophages that borrow their respective classification from Th1 and Th2 helper cells [18].

The pro-inflammatory M1 macrophages will respond to mostly tumour, viral or bacterial stimulation and cause acute inflammation. TLRs, as mentioned before, are instrumental to this response and lead to the induction of two pathways. Upon

stimulation with bacterial agents, TLRs form dimers which, through a series of adapters, induce MYD88 (myeloid differentiation primary response 88) [19]. Further downstream, MYD88 activates TRAF6 (TNF receptor-associated factor 6) that induces a MAP kinase leading to the inhibition of Inhibitor κB (IκB) [19]. The MYD88-independent pathway induces a delayed NF-κB (Nuclear factor-κB) response [20] through IRF3 signalling [21] leading to the activation IFN-β and its downstream genes [21].

Nuclear factor-κB (NF-κB) is now accepted to be the main player involved in the pro-inflammatory signalling cascade [22]. NF-κB is the name of a family of transcription factors that all share a Rel homology domain. NF-κB is found in a complex with IκB in the cytoplasm, where upon a pro-inflammatory stimulus, such as LPS or cytokines, NF-κB is released after phosphorylation of IκB and immediately translocates to the nucleus. Being a transcription factor, NF-κB binds to various pro-inflammatory target genes which result in up-regulated inflammatory responses by the cell, such as the release of cytokines TNF-α (tumor necrosis factor), interleukin (IL) 6 and IL12-p40 (interleukin 12, subunit p40) [23] which in turn recruit other cells and induce the pro-inflammatory macrophage differentiation into the M1 phenotype.

The anti-inflammatory M2 macrophages are induced by parasitic worm infections or tissue damage and facilitate wound healing by preventing inflammation [24]. However, the induction of this phenotype has been somewhat more difficult to elucidate [25]. The stimulation with IL-4 and IL-13 induces an increased expression of the mannose receptor [26] which is conversely down-regulated by IFNγ [27]. IL-4

and IL-13 can be produced by a variety of non-haematopoietic cells signalling innate problems such as cell death, injury but not infection [25]. Finally, downstream of IL-4 and IL-13, STAT6 (signal transducer and activator of transcription 6), along with other epigenetic modulators, induces the transcription of various anti-inflammatory responses, such as the expression of chemokines and associated receptors, some of which lead to stimulation of Th2 cells [28].

There is a number of methods established for monocyte to macrophage differentiation in vitro [29]. The development into these two distinctive states (classically and alternatively activated macrophages) is achieved by culturing the monocytes with GM-CSF for M1 or MCSF for M2 [30]. On the other hand, there have been contradictory claims that the latter differentiation protocol would only yield resting macrophages, rather than induced ones, and further induction with Interferon gamma (IFNγ) is required for M1 while Interleukin (IL) 4 or IL-13 is necessary for the M2 phenotype [31]. This ongoing debate supports the idea that macrophages are a heterogeneous population of cells even under uniform cell culture conditions and it requires careful study of various sub-populations and in-depth analysis of underlying signalling pathways and epigenetic mechanisms that constitute the differentiation processes.

3. Why epigenetics?

Epigenetics was originally mentioned in 1942 by Conrad Hal Waddington as a biologic phenomenon leading to gene regulation that affects development and cell

fate determination. Interestingly, the term was defined based on the general observation of cell differentiation, even before the discoveries of the DNA double helix and genes were made. Later, the term re-surfaced again when DNA methylation was described as an interesting biological phenomenon that allows passing information related to the phenotype to the next generation without any alterations in the genetic code itself [32].

The current definition of the term epigenetics covers more broadly the regulation of gene expression that is caused by various covalent modifications of the histones as well as DNA methylation. Initially DNA wrapping around the nucleosomes was thought to be just a simple mechanism to tightly pack DNA inside the nucleus (Figure 3A), but it has become evident that it plays a fundamental role in gene expression [33]. The diversity of histone modifications led to the theory of a "histone code" that surrounds the genetic code and exerts its control on the expression of genes in close proximity through combinations of histone tail modifications that attract various enzyme complexes as well as reveal transcription factor consensus binding sites [34].

Earlier in this chapter a number of transcription factors influencing the determination of cell fate were described. However, there is evidence that genes may ultimately be expressed due to a supportive histone code, which only then allows binding of transcription factors [35]. Certain histone marks fall into categories of inducers and repressors, i.e. they are known to be found around genes that are expressed or silenced. By using spatial information, the distance from the transcription start site, it was discovered that individual histone marks or a small

selection of them, can accurately predict the state of expression of any gene [36]. Using machine learning algorithms, a mathematical model was built to describe this phenomenon. By employing similar methods, a different model was built based on the same spatial information for transcription factors [35]. This revealed that the model based on histone marks performs significantly better than that based on transcription factors. However, combining those two models into one did not improve the prediction accuracy [35], implying that transcription factors do not provide any additional information required to predict the degree of gene expression. As a result, while it is obvious that transcription factors are required for the expression of genes, chromatin modifications may hold the key to the basis of regulation of gene expression. Interestingly, models built on cells from the same developmental stage or cell type have about 90% prediction accuracy within the same category and the same models applied on a different cell type or developmental stage have prediction power of around 80% [36], indicating that while the processes are mostly the same, other factors may be playing a role, or simply different histone marks become more significant within different cell types.

There is a number of different DNA and histone modifications as well as RNA constructs, that all fall under the umbrella of epigenetic regulation. Furthermore, there are various effector proteins that modify or read those marks, or, in conjunction with RNA constructs, degrade the target transcripts. The field is evolving rapidly with new proteins and entire mechanisms being discovered on a regular basis. Next only the most studied histone marks, acetylation and methylation, will be reviewed, as the latter modification is also the focus of this thesis.

4. Epigenetic modifications: histone lysine acetylation

Histone acetylation is probably the most studied histone mark due to the early discovery of histone acetyl-transferases (HATs) and histone deacetylases (HDACs) [37]. Most N-terminal lysines of histones 3 and 4 (H3 and H4, respectively) can be acetylated (Figure 3B) and it was observed that acetylated histones are associated with euchromatin while hypoacetylation is a signature of heterochromatin. Further investigations showed that it does not just associate with an open chromatin state, but it is directly linked to active transcription of genes [38].

4.1. Histone lysine acetyl-transferases

Five protein families are implicated in the acetylation of lysine residues, however only three of those are directly linked to histone acetylation leading to regulation of gene expression [39]. A family of HATs related to GCN5 (general control non-derepressable 5, also known as KAT2A) was first described in yeast as an important

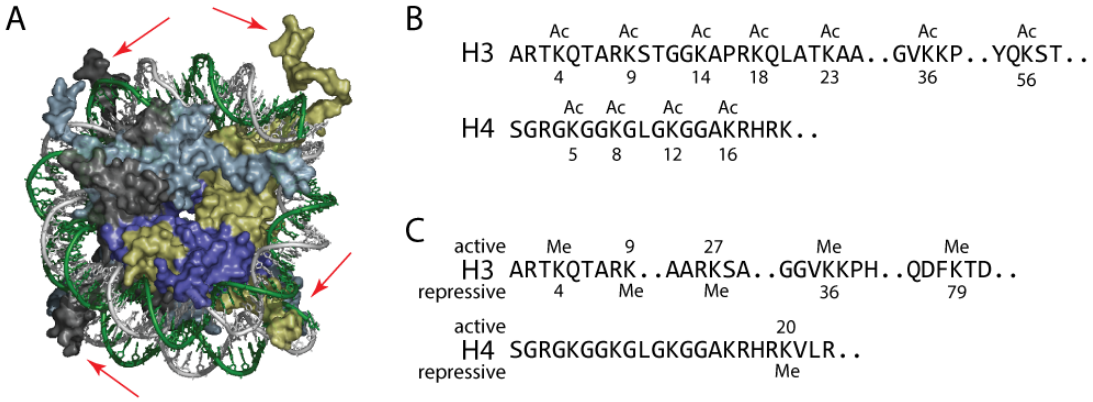


Figure 3 A. A top view of nucleosome shown as surface coloured by subunit type with DNA wrapped around it, shown as sticks with cartoon tube coloured by strand. Red arrows indicate the N-terminal ends of histone proteins that can be modified. Protein data bank accession number 1EQZ. B. The lysine residues on Histone 3 and 4 N-terminal ends that can be acetylated which leads to active transcription. C. The lysine residues on Histone 3 and 4 N-terminal ends that can be methylated which leads to either active or silenced transcription as indicated on the image.

activator of gene expression [40]. HATs from this superfamily, while being extremely similar to each other, have specific consensus sequences of amino acids that they recognise on the histone, including other surrounding modified residues, which leads to individual HATs acetylating only specific residues on the histone tails [33]. Another family called MYST (based on MOZ, Ybf2/Sas3, Sas2 and Tip60) is known to have a wider range of cellular function, including silencing of gene expression. Genes SAS2 and SAS3 (something about silencing 2 and 3, respectively) were shown to be fundamental to telomere silencing in yeast. SAS2/3 homologues in humans are Tip60 (HIV-1 Tat interacting protein 60) and MOZ-CBP (monocytic leukaemia zinc finger – CREB binding protein). The latter is known to be associated to chromosomal region chr8p11 which is frequently observed as a breakpoint of a chromosomal translocation in acute myeloid leukaemia [41]. On the other hand, MORF, a homologue of MOZ, and a member of the MYST family, was shown to be important for Runx2 activation despite the presence of a weak translational silencing domain [42]. The latter observations lead to the assumption of a complex transcriptional co-regulatory mechanism. Finally, p300/CBP family contains only those two members, which have been observed to have tumour suppressor effects in cells and they are also responsible for transcriptional activation [43].

4.2. Acetyl-lysine "reader" proteins

Once acetyl marks are established on histone proteins, a number of "reader" molecules can recognise those and then recruit other effector proteins, effectively forming large gene regulatory complexes around those epigenetic marks. Proteins that recognise acetyl marks, all harbour a bromodomain that specifically recognises

this modification [44]. In fact, there is a huge diversity of bromodomain-containing proteins and with 61 different domains available in the human genome, it has evolved to recognise a whole range of flanking modifications or nearby residues [45]. Furthermore, these are normally large proteins containing multiple domains, some of which may be recognising other modifications, such as histone methylation. For example, BPTF (Bromodomain PHD Finger Transcription) binds specifically to an acetylated lysine in the 16th position from the N-terminal end of histone 4 (H4K16ac), but it also contains a PHD domain that recognises tri-methylated lysine 4 on histone 3 (H3K4me3). It was shown to have much higher binding affinity to a peptide that has both modifications present, rather than just one [46]. This is just one example of many, showing how a previously mentioned "histone code" might work.

4.3. Histone deacetylases

HDACs are responsible for removing the acetyl marks from histone lysines, and they also carry a bromodomain which is initially used to find these acetylated residues. There are four classes of HDACs that were assigned by homology and phylogeny. Classes I, II and IV, the latter found only in higher organisms, rely on Zn for acetyl hydrolysis while Class III HDACs employ NAD as a cofactor for deacetylation [47]. Once again, HDACs, just like HATs and bromodomains, are specific for certain lysines and surrounding residues, and being multi-domain proteins they can harbour other domains or recruit proteins that further affect the histone marks. HDACs are known to interact with HP1 domains which after deacetylation further induce heterochromatin formation by recruiting lysine methyl-transferases specific for the

repressive H3K9 residue, which can eventually lead to even more long-term repression through DNA methylation [48].

To sum up, histone lysine acetylation is usually seen as a transcription activating mark. The diversity of domains and combinations of various domains within a protein support the "histone code" hypothesis and let us assume a more complex interaction between the different epigenetic marks.

5. Epigenetic modifications: histone lysine methylation

Histone lysine methylation, in comparison to acetylation, has much more varied effects on regulation of transcription (Figure 3C). The complexity arises from the fact that a lysine can have up to three methyl groups added onto a single lysine residue. Also, different methylated residues have been shown have either activating or repressing outcomes, contrary to the acetylated lysines which are assumed to be activating in most scenarios, as discussed before.

5.1. Lysine methyl-transferases

Of particular importance and probably the best studied histone lysine methyl marks are on H3K4, H3K9, H3K27 and H3K36. One of the first lysine methyl-transferases (HMT) discovered were proteins containing the SET domain (names after three founding members: SUV3-9, E(Z) and TRX). In *Drosophila* SUV3-9, containing what was later defined as the SET domain, was a suppressor of position effect variegation [49] – a phenomenon which gives rise to mosaic-coloured eyes due to misplacement

of a gene next to heterochromatin. The human homologues SUV39H1 and SUV39H2 also contain the SET domain and were shown to be specific for H3K9 tri-methylation [50]. Furthermore, H3K9me3 was shown to be associated with heterochromatin and there was tight cross-talk identified with a previously mentioned HDACs.

SETD1A (SET domain containing 1A), a homologue of the yeast SET1, is another protein harbouring a SET domain, and it was shown to methylate H3K4 residues instead [51]. The same study showed through co-immunoprecipitation experiments that while SETD1A methylates H3K4 residues, only H3K4me3 was associated with increased gene transcription, but not H3K4me2, despite retaining an open chromatin state. Some other H3K4-specific HTMs, such as KMT2A, KMT2B or MLLT6, are all homologues of *Drosophila* trithorax protein which is the founding member of the trxG (trithorax group) complex. TrxG can contain a range of different effector molecules, but all are associated with activation of gene expression [52]. Even more, acting together with the PcG complex, covered later among the H3K27 HMTs, the PcG/trxG formation can act as a complex switch between active and repressive chromatin states leading to dynamic gene expression.

H3K36me3 is another histone mark that is associated with active transcription, and, like the rest, originally discovered in yeast as SET2 [53], it was found to have multiple homologues in mammals, called Nuclear Receptor Binding SET Domain Protein 1, 2 and 3 (NSD1, NSD2 and NSD3) [54]. The latter study found H3K36me3 marks to be associated with coding regions of actively transcribed genes, rather than 3' ends or promoters of genes as seen with other marks. Also, RNAPII (RNA polymerase 2) directly interacts with SET2 in yeast and while deletion of SET2 reduces the

expression of genes, it does not abolish it completely [54], indicating that other factors may be at play.

Finally, H3K27me3 is another mark of significant importance and it is known for being associated with repressed genes and silenced loci. In yeast, E(Z) (Enhancer of Zeste), also one of the founding members of the SET domain name, was shown to be the effector causing H3K27 tri-methylation, among others, within the polycomb group (PcG) gene repressor complex [55]. Human homologues EZH1 and EZH2 (Enhancer of Zeste homologue 1 and 2, respectively) were shown to perform the same function, although EZH2 much more so than EZH1 [56]. Interestingly, it appears that EZH2 is only able to catalyse methylation of mono-methylated, not unmethylated, H3K27 residues. If EED (Embryonic Ectoderm Development), a H3K27me1 reader protein, is knocked-out, H3K27me3 marks are also eliminated [57]. EZH1, EZH2 and EED along with other proteins for PcG repressive complexes (PRC) are linked with long-term gene silencing through DNA methylation [58].

To sum up, SET domain containing proteins are responsible for adding methyl groups on histone lysines. In fact, only one other enzyme without the SET domain, DOT1L (disruptor of telomeric silencing 1 like) is known to be a HMT, specific for H3K79 [59]. H3K4me3 and H3K36me3 are known to be linked to actively transcribed genes, while H3K9me3 and H3K27me3 are associated with silencing. More complicated regulatory switches can be implemented through the recruitment of the PcG/trxG complexes.

5.2. Histone methyl-lysine reader domains

While methyl marks on lysine residues do have an impact on nucleosome compaction alone, by introducing hydrophobicity to the residues and a positive charge, the main effects on gene expression are exercised through the binding of methyl-lysine recognising proteins [60]. Contrary to the lack of diversity between HMT domains, there is a whole range of methyl-lysine recognising protein domains. Such variety facilitates the recognition of a combination of modifications and takes into consideration the surrounding epigenetic marks which lead to a coherent regulatory programme, as shown next.

The first protein containing the PHD (plant homeo domain) finger was discovered in plants and it was shown to have a strong affinity for H3K4me3 marks, less so for H3K4me2 and no affinity for a multitude of other tested peptides [61]. Soon, a number of proteins containing PHD fingers were found in animals and they were all shown to have the same affinity. Interestingly, despite the H3K4me3 mark being commonly accepted as the activating mark, different proteins can have opposite outcomes. ING (inhibitor of growth) is a family of proteins that was shown to have such divergence between family members [62]. For example, ING2 was shown to be part of a complex containing the HDAC RBP1 (retinol binding protein 1) while ING3 and ING4, on the other hand, were found in HAT-containing protein complexes [62]. The ING family is implicated in regulation of p53 activity and its downstream genes, so it is heavily implicated in cancers, despite having opposing effects. Thus, the mark by itself may not always provide sufficient information about the state of

the expression of genes and it may be necessary to look into the binding of effector proteins and even the complexes they are found in.

Tudor domains are known to bind mainly H3K4me3, H4K20me2 and H3K36me3 of all the methylated histone lysines [63]. KDM4A (Lysine demethylase 4A, also known as JmjD2A) is particularly interesting, because while its double Tudor domain is specific for H3K4me3 and H4K20me2 marks, it also harbours H3K9me3 and H3K36me3 demethylase domains [64] which hints at an intricate targeting mechanism to histone marks to adopt an active gene transcription state. Another important scenario is a domain cassette containing Tudor and bromodomains separated by a linker. An example of such cassette-containing protein is BPTF (bromodomain PHD finger transcription factor) which recognises the H3K4me3 mark through the Tudor domain and the H4K16ac mark through the bromodomain [65]. Considering the two marks are on separate histones, the steric positioning of the two binding pockets is rigidly controlled by the linker consisting of an α-helix. It is also predicted that further DNA-binding to the major groove contributes to the binding affinity [65].

Chromodomain is another type of protein domain that is known to bind methyl-lysine residues. Canonical chromodomains, based on the homology to HP1 (heterochromatin protein 1) and Pc (polycomb protein) [66], are the main focus because of their direct interaction with methylated histone lysines and their interaction with other epigenetic complexes. Curiously, a simple non-canonical chromodomain was found in a vertebrate transposase specifically targeting H3K9me3 marks, also earning it the name of a chromovirus [67].

PcG was already mentioned before as strong gene silencer and the founding member Pc has five variants in mammals, all of which contain a chromodomain that is specific to H3K27me3 [66]. HP1, originally identified in *Drosophila*, has three isotypes in mammalians, namely CBX1, CBX3 and CBX5, all of which bind directly to methylated H3K9 [68]. As mentioned previously, H3K9me3 is a silent chromatin mark, and to achieve this, HP1 analogues mediate this process by binding to small protein complexes [69]. Using tandem affinity purification coupled with mass spectroscopy, authors in the latter publication identified the transcriptional repressor TRIM24 (Tripartite Motif Containing 24, also known as TIF1α) as a major binding partner to all three isotypes of HP1. A number of components of the NuRD (nucleosome remodelling and histone deacetylase) complex, due to presence of active deacetylases and demethylases normally seen as a potent gene silencer [70], were also identified as binding partners, indicating HP1 proteins are a vital link between the epigenetic mark and the inhibitory machinery. To sum up, both canonical chromodomain-containing protein families, HP1 and Pc, participate in large gene silencing complexes that specifically recognise H3K9me3 or H3K27me3, respectively.

Finally, there is a number of other protein domains, such as WD40 and MBT repeats, PWWP and CW domains and tandem combinations of the already reviewed domains [60]. They all have distinct recognition sites and in combination with other protein complexes they exhibit gene silencing or activating function, depending on the respective histone mark as well as the surrounding epigenetic code.

5.3. Histone lysine demethylases

Originally it was thought that histone lysine methylation is reversed only by replacing the entire histone in a nucleosome [71] or through multiple cycles of cell division which lead to the dilution of histone methyl marks [72]. However, the discovery of a novel family of amine oxidases and the founding member LSD1 (also known as KDM1A) [73], gave the first indication that lysine methyl marks are capable of regulating gene expression just as dynamically as acetyl marks. While a lysine can harbour up to three methyl residues, LSD1 and LSD2 (also known as KDM1B) are only able to eliminate methyl marks from mono- and di-methylated H3K4 and H3K9 residues [73, 74].

Soon after the discovery of LSD1, a much larger family of histone lysine demethylases, all containing the Jumonji C domain, was identified. The founding member JHDM1A (also known as FBXL11 or KDM2A) was shown to be a potent H3K36me2 demethylase [75]. Many more JmjC-domain containing proteins have been discovered since, and the family now contains 30 members of which 18 have been shown to be potent demethylases [76]. The JmjC domain is part of the 2-oxoglutarate dependent oxygenases which were discovered as part of a larger family of iron dependent oxygenases [77]. Jumonji demethylases affect histone lysines, as well as potentially arginines, on a number of different locations on the histone tails. Interestingly, a lot of them are limited to di- and tri-methylated lysines while only a few can demethylate mono- and di-methylated lysines while KDM2B can remove methyl residues from a whole spectrum of mono-, di- or tri-methylated lysines [76].

Considering the particular interest in H3K4me3 and H3K27me3, which are some of the most prominent marks for active gene expression and silencing, respectively, it is important to mention enzymes belonging to KDM5 and KDM6 sub-families. KDM5A-D proteins (also known as JARID1A-D) are all sole demethylases of H3K4me3 marks. Considering the KDM5 sub-family removed the methyl marks from H3K4me3 residue which signifies active gene expression, it is considered to be a family of repressor proteins. For example, KDM5A was shown to be involved in cell cycle regulation and cell differentiation which is exercised through a complex with retinoblastoma protein (Rb) [78]. KDM5B, related to cellular differentiation and senescence [79], was shown to be specifically expressed in melanoma [80] and breast cancer [81].

The sub-family of KDM6 proteins, on the other hand, are the only enzymes capable of demethylating H3K27me3 marks, and hence are often seen acting on genes that are in a poised state due to binding of PRC2 complex [82]. Unsurprisingly, KDM6 sub-family is also involved in early development, where, by demethylating H3K27me3, poised developmental transcription factors, such as HOX and T-box, become expressed [83]. Interestingly, KDM6B (also known as JMJD3) was shown to be particularly involved in inflammation, as NF-κB upregulates expression of KDM6B [84] which in turn demethylates H3K27me3 marks around the promoter of the pro-inflammatory cytokine TNF-α [85]. Furthermore, more than 70% of LPS-inducible genes were shown to be associated with KDM6B in a human monocytic cell line [86].

To sum up, KDM5 and KDM6 sub-families are the main focus in this thesis because of particular importance in cell differentiation and inflammation which therefore

relate strongly to the study of epigenetics in inflammatory cells and myeloid cancers. Through the interplay of H3K27me3 and H3K4me3 marks, and the interaction of KDM5 and KDM6 enzymes within macrophages, osteoclasts and multiple myeloma cell line, an attempt will be made to explain the epigenetic mechanisms that govern the fate of these cell types.

6. Epigenetics of macrophage differentiation

While there is a substantial amount of literature available on the involvement of transcription factors in differentiation of haematopoietic cells [87], there is only scant evidence describing specific epigenetic mechanisms that lead to cell specialisation [88]. Available studies rely on the global chromatin states assessed by DNase 1 sensitivity to determine transcriptionally active genomic regions [89]. Others offer an insight into the acetylated state of histones around genes of interest [90]. These studies provide evidence of certain genomic regions, such as the globin cluster, adopting an open chromatin conformation in the downstream differentiation path and maintaining it for several stages.

Interferons (IFN), particularly IFNα is a major factor in T-cell activation and subsequent collaboration with B-cells, both of which are implicated in the pathogenesis of rheumatoid arthritis. IFNs induce one of nine interferon regulatory factors (IRFs), a family of transcription factors that induce a subsequent pro-inflammatory response [91]. However, IRF induction as well as its mode of action has a tight interplay with histone marks. For example, IFNγ binding to the cell

surface induces transphosphorylation of JAK1/JAK2 which leads to phosphorylation of STAT1 and this allows dimerization of STAT1 and nuclear translocation [14]. However, under presence of HDAC inhibitors STAT1 signalling is impaired and IFNγ–induced IRF1 activation does not occur [92]. On the other hand, STAT1 double knock-outs still can respond with inflammatory signatures through the alternative STAT2 signalling pathway.

IRFs also play a role in the pro-inflammatory response in macrophages. Classically activated, or pro-inflammatory M1 macrophages can be induced from monocytes with the stimulation of IFNγ, lipopolysaccharide (LPS) or granulocyte macrophage colony stimulating factor (GMCSF) [93]. IRF5 upregulation within 2h of treatment was found to be the predominant factor in macrophage subtype determination compared to alternatively activated, or anti-inflammatory M2 macrophages [94]. Even more, the same study reports that M2 cells treated with IFNγ start producing IRF5 too and end up exhibiting a pro-inflammatory phenotype. Direct introduction and constitutive expression of IRF5 induces a pro-inflammatory response in macrophages without any of the aforementioned stimulants [94]. An epigenetic modulator, histone lysine demethylase specific for H3K27me3 (KDM6B or JmjD3), was also shown to be highly increased in LPS-stimulated pro-inflammatory macrophages through NF-κB induced expression [84]. However, despite an over-expression or knock-out of KDM6B, only a handful of genes were affected without a global decrease or increase of the H3K27me3 histone mark. Bone morphogenetic protein 2 (BMP2) was the most down-regulated gene upon KDM6B inhibition and

this gene is firstly known as a growth factor promoting bone formation, in addition it has been noted to be a strong pro-inflammatory stimulant [95].

As mentioned before, NF-κB is a major player in macrophage differentiation, however it is not fully understood what triggers chromatin remodelling to a more open state, required for NF-κB binding. Even though NF-κB cannot bind to tightly packed heterochromatin, there have been cases when p65, one of the possible constituents of NF-κB, has been shown to interact with histone acetyltransferases and other parts of the epigenetic machinery [96] to open up the chromatin around the IL-6 promoter where NF-κB can then bind and induce transcription of IL-6. In fact, it is becoming clear that there is an interplay of multiple transcription factors, each changing the epigenetic marks around NF-κB response elements in favour of a more open or closed chromatin state [97]. Furthermore, different cell types have a unique set of cofactors [98-100] required to trigger the same outcome hence allowing tighter regulation of an inflammatory stimulus.

In a similar way, anti-inflammatory signalling pathways have been described in some detail, but many questions remain [101]. For example, histone 3 lysine 27 demethylation by JmjD3 in resting macrophages has been described as vital for polarization to the M2 phenotype upon stimulation by chitin, a component found in helminths and known to be an M2 macrophage stimulus [102]. However, the derivation of macrophages from monocytes has been studied only by comparing the transcriptomes of pre-cursor monocytes and mature macrophages [31] rather than looking into various epigenetic and signalling pathways. Similarly, other previously mentioned studies have been carried out mostly using targeted techniques studying

a specific locus or modification, rather than high-throughput screens. As such, there is plenty of scope for new findings if genome-wide or transcriptome-wide methods are employed.

7. Epigenetic regulation in inflammation and cancer

As it was already pointed out before, epigenetic regulation plays a significant role in cell cycle, differentiation and response to stimuli and quite naturally, any miss-regulation of this intricate system, would lead to the development of disease. It is common that once a genetic link is difficult to identify, epigenetic regulation is the next most likely explanation for a complex disease, such as cancer, arthritis and many others.

Epigenetics, and specifically the interplay of various histone marks, is frequently implicated in adverse inflammatory conditions. In rheumatoid arthritis (RA), for example, a strong self-propagating pro-inflammatory environment is established by a range of myeloid and lymphoid cells [103]. While genetic inheritance can contribute to the disease, purely genetic reasoning does not explain the full cause and environmental factors are frequently implicated, such as smoking. The effect also appears to be imprinted as cells taken from RA joints and transplanted to healthy joints will cause RA, however it is not uncommon to have only one joint affected. In support of the hypothesis that epigenetics may be involved, histone deacetylase (HDAC) inhibitors were shown to have a significant effect in animal models of RA, relieving the animals of swelling and bone erosion [104]. While the mechanism of

action remains largely unknown, contrary to the previous observation, one study hints at the importance of hyperacetylation of MMP-1 promoter in RA synovial fibroblasts, in which case HDACs would be beneficial in treating the condition [105]. In a different study, peripheral blood mononuclear cells (PBMCs) from healthy volunteers and RA patients were incubated with an HDAC3-specific or a wide-range HDAC inhibitor before stimulating with LPS [106]. IL-6, a pro-inflammatory cytokine normally seen at high levels in RA patients, was significantly reduced upon inhibition of HDAC3, but not with the broad-range HDAC inhibitor. Considering HDAC3 is an H3K27ac deacetylase, it would be an interesting next step to investigate which genes have their acetylation patterns affected that could lead to such a drastic and specific reduction in pro-inflammatory cytokine production.

The role of epigenetics in cancer has been studied to a much greater extent. Cancer has always been thought to result from the accumulation of genetic mutations and chromosomal translocations, but recent evidence is pointing to a significant role of epigenetics in cancer progression [107]. Changes in gene expression induced by hypoxia as well as suppression of apoptotic genes require large-scale epigenetic remodelling and this is frequently observed in a range of cancers [108]. Furthermore, it is particularly striking how inhibition of epigenetic modulators with opposing effects always has a tumour suppressive effect. For example, inhibition of Ezh2, and subsequently the prevention of new H3K27me3 marks, stops the progression of prostate cancer [109] or the inhibition of KDM1A, which leads to a build-up of H3K4me2, and prevents mixed-lineage leukaemia as shown in ex-vivo experiments [110]. On the other hand, inhibition of KDM6A and B results in the build-up of

H3K27me3 marks and this leads to tumour shrinkage in xenograft mouse models [111].

Multiple myeloma (MM) is a cancer of blood plasma cells; it starts in the bone marrow and is of particular interest due to its connection to the haematopoiesis. As with all cancers, there is a number of genetic mutations and chromosomal translocations identified through whole-genome sequencing, and even more importantly, a lot of them are associated with histone modifying enzymes [112]. Among the most notable ones are mutations in KDM6A and the family of histone methyl-transferases KMT2, all of which turn out to result in the overexpression of a transcription factor oncogene HOXA9.

Another notable alteration in MM is the overexpression of a SET domain-containing protein NSD1 [113]. It leads to increased H3K36me2 marks, and also acts together with HDAC1, HDAC2 and KDM1A, all of which increase the expression of oncogenic loci. Furthermore, the anti-apoptotic properties of MM are gained through the activation of JAK/STAT3 pathway and stimulation by IL-6 [114], all of which have well defined epigenetic regulatory networks. The previously mentioned HAT p300/CBP acetylates histones around the promoter region of STAT3 leading to increased expression, while STAT3 recruits HDAC1 which inhibits further expression in a negative feedback loop [115].

High similarity of MM signalling pathways to PBMCs and the tightly inter-linked epigenetic regulation may hint at further opportunities for novel drug targets if previously seemingly disconnected studies on H3K27me3 and H3K4me3 in

inflammatory conditions and MM epigenetics were to be considered as a bulk. The search for such links and relevance of inflammatory epigenetics in MM is one of the underlying topics of this thesis.

8. Small molecule inhibitors for epigenetic modulators

Hopefully it has become evident to the reader that epigenetics is a fundamental component of disease progression, if not the key to it. As such, academia as well as the pharmaceutical industry has recently shifted their focus towards developing small molecule inhibitors specifically targeting individual epigenetic modulators. One impressive example of such public-private partnerships is the Structural Genomics Consortium, an international collaboration between a number of universities and biotechnology companies with the aim to design and discover potent epigenetic inhibitors. While there is already a number of inhibitors available for various epigenetic targets, active even at nanomolar concentrations in ex-vivo and in-vitro systems, those are only suitable for research activities and an extra effort needs to go into those validated molecules to improve their pharmacological properties to make them suitable for administration in patients.

Out of all the classes of epigenetic modulators that were reviewed so far, most of them have small molecule inhibitors, some even approved for treatment in humans. An *in silico* screen of commercially available compounds against the p300/CBP HAT yielded a potent molecule targeting the active site of the p300 component [116]. There is a range of non-specific HDAC inhibitors, of which varinostat (also known

as SAHA) [117] inhibits a broad spectrum of HDACs whereas romidepsin [118] is somewhat more specific by affecting only class I HDACs. Both are approved for treatment of cancers, however there are known side effects [119] that arise from their mode of action related to their chelating properties.

Out of the acetyl-lysine reader domain inhibitors, I-BET and JQ1 are probably the most interesting and promising. I-BET was designed as a mimick of an acetylated histone, and it was shown to greatly reduce the expression of pro-inflammatory genes in immune system cells [120]. Thanks to a public-private drug discovery partnership between the Structural Genomics Consortium and GlaxoSmithKline, the structure and mode of action of I-BET was made publicly available, and this facilitated the discovery of another structurally similar inhibitor of the acetyl-lysine reader domains, JQ1. JQ1 is specific only to the BET family of bromodomains, having the highest binding affinity for BRD4 [121]. A pharmacologically improved version of JQ1 is already in phase I clinical trials.

Histone methyl-lysine affecting inhibitors have been explored quite broadly too. While histone methyl-lysine reader domains have only recently captured the attention of scientists and very few examples of ligands exist [122], there are plenty of writer and eraser domain inhibitors. DOT1L, as previously mentioned, is a unique HMT without a SET domain, so it makes it fairly easy to establish a highly specific and potent inhibitor molecule. Indeed, the most recent variant is more than 37,000 times more selective against DOT1L than any other protein lysine transferase and it has been shown to effectively reduce tumours in xenograft mice with leukaemia cells that had an MLL rearrangement [123]. EZH2 is another particularly interesting target

because of its involvement in the PRC complex, and there are two inhibitors already in clinical trials [124]. Again, both have been shown to significantly reduce tumours in mouse xenograft models of lymphomas [125].

Finally, the most relevant to this thesis are KDM6A and KDM6B inhibitors, due to their involvement in immune cell differentiation, inflammation and haematopoietic cancers, as covered before. An inhibitor, GSK-J1 and the pro-drug version GSK-J4, targeting both KDMs is available and, while it is not as specific as the DOT1L inhibitor, the IC50 is at 9μM and it has been shown to greatly affect the inflammatory response of primary macrophages [85] as well reduce the progression of cancers. Despite a recent discovery that GSK-J4 is also affecting KDM5B [126], the effects in cells are undeniable and simply ask for a more thorough investigation including H3K4me3 histone marks in addition to the H3K27me3.

9. Summary and aims

To sum up, myeloid cells, especially in the context of the whole haematopoietic cell lineage, exhibit a complex diversity of specialised cells arising from one progenitor. In order to maintain this diversity, a network of transcription factors as well as epigenetic modifications are employed, enabling to choose a path towards fully differentiated cells. The large number of epigenetic modifications and an even more diverse range of epigenetic modulators introduces an additional level of flexibility, redundancy and regulation. Sometimes this intricate machinery can go wrong and this leads to auto-inflammatory conditions and cancer.

Given the background of epigenetics in immune system cells and cancer, the main aim of this thesis was to explore this tight interplay of factors that lead to inflammation or cancer by employing small molecule inhibitors, genetic knock-down methods and high-throughput sequencing technology. Initially, a set of biomarkers to identify macrophage subtypes was defined. Subsequently, a range of gene knock-down techniques was established that was most suitable to the cell types being used. Finally, high-throughput sequencing techniques were used to dissect the effects GSK-J4, a KDM6A and KDM6B inhibitor, has on osteoclast differentiation as well as multiple myeloma cells.

Chapter II

Macrophage Biomarkers

1. Introduction

The discovery and study of biomarkers is an important task that facilitates the identification of diverse cellular phenotypes and even progression of disease. The concept of a biomarker was first coined by Brotman *et al* in 1988, and it was called a "surrogate marker" for the condition they were studying [127]. This marker, in the context of disease, is meant to replace the "clinical endpoint" and allows early diagnosis of a condition without waiting until the disease manifests itself in the full spectrum of clinical outcomes, some of which can be fatal. The perfect biomarker is one which is involved in the biochemistry of disease and which can be involved in possible ways of treatment [128].

Similarly, biomarkers for cell types, can be early indicators that the cells are entering a signalling programme leading to the endpoint of full differentiation. Cell type biomarkers can also be useful for identifying already differentiated cells in an indirect manner, because direct methods of identification may require cell fixation and those fully characterised cells would not be available for any further experiments.

Finally, Bradford Hill criteria which are normally used to determine causality [129] and are widely accepted in epidemiology, can also be implemented in biomarker discovery and validation. Biological plausibility, consistency and dose-responsiveness

have been shown to be particularly important in biomarker discovery and a combination of biomarkers can provide better predictions than individual biomarkers [130]. These principles are equally important in cell type biomarkers and have been used in experiments described within this chapter.

Originally, human macrophage nomenclature was based on the better studied mouse model, where macrophage differentiation from monocytes is induced by stimulation with cytokines produced by T-helper (Th) cells. Th1 cells were known to induce the pro-inflammatory macrophages M1 through the production of IFNγ. Th2 cells induce the anti-inflammatory macrophages M2 through the production of IL-4 and IL-13, or IL-10 [131]. Once induced, M1 cells are known to produce large amounts of IL-1, IL-6 and IL-23, all of which are known to induce the proliferation of Th17 cells [132]. IL-4- and IL-13-induced M2 cells, on the other hand, express chitinase, CHIA (Acidic Chitinase, also known as Ym1) [133], IGF1 (insulin-like growth factor 1) [134] among others, all of which reduce the pro-inflammatory response which is normally observed by M1 macrophages. IL-10 induces a slightly different M2 macrophage phenotype, characterised by high levels of TGFβ [135] and also IL-10 which leads to a self-propagating M2 cycle [132].

While pro- and anti-inflammatory macrophages have distinct cytokine profiles, they can only be used to determine the state of an entire population of cells by analysing the media. However, with the field moving to assays based on single-cell analysis this crude method is becoming obsolete. It is becoming desirable to find more accurate ways of classifying macrophages into M1s and M2s through surface markers or expression of specific genes. Such biomarkers would facilitate the elucidation of

whether macrophages can polarise from one type to another, which enzymes are fundamental to macrophage maturation and polarisation, or enable scientists to dissect the heterogeneity within macrophage populations.

Further discoveries of intracellular, rather than secretory, biomarkers were made only recently by employing high-throughput techniques on various stimuli that lead to mature macrophages. It has been postulated that H3K4me1 marks along with the binding of HAT p300 indicate a putative enhancer region on the genome [136]. Consequently, a study, employing the latter observation, identified a number of enhancers in LPS-stimulated macrophages [137]. Quite expectedly, by comparing the families of consensus sequences of the various enhancer regions, it was determined that SPI1 is the most significant TF that is induced upon LPS stimulation, despite SPI1 already being a macrophage-specific TF1 required for differentiation. Other enhancer regions that are induced upon LPS addition with high sequence similarity to NF-kB, IRF, STAT1, and AP-1 consensus binding sequences were also identified in the same study [137]. Such analyses can contribute to the explanation why IL-12, a pro-inflammatory cytokine highly dependent on NF-κB signalling is only released by macrophages, but not other NF-κB expressing cells.

A different approach to study enhancers was taken with the genome-wide localisation of RNA polymerase II (RNApolII) [138]. This method allows identification of non-coding RNAs (ncRNAs) that may still be vital for the function of cells but may never be translated into a functional protein. As a result, a number of genes were identified that indicate expression by utilising the enhancers of other

nearby genes [138]. Again, SPI1 was the dominant TF which had the target consensus sequence for the enhancer laying near those ncRNAs identified by RNApolII.

Finally, the differential analysis of the transcriptomes of differently treated macrophages can be utilised to dissect the changes of mRNA. Apart from the expected secretory proteins and cell surface receptors (see Chapter 1 – "Macrophages and their subtypes"), a variety of intracellular enzymes were identified to be different between GMCSF- and MCSF-induced macrophages [31]. A gene ontology (GO) analysis revealed a large number of genes involved in the cell cycle regulation as well as solute molecule transport.

In such high-throughput analyses particular attention needs to be paid to the timing of treatments. Macrophage intracellular signalling is known to be oscillating due to NF-kB translocation to the nucleus and the subsequent refraction time, forming a period of approximately 100 minutes [139]. As a result, if the cells are fixed or lysed at 30 min or 90 min time-points post-stimulation with LPS, very different results may be observed due to NF-κB being at peak activity or in refractory state, respectively. Donor-to-donor variation is another important factor to consider, as seen in our unpublished observations. Finally, even cytokines used to induce macrophage differentiation from monocytes have a high degree of variation between batches and especially between different manufacturers.

One of the research objects throughout this thesis was monocytes and macrophages in their distinctive states – resting, pro- or anti-inflammatory. This chapter focuses on the procedures undertaken to validate biomarker candidates found in the

literature. Publicly available high-throughput sequencing data was also used to facilitate novel biomarker discovery by re-analysis using different or improved algorithms compared to the original publication. This biomarker work is instrumental in laying ground for future exploration of the epigenetic regulation of macrophage differentiation and response to inflammatory stimuli, which was briefly mentioned before.

2. MCP1, CCL1 and MMR can serve as specific macrophage biomarkers

At the time when these experiments were designed, there was only limited knowledge on macrophage biomarkers and there were no obvious biomarkers readily available to define the M1 and M2 macrophages. A number of biomarker candidates were picked from published microarray studies [31, 140]. INHBA (Inhibin, beta A, also known as Activin A), MCP-1 (monocyte chemotactic protein 1, also known as CCL2 or chemokine (C-C motif) ligand 2), CCL1 and MMR (Macrophage Mannose Receptor, also known as MRC1 or mannose receptor, C type 1) were selected to be further validated by qPCR (quantitative polymerase chain reaction), immunocytochemistry and ELISA (enzyme-linked immunosorbent assay). COX2 (Cyclooxygenase 2) was included as a biomarker for stimulated monocytes and M1 cells [141].

The initial qPCR assay, carried out by our collaborator Dr Palwinder Mander at GlaxoSmithKline, indicated Activin A and MCP-1 to be specific markers for M1 and M2 macrophages, respectively. While Activin A peaked at 24h of differentiation and

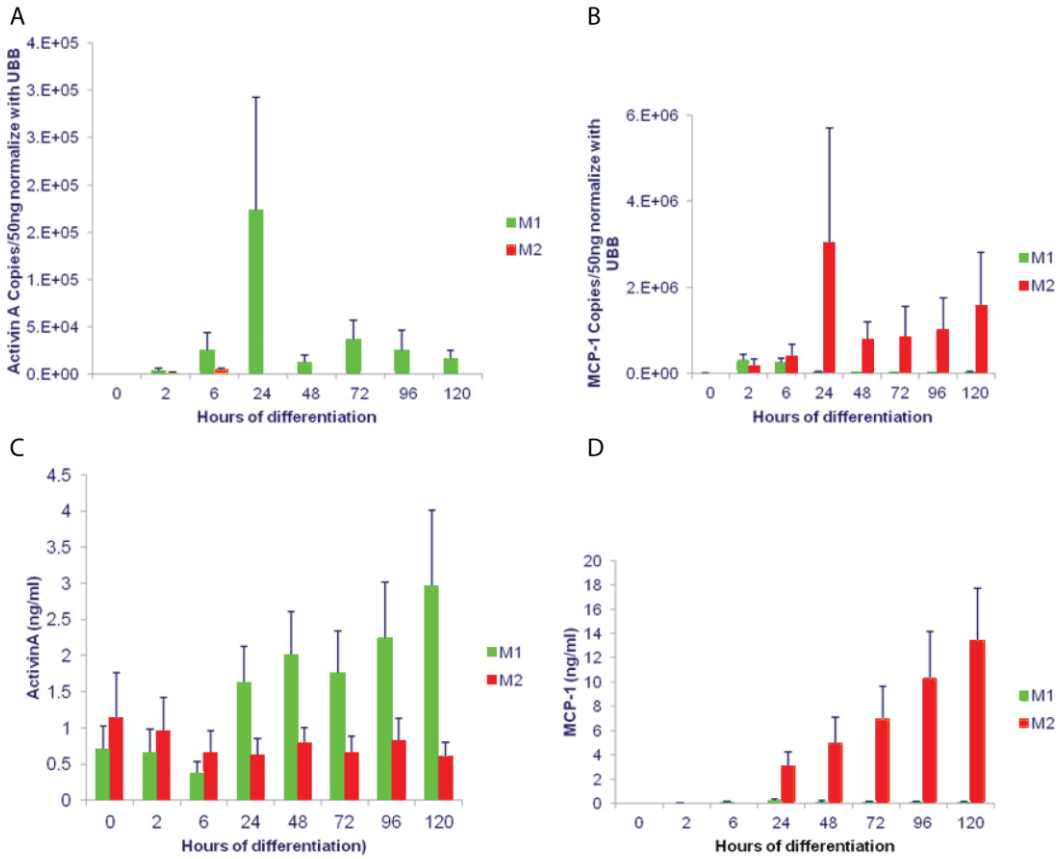


Figure 4 M1 and M2 biomarker levels determined by qPCR and ELISA. qPCR results were normalised to endogenous control UBB (Ubiquitin B). ELISA assays were carried out on monocyte cell culture supernatants treated with GMCSF (M1) or MCSF (M2) for 5 days. Experiments were carried out by Dr Mander and results (from n=3 experiments) are presented as the mean ± s.d.

then slowly levelled off (Figure 4A), MCP-1 showed a clear increasing expression pattern throughout differentiation (Figure 4B). An early peak of Activin A is inconvenient for an end-point macrophage subtyping assay, as low mRNA levels would be hard to detect in a fully differentiated M1 cells. On the other hand, ELISA results, also carried out by Dr Mander, showed a consistent rise of Activin A in M1 cells (Figure 4C) and MCP-1 in M2 macrophages (Figure 4D). This indicates an accumulation of the protein and slow turn-over rates of Activin A, as the pattern is different to the observed mRNA data. Also, significant levels of Activin A in M2 cells suggest that the protein might be an unsuitable biomarker.

To further validate these biomarkers, especially keeping in mind their utility for cell sorting and single cell analysis, immunocytochemistry was performed on monocytes, M1 and M2 macrophages and, by acquiring tiled images, image analysis scripts were used to quantify the data based on individual cell staining. Firstly, the average fluorescence intensity in each cell was measured, which indicates the concentration of the biomarker found in individual cells. Activin A, consistently with the ELISA data shows only slightly larger concentrations in M1 macrophages compared to M2 (Figure 5A). On the other hand, there is, on average, significantly more of MCP-1 found in M2 cells than in M1s (Figure 5B), while CCL1 and, surprisingly, MMR are found at much higher concentrations in M1 cells than M2 (Figure 5C,D). While this indicates those cells express much higher amounts of the biomarker of interest than the other tested cell types, there is still residual expression of all biomarkers in both types of macrophages. In a cellular screen individual cells could be characterised by their respective biomarker if stringent gating is applied.

Consequently, in compliance with the Hill criteria, it was necessary to assess the specificity of these potential biomarkers. For a biomarker to be specific it is important to minimise false positives. A good way of estimating the latter is to find out what proportion of cells stain for the biomarker at fluorescence intensities higher than the average negative control – i.e. a cell type not expected to contain the biomarker. Expectedly, it was demonstrated that even early in differentiation, when Activin A expression peaks, it is only found in a small fraction of cells in the culture which is not distinctive from the M2 population (Figure 6A). In contrast, over 80% of cells in M2 population are expressing MCP1 at a significantly higher concentration

than average M1 cells (Figure 6B). This suggests under 30% false positives and under 20% false negatives, MCP-1 is distinctively expressed at much higher concentrations in M2 populations. On the other hand, 40% of M1 macrophages express CCL1 at a higher concentration than the average M2 cells (Figure 6C). Finally, MMR was shown to be a significant marker for the M1 population with over 90% of cells expressing

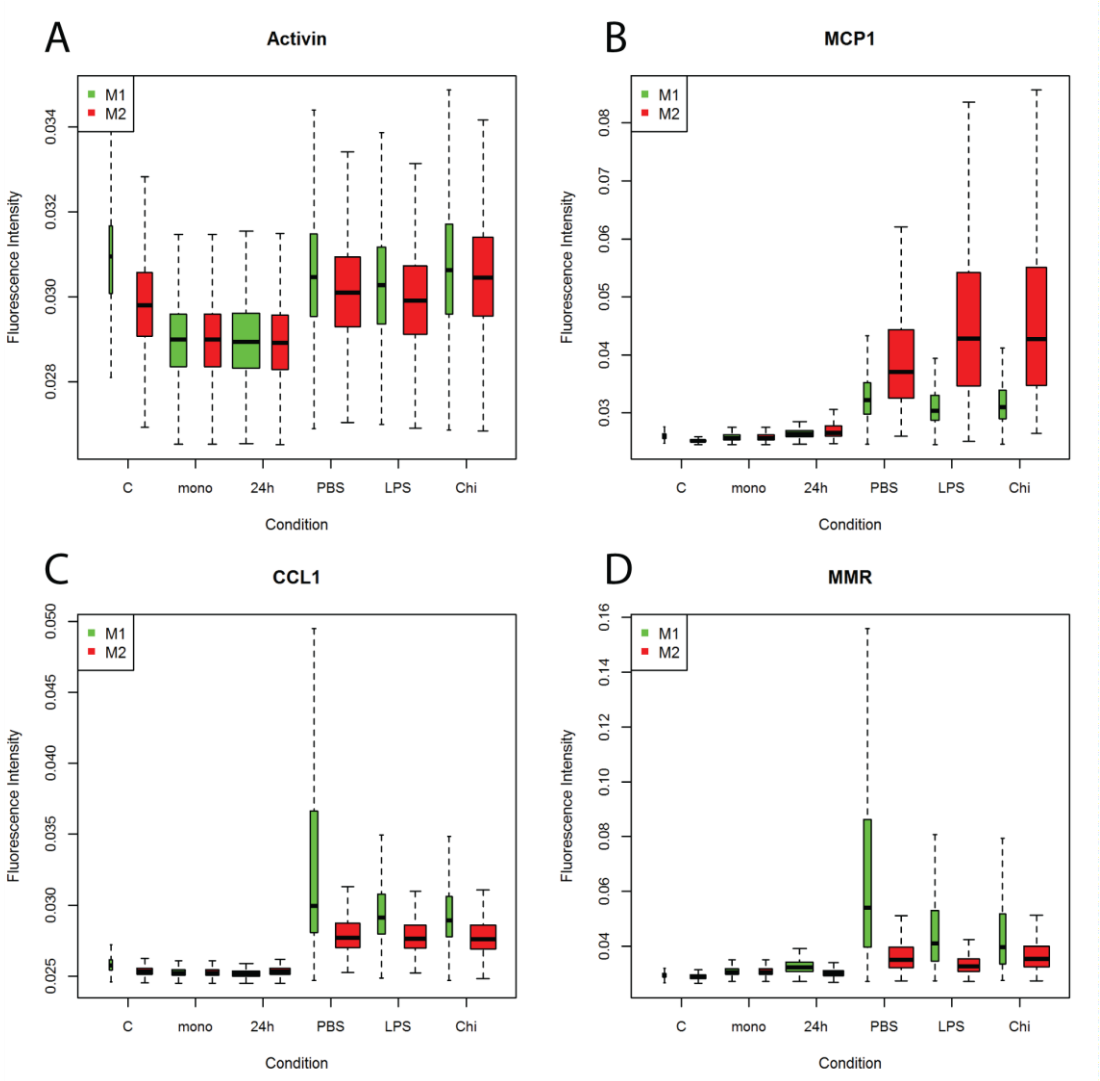


Figure 5 Average cell fluorescence intensity, based on fluorescence intensity measurements within individual cells using high-throughput microscopy (but not fluorescence-activated cell sorting or FACS). The width of the bars reflect the number of cells determined within samples. C stands for negative control – cells incubated with secondary, but no primary antibody; mono stands for monocytes, fixed immediately after isolation and not treated with any cytokines. 24h stands for cells treated with either GMCSF (M1) or MCSF (M2) for 24h. PBS, LPS and Chi are different treatments on cells after 7 days of differentiation either with GMCSF (M1) or MCSF (M2) – phosphate buffer saline (negative control), lipopolysaccharide and chitin particles, respectively. Results (from n=3 experiments) are presented as the mean ± s.d.

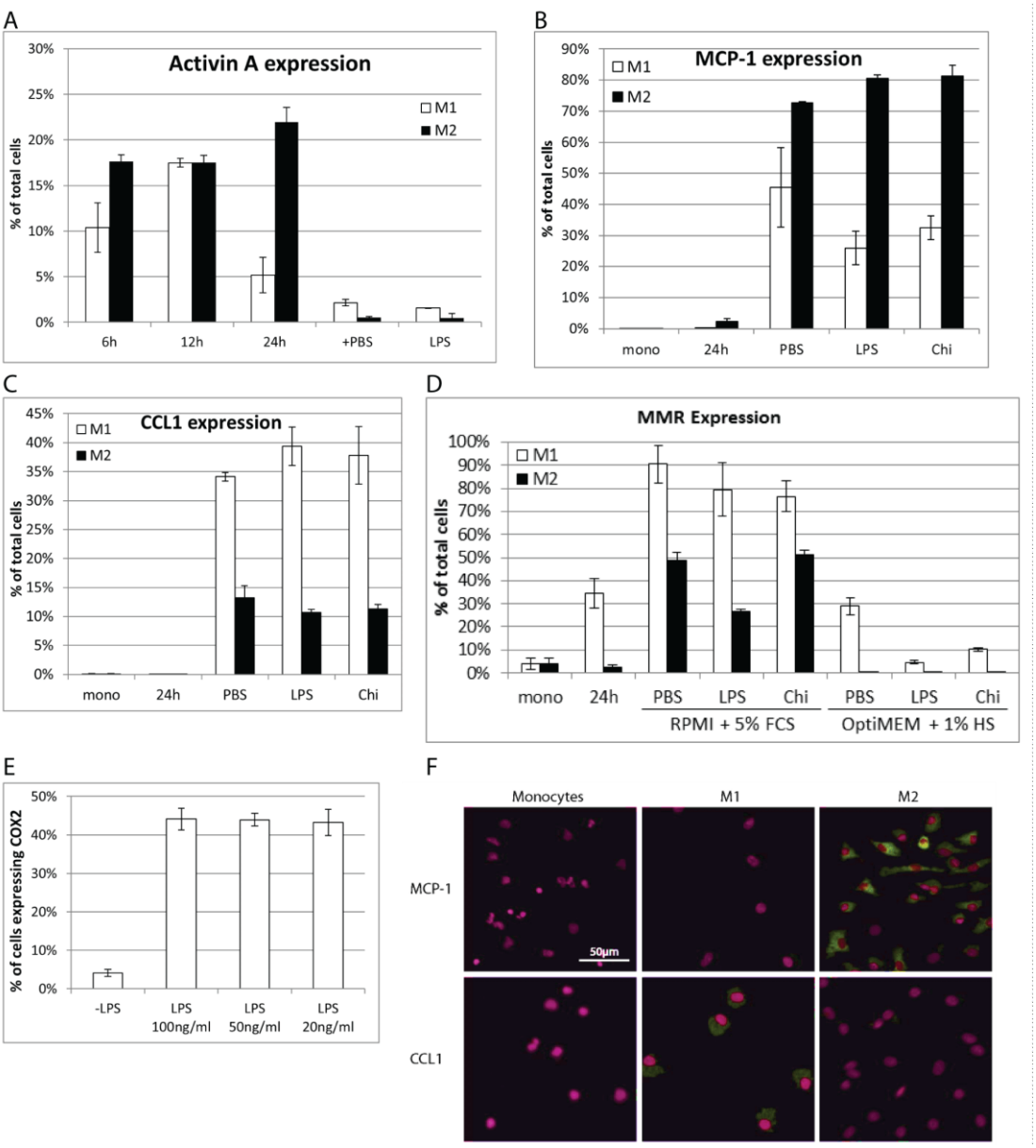


Figure 6 Validation of M1 and M2 biomarkers. Graphs A-D show the fraction of cells that express the biomarker, as determined by the fluorescence intensity, to a higher level than cells that are not expected to have this biomarker. E. The proportion of monocytes expressing COX2, as determined by fluorescence 2 standard deviations higher than the negative control, after 24 hours treatment with LPS. F. Fluorescent microscopy images showing expression of MCP-1 or CCL1 in monocytes, GMCSF (M1) or MCSF (M2) treated cells for 7 days. Pink are the cell nuclei stained with DAPI, yellow fluorescence of the biomarker. Results (from n=3 experiments) are presented as the mean ± s.d.

MMR at a higher concentration than average M2 cells, and the specificity, although to a lesser extent, was also confirmed by culturing cells under varied differentiation conditions, such as OptiMEM supplemented with 1% human serum. A sample image

of what the cells with fluorescently labelled biomarkers look like shows a clear distinction between the two cell types (Figure 6F).

Finally, COX2 was shown to be a reliable biomarker for monocyte stimulation with LPS. Again, tiled images were taken of a fluorescent probe that is only taken up into cells when COX2 is expressed [142]. Cells with statistically significant expression, as compared to background noise, were counted. Over 40% of cells were identified as positive for COX2 with just under 5% of false positive signal observed in untreated cells (Figure 6E).

To sum up, Activin A, initially thought to be a specific biomarker for M1 macrophages was determined to be unfit for the purpose because it appears early in differentiation and hence cannot serve at later time-points. Also, M2 macrophages were shown to produce significant levels of Activin A and as such, the biomarker is not specific enough for M1 macrophages. On the other hand, CCL1 and MMR were shown to be specific biomarkers for the M1 cell type, with MMR being expressed in up to 90% of the cell population. MCP-1 was expressed in up to 80% of the M2 macrophage population and with only under a 30% false positive rate it is also a specific biomarker for this cell subtype. Finally, COX2 was shown to be an early-immediate marker of an inflammatory stimulus in monocyte populations.

3. Re-analysis of published data on high-throughput sequencing of macrophage transcriptome yields new biomarkers

There is a rapidly increasing amount of data available on various databanks, freely available to the research community. Instead of carrying out a simple transcriptomics experiment, aiming to identify unique biomarkers and comparing M1 and M2 cells, published literature and data available online was used. A study (published after the experiments in the previous section were carried out) performed an RNA-Seq experiment on macrophages treated identically to our protocol [143]. While the original paper describes novel biomarkers, most, if not all, appear to have highly overlapping spectra of samples and negative controls in their confirmatory FACS analyses, indicating that the likelihood of false positives and negatives is high, as seen in the previous section with different biomarkers.

The raw reads were re-analysed, with the hope that different parameters and much improved bioinformatic pipelines will offer novel insights into the dataset. The re-analysis yielded much the same outcome as the original paper, along with a number of other previously described macrophage biomarkers, such as GBP1 (guanylate binding protein 1), TNFAIP6 (tumor necrosis factor, alpha-induced protein 6), CHI3L1 (chitinase 3-like 1), SLAMF7 (SLAM Family Member 7) for M1 macrophages and F13A1 (coagulation factor XIII, A1 polypeptide), CD1A, CCL17, CCL23 for M2 cells. Also, looking at the overall gene expression patterns, there are clear distinctions between M1 and M2 macrophages. Firstly, genes cluster into distinctive functional groups that are reflective of macrophage function within these two subtypes (Figure 7). Their relatedness, based on the similarity of gene expression between each sample,

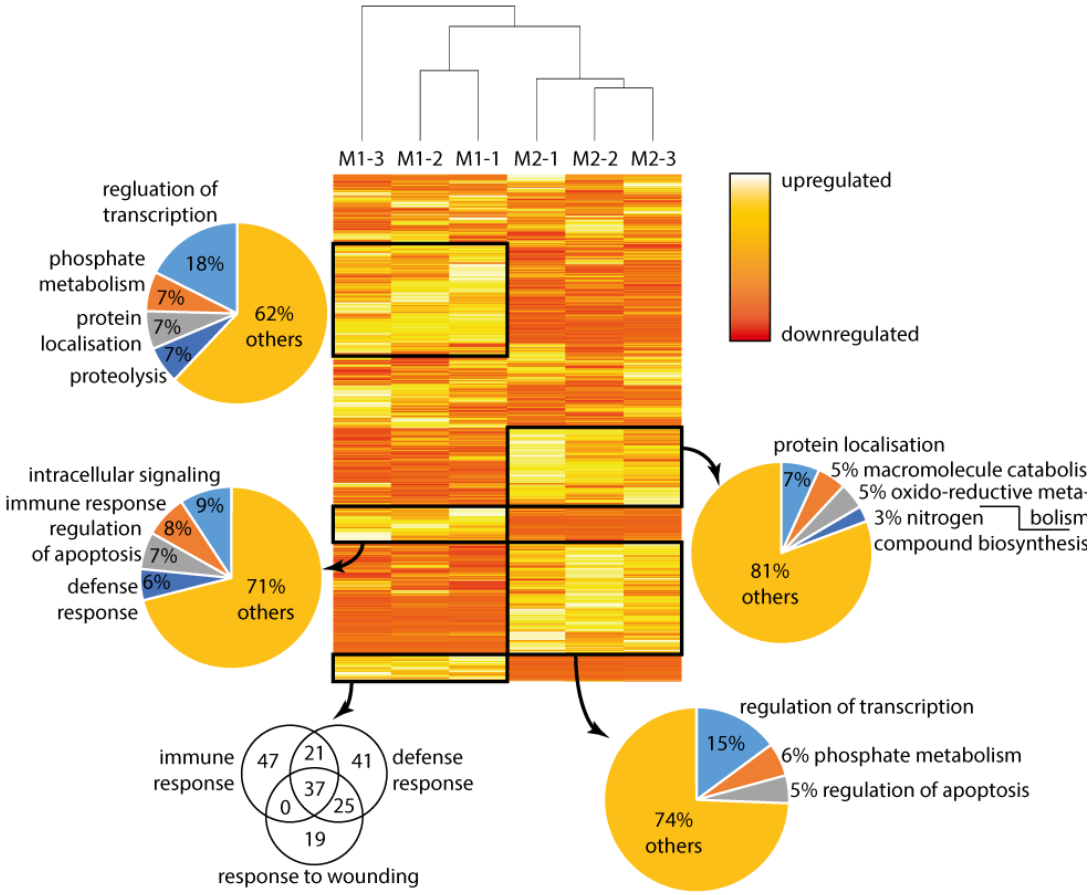


Figure 7 A heatmap showing the differential expression of genes for each replicate. White colour represents upregulation, red – downregulation. Each line represents one of the over 20,000 genes from the reference genome. Dendrogram at the top of the heatmap cluster samples together by similarity of gene expression and the length of the branches represents the size of differences (distance) between the samples. Genes within black squares were annotated using gene ontology (GO) algorithms and the most prevalent terms within those groups of genes are provided on the pie charts. Only the bottom left group of genes are shown as a Venn diagram, because the majority of genes were inflammatory and within the different GO terms, a lot of the genes were overlapping.

was represented with a dendrogram above the heat map (Figure 7), which indicates how distant each sample is from one another. All M2 replicates group into a tight cluster, while M1 samples are showing greater differences compared to M2 cells based on the length of the branches of the dendrogram. Even though the distance between two M1 replicates is nearly the same as between all M2 replicates, one M1 replicate (M1-3) stands out significantly and the algorithm denoted it as a separate unique group, indicating how significant donor-to-donor variability can be.

Using DAVID, a gene ontology (GO) analysis platform [144], a number of terms and pathways were fitted to the genes in clusters of similar expression, denoted by black quadrangles (Figure 7). As expected, M1 macrophages have the genes upregulated that fall under GO terms such as immune and defence response, response to wounding and proteolysis. M2 cells, in contrast to M1, exhibit upregulation in oxido-reductive metabolism, nitrogen compound biosynthesis and macromolecule catabolism, all of which reflect the known M2 phenotype in helminth infection. Interestingly, both types of macrophages have different sets of genes upregulated for functions such as phosphate metabolism, regulation of transcription, apoptosis and intracellular signalling.

Next, individual genes that could serve as biomarkers specific for the type of macrophage were identified. Normally, a cut-off FPKM (Fragments Per Kilobase of exon model per Million mapped fragments) of 0.5 is used to filter out all other exons, because it is assumed that if there are no reads detected (i.e. FPKM<0.5) it is uncertain whether it is because they are indeed absent, or whether it is because of a failure to capture those exons. Also, at very low FPKM values there is much higher variation between replicates than at higher FPKM values (Figure 8), and, as a result, a weighing coefficient would need to be introduced that would lower the significance of differential expression observed if FPKM was low for both samples. However, because the study in question used 3 biological replicates for each condition, all reads, even those with FPKM<0.5, can be confidently included in a subsequent differential

expression analysis. While this increases the likelihood of false positives, a manual inspection of individual tracks is often sufficient to identify an artefact.

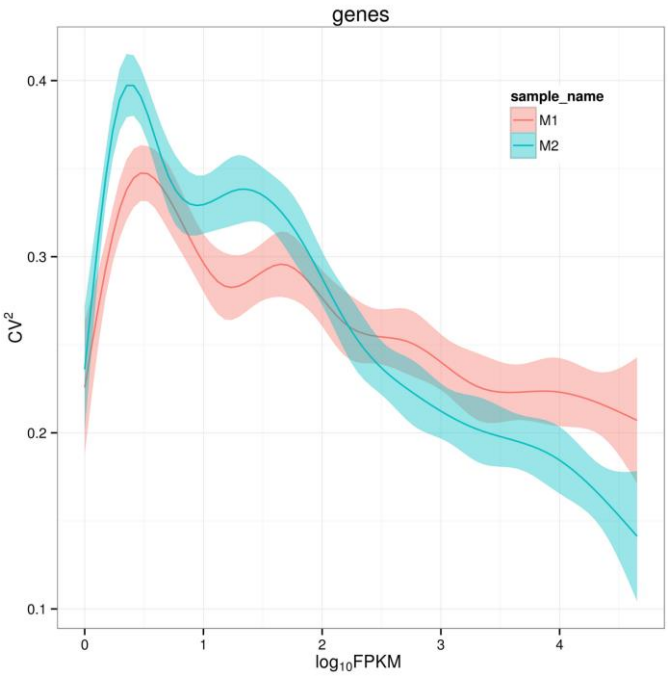


Figure 8 The squared coefficient of variation shows the variation of FPKM values between biological replicates of the same sample. High differential expression determined for genes with low FPKM values can be unreliable due to this variance.

As a result, a set of biomarkers was identified that highly overlaps with the findings of the original publication. Additionally, 4 new biomarkers were obtained with exclusive or near-exclusive expression in one macrophage subtype, but not the other. The biomarkers for M2 cells, namely DHRS2 (Dehydrogenase/Reductase (SDR Family) Member 2) and CCL26 (Figure 9A, B), appear to be completely exclusive to the cell type of interest. The finding is ever more plausible due to the FPKM in expressing cells varying between 200 and 2000 for CCL26 and 100-800 for DHRS2 while the signal is completely absent in M1 cells across all three biological replicates. On the other hand, M1 markers APOBEC3A and GBP5, while being weakly expressed in M2 cells, are orders of magnitude higher in M1 macrophages (Figure 9C, D). It is important to note that the scale on M1 tracks goes up to 970FPKM for APOBEC3A and 6200FPKM for GBP5 compared to only

70FPKM or 65FPKM on M2 tracks, respectively, which makes the expression significantly different.

To further confirm these new findings that were not previously identified in the original publication, expression levels of the new biomarkers were confirmed by

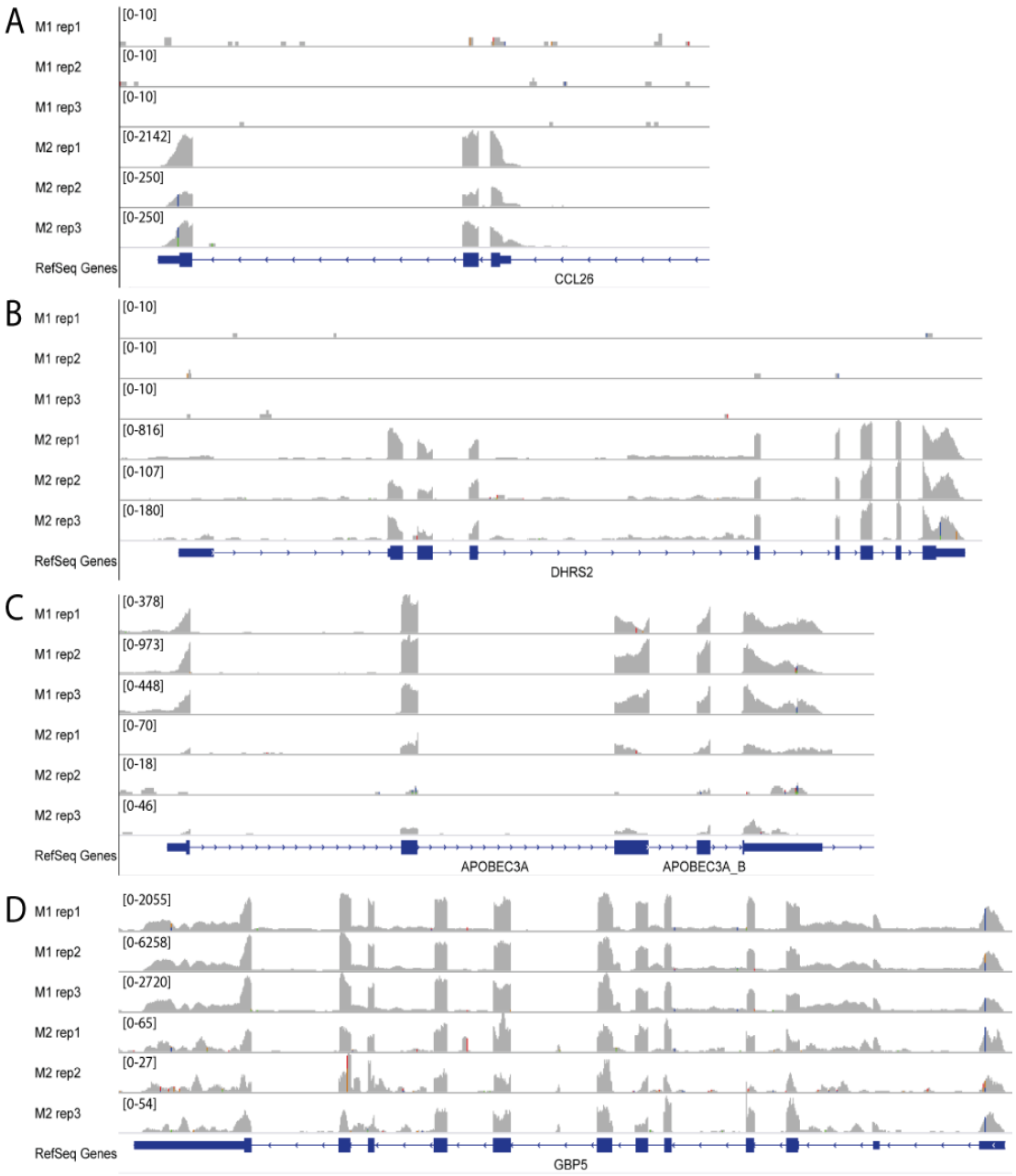


Figure 9 Track data provided for each replicate for the top differentially expressed genes between M1 and M2 cells. Please refer to the scale of each track on the top left for the approximate height of peaks.

qPCR. Monocytes were differentiated into macrophages using the same protocol with cells from 3 different donors. The expression levels of the target genes were determined by qPCR. Both M2 markers were around 20-fold higher than in M1 cells, the M1 marker GBP5 was approximately 24-fold higher than in M2 cells while the APOBEC3A varied from a 31 to 88 fold change between donors as compared to M2 cells (Figure 10). Based on the FPKM values observed in the RNA-Seq experiment, about a 30-fold difference in GBP5 expression and up to a 60-fold difference in APOBEC3A expression between M1 and M2 cells was expected, which is indeed very similar to the obtained qPCR values. However, the expected enrichment by a hundred fold or more for M2 biomarkers was not observed, possibly due to some artefact that prevented reliable capture or alignment of the two genes in RNA-Seq M1 samples.

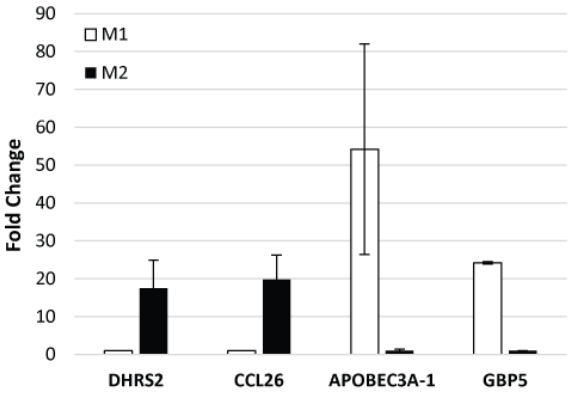


Figure 10 Expression of newly identified macrophage biomarkers as measured by qPCR. The fold change values are calculated in relation to the opposite macrophage subtype, i.e. expression levels in M1 cells are used as a reference point for M2 specific biomarkers (DHRS2 and CCL26), and the opposite was done for M1 biomarkers APOBEC3A and GBP5. Results (from n=2 experiments) are presented as the mean ± s.d.

4. Discussion

Monocyte differentiation into macrophages is a fundamental element of the immune system that acts as a first response to infection as well as a clearing mechanism for dying cells. Briefly, M1 or pro-inflammatory macrophages differentiate from

monocytes upon exposure to IFNγ while IL-4 and IL-13 will lead to monocyte differentiation towards an anti-inflammatory or M2 macrophage subtype (Figure 11A). The M1 and M2 nomenclature arises from tight links with the Th1 and Th2 cell function, as discussed previously.

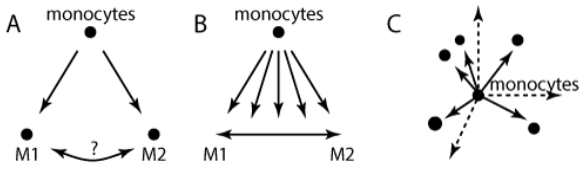


Figure 11 The evolution of the understanding of macrophage polarity. A. Originally thought that monocytes can differentiate into two distinctive macrophage states, corresponding to the Th1 and Th2 cells. B. Study of a wider range of stimulants and availability of high-throughput techniques led to the more fluid understanding of macrophage polarity. C. Application of systems biology to a wide range of treatments and cells coming from different environments enabled to build a spectrum of different macrophage types that are closer or further apart within the space depending on the overlap of their gene networks.

To study macrophage responses in vitro, primary monocytes or immortalised monocyte cell lines can be differentiated to macrophages in a number of ways, most commonly by treating with IFNγ or GMCSF to produce M1 cells, or IL-4 and IL-13 or MCSF to yield M2 cells. These cells can be further stimulated with other cytokines, LPS or other agents and their response can be measured. Biomarkers, some of which were presented in this chapter, play a fundamental role in this area, because they help to verify the cell type and allow to predict the response to stimulus.

Recently, however, an argument against this polar understanding of macrophage biology has been growing more vocal [145]. The authors argue that previously, with few macrophage stimulants and biomarkers known, it was understandable to have a binary response, but the breadth of recent studies involving a much broader range of stimulants and genome-wide analyses point to a spectrum between M1 and M2 responses (Figure 11B).

The pro-inflammatory macrophages can be induced with three different stimulants, namely IFNγ, LPS and GMCSF. While all three induce a pro-inflammatory environment, they act through different signalling mechanisms and induce expression of certain cytokines, only some of which are shared between all three stimulants [146]. Anti-inflammatory macrophages can be induced in an even more diverse range of ways: IL-4, IL-10, glucocorticoids (GCs), IgG and MCSF. Just like in the case of pro-inflammatory stimulants, they all induce distinct signalling pathways and responses [145].

GCs are steroid hormones produced by the adrenal glands and are part of the anti-inflammatory feedback mechanism of the immune system [147]. GCs have also been used in therapies used to dampen down the immune response in various conditions. As such, their effect on monocytes and macrophages has a high biological as well as clinical relevance. GC stimulation of monocytes induces a distinctive subtype of macrophages, significantly different from the previously discussed M1 and M2 cell types. A microarray study along with the known down-regulation of an inflammatory response, revealed a reduction in genes related to cell motility, cell-to-cell communication, cell death and particularly chemokine activity [148]. On the other hand, the same study shows an increased capacity of phagocytosis of dead cells which further dampens down inflammation through removal of pro-inflammatory stimulants.

In addition to the macrophages developing from a circulating monocyte population in response to a stimulus, it is also known there are resident macrophages found in tissues [149]. It has long been recognised that macrophages from haematopoietic

stem cells in early mammalian development can inhabit a range of developing tissue and persist, [150] which lays the basis for the study of monocyte-independent tissue-macrophages [151]. Across literature it is observed that tissue macrophages retain an M2-like profile, because IL-4 and MCSF released by surrounding cells helps maintain homeostasis, although the response is markedly different from monocyte-derived IL-4 or MCSF-induced macrophages [152].

The rise of systems biology has recently helped to construct a more integrated picture of macrophage subtypes, incorporating cells from environments as diverse as cancers, chronic inflammatory conditions and others, totalling up to 28 different types of treatment [153]. A complex network of genes emerges that proves the intricacy of this large network of signalling pathways, enabling subtle variations of response depending on the environment the cells are found in. The latter approach evolves the model of macrophage polarisation from a bipolar linear scale to a three-dimensional spectrum, each node indicating a distinctive type of macrophage that can be characterised by a set of transcription factors, cytokines and cell surface receptors (Figure 11C).

To sum up, the field of immunology with all its diverse cell types and signalling patterns is evolving at an ever faster pace. Macrophages constitute a significant part of it and the research community is currently undergoing an important paradigm shift - from a linear differentiation path from monocytes to macrophages and a polar view of cell types, to an intricate spectrum of macrophage subtypes that can change along with the evolving environment, and not all of them are necessarily monocyte-derived macrophages. In the future, a complex unified model incorporating a large

number of genes should emerge that would allow better characterisation of cell types and their responses to treatment. Eventually, a better understanding of this interplay will lead to better understanding of disease and better healthcare.

Chapter III

Gene Knock-Down Methods

1. Introduction

Manipulation of gene expression is fundamental to the study of gene function and disease. While small molecule inhibitors mostly bind to proteins to reduce their binding affinity or impair an enzymatic function, such observations need to be confirmed by knocking-down the protein of interest, or even eliminating it completely. Preferably, such changes in protein levels or gene expression would have an on/off switch, especially if they are lethal. Finally, such manipulation must be carried out through minimising the effects on the remaining proteins and their complexes, regulatory genetic loci and genes themselves.

There is a vast range of techniques available to achieve altered gene expression. Originally, there was a time- and resource-intense method based on cloned disruption cassettes [154]. Those were instrumental in elucidating the function of newly discovered genes that were identified in whole genome sequences. However, the difficulty of cloning whole ORFs, finding homology regions within target genes and challenging transfection in cells from higher organisms are all limiting factors for convenient screening procedures.

The discovery of RNA interference as an evolutionary pathway to gene silencing in *Caenorhabditis elegans* [155] paved the way for siRNA and shRNA technology.

Essentially, the technique works by introducing a short strand of RNA into the cells which is complementary to the target mRNA. Once a double-stranded RNA is formed, it activates the endogenous cellular machinery to degrade the RNA, while keeping the interfering RNA within a protein complex called RNA-induced silencing complex (RISC). It recycles the siRNA so it can target other RNA molecules with the same target sequence [156].

There are two methods of delivering the interfering RNA into cells. One way is to directly transfect cells with synthesised siRNAs. Another way is to employ lentiviral systems to deliver constructs that integrate into the host genome and express a siRNA precursor called shRNA under the control of U6 promoter. The former method is the easiest, however very short-term, as even the most stable siRNAs remain active in the cells for up to 72 hours. This requires frequent transfection with siRNA in order to observe long-term effects of the knock-down, but that may be highly affected by the exposure to transfection reagents. The lentiviral delivery method is much more stable, as the shRNA is constitutively expressed from the genome and conditional expression switches can also be incorporated. However, there is a risk of integrating the construct in a genomic region of significant importance to the experiment. Also viral infection may be difficult in certain cell types or it may trigger adverse effects, as will be later demonstrated in this chapter.

One very recently developed technique, similar to the siRNA knock-down mechanism, is based on locked nucleic acids (LNAs). These are modified oligonucleotides that have a methylene bridge between the hydroxyl group sitting on the 2' carbon and the 4' carbon in the ribose ring (Figure 12). The LNAs are normally

placed at the ends of the oligonucleotide, thus preventing the digestion by endogenous exonucleases. While LNAs form double-stranded RNAs with their targets, instead of inducing the RISC complex as with siRNAs, the target is degraded by the RNase H and the LNA oligonucleotide remains intact and can bind to the next target [157].

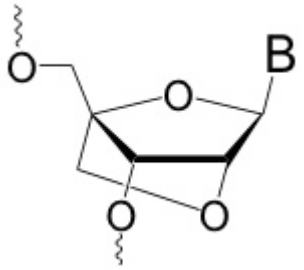


Figure 12 The structure of a locked nucleic acid. B stands for one of the RNA bases. Image adapted from Santaris Pharma website.

In this chapter a number of techniques are discussed that were tried for screening and studying genes significant for macrophage inflammatory response. The aim was to facilitate the discovery of new anti-inflammatory treatments for auto-immune conditions. The reader must consider the complexity of issues that arise when trying to manipulate primary inflammatory cells, which over millennia have evolved to be extremely resilient and sensitive to any such attempts by invading pathogens.

2. Production of lentiviral particles

Firstly, opportunity arose to use a library of lentiviral constructs with shRNAs targeting individual epigenetic proteins. While the library was already available for use and looked after by our collaborators in Vienna, at the Centre for Molecular Medicine, lentiviral particles were produced as part of this thesis for optimisation. A pipeline was established for the production of lentiviral particles using a commercially available plasmid sh003, which is the exact same construct as the rest

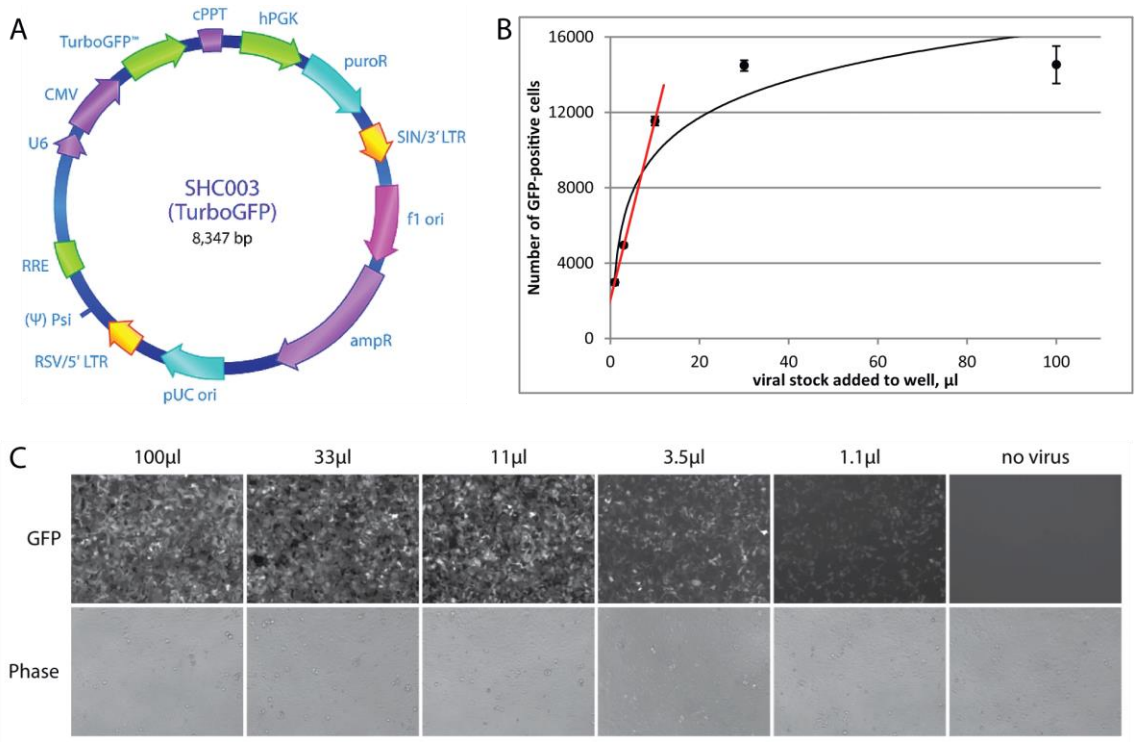


Figure 13 A. Structure of the commercially available control plasmid shc003 for lentiviral infection optimisation. Reproduced from www.sigmaaldrich.com. B. The viral titre was determined by infecting HEK cells with serial dilutions of the produced viral stock and by estimating the differential of the linear phase of the curve. Results (from n=3 experiments) are presented as the mean ± s.d. C. Microscopy image showing the saturation of GFP positive cells at high volumes of viral stock.

of the library, just in addition to puromycin it contains GFP for the ease of visualisation of infected cells (Figure 13A). Briefly, the lentiviral particles were produced using the second generation system which involves transfecting Phoenix cell line with three plasmids: one being the shc003 and the other two containing VSV-G viral envelope protein and packaging genes. After 60h of culture in 20% BSA, the supernatant was filtered through a 0.4µm membrane and stored at -80°C.

The viral titre was estimated by infecting HEK cells with serial dilutions of the lentiviral stock, expressed as the volume of original stock added to wells (Figure 13B). At high viral concentrations, the number of GFP-positive cells levels off as it forms a logarithmic curve due to saturation at around 95%. To estimate the viral titre the linear part of the curve was used, which results in 3×10^3 TU/µl. Further stocks of

lentiviral particles, containing modifications to the shc003 plasmid, but using the production method and titration assays that were established as part of this thesis, were produced by a member of our group, Dr Peter Cain.

3. Lentiviral delivery methods induce pro-inflammatory response in inflammatory cells and have low infection yields

The interest in using a lentiviral delivery system to induce a knock-down arose when we were offered the opportunity to screen a library of shRNAs targeting over 300 individual epigenetic enzymes. There is evidence available of successful primary monocyte transfection with lentiviral constructs [158]. As such, the system was optimised by taking into consideration a range of different factors, including the multiplicity of infection (MOI), transfection reagents and their concentrations and the concentration of the selective agent puromycin, to name a few.

Firstly, an optimisation using the shc003 containing lentiviral particles was conducted. Due to the presence of both, a GFP reporter gene and a puromycin resistance gene, the cells were checked for successful infection and integration by imaging and viability measurements. Puromycin IC_{50} on macrophages was determined to be 0.33µg/ml (Figure 14A); however concentrations in the range of 1-20µg/ml were used, because upon lentiviral integration, the puromycin resistance gene is expected, hence positively selecting successfully transfected cells. It was also noted that the recommended concentration of transfection reagent polybrene has a significant effect on macrophage viability (Figure 14B), but it is meant to facilitate

viral infection, so optimisation was carried out with and without polybrene. The optimisation was carried out by adding lentivirus up to MOI of 10, either with or without polybrene, selecting with a range of puromycin concentrations and on both, MCSF and GMCSF treated macrophages. However, no condition appeared to be leading to a successful transfection, based on very low viability even at highest viral loads, indicating the puromycin resistance gene was not present or active (Figure 14C).

Interestingly, the viability is consistently higher in the sample with MOI of 5 and 10, across both types of macrophages and with or without puromycin, indicating that

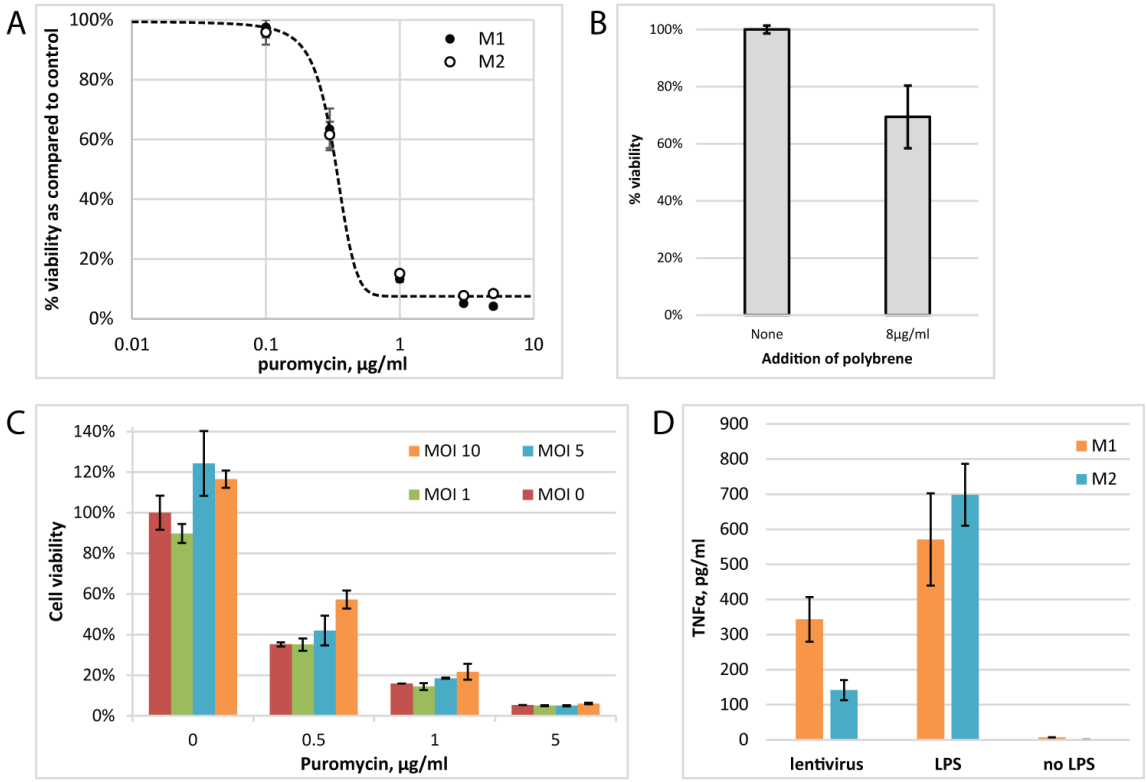


Figure 14 Optimisation of lentiviral infection. A. The toxicity of puromycin as determined by the IC50 curve and further used for positive successfully transfected macrophage selection. B. The effect of the recommended concentration of transfection reagent polybrene on macrophage viability. C. Lack of successful lentiviral transfection is evident by reduction of viability with increasing concentrations of puromycin despite the high lentiviral loads that carry the puromycin resistance gene. D. Macrophage incubation with lentivirus induces a pro-inflammatory effect, similar to the levels obtained by stimulating with LPS. Results (from n=2 experiments) are presented as the mean ± s.d.

the cells instead of getting transfected, sense the virus in the media and, due to this pro-inflammatory stimulus, start proliferating. Despite conflicting views on the ability of macrophage cells to proliferate, there have been a number of studies showing that a pro-inflammatory stimulus can induce proliferation by increasing expression of cyclins, particularly cyclin D2 [159, 160]. To follow up this observation, the production of pro-inflammatory cytokine TNFα was measured in the media. The cells incubated with lentivirus show a significant increase in TNFα which indicates an inflammatory response to the treatment, even though it does not reach the peak amounts as seen in LPS-treated macrophages (Figure 14D). This further supports the suspicion that rather than infecting the cells and integrating the DNA it may instead be simply causing an inflammatory response that also causes the cells to proliferate.

4. Human monocytes do not express genes under the control of the CMV promoter

Originally, due to lack of available information from the company providing the lentiviral constructs, it was thought that the puromycin resistance gene is controlled by the CMV promoter. It was also assumed the control plasmid has the puromycin resistance gene replaced by GFP, so it was also controlled by the CMV promoter. Following a literature search it turned out that the CMV promoter is not expressed in monocytes and macrophages, despite it being an active early-immediate promoter in most other cell types [161, 162]. Using high sensitivity fluorescence microscopy, imaging was carried out by measuring the emission wavelength in a step-wise fashion

of 9nm and observing the number of components that can be separated into separate peaks in the emission spectrum. It was determined that there is a residual signal of GFP in monocytes treated with lentivirus at MOI of 10 as the emission clearly peaks at different wavelengths – 470nm for blue excitation and 502nm for green (Figure 15A). This separation is not observed in monocytes cultured without lentivirus, where the peak emission for green and blue excitation overlap at 468nm (Figure 15B).

To further confirm this finding, a lentiviral construct was tried, that had the GFP gene under a different promoter, specifically hPGK. Primary human monocytes were transfected, along with the Thp1 human monocyte cell line and HEK cells for positive control. A range of polybrene concentrations and lentiviral MOIs were used. Firstly, just by comparing the monocyte Thp1 cell line and HEK cells it is evident

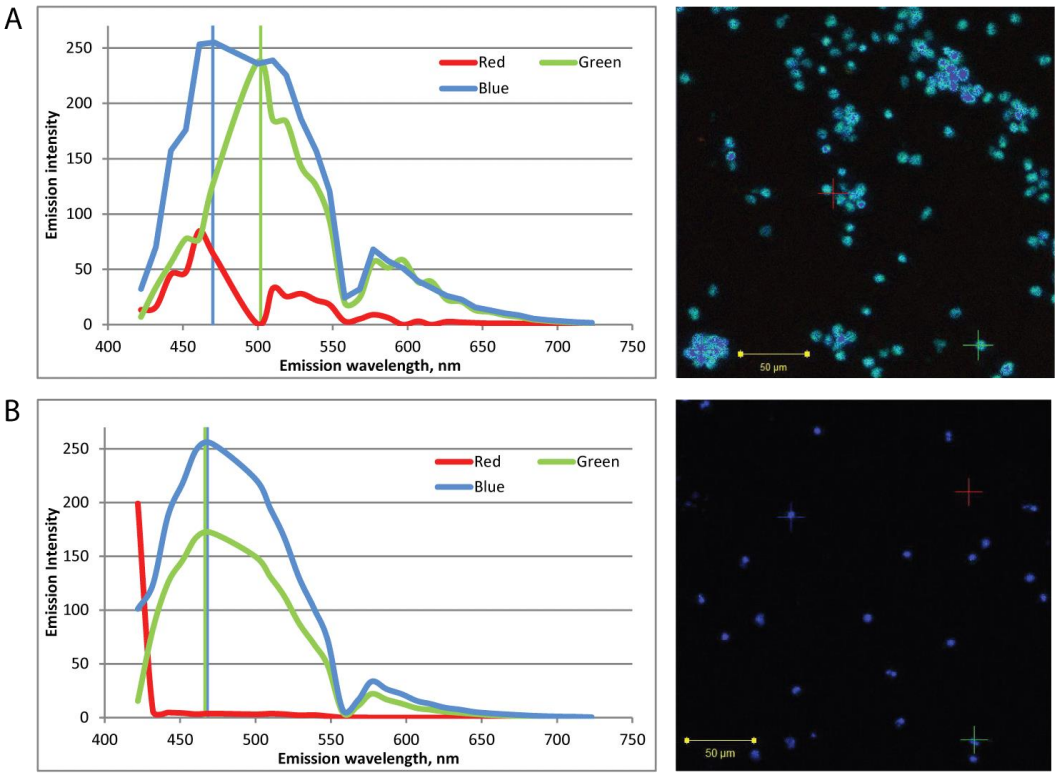


Figure 15 Primary human monocytes express only traces of GFP under the control of CMV promoter. The peaks of emission for each excitation wavelength are denoted by a vertical line. The crosses on the images represent the pixels that were used to obtain the emission data from.

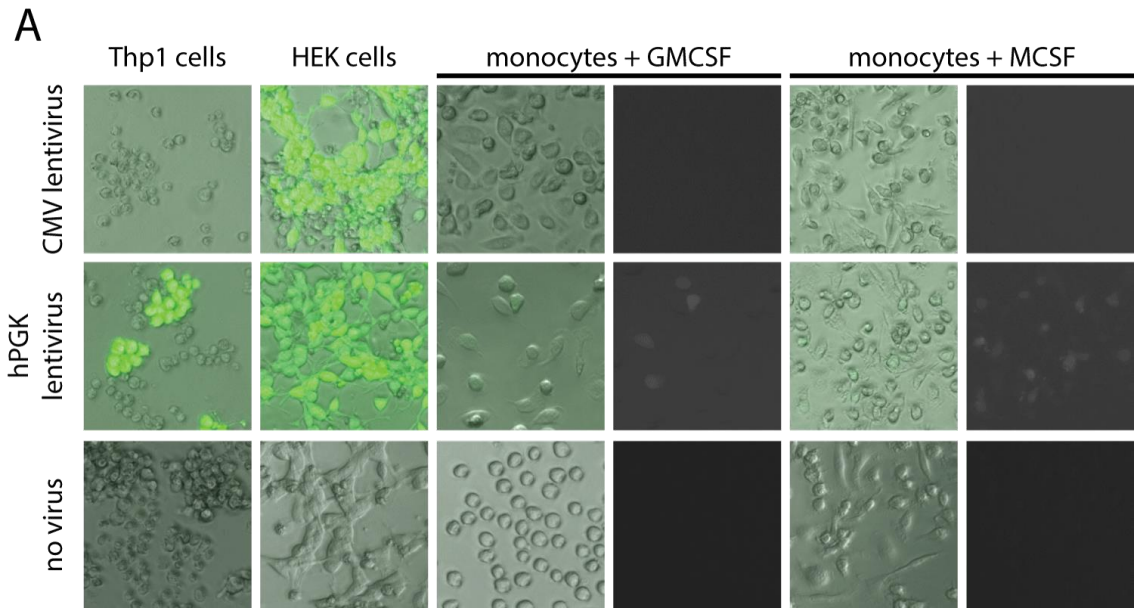


Figure 16 Comparison of GFP reporter gene expression between different promoters and cell types. The GFP channel and merged channels are provided only for the macrophages because of low but visible fluorescence within those channels.

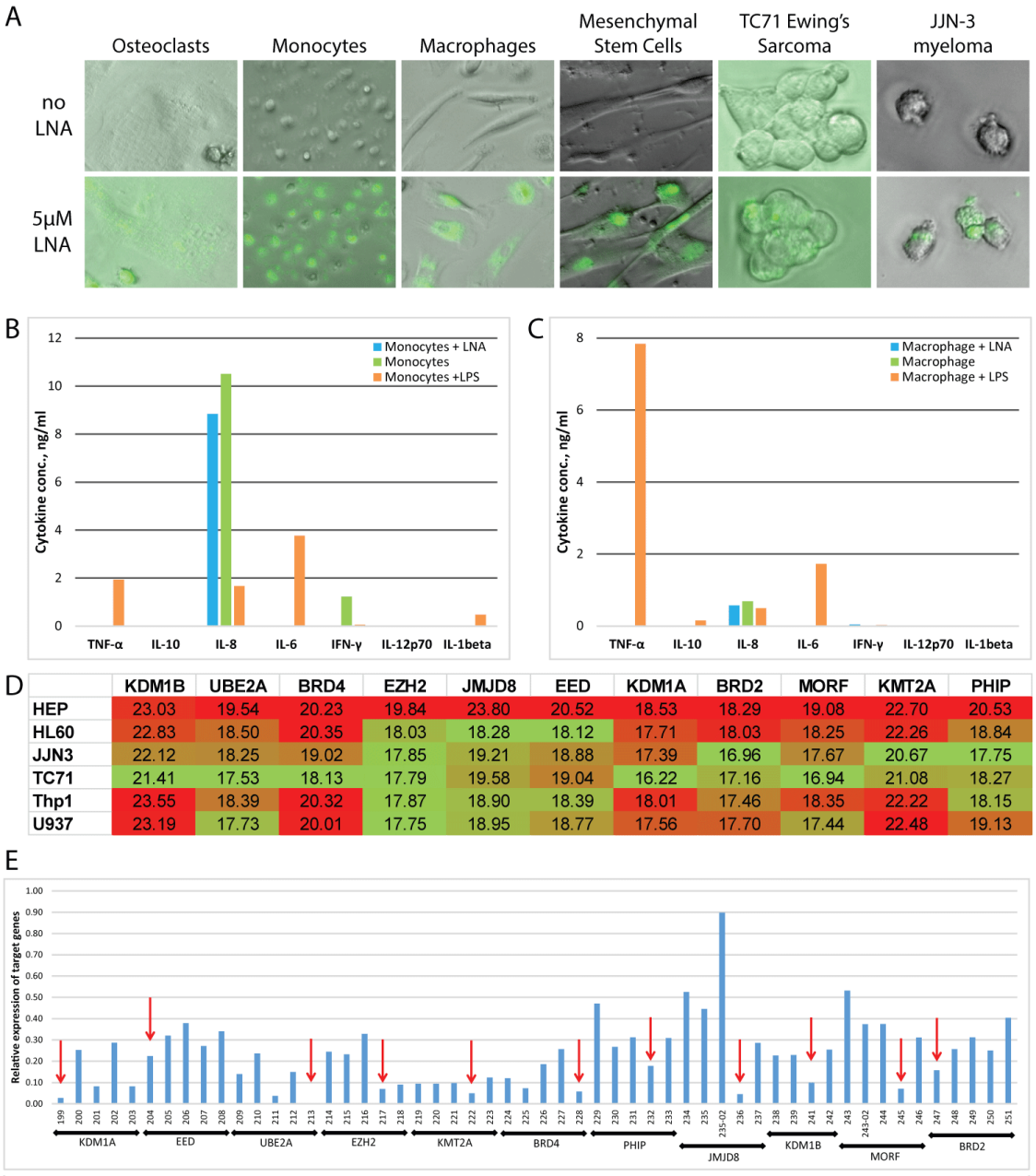
that the GFP reporter gene under the control of the CMV promoter is not expressed while if the promoter is replaced by hPGK, it yields expression levels comparable to those seen in HEK cells (Figure 16). Secondly, primary monocytes appear to have reduced levels of GFP even with the hPGK promoter. Finally, GMCSF treated monocytes are less likely to become transfected than MCSF treated monocytes, possibly due to the pre-existing pro-inflammatory nature of the cells.

5. Locked nucleic acids can penetrate all tried cell types without any observed adverse effects

Firstly, several series of imaging experiments were conducted in a range of cell types that are known to be difficult to transfect with common tools, such as lentiviral constructs or siRNAs. All investigated cell types exhibit high transfection efficiency

using FAM (fluorescein) labeled LNA molecule even without the addition of transfection reagents (Figure 17A). Addition of LNAs also does not induce any adverse effects, such as the expression of pro-inflammatory cytokines, as seen with lentiviral vectors, in monocytes (Figure 17B) and macrophages (Figure 17C). In the monocyte sample only, the non-stimulated monocytes without LNA showed a small increase in IFNγ which could be attributed to an experimental error, but elsewhere the monocytes and macrophages without LNA and with LNA exhibited the exact same pattern of cytokine expression. In the remaining treatments there is a sharp contrast between the stimulated monocytes or macrophages with LPS and those treated with LNAs or left untreated.

Next, a set of LNAs against a number of identified targets from a different screen [163]were synthesised. Their knock-down efficiency had to be validated and, as such, a cell line with high expression levels of most targets, if not all, had to be picked. Out of six cell lines tested, two, namely myeloma cell line JJN3 and Ewings sarcoma cell line TC71, had on average the highest expression of all targets of interest (Figure 17D). JJN3 was picked due to ease of handling of non-adherent cell culture and was incubated with all the LNAs against individual targets. Each target of interest had 5 different LNA molecules synthesised against it, and those LNAs with the highest knock-down efficiency (Figure 17E) were further validated in the experiments of other members of the research group. Importantly, the majority of the LNAs reduced the availability of target RNA to 10% or less which is a superior knock-down level even for well-established lentiviral or siRNA techniques.



	KDM1B	UBE2A	BRD4	EZH2	JMJD8	EED	KDM1A	BRD2	MORF	KMT2A	PHIP
HEP	23.03	19.54	20.23	19.84	23.80	20.52	18.53	18.29	19.08	22.70	20.53
HL60	22.83	18.50	20.35	18.03	18.28	18.12	17.71	18.03	18.25	22.26	18.84
JJN3	22.12	18.25	19.02	17.85	19.21	18.88	17.39	16.96	17.67	20.67	17.75
TC71	21.41	17.53	18.13	17.79	19.58	19.04	16.22	17.16	16.94	21.08	18.27
Thp1	23.55	18.39	20.32	17.87	18.90	18.39	18.01	17.46	18.35	22.22	18.15
U937	23.19	17.73	20.01	17.75	18.95	18.77	17.56	17.70	17.44	22.48	19.13

Figure 17 A. LNA transfection is fully efficient across all tried cell types. B-C. LNA transfection does not induce adverse effects in monocytes or macrophages. The LNA used for the experiment was a Fam-tagged scrambled LNA. D. A summary of target gene expression across available cell lines for LNA validation. Red represent the highest and green represents the lowest qPCR Ct values for each gene across the different cell lines. The cell line with the largest number of low Ct values will be the most suitable for validating LNAs because it is expressing all genes of interest to the highest extent. E. The validation of LNAs for each individual target in JJN3 cells. Arrows indicate the best knock-down efficiency. Results (from n=2 experiments) are presented as the mean.

Finally, the LNAs were employed to support the elucidation of histone demethylase KDM6B function in macrophage differentiation, as studied in collaboration with our industrial collaborators. It is known that KDM6B is vital for monocyte

differentiation into an M2 phenotype in mice, but KDM6B is dispensable for differentiation into M1 cells [102]. In M2 cells KDM6B induces differentiation into M2 cells by demethylating the inhibitory H3K27me3 marks around the transcription start site of IRF4. However, it was noted by our industrial collaborator Dr Palwinder Mander that the knock-down of KDM6A and KDM6B in human monocyte cell line Thp1 prevents their maturation into macrophages upon stimulation with PMA (Phorbol 12-myristate 13-acetate), which has a similar effect on Thp1 cells as GMCSF has on primary cells. It was decided to employ LNA technology to target KDM6B to assess their individual contributions to macrophage maturation from Thp1 cells.

Briefly, Thp1 cells were incubated for 8 days with LNAs against KDM6B or a control LNA that does not have a target in the human genome. Cells were then stimulated with PMA and samples for RNA were collected 6 hours later while media samples for the MMP9 (matrix metallopeptidase 9) assay were collected after 24 hours of PMA addition. MMP9 in this system served as a biomarker for mature macrophages.

KDM6B was effectively knocked-down by the specific LNA with 20% residual levels across all samples, as compared to the control LNA (Figure 18). Levels of H3K27me3 demethylase KDM6A were also measured because unpublished observations suggested that KDM6A has been upregulated to compensate for the reduction of KDM6B. Upon treatment with LNA against KDM6B, KDM6A is upregulated by 1.75-fold in the sample without PMA, but it remains at baseline in the PMA-treated sample. Finally, levels of MMP9 protein in the media rise by 3-fold with the addition of PMA and control LNA, however they remain at baseline levels if the cells had KDM6B knocked-down. This result indicates that KDM6B, without the need for

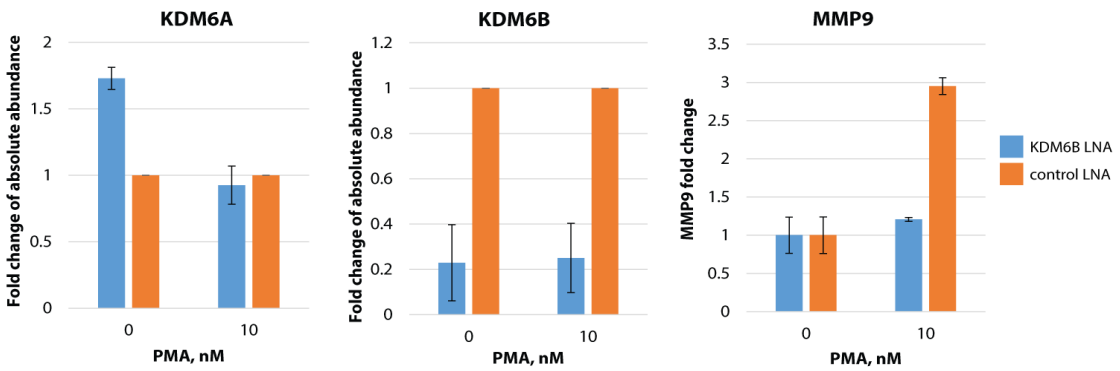


Figure 18 Expression levels of KDM6A and KDM6B were quantified by qPCR and normalised to control LNA within each treatment. MMP9 was measured by MSD (meso scale discovery), a technique similar to ELISA, and normalised to the samples without PMA. Results (from n=2 experiments) are presented as the mean ± s.d.

KDM6A can abolish Thp1 cell maturation into macrophages upon stimulation with PMA. The result was further confirmed by our collaborators with Western blots and similar experimental designs using small molecule inhibitors, such as GSK-J4 and targeted siRNAs.

6. Discussion

A number of techniques were applied in order to facilitate reliable gene knock-down in a variety of cell types. Not a single technique appears to be perfect: while offering high knock-down efficiency or while capable of knock-down in a large proportion of cells, it may cause unintended adverse side-effects, such as seen in macrophages where lentiviral transfection caused an inflammatory response.

Lentiviral delivery of shRNA constructs can indeed yield high transfection rates in certain cell types, as shown with HEK cells. However, care needs to be taken when designing the delivery plasmids. One important aspect to consider is the reporter gene which will be utilised to facilitate the selection of transfected cells. GFP and

puromycin resistance genes are the most common options, enabling researches to select transfected cells by fluorescence-activated cell sorting or by eliminating the non-transfected cells with antibiotics. Another aspect to consider when designing lentiviral shRNA delivery plasmids is the promoter used to transcribe the reporter gene. As shown in this chapter, different cell types exhibit different affinity for promoters and even weak expression of a puromycin resistance gene may not be sufficient to successfully select transfected cells. All these considerations pose a particularly difficult task to projects that design libraries of constructs to screen families of genes or even entire genomes.

As an alternative solution, locked nucleic acids (LNAs) performed particularly well in experiments shown in this chapter. They did not induce any adverse effects in immune cells, and the ease of transfection is unprecedented. The ability to simply add the LNAs into the growth media without any additional carrier or transfection reagent reduces the likelihood of side effects and simplifies treatments, eliminating potential errors in the process. However, the required incubation time of at least 5-7 days is a significant draw-back for primary cell cultures, such as monocytes, which require immediate growth or survival stimuli once isolated. On the other hand, LNAs seem to hold an enormous potential as therapeutic agents when a well characterised target for knocking-down is available.

Finally, a genome editing technique, discovered only in 2012, is rapidly gaining popularity and is called CRISPR-Cas9. It utilises a bacterial protein Cas9 that is used as a self-defence mechanism against viruses and, along with a guide RNA, it can target specific regions in the genome and introduce nicks or double-stranded cuts that will

end up being repaired by non-homologous end-joining that can result in insertions or deletions [164]. The latter results in a frame-shift or nonsense mutations that will impair the expression of the protein, hence resulting in a stable knock-out.

While not covered in this chapter, CRISPR-Cas9 constructs have been designed and used in experiments. Applications of this technique are rising in numbers as the Cas9 enzyme is modified by inducing only nicks, rather than double-stranded breaks, in the DNA, or by conjugating Cas9 to proteins that can then be targeted to user-defined sequences. The difficulty with this technique arises when a homogeneous population of mutants needs to be obtained. Because of the random nature of non-homologous end-joining, the gene-impairing mutation could be introduced in a number of locations, and it gives rise to a diverse set of cells. Some mutations may be less prominent than others, or heterozygous mutant cells can be established, that, if cultured in such a diverse population, will soon outcompete the double knock-outs due to their competitive advantages, if such exist because of the nature of the mutation. One solution is single-cell sorting, but raising a homozygous culture from a single cell can take months. Furthermore, validation of such knock-out cell cultures is only possible when there are sufficient cells, so multiple samples may need to be grown before one validated sample is chosen. Also, certain cell types will not proliferate if there is lack of cell-to-cell contact or missing cytokines in the media, which would otherwise be produced by neighbouring cells.

To sum up, gene knock-down techniques are evolving at a rapid pace. Mutual effort by the scientific community to improve existing methods and create new gene knock-down techniques is offering ever better opportunities to study gene function.

The lentiviral shRNA delivery system, LNAs and CRISPR-Cas9 have all been established in our research group through optimisation described in this chapter and is currently utilised in a number of independent projects.

Chapter IV

Epigenetics of Osteoclast Differentiation

1. Introduction

From the Discussion in Chapter II on macrophage biomarkers it may have become clear that macrophages are a diverse and versatile cell type with a spectrum of subtypes. Furthermore, upon additional stimulation macrophages can differentiate into osteoclasts, a cell type involved in bone remodelling and responsible for demineralisation of the bone tissue. Osteoclasts are central to clinical conditions of bone loss in which the balance between bone formation and resorption is lost. Osteoclasts arise from the myeloid cell lineage unlike the rest of the bone micro-environment, and were studied in this chapter using a small molecule inhibitor of histone lysine demethylases, noted to prevent osteoclast differentiation in vitro [165].

The bone micro-environment consists of a comparatively small number of different cell types and it is an ever regenerating tissue (Figure 19). Firstly, bone formation is a result of osteoblast activity. Mesenchymal stem cells upon stimulation with bone morphogenic proteins (BMPs), or indirectly by leptin, differentiate into osteoblasts, and while they have a highly specialised function of laying mineralised bone tissue, they differ from fibroblasts mainly by enhanced signalling of transcription factor CBFA1 (Core-Binding Factor Subunit Alpha-1, also known as RUNX2; runt-related transcription factor 2) [166]. Secondly, osteoblasts can turn into osteocytes by rapidly

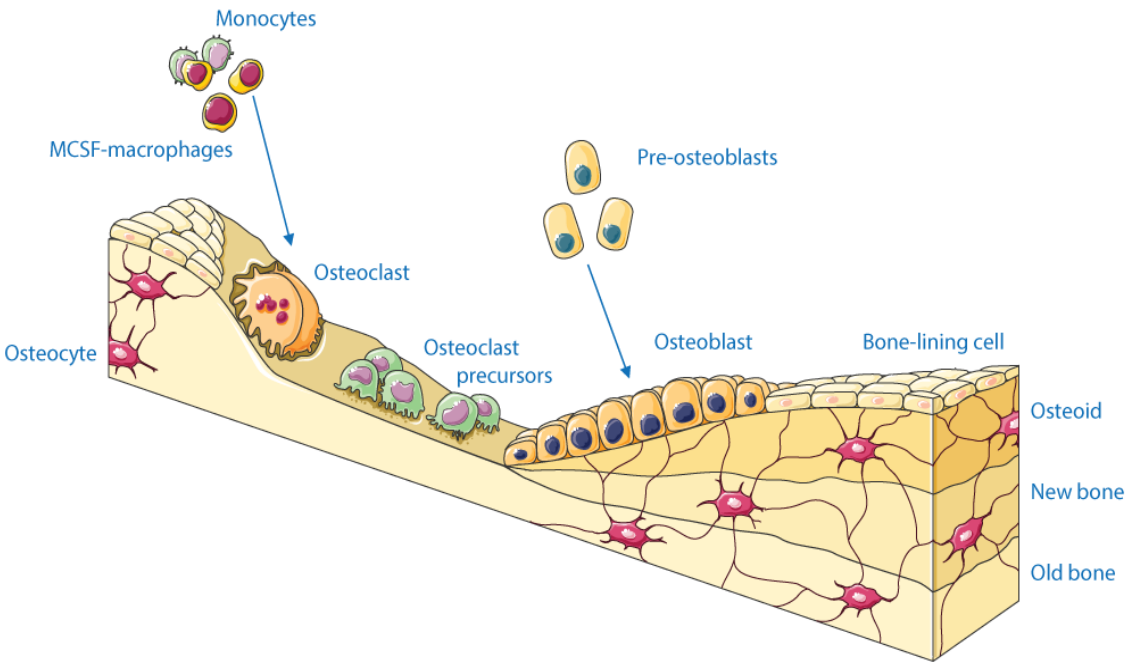


Figure 19 The bone micro-environment. New bone is laid by osteoblasts. They can get entrapped in the newly laid bone matrix and become osteocytes. Bone is resorbed my osteoclasts that differentiate from MCSF-induced macrophages. Before editing, image sourced from Servier Medical Art under a Creative Commons Attribution 3.0 Unported License.

producing three times the cell volume of osteoid, an extracellular cell matrix consisting of proteinaceous (such as collagen) and mineral components, and as a result they become trapped in the newly laid bone matrix [167]. Finally, osteoclasts arise from a monocyte progenitor that due to stimulation with MCSF is expressing RANK (receptor for activation of NF-κB). Upon stimulation with RANKL (RANK ligand) they fuse to form a multi-nucleated polykaryon that induces bone resorption by acidifying an isolated area under the surface of the osteoclast [168] (Figure 20). Bone resorption is achieved by employing an ATPase proton pump, necessary to produce the required hydrochloric acid (HCl) for resorption [169]. This leads to the pH dropping as low as 4.0, release of the inorganic component of the bone matrix, and collagen, digested by proteases, can be taken up by endocytosis and processed in the lysosome by CTSK (Cathepsin K) [170].

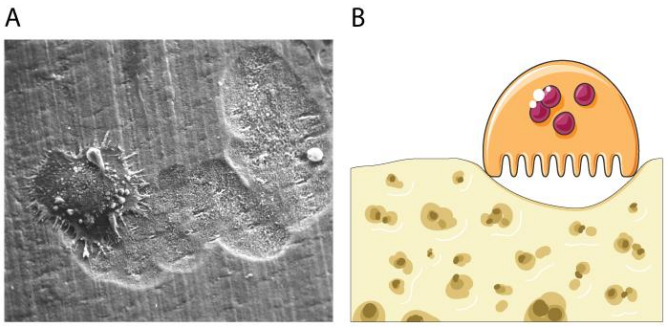


Figure 20 A. An electron microscopy image of an osteoclast cell moving along the surface of a bone and leaving a groove of degraded bone behind. Reproduced from www.ucl.ac.uk/cdb/research/arnett where it is freely available for non-commercial use. B. A schematic showing the osteoclast cell on top of bone tissue and establishing an enclosed area underneath which is degraded by acidic pH and proteases. Image sourced from Servier Medical Art under a Creative Commons Attribution 3.0 Unported License.

The main players for osteoclast differentiation are RANK, RANKL and the decoy receptor OPG (osteoprotegerin, also known as TNFRSF11B, tumor necrosis factor receptor superfamily, member 11b). RANKL is expressed by osteoblasts as well as T-cells, with the latter involvement showing a possible link to inflammatory conditions that could lead to increased bone loss [171]. Mice homozygous for RANKL knock-out lack osteoclasts completely, despite having healthy precursor macrophages [172]. An identical phenotype is observed in mice lacking RANK [173], indicating that this specific ligand-receptor combination is indeed fundamental for osteoclastogenesis. While RANK is a membrane-bound receptor, OPG, a decoy receptor for RANKL, is secreted by bone marrow stromal cells and it can quench the effects of RANKL hence reducing the number of osteoclasts as witnessed in *in vitro* and *in vivo* experiments [174].

Notably, osteoclasts are multi-nucleated cells that form by numerous macrophages fusing into one. DC-STAMP (dendritic cell–specific transmembrane protein) was identified as one factor that is involved in cell fusion [175]. Knock-out mice for DC-STAMP were completely lacking any multi-nucleated TRAP-positive (Tartrate-

resistant acid phosphatase stain is a specific marker for osteoclasts) cell, although many mononuclear TRAP-positive cells were observed. The same study also indicates that while there is resorption observed, it is highly reduced compared to wild-type multi-nucleated osteoclasts. The latter is probably because of a much smaller surface volume that a single macrophage can cover, compared to a large osteoclast, in order to establish an isolated resorptive area.

Finally, on a gene regulatory level, RANK stimulation induces a signalling cascade controlled by the transcription factors NFATc1 and AP1 (activated protein 1), a heterodimer of transcription factors c-Jun and c-Fos [176]. The previously mentioned TRAF6 appears to be the vital molecule to transduce the signal from RANK, leading to activation of NF-κB and JNK (Jun N-terminal kinase) pathways [177]. A different study showed c-Fos, a member of the Fos family, along with members of the Jun family to be induced to form the dimeric AP-1 transcription factor [178]. Individual knock-outs of both, TRAF6 and c-Fos, showed nearly complete abrogation of NFATc1 expression [176]. A search using MotifMap [179] upstream of NFATc1 identified a binding site for NF-κB at -504bp and a different one for AP family members at -466bp. While the exact mechanism of action remains unknown, clearly either one of the two, or both transcription factors are required to induce an NFATc1 transcriptional programme resulting in osteoclastogenesis.

Aside from RANKL stimulation, there also exists an alternative path to induce NFATc1. Adapters harbouring ITAM (immunoreceptor tyrosine based activation motif) can associate with immunoglobulin-like receptors which undergo phosphorylation. Subsequently, this induces calcium signalling and activates PLCγ

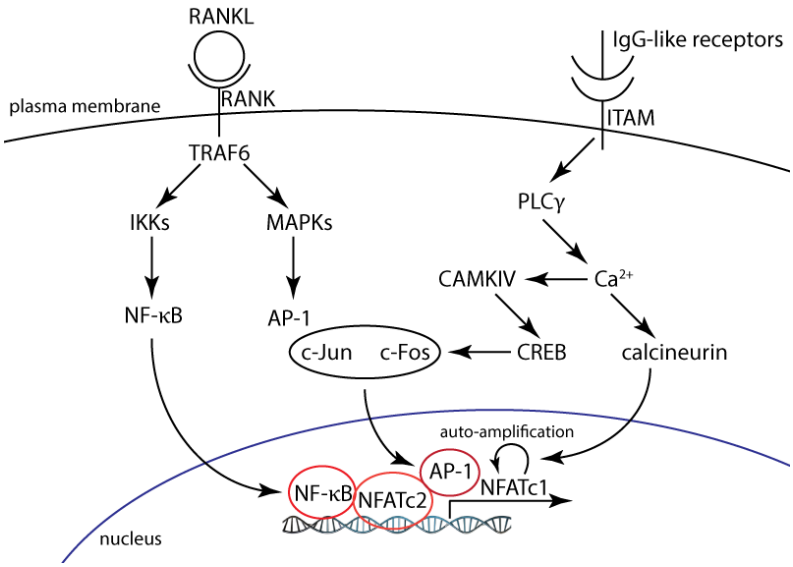


Figure 21 Signalling through RANK and ITAM leads to the activation of NF-κB, AP-1 and calcineurin which induce the expression of NFATc1. NF-κB and AP-1 bind upstream of NFATc1 transcription start site while calcineurin enables auto-amplification of NFATc1 expression.

(phospholipase γ). Calcium signalling activates CaMK/CREB and calcineurin, both of which lead to the activation of NFATc1 expression [180]. Interestingly, osteoclastogenesis can be completely abolished by inhibiting calcineurin [181], so instead of there being alternative pathways to the activation of osteoclast differentiation, they are all complementary and reinforce each other (Figure 21).

As mentioned in the beginning of this chapter, excessive numbers of osteoclasts are frequently implicated in the increased bone loss in a number of conditions. Inhibiting osteoclast function and formation is desirable in clinical conditions of excessive bone resorption (such as osteoporosis, multiple myeloma and others). However, only few therapeutic principles are established, such as bisphosphonates which have high affinity for bone and they inhibit farnesyl pyrophosphate synthase, a key step in isoprenoid synthesis, or anti-RANKL therapeutic denosumab [182]. The latter is an antibody based therapy targeting specifically RANKL, which are known to be

expensive and difficult to manage due to properties of antibodies and their sensitivity to temperature. Clearly, there is a need for an increased diversity of treatment options, and particularly a single treatment that is anti-resorptive, anti-tumour and anti-inflammatory, as will be discussed in Chapter V in relation to multiple myeloma.

In this chapter a previously mentioned small molecule GSK-J4, inhibitor of histone lysine demethylases KDM6A, KDM6B and KDM5B, was used to prevent osteoclast differentiation. Treated cells were used to determine differential enrichment of H3K4me3 and H4K27me3 histone marks as compared to RANKL-induced osteoclasts and differential gene expression was measured and compared to chromatin changes. Additionally, new biomarkers were identified for early and late osteoclast differentiation which are now used for high-throughput small molecule screens.

2. Osteoclast biomarkers

Firstly, when whole transcriptome data was obtained for RANKL-treated MCSF-macrophages, it was confirmed that the differentiated cells are expressing osteoclast biomarkers. There is a number of previously described markers specific for osteoclasts. For example, MMP9 (matrix metalloproteinase-9), a type IV collagenase important for extracellular matrix degradation, is upregulated through by RANKL signalling via the TRAF6 pathway [183]. CA2 (carbonic anhydrase 2) is an early osteoclast marker that is important for differentiation as well as maintenance of low pH and low levels of calcium ions in the cell [184]. The previously mentioned CTSK

and DC-STAMP are also among a few other biomarkers used to confirm osteoclast differentiation.

The majority of markers mentioned here and in the literature were identified after 7 days of RANKL treatment (Table 1). On the other hand, markers such as NFATc1 are absent from this table, because transcription factors are potent inducers of down-stream effects, and an increase of only 6-fold over macrophages without RANKL stimulation, as is the case with NFATc1, is required to have a large effect on cells.

Gene	MCSF	MCSF+ RANKL	Fold Change
FAM89A	1.34	714	531
CKB	6.47	2036	315
COL6A2	5.03	1424	283
ITGB3	0.51	128	251
COL6A1	3.90	846	217
CTSK	18.78	3737	199
AK5	0.74	127	171
NCS1	0.51	66	129
MYO1D	1.67	205	122
COL11A2	1.15	135	117
CA2	0.81	93	115
TMEM132A	0.70	79	112
CSPG4	0.50	48	96
CDC42EP5	0.86	78	91
C19orf71	18.70	1569	84
HOXC12	1.79	139	78
SLAMF7	0.55	40	72
TNFSF15	1.07	74	69
SLCO4A1	0.93	62	67
ATP6V0D2	0.54	28	52
COL18A1	0.67	34	51
TRMT61A	1.57	79	50
NRIP3	2.31	116	50
NEURL3	1.44	71	50

Table 1 A list of top 24 genes that are expressed at least 50-fold more in osteoclast samples compared to macrophages. The list is ranked by fold change and the values for MCSF and RANKL samples are provided in FPKM (fragments per kilobase of Exon per million fragments mapped).

Furthermore, a number of genes that are highly up-regulated, such as NEURL3 (neuralized E3 ubiquitin protein ligase 3), TMEM132A (transmembrane protein 132A), CDC42EP5 (cell division cycle 42 effector protein, Rho GTPase binding, 5) or FAM89A (family with sequence similarity 89, member A), to mention a few, have not been previously described as being involved in osteoclast differentiation or function.

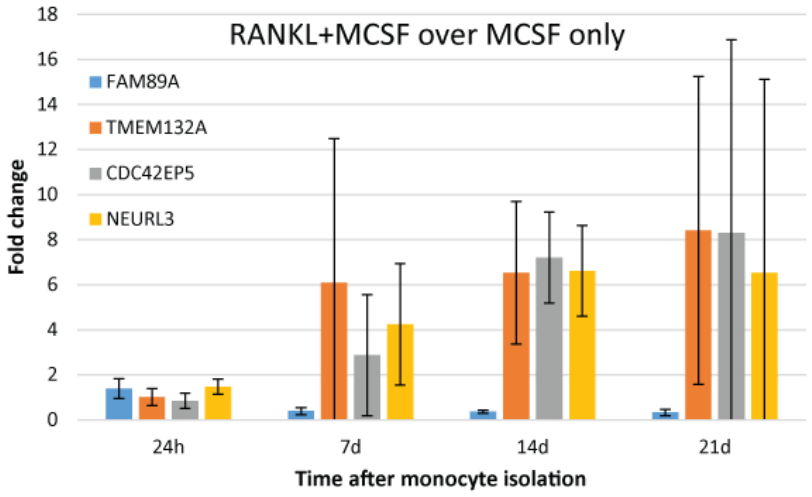


Figure 22 Fold change expression of possible new biomarkers in osteoclasts as compared to MCSF-induced macrophages, measured at various time-points after monocyte isolation and immediate treatment with RANKL+MCSF or MCSF alone. Results (from n=3 experiments) are presented as the mean ± s.d.

As such, an effort has been made by a member of our research group, Na Wu, to validate some of these newly identified putative biomarkers by qPCR. Briefly, $CD14^+$ monocytes from three different donors were isolated and treated either with MCSF alone or with RANKL and MCSF for 21 days. RNA was isolated after 1, 7, 14 and 21 days of treatment. Expression levels of genes of interest were determined by qPCR and shown as fold change of expression in RANKL and MCSF treatment over expression in the cells treated with MCSF alone (Figure 22).

Out of the putative biomarker genes, only FAM89A appears to be nearly at baseline across all time-points. Although it must be noted that the Ct values were above 26, so there is a chance that the gene is expressed at very low levels and larger cell numbers may be required to confidently show differential expression between the treatments. The remaining biomarkers are expressed on average more than 6-fold in osteoclasts as compared to MCSF-treated macrophages. It is also noticeable how

there is an increase of expression leading up to the 14 day time-point, when all three donors appear to reach peak expression of the biomarkers. Later, there can be a further increase or reduction in all markers, depending on the donor. The latter fact illustrates the difficulty of working with primary tissue, because some donors are slow responders, while others can be even complete non-responders. In this chart (Figure 22) day 14 is the time-point when expression evens out across donors and because of that, this time-point along with the 24h time-point post-RANKL treatment will be later used to better understand osteoclast differentiation.

To sum up, based on a number of already known biomarkers, it was confirmed that the cells investigated are osteoclasts. Furthermore, a comparison of the transcriptional profile of RANKL treated cells against MCSF-treated macrophages allowed to identify a set of new biomarkers for osteoclast differentiation, although their exact functions remain to be studied. Finally, the latter finding was confirmed by qPCR and due to large donor-to-donor variation, samples treated for 7 days and 14 days were chosen for the next set of experiments as time-points when the variation between samples could be reduced.

3. GSK-J4 affects a specific set of transcription factors important for osteoclastogenesis and otherwise induced by RANKL

It has been previously seen in unpublished observations [165] that the small molecule inhibitor GSK-J4, specific for histone demethylases KDM6A, KDM6B and KDM5B [85, 126], effectively inhibits osteoclast differentiation from MCSF-treated

macrophages (Figure 23A). Furthermore, with the help offered by Dr Clarence Yapp on a novel imaging technique based on second-harmonic generation [185], it was possible to accurately measure the area of resorbed bone on dentine slices, and it was shown that cells treated with RANKL and GSK-J4 have a dramatically reduced capacity to resorb bone, as compared to RANKL only, or RANKL and GSK-J5, an inactive isomer of GSK-J4, treated cells (Figure 23B). As such, a more elaborate effort was launched in order to understand the changes occurring on a transcriptomic and epigenetic level (more about the latter in the next subchapter).

Briefly, there have been two independent attempts carried out with different donors. Primary $CD14^+$ monocytes were isolated and allowed to mature to macrophages by treating them with MCSF for 7 days. Next, macrophages were pre-treated for 1 hour with either DMSO (carrier control) or GSK-J4 before adding RANKL. Samples for

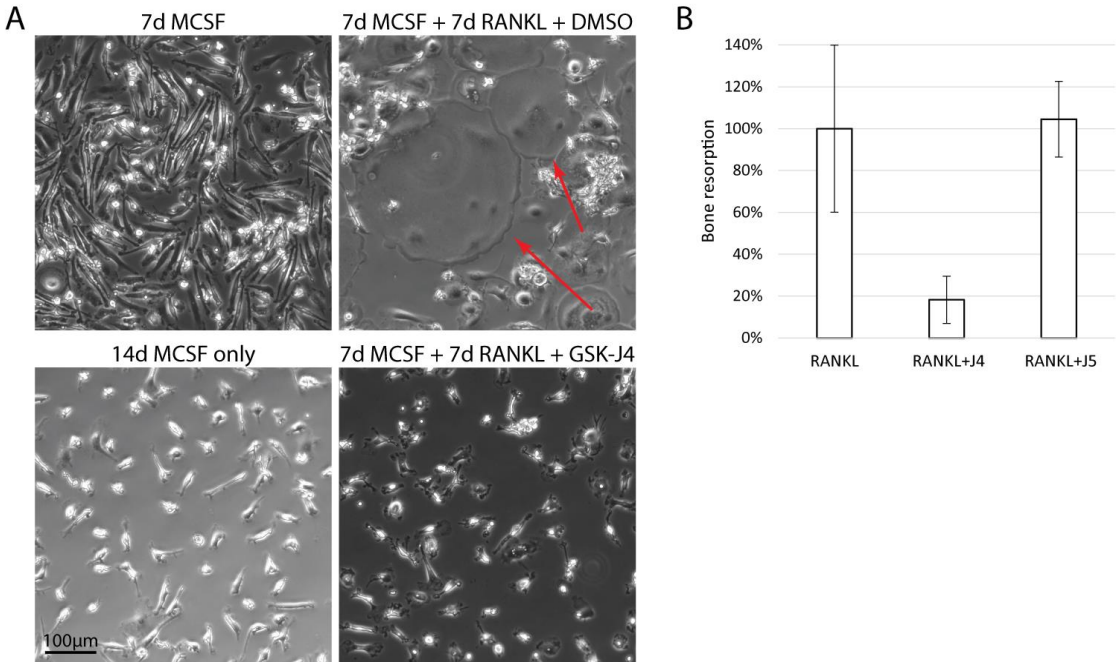


Figure 23 A. Large multi-nucleated osteoclast cells (indicated with red arrows) are visible after 7 days of treatment with RANKL-DMSO, but only macrophage-like cells are seen after 7 days of treatment with RANKL+GSK-J4. B. Bone resorption is significantly reduced in RANKL and GSK-J4 treated cells, as determined using second-harmonic generation imaging technique. The samples are normalised to the volume of boned resorbed by cells treated with RANKL only.

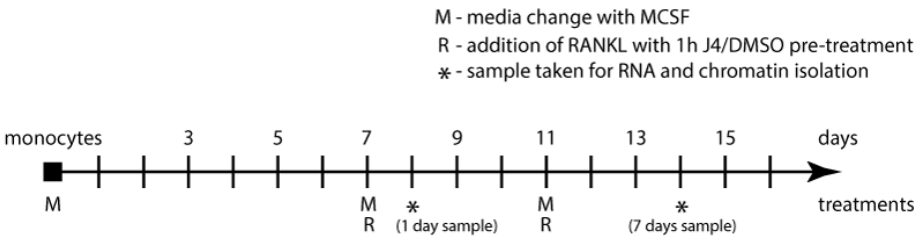


Figure 24 Timeline of osteoclast differentiation and treatments. Cell growth media contained MCSF at all times and it was replaced on day 7 and 11. RANKL was added on day 7 and 11 after 1 hour pre-treatment with GSK-J4 or DMSO. Samples for RNA were taken on days 8 and 14 which is 1 day and 7 days respectively after RANKL treatment. The latter annotation of 1 and 7 days of RANKL treatment is used in the remainder of this chapter.

RNA isolation were taken at 24h and 7 days after treatment with RANKL (Figure 24). Ribosomal RNA-depleted total RNA samples were used to produce indexed libraries for sequencing with an Illumina high-throughput sequencer. 80-100bp long paired-ended reads were obtained and further processed and normalised using the *bowtie* and *cufflinks* pipeline. Genes with at least a 2-fold change over negative control were used in the subsequent analysis.

Firstly, sample relatedness was assessed by employing hierarchical clustering in a dendrogram (Figure 25A). Samples of different time-points are clustering in separate branches of the highest order α. As seen in the microscopy images (Figure 23), samples that were treated for 7 days with RANKL+GSK-J4 and 14 day MCSF-treated macrophages have identical cell morphology, and hence it may not be surprising to see them cluster together in the branch β, while the RANKL+DMSO sample, which was confirmed to be a differentiated osteoclast, clusters on a separate branch. On the other hand, somewhat surprisingly, all day 1 samples, whether treated with RANKL and J4, or DMSO, or MCSF only, cluster all together into a single branch γ. Statistically the algorithm only allows clustering of pairs, but because the dispersion

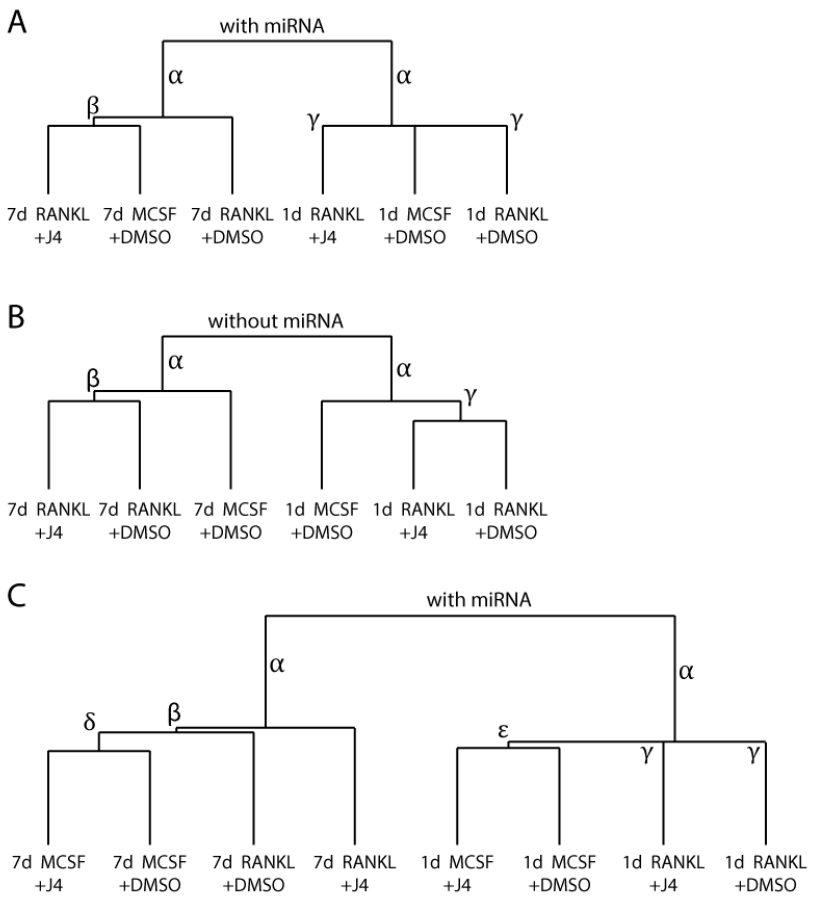


Figure 25 Sample hierarchical clustering expressed via dendrograms showing sample relatedness, based on normalised RPKM expression values. Only genes with variance of at least 2 across all samples are included. A. Clustering of samples if all genes are included. B. Clustering of samples if miRNA genes are excluded. C. Clustering of all previously used samples as well as the MCSF+J4 sample when all genes are included in the analysis.

between the RANKL-treated sample branch and MCSF is so small, visually it appears as a trident branch. This could indicate that during the early stages of differentiation the effects of J4 treatment are not global and rather affect a very specific subset of genes.

Curiously, if a hierarchical clustering model is built by excluding the micro RNA (miRNA) genes (Figure 25B), the lower level β and γ branches are affected in such a way that RANKL samples cluster together leaving out the MCSF sample on its own branch. Out of 1596 miRNA genes that are available in the annotation, only a small

subset of highly overlapping 62 miRNAs on day 1 (Supplementary Table 5) and 55 miRNAs on day 7 (Supplementary Table 6) seem to be exhibiting such a large effect on clustering. This strongly suggests miRNAs are fundamentally involved in osteoclast differentiation and the effects of GSK-J4 may manifest themselves through the miRNA machinery.

Finally, the inclusion of an MCSF only sample treated with J4 at both time-points, significantly alters the landscape of the entire dendrogram (Figure 25C). As usual, top level α level branches separate the time-points with high confidence dispersion. However, now the MCSF samples cluster together in the lowest level branches δ and ε, indicating that GSK-J4 effects on MCSF-treated macrophages are fairly insignificant compared to RANKL treatment. On the other hand, while preventing osteoclast differentiation as seen before, this also shows that GSK-J4 treated RANKL-stimulated MCSF-macrophages are not remaining completely idle to RANKL stimulation and have a differential response that is somewhere in between fully differentiated osteoclasts and MCSF-macrophages, as indicated by branch β.

Next, a gene ontology (GO) analysis was carried out to better understand the role of genes that have the largest differences in expression between samples (Figure 26). Within the 1 day branch, approximately 36% of genes upregulated in the RANKL+DMSO samples are related to regulation of transcription, cell cycle control and chromatin remodelling. This is expected because, as already shown in the literature, RANKL stimulation induces a specific set of transcription factors, with NFATc1 being the most important one that starts osteoclastogenesis. The latter requires expression of a different set of genes, so transcriptional activation and

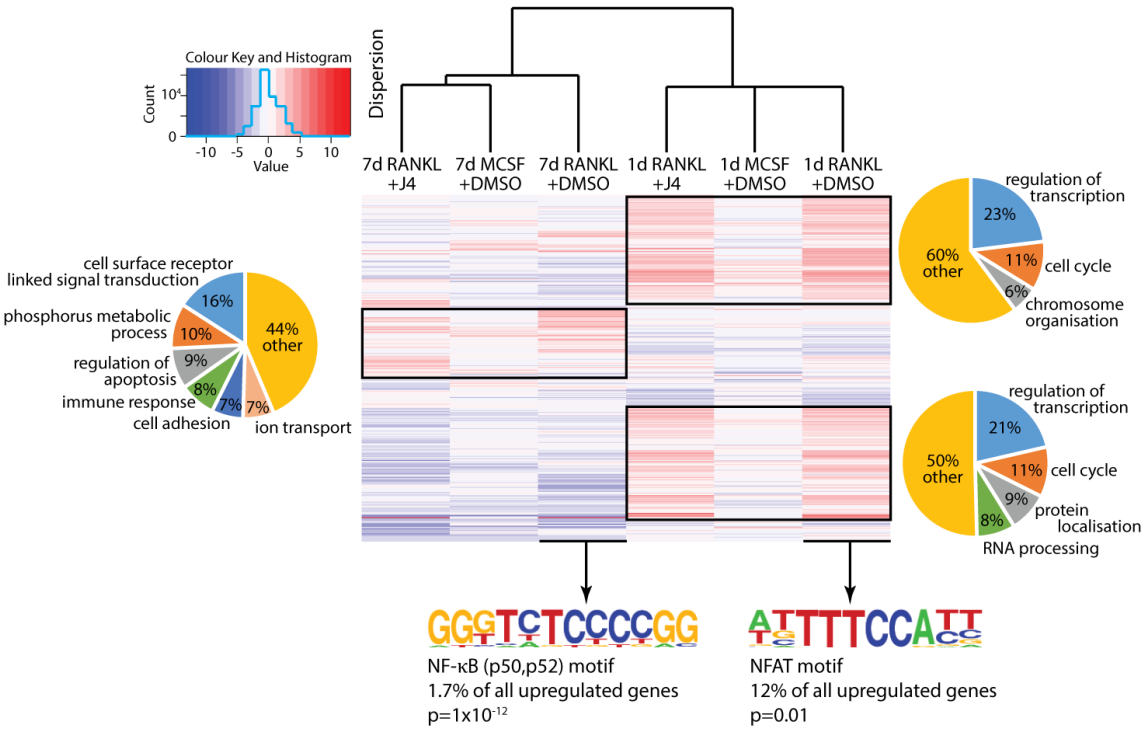


Figure 26 A heatmap of expression of 23504 genes. Each line represents one gene with its abundance value expressed as a log2 fold change over baseline (either 1 or 7 days MCSF only sample, without DMSO). The distribution of values is shown in a histogram visible on top of the heatmap. Pie charts represent the percentage of genes falling into the shown categories of GO annotations. The consensus motifs at the bottom of RANKL+DMSO columns indicate what top regulatory sequences are present upstream of the up-regulated genes within RANKL+DMSO sample as compared to MCSF+DMSO.

chromatin remodelling are the top upregulated mechanisms after 1 day of RANKL treatment. A notable subset of upregulated genes are epigenetic enzymes affecting histone marks, such as KAT5 (lysine acetyltransferase 5), BRD8 (bromodomain containing 8), EZH2, SET domain containing proteins and, importantly, KDM5B (Supplementary Table 7).

On the other hand, after 7 days of treatment with RANKL, the majority of genes upregulated within the first day are down-regulated and genes with an osteoclast-related function are increased significantly. Genes coding for ADAM (a disintegrin and metalloproteinase) domains 2, 7 and 18, ACTN1 (actinin, alpha 1), CHST10 (carbohydrate sulfotransferase 10) and CADM1 (cell adhesion molecule 1), among

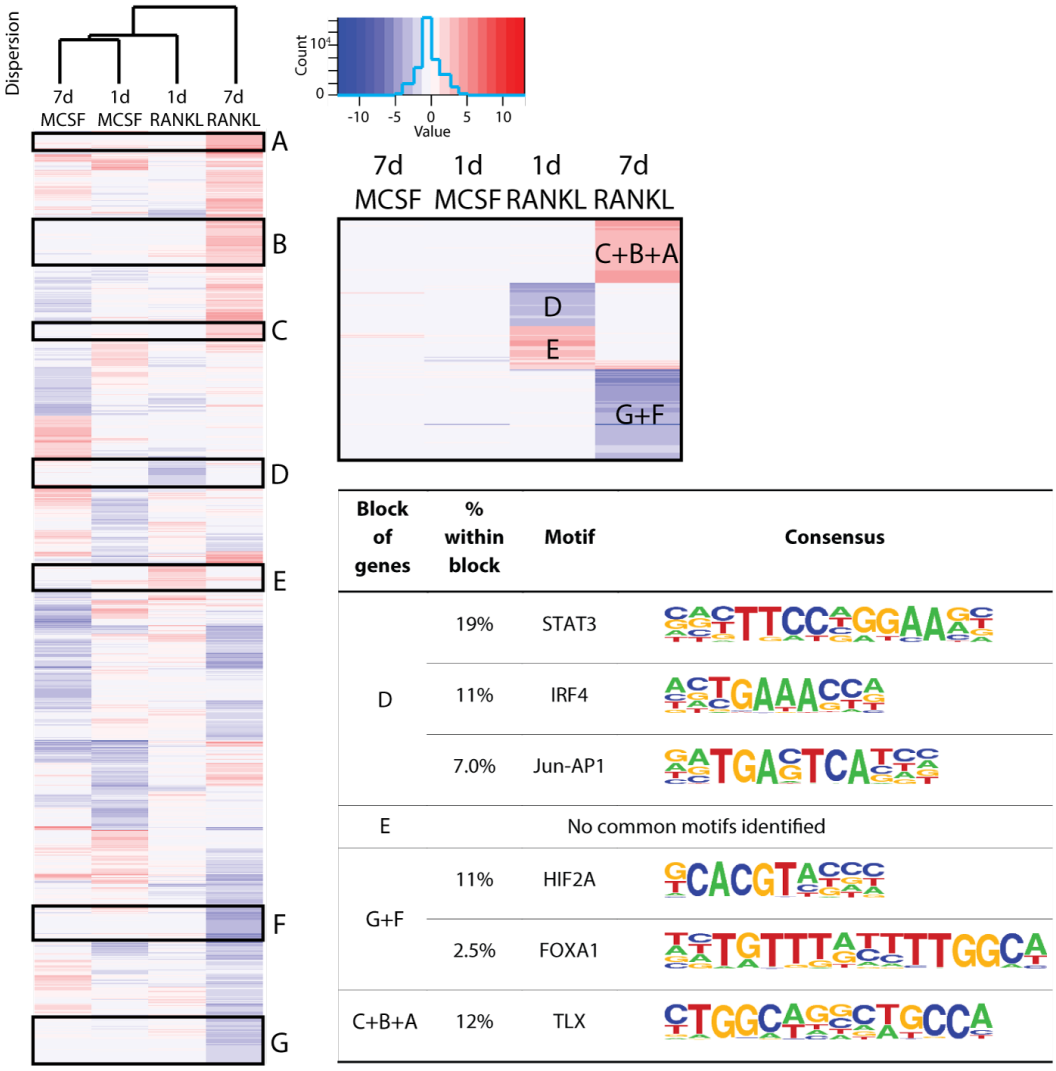
many others, are responsible for cell adhesion (Supplementary Table 8), and are fundamental to osteoclast function [186, 187]. Another set of genes involved in cell surface receptor signalling is also seen as fundamentally different from MCSF macrophages taken at the same time-point. Differentially expressed genes include ATP6AP1 (ATPase, H+ transporting, lysosomal accessory protein 1), NFAM1 (NFAT activating protein with ITAM (immunoreceptor tyrosine-based activation motif) 1), BAI1 (brain-specific angiogenesis inhibitor 1) and a list of chemokine receptors and ligands (Supplementary Table 9), most of which have important functions in osteoclast homeostasis and signalling.

It was of interest to find out if the genes, upregulated at least 2-fold as compared to MCSF-treated cells, were under control by any common transcription factors. This was determined by analysing the abundance of a known transcription factor binding consensus motif upstream of the gene transcription start site (TSS). Expectedly, an NFAT motif was found to be upstream of 12% of all upregulated genes after 1 day of treatment with RANKL (Supplementary Table 10). This is very much in agreement with available literature and known RANKL signalling pathways leading to the activation of NFATc1, as discussed previously. On day 7 of RANKL treatment the profile of gene expression switches to maintain the function of an osteoclast, and as a result there are more diverse gene sets. 1.2% of all up-regulated genes as compared to MCSF treatment have an NF-κB binding site upstream of their TSS. Because osteoclasts arise from immune cells and NF-κB is known to be one of the signalling paths leading to NFATc1 activation, it appears like this is the signalling pathway that

is sustained throughout osteoclast differentiation and later cell maintenance by expressing related cell cycle genes, relevant cytokines and cell surface receptors.

Interestingly, treatment with GSK-J4 visually appears to upregulate a very similar set of genes as with DMSO, which suggests that while morphologically the cells do not form large multi-nucleated cells like osteoclasts (Figure 23A) and they also have up to 90% reduced capacity to resorb bone (Figure 23B), as discussed before, transcriptionally they are still similar to osteoclasts. To better understand what exactly could be impaired to result in such outcomes, a more focused comparison of RANKL-treated cells with DMSO and GSK-J4 was carried out (Figure 27). As a negative control, MCSF-treated macrophages with DMSO and GSK-J4 were used, as these samples would eliminate false positive signal – genes affected by J4 non-specifically in both cell types, while only specific effects within the RANKL treated cells were of interest.

Cluster D (Figure 27) represents all genes that are down-regulated with GSK-J4 treatment as compared to the DMSO control in RANKL-treated cells on day 1, but are unaffected by GSK-J4 treatment in any of the other samples or time-points. Importantly, the most prevalent TF motifs were identified upstream of the genes that are up-regulated with RANKL treatment but down-regulated if pre-treated with GSK-J4 before stimulation with RANKL. As a result, STAT3, IRF4 and Jun-AP1 were determined to be the transcription factors whose signalling pathways are affected by GSK-J4 treatment. Jun-AP1 is of particular interest, because as discussed before, a binding site for members of the AP1 family exists upstream of the NFATc1 TSS. Similarly, IRF4 was shown to be directly implicated in the increased expression



Block of genes	% within block	Motif	Consensus
D	19%	STAT3	
	11%	IRF4	
	7.0%	Jun-AP1	
E	No common motifs identified		
G+F	11%	HIF2A	
	2.5%	FOXA1	
C+B+A	12%	TLX	

Figure 27 The heatmap shows the log2 fold change of gene expression in GSK-J4 treated samples over DMSO control within the different time-points and cytokine treatments. Each line represents one gene and there are 16538 genes in total. Genes that are up- or down-regulated in GSK-J4 as compared to the respective DMSO control were pooled into 4 clusters, as shown in the separate heatmap. Each cluster was analysed for common transcription factor consensus motifs upstream of the gene TSS and the results are summarised in a table showing the cluster, % of genes with that motif within the gene cluster, a transcription factor name known to bind to that consensus sequence and the sequence itself.

of NFATc1 during osteoclastogenesis [188]. On the other hand STAT3 was shown to affect osteoclastogenesis negatively [189], so it is slightly surprising to see STAT3-regulated genes go up in the RANKL-treated cells but be inhibited in the GSK-J4 treatment.

Cluster E (Figure 27) represents genes that are upregulated by GSK-J4 treatment in cells stimulated with RANKL for 1 day. However, none of these genes would form a subset for any known transcription factor signalling pathway, like it did in Cluster D.

Consistent with the observation that RANKL treatment over time induces maturation of osteoclasts, after 7 days of treatment with RANKL, changes in the transcriptional program were observed. The binding motif for HIF2A (hypoxia-inducible transcription factor 2A) is significantly enriched among 11% of genes upregulated in RANKL-treated cells, but is inhibited by GSK-J4 addition. HIF2α has previously been demonstrated to play a role in osteoclast function [190]. It is also noted in the same study that while mRNA levels are high, the protein only becomes detectable under hypoxic conditions. Hypoxia of up to 0.2% O_2 increases the size of osteoclasts. Among genes upregulated by HIF2A are HOXA13 (homeobox A13), WISP3 (WNT1 inducible signaling pathway protein 3) and CCNA1 (cyclin A1) all of which play important roles in cell cycle regulation in osteoclasts [191]. Other transcription factors identified in this analysis, such as FOXA1 (forkhead box A1) and TLX (T-cell leukaemia homeobox) have a less understood role in osteoclasts but that in itself could be an interesting research direction to explore in the future.

To sum up, within the first day of treatment with RANKL, Jun-AP1 and IRF4 are transcription factors that induce down-stream signalling pathways fundamental to osteoclastogenesis, but GSK-J4 treatment appears to prevent their effect.

4. Inhibition of Histone Lysine Demethylases Causes Changes in Histone Marks of Genes Important in Osteoclast Differentiation

Considering that GSK-J4 is targeting epigenetic modulators related to H3K4me3 and H3K27me3, the next step was to explore the changes in relevant histone marks. Specifically, the analysis was focused on the connection to the effects seen with GSK-J4 treatment in the transcriptomic data after 1 day of RANKL stimulation, because this was a critical time point where the transcriptional and epigenetic changes, leading to osteoclastogenesis, occur. Chromatin enriched for H3K4me3 and H3K27me3 marks was isolated using chromatin immunoprecipitation (ChIP), followed by high-throughput sequencing, on the same samples as the previously discussed transcriptomics dataset.

Firstly, it is important to point out the differences in the appearance of the two different histone marks. H3K4me3 marks are distinct narrow peaks that normally cluster around the TSS of genes. On the other hand, H3K27me3 marks are much broader and the signal arising from those samples always has more background signal, probably due to the technical limitations of the antibodies used for ChIP. As a result, the number of peaks is smaller, because it is common to allow multiple smaller peaks to be merged into broad H3K27me3-enriched genomic loci. Because of high noise levels there is a range of peak-calling algorithms that employ different strategies [192] and after comparing multiple outputs arising from the same dataset (for example, the output from the algorithm of choice, SoleSearchV2 is compared against the most popular peak caller MACS 2.0 in Figure 28A) it was decided to provide processed

peak data along with the raw peak coverage in order to confirm that the peaks are real and not false positives.

Secondly, cross-sample statistics were performed. Peak distribution around the TSS was used to confirm the validity of results (Figure 28A). H3K4me3 peaks cluster tightly around the TSS and most peaks are found within 1.5kb of it. H3K27me3 peaks are broader and they spread out up to 5kb to each side of the TSS. It is also of interest to note how different the output for the same dataset can be, if the bioinformatics pipeline is set up incorrectly or by not using the most optimal tools, as can be seen by comparing the output of MACS 2.0 and SoleSearch V2. Throughout the remainder of this thesis, only ChIP marks that are within those distances from TSS, ±2.5kb for H3K4me3 and ±5kb for H3K27me3, will be used for analysis. Additionally, only peaks with a peak score of more than 3 were used for differential chromatin binding analysis.

Because GSK-J4 is an inhibitor of H3K4me3 and H3K27me3 demethylases, one could expect an increase in number of peaks upon treatment. Indeed, the number of H3K4me3 marks increases with treatment within the same time-point, and the number of the same histone marks also increases by almost 4-fold between day 1 and day 7 of RANKL treatment (Figure 28B). On the other hand, the number of H3K27me3 remains fairly stable. This could be because of the way the peaks are called – by joining in smaller peaks into broad regions, so the signal might appear stronger and called with higher confidence.

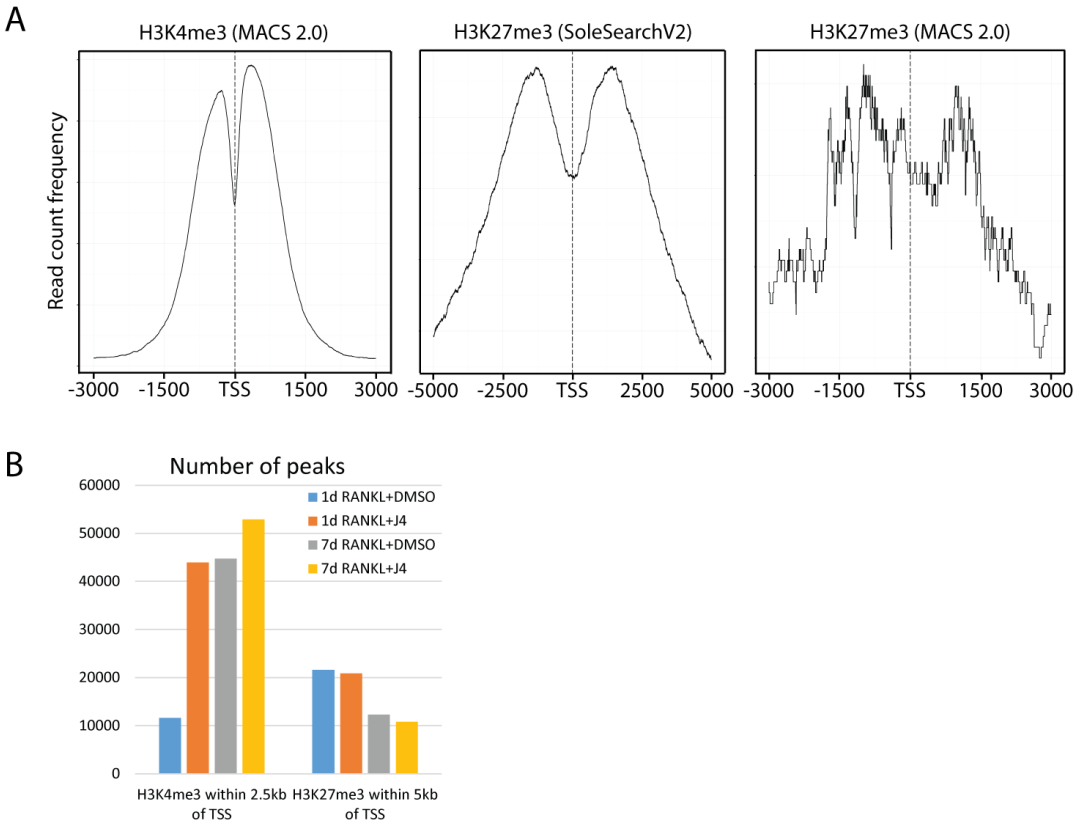


Figure 28 A. The distribution of H3K4me3 and H3K27me3 peaks around gene transcription start sites. B. Distribution of H3K4me3 and H3K27me3 peaks identified in RANKL treated samples.

Next, it was determined what genes have increased H3K4me3 marks with the stimulation of RANKL for 1 day as compared to MCSF macrophages. Of the 271 genes with increased H3K4me3 marks, 100 overlap with genes that were shown to have increased expression (Figure 29A). A GO analysis revealed that 19% of them fall under the term "regulation of transcription" (Figure 29B). Clearly, both, RNA data and the increase in activating histone mark H3K4me3, signify an upregulation of transcriptional activity as a result of induced osteoclastogenesis. It is also likely that H3K27me3 marks are present, which first need to be removed in order to allow gene expression. Because of that, it was expected that GSK-J4 will inhibit H3K27me3

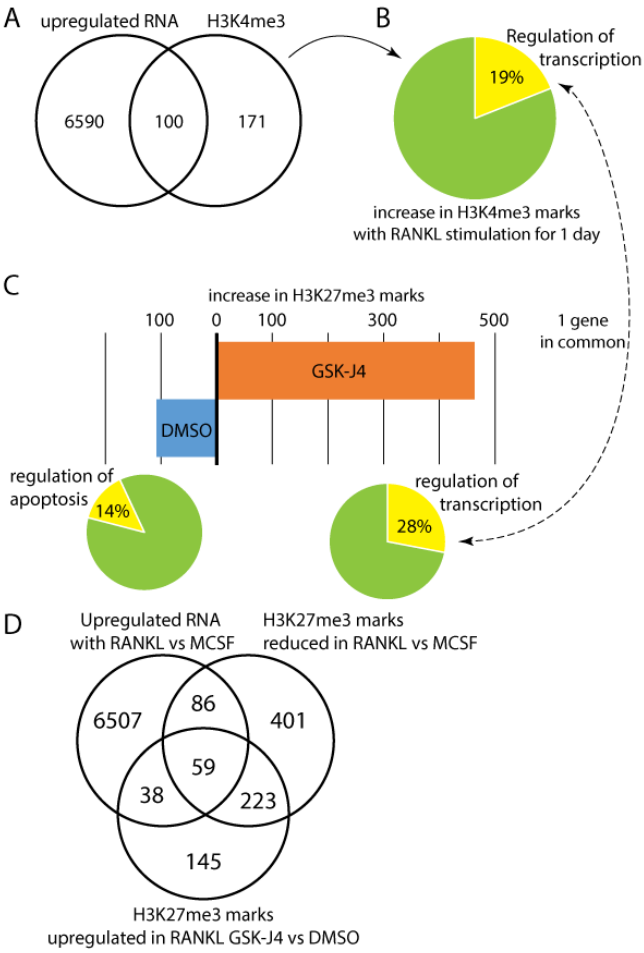


Figure 29 A. The number of genes that have increased expression and increased H3K4me3 marks upon RANKL stimulation. B. The fraction of genes that fall into the largest GO category identified among the subset of genes with increased H3K4me3 marks after 1 day of stimulation with RANKL. The arrow indicates that the same subset of genes is used in the GO analysis. C. The difference in the number of genes with increased H3K27me3 marks as compared between RANKL and GSK-J4 or DMSO treatment, and their main GO annotation terms. The arrow indicates there is only 1 gene in common between the two identical GO terms, so there are different genes inhibited by increased H3K27me3 abundance than those increased by H3K4me3 appearance upon RANKL treatment. D. The overlap of genes identified as having increased expression, reduced H3K27me3 marks with RANKL+DMSO treatment or retained H3K27me3 marks with RANKL+GSK-J4 treatment after 1 day.

demethylation and as a result, genes that would increase with RANKL stimulation will be silenced (Figure 30).

The comparison of H3K27me3 marks after 1 day of treatment with RANKL between GSK-J4 samples and DMSO yielded two subsets of genes (Figure 29C). H3K27me3 marks are differentially enriched around 108 genes in DMSO treatment as compared to GSK-J4, while there are over 4-fold more, 463 genes, enriched for the silencing H3K27me3 mark with GSK-J4 treatment as compared to DMSO. Interestingly, 28% of the latter genes fall under the same GO term of "regulation of transcription" as seen in upregulation of H3K4me3 marks with RANKL treatment (indicated by a dashed arrow, Figure 29C). An increase of H3K27me3 marks in the

DMSO sample (or, alternatively, a decrease of H3K27me3 marks in GSK-J4 treated cells) seems to affect a less defined set of genes, with the largest subset falling under the term of "regulation of apoptosis", which was also seen in the RNA data. Even though genes within the same GO annotation term have increased H3K4me3 marks (i.e. open chromatin leading to osteoclast differentiation) with RANKL treatment and increased H3K27me3 marks (i.e. silenced genes that if activated could lead to osteoclast differentiation) with GSK-J4 treatment, there is only 1 gene overlapping and the rest is different.

If both, H3K27me3 and H3K4me3 marks are present at the same time, the configuration is termed bivalent chromatin state and the gene is considered to be transcriptionally poised [193]. Osteoclastogenesis appears to be mostly induced by the up-regulation of gene expression, as witnessed by a number of transcription factors (AP-1, NFAT, NF-κB). As such, a particular focus was given to the H3K27me3 marks, their removal with RANKL stimulation to allow gene expression, and the prevention of this process by inhibiting H3K27me3 demethylases (Figure 30).

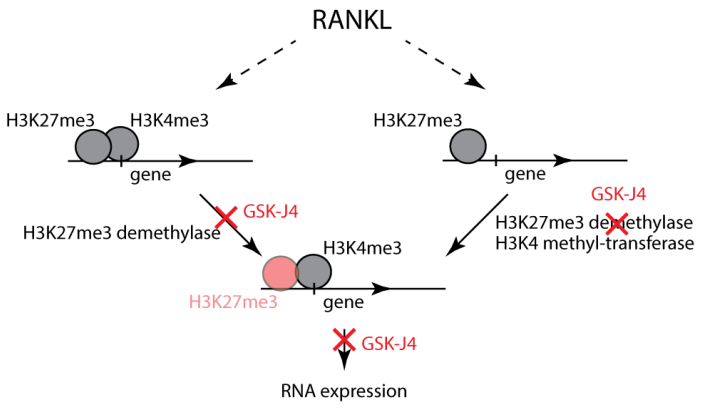


Figure 30 The mechanism of action of GSK-J4 treatment that prevents osteoclastogenesis. Osteoclastogenesis is induced by RANKL through signalling pathways (as discussed in text) acting on (dashed arrows) genes with two types of chromatin figuration, either bivalent (both H3K4me3 and H3K27me3) or only H3K27me3 marks. GSK-J4 treatment induces an increase in H3K27me3 marks (red sphere), leading to the reduction or silencing of transcription of essential osteoclast genes.

RANKL stimulation of MCSF macrophages significantly affects the H3K27me3 marks by yielding 767 genes with reduced or completely removed H3K27me3 marks from the TSS within the first day of osteoclastogenesis (Figure 29D). Interestingly, if treated with GSK-J4, 282 of those genes have their H3K27me3 marks significantly enriched, as compared to RANKL+DMSO. Furthermore, 59 of the latter subset have increased expression with RANKL stimulation (Figure 29D). Finally, out of those, 9 were also shown to be significantly down-regulated in GSK-J4 treatment (Table 2). In other words, first the genes were identified that are upregulated on the RNA level and had their H3K4me3 marks increased when MCSF-macrophages were treated with RANKL alone. Then, that set was compared to the genes that had their RNA down-regulated and H3K27me3 marks increased with GSK-J4 treatment of RANKL-stimulated cells. The set of genes that is upregulated with RANKL treatment, but down-regulated with GSK-J4, possibly holds the key to deciphering the mechanism of GSK-J4 action.

Gene ID	Distance to TSS	Peak Score	Gene Name
RPP38	-3775	5.02897	ribonuclease P/MRP 38kDa subunit
TASP1	-1211	3.82704	taspase, threonine aspartase, 1
MNX1	94	3.82704	motor neuron and pancreas homeobox 1
FLJ30679	3778	3.82704	uncharacterized protein FLJ30679
MNX1	4504	3.82704	motor neuron and pancreas homeobox 1
OSR2	113	3.81851	odd-skipped related 2 (Drosophila)
FAM171A1	1954	3.55252	family with sequence similarity 171, member A1
TOX2	2118	3.30066	TOX high mobility group box family member 2
HOMER1	3860	3.30066	homer homolog 1 (Drosophila)
BATF3	2928	3.2686	basic leucine zipper transcription factor, ATF-like 3

Table 2 A list of genes that have H3K27me3 marks removed and their expression upregulated after 1 day of stimulation with RANKL, but have retained H3K27me3 marks and differentially reduced expression in GSK-J4 treated RANKL-stimulated cells.

The majority of genes within this list (Table 2) are transcription factors or related to the regulation of gene expression. For example, TASP1 (taspase, threonine aspartase, 1) is a protease vital for the activity of KMT2A [194] which in turn regulates HOXC8 (homeobox C8) which has links to osteoclast differentiation [195]. OSR2 (Odd-Skipped Related 2) is an ill-defined transcription factor with a known role in osteoblast function [196], but it may also have some undiscovered function in osteoclastogenesis. Finally, BATF3 (Basic Leucine Zipper Transcription Factor, ATF-Like 3) belongs to the AP-1 family of transcription factors and dimerises with Jun [197], which was already shown before to be strongly reduced with GSK-J4 treatment.

On the other hand, the changes in H3K27me3 peak enrichments for this narrow list of genes are very low. There is a number of other genes that fall into the subset of the 59 genes (Supplementary Table 11), which are upregulated in osteoclastogenesis as indicated by reduced H3K27me3 marks and increased expression, but which are inhibited with GSK-J4 treatment, possibly by inhibiting the removal of H3K27me3 marks. Some of the top genes with the most up-regulated H3K27me3 marks are LHX4 (LIM (Lin11, Isl-1 & Mec-3) Homeobox 4), RORA (Retinoic Acid Receptor-Related Orphan Receptor A), STMN2 (Stathmin 2), STK3 (Serine/Threonine Kinase 3) and NRG1 (Neuregulin 1) which have only few known links with osteoclast differentiation but they are mostly transcription factors, or serine/threonine kinases.

On day 7 of RANKL treatment there are other genes that are important for mature osteoclast function. CTSK, LIF (Leukemia Inhibitory Factor) and ILF3 (Interleukin Enhancer Binding Factor 3), among many others, were shown to have their

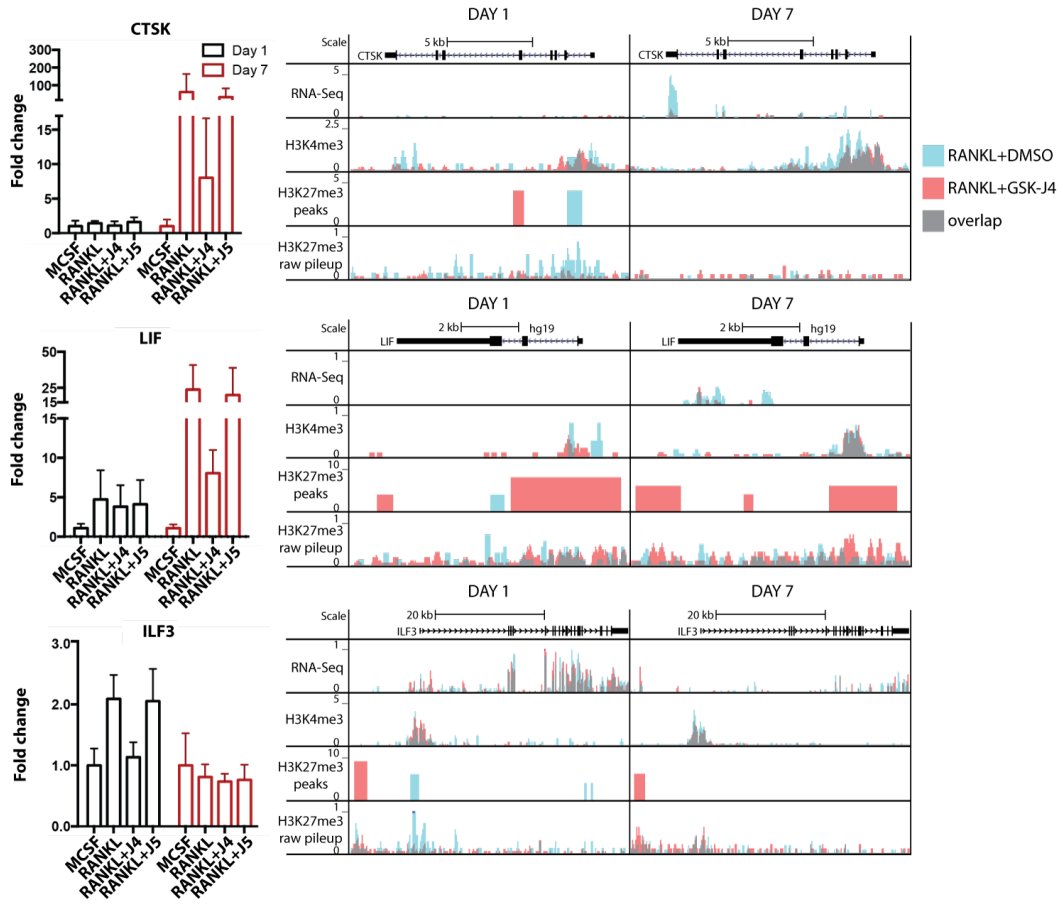


Figure 31 Expression of CTSK, LIF and ILF3 confirmed by qPCR at the same time-points as transcriptomics and ChIP data. The increase in H3K4me3 marks is shown to be related to the increase in transcription, increase of H3K27me3 alone or along with H3K27me3 marks is linked to a silenced genes state. H3K27me3 peaks represent the peaks identified by SoleSearch V2 while the H3K27me3 raw pileup tracks represent the unprocessed signal arising from piling up individual reads mapped onto the genome.

expression and histone marks affected by GSK-J4 at this later time point in such a way that would prevent osteoclast function (Figure 31). The expression was previously confirmed by Dr Ka Hing Che using qPCR.

5. Discussion

Data showing the inhibition of RANKL-induced osteoclastogenesis from MCSF-macrophages by treating with the histone demethylase inhibitor GSK-J4 was briefly

reviewed in this chapter. RNA and chromatin samples enriched for H3K4me3 and H3k27me3 histone marks were sequenced thus allowing differences to be observed between various treatments on a genome-wide scale.

Firstly, RANKL-treated cells were confirmed to be osteoclast-like by observing that a number of known osteoclast markers increase upon stimulation. Due to the high-throughput nature of this study, it was also possible to identify new biomarkers that were confirmed by qPCR and are now employed by our research group for various screens in osteoclasts.

Secondly, a thorough analysis of transcriptome changes between the samples treated with or without RANKL and with GSK-J4 or DMSO allowed the deciphering of an intricate inter-play of factors involved in osteoclastogenesis. While NFATc1 expectedly appears as the main regulator within the first day of RANKL treatment leading to osteoclast differentiation, it was shown that GSK-J4 specifically affects the Jun-AP1 signalling pathway, among others, in early osteoclastogenesis, that prevents osteoclast function and morphology. Despite the observation that GSK-J4-treated cells exhibit a strikingly similar global expression profile to osteoclasts, when compared to DMSO control, the comparably few changes observed hold the key for the inhibition of osteoclast differentiation.

The study of GSK-J4 effects were further complemented by H3K4me3 and H3K27me3 chromatin immunoprecipitation data, with the main focus laying on the epigenetic changes observed within the first day, when the largest transcriptional effects take place leading to induction of osteoclasts. Indeed, H3K27me3 marks,

retained or increased with the treatment of GSK-J4, closely overlap with the genes that are prevented from increasing upon RANKL stimulation. A large fraction of those genes are related to regulation of transcription (BATF3, TASP1) as well as other important genes to osteoclast function, such as CAV2 (caveolin 2) or ATP1A3 (ATPase, Na^+/K^+ transporting, alpha 3 polypeptide).

To sum up, it is difficult to point out a single gene or two, which have their H3K27me3 marks increased and which would lead to the blockage of the entire pathway leading to osteoclastogenesis. Instead, it appears as if smaller effects of inhibition of gene expression through the inhibition of H3K27me3 demethylation add up and lead to the lack of a transcriptional stimulus to start osteoclast differentiation. Such mechanism also sounds more plausible, considering the large number of genes that were observed to be up-regulated with RANKL. This global upregulation of expression is reduced due to retained H3K27me3 marks upon GSK-J4 treatment, and as a result osteoclastogenesis is inhibited.

Furthermore, this study also revealed a significant role of microRNAs in osteoclast differentiation, especially at a later time-point of 7 day treatment with RANKL. Unfortunately, the study of microRNA still remains a complicated procedure due to missing annotations and unknown functions. However, this data could facilitate a larger effort to discover the targets of the identified microRNAs *in silico* and subsequently validating them by probe-based qPCR (TaqMan assays).

It must be noted that particular challenges arise already within the data processing stages of sequencing data. There are an increasing number of algorithms and

bioinformatics pipelines that employ different statistical principles and ideas that could potentially yield slightly different results. Future versions of bioinformatics tools may allow re-analysis of the raw data and possibly identify new hits and targets.

The richness of the datasets due to the large number of samples under different conditions and time-points could also give rise to a number of future projects to discover and validate new targets within their own signalling pathways, instead of looking at a genome-wide picture. The number of ways in which the data can be cross-correlated is also astoundingly large and could still be hiding interesting discoveries if novel approaches are to be applied in the future.

Chapter V

Epigenetic Regulation of Multiple Myeloma

1. Introduction

Multiple myeloma (MM) is a cancer of bone marrow derived plasma cells. White blood cells are regenerated in the bone marrow but once they become malignant, the sudden uncontrolled propagation of the plasma cells causes multiple lesions in the bone, giving rise to the name multiple myeloma. Multiple myeloma is an incurable disease and constitutes 2% of all cancers in the United States [198]. New treatments have been recently introduced that are able to extend the survival rate up to 7-8 years after diagnosis [198].

The understanding of the precise mechanism of developing MM is incomplete. It is known that the progenitor cells for MM are terminally differentiated plasma cells that express IgM (Immunoglobulin M) and IgG; they reside in the bone marrow, do not proliferate, and are sustained by IL-6 serving as a survival signal produced by stromal cells [199]. These plasma cells are responsible for the first antibody-based response to antigens and they arise from B cells that have undergone isotype switch recombination in a germinal centre, induced by antigen-presenting T-cells. The first step in the direction of MM is development of MGUS (monoclonal gammopathy of undetermined significance) which manifests itself in a stable increase of monoclonal Ig in serum [200]. Interestingly, chromosomal abnormalities are already present in

50% of MGUS cases but only a small minority of those, around 1% develop into MM, although the number increases significantly with age [200].

There are three main transcription factors vital for plasma cell differentiation [201]. PRDM1 (Positive-Regulatory Domain Containing 1, also known as BLIMP1 or B-Lymphocyte-Induced Maturation Protein 1) was shown to be of vital importance for plasma cell differentiation in mice [202]. It inhibits cell proliferation by repressing MYC (Avian Myelocytomatosis Viral Oncogene Homolog, also known as c-Myc) and E2F1 (E2 family transcription factor 1), among many others [203]. Genes involved in the production and secretion of IgG are up-regulated via XBP1 (X-Box Binding Protein 1), DDIT3 (DNA-Damage-Inducible Transcript 3, also known as CHOP or C/EBP-Homologous Protein) and HSP70 (heat shock protein 70) [204].

The role of transcription factor XBP1 is particularly important, even though it is signalling down-stream of PRDM1. Because of the induction of production of large amounts of IgG, stress thresholds are reached in the endoplasmic reticulum which in turn induces the unfolded protein response (UPR). The latter induces ERN1 (Endoplasmic Reticulum to Nucleus Signalling 1, also known as IRE1 or Inositol-Requiring Enzyme 1) which is capable of producing an alternative splice form of XBP1 [205]. This new splice form is fundamental in regulating protein folding and maturation [206].

Finally, IRF4 was also recently discovered to be fundamental for plasma cell differentiation and function by using a conditional IRF4 double-knockout construct in mice [207]. In this study it was noted that IRF4 is acting independently and

upstream of XBP1, however the exact function remains unclear. The authors speculate that IRF4 may provide a pro-survival function, as seen in other myeloid and lymphoid cells [208].

Multiple myeloma frequently develops due to various chromosomal abnormalities and translocations. The misregulation of IRF4, MYC, CCND1 (Cyclin D1), MMSET (Multiple Myeloma SET Domain Containing Protein), FGFR3 (fibroblast growth factor receptor 3), among many others, through translocations is thought to be important for oncogenesis [200]. Furthermore, signalling by cytokines IL-6, TNFα and VEGF (vascular endothelial growth factor) are thought to contribute to MM proliferation and migration [209]. The expression of plasma cell transcription factors PRDM1 and IRF4 remained largely unchanged and characteristic of plasma cells while expression of XBP1 can vary [210].

Multiple array expression experiments have identified several down-stream pathways, such as Notch and Wnt (wingless type), that are utilised in plasma cell function but are highly misregulated in MM cells [201]. Specifically FRZB (Frizzled-Related Protein) and DKK1 (Dickkopf WNT Signaling Pathway Inhibitor 1) were identified as potential therapeutic targets for MM [211, 212] as they were differentially overexpressed in MM patients.

Furthermore, epigenetic misregulation in MM cells has also been implicated in the progression of disease. There is a substantial amount of literature available, and even a number of inhibitors in clinical trials, targeting, for example, HDACs in relation to progression of MM [213]. However, histone lysine methyl marks may have even more

profound effects in MM. MMSET is directly implicated in the methylation of H3K36 marks, but it was also shown to indirectly affect H3K4, H3K27 and H4K20 histone marks [214]. Interestingly, a transcript variant starting at exon 15 of MMSET gene, RE-IIBP (response element II binding protein), was shown to be a HMT specific for the methylation of H3K27 residues and it is largely upregulated in MM cells [215]. Additionally, the H3K27me3 demethylase KDM6A was shown to have a loss of function mutation in up to 10% of patients [216], which would add to the build-up of H3K27me3 marks in MM cells. As such, a build-up of H3K27me3 is largely expected in cells and thus the cells may be highly sensitive to the levels of the H3K4me3 mark that can lead to bivalent chromatin states or enable activation of gene expression.

In light of existing literature on genetic regulation in MM and the substantial involvement of epigenetics, GSK-J4 emerged out of a focused screen of epigenetic inhibitors targeting myeloma proliferation [165]. Unpublished results from our research group indeed showed apoptosis levels of a MM cell line JJN3 of up to 85% within 48 hours of treatment. Furthermore, as shown in the previous chapter, GSK-J4 is a potent inhibitor of osteoclast differentiation and bone loss, which is strongly associated with MM and hence GSK-J4 would have a double beneficial effect on eliminating the cancer as well as preventing cancer-induced bone loss.

It was therefore of interest to investigate further global epigenomic and transcriptional effects in multiple myeloma. These results with GSK-J4 treatment of the multiple myeloma cell line JJN3 under various conditions will be discussed in this chapter. Total RNA and chromatin enriched for H3K4me3 and H3K27me3 marks was isolated and sequenced. A number of other supporting experiments were carried

out by other members of the research group with the aim to validate these high-throughput genomic and transcriptomic screens.

2. Transcriptome analysis reveals an upregulation of a number of pro-apoptotic and stress-related genes with GSK-J4 treatment

Firstly the aim was to identify transcriptional changes in the MM cell line that occur with GSK-J4 treatment as compared to the DMSO negative control. JJN3 cells were incubated for 2, 6 or 24 hours with DMSO or GSK-J4 (4 μM). Briefly, total RNA was isolated by depleting ribosomal RNA and after preparing libraries, they were sequenced on an Illumina NextSeq 500 high-throughput sequencer. The resulting reads were mapped, quantified and annotated using the *bowtie* and *cufflinks* pipelines, as explained in more detail in the Materials and Methods section.

Differential expression was assessed by normalising GSK-J4 treatment to the DMSO samples. Samples that had a fold change of more than 1.5-fold were used in the analysis and compared via a heatmap (Figure 32). In total there were 4152 genes with differential expression matching the chosen criteria. The clustering algorithm indicates the 2h samples as having the largest dispersion compared to the nearest branch of later time-points. This indicates there are significant changes happening already within only 2 hours of GSK-J4 treatment. The vast majority of the genes have fold changes of up to 6-7 up- or down-regulated.

Next, 6 categories of changes were identified, as denoted by the Greek letters (Figure 32). The sets of genes within clusters α, β and γ represent different patterns of up-

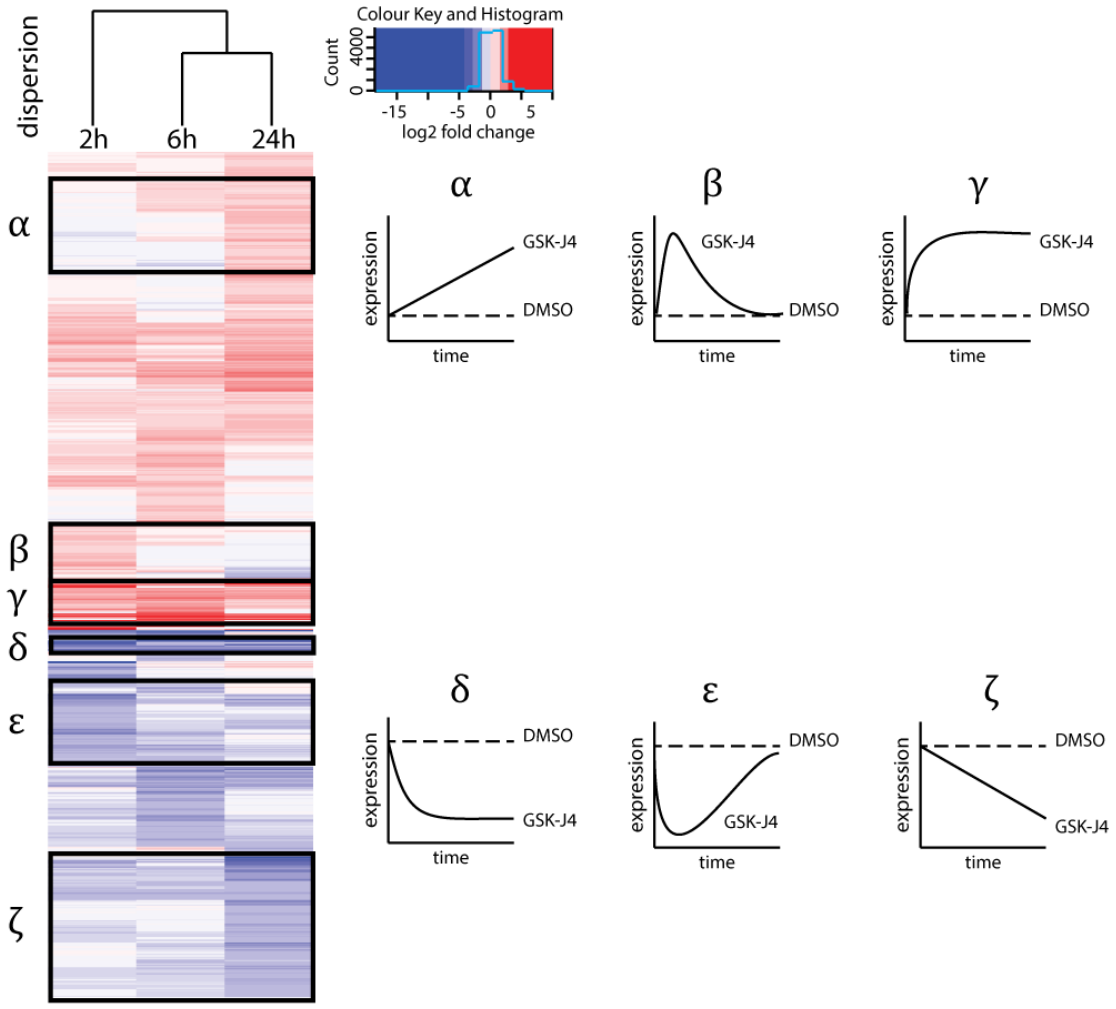


Figure 32 Heatmap showing the fold change in gene expression at different time-points after treatment with GSK-J4 as compared to DMSO. The dispersion shows how far apart the samples are based on the differences in genes expression. The colour scale is based on a logarithmic distribution, enabling better resolution for the samples with most abundant expression values while binning all outliers into one colour. Gene expression is denoted in log2 of fold-change. Charts α to ζ represent schematic trends that were identified within the heatmap.

regulation when treated with GSK-J4. Those can be briefly defined as a gradual increase with time (cluster α), a sudden increase at the very early time-points and then slow reduction to basal levels (cluster β) and rapid increase in expression and sustained expression at those high levels (cluster γ). The same trends across the time-scale but in the opposite direction, i.e. reduction in expression, were also defined as δ, ε and ζ. Each of these clusters was analysed using DAVID gene ontology (GO) [144] and Homer motif discovery [217].

There are 281 genes that show a gradual increase in expression (Figure 32α). A GO analysis reveals that 9.8% of them are involved in intracellular signalling (Supplementary Table 18) while another 7.5% are involved in regulating cell proliferation (Supplementary Table 19). 3.1% of genes have a TP53 (tumour protein 53, also known as p53) binding site upstream of their TSS. TP53 is one of the most studied tumour suppressor genes that is frequently mutated in cancers and so the cells can escape its anti-proliferative effects [218]. Considering that genes, normally under control of TP53, are expressed in MM upon GSK-J4 treatment, shows the cells are undergoing transcriptional changes that will lead to eventual apoptosis.

The subset of genes that are up-regulated at an early time-point and then reduce to basal levels contains 232 members (Figure 32β). 7.8% of them, such as BAG3 (BCL2-associated athanogene 3), CRADD (CASP2 and RIPK1 domain containing adaptor with death domain) or TNFRDF4 (tumor necrosis factor receptor superfamily, member 4), are implicated in the regulation of apoptosis (Supplementary Table 20). In fact, this subset of genes may initiate pathways leading to apoptosis, such as through TNFα signalling, which is further executed by the set of genes induced later, as seen in group α. Motif analysis indicates an enrichment of binding sites for E2F within this subset, which was already previously discussed to play a role in plasma cell differentiation by inhibiting proliferation.

Finally, the cluster of genes that are up-regulated early on with GSK-J4 treatment and then remain constitutively expressed at those high levels (Figure 32γ) contain 3.9% of genes with an up-stream MTF1 (metal-regulatory transcription factor 1) binding site (Supplementary Table 21). Curiously, this transcription factor primarily

regulates the intracellular concentration of metal ions and is normally up-regulated in response to increased levels of divalent metal ions as well as oxidative stress [219]. Indeed, nearly all metallothioneins are highly up-regulated already at 2h and retain elevated levels with the expression peaking at 6h. This is a strong indication that GSK-J4 effects may manifest themselves by causing misregulation of intracellular metal ion concentrations or by causing oxidative stress.

Furthermore, the same cluster contains a set of genes involved in amino acid synthesis and conversion. Those genes are namely NAGS (N-acetylglutamate synthase), ASS1 (argininosuccinate synthase 1), GSTO1 (glutathione S-transferase omega 1), LIAS (lipoic acid synthetase) and MTHFD2L (methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like). NAGS converts glutamate into N-acetylglutamate which is also the first step of the urea cycle, and ASS1 is fundamental in the biosynthetic pathway of arginine. Upregulation of a number of ATF4-inducible genes within the same cluster, such as DDIT3 and CHAC1 (ChaC, cation transport regulator homolog 1) further hints at the possible induction of oxidative stress upon treatment with GSK-J4, however it may not be a direct causal effect. This data led to an in-depth metabolomic assay discussed later, which revealed that there is a misbalance of amino acids and there is a build-up of intermediate compounds of the citric acid cycle. Increased oxidative stress, as indicated by the up-regulation ATF4-inducible genes, is another potential reason for the ultimate outcome, albeit indirect, of GSK-J4 treatment.

On the other hand, the genes that are reduced do not exhibit the same strong functional relevance as seen in each subset of the up-regulated genes. The most

notable subset ζ contains 7.6% of genes involved in the GO term oxidation reduction (Supplementary Table 22). This subset contains a number of NADH dehydrogenase subunits, cytochrome genes, as well as FASN (fatty acid synthase), PDHB (pyruvate dehydrogenase beta), LDHA (lactate dehydrogenase A), and GPX1 (glutathione peroxidase 1) among many others. This indicates a reduced mitochondrial function which may further link to the notion of possible GSK-J4-induced oxidative stress.

The same ζ subset is enriched for the MYC binding site upstream of 17% of genes (Supplementary Table 23). MYC itself is gradually reduced over time, and so are its target genes, among which of particular interest are LDHA (lactate dehydrogenase A), PDHB (pyruvate dehydrogenase (lipoamide) beta), ATPAF2 (ATP synthase mitochondrial F1 complex assembly factor 2) and NDUFS8 (NADH dehydrogenase (ubiquinone) Fe-S protein 8). MYC was previously shown to be fundamental in MM progression and an siRNA against MYC killed the cells [220]. Hence, this finding further adds to the growing list of genes, contributing to GSK-J4-induced apoptosis of MM.

However, as previously discussed with osteoclasts, transcriptomics only gives an indication of what processes GSK-J4 might affect downstream of its targets. Assuming inhibition of chromatin demethylation, the mechanism of GSK-J4 on MM must be lying within the epigenetic processes that GSK-J4 is controlling.

3. Global changes in chromatin marks observed within 24 hours of treatment with GSK-J4

Chromatin immunoprecipitation (ChIP) was carried out on MM samples taken from the same conditions as was discussed in the previous section. Chromatin was enriched for H3K4me3 and H3K27me3 marks, sequenced and mapped using bioinformatic algorithms outlined in the methods and discussed in the Osteoclast chapter (Chapter IV, p. 99).

Firstly, the total number of identified peaks varied only very little between DMSO and GSK-J4 treatment, with a minor non-significant increase in GSK-J4 treatment for H3K4me3 marks. The number of peaks was very similar in H3K4me3 and H3K27me3 samples and varied between 12,801 and 24,296 with the mean around 20,000 peaks. Interestingly, though, the number of differentially enriched peaks in the H3K4me3 samples was much higher than in the H3K27me3 (Figure 33), which might suggest that GSK-J4 mostly affects the KDM5B and hence the differential number of H3K4me3 rises sharply. It was also mentioned before that MM overexpresses a transcript variant of MMSET, RE-IIBP, which induces H3K27me3 marks [215] and in up to 10% of cases of MM have an impaired KDM6A [216]. This could lead to the speculation that MM samples naturally contain a large number of H3K27me3 peaks and inhibition of demethylases specific for this mark do not have a major effect. Conversely, the full-size transcript MMSET was shown to indirectly reduce global H3K27me3 levels but it induces hypermethylation of H3K27me3 marks at specific loci by recruiting HMT EZH2 [221].

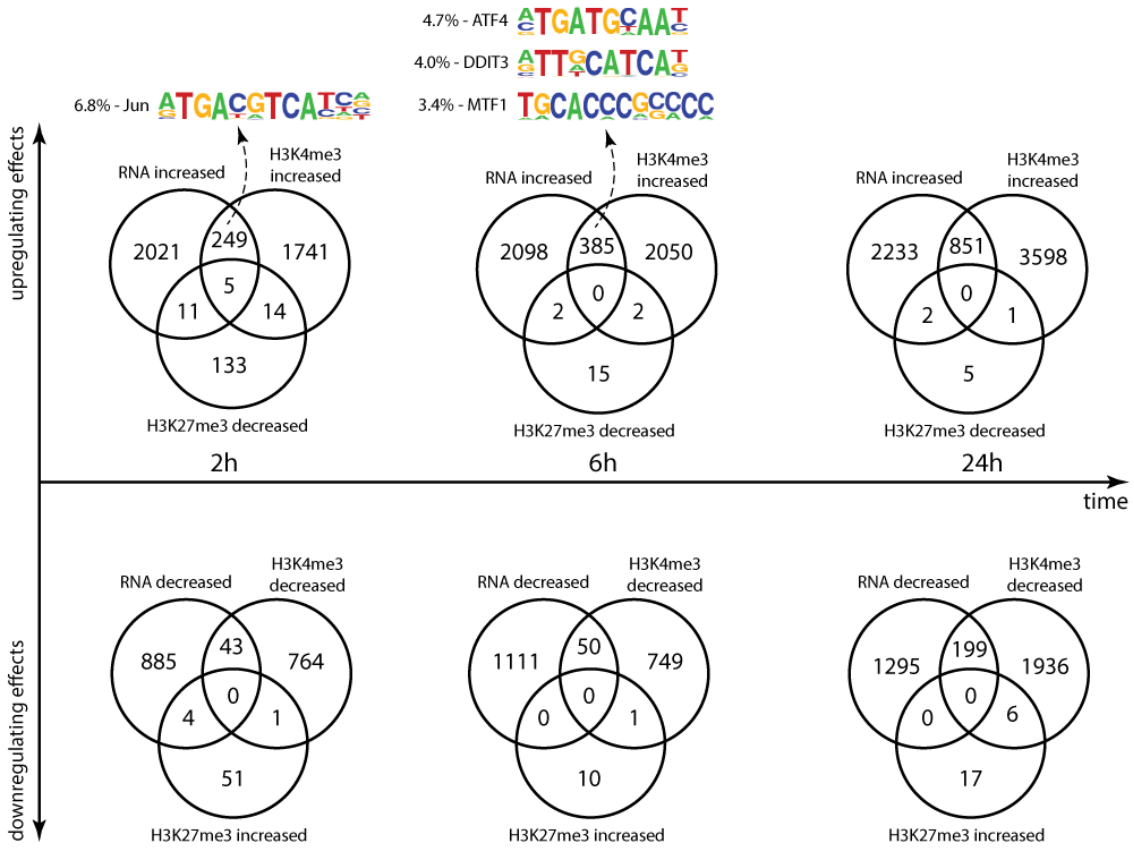


Figure 33 Comparison of differentially expressed genes or genes with differentially enriched histone marks across different time-points. Time-points and samples clustered into either down-regulated gene expression denoted by reduced RNA levels, decreased H3K4me3 marks and increased H3K27me3 marks, or up-regulated effects by increased H3K4me3 marks, decreased H3K27me3 marks and validated by increased levels of RNA. The Venn diagrams represent the number of genes that are identified within each treatment. Only top relevant motifs are displayed which were identified in the subset of genes overlapping from different treatments, as indicated by arrows.

The histone marks and RNA expression were compared in combinations of either activating or repressive conditions, i.e. activating assumes RNA expression is increased, H3K4me3 marks are increased while H3K27me3 marks are reduced in GSK-J4 samples as compared to DMSO (Figure 33). Such a comparison was carried out for each time-point and those genes that were common for multiple conditions were further inspected. The number of differentially enriched H3K4me3 marks increased over time and nearly tripled between 2h and 24h, while the number of differentially enriched H3K27me3 marks remained very low.

4. Identification of transcription factor binding motifs reveals distinct patterns of stress responses

Among the genes up-regulated at 2h and overlapping with increased H3K4me3 marks there are 6.8% with a c-Jun binding site up-stream of the TSS (Supplementary Table 24). In support of the observed anti-proliferative phenotype in JJN3 treated with GSK-J4, it was previously noted that c-Jun has a strong apoptotic effect on MM if up-regulated [222]. EGR1 (early growth response 1), down-stream of c-Jun, is the important gene that if upregulated causes apoptosis in MM, as noted by the authors. EGR1 is also upregulated 2.7-fold already at 2h and it follows a steady increase with time and is 5.8-folds higher in the GSK-J4 sample after 24h of treatment as compared to DMSO.

Of particular interest are the genes that have all three factors changed – i.e. RNA and H3K4me3 marks are up while H3K27me3 marks are reduced (Table 3). GRIN2D (glutamate receptor, ionotropic, N-methyl D-aspartate 2D) is known to play an important role in neuronal signalling, however it is interesting to note that it is also sensing glutamate [223], which maybe be implicated in the previously seen signs of misbalance of amino acid synthesis and conversion. UBE2QL1 (ubiquitin-

Gene	Gene Name
GRIN2D	glutamate receptor, ionotropic, N-methyl D-aspartate 2D
UBE2QL1	ubiquitin-conjugating enzyme E2Q family-like 1
ANKRD19P	ankyrin repeat domain 19, pseudogene
N4BP3	NEDD4 binding protein 3
TTYH1	tweety homolog 1 (Drosophila)

Table 3 Genes that were upregulated as determined by increased RNA levels and which had increased H3K4me3 enrichment and reduced H3K27me3 marks at 2h time-point.

conjugating enzyme E2Q family-like 1) is related to proteasomal degradation and this may link into the amino acid biosynthesis and protein folding mechanisms that are particularly active in MM. Finally, TTYH1 (tweety homolog 1) is a chloride ion channel which is said to regulate cell size through osmotic pressures. Considering there may be a misregulation of cations in the cell, as suspected based on the MTF1 upregulation of metallothioneins, the upregulation of an anion channel (Figure 34) may be required to balance the cross-membrane charge.

Removal of H3K27me3 alone may contribute to an increased expression of a number of other genes relevant to the eventual apoptosis of MM cells (Table 4). The previously mentioned metallothionein genes, caspase-related CARD8 (caspase recruitment domain family, member 8) and PRDM11 (PR domain containing 11), closely related to the main factor inducing plasma cell differentiation, all further support previously aired speculations about the possible mechanism.

Genes upregulated at 6h on the RNA level and with H3K4me3 marks differentially increased, contain a number of interesting transcription factor binding sites upstream of their TSS. ATF4 (activating transcription factor 4) is a transcription factor induced by oxidative stress, endoplasmic reticulum stress, amino acid

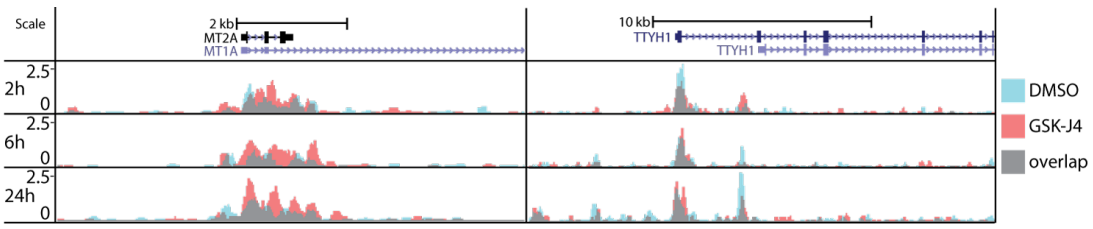


Figure 34 ChIP tracks of metallothionein MT2A and TTYH1 loci showing the gradual increase of H3K4me3 around their transcription start sites, correlating with increased expression, upon GSK-J4 treatment.

deprivation and a few other mechanisms [224], so this further reinforces the notion of an ongoing stress related to amino acid biosynthesis and oxidative stress (Supplementary Table 25). The ATF4 gene itself is also mildly upregulated from 6h at 1.7-fold as compared to DMSO. Additional validations carried out by Dr Edward Hookway using qPCR and by Federica Lari using Western blotting confirm ATF4 up-regulation at the latter time-points. DDIT3 (also known as CHOP) was already mentioned to be implicated in plasma cell differentiation, is down-stream of ATF4 signalling, and further enhances the stress related response. Finally, genes containing an up-stream MTF1 binding site, already previously seen in RNA samples alone, are further supported by an increase in H3K4me3 marks. All three transcription factors and upregulation of their down-stream signalling genes hint at an increasingly stressful intracellular environment occurring within 6h of treating with GSK-J4.

The remaining samples did not contain any relevant motif enrichments within the sets of overlapping genes identified via different treatments. In the 24h upregulated condition, a lot of the genes previously seen upregulated in the 6h sample, remain at increased levels and maintain differentially enriched H3K4me3 marks. 17.5% of the genes in the cluster are linked to the GO term "regulation of transcription"

Gene	Gene Name
SEMA3F	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F
CARD8	caspase recruitment domain family, member 8
VWC2	von Willebrand factor C domain containing 2
PRODH2	proline dehydrogenase (oxidase) 2
PRDM11	PR domain containing 11
MT1M	metallothionein 1M
SNX29P2	sorting nexin 29 pseudogene 2
RDH13	retinol dehydrogenase 13 (all-trans/9-cis)
RPP25	ribonuclease P/MRP 25kDa subunit
KLHDC9	kelch domain containing 9
DGKZ	diacylglycerol kinase, zeta

Table 4 Genes that had increased RNA levels and reduced H3K27me3 marks at 2h time-point with GSK-J4 treatment.

(Supplementary Table 26) and genes implicated in apoptosis are particularly abundant, such as CDK7 (cyclin-dependent kinase 7), CEBPB (CCAAT/enhancer binding protein beta) and EAF2 (ELL associated factor 2) [225-227].

Ultimately cells undergo apoptosis and that is apparent from the gene set upregulated at 24h on the RNA level as well as with increase H3K4me3 marks (Supplementary Table 27). JAK2 (Janus kinase 2) tyrosine kinase was shown to be an activator of oxidative stress-induced apoptosis [228]. Others, such as DAPK2 (death-associated protein kinase 2) are also upregulated and they are part of a gene signature of a dying cell [229].

5. Discussion

Global analysis of transcriptional and epigenomic effects suggests that the main effects of GSK-J4 on MM JJN3 cells might arise due to an increase of H3K4me3 marks, and not H3K27me3, which is the opposite of what was observed in osteoclast samples. On the other hand, changes in transcription are accompanied by changes in H3K4me3 marks, but are not necessarily the reason for transcriptional changes. A precise mechanism of how GSK-J4 may induce apoptosis in MM still remains unclear and may only provide hints at what pathways are involved. In fact, it may be necessary to analyse samples in a similar fashion as discussed in this chapter, only at much earlier time-points. This is further supported by the fact that knockdown or selective chemical inhibition of KDM5 enzymes are not sufficient to induce the complex phenotype observed with GSK-J4, rather a combined inhibition or knockdown of

KDM5 and KDM6 enzymes creates the same pattern of effects (apoptosis, ATF4 induction, metallothionein upregulation), as shown in unpublished work carried out by Dr Edward Hookway and Federica Lari.

So far it has been shown that within the earliest 2h time-point a set of genes regulated by the transcription factor c-Jun are highly increased. RNA levels and H3K4me3 marks around the TSS of JUN are also differentially increased at 2 and 6 hours. It's main down-stream effector EGR1, previously shown to lead to apoptosis of MM cells [222], is also gradually upregulated with the increased treatment time with GSK-J4. Although, while H3K4me3 marks are not differentially enriched, they are present in all treatments and time-points. An increased expression of c-Jun in a transcriptionally favourable chromatin conformation is sufficient to support the observed increase in EGR1.

Another significant effect observed is the dramatic early up-regulation of metallothionein genes which respond to metal ion influx into the cell. Finally, at the same time point, signalling pathways involved in oxidative stress, unfolded protein response and amino acid deprivation manifest themselves through the upregulation of ATF4 and DDIT3 and a large number of their down-stream effector genes. Combined, this offers an indication of what causes the cells to eventually undergo apoptosis. However, a causal mechanism of how these effects are induced is still lacking, and investigations into even earlier time-points may provide the answer.

Interestingly, recent unpublished results by Dr Edward Hookway from our research group have provided possible answers to the metallothionein induction. Using

fluorescent probes and time lapse microscopy it was shown that with GSK-J4 treatment cells take up large amounts of zinc ions within the first 30 minutes, and then the fluorescence signal returns to basal levels within the measured time-frame of 160 minutes. Such a sudden influx of zinc could induce the metallothionein genes which we observe in RNA-Seq and ChIP-Seq data analysis.

Furthermore, zinc could impair mitochondrial function by inhibiting glutamate and/or α-ketoglutarate dehydrogenase enzymes [230]. This could indeed explain the metabolic phenotype which manifests itself in amino acid deprivation and induction of oxidative stress in JJN3 cells. However, this is a hypothesis that still needs experimental proof.

Additionally, elaborate metabolomic experiments were carried out by Dr Edward Hookway in collaboration with Professor Ulrich Günther, which showed that there is a large accumulation of glutamine but a significant decrease in asparagine and aspartic acid when JJN3 cells are treated with GSK-J4 indicating a blockade of the tricarboxylic acid cycle. This observation could also explain the stress response indicated by ATF4 and DDIT3. Further knockdown experiments revealed that the EIF2α/ATF4 pathway, as well as the amino acid stress response kinase GCN2 are essential to mediate the GSK-J4 stress response and apoptosis in JJN3 cells. These findings, as well as a substantial number of genes that were identified to be up- or down-regulated in the high-throughput sequencing experiments, were further confirmed by Western blotting, carried out by Federica Lari, and by qPCR, carried out by Dr Edward Hookway.

This chapter once more provides an example of how high-throughput sequencing can offer a large-scale insight and provide testable hypotheses into possible molecular mechanisms which can then be individually investigated with more accurate techniques.

Chapter VI

Discussion and Future Perspectives

A diverse set of experiments was discussed in this thesis, all of which contributed to a number of different projects. The main focus across all these experiments carried out in a range of cell types of the myeloid cell lineage was to understand the epigenetic regulation of observed phenotypic responses. The findings were further supported and validated by carrying out additional microscopy experiments, quantitative PCR and other assays. While not explicitly shown, a large proportion of these techniques were employed in our research group and even the whole department for the first time, so a great effort was made to optimise and validate the methods.

Firstly, the study of macrophages and their biomarkers played a fundamental role in this research, because it laid the basis for the understanding of factors involved in the immune system, as revealed by the validation of putative biomarkers and re-analysis of existing datasets related to macrophage differentiation. It also highlighted donor-to-donor variation and the implications it has on experimental designs. Furthermore, the growing understanding of macrophage polarisation into a spectrum of cell subtypes is helping to better define possible targets for treatments in disease and inflammatory conditions.

Originally, the aim was to use the lentiviral shRNA delivery system to identify vital elements of the epigenetic machinery in macrophage differentiation and response to

pro-inflammatory stimulation. This aim remains to be an important future direction for the study of macrophages in our research group, with a significantly expanded picture which includes a diverse spectrum of monocyte to macrophage differentiation pathways. While screening of monocytes with established libraries of lentiviral constructs or siRNAs remains elusive due to technical limitations [231], there has been progress made using adenoviral vectors [232] and LNAs, as shown in this thesis. Unfortunately, only a limited number of libraries exist for adenoviral shRNA gene knock-down [233], and they are largely incomplete [234]. LNAs have so far proven to be difficult to obtain due to the required time to design LNAs, limited production capacity from our collaborators and the need to test multiple LNAs until an efficient design for each target is found. On the other hand, increasingly automated and high-throughput laboratory processes will enable the creation of such comprehensive gene knock-down libraries in the future.

Secondly, GSK-J4, a potent and selective inhibitor for H3K27me3 demethylases KDM6A and KDM6B as well as for H3K4me3 demethylase KDM5B [85] was investigated with regards to transcriptional and epigenomic effects in two cell systems, of importance to bone biology. The inhibitor induces apoptosis in malignant multiple myeloma cells while potentially reducing associated bone loss by inhibiting osteoclast differentiation and function. The overall comparison between the two cell types points at little overlap between the mechanisms and hence points to two modes of action. In fact, it appears that GSK-J4 inhibits osteoclast function by inhibiting H3K27me3 demethylases and hence inhibiting the activation of the transcriptional program required to induce differentiation. Conversely, in multiple myeloma GSK-

J4 differentially upregulates a number of genes by largely inhibiting KDM5B which dramatically increases the number of H3K4me3 marks.

In multiple myeloma, after 2 hours of treatment with GSK-J4, c-Jun is the main differentially upregulated transcription factor which appears to induce a down-stream pro-apoptotic gene EGR1 which could initiate the apoptotic response seen at a later time-point. The upregulated sets of genes after 6 hours of GSK-J4 treatment, along with the results obtained from other experiments by members of our group, hint at oxidative stress or protein synthesis impairment due to amino acid deprivation, as indicated by the upregulation of ATF4 and DDIT3 signalling cascades. The latter findings were confirmed by Federica Lari using Western blotting of stress- and apoptosis-related proteins, such as cleavage of PARP (Poly-(ADP-ribose) polymerase) and upregulation of Caspase 9 and EIF2AK4 (eukaryotic translation initiation factor 2 alpha kinase 4). A metabolomic assay showed the build-up of glutamine, indicating a breakdown of the tricarboxylic acid cycle cycle and deprivation of asparagine and aspartate, hence further supporting amino acid deprivation signalling arising due to the apparent inability to synthesise them.

In osteoclast precursors, on the other hand, the build-up of H3K27me3 marks appears to inhibit the induction of the osteoclast-specific transcription factor NFATc1 by possibly reducing c-Jun signalling, which is a component of the AP-1 dimer required for NFATc1 transcription. In samples treated for 7 days with RANKL and GSK-J4 the differences in the transcriptional programmes overlapping with H3K27me3 marks show inhibition of HIF2A signalling as well as other genes

important for osteoclast function, such as ATPase proton pumps or cell adhesion genes.

Despite acting on apparently different histone marks in the two cells systems and hence inducing unrelated mechanisms leading to the observed outcomes, there is one effect found in both cell systems. Metallothionein genes are among the most increased genes in multiple myeloma cells as well as in osteoclasts in response to GSK-J4 treatment. In multiple myeloma, additional experiments showed zinc ion levels to rise dramatically inside the cells within the first 30 minutes of treatment with GSK-J4. This observation could explain the dramatic increase in metallothionein gene expression and might allow to hypotheses about the possibility of zinc effects leading to eventual apoptosis in myeloma. In osteoclasts changes in intracellular zinc levels have not been investigated yet, but the observed increase in the metallothionein genes indicates a similar zinc influx mechanism could be taking place. Interestingly, a previously published study suggested an increase in zinc ions as a potent inhibiting factor of osteoclast differentiation by blocking the calcineurin differentiation pathway [181].

These assumptions stipulate new sets of experiments to be carried out in order to test the aforementioned hypotheses. Firstly, the dynamics of the intracellular zinc ion should be measured in osteoclast and macrophage samples immediately after the start of incubation with GSK-J4. Secondly, RNA and chromatin should be isolated after short time points (e.g. 20 minutes) of adding GSK-J4, while the zinc ion levels are possibly still rising, in order to determine if there is a particular transcriptional change that could explain this effect. On the other hand, considering the short time-

frame of only 20 minutes it is not unlikely that a non-transcriptional event, such as inhibition of an enzyme or transport activity, is occurring that causes the effect. Further investigations of potential ion transporters or other candidate proteins and their known regulation sites could provide insights into this possibility.

Additionally, all gene knock-out techniques discussed in this thesis are being utilised to knock-down or knock-out the genes individually or in combinations, to understand the complete mechanism and to identify the fundamental players leading to the outcomes observed. GSK-J4 effects have been validated by using LNA targeting KDM5B, KDM6A and KDM6B individually or all at once in myeloma, indicating a complex interplay between demethylases to cause a phenotypic effect. Other genes identified in the high-throughput sequencing experiment are also being tested using LNAs, lentiviral shRNA delivery techniques as well as CRISPR-Cas9.

To summarise, the ground-laying work carried out on macrophage biomarkers and suitable gene knock-down techniques in this particular cell type enables to screen and identify epigenetic factors playing a role in a multitude of macrophage functions, such differentiation, pro- and anti-inflammation, as well as polarisation across different sub-types. Furthermore, the optimised and established gene knock-down techniques enabled a number of projects that are aiming to identify key epigenetic players in other cell types, such as differentiation of mesenchymal stem cells into osteoblasts. Finally, despite having a substantial amount of data from different experiments, the precise mechanism of GSK-J4 action still remains elusive. However, the number of possibilities is reducing and soon this may provide a new avenue for better insights into multiple myeloma. The amount of high-throughput

sequencing data produced from these experiments and the numerous ways one can analyse it, offers plenty of opportunities for future discoveries.

Materials and Methods

1. Cell culture

Macrophages were generated from blood samples of healthy human volunteers. Variable amounts of blood were diluted 1:1 with PBS and 1 volume of the resulting mix added on top of 1 volume of Histopaque-1077 (Sigma, 10771) without disturbing the layers. The mixture was centrifuged at 500g for 20min at 18°C with brakes off and the middle layer of mononuclear cells was carefully transferred to a new tube with PBS. Cells were washed in PBS by centrifuging at 300g for 10min twice and then the pellet was re-suspended in MACS buffer (0.5% BSA, 2mM EDTA in PBS). 1/4th volume of CD14 beads (MACS Miltenyi Biotec, 17-1440-03) were added for positive selection and after incubating for 15min at 4°C the sample was processed through a magnetic column according to manufacturer's instructions. Monocytes were eluted in MACS buffer which is exchanged for RPMI with 10% FCS or OptiMEM with 1% human serum, or MEMα with 10% FCS. Monocytes in RPMI supplemented with 5ng/ml GMCSF for M1 differentiation or 100ng/ml MCSF for M2 differentiation. Monocytes in MEMα are supplemented with 20ng/ml MCSF. Monocytes in OptiMEM were either not supplemented with cytokines, or 50ng/ml GMCSF was added. Cells were plated out at 1.5x10^5 cells/cm^2 and allowed to differentiate for 5-7 days after which they were stimulated with 100ng/ml LPS, 0.5 nanoparticle/cell (<30µm) of Chitin, or an equal volume of PBS (negative control) for 6-12h.

Osteoclasts are produced by further stimulating mature macrophages raised in MEMα with MCSF. In fresh media cells were pre-treated with compounds of interest for 1-6h, then 50ng/ml of RANKL was added. Multi-nucleated osteoclast cells start appearing within a week.

2. Immunocytochemistry

Cells were fixed by replacing medium with 4% formaldehyde in PBS for 15min. Cells were washed with PBS, then permeabilized with 0.5% Triton X-100 in PBS for 10min. Cells were washed with PBS, then pre-blocked with 10% FCS, 10% BSA or 1% human IgG in PBS for 1h. Finally, 1:500 dilution (1µg/ml) of the following primary antibody dilutions in PBS with 5% FCS was added: mouse monoclonal IgG1 to human Activin (Abcam, ab89307, Lot GR67257-1), rabbit monoclonal IgG to human CCL1 (Invitrogen, 700654, Lot 900433A), mouse monoclonal IgG1 to human TNFα (R&D, MAB610), rabbit polyclonal IgG to human MCP-1 (Abcam, ab9669, Lot GR4808-5), mouse monoclonal IgG2B to human MMR/CD206 (R&D, MAB25341, Lot CEZR0111051), goat polyclonal IgG to human IL-10 (R&D, AF-217-NA, Lot QW0811071). Samples were incubated overnight at 4°C. The next day samples were washed 3 times 5 min each with PBS. The following secondary antibody dilutions of 1:1000 (2µg/ml) were added (emission wavelength shown): goat anti-mouse IgG1 (γ 1) 488nm (Invitrogen, A21121, Lot 997812), goat anti-mouse IgG (H+L) 488nm (Invitrogen, A11029, Lot 1073083), goat anti-rabbit IgG (H+L) 568nm (Invitrogen, A11011, Lot 898234), rabbit anti-goat IgG (H+L) 488nm (Invitrogen, A11078, Lot

454022), rabbit anti-mouse IgG (H+L) 568nm (Invitrogen, A11061, Lot 43458A). Samples were incubated for 2h at room temperature. Finally, 1:1000 DAPI or Hoechst in PBS was added to cells for 3min and samples washed twice with PBS. Samples were stored at 4˚C until ready for image acquisition with either BD Pathway 800 or GE InCell 2000 automated microscopy systems. A tile scan of the entire well surface was performed with either system.

3. Bioinformatic analysis of microscopy data

Images collected with the BD Pathway system had 12 bit pixel (px) depth (4096 levels). Image processing was carried out with AttoVision software. No background correction or shading was applied. Cell nuclei were identified in the DAPI channel by picking objects of size between 20-180px and above intensity of 1000. The nuclei were then expanded by 7px using a polygonal cell shape predictor and the overall and maximum intensities for each cell across all channels was measured.

Images collected with the InCell 2000 system had 16 bit pixel depth (65536 levels). Image processing was carried out with Cell Profiler. Due to a wide-angle lens used in the microscope, the illumination around the edges was decreased, hence, 300px from each edge was omitted. Nuclei were identified in the DAPI channel by selecting objects 10-25px in size, with Otsu Global thresholding algorithm. Cellular outlines were drawn by expanding nuclei by 10px and measuring mean, median, integrated and maximum signal intensities across all channels.

Signal intensities were analysed using R statistical software. Fluorescence was considered to be present when the intensities were above mean signal of negative control plus four standard deviations.

4. Lentivirus production and infection

The Phoenix cell line was cultured in 15-cm dishes in RPMI with 10% FCS until 90% confluent. Lipofectamine 2000 solution was prepared according to manufacturer's instructions with 140µl of reagent and 28µg lentiviral plasmid of interest, 21µg pCMV-dR8.2 dvpr (Addgene, 8455) and 7µg pCMV-VSV-G (Addgene, 8454). Lipofectamine and plasmid mixture was added to cells for 12h in antibiotic-free RPMI with 10% FCS. Then media was changed and cultured for another 48h. Lentivirus was harvested by aspirating medium, filtering through 0.4µm membrane filter and snap-freezing in dry ice. Aliquots were stored at -80˚C.

To infect cells, up to 8µg/ml of polybrene was added to the cell culture. An aliquot of lentivirus was thawed from -80˚C and diluted to required concentration with OptiMEM. Viral dilutions were added to cells, sealed and centrifuged at 1000g for 1h at 37˚C. Samples were left for 18-48h at cell culture conditions and inspected for fluorescence or other markers of transfection.

5. Chromatin Immunoprecipitation

Cells were fixed in 1% formaldehyde solution for 8 min followed by quenching of the reaction with glycine (125 mM) for 5 min. Cell were lysed in ChIP Lysis Buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH 8.1) and sonicated using Bioruptor Pico (Diagenode). Sonication was performed at 4˚C using 3-5 sets of 10 cycles (each consisting of 30s sonication, 30s refraction). Protein A Sepharose beads (Sigma, P9424) were washed and pre-blocked in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167mM NaCl, 16.7 mM Tris-HCl pH 8.1) with 5mg/ml BSA and 0.2mg/ml yeast tRNA (Sigma, R5636) for 2 hours at 4˚C. Once sonication was complete, samples were diluted 10 times with ChIP dilution buffer and incubated with pre-blocked beads for 1 hour. Beads were then removed and the pre-cleared samples were incubated overnight with 2µg/sample of anti-H3K4me3 (Millipore, 07-473 Lot 2207275) or anti-H3K27me3 (Abcam, ab108245 Lot GR161273-1) at 4˚C. Beads were washed twice with ChIP Low Salt Buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 150mM NaCl, 20mM Tris-HCl pH 8.1), once with ChIP High Salt Buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 500mM NaCl, 20mM Tris-HCl pH 8.1), once with ChIP LiCl Buffer (0.25M LiCl, 1% NP40, 1% deoxycholate, 1mM EDTA, 10mM Tris-HCl pH 8.1) and twice with TE buffer (1mM EDTA, 10mM Tris-HCl pH 8.0). Chromatin was eluted with freshly prepared ChIP Elution buffer (1% SDS, 100mM $NaHCO_3$, 250mM NaCl) and de-cross-linked at 65˚C for at least 4 hours. Proteins were digested by adding 0.7mg/ml of Proteinase K and heating at 55˚C for at least 1 hour. DNA was isolated using 1.8 volumes of AMPureXP beads (Beckman Coulter, A63881). The concentration was measured with Qubit (Life

Technologies). Enrichment of genomic fragments was determined by qPCR with primers specifically designed for genes known to have the histone marks that were pulled down. For H3K4me3 marks primers were designed around the transcription start site of ActB and for H3K27me3 marks primers were targeted at MYOG 3' untranslated region (Table 5).

6. Library preparation for next-generation sequencing

Chromatin immunoprecipitation (ChIP) libraries were prepared using NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (NEB, E7370) and NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) (NEB, E7335).

RNA samples were first depleted of ribosomal RNA (rRNA) with the GeneRead rRNA Depletion Kit (Qiagen, 180211) that utilises magnetic beads with conjugated oligonucleotides specific for rRNA. RNA-Seq libraries were then prepared using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, E7420) and NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) (NEB, E7335).

Target	Orientation	Sequence
ActB	Fwd	5'-GTGGACATCTCTTGGGCACT-3'
	Rev	5'-TCTGCAGGAGCGTACAGAAC-3'
MYOG	Fwd	5'-GGAGAAAGAAGGGGAATCACA-3'
	Rev	5'-GATAAATATAGCCAACGCCACA-3'

Table 5 Primer sequences used to quantify relative enrichment for H3K4me3 marks (ActB) and H3K27me3 (MYOG).

Importantly, at the size selection step, fragments were selected with AMPureXP beads (Beckman Coulter, A63881) first at 0.45X, then at a final 0.9X volume of beads which results in 150-400bp band size. DNA was amplified with 15 PCR cycles. Samples were submitted for sequencing aiming for 20 million reads for H3K4me3 ChIP samples, 40 million reads for H3K27me3 ChIP samples and 50 million reads for RNA samples. 50 base pair (bp) single-ended reads were obtained for ChIP samples and 80-100 bp pair-ended reads were obtained for RNA samples. If an Illumina HiSeq instrument is used, only the following groups of primers were allowed to be multiplexed: either 1, 2, 4, 10 or 3, 5, 7, 11 which allows for simultaneous calibration for all 4 bases during sequencing initiation.

7. Bioinformatic analysis of sequencing data

The quality of the reads was assessed with FastQC v0.9.2. Reads below quality scores of 25 (Illumina 1.5 encoding) and any adapter or universal primer sequences were trimmed and only reads longer than 24bp were retained by utilising cutadapt v1.2.1.

For ChIP-Seq, reads were aligned with bowtie v2.1.0 using the "`--sensitive`" parameter. H3K4me3 peaks were called using MACS v2.0.10.20131028 [235] by pre-estimating the parameter "`--extsize`" with predictd and then defining the latter value as well as additional parameters "`--keep-dup auto --nomodel`". H3K27me3 peaks were called using SoleSearchV2 [236] and additional recommended parameters for H3K27me3 peaks "`-s 0.1 -j 300`" were defined. Differential enrichment was assessed with MACS2 bdgdiff by normalising with sequencing depth and providing

the additional parameters "`--min-len 100 --max-gap 80`". The peaks were annotated using homer v4.3.

For RNA-Seq the reads were mapped onto the genome using TopHat v2.0.13 [237] using directional library parameter "`--library-type=fr-firststrand`" and additional parameters "`--mate-inner-dist 100 --mate-std-dev 80 --b2-sensitive`". The transcriptome was assembled using Cufflinks v.2.2.1 [238] and the parameters "`--multi-read-correct --library-type fr-firststrand --compatible-hits-norm`". A masking file was supplied containing tRNA and rRNA sequences generated on UCSC Table Browser. Human transcriptome annotation version hg19 was used from UCSC. Expression data was normalised across all samples and displayed in Fragments per Kilobase of Exon per Million Fragments Mapped (FPKM).

The subsequent differential expression and chromatin enrichment analysis was carried out using a range of tools. Heat maps were produced using function "heatmap.2" within the R package "gplots". Motif enrichment analysis was carried out using homer v4.3 [217] function "findMotifs.pl". Gene identification based on motif consensus sequence was carried out by employing the same function with an additional parameter "`-find`" followed by the motif binding consensus sequence. GO annotation was carried out using an online tool called DAVID [144]. Global motif enrichment data was sought on MotifMap [179].

References

1. Ogawa, M., *Differentiation and proliferation of hematopoietic stem cells.* Blood, 1993. **81**(11): p. 2844-53.
2. Till, J.E., E.A. McCulloch, and L. Siminovitch, *A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells.* Proc Natl Acad Sci U S A, 1964. **51**: p. 29-36.
3. Laiosa, C.V., M. Stadtfeld, and T. Graf, *Determinants of lymphoid-myeloid lineage diversification.* Annual review of immunology, 2006. **24**: p. 705-38.
4. Morrison, S.J., et al., *The purification and characterization of fetal liver hematopoietic stem cells.* Proc Natl Acad Sci U S A, 1995. **92**(22): p. 10302-6.
5. Christensen, J.L. and I.L. Weissman, *Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells.* Proc Natl Acad Sci U S A, 2001. **98**(25): p. 14541-6.
6. Akashi, K., et al., *A clonogenic common myeloid progenitor that gives rise to all myeloid lineages.* Nature, 2000. **404**(6774): p. 193-7.
7. Kondo, M., I.L. Weissman, and K. Akashi, *Identification of clonogenic common lymphoid progenitors in mouse bone marrow.* Cell, 1997. **91**(5): p. 661-72.
8. Metcalf, D., *Cell-cell signalling in the regulation of blood cell formation and function.* Immunol Cell Biol, 1998. **76**(5): p. 441-7.
9. Liu, F., et al., *Impaired production and increased apoptosis of neutrophils in granulocyte colony-stimulating factor receptor-deficient mice.* Immunity, 1996. **5**(5): p. 491-501.
10. Nishinakamura, R., et al., *Mice deficient for the IL-3/GM-CSF/IL-5 beta c receptor exhibit lung pathology and impaired immune response, while beta IL3 receptor-deficient mice are normal.* Immunity, 1995. **2**(3): p. 211-22.
11. Scott, E.W., et al., *Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages.* Science, 1994. **265**(5178): p. 1573-7.
12. McKercher, S.R., et al., *Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities.* EMBO J, 1996. **15**(20): p. 5647-58.
13. Zhang, D.E., et al., *Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 569-74.
14. Darnell, J.E., Jr., I.M. Kerr, and G.R. Stark, *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins.* Science, 1994. **264**(5164): p. 1415-21.
15. Holtschke, T., et al., *Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene.* Cell, 1996. **87**(2): p. 307-17.
16. Ziegler-Heitbrock, L., et al., *Nomenclature of monocytes and dendritic cells in blood.* Blood, 2010. **116**(16): p. e74-e80.
17. Geissmann, F., et al., *Blood monocytes: distinct subsets, how they relate to dendritic cells, and their possible roles in the regulation of T-cell responses.* Immunol Cell Biol, 2008. **86**(5): p. 398-408.
18. Goerdt, S. and C.E. Orfanos, *Other functions, other genes: Alternative activation of antigen-presenting cells.* Immunity, 1999. **10**(2): p. 137-142.
19. Takeda, K. and S. Akira, *Toll-like receptors in innate immunity.* Int Immunol, 2005. **17**(1): p. 1-14.

20. Kawai, T., et al., *Unresponsiveness of MyD88-deficient mice to endotoxin.* Immunity, 1999. **11**(1): p. 115-22.
21. Kawai, T., et al., *Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes.* J Immunol, 2001. **167**(10): p. 5887-94.
22. Vanden Berghe, W., et al., *Keeping up NF-kappaB appearances: epigenetic control of immunity or inflammation-triggered epigenetics.* Biochemical pharmacology, 2006. **72**(9): p. 1114-31.
23. Colonna, M., *TLR pathways and IFN-regulatory factors: To each its own.* Eur J Immunol, 2007. **37**(2): p. 306-309.
24. Mantovani, A., A. Sica, and M. Locati, *Macrophage polarization comes of age.* Immunity, 2005. **23**(4): p. 344-6.
25. Gordon, S. and F.O. Martinez, *Alternative Activation of Macrophages: Mechanism and Functions.* Immunity, 2010. **32**(5): p. 593-604.
26. Stein, M., et al., *Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation.* J Exp Med, 1992. **176**(1): p. 287-92.
27. Mokoena, T. and S. Gordon, *Human macrophage activation. Modulation of mannosyl, fucosyl receptor activity in vitro by lymphokines, gamma and alpha interferons, and dexamethasone.* The Journal of clinical investigation, 1985. **75**(2): p. 624-31.
28. Hebenstreit, D., et al., *Signaling mechanisms, interaction partners, and target genes of STAT6.* Cytokine Growth Factor Rev, 2006. **17**(3): p. 173-188.
29. Martinez, F.O., *Analysis of gene expression and gene silencing in human macrophages.* Current protocols in immunology / edited by John E. Coligan, 2012. **Chapter 14**: p. Unit 14 28 1-23.
30. Fleetwood, A.J., et al., *GM-CSF- and M-CSF-dependent macrophage phenotypes display differential dependence on type I interferon signaling.* J Leukoc Biol, 2009. **86**(2): p. 411-21.
31. Martinez, F.O., et al., *Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression.* J Immunol, 2006. **177**(10): p. 7303-7311.
32. Holliday, R., *Epigenetics: A Historical Overview.* Epigenetics, 2006. **1**(2): p. 76-80.
33. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.
34. Turner, B.M., *Defining an epigenetic code.* Nat Cell Biol, 2007. **9**(1): p. 2-6.
35. Cheng, C. and M. Gerstein, *Modeling the relative relationship of transcription factor binding and histone modifications to gene expression levels in mouse embryonic stem cells.* Nucleic Acids Res, 2012. **40**(2): p. 553-68.
36. Cheng, C., et al., *A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets.* Genome Biol, 2011. **12**(2): p. R15.
37. Grunstein, M., *Histone acetylation in chromatin structure and transcription.* Nature, 1997. **389**(6649): p. 349-52.
38. Durrin, L.K., et al., *Yeast histone H4 N-terminal sequence is required for promoter activation in vivo.* Cell, 1991. **65**(6): p. 1023-31.
39. Carrozza, M.J., et al., *The diverse functions of histone acetyltransferase complexes.* Trends Genet, 2003. **19**(6): p. 321-9.
40. Brownell, J.E., et al., *Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation.* Cell, 1996. **84**(6): p. 843-51.
41. Tercero, J.C., L.E. Riles, and R.B. Wickner, *Localized mutagenesis and evidence for post-transcriptional regulation of MAK3. A putative N-acetyltransferase required for double-stranded RNA virus propagation in Saccharomyces cerevisiae.* J Biol Chem, 1992. **267**(28): p. 20270-6.

42. Pelletier, N., et al., *MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2.* Oncogene, 2002. **21**(17): p. 2729-40.
43. Goodman, R.H. and S. Smolik, *CBP/p300 in cell growth, transformation, and development.* Genes Dev, 2000. **14**(13): p. 1553-77.
44. Zeng, L. and M.M. Zhou, *Bromodomain: an acetyl-lysine binding domain.* FEBS Lett, 2002. **513**(1): p. 124-8.
45. Filippakopoulos, P., et al., *Histone recognition and large-scale structural analysis of the human bromodomain family.* Cell, 2012. **149**(1): p. 214-31.
46. Ruthenburg, A.J., et al., *Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions.* Cell, 2011. **145**(5): p. 692-706.
47. Gallinari, P., et al., *HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics.* Cell research, 2007. **17**(3): p. 195-211.
48. Martin, C. and Y. Zhang, *The diverse functions of histone lysine methylation.* Nat Rev Mol Cell Biol, 2005. **6**(11): p. 838-49.
49. Tschiersch, B., et al., *The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes.* EMBO J, 1994. **13**(16): p. 3822-31.
50. Krouwels, I.M., et al., *A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain.* J Cell Biol, 2005. **170**(4): p. 537-49.
51. Santos-Rosa, H., et al., *Active genes are tri-methylated at K4 of histone H3.* Nature, 2002. **419**(6905): p. 407-11.
52. Schuettengruber, B., et al., *Genome regulation by polycomb and trithorax proteins.* Cell, 2007. **128**(4): p. 735-45.
53. Strahl, B.D., et al., *Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression.* Molecular and Cellular Biology, 2002. **22**(5): p. 1298-306.
54. Krogan, N.J., et al., *Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II.* Molecular and Cellular Biology, 2003. **23**(12): p. 4207-18.
55. Czermin, B., et al., *Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites.* Cell, 2002. **111**(2): p. 185-96.
56. Margueron, R., et al., *Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms.* Molecular cell, 2008. **32**(4): p. 503-18.
57. Schoeftner, S., et al., *Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing.* EMBO J, 2006. **25**(13): p. 3110-22.
58. Vire, E., et al., *The Polycomb group protein EZH2 directly controls DNA methylation.* Nature, 2006. **439**(7078): p. 871-874.
59. Feng, Q., et al., *Methylation of H3-Lysine 79 Is Mediated by a New Family of HMTases without a SET Domain.* Curr Biol, 2002. **12**(12): p. 1052-1058.
60. Greer, E.L. and Y. Shi, *Histone methylation: a dynamic mark in health, disease and inheritance.* Nat Rev Genet, 2012. **13**(5): p. 343-57.
61. Schindler, U., H. Beckmann, and A.R. Cashmore, *HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region.* Plant J, 1993. **4**(1): p. 137-50.
62. Doyon, Y., et al., *ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation.* Molecular cell, 2006. **21**(1): p. 51-64.
63. Lu, R. and G.G. Wang, *Tudor: a versatile family of histone methylation 'readers'.* Trends Biochem Sci, 2013. **38**(11): p. 546-555.
64. Klose, R.J., et al., *The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine[thinsp]9 and lysine[thinsp]36.* Nature, 2006. **442**(7100): p. 312-316.

65. Ruthenburg, A.J., et al., *Multivalent engagement of chromatin modifications by linked binding modules.* Nat Rev Mol Cell Biol, 2007. **8**(12): p. 983-994.
66. Blus, B.J., K. Wiggins, and S. Khorasanizadeh, *Epigenetic virtues of chromodomains.* Crit Rev Biochem Mol Biol, 2011. **46**(6): p. 507-526.
67. Gorinsek, B., F. Gubensek, and D. Kordis, *Evolutionary genomics of chromoviruses in eukaryotes.* Mol Biol Evol, 2004. **21**(5): p. 781-98.
68. Zeng, W., A.R. Ball Jr, and K. Yokomori, *HP1: Heterochromatin binding proteins working the genome.* Epigenetics, 2010. **5**(4): p. 287-292.
69. Rosnoblet, C., et al., *Analysis of the human HP1 interactome reveals novel binding partners.* Biochem Biophys Res Commun, 2011. **413**(2): p. 206-11.
70. Lai, A.Y. and P.A. Wade, *Cancer biology and NuRD: a multifaceted chromatin remodelling complex.* Nat Rev Cancer, 2011. **11**(8): p. 588-96.
71. Allis, C.D., et al., *Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in Tetrahymena micronuclei.* Cell, 1980. **20**(1): p. 55-64.
72. Ahmad, K. and S. Henikoff, *The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly.* Molecular cell, 2002. **9**(6): p. 1191-200.
73. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell, 2004. **119**(7): p. 941-53.
74. Ciccone, D.N., et al., *KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints.* Nature, 2009. **461**(7262): p. 415-8.
75. Tsukada, Y., et al., *Histone demethylation by a family of JmjC domain-containing proteins.* Nature, 2006. **439**(7078): p. 811-6.
76. Kooistra, S.M. and K. Helin, *Molecular mechanisms and potential functions of histone demethylases.* Nat Rev Mol Cell Biol, 2012. **13**(5): p. 297-311.
77. Clissold, P.M. and C.P. Ponting, *JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2 β.* Trends Biochem Sci, 2001. **26**(1): p. 7-9.
78. Benevolenskaya, E.V., et al., *Binding of pRB to the PHD protein RBP2 promotes cellular differentiation.* Molecular cell, 2005. **18**(6): p. 623-35.
79. Chicas, A., et al., *H3K4 demethylation by Jarid1a and Jarid1b contributes to retinoblastoma-mediated gene silencing during cellular senescence.* Proc Natl Acad Sci U S A, 2012. **109**(23): p. 8971-6.
80. Kuzbicki, L., et al., *JARID1B expression in human melanoma and benign melanocytic skin lesions.* Melanoma Res, 2013. **23**(1): p. 8-12.
81. Lu, P.J., et al., *A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer.* J Biol Chem, 1999. **274**(22): p. 15633-45.
82. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.
83. Agger, K., et al., *UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development.* Nature, 2007. **449**(7163): p. 731-4.
84. De Santa, F., et al., *The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing.* Cell, 2007. **130**(6): p. 1083-94.
85. Kruidenier, L., et al., *A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response.* Nature, 2012. **488**(7411): p. 404-8.
86. Das, N.D., et al., *Gene networking and inflammatory pathway analysis in a JMJD3 knockdown human monocytic cell line.* Cell Biochem Funct, 2012. **30**(3): p. 224-32.
87. Iwasaki, H. and K. Akashi, *Myeloid lineage commitment from the hematopoietic stem cell.* Immunity, 2007. **26**(6): p. 726-40.
88. Brown, K., *The Epigenetics of Haematopoiesis*. Cell Determination during Hematopoiesis, ed. G. Brown and R. Ceredig. 2009. 285-309.

89. Bender, M.A., et al., *Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region.* Molecular cell, 2000. **5**(2): p. 387-93.
90. Yin, W., et al., *Histone acetylation at the human beta-globin locus changes with developmental age.* Blood, 2007. **110**(12): p. 4101-7.
91. Rogatsky, I., et al., *Epigenetics and the IRFs: a complex interplay in the control of immunity and autoimmunity.* Autoimmunity, 2014. **47**(4): p. 242-55.
92. Klampfer, L., et al., *Requirement of histone deacetylase activity for signaling by STAT1.* J Biol Chem, 2004. **279**(29): p. 30358-68.
93. Verreck, F.A., et al., *Phenotypic and functional profiling of human proinflammatory type-1 and anti-inflammatory type-2 macrophages in response to microbial antigens and IFN-gamma- and CD40L-mediated costimulation.* J Leukoc Biol, 2006. **79**(2): p. 285-93.
94. Krausgruber, T., et al., *IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses.* Nat Immunol, 2011. **12**(3): p. 231-8.
95. Shen, J., et al., *BMP2-induced inflammation can be suppressed by the osteoinductive growth factor NELL-1.* Tissue engineering. Part A, 2013. **19**(21-22): p. 2390-401.
96. Vermeulen, L., et al., *Transcriptional activation of the NF-kappa B p65 subunit by mitogen- and stress-activated protein kinase-1 (MSK1).* EMBO J, 2003. **22**(6): p. 1313-1324.
97. Natoli, G., et al., *Interactions of NF-kappa B with chromatin: the art of being at the right place at the right time.* Nat Immunol, 2005. **6**(5): p. 439-445.
98. Saccani, S., et al., *Degradation of promoter-bound p65/RelA is essential for the prompt termination of the nuclear factor kappaB response.* J Exp Med, 2004. **200**(1): p. 107-13.
99. Saccani, S., S. Pantano, and G. Natoli, *Two waves of nuclear factor kappaB recruitment to target promoters.* J Exp Med, 2001. **193**(12): p. 1351-9.
100. Saccani, S., S. Pantano, and G. Natoli, *p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment.* Nat Immunol, 2002. **3**(1): p. 69-75.
101. Saraiva, M. and A. O'Garra, *The regulation of IL-10 production by immune cells.* Nature reviews. Immunology, 2010. **10**(3): p. 170-181.
102. Satoh, T., et al., *The Jmjd3-Irf4 axis regulates M2 macrophage polarization and host responses against helminth infection.* Nat Immunol, 2010. **11**(10): p. 936-44.
103. Ospelt, C. and S. Gay, *The role of resident synovial cells in destructive arthritis.* Clin Rheumatol, 2008. **22**(2): p. 239-52.
104. Lin, H.S., et al., *Anti-rheumatic activities of histone deacetylase (HDAC) inhibitors in vivo in collagen-induced arthritis in rodents.* British journal of pharmacology, 2007. **150**(7): p. 862-72.
105. Maciejewska-Rodrigues, H., et al., *Epigenetics and rheumatoid arthritis: the role of SENP1 in the regulation of MMP-1 expression.* J Autoimmun, 2010. **35**(1): p. 15-22.
106. Gillespie, J., et al., *Histone deacetylases are dysregulated in rheumatoid arthritis and a novel histone deacetylase 3-selective inhibitor reduces interleukin-6 production by peripheral blood mononuclear cells from rheumatoid arthritis patients.* Arthritis Rheum, 2012. **64**(2): p. 418-22.
107. Baxter, E., et al., *Epigenetic regulation in cancer progression.* Cell and Bioscience, 2014. **4**.
108. Bert, Saul A., et al., *Regional Activation of the Cancer Genome by Long-Range Epigenetic Remodeling.* Cancer Cell, 2013. **23**(1): p. 9-22.
109. Dalgliesh, G.L., et al., *Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes.* Nature, 2010. **463**(7279): p. 360-363.
110. Harris, William J., et al., *The Histone Demethylase KDM1A Sustains the Oncogenic Potential of MLL-AF9 Leukemia Stem Cells.* Cancer Cell. **21**(4): p. 473-487.
111. Helin, K. and D. Dhanak, *Chromatin proteins and modifications as drug targets.* Nature, 2013. **502**(7472): p. 480-488.

112. Chapman, M.A., et al., *Initial genome sequencing and analysis of multiple myeloma.* Nature, 2011. **471**(7339): p. 467-472.
113. Martinez-Garcia, E., et al., *The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells*. Vol. 117. 2011. 211-220.
114. Heinrich, P.C., et al., *Principles of interleukin (IL)-6-type cytokine signalling and its regulation.* Biochem J, 2003. **374**(Pt 1): p. 1-20.
115. Xiong, H., et al., *Trichostatin A, a histone deacetylase inhibitor, suppresses JAK2/STAT3 signaling via inducing the promoter-associated histone acetylation of SOCS1 and SOCS3 in human colorectal cancer cells.* Molecular Carcinogenesis, 2012. **51**(2): p. 174-184.
116. Bowers, E.M., et al., *Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor.* Chem Biol, 2010. **17**(5): p. 471-82.
117. Finnin, M.S., et al., *Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors.* Nature, 1999. **401**(6749): p. 188-93.
118. Furumai, R., et al., *FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases.* Cancer research, 2002. **62**(17): p. 4916-21.
119. Mann, B.S., et al., *FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma.* Oncologist, 2007. **12**(10): p. 1247-52.
120. Nicodeme, E., et al., *Suppression of inflammation by a synthetic histone mimic.* Nature, 2010. **468**(7327): p. 1119-23.
121. Filippakopoulos, P., et al., *Selective inhibition of BET bromodomains.* Nature, 2010. **468**(7327): p. 1067-73.
122. James, L.I. and S.V. Frye, *Targeting chromatin readers.* Clin Pharmacol Ther, 2013. **93**(4): p. 312-4.
123. Daigle, S.R., et al., *Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor.* Cancer Cell, 2011. **20**(1): p. 53-65.
124. Finley, A. and Robert A. Copeland, *Small Molecule Control of Chromatin Remodeling.* Chemistry & Biology, 2014. **21**(9): p. 1196-1210.
125. Knutson, S.K., et al., *Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma.* Mol Cancer Ther, 2014. **13**(4): p. 842-54.
126. Heinemann, B., et al., *Inhibition of demethylases by GSK-J1/J4.* Nature, 2014. **514**(7520): p. E1-E2.
127. Brotman, B. and A.M. Prince, *Gamma-glutamyltransferase as a potential surrogate marker for detection of the non-A, non-B carrier state.* Vox Sang, 1988. **54**(3): p. 144-7.
128. Aronson, J.K., *Biomarkers and surrogate endpoints.* Br J Clin Pharmacol, 2005. **59**(5): p. 491-494.
129. Hill, A.B., *The Environment and Disease: Association or Causation?* Proceedings of the Royal Society of Medicine, 1965. **58**: p. 295-300.
130. Dumont, G.J., et al., *Biomarkers for the effects of selective serotonin reuptake inhibitors (SSRIs) in healthy subjects.* Br J Clin Pharmacol, 2005. **59**(5): p. 495-510.
131. Gordon, S., *Alternative activation of macrophages.* Nature reviews. Immunology, 2003. **3**(1): p. 23-35.
132. Mosser, D.M. and J.P. Edwards, *Exploring the full spectrum of macrophage activation.* Nature reviews. Immunology, 2008. **8**(12): p. 958-969.
133. Raes, G., et al., *Differential expression of FIZZ1 and Ym1 in alternatively versus classically activated macrophages.* J Leukoc Biol, 2002. **71**(4): p. 597-602.
134. Wynes, M.W. and D.W. Riches, *Induction of macrophage insulin-like growth factor-I expression by the Th2 cytokines IL-4 and IL-13.* J Immunol, 2003. **171**(7): p. 3550-9.
135. Fadok, V.A., et al., *Macrophages that have ingested apoptotic cells in vitro inhibit proinflammatory cytokine production through autocrine/paracrine mechanisms involving*

TGF-beta, PGE2, and PAF. The Journal of clinical investigation, 1998. **101**(4): p. 890-8.
136. Heintzman, N.D., et al., *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome.* Nat Genet, 2007. **39**(3): p. 311-8.
137. Ghisletti, S., et al., *Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages.* Immunity, 2010. **32**(3): p. 317-28.
138. De Santa, F., et al., *A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers.* PLoS Biol, 2010. **8**(5).
139. Nelson, D.E., et al., *Oscillations in NF-κB Signaling Control the Dynamics of Gene Expression.* Science, 2004. **306**(5696): p. 704-708.
140. Lacey, D.C., et al., *Defining GM-CSF- and macrophage-CSF-dependent macrophage responses by in vitro models.* J Immunol, 2012. **188**(11): p. 5752-65.
141. Barrios-Rodiles, M. and K. Chadee, *Novel regulation of cyclooxygenase-2 expression and prostaglandin E2 production by IFN-gamma in human macrophages.* J Immunol, 1998. **161**(5): p. 2441-8.
142. Uddin, M.J., et al., *Selective visualization of cyclooxygenase-2 in inflammation and cancer by targeted fluorescent imaging agents.* Cancer research, 2010. **70**(9): p. 3618-27.
143. Beyer, M., et al., *High-resolution transcriptome of human macrophages.* PLoS One, 2012. **7**(9): p. e45466.
144. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
145. Martinez, F.O. and S. Gordon, *The M1 and M2 paradigm of macrophage activation: time for reassessment.* F1000Prime Rep, 2014. **6**: p. 13.
146. Sica, A. and A. Mantovani, *Macrophage plasticity and polarization: in vivo veritas.* The Journal of clinical investigation, 2012. **122**(3): p. 787-95.
147. Pelt, A.C., *Glucocorticoids : effects, action mechanisms, and therapeutic uses*. Biochemistry research trends. 2011, Hauppauge, N.Y.: Nova Science. xiii, 250 p.
148. Ehrchen, J., et al., *Glucocorticoids induce differentiation of a specifically activated, anti-inflammatory subtype of human monocytes.* Blood, 2007. **109**(3): p. 1265-74.
149. Lawrence, T. and G. Natoli, *Transcriptional regulation of macrophage polarization: enabling diversity with identity.* Nature reviews. Immunology, 2011. **11**(11): p. 750-761.
150. Takahashi, K. and M. Naito, *Development, differentiation, and proliferation of macrophages in the rat yolk sac.* Tissue Cell, 1993. **25**(3): p. 351-62.
151. Gomez Perdiguero, E. and F. Geissmann, *Myb-independent macrophages: a family of cells that develops with their tissue of residence and is involved in its homeostasis.* Immunology, 2013. **140**: p. 79-79.
152. Haldar, M. and K.M. Murphy, *Origin, development, and homeostasis of tissue-resident macrophages.* Immunol Rev, 2014. **262**(1): p. 25-35.
153. Xue, J., et al., *Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation.* Immunity, 2014. **40**(2): p. 274-288.
154. Wendland, J., *PCR-based methods facilitate targeted gene manipulations and cloning procedures.* Curr Genet, 2003. **44**(3): p. 115-23.
155. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.
156. Karagiannis, T.C. and A. El-Osta, *RNA interference and potential therapeutic applications of short interfering RNAs.* Cancer Gene Ther, 2005. **12**(10): p. 787-95.
157. Wahlestedt, C., et al., *Potent and nontoxic antisense oligonucleotides containing locked nucleic acids.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5633-8.
158. Leyva, F.J., et al., *Evaluation of transduction efficiency in macrophage colony-stimulating factor differentiated human macrophages using HIV-1 based lentiviral vectors.* BMC Biotechnol, 2011. **11**: p. 13.

159. Robbins, C.S., et al., *Local proliferation dominates lesional macrophage accumulation in atherosclerosis.* Nat Med, 2013. **19**(9): p. 1166-1172.
160. Vadiveloo, P.K., *Macrophages--proliferation, activation, and cell cycle proteins.* J Leukoc Biol, 1999. **66**(4): p. 579-82.
161. Kondo, K., H. Kaneshima, and E.S. Mocarski, *Human Cytomegalovirus Latent infection of Granulocyte-Macrophage Progenitors.* Proc Natl Acad Sci U S A, 1994. **91**(25): p. 11879-11883.
162. Taylor-Wiedeman, J., et al., *Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells.* Journal of General Virology, 1991. **72**(9): p. 2059-2064.
163. Kramm, A., DPhil Thesis: *Identification and characterisation of epigenetic mechanisms in osteoblast differentiation of human mesenchymal stem cells*, at *Medical Sciences Division - NDORMS (Nuffield Department of Orthopaedics,Rheumatology and Musculoskeletal Sciences)* 2014, University of Oxford
164. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
165. Che, K.H., DPhil Thesis: *Development of biochemical tools to characterise human H3K27 histone demethylase JmjD3*, at *Medical Sciences Division - NDORMS (Nuffield Department of Orthopaedics,Rheumatology and Musculoskeletal Sciences)* 2013, University of Oxford
166. Ducy, P., T. Schinke, and G. Karsenty, *The osteoblast: a sophisticated fibroblast under central surveillance.* Science, 2000. **289**(5484): p. 1501-4.
167. Knothe Tate, M.L., et al., *The osteocyte.* Int J Biochem Cell Biol, 2004. **36**(1): p. 1-8.
168. Teitelbaum, S.L., *Bone resorption by osteoclasts.* Science, 2000. **289**(5484): p. 1504-8.
169. Schlesinger, P.H., et al., *Characterization of the osteoclast ruffled border chloride channel and its role in bone resorption.* J Biol Chem, 1997. **272**(30): p. 18636-43.
170. Gowen, M., et al., *Cathepsin K knockout mice develop osteopetrosis due to a deficit in matrix degradation but not demineralization.* J Bone Miner Res, 1999. **14**(10): p. 1654-63.
171. Theill, L.E., W.J. Boyle, and J.M. Penninger, *RANK-L AND RANK: T cells, bone loss, and mammalian evolution.* Annual review of immunology, 2002. **20**: p. 795-823.
172. Yasuda, H., et al., *Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL.* Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3597-3602.
173. Li, J., et al., *RANK is the intrinsic hematopoietic cell surface receptor that controls osteoclastogenesis and regulation of bone mass and calcium metabolism.* Proc Natl Acad Sci U S A, 2000. **97**(4): p. 1566-1571.
174. Simonet, W.S., et al., *Osteoprotegerin: A Novel Secreted Protein Involved in the Regulation of Bone Density.* Cell, 1997. **89**(2): p. 309-319.
175. Yagi, M., et al., *DC-STAMP is essential for cell–cell fusion in osteoclasts and foreign body giant cells.* J Exp Med, 2005. **202**(3): p. 345-351.
176. Takayanagi, H., et al., *Induction and Activation of the Transcription Factor NFATc1 (NFAT2) Integrate RANKL Signaling in Terminal Differentiation of Osteoclasts.* Dev Cell, 2002. **3**(6): p. 889-901.
177. Kobayashi, N., et al., *Segregation of TRAF6-mediated signaling pathways clarifies its role in osteoclastogenesis.* EMBO J, 2001. **20**(6): p. 1271-1280.
178. Matsuo, K., et al., *Fosl1 is a transcriptional target of c-Fos during osteoclast differentiation.* Nat Genet, 2000. **24**(2): p. 184-187.
179. Daily, K., et al., *MotifMap: integrative genome-wide maps of regulatory motif sites for model species.* BMC Bioinformatics, 2011. **12**: p. 495.
180. Shinohara, M. and H. Takayanagi, *Novel osteoclast signaling mechanisms.* Curr Osteoporos Rep, 2007. **5**(2): p. 67-72.

181. Park, K.H., et al., *Zinc inhibits osteoclast differentiation by suppression of Ca2+-Calcineurin-NFATc1 signaling pathway.* Cell Commun Signal, 2013. **11**: p. 74.
182. Pageau, S.C., *Denosumab.* mAbs, 2009. **1**(3): p. 210-215.
183. Sundaram, K., et al., *RANK ligand signaling modulates the matrix metalloproteinase-9 gene expression during osteoclast differentiation.* Exp Cell Res, 2007. **313**(1): p. 168-178.
184. Lehenkari, P., et al., *Carbonic Anhydrase II Plays a Major Role in Osteoclast Differentiation and Bone Resorption by Effecting the Steady State Intracellular pH and Ca2+.* Exp Cell Res, 1998. **242**(1): p. 128-137.
185. Stoller, P., et al., *Polarization-modulated second harmonic generation in collagen.* Biophys J, 2002. **82**(6): p. 3330-42.
186. Verrier, S., et al., *ADAM gene expression and regulation during human osteoclast formation.* Bone, 2004. **35**(1): p. 34-46.
187. Väänänen, H.K. and M. Horton, *The osteoclast clear zone is a specialized cell-extracellular matrix adhesion structure.* J Cell Sci, 1995. **108**(8): p. 2729-2732.
188. Nakashima, Y. and T. Haneji, *Stimulation of Osteoclast Formation by RANKL Requires Interferon Regulatory Factor-4 and Is Inhibited by Simvastatin in a Mouse Model of Bone Loss.* PLoS One, 2013. **8**(9): p. e72033.
189. Zhang, H., et al., *Crosstalk between STAT3 and RANK signaling pathways during osteoclastogenesis.* J Immunol, 2012. **188**.
190. Utting, J.C., et al., *Hypoxia stimulates osteoclast formation from human peripheral blood.* Cell Biochem Funct, 2010. **28**(5): p. 374-380.
191. Wang, W.a., DPhil Thesis: *Investigating the role of CCN1, CCN2, and CCN6 in osteoclast and osteoblast physiology*, at 2012, University of Aberdeen
192. Bailey, T., et al., *Practical Guidelines for the Comprehensive Analysis of ChIP-seq Data.* PLoS Comput Biol, 2013. **9**(11): p. e1003326.
193. Mendenhall, E.M. and B.E. Bernstein, *Chromatin state maps: new technologies, new insights.* Curr Opin Genet Dev, 2008. **18**(2): p. 109-115.
194. Hsieh, J.J., E.H. Cheng, and S.J. Korsmeyer, *Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression.* Cell, 2003. **115**(3): p. 293-303.
195. Wan, M., et al., *Transcriptional mechanisms of bone morphogenetic protein-induced osteoprotegrin gene expression.* J Biol Chem, 2001. **276**(13): p. 10119-25.
196. Kawai, S., et al., *Zinc-Finger Transcription Factor Odd-Skipped Related 2 Is One of the Regulators in Osteoblast Proliferation and Bone Formation.* J Bone Miner Res, 2007. **22**(9): p. 1362-1372.
197. Iacobelli, M., W. Wachsman, and K.L. McGuire, *Repression of IL-2 Promoter Activity by the Novel Basic Leucine Zipper p21SNFT Protein.* J Immunol, 2000. **165**(2): p. 860-868.
198. Graham-Rowe, D., *Overview: Multiple lines of attack.* Nature, 2011. **480**(7377): p. S34-S35.
199. Slifka, M.K. and R. Ahmed, *Long-lived plasma cells: a mechanism for maintaining persistent antibody production.* Curr Opin Immunol, 1998. **10**(3): p. 252-8.
200. Shapiro-Shelef, M. and K. Calame, *Plasma cell differentiation and multiple myeloma.* Curr Opin Immunol, 2004. **16**(2): p. 226-34.
201. Anderson, K.C. and R.D. Carrasco, *Pathogenesis of myeloma.* Annual review of pathology, 2011. **6**: p. 249-74.
202. Shapiro-Shelef, M., et al., *Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells.* Immunity, 2003. **19**(4): p. 607-20.
203. Lin, Y., K. Wong, and K. Calame, *Repression of c-myc transcription by Blimp-1, an inducer of terminal B cell differentiation.* Science, 1997. **276**(5312): p. 596-9.
204. Gass, J.N., N.M. Gifford, and J.W. Brewer, *Activation of an unfolded protein response during differentiation of antibody-secreting B cells.* J Biol Chem, 2002. **277**(50): p. 49047-54.

205. Iwakoshi, N.N., et al., *Plasma cell differentiation and the unfolded protein response intersect at the transcription factor XBP-1.* Nat Immunol, 2003. **4**(4): p. 321-9.
206. Iwakoshi, N.N., A.H. Lee, and L.H. Glimcher, *The X-box binding protein-1 transcription factor is required for plasma cell differentiation and the unfolded protein response.* Immunol Rev, 2003. **194**: p. 29-38.
207. Klein, U., et al., *Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination.* Nat Immunol, 2006. **7**(7): p. 773-82.
208. Marecki, S. and M.J. Fenton, *The role of IRF-4 in transcriptional regulation.* J Interferon Cytokine Res, 2002. **22**(1): p. 121-33.
209. Hideshima, T., P. Richardson, and K.C. Anderson, *Novel therapeutic approaches for multiple myeloma.* Immunol Rev, 2003. **194**: p. 164-76.
210. Zhan, F., et al., *Gene expression profiling of human plasma cell differentiation and classification of multiple myeloma based on similarities to distinct stages of late-stage B-cell development.* Blood, 2003. **101**(3): p. 1128-40.
211. Davies, F.E., et al., *Insights into the multistep transformation of MGUS to myeloma using microarray expression analysis.* Blood, 2003. **102**(13): p. 4504-11.
212. Tian, E., et al., *The role of the Wnt-signaling antagonist DKK1 in the development of osteolytic lesions in multiple myeloma.* The New England journal of medicine, 2003. **349**(26): p. 2483-94.
213. Smith, E.M., K. Boyd, and F.E. Davies, *The potential role of epigenetic therapy in multiple myeloma.* Br J Haematol, 2010. **148**(5): p. 702-13.
214. Pawlyn, C., et al., *Current and potential epigenetic targets in multiple myeloma.* Epigenomics, 2014. **6**(2): p. 215-228.
215. Kim, J.Y., et al., *Multiple-myeloma-related WHSC1/MMSET isoform RE-IIBP is a histone methyltransferase with transcriptional repression activity.* Molecular and Cellular Biology, 2008. **28**(6): p. 2023-34.
216. van Haaften, G., et al., *Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer.* Nat Genet, 2009. **41**(5): p. 521-3.
217. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Molecular cell, 2010. **38**(4): p. 576-89.
218. Bieging, K.T., S.S. Mello, and L.D. Attardi, *Unravelling mechanisms of p53-mediated tumour suppression.* Nat Rev Cancer, 2014. **14**(5): p. 359-370.
219. Laity, J.H. and G.K. Andrews, *Understanding the mechanisms of zinc-sensing by metal-response element binding transcription factor-1 (MTF-1).* Arch Biochem Biophys, 2007. **463**(2): p. 201-210.
220. Holien, T., et al., *Addiction to c-MYC in multiple myeloma*. Vol. 120. 2012. 2450-2453.
221. Popovic, R., et al., *Histone methyltransferase MMSET/NSD2 alters EZH2 binding and reprograms the myeloma epigenome through global and focal changes in H3K36 and H3K27 methylation.* PLoS Genet, 2014. **10**(9): p. e1004566.
222. Chen, L., et al., *Identification of early growth response protein 1 (EGR-1) as a novel target for JUN-induced apoptosis in multiple myeloma.* Blood, 2010. **115**(1): p. 61-70.
223. Hess, S.D., et al., *Functional characterization of human N-methyl-D-aspartate subtype 1A/2D receptors.* J Neurochem, 1998. **70**(3): p. 1269-79.
224. Ameri, K. and A.L. Harris, *Activating transcription factor 4.* Int J Biochem Cell Biol, 2008. **40**(1): p. 14-21.
225. Morishita, J., et al., *CDK7 regulates the mitochondrial localization of a tail-anchored proapoptotic protein, Hid.* Cell Rep, 2013. **5**(6): p. 1481-8.
226. Gade, P., et al., *Critical role for transcription factor C/EBP-beta in regulating the expression of death-associated protein kinase 1.* Molecular and Cellular Biology, 2008. **28**(8): p. 2528-48.

227. Hahn, J., et al., *Apoptosis Induction and Growth Suppression by U19/Eaf2 Is Mediated Through its ELL-Binding Domain.* Prostate, 2007. **67**(2): p. 146-153.
228. Sandberg, E.M. and P.P. Sayeski, *Jak2 Tyrosine Kinase Mediates Oxidative Stress-induced Apoptosis in Vascular Smooth Muscle Cells.* J Biol Chem, 2004. **279**(33): p. 34547-34552.
229. Tur, M.K., et al., *Targeted restoration of down-regulated DAPK2 tumor suppressor activity induces apoptosis in Hodgkin lymphoma cells.* J Immunother, 2009. **32**(5): p. 431-41.
230. Brown, A.M., et al., *Zn2+ Inhibits α-Ketoglutarate-stimulated Mitochondrial Respiration and the Isolated α-Ketoglutarate Dehydrogenase Complex.* J Biol Chem, 2000. **275**(18): p. 13441-13447.
231. Burke, B., et al., *Macrophages in gene therapy: cellular delivery vehicles and in vivo targets.* J Leukoc Biol, 2002. **72**(3): p. 417-428.
232. Kluth, D.C., et al., *Macrophages Transfected with Adenovirus to Express IL-4 Reduce Inflammation in Experimental Glomerulonephritis.* J Immunol, 2001. **166**(7): p. 4728-4736.
233. Bernards, R., T.R. Brummelkamp, and R.L. Beijersbergen, *shRNA libraries and their use in cancer genetics.* Nat Meth, 2006. **3**(9): p. 701-706.
234. Michiels, F., et al., *Arrayed adenoviral expression libraries for functional screening.* Nat Biotech, 2002. **20**(11): p. 1154-1157.
235. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
236. Blahnik, K.R., et al., *Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data.* Nucleic Acids Res, 2010. **38**(3): p. e13-e13.
237. Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biology, 2013. **14**(4): p. R36.
238. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.

Supplementary data

Supplementary Table 1 The top 100 up-regulated genes in 24h GSK-J4 treated and RANKL-stimulated MCSF-macrophages as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene Name	1d	7d	Gene Name
MIR5047	2575.13	1	microRNA 5047
SCARNA23	396.807	1	small Cajal body-specific RNA 23
MIR612	322.766	612.486	microRNA 612
MIR6740	213.717	1	microRNA 6740
TSTD1	53.9766	1	thiosulfate sulfurtransferase (rhodanese)-like domain containing 1
PITPNA-AS1	35.1346	1	PITPNA antisense RNA 1
AAMDC	28.5336	1	adipogenesis associated, Mth938 domain containing
CIDECP	27.1688	3.423352	cell death-inducing DFFA-like effector c pseudogene
C1orf54	23.7466	1	chromosome 1 open reading frame 54
HIST3H2BB	23.1613	1	histone cluster 3, H2bb
LY86	22.8679	3.251458	lymphocyte antigen 86
SPATA45	22.415	1	Spermatogenesis Associated 4
HIST1H3E	22.3116	4.650479	histone cluster 1, H3e
NAPA-AS1	21.3687	0.031141	NAPA antisense RNA 1
PQLC3	17.5142	33.5615	PQ loop repeat containing 3
PPIL3	17.3216	1	peptidylprolyl isomerase (cyclophilin)-like 3
MT1F	16.4519	1	metallothionein 1F
ZSCAN9	16.388	1	zinc finger and SCAN domain containing 9
S100A8	15.3934	1	S100 calcium binding protein A8
AZU1	14.0329	0.070071	azurocidin 1
GAL	13.8485	0.008087	galanin/GMAP prepropeptide
EAF2	13.4453	1	ELL associated factor 2
TP53I3	13.3222	1	tumor protein p53 inducible protein 3
RPS27	13.3059	1	ribosomal protein S27
RAB20	13.0421	1.043898	RAB20, member RAS oncogene family
TESC	12.3851	1	tescalcin
HIATL2	12.0886	1	hippocampus abundant transcript-like 2
RPL39L	12.0766	1	ribosomal protein L39-like
SMCR9	11.8409	1	Smith-Magenis syndrome chromosome region, candidate 9
SNHG11	11.4582	0.938878	small nucleolar RNA host gene 11 (non-protein coding)
C5AR2	11.4192	0.080142	complement component 5a receptor 2
AKIP1	11.34823	0.067596	A kinase (PRKA) interacting protein 1
OR5D14	10.96377	1	olfactory receptor, family 5, subfamily D, member 14
ELAVL2	10.92249	1	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 (Hu antigen B)
RRN3P1	10.87061	1	RNA polymerase I transcription factor homolog (S. cerevisiae) pseudogene 1
HIGD1C	10.52533	1	HIG1 hypoxia inducible domain family, member 1C
TIFAB	10.45305	10.85544	TRAF-interacting protein with forkhead-associated domain, family member B
ZFPL1	10.3847	2.109157	zinc finger protein-like 1
CDK5	10.35665	1.728418	cyclin-dependent kinase 5
LPPR2	9.753992	0.696055	lipid phosphate phosphatase-related protein type 2
SIVA1	9.752924	0.575302	SIVA1, apoptosis-inducing factor
ULBP3	9.67716	1	UL16 binding protein 3
IL18	9.423752	5.115506	interleukin 18 (interferon-gamma-inducing factor)
N6AMT2	9.28604	1	N-6 adenine-specific DNA methyltransferase 2 (putative)
DIP2A-IT1	9.20476	1	DIP2A intronic transcript 1 (non-protein coding)
LLGL2	9.16949	41.9121	lethal giant larvae homolog 2 (Drosophila)
FAM177A1	9.15966	1.076051	family with sequence similarity 177, member A1
TTC9C	9.120738	1	tetratricopeptide repeat domain 9C
TSHR	9.101001	1	thyroid stimulating hormone receptor
HSPB1	8.98833	0.035569	heat shock 27kDa protein 1
RPS4Y1	8.98833	0.035569	ribosomal protein S4, Y-linked 1
ZNF669	8.91884	13.7338	zinc finger protein 669
ECM1	8.781058	1.59345	extracellular matrix protein 1
GNRH1	8.72132	0.292373	gonadotropin-releasing hormone 1 (luteinizing-releasing hormone)
TRAF3IP2-AS1	8.623318	2.18693	TRAF3IP2 antisense RNA 1
C4orf48	8.416332	0.004448	chromosome 4 open reading frame 48
KLRAP1	8.371477	1	killer cell lectin-like receptor subfamily A pseudogene 1
KCNE1L	8.33019	1	KCNE1-like

IPO9-AS1	8.22829	1	IPO9 antisense RNA 1
TMEM42	8.22044	1	transmembrane protein 42
PMS2P4	8.11109	1	postmeiotic segregation increased 2 pseudogene 4
C2orf48	8.03397	1	chromosome 2 open reading frame 48
ZC3H10	7.75565	1	zinc finger CCCH-type containing 10
KHSRP	7.651493	1.37723	KH-type splicing regulatory protein
YBEY	7.61846	1	ybeY metallopeptidase (putative)
TM2D3	7.57128	4.12179	TM2 domain containing 3
ACP5	7.427952	0.35875	acid phosphatase 5, tartrate resistant
IFT27	7.41307	4.04688	intraflagellar transport 27 homolog (Chlamydomonas)
SERPINA6	7.403313	1	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 6
GALT	7.322707	1.401257	galactose-1-phosphate uridylyltransferase
NUDT14	7.267035	0.356641	nudix (nucleoside diphosphate linked moiety X)-type motif 14
FAM174A	7.26663	0.399602	family with sequence similarity 174, member A
EPM2A	7.24447	1.848222	epilepsy, progressive myoclonus type 2A, Lafora disease (laforin)
FKBP1B	7.17711	1	FK506 binding protein 1B, 12.6 kDa
PRKAG1	7.142189	0.995803	protein kinase, AMP-activated, gamma 1 non-catalytic subunit
PDRG1	7.117331	0.199	p53 and DNA-damage regulated 1
CBR3	7.10267	12.5975	carbonyl reductase 3
ZIK1	6.9893	1	zinc finger protein interacting with K protein 1 homolog (mouse)
CFP	6.96885	1.198397	complement factor properdin
MEF2B	6.809097	1.41993	myocyte enhancer factor 2B
FEZ1	6.74756	1	fasciculation and elongation protein zeta 1 (zygin I)
SPOCK3	6.73261	1	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3
C6orf57	6.72178	1	chromosome 6 open reading frame 57
CTF1	6.70557	8.22859	cardiotrophin 1
MIR3651	6.666231	0.001607	microRNA 3651
GHDC	6.616535	2.86206	GH3 domain containing
GPR52	6.58421	0.212243	G protein-coupled receptor 52
GDF15	6.56499	0.040631	growth differentiation factor 15
TDO2	6.55168	1	tryptophan 2,3-dioxygenase
RRAS2	6.54169	1	related RAS viral (r-ras) oncogene homolog 2
HLA-DOB	6.53871	1	major histocompatibility complex, class II, DO beta
CBX8	6.51094	0.221861	chromobox homolog 8
NBR2	6.474415	0.198047	neighbor of BRCA1 gene 2 (non-protein coding)
RARRES1	6.44963	7.580187	retinoic acid receptor responder (tazarotene induced) 1
PFDN1	6.341476	1.613132	prefoldin subunit 1
HHLA3	6.31029	0.112386	HERV-H LTR-associating 3
RASL11B	6.27618	1	RAS-like, family 11, member B
SLC29A3	6.27318	0.857594	solute carrier family 29 (nucleoside transporters), member 3
RPS15AP10	6.25755	1	ribosomal protein S15a pseudogene 10
PCGF1	6.241982	0.575734	polycomb group ring finger 1

Supplementary Table 2 The top 100 up-regulated genes in 7 days GSK-J4 treated and RANKL-stimulated MCSF-macrophages as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene Name	1d	7d	Gene Name
RNU6ATAC	1.143314	2294.42	RNA, U6atac small nuclear (U12-dependent splicing)
SCARNA11	0.019945	1048.74	small Cajal body-specific RNA 11
SCARNA22	1	861.034	small Cajal body-specific RNA 22
MIR612	322.766	612.486	microRNA 612
HIST1H4I	1.040129	188.331	histone cluster 1, H4i
BMP8B	0.681905	144.433	bone morphogenetic protein 8b
HIST1H2BH	0.56287	142.584	histone cluster 1, H2bh
HIST1H4H	0.972378	102.702	histone cluster 1, H4h
MT1G	2.184996	91.326	metallothionein 1G
NDUFS6	0.833542	87.297	NADH dehydrogenase (ubiquinone) Fe-S protein 6, 13kDa (NADH-coenzyme Q reductase)
UXT	2.805988	85.0844	ubiquitously-expressed, prefoldin-like chaperone
NDUFS5	0.511546	73.1846	NADH dehydrogenase (ubiquinone) Fe-S protein 5, 15kDa (NADH-coenzyme Q reductase)
MZT2A	1.258032	72.5933	mitotic spindle organizing protein 2A
CKS2	1.758577	69.401	CDC28 protein kinase regulatory subunit 2
CCDC12	1.031705	65.7428	coiled-coil domain containing 12
UCHL1-AS1	1	64.0669	UCHL1 antisense RNA 1 (head to head)
TOMM5	0.813152	62.3971	translocase of outer mitochondrial membrane 5 homolog (yeast)
ZNRD1	0.962677	62.2604	zinc ribbon domain containing 1
CHCHD5	1.702897	60.7025	coiled-coil-helix-coiled-coil-helix domain containing 5
PSMB6	0.365029	60.5118	proteasome (prosome, macropain) subunit, beta type, 6
PFN3	1	59.896	profilin 3
ZNF32	0.479438	59.8179	zinc finger protein 32
SPINT2	0.417819	59.288	serine peptidase inhibitor, Kunitz type, 2
HIST1H3I	0.236746	58.4472	histone cluster 1, H3i
MT2A	0.762026	57.575	metallothionein 2A
NCOR1P1	0.140514	57.575	nuclear receptor corepressor 1 pseudogene 1
AIF1	4.217094	57.2966	allograft inflammatory factor 1
HIST1H1D	1.191847	56.0344	histone cluster 1, H1d
GAMT	1.757069	53.6197	guanidinoacetate N-methyltransferase
RERG-AS1	1	53.1148	RERG antisense RNA 1
MRPL51	0.512179	52.8005	mitochondrial ribosomal protein L51
S100P	1	52.3084	S100 calcium binding protein P
UBE2MP1	0.769947	51.9752	ubiquitin-conjugating enzyme E2M pseudogene 1
MIR7703	0.888413	51.476	microRNA 7703
DNAJC9	0.75424	50.8608	DnaJ (Hsp40) homolog, subfamily C, member 9
POU3F3	1.969563	49.6853	POU class 3 homeobox 3
SUSD3	0.39655	49.4913	sushi domain containing 3
TIMM10	0.870564	49.4057	translocase of inner mitochondrial membrane 10 homolog (yeast)
CCL2	1.271029	49.3827	chemokine (C-C motif) ligand 2
LSM3	1.540181	47.3472	LSM3 homolog, U6 small nuclear RNA associated (S. cerevisiae)
SERPINF1	0.422418	47.2667	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1
FKBP3	0.052464	47.0947	FK506 binding protein 3, 25kDa
RPS18P9	0.59514	46.8583	ribosomal protein S18 pseudogene 9
MT1X	4.448769	44.3905	metallothionein 1X
CARD16	0.052525	43.7818	caspase recruitment domain family, member 16
TSEN15	1.570057	42.429	tRNA splicing endonuclease 15 homolog (S. cerevisiae)
TMEM55B	1.465661	42.1479	transmembrane protein 55B
LLGL2	9.16949	41.9121	lethal giant larvae homolog 2 (Drosophila)
GPR162	0.703618	41.2311	G protein-coupled receptor 162
SUMO1	0.908713	40.41	SMT3 suppressor of mif two 3 homolog 1 (S. cerevisiae)
HSD17B14	5.47184	40.0922	hydroxysteroid (17-beta) dehydrogenase 14
CCDC37-AS1	1	38.6706	CCDC37 antisense RNA 1 (head to head)
TSSK6	1.746606	37.8902	testis-specific serine kinase 6
IGLL1	2.961651	37.2488	immunoglobulin lambda-like polypeptide 1
HIST1H2BG	0.642189	37.1039	histone cluster 1, H2bg
CCDC124	0.691263	36.846	coiled-coil domain containing 124
E2F5	0.41784	36.6224	E2F transcription factor 5, p130-binding
ANXA2P1	0.460471	36.4631	annexin A2 pseudogene 1
IMP4	1.092502	36.3014	IMP4, U3 small nucleolar ribonucleoprotein, homolog (yeast)
TCEAL8	0.29642	34.7907	transcription elongation factor A (SII)-like 8
YAE1D1	0.317477	33.798	Yae1 domain containing 1
PQLC3	17.5142	33.5615	PQ loop repeat containing 3

HBG1	1	33.1924	hemoglobin, gamma A
SIGLEC7	1.191472	32.7881	sialic acid binding Ig-like lectin 7
CCNL2	0.308107	32.5032	cyclin L2
HIST1H2BJ	0.653441	32.4589	histone cluster 1, H2bj
SCN1B	0.505318	31.9458	sodium channel, voltage-gated, type I, beta subunit
NUTF2	0.953482	31.8498	nuclear transport factor 2
GGCT	1.201962	31.7599	gamma-glutamylcyclotransferase
ZSWIM1	0.821772	31.6379	zinc finger, SWIM-type containing 1
DYNLL1	1.187599	31.4218	dynein, light chain, LC8-type 1
RBP1	3.16143	31.3192	retinol binding protein 1, cellular
ZNF394	2.132797	30.8885	zinc finger protein 394
LST1	2.618818	30.8547	leukocyte specific transcript 1
HSCB	0.284605	30.7785	HscB iron-sulfur cluster co-chaperone homolog (E. coli)
VAV3-AS1	1	30.4957	VAV3 antisense RNA 1
NFU1	0.518189	30.2505	NFU1 iron-sulfur cluster scaffold homolog (S. cerevisiae)
INSL3	1	30.1567	insulin-like 3 (Leydig cell)
C19orf60	1.957149	29.7284	chromosome 19 open reading frame 60
IRF9	1.135919	29.6108	interferon regulatory factor 9
CAMK1D	1.115333	29.4958	calcium/calmodulin-dependent protein kinase ID
HIST1H2BN	1.172763	29.4169	histone cluster 1, H2bn
RPRML	1	28.6374	reprimo-like
TMEM101	0.638335	28.6355	transmembrane protein 101
OIP5-AS1	1.272038	28.0401	OIP5 antisense RNA 1
CDK2AP1	2.416171	27.9943	cyclin-dependent kinase 2 associated protein 1
POLG2	1.229472	27.5442	polymerase (DNA directed), gamma 2, accessory subunit
RNF114	0.83842	27.4179	ring finger protein 114
FTSJ1	1.01107	27.0844	FtsJ RNA methyltransferase homolog 1 (E. coli)
COPS5	0.653443	26.9743	COP9 constitutive photomorphogenic homolog subunit 5 (Arabidopsis)
MAP2K5	1.23448	26.5986	mitogen-activated protein kinase kinase 5
COQ5	0.513563	26.5771	coenzyme Q5 homolog, methyltransferase (S. cerevisiae)
POLR1C	0.300391	26.1771	polymerase (RNA) I polypeptide C, 30kDa
IAH1	0.368299	25.9787	isoamyl acetate-hydrolyzing esterase 1 homolog (S. cerevisiae)
DEFB129	1	25.8055	defensin, beta 129
SNRNP40	0.776661	25.7472	small nuclear ribonucleoprotein 40kDa (U5)
MILR1	1.027209	25.5762	mast cell immunoglobulin-like receptor 1
ZNF688	0.591962	25.4045	zinc finger protein 688
MIR133A2	1	25.1923	microRNA133A2
OR2J3	1	25.1772	olfactory receptor, family 2, subfamily J, member 3
NDUFAF4	0.458762	24.9007	NADH dehydrogenase (ubiquinone) complex I, assembly factor 4

Supplementary Table 3 The top 100 down-regulated genes in 24h GSK-J4 treated and RANKL-stimulated MCSF-macrophages as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene Name	1d	7d	Gene Name
MIR320A	0.000421	1	microRNA 320a
MIR3687	0.000465	0.000114	microRNA 3687
MIR4750	0.000742	1	microRNA 4750
MIR3689C	0.001457	1	microRNA 3689c
MIR4792	0.001573	1	microRNA 4792
MIR590	0.002298	1	microRNA 590
VTRNA1-2	0.005148	1.119723	vault RNA 1-2
TPRA1	0.005347	0.797437	transmembrane protein, adipocyte asscociated 1
MIR589	0.007294	1	microRNA 589
MIR6892	0.011516	1	microRNA 6892
MIR3655	0.015654	12.1695	microRNA 3655
NDUFA3	0.016223	0.031515	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa
SCARNA9L	0.019225	1	small Cajal body-specific RNA 9-like
RNVU1-19	0.019428	1	RNA, variant U1 small nuclear 19
SCARNA11	0.019945	1048.74	small Cajal body-specific RNA 11
INHBB	0.024049	1	inhibin, beta B
PRKCZ	0.028723	1	protein kinase C, zeta
RNVU1-19	0.037404	1	RNA, variant U1 small nuclear 19
MPC1	0.044837	0.782235	mitochondrial pyruvate carrier 1
IFITM3	0.045024	12.4327	interferon induced transmembrane protein 3
IFI27L1	0.045573	1	interferon, alpha-inducible protein 27-like 1
C8orf88	0.045591	1	chromosome 8 open reading frame 88
CRNDE	0.046243	0.120927	colorectal neoplasia differentially expressed (non-protein coding)
SH2D2A	0.051268	1	SH2 domain containing 2A
AUH	0.051727	1.104417	AU RNA binding protein/enoyl-CoA hydratase
FKBP3	0.052464	47.0947	FK506 binding protein 3, 25kDa
NCR1	0.052496	1	natural cytotoxicity triggering receptor 1
CARD16	0.052525	43.7818	caspase recruitment domain family, member 16
RPP25L	0.052887	0.046934	ribonuclease P/MRP 25kDa subunit-like
TMEM238	0.057378	1	transmembrane protein 238
HERC2P10	0.057773	1	hect domain and RLD 2 pseudogene 10
MFI2-AS1	0.063047	0.137165	MFI2 antisense RNA 1
IFT20	0.067444	1	intraflagellar transport 20 homolog (Chlamydomonas)
C1QB	0.067786	9.3555	complement component 1, q subcomponent, B chain
RNF113A	0.069305	1	ring finger protein 113A
EFNA4	0.071296	1	ephrin-A4
RNF11	0.071927	0.285122	ring finger protein 11
CYTL1	0.072281	1	cytokine-like 1
GPN3	0.073989	1	GPN-loop GTPase 3
ZNF692	0.074493	1	zinc finger protein 692
PIGC	0.074935	1	phosphatidylinositol glycan anchor biosynthesis, class C
ZBTB8OS	0.076123	1	zinc finger and BTB domain containing 8 opposite strand
TMEM170A	0.076787	1	transmembrane protein 170A
DDC	0.078151	1	dopa decarboxylase (aromatic L-amino acid decarboxylase)
SPAG7	0.082054	0.148689	sperm associated antigen 7
PXMP2	0.082591	0.845995	peroxisomal membrane protein 2, 22kDa
C16orf91	0.084016	1	chromosome 16 open reading frame 91
FBXO22-AS1	0.084723	1	FBXO22 antisense RNA 1
IL15RA	0.086754	0.124968	interleukin 15 receptor, alpha
COMMD6	0.089497	1	COMM domain containing 6
EFNA3	0.089619	2.11369	ephrin-A3
NSDHL	0.089664	1	NAD(P) dependent steroid dehydrogenase-like
RPARP-AS1	0.09024	1	RPARP antisense RNA 1
MLLT10P1	0.090402	1	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 10 pseudogene 1
FAM89A	0.094075	1	family with sequence similarity 89, member A
DYRK3	0.097045	0.172458	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 3
RAMP2	0.098598	1	receptor (G protein-coupled) activity modifying protein 2
C4orf33	0.098642	1	chromosome 4 open reading frame 33
PDE9A	0.099602	1	phosphodiesterase 9A
UNC50	0.099958	1	unc-50 homolog (C. elegans)
UBXN8	0.100503	1	UBX domain protein 8
TNFRSF12A	0.104329	0.252464	tumor necrosis factor receptor superfamily, member 12A
C20orf196	0.106223	1	chromosome 20 open reading frame 196
CCL24	0.106991	23.0973	chemokine (C-C motif) ligand 24

FLJ27354	0.107093	1	uncharacterized LOC400761
EP400NL	0.107127	1	EP400 N-terminal like
KBTBD4	0.107285	10.61036	kelch repeat and BTB (POZ) domain containing 4
RETN	0.107555	1	resistin
REEP4	0.107726	3.681468	receptor accessory protein 4
TRAPPC2L	0.109201	0.043384	trafficking protein particle complex 2-like
IFI27	0.109683	1	interferon, alpha-inducible protein 27
FAM222A-AS1	0.11001	1	FAM222A antisense RNA 1
SLC25A37	0.112248	1	solute carrier family 25 (mitochondrial iron transporter), member 37
C21orf119	0.112901	1	chromosome 21 open reading frame 119
ZNF295-AS1	0.11423	1	ZNF295 antisense RNA 1
COX17	0.114895	0.044917	cytochrome c oxidase assembly homolog 17 (yeast)
HOXB7	0.115511	1	homeobox B7
TSR2	0.117196	5.69553	TSR2, 20S rRNA accumulation, homolog (S. cerevisiae)
C16orf59	0.117611	0.13138	chromosome 16 open reading frame 59
TMEM55A	0.117684	5.153142	transmembrane protein 55A
VIM-AS1	0.118172	9.41448	VIM antisense RNA 1
CFD	0.11837	5.852643	complement factor D (adipsin)
METTL24	0.118594	1	methyltransferase like 24
GPN2	0.119511	1.199258	GPN-loop GTPase 2
RMI2	0.119707	1	RMI2, RecQ mediated genome instability 2, homolog (S. cerevisiae)
ITGA1	0.121199	0.493454	integrin, alpha 1
ZFAND2B	0.122223	1.793	zinc finger, AN1-type domain 2B
GPBAR1	0.122314	14.4184	G protein-coupled bile acid receptor 1
LYRM1	0.122965	1	LYR motif containing 1
TAS2R14	0.123381	0.137638	taste receptor, type 2, member 14
LEAP2	0.124306	1	liver expressed antimicrobial peptide 2
ZNHIT2	0.126194	0.188387	zinc finger, HIT-type containing 2
IGSF6	0.126965	0.082016	immunoglobulin superfamily, member 6
PDK2	0.127095	1.449343	pyruvate dehydrogenase kinase, isozyme 2
RPL22L1	0.127544	0.157297	ribosomal protein L22-like 1
HIST2H4B	0.127681	1	histone cluster 2, H4b
TIPIN	0.127754	1	TIMELESS interacting protein
OASL	0.128313	0.523994	2'-5'-oligoadenylate synthetase-like
CCRN4L	0.129223	1	CCR4 carbon catabolite repression 4-like (S. cerevisiae)
ST13P4	0.129775	2.26039	suppression of tumorigenicity 13 (colon carcinoma) (Hsp70 interacting protein) pseudogene 4

Supplementary Table 4 The top 100 down-regulated genes in 7 days GSK-J4 treated and RANKL-stimulated MCSF-macrophages as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene Name	1d	7d	Gene Name
MIR3687	0.000465	0.000114	microRNA 3687
SCARNA8	0.566475	0.000764	small Cajal body-specific RNA 8
MIR3651	6.666231	0.001607	microRNA 3651
SCARNA10	0.676604	0.003062	small Cajal body-specific RNA 10
C4orf48	8.416332	0.004448	chromosome 4 open reading frame 48
CSTB	0.978447	0.005563	cystatin B (stefin B)
ROMO1	4.627435	0.006083	reactive oxygen species modulator 1
GAL	13.8485	0.008087	galanin/GMAP prepropeptide
ATP6V1E1	2.39283	0.01112	ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E1
RABAC1	0.488341	0.011542	Rab acceptor 1 (prenylated)
POLR1D	0.781477	0.011694	polymerase (RNA) I polypeptide D, 16kDa
HIST1H4C	0.829194	0.013177	histone cluster 1, H4c
SCARNA7	0.366879	0.013631	small Cajal body-specific RNA 7
TTTY20	1	0.014249	testis-specific transcript, Y-linked 20 (non-protein coding)
C12orf57	0.571159	0.014287	chromosome 12 open reading frame 57
MIR611	0.895282	0.014613	microRNA 611
PQLC2	1.677432	0.015054	PQ loop repeat containing 2
ATP5O	0.637279	0.015776	ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit
MEMO1	1.220683	0.016914	Methylation modifier for class I HLA
ARF5	0.858942	0.016954	ADP-ribosylation factor 5
TERC	0.557053	0.017348	telomerase RNA component
HIST1H4F	3.204602	0.017415	histone cluster 1, H4f
NDUFS3	1.443254	0.017539	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase)
GSTK1	1.062525	0.018017	glutathione S-transferase kappa 1
APOC1	1	0.01803	apolipoprotein C-I
PRKCH	2.68191	0.018105	protein kinase C, eta
DPM3	0.791776	0.019131	dolichyl-phosphate mannosyltransferase polypeptide 3
IGBP1P1	0.456494	0.019428	immunoglobulin (CD79A) binding protein 1 pseudogene 1
SNRNP25	0.659905	0.020418	small nuclear ribonucleoprotein 25kDa (U11/U12)
SLIRP	0.713106	0.020853	SRA stem-loop interacting RNA binding protein
PRKCDBP	0.272717	0.02167	protein kinase C, delta binding protein
S100A10	1.38027	0.022127	S100 calcium binding protein A10
ANG	0.276181	0.022418	angiogenin, ribonuclease, RNase A family, 5
RPL39	0.266554	0.02257	ribosomal protein L39
TREM2	1.054222	0.022784	triggering receptor expressed on myeloid cells 2
ZDHHC4	0.543438	0.023393	zinc finger, DHHC-type containing 4
SSSCA1	0.235169	0.023438	Sjogren syndrome/scleroderma autoantigen 1
RGCC	2.586381	0.023468	regulator of cell cycle
KRTCAP2	1.506594	0.023856	keratinocyte associated protein 2
ZNF410	0.41599	0.024108	zinc finger protein 410
ZNF768	2.128359	0.024293	zinc finger protein 768
UBL7	1.458741	0.02559	ubiquitin-like 7 (bone marrow stromal cell-derived)
CYB5A	1.191144	0.026293	cytochrome b5 type A (microsomal)
CCDC58	1.089176	0.026662	coiled-coil domain containing 58
C12orf75	0.379959	0.026707	chromosome 12 open reading frame 75
MAP1LC3A	1.59308	0.02695	microtubule-associated protein 1 light chain 3 alpha
NUPR1	1	0.027096	nuclear protein, transcriptional regulator, 1
CRCP	1.48911	0.027144	CGRP receptor component
FAM63A	3.021845	0.027264	family with sequence similarity 63, member A
DYNLT1	0.209509	0.027265	dynein, light chain, Tctex-type 1
SNHG12	1.070034	0.027716	small nucleolar RNA host gene 12 (non-protein coding)
ACYP1	0.301495	0.02788	acylphosphatase 1, erythrocyte (common) type
TUBA8	0.680075	0.02788	tubulin, alpha 8
MDP1	1.646899	0.028332	magnesium-dependent phosphatase 1
MVB12A	2.420998	0.029164	multivesicular body subunit 12A
EMC7	0.663534	0.02922	ER membrane protein complex subunit 7
PDE8A	0.222864	0.029448	phosphodiesterase 8A
EAPP	2.117402	0.029496	E2F-associated phosphoprotein
NSMCE1	0.561667	0.029519	non-SMC element 1 homolog (S. cerevisiae)
DEXI	3.493591	0.02971	Dexi homolog (mouse)
DNAJC1	1.018144	0.029899	DnaJ (Hsp40) homolog, subfamily C, member 1
BACE1-AS	1	0.030564	BACE1 antisense RNA
MED7	3.313185	0.030782	mediator complex subunit 7
NAPA-AS1	21.3687	0.031141	NAPA antisense RNA 1

CD52	0.405852	0.031201	CD52 molecule
BLOC1S1	1.46942	0.031202	biogenesis of lysosomal organelles complex-1, subunit 1
DPPA5	1	0.031407	developmental pluripotency associated 5
NDUFA3	0.016223	0.031515	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa
MRPL49	0.685071	0.032095	mitochondrial ribosomal protein L49
MRPL28	0.358801	0.032108	mitochondrial ribosomal protein L28
NGFRAP1	1	0.032129	nerve growth factor receptor (TNFRSF16) associated protein 1
C11orf84	0.779758	0.032806	chromosome 11 open reading frame 84
PSMB10	0.509874	0.033	proteasome (prosome, macropain) subunit, beta type, 10
LSM10	0.829099	0.033246	LSM10, U7 small nuclear RNA associated
CINP	0.875066	0.033292	cyclin-dependent kinase 2 interacting protein
TRPT1	2.384439	0.033295	tRNA phosphotransferase 1
PPP2R3C	1.119528	0.033459	protein phosphatase 2, regulatory subunit B'', gamma
DYNLRB1	0.753669	0.033793	dynein, light chain, roadblock-type 1
NDUFB10	0.460667	0.033854	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22kDa
CUEDC2	2.052962	0.033959	CUE domain containing 2
TRAPPC6A	1.786034	0.034118	trafficking protein particle complex 6A
ATP5EP2	0.634013	0.034234	ATP synthase, H+ transporting, mitochondrial F1 complex, epsilon subunit pseudogene 2
UPP1	3.725688	0.034868	uridine phosphorylase 1
HNRNPA1P33	1	0.034869	heterogeneous nuclear ribonucleoprotein A1 pseudogene 33
HSPB1	8.98833	0.035569	heat shock 27kDa protein 1
RPS4Y1	8.98833	0.035569	ribosomal protein S4, Y-linked 1
GH1	1	0.035791	growth hormone 1
ZNF511	1.310585	0.035925	zinc finger protein 511
ALOX5AP	1	0.036012	arachidonate 5-lipoxygenase-activating protein
RPP21	0.350798	0.036253	ribonuclease P/MRP 21kDa subunit
SNX33	2.128941	0.036433	sorting nexin 33
CDC42EP5	1	0.036724	CDC42 effector protein (Rho GTPase binding) 5
RPS3A	0.787374	0.037389	ribosomal protein S3A
OR2M7	1	0.037635	olfactory receptor, family 2, subfamily M, member 7
OR4L1	1	0.037635	olfactory receptor, family 4, subfamily L, member 1
RPL36AL	0.569196	0.038313	ribosomal protein L36a-like
SDR39U1	1.758215	0.038365	short chain dehydrogenase/reductase family 39U, member 1
EXOSC1	0.738894	0.038676	exosome component 1
PLEKHG3	1.441267	0.03889	pleckstrin homology domain containing, family G (with RhoGef domain) member 3
TGIF1	3.579919	0.039141	TGFB-induced factor homeobox 1
TOR2A	0.873731	0.039388	torsin family 2, member A
S100A1	1	0.039976	S100 calcium binding protein A1

Supplementary Table 5 RPKM values of detectable microRNA genes on day 8 of treatments.

Gene Name	1d MCSF+DMSO	1d RANKL+DMSO	1d RANKL+J4
MIR1248	104697	21483.5	40924.7
MIR1256,SLC25A53	0	2.29714	1.9996
MIR1273H,WDR19	0	1.85995	0.858049
MIR1292,NOP56	7.74708	157.808	144.518
MIR1307,USMG5	0	104.677	30.6057
MIR155HG	0	6.30719	0
MIR17HG	0	6.85609	11.4849
MIR181A1HG	0	27.6283	22.9981
MIR1909,REXO1	22.4336	26.4919	23.4783
MIR210HG	0	1.2736	0
MIR22HG	0	12.7975	4.49954
MIR3134,PTBP3	9.6703	28.8329	25.9398
MIR320A	0	2372.07	0
MIR3605,PHC2	93.0994	31.5656	49.7711
MIR3648	33576	42793.5	45971.2
MIR3651	0	54.0686	366.1
MIR3656,TRAPPC4	6.01673	15.163	22.2377
MIR3687	0	2150.34	0
MIR3689C	0	685.121	0
MIR3917,STMN1	13.6731	150.516	126.717
MIR4315-1,PLEKHM1P	5.8959	4.42785	4.08583
MIR4435-1HG	0	12.6723	37.9433
MIR4458HG	0	0.879879	1.82673
MIR4469,RNF170	3.33679	3.87934	3.58964
MIR4500HG	0	0	0.628573
MIR4697HG	0	0.349696	0
MIR4721,TUFM	28.955	99.4915	109.99
MIR4750	0	1346.68	0
MIR4792	0	634.746	0
MIR5001,TIGD1	0.892466	4.56722	2.88622
MIR5004,SYNGAP1	0	1.53661	2.39223
MIR5006,VWA8	0.86448	9.32521	5.59145
MIR5047	0	0	2574.13
MIR5187,TOMM40L	14.6103	14.9289	15.0422
MIR5193,UBA7	3.22192	1.00377	2.08393
MIR548N,OSBPL6	0	2.51757	2.27787
MIR564,TMEM42	5.20958	0	7.22044
MIR589	1441.01	136.108	0
MIR590	0	434.227	0
MIR600HG	0	3.70445	2.40298
MIR611,TMEM258	9.55553	88.8748	79.4633
MIR612	0	0	321.766
MIR632,ZNF207	23.7205	82.0099	61.3461
MIR636,SRSF2	42.1042	169.705	122.117
MIR639,TECR	6.23443	19.8714	44.1684
MIR6511B1,PKD1	9.92873	7.84732	9.37353
MIR6513,TMBIM1	112.966	19.7337	30.1925
MIR663A	57822.9	90780.9	74799.3
MIR6716,PHLDB1	1.96432	0.961746	0
MIR6740	0	0	212.717
MIR6789,PLEKHJ1	31.6461	15.8335	4.56888
MIR6838,POLM	3.30348	7.68083	7.24973
MIR6852,TLN1	263.048	94.1125	107.687
MIR6892	0	85.8325	0
MIR7109,PISD	5.66485	11.8847	12.338
MIR761,NRD1	11.6362	48.9546	43.7753
MIR7703,PSME2	3.19394	38.3271	33.9387
MIR7846,TNFRSF1B	102.748	62.6873	68.8037
MIR7851,UBQLN4	12.738	17.4129	14.8524
MIR937,SCRIB	17.8533	52.3777	34.4927
MIR943,NELFA	0	24.4061	23.7034

Supplementary Table 6 RPKM values of detectable microRNA genes on day 14 of treatments.

Gene Name	7d MCSF+DMSO	7d RANKL+DMSO	7d RANKL+J4
MIR100HG	0	1.63728	0
MIR1248	127132	60648.7	53335.4
MIR1256,SLC25A53	0.900386	0	0
MIR1273H,WDR19	2.92982	0	0
MIR1287,PYROXD2	0	2.53067	0
MIR1292,NOP56	27.4257	13.0164	6.80177
MIR1307,USMG5	8.69655	0	0
MIR133A1HG	0	1.16958	0
MIR137HG	0	2.07984	0.776313
MIR17HG	0	0	0.365498
MIR1909,REXO1	19.8104	26.9229	28.1376
MIR210HG	0	2.24411	0
MIR22HG	3.14398	0	10.697
MIR3134,PTBP3	14.4904	11.3738	2.12303
MIR31HG	0	0	4.52158
MIR3605,PHC2	35.9524	82.6456	65.5817
MIR3648	42815.4	86942.5	127413
MIR3651	952.462	621.323	0
MIR3656,TRAPPC4	5.42386	4.71757	10.5652
MIR3687	3350.79	8743.36	0
MIR3917,STMN1	16.6745	3.62517	0
MIR4315-1,PLEKHM1P	9.30114	1.73356	25.2354
MIR4435-1HG	0	13.799	0
MIR4469,RNF170	5.06891	2.64603	0
MIR4500HG	0	3.20033	0
MIR4721,TUFM	22.235	25.2255	9.41555
MIR497HG	3.04958	0	0
MIR5004,SYNGAP1	0.933557	0.811989	0
MIR5006,VWA8	4.10793	0.677801	2.17429
MIR5187,TOMM40L	2.87495	9.37762	2.1
MIR5193,UBA7	489.388	7.57869	11.8809
MIR548AA1,PALMD	0	1.98446	0
MIR548N,OSBPL6	1.09389	1.42709	0
MIR564,TMEM42	7.04437	0	0
MIR611,TMEM258	51.6838	67.4303	0
MIR612	0	0	611.486
MIR632,ZNF207	19.601	4.64951	4.33841
MIR636,SRSF2	11.2539	2.66943	29.8819
MIR639,TECR	11.2407	0	20.0721
MIR6511B1,PKD1	12.1933	17.2621	14.7814
MIR6513,TMBIM1	62.2791	72.4696	91.8049
MIR663A	45709.9	92070	99972.9
MIR6716,PHLDB1	0.708235	4.62065	4.13923
MIR6717,NDRG2	0	0	6.0127
MIR6789,PLEKHJ1	7.13195	4.65242	27.7847
MIR6838,POLM	4.1478	3.60767	11.25
MIR6852,TLN1	223.515	236.613	239.077
MIR7109,PISD	30.6399	26.65	36.2365
MIR761,NRD1	14.3681	18.2473	27.3561
MIR7703,PSME2	5.75845	0	50.476
MIR7846,TNFRSF1B	114.999	126.274	143.425
MIR7851,UBQLN4	6.0148	8.56072	6.92333
MIR937,SCRIB	10.9732	29.1457	21.7312
MIR943,NELFA	3.18971	14.5653	3.10662

Supplementary Table 7 List of genes that fall under the GO annotation category of "chromosome organisation" within the osteoclast samples treated with RANKL for 1 day.

ID	Gene Name
ASF1B	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)
BRCC3	BRCA1/BRCA2-containing complex, subunit 3
BLM	Bloom syndrome, RecQ helicase-like
CTCF	CCCTC-binding factor (zinc finger protein)
CSRP2BP	CSRP2 binding protein
DDX11	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, S. cerevisiae)
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta
FBXO4	F-box protein 4
FANCD2	Fanconi anemia, complementation group D2
H2AFX	H2A histone family, member X
H2AFZ	H2A histone family, member Z
KAT5	K(lysine) acetyltransferase 5
LEO1	Leo1, Paf1/RNA polymerase II complex component, homolog (S. cerevisiae)
MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)
NDC80	NDC80 homolog, kinetochore complex component (S. cerevisiae)
NEK2	NIMA (never in mitosis gene a)-related kinase 2
PDS5B	PDS5, regulator of cohesion maintenance, homolog B (S. cerevisiae)
PAF1	Paf1, RNA polymerase II associated factor, homolog (S. cerevisiae)
RAD50	RAD50 homolog (S. cerevisiae)
RAD51C	RAD51 homolog C (S. cerevisiae)
RAD54L	RAD54-like (S. cerevisiae)
RTF1	Rtf1, Paf1/RNA polymerase II complex component, homolog (S. cerevisiae)
RUVBL2	RuvB-like 2 (E. coli)
SMYD3	SET and MYND domain containing 3
SETDB1	SET domain, bifurcated 1
SHPRH	SNF2 histone linker PHD RING helicase
SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5
SMARCE1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1
SMARCAD1	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1
TAF12	TAF12 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 20kDa
TAF5	TAF5 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 100kDa
TAF9	TAF9 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 32kDa
WHSC1L1	Wolf-Hirschhorn syndrome candidate 1-like 1
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining)
YEATS4	YEATS domain containing 4
ZWINT	ZW10 interactor
ACTL6A	actin-like 6A
ACD	adrenocortical dysplasia homolog (mouse)
AIFM1	apoptosis-inducing factor, mitochondrion-associated, 1
AIFM2	apoptosis-inducing factor, mitochondrion-associated, 2
BPTF	bromodomain PHD finger transcription factor
BRD8	bromodomain containing 8
BUB3	budding uninhibited by benzimidazoles 3 homolog (yeast)
CDC23	cell division cycle 23 homolog (S. cerevisiae)
CENPA	centromere protein A
CENPF	centromere protein F, 350/400ka (mitosin)
CHAF1B	chromatin assembly factor 1, subunit B (p60)
CBX1	chromobox homolog 1 (HP1 beta homolog Drosophila)
CBX2	chromobox homolog 2 (Pc class homolog, Drosophila)
CBX5	chromobox homolog 5 (HP1 alpha homolog, Drosophila)
CBX7	chromobox homolog 7
CHD1L	chromodomain helicase DNA binding protein 1-like
CHD9	chromodomain helicase DNA binding protein 9
MEAF6	chromosome 1 open reading frame 149
DSCC1	defective in sister chromatid cohesion 1 homolog (S. cerevisiae)
DLGAP5	discs, large (Drosophila) homolog-associated protein 5
DKC1	dyskeratosis congenita 1, dyskerin
EED	embryonic ectoderm development
EZH1	enhancer of zeste homolog 1 (Drosophila)
EZH2	enhancer of zeste homolog 2 (Drosophila)
ERCC4	excision repair cross-complementing rodent repair deficiency, complementation group 4
GSG2	germ cell associated 2 (haspin)
HELLS	helicase, lymphoid-specific

HLTF	helicase-like transcription factor
HP1BP3	heterochromatin protein 1, binding protein 3
H3F3C	histone H3-like
HIST1H1D	histone cluster 1, H1d
HIST1H1E	histone cluster 1, H1e
HIST1H2AC	histone cluster 1, H2ac
HIST1H2AB	histone cluster 1, H2ae; histone cluster 1, H2ab
HIST1H2AL, HIST1H2AK, HIST1H2AH, HIST1H2AI	histone cluster 1, H2ag; histone cluster 1, H2ah; histone cluster 1, H2ai; histone cluster 1, H2ak; histone cluster 1, H2al; histone cluster 1, H2am
HIST1H2AJ	histone cluster 1, H2aj
HIST1H2BB	histone cluster 1, H2bb
HIST1H2BD	histone cluster 1, H2bd
HIST1H2BH	histone cluster 1, H2bh
HIST1H2BF, HIST1H2BG, HIST1H2BI	histone cluster 1, H2bi; histone cluster 1, H2bg; histone cluster 1, H2be; histone cluster 1, H2bf; histone cluster 1, H2bc
HIST1H2BJ	histone cluster 1, H2bj
HIST1H2BL	histone cluster 1, H2bl
HIST1H2BM	histone cluster 1, H2bm
HIST1H2BN	histone cluster 1, H2bn
HIST1H2BO	histone cluster 1, H2bo
HIST1H3D, HIST1H3H, HIST2H3D, HIST1H2AD, HIST1H3F, HIST1H3B	histone cluster 1, H3j; histone cluster 1, H3i; histone cluster 1, H3h; histone cluster 1, H3g; histone cluster 1, H3f; histone cluster 1, H3e; histone cluster 1, H3d; histone cluster 1, H3c; histone cluster 1, H3b; histone cluster 1, H3a; histone cluster 1, H2ad; histone cluster 2, H3a; histone cluster 2, H3c; histone cluster 2, H3d
HIST1H4I, HIST1H4A, HIST1H4H, HIST1H4C, HIST1H4K, HIST1H4F	histone cluster 1, H4l; histone cluster 1, H4k; histone cluster 4, H4; histone cluster 1, H4h; histone cluster 1, H4j; histone cluster 1, H4i; histone cluster 1, H4d; histone cluster 1, H4c; histone cluster 1, H4f; histone cluster 1, H4e; histone cluster 1, H4b; histone cluster 1, H4a; histone cluster 2, H4a; histone cluster 2, H4b
HIST2H2AC	histone cluster 2, H2ac
HIST2H2BE	histone cluster 2, H2be
HDAC2	histone deacetylase 2
HDAC8	histone deacetylase 8
RBBP4	hypothetical LOC642954; retinoblastoma binding protein 4
ING4	inhibitor of growth family, member 4
IRF4	interferon regulatory factor 4
KIF18A	kinesin family member 18A
LIG4	ligase IV, DNA, ATP-dependent
KDM4A	lysine (K)-specific demethylase 4A
KDM4C	lysine (K)-specific demethylase 4C
MCM2	minichromosome maintenance complex component 2
MORF4L2	mortality factor 4 like 2
MSH3	mutS homolog 3 (E. coli)
NBN	nibrin
NCAPG	non-SMC condensin I complex, subunit G
NCAPH	non-SMC condensin I complex, subunit H
NCAPG2	non-SMC condensin II complex, subunit G2
NASP	nuclear autoantigenic sperm protein (histone-binding)
NUSAP1	nucleolar and spindle associated protein 1
NAP1L1	nucleosome assembly protein 1-like 1
NAP1L2	nucleosome assembly protein 1-like 2
PBRM1	polybromo 1
PCGF2	polycomb group ring finger 2
PAPD7	polymerase (DNA directed) sigma
PHB	prohibitin
PRMT6	protein arginine methyltransferase 6
PRMT7	protein arginine methyltransferase 7
RCBTB1	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 1
RFC1	replication factor C (activator 1) 1, 145kDa
RPA1	replication protein A1, 70kDa
RB1	retinoblastoma 1
RBBP7	retinoblastoma binding protein 7
RBL2	retinoblastoma-like 2 (p130)
SAFB	scaffold attachment factor B
KDM5B	similar to Jumonji, AT rich interactive domain 1B (RBP2-like); lysine (K)-specific demethylase 5B
WRN	similar to Werner syndrome protein; Werner syndrome, RecQ helicase-like
CBX3	similar to chromobox homolog 3; chromobox homolog 3 (HP1 gamma homolog, Drosophila)
PRKDC	similar to protein kinase, DNA-activated, catalytic polypeptide; protein kinase, DNA-activated, catalytic polypeptide
SIRT1	sirtuin (silent mating type information regulation 2 homolog) 1 (S. cerevisiae)
SMC1A	structural maintenance of chromosomes 1A
SMC2	structural maintenance of chromosomes 2

SMC4	structural maintenance of chromosomes 4
SMCHD1	structural maintenance of chromosomes flexible hinge domain containing 1
SUPT4H1	suppressor of Ty 4 homolog 1 (S. cerevisiae)
SUPT6H	suppressor of Ty 6 homolog (S. cerevisiae)
SUV39H1	suppressor of variegation 3-9 homolog 1 (Drosophila)
SUV420H2	suppressor of variegation 4-20 homolog 2 (Drosophila)
SUZ12	suppressor of zeste 12 homolog (Drosophila)
TERF2	telomeric repeat binding factor 2
TERF2IP	telomeric repeat binding factor 2, interacting protein
TOP2A	topoisomerase (DNA) II alpha 170kDa
TBL1XR1	transducin (beta)-like 1 X-linked receptor 1
TRIM16	tripartite motif-containing 16
UIMC1	ubiquitin interaction motif containing 1
USP16	ubiquitin specific peptidase 16
UBE2N	ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)
VPS72	vacuolar protein sorting 72 homolog (S. cerevisiae)

Supplementary Table 8 List of genes that fall under the GO annotation category of "cell adhesion" within the osteoclast samples treated with RANKL for 7 days.

ID	Gene Name
ADAM17	ADAM metallopeptidase domain 17
ADAM2	ADAM metallopeptidase domain 2
ADAM8	ADAM metallopeptidase domain 8
ALX1	ALX homeobox 1
BCL2	B-cell CLL/lymphoma 2
CLEC4A	C-type lectin domain family 4, member A
CD22	CD22 molecule
CD33	CD33 molecule
CD72	CD72 molecule
CD84	CD84 molecule
FAT4	FAT tumor suppressor homolog 4 (Drosophila)
LYPD3	LY6/PLAUR domain containing 3
RAPH1	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1
SLAMF7	SLAM family member 7
THY1	Thy-1 cell surface antigen
ACTN1	actinin, alpha 1
AMIGO3	adhesion molecule with Ig-like domain 3
APOA4	apolipoprotein A-IV
ARVCF	armadillo repeat gene deletes in velocardiofacial syndrome
BAI1	brain-specific angiogenesis inhibitor 1
CELSR1	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, Drosophila)
CDH23	cadherin-like 23
CDH24	cadherin-like 24
CHST10	carbohydrate sulfotransferase 10
CADM1	cell adhesion molecule 1
CERCAM	cerebral endothelial cell adhesion molecule
CCL4	chemokine (C-C motif) ligand 4
CCL5	chemokine (C-C motif) ligand 5
CX3CR1	chemokine (C-X3-C motif) receptor 1
CLDN14	claudin 14
CLDN7	claudin 7
COL8A2	collagen, type VIII, alpha 2
CYR61	cysteine-rich, angiogenic inducer, 61
DCHS1	dachsous 1 (Drosophila)
DSC2	desmocollin 2
DDR2	discoidin domain receptor tyrosine kinase 2
DLG5	discs, large homolog 5 (Drosophila)
ENG	endoglin
FN1	fibronectin 1
FNDC3A	fibronectin type III domain containing 3A
HAPLN4	hyaluronan and proteoglycan link protein 4
ALCAM	hypothetical protein LOC100133690; activated leukocyte cell adhesion molecule
INPPL1	inositol polyphosphate phosphatase-like 1
ITGA2	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)
ITGA5	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)
ITGAX	integrin, alpha X (complement component 3 receptor 4 subunit)
ITGB3	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)
LYVE1	lymphatic vessel endothelial hyaluronan receptor 1
LY9	lymphocyte antigen 9
MFGE8	milk fat globule-EGF factor 8 protein
MYBPH	myosin binding protein H
NRP2	neuropilin 2
PARVB	parvin, beta
PXN	paxillin
PVRL2	poliovirus receptor-related 2 (herpesvirus entry mediator B)
PVRL3	poliovirus receptor-related 3
PVRL4	poliovirus receptor-related 4
PSEN1	presenilin 1
PTPRF	protein tyrosine phosphatase, receptor type, F
PTPRM	protein tyrosine phosphatase, receptor type, M
PCDH1	protocadherin 1
PCDH18	protocadherin 18
RS1	retinoschisin 1
SCARF1	scavenger receptor class F, member 1
SPP1	secreted phosphoprotein 1
SELPLG	selectin P ligand
SIGLEC16	sialic acid binding Ig-like lectin 16 (gene/pseudogene)

ROCK1P1	similar to Rho-associated, coiled-coil containing protein kinase 1; Rho-associated, coiled-coil containing protein kinase 1
BCAR1	similar to breast cancer anti-estrogen resistance 1; breast cancer anti-estrogen resistance 1
S1PR1	sphingosine-1-phosphate receptor 1
THBS1	thrombospondin 1
TMEM8A	transmembrane protein 8A
SRC	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)
ZYX	zyxin

Supplementary Table 9 List of genes that fall under the GO annotation category of "cell surface receptor linked signal transduction" within the osteoclast samples treated with RANKL for 7 days.

ID	Gene Name
ADAM17	ADAM metallopeptidase domain 17
ADAM8	ADAM metallopeptidase domain 8
ATP6AP1	ATPase, H+ transporting, lysosomal accessory protein 1
CLEC4A	C-type lectin domain family 4, member A
CD79A	CD79a molecule, immunoglobulin-associated alpha
CD8A	CD8a molecule
DIXDC1	DIX domain containing 1
FYN	FYN oncogene related to SRC, FGR, YES
GPR113	G protein-coupled receptor 113
GPR153	G protein-coupled receptor 153
GPR162	G protein-coupled receptor 162
GPR18	G protein-coupled receptor 18
GPR182	G protein-coupled receptor 182
GPR55	G protein-coupled receptor 55
GPR61	G protein-coupled receptor 61
GPR68	G protein-coupled receptor 68
GRK1	G protein-coupled receptor kinase 1
GLI1	GLI family zinc finger 1
GAB2	GRB2-associated binding protein 2
JAK2	Janus kinase 2
MICA	MHC class I polypeptide-related sequence A
NDST1	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 1
NFAM1	NFAT activating protein with ITAM motif 1
NOTCH3	Notch homolog 3 (Drosophila)
PDZD3	PDZ domain containing 3
REPS2	RALBP1 associated Eps domain containing 2
SMAD7	SMAD family member 7
TIAM1	T-cell lymphoma invasion and metastasis 1
TAX1BP3	Tax1 (human T-cell leukemia virus type I) binding protein 3
THY1	Thy-1 cell surface antigen
ACVRL1	activin A receptor type II-like 1
ADCY6	adenylate cyclase 6
BAI1	brain-specific angiogenesis inhibitor 1
CELSR1	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, Drosophila)
CALCR	calcitonin receptor
CACNB3	calcium channel, voltage-dependent, beta 3 subunit
CLCF1	cardiotrophin-like cytokine factor 1
CERK	ceramide kinase
CER1	cerberus 1, cysteine knot superfamily, homolog (Xenopus laevis)
CLN3	ceroid-lipofuscinosis, neuronal 3
CCL5	chemokine (C-C motif) ligand 5
CCR5	chemokine (C-C motif) receptor 5
CCRL2	chemokine (C-C motif) receptor-like 2
CXCL1	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
CXCL10	chemokine (C-X-C motif) ligand 10
CXCL16	chemokine (C-X-C motif) ligand 16
CXCL3	chemokine (C-X-C motif) ligand 3
CXCL5	chemokine (C-X-C motif) ligand 5
CXCL9	chemokine (C-X-C motif) ligand 9
CX3CR1	chemokine (C-X3-C motif) receptor 1
C5AR1	complement component 5a receptor 1
CCND1	cyclin D1
CCR6	cyclin L2; chemokine (C-C motif) receptor 6
DLL1	delta-like 1 (Drosophila)
DGKI	diacylglycerol kinase, iota
DGKZ	diacylglycerol kinase, zeta 104kDa
DDR2	discoidin domain receptor tyrosine kinase 2
EMR2	egf-like module containing, mucin-like, hormone receptor-like 2
EMR3	egf-like module containing, mucin-like, hormone receptor-like 3
ENG	endoglin
EFNB3	ephrin-B3
FGFR2	fibroblast growth factor receptor 2
FLNA	filamin A, alpha (actin binding protein 280)
FPR3	formyl peptide receptor 3
FZD8	frizzled homolog 8 (Drosophila)
GABRG1	gamma-aminobutyric acid (GABA) A receptor, gamma 1
GRP	gastrin-releasing peptide

GHRL	ghrelin/obestatin prepropeptide
GH1	growth hormone 1
GNB2	guanine nucleotide binding protein (G protein), beta polypeptide 2
GNG12	guanine nucleotide binding protein (G protein), gamma 12
HIPK2	homeodomain interacting protein kinase 2; similar to homeodomain interacting protein kinase 2
APOE	hypothetical LOC100129500; apolipoprotein E
IKBKG	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma
ITGA2	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)
ITGA5	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)
ITGAX	integrin, alpha X (complement component 3 receptor 4 subunit)
ITGB3	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)
IL1R1	interleukin 1 receptor, type I
IL7R	interleukin 7 receptor
IRAK1	interleukin-1 receptor-associated kinase 1
IFT52	intraflagellar transport 52 homolog (Chlamydomonas)
KL	klotho
KREMEN1	kringle containing transmembrane protein 1
LTBP2	latent transforming growth factor beta binding protein 2
LGR4	leucine-rich repeat-containing G protein-coupled receptor 4
LIF	leukemia inhibitory factor (cholinergic differentiation factor)
LILRB3	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 3
LTB4R	leukotriene B4 receptor
LTB4R2	leukotriene B4 receptor 2
LY6E	lymphocyte antigen 6 complex, locus E
LCP2	lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa)
LPAR5	lysophosphatidic acid receptor 5
MTNR1A	melatonin receptor 1A
MITF	microphthalmia-associated transcription factor
MYO1E	myosin IE
NLK	nemo-like kinase
NPFFR1	neuropeptide FF receptor 1
NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
OR1C1	olfactory receptor, family 1, subfamily C, member 1
OR1D5	olfactory receptor, family 1, subfamily D, member 4; olfactory receptor, family 1, subfamily D, member 5
OR1F2P	olfactory receptor, family 1, subfamily F, member 2
OR1J2	olfactory receptor, family 1, subfamily J, member 2
OR10AG1	olfactory receptor, family 10, subfamily AG, member 1
OR10G9	olfactory receptor, family 10, subfamily G, member 9
OR10P1	olfactory receptor, family 10, subfamily P, member 1
OR10Z1	olfactory receptor, family 10, subfamily Z, member 1
OR2AT4	olfactory receptor, family 2, subfamily AT, member 4
OR2M7	olfactory receptor, family 2, subfamily M, member 7
OR2T4	olfactory receptor, family 2, subfamily T, member 4
OR2W5	olfactory receptor, family 2, subfamily W, member 5
OR4E2	olfactory receptor, family 4, subfamily E, member 2
OR4L1	olfactory receptor, family 4, subfamily L, member 1
OR4N5	olfactory receptor, family 4, subfamily N, member 5
OR5H6	olfactory receptor, family 5, subfamily H, member 6
OR51B2	olfactory receptor, family 51, subfamily B, member 2
OR51I1	olfactory receptor, family 51, subfamily I, member 1
OR51S1	olfactory receptor, family 51, subfamily S, member 1
OR52B4	olfactory receptor, family 52, subfamily B, member 4
OR6S1	olfactory receptor, family 6, subfamily S, member 1
OR6Y1	olfactory receptor, family 6, subfamily Y, member 1
OR7G2	olfactory receptor, family 7, subfamily G, member 2
OR8A1	olfactory receptor, family 8, subfamily A, member 1
OR9G9	olfactory receptor, family 9, subfamily G, member 9; olfactory receptor, family 9, subfamily G, member 1
OPN5	opsin 5
OXGR1	oxoglutarate (alpha-ketoglutarate) receptor 1
PXN	paxillin
PDGFB	platelet-derived growth factor beta polypeptide (simian sarcoma viral (v-sis) oncogene homolog)
PLXNA2	plexin A2
PORCN	porcupine homolog (Drosophila)
PSEN1	presenilin 1
PRLR	prolactin receptor
PTGER2	prostaglandin E receptor 2 (subtype EP2), 53kDa
PTGIR	prostaglandin I2 (prostacyclin) receptor (IP)
PTPN22	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
PTPRE	protein tyrosine phosphatase, receptor type, E
PTPRF	protein tyrosine phosphatase, receptor type, F
PTPRG	protein tyrosine phosphatase, receptor type, G

PTPRJ	protein tyrosine phosphatase, receptor type, J
P2RY2	purinergic receptor P2Y, G-protein coupled, 2
RIPK2	receptor-interacting serine-threonine kinase 2
RBPJ	recombination signal binding protein for immunoglobulin kappa J region
RXFP3	relaxin/insulin-like family peptide receptor 3
RARG	retinoic acid receptor, gamma
RPS6KA5	ribosomal protein S6 kinase, 90kDa, polypeptide 5
SCG5	secretogranin V (7B2 protein)
SPN	sialophorin
SIRPB1	signal-regulatory protein beta 1
BCAR1	similar to breast cancer anti-estrogen resistance 1; breast cancer anti-estrogen resistance 1
SSTR1	somatostatin receptor 1
SPHK1	sphingosine kinase 1
S1PR1	sphingosine-1-phosphate receptor 1
SUCNR1	succinate receptor 1
TAS2R14	taste receptor, type 2, member 14
TAAR3	trace amine associated receptor 3 (gene/pseudogene)
TAAR5	trace amine associated receptor 5
TGFBR1	transforming growth factor, beta receptor 1
TGFBR2	transforming growth factor, beta receptor II (70/80kDa)
TGM2	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)
TRPV4	transient receptor potential cation channel, subfamily V, member 4
TNFRSF13B	tumor necrosis factor receptor superfamily, member 13B
TNFRSF14	tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator)
TP63	tumor protein p63
AKT1	v-akt murine thymoma viral oncogene homolog 1
FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog
SRC	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)
VIPR2	vasoactive intestinal peptide receptor 2
WNT2B	wingless-type MMTV integration site family, member 2B
XPR1	xenotropic and polytropic retrovirus receptor
ZNF219	zinc finger protein 219
ZFYVE16	zinc finger, FYVE domain containing 16

Supplementary Table 10 Genes, that were differentially upregulated after 1 day of treatment with RANKL+DMSO and that were found to contain an NFAT binding consensus motif close to their TSS.

Name	Description	Distance to TSS	Sequence
ANXA2P1	annexin A2 pseudogene 1	-188	TTTTTCCACT
AXDND1	axonemal dynein light chain domain containing 1	-84	GATGGAAACA
BBS10	Bardet-Biedl syndrome 10	-98	CTTTTCCACC
BZRAP1	benzodiazepine receptor (peripheral) associated protein 1	-91	CATGGAAACC
C14orf105	chromosome 14 open reading frame 105	-38	CATGGAAACC
C2orf71	chromosome 2 open reading frame 71	-38	CTTGGAAAAT
C6orf132	chromosome 6 open reading frame 132	-131	TTTTTCCACC
CAPN6	calpain 6	-92	ACTTTCCACA
CCDC39	coiled-coil domain containing 39	-183	CGTGGAAAGT
CD8A	CD8a molecule	35	AATTTCCATC
CFH	complement factor H	-54	ATTTTCCAAC
CLEC11A	C-type lectin domain family 11, member A	-127	AGTGGAAAGA
CORO2A	coronin, actin binding protein, 2A	-128	TTTTTCCAGC
CXCR2P1	chemokine (C-X-C motif) receptor 2 pseudogene 1	-113	AATGGAAAAT
DACT2	dishevelled-binding antagonist of beta-catenin 2	-241	TTGTTCCACT
DNAH7	dynein, axonemal, heavy chain 7	-67	AATGGAACCT
DYTN	dystrotelin	-29	ACTTTCCACA
ENAM	enamelin	-32	AATGGAACCT
FAM163A	family with sequence similarity 163, member A	-169	ATTTTCCGTC
FAM163A	family with sequence similarity 163, member A	-223	TCTTTCCAGT
FAM181A-AS1	FAM181A antisense RNA 1	-65	AGTTTCCATG
FAM71F2	family with sequence similarity 71, member F2	-125	GGTGGAAAGA
FAM71F2	family with sequence similarity 71, member F2	-203	TTTTTCCATG
FCAMR	Fc receptor, IgA, IgM, high affinity	-140	TCTTTCCATG
FOXN3-AS1	FOXN3 antisense RNA 1	-153	TGTTTCCACT
GAL3ST4	galactose-3-O-sulfotransferase 4	-128	TCTTTCCACC
GFRA3	GDNF family receptor alpha 3	-205	AGTTTCCAGC
GIMAP2	GTPase, IMAP family member 2	-139	ACTTTCCATT
GIMAP2	GTPase, IMAP family member 2	-234	TCTTTCCAGT
GRIN3A	glutamate receptor, ionotropic, N-methyl-D-aspartate 3A	-15	GGTGGAAAAA
GRM4	glutamate receptor, metabotropic 4	-159	TGTGGAAAGT
GSPT2	G1 to S phase transition 2	-130	AATGGAAATT
HIST1H3I	histone cluster 1, H3i	-224	TGTTTCCGTT
IGBP1P1	immunoglobulin (CD79A) binding protein 1 pseudogene 1	-64	TGTGGAAACT
IGSF6	immunoglobulin superfamily, member 6	-136	CATGGAAAAA
INHBA	inhibin, beta A	-100	ACTGGAAAAA
INHBA	inhibin, beta A	-145	GCTTTCCATT
KALRN	kalirin, RhoGEF kinase	-230	GCTGGAAAAA
KRBA2	KRAB-A domain containing 2	-246	ATTGGAAACA
KRBA2	KRAB-A domain containing 2	-262	TATTTCCATT
LOC101927181	uncharacterized LOC101927181	-203	ATTGGAAAAA
LOC644656	uncharacterized LOC644656	-220	GTTTTCCATT
MAP7	microtubule-associated protein 7	-258	AACGGAAACT
MB	myoglobin	-229	TCTTTCCACA
METTL25	methyltransferase like 25	-183	AGTGGAAACC
MYH15	myosin, heavy chain 15	-92	TGTTTCCACC
NOS1	nitric oxide synthase 1 (neuronal)	-8	AATGGAAATT
NPTN-IT1	NPTN intronic transcript 1 (non-protein coding)	-15	GCTTTCCATC
OCSTAMP	osteoclast stimulatory transmembrane protein	-91	AGTGGAAAAG
OPRM1	opioid receptor, mu 1	-35	GCTGGAAAAT
OSM	oncostatin M	-135	CGTGGAAATT
OSMR	oncostatin M receptor	-25	GTTTTCCAGT
P2RY14	purinergic receptor P2Y, G-protein coupled, 14	-50	TGTTTCCACG
PAX7	paired box 7	-115	CTTTTCCATT
PCED1B-AS1	PCED1B antisense RNA 1	-262	AATGGAAAAA
PRLR	prolactin receptor	-80	CCTGGAAAAA
PURG	purine-rich element binding protein G	27	GATGGAAAGA

RGL4	ral guanine nucleotide dissociation stimulator-like 4	-60	ACTTTCCAGC
RIMBP3	RIMS binding protein 3	-5	ATTTTCCAGC
RSAD2	radical S-adenosyl methionine domain containing 2	-11	TGTGGAAAAA
SLCO6A1	solute carrier organic anion transporter family, member 6A1	-140	TTTTTCCATG
SNAP47	synaptosomal-associated protein, 47kDa	-279	TCTTTCCAGT
SP140	SP140 nuclear body protein	-161	CTTTTCCACC
ST8SIA1	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	-51	CGTGGAAATT
SUSD3	sushi domain containing 3	-199	ACTTTCCGCT
SUSD3	sushi domain containing 3	-290	CATGGAAAAA
TJP1	tight junction protein 1	-149	AGCGGAAACT
TMEM144	transmembrane protein 144	-250	AGTGGAAAAA
TTC22	tetratricopeptide repeat domain 22	-83	AGCGGAAACA
UGT3A1	UDP glycosyltransferase 3 family, polypeptide A1	-117	ATTTTCCAGA
USP51	ubiquitin specific peptidase 51	-75	AGTTTCCATC
VSTM2A	V-set and transmembrane domain containing 2A	-155	AGTGGAAAAA
ZFHX2	zinc finger homeobox 2	36	AATGGAAACC
ZNF214	zinc finger protein 214	-18	CATGGAAACA
ZNF483	zinc finger protein 483	-265	TTTTTCCACC
ZNF699	zinc finger protein 699	-226	TTGTTCCACT

Supplementary Table 11 A list of genes that have removed H3K27me3 marks and increased expression after 1 day of stimulation with RANKL, but retain their H3K27me3 marks if treated with GSK-J4.

Gene ID	Distance to TSS	Peak Score	Gene name
LHX4	2287	8.00694	LIM homeobox 4
RORA	-368	7.41713	RAR-related orphan receptor A
ANKRD18B	1489	6.33326	ankyrin repeat domain 18B
STMN2	4691	6.33326	stathmin-like 2
FOXA1	-2659	6.18011	forkhead box A1
NCKAP1	1085	6.17367	NCK-associated protein 1
EPHA7	-2224	6.13187	EPH receptor A7
EPHA7	4691	6.13187	EPH receptor A7
CDC20B	-3314	5.96073	cell division cycle 20 homolog B (S. cerevisiae)
HNF1B	1342	5.72921	HNF1 homeobox B
NRG1	4048	5.51978	neuregulin 1
SNCAIP	-1766	5.44172	synuclein, alpha interacting protein
OTUD7B	-2689	5.20016	OTU domain containing 7B
BHLHE22	2529	5.02897	basic helix-loop-helix family, member e22
STK3	2899	5.02897	serine/threonine kinase 3
TFAP2C	-1688	5.02897	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)
ZDBF2	-324	5.02897	zinc finger, DBF-type containing 2
SPAG6	2830	4.92857	sperm associated antigen 6
PPIC	-1894	4.75786	peptidylprolyl isomerase C (cyclophilin C)
PVRL3	3440	4.67957	poliovirus receptor-related 3
ELOVL4	-1750	4.41443	ELOVL fatty acid elongase 4
GPR114	89	4.41443	G protein-coupled receptor 114
LRIF1	-3013	4.41443	ligand dependent nuclear receptor interacting factor 1
MPPED2	-1864	4.41443	metallophosphoesterase domain containing 2
MPPED2	2791	4.41443	metallophosphoesterase domain containing 2
MPPED2	1496	4.41443	metallophosphoesterase domain containing 2
ONECUT2	-1677	4.41443	one cut homeobox 2
ONECUT2	2223	4.41443	one cut homeobox 2
SFRP2	-3551	4.41443	secreted frizzled-related protein 2
SELK	-4245	4.3619	selenoprotein K
CDKN2A	1956	3.82704	cyclin-dependent kinase inhibitor 2A
CDKN2A	-4082	3.82704	cyclin-dependent kinase inhibitor 2A
FLJ30679	3778	3.82704	uncharacterized protein FLJ30679
INTS4L2	-1318	3.82704	integrator complex subunit 4-like 2
PCDH9	-2866	3.82704	protocadherin 9
PDE4B	-704	3.82704	phosphodiesterase 4B, cAMP-specific
SCRN2	-3831	3.82704	secernin 2
TASP1	-1211	3.82704	taspase, threonine aspartase, 1
ZBTB8A	-4608	3.82704	zinc finger and BTB domain containing 8A
KIF3A	36	3.81851	kinesin family member 3A
OSR2	113	3.81851	odd-skipped related 2 (Drosophila)
SETBP1	3571	3.81851	SET binding protein 1
STON2	2707	3.81851	stonin 2
TFAP2C	2061	3.81851	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)
FAM171A1	1954	3.55252	family with sequence similarity 171, member A1
FGF2	666	3.30066	fibroblast growth factor 2 (basic)
LRIF1	-104	3.30066	ligand dependent nuclear receptor interacting factor 1
SKAP1	4204	3.30066	src kinase associated phosphoprotein 1
SLC4A4	3606	3.30066	solute carrier family 4, sodium bicarbonate cotransporter, member 4
TBX20	2012	3.30066	T-box 20
ZYG11A	3452	3.30066	zyg-11 family member A, cell cycle regulator
ATP1A3	-4057	3.2686	ATPase, Na+/K+ transporting, alpha 3 polypeptide
CACNA1C	-2371	3.2686	calcium channel, voltage-dependent, L type, alpha 1C subunit
CAV2	2201	3.2686	caveolin 2
GLIS1	-2352	3.2686	GLIS family zinc finger 1
PCDH7	3812	3.2686	protocadherin 7
SEMA6D	1014	3.2686	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D
SEMA6D	-3876	3.2686	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D
TMEM132E	-274	3.2686	transmembrane protein 132E

Supplementary Table 12 The top 100 up-regulated genes at 2h GSK-J4 treated JJN3 cells as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene	2h/DMSO	6h/DMSO	24h/DMSO	Gene name
MIR4664	560.06	1.00	1.00	microRNA 4664
MIR1291,SNORA34	423.58	299.54	596.45	small nucleolar RNA, H/ACA box 2B
SCARNA4	223.50	123.26	1.00	small Cajal body-specific RNA 4
MIR330	218.05	1.00	1.00	microRNA 330
RNU4ATAC	217.70	120.07	179.13	RNA, U4atac small nuclear (U12-dependent splicing)
MIR3662	210.44	1.00	1.00	microRNA 3662
MIR647	203.23	334.35	1.00	UCKL1 antisense RNA 1
MIR635	189.60	1.00	1.00	microRNA 635
VTRNA2-1	177.16	1.00	1.00	vault RNA 2-1
MIR3939	145.66	239.47	1.00	microRNA 3939
SCARNA8	141.73	1.00	1.00	small Cajal body-specific RNA 8
MT2A	131.39	634.30	288.08	metallothionein 2A
MT1G	119.81	763.45	248.18	metallothionein 1G
MIR320B2	118.77	1.00	1.00	microRNA 320b-2
KISS1R	118.15	166.52	372.41	KISS1 receptor
MIR663B	110.72	1.00	1.00	microRNA 663b
MIR769	101.52	1.00	1.00	microRNA 769
MIR5194	95.94	1.00	1.00	family with sequence similarity 49, member B
SCARNA18	66.12	1.00	161.58	small Cajal body-specific RNA 18
FAM172A,MIR2277	51.98	0.46	0.80	POU domain class 5, transcription factor 2
SCARNA3	51.90	1.00	1.00	small Cajal body-specific RNA 3
MIR3689B	47.34	1.00	1.00	microRNA 3689d-1
SYCP3	33.54	7.71	1.00	synaptonemal complex protein 3
GP6	29.02	1.00	7.91	glycoprotein VI (platelet)
ASF1A	28.23	13.83	20.19	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)
CELSR3,MIR4793	26.25	1.16	1.39	microRNA 4793
KCNRG	25.78	1.00	11.18	microRNA 15a
C19orf71	25.76	31.61	12.45	chromosome 19 open reading frame 71
GABRQ	25.55	5.50	7.73	gamma-aminobutyric acid (GABA) A receptor, theta
MTHFD2L	25.07	19.44	12.12	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like
TP53INP1	24.01	7.32	105.02	tumor protein p53 inducible nuclear protein 1
ACPT	22.86	6.15	4.85	acid phosphatase, testicular
KCNV2	21.14	7.64	1.00	potassium channel, subfamily V, member 2
C21orf2	21.06	10.92	1.00	chromosome 21 open reading frame 2
SLC9A3	20.37	20.16	1.00	solute carrier family 9, subfamily A (NHE3, cation proton antiporter 3), member 3
RBM12B-AS1	19.92	15.67	17.46	RBM12B antisense RNA 1
NRF1,RNA5SP244	19.61	0.85	0.89	nuclear respiratory factor 1
FILIP1L	18.50	23.44	1.00	filamin A interacting protein 1-like
VASH1	17.92	1.00	12.38	vasohibin 1
HIST1H4K	17.52	39.13	33.59	histone cluster 1, H4k
MT1H	16.53	372.30	103.15	metallothionein 1H
NOP14-AS1	16.51	1.00	4.48	NOP14 antisense RNA 1
PRSS21	16.36	11.91	9.90	protease, serine, 21 (testisin)
SMKR1	15.60	12.03	8.50	small lysine-rich protein 1
C2orf47	15.21	9.30	6.73	chromosome 2 open reading frame 47
GRIN2D	13.69	5.18	10.38	glutamate receptor, ionotropic, N-methyl D-aspartate 2D
WARS2	13.48	18.82	7.15	tryptophanyl tRNA synthetase 2, mitochondrial
PROB1	13.29	5.34	9.66	proline-rich basic protein 1
TSSK3	12.94	4.28	5.91	testis-specific serine kinase 3
RPS27	11.11	1.00	9.31	ribosomal protein S27
THBS3	11.05	7.37	4.81	thrombospondin 3
PCGF3	10.49	28.37	1.00	polycomb group ring finger 3
PIGP	10.29	16.26	7.22	phosphatidylinositol glycan anchor biosynthesis, class P
C5orf45	10.11	11.02	12.26	sequestosome 1
IFITM3	9.69	6.43	6.41	interferon induced transmembrane protein 3
MT1P2	9.65	72.33	22.34	metallothionein 1H-like 1
STK19	9.51	11.73	8.37	serine/threonine kinase 19
KCTD6	9.22	2.02	1.92	potassium channel tetramerisation domain containing 6
S100A4	9.16	17.14	4.45	S100 calcium binding protein A4
FAM65A	9.15	14.43	23.96	family with sequence similarity 65, member A
CCS	8.94	3.01	7.02	copper chaperone for superoxide dismutase
NAPA-AS1	8.81	2.84	6.50	NAPA antisense RNA 1
TMEM216	8.47	12.55	5.61	transmembrane protein 216
PIGL	8.47	2.76	8.89	microRNA 1288
S100A6	8.37	4.47	16.57	S100 calcium binding protein A6

STARD5	7.82	4.75	3.24	interleukin 16
CHAC1	7.78	17.73	14.91	ChaC, cation transport regulator homolog 1 (E. coli)
GPX7	7.42	5.88	3.43	glutathione peroxidase 7
GTF2H4	7.36	9.83	7.61	general transcription factor IIH, polypeptide 4, 52kDa
WDFY1	7.15	3.38	6.55	WD repeat and FYVE domain containing 1
C11orf94	7.07	6.00	1.00	chromosome 11 open reading frame 94
BLK	6.95	5.55	3.62	B lymphoid tyrosine kinase
NMU	6.77	5.08	9.14	neuromedin U
TTC32	6.75	1.00	3.02	tetratricopeptide repeat domain 32
HIBCH	6.66	1.98	6.14	3-hydroxyisobutyryl-CoA hydrolase
CRYBB1	6.64	4.48	16.64	crystallin, beta B1
HIST1H3J	6.50	1.00	1.00	histone cluster 1, H3j
C2orf74	6.39	3.54	21.88	chromosome 2 open reading frame 74
G0S2	6.33	7.58	10.85	G0/G1switch 2
PAOX	6.31	3.24	4.02	polyamine oxidase (exo-N4-amino)
PMS2P4	6.19	9.56	3.59	postmeiotic segregation increased 2 pseudogene 4
MPZL3	6.19	3.14	2.60	myelin protein zero-like 3
MT1X	6.17	26.76	4.68	metallothionein 1X
TPPP	6.16	22.28	4.18	tubulin polymerization promoting protein
C1orf86	6.15	5.75	22.33	protein kinase C, zeta
FAHD1	6.13	4.38	6.06	meiosis specific with OB domains
SLFNL1	6.13	0.11	4.56	sex comb on midleg homolog 1 (Drosophila)
B9D1	6.06	11.90	2.99	B9 protein domain 1
TNFRSF12A	5.88	5.03	2.51	tumor necrosis factor receptor superfamily, member 12A
ADM5	5.78	4.50	1.65	adrenomedullin 5 (putative)
S100A10	5.75	1.00	1.00	S100 calcium binding protein A10
C3orf37	5.74	4.03	5.94	chromosome 3 open reading frame 37
SLC9A5	5.74	4.08	6.66	solute carrier family 9, subfamily A (NHE5, cation proton antiporter 5), member 5
GSTT1	5.72	2.95	3.91	glutathione S-transferase theta 1
IFI6	5.70	3.63	12.82	interferon, alpha-inducible protein 6
NCOR1P1	5.66	4.84	1.00	nuclear receptor corepressor 1 pseudogene 1
MIR564,TMEM42	5.61	8.67	5.87	transmembrane protein 42
NAGS	5.61	6.81	2.34	N-acetylglutamate synthase
SLC25A14	5.51	3.03	6.06	solute carrier family 25 (mitochondrial carrier, brain), member 14
NDUFAF5	5.37	4.33	3.49	NADH dehydrogenase (ubiquinone) complex I, assembly factor 5

Supplementary Table 13 The top 100 up-regulated genes at 6h GSK-J4 treated JJN3 cells as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene	2h	6h	24h	Gene name
MT1G	119.81	763.45	248.18	metallothionein 1G
MT2A	131.39	634.30	288.08	metallothionein 2A
MT1H	16.53	372.30	103.15	metallothionein 1H
MIR647	203.23	334.35	1.00	UCKL1 antisense RNA 1
MIR1291,SNORA34	423.58	299.54	596.45	small nucleolar RNA, H/ACA box 2B
MIR3939	145.66	239.47	1.00	microRNA 3939
SCARNA11	1.00	200.03	1.00	small Cajal body-specific RNA 11
KISS1R	118.15	166.52	372.41	KISS1 receptor
SCARNA4	223.50	123.26	1.00	small Cajal body-specific RNA 4
RNU4ATAC	217.70	120.07	179.13	RNA, U4atac small nuclear (U12-dependent splicing)
MT1P2	9.65	72.33	22.34	metallothionein 1H-like 1
HIST1H4K	17.52	39.13	33.59	histone cluster 1, H4k
C19orf71	25.76	31.61	12.45	chromosome 19 open reading frame 71
MT1M	5.25	29.01	11.47	metallothionein 1M
PCGF3	10.49	28.37	1.00	polycomb group ring finger 3
MT1X	6.17	26.76	4.68	metallothionein 1X
FILIP1L	18.50	23.44	1.00	filamin A interacting protein 1-like
TPPP	6.16	22.28	4.18	tubulin polymerization promoting protein
MT1E	2.66	20.20	5.10	metallothionein 1E
SLC9A3	20.37	20.16	1.00	solute carrier family 9, subfamily A (NHE3, cation proton antiporter 3), member 3
MTHFD2L	25.07	19.44	12.12	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like
WARS2	13.48	18.82	7.15	tryptophanyl tRNA synthetase 2, mitochondrial
CHAC1	7.78	17.73	14.91	ChaC, cation transport regulator homolog 1 (E. coli)
S100A4	9.16	17.14	4.45	S100 calcium binding protein A4
PIGP	10.29	16.26	7.22	phosphatidylinositol glycan anchor biosynthesis, class P
RBM12B-AS1	19.92	15.67	17.46	RBM12B antisense RNA 1
CD151	1.00	14.62	1.00	CD151 molecule (Raph blood group)
FAM65A	9.15	14.43	23.96	family with sequence similarity 65, member A
PXN-AS1	3.71	14.39	7.68	paxillin
INHBE	1.00	14.09	16.66	inhibin, beta E
ASF1A	28.23	13.83	20.19	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)
MIR1304, SNORA18, SNORA40,SNORA8	2.74	12.94	5.21	microRNA 1304
TMEM216	8.47	12.55	5.61	transmembrane protein 216
SMKR1	15.60	12.03	8.50	small lysine-rich protein 1
PRSS21	16.36	11.91	9.90	protease, serine, 21 (testisin)
B9D1	6.06	11.90	2.99	B9 protein domain 1
STK19	9.51	11.73	8.37	serine/threonine kinase 19
C5orf45	10.11	11.02	12.26	sequestosome 1
C21orf2	21.06	10.92	1.00	chromosome 21 open reading frame 2
APPBP2	2.79	10.76	4.74	amyloid beta precursor protein (cytoplasmic tail) binding protein 2
C2orf82	2.92	10.51	24.70	chromosome 2 open reading frame 82
CA5BP1	2.92	10.49	5.73	carbonic anhydrase VB pseudogene 1
ANXA2R	2.43	10.42	9.22	annexin A2 receptor
GTF2H4	7.36	9.83	7.61	general transcription factor IIH, polypeptide 4, 52kDa
CPSF1,MIR1234, MIR939	1.28	9.67	1.06	microRNA 939
PMS2P4	6.19	9.56	3.59	postmeiotic segregation increased 2 pseudogene 4
C2orf47	15.21	9.30	6.73	chromosome 2 open reading frame 47
RNF187	1.00	9.06	31.15	ring finger protein 187
TGIF1	4.83	8.98	6.42	TGFB-induced factor homeobox 1
FOXD1	3.94	8.69	5.84	forkhead box D1
MIR564,TMEM42	5.61	8.67	5.87	transmembrane protein 42
ERLEC1	5.01	8.62	10.23	ChaC, cation transport regulator homolog 2 (E. coli)
VPREB1	3.31	8.60	1.00	pre-B lymphocyte 1
CEACAM21	3.04	8.41	4.25	carcinoembryonic antigen-related cell adhesion molecule 21
KPTN	4.07	8.31	2.68	kaptin (actin binding protein)
CCL4	3.18	8.18	3.68	chemokine (C-C motif) ligand 4
HHLA3	4.91	8.00	5.19	HERV-H LTR-associating 3
GALT	5.25	7.97	4.48	galactose-1-phosphate uridylyltransferase
OR5B21	3.09	7.88	2.72	olfactory receptor, family 5, subfamily B, member 21
SELM	5.14	7.83	1.00	selenoprotein M
RIT1	3.97	7.71	4.64	Ras-like without CAAX 1
SYCP3	33.54	7.71	1.00	synaptonemal complex protein 3

GSTM2	4.04	7.69	6.00	glutathione S-transferase mu 2 (muscle)
KCNV2	21.14	7.64	1.00	potassium channel, subfamily V, member 2
PDLIM2	3.58	7.63	1.99	PDZ and LIM domain 2 (mystique)
G0S2	6.33	7.58	10.85	G0/G1switch 2
TSNARE1	4.87	7.38	14.95	t-SNARE domain containing 1
THBS3	11.05	7.37	4.81	thrombospondin 3
TP53INP1	24.01	7.32	105.02	tumor protein p53 inducible nuclear protein 1
SPINT1	3.14	7.28	1.00	serine peptidase inhibitor, Kunitz type 1
METAP1D	5.19	7.27	1.00	methionyl aminopeptidase type 1D (mitochondrial)
VMAC	3.76	7.26	7.81	vimentin-type intermediate filament associated coiled-coil protein
ZBED3	1.83	7.18	13.32	zinc finger, BED-type containing 3
SMIM3	4.75	7.17	51.80	small integral membrane protein 3
PLEKHF2	4.46	7.12	10.04	pleckstrin homology domain containing, family F (with FYVE domain) member 2
CAPS	2.08	6.94	6.33	calcyphosine
RPL13P5	2.20	6.93	1.00	ribosomal protein L13 pseudogene 5
CCDC159	2.18	6.82	8.26	coiled-coil domain containing 159
NAGS	5.61	6.81	2.34	N-acetylglutamate synthase
SCARNA1	3.70	6.64	3.31	protein phosphatase 1, regulatory subunit 8
LENEP	3.27	6.60	1.00	lens epithelial protein
HEXIM1	1.96	6.55	0.88	hexamethylene bis-acetamide inducible 1
GTSF1	2.66	6.48	9.19	gametocyte specific factor 1
IGBP1	3.92	6.43	4.72	immunoglobulin (CD79A) binding protein 1
IFITM3	9.69	6.43	6.41	interferon induced transmembrane protein 3
SMOX	1.54	6.35	1.00	spermine oxidase
DKFZP586I1420	2.99	6.26	2.97	uncharacterized protein DKFZp586I1420
SESN2	2.45	6.26	6.81	sestrin 2
ACPT	22.86	6.15	4.85	acid phosphatase, testicular
BOLA1	2.90	6.13	4.37	bolA homolog 1 (E. coli)
SPSB2	3.89	6.00	5.99	triosephosphate isomerase 1
C11orf94	7.07	6.00	1.00	chromosome 11 open reading frame 94
BBC3	1.77	5.96	4.49	microRNA 3190
RBM12B-AS1	2.00	5.94	1.00	RBM12B antisense RNA 1
GPX7	7.42	5.88	3.43	glutathione peroxidase 7
SMIM1	1.00	5.85	1.00	small integral membrane protein 1
RAB39B	4.86	5.84	5.95	RAB39B, member RAS oncogene family
ZNF584	2.45	5.77	3.01	zinc finger protein 584
C1orf86	6.15	5.75	22.33	protein kinase C, zeta
TMEM234	4.45	5.75	3.84	transmembrane protein 234
SWI5	2.42	5.73	6.11	SWI5 recombination repair homolog (yeast)

Supplementary Table 14 The top 100 up-regulated genes at 24h GSK-J4 treated JJN3 cells as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene	2h	6h	24h	Gene name
MIR1291,SNORA34	423.58	299.54	596.45	small nucleolar RNA, H/ACA box 2B
KISS1R	118.15	166.52	372.41	KISS1 receptor
MT2A	131.39	634.30	288.08	metallothionein 2A
MT1G	119.81	763.45	248.18	metallothionein 1G
RNU4ATAC	217.70	120.07	179.13	RNA, U4atac small nuclear (U12-dependent splicing)
SCARNA18	66.12	1.00	161.58	small Cajal body-specific RNA 18
TP53INP1	24.01	7.32	105.02	tumor protein p53 inducible nuclear protein 1
MT1H	16.53	372.30	103.15	metallothionein 1H
DGCR8,MIR1306, MIR3618	1.17	1.40	68.06	microRNA 1306
SMIM3	4.75	7.17	51.80	small integral membrane protein 3
HIST1H4K	17.52	39.13	33.59	histone cluster 1, H4k
RNF187	1.00	9.06	31.15	ring finger protein 187
C2orf82	2.92	10.51	24.70	chromosome 2 open reading frame 82
FAM65A	9.15	14.43	23.96	family with sequence similarity 65, member A
MT1P2	9.65	72.33	22.34	metallothionein 1H-like 1
C1orf86	6.15	5.75	22.33	protein kinase C, zeta
C2orf74	6.39	3.54	21.88	chromosome 2 open reading frame 74
IRAK1BP1	3.12	2.75	20.63	interleukin-1 receptor-associated kinase 1 binding protein 1
ASF1A	28.23	13.83	20.19	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)
RBM12B-AS1	19.92	15.67	17.46	RBM12B antisense RNA 1
INHBE	1.00	14.09	16.66	inhibin, beta E
CRYBB1	6.64	4.48	16.64	crystallin, beta B1
S100A6	8.37	4.47	16.57	S100 calcium binding protein A6
SAT1	1.59	3.91	15.50	spermidine/spermine N1-acetyltransferase 1
MLLT10P1	2.43	1.00	15.14	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 10 pseudogene 1
TSNARE1	4.87	7.38	14.95	t-SNARE domain containing 1
CHAC1	7.78	17.73	14.91	ChaC, cation transport regulator homolog 1 (E. coli)
ZBED3	1.83	7.18	13.32	zinc finger, BED-type containing 3
IFI6	5.70	3.63	12.82	interferon, alpha-inducible protein 6
C19orf71	25.76	31.61	12.45	chromosome 19 open reading frame 71
VASH1	17.92	1.00	12.38	vasohibin 1
C5orf45	10.11	11.02	12.26	sequestosome 1
MTHFD2L	25.07	19.44	12.12	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like
MT1M	5.25	29.01	11.47	metallothionein 1M
KCNRG	25.78	1.00	11.18	microRNA 15a
SORT1	1.02	2.79	10.85	sortilin 1
G0S2	6.33	7.58	10.85	G0/G1switch 2
GRIN2D	13.69	5.18	10.38	glutamate receptor, ionotropic, N-methyl D-aspartate 2D
DUSP28	1.37	1.62	10.24	dual specificity phosphatase 28
ERLEC1	5.01	8.62	10.23	ChaC, cation transport regulator homolog 2 (E. coli)
PLEKHF2	4.46	7.12	10.04	pleckstrin homology domain containing, family F (with FYVE domain) member 2
PRSS21	16.36	11.91	9.90	protease, serine, 21 (testisin)
PROB1	13.29	5.34	9.66	proline-rich basic protein 1
TMEM88B	1.00	1.00	9.60	transmembrane protein 88B
RPS27	11.11	1.00	9.31	ribosomal protein S27
ANXA2R	2.43	10.42	9.22	annexin A2 receptor
GTSF1	2.66	6.48	9.19	gametocyte specific factor 1
NMU	6.77	5.08	9.14	neuromedin U
PIGL	8.47	2.76	8.89	microRNA 1288
C5orf63	2.52	2.46	8.77	chromosome 5 open reading frame 63
FAM122C	2.39	3.87	8.63	family with sequence similarity 122C
PTOV1	3.35	4.95	8.57	prostate tumor overexpressed 1
SMKR1	15.60	12.03	8.50	small lysine-rich protein 1
CDC42EP5	4.01	1.00	8.42	CDC42 effector protein (Rho GTPase binding) 5
STK19	9.51	11.73	8.37	serine/threonine kinase 19
CCDC159	2.18	6.82	8.26	coiled-coil domain containing 159
ANKRD30A	3.69	5.43	8.24	ankyrin repeat domain 30A
COPZ2	1.72	1.00	8.13	microRNA 152
GSDMB	2.22	2.60	8.10	gasdermin B
ACYP1	4.21	5.04	8.10	acylphosphatase 1, erythrocyte (common) type
GP6	29.02	1.00	7.91	glycoprotein VI (platelet)
PCGF5	4.11	2.63	7.88	polycomb group ring finger 5
MAPK11	4.02	2.53	7.86	mitogen-activated protein kinase 11
VMAC	3.76	7.26	7.81	vimentin-type intermediate filament associated coiled-coil protein

DDIT3	1.11	3.67	7.73	methionyl-tRNA synthetase
GABRQ	25.55	5.50	7.73	gamma-aminobutyric acid (GABA) A receptor, theta
SOCS2-AS1	1.00	5.49	7.71	suppressor of cytokine signaling 2
LYRM5	1.00	1.00	7.70	LYR motif containing 5
PXN-AS1	3.71	14.39	7.68	paxillin
S100A9	1.00	3.22	7.66	S100 calcium binding protein A9
GTF2H4	7.36	9.83	7.61	general transcription factor IIH, polypeptide 4, 52kDa
NCR1	3.47	5.07	7.57	natural cytotoxicity triggering receptor 1
RGS6	2.10	3.95	7.48	regulator of G-protein signaling 6
DNAJC12	3.62	3.48	7.47	DnaJ (Hsp40) homolog, subfamily C, member 12
PPP1R15A	2.38	2.47	7.34	protein phosphatase 1, regulatory subunit 15A
ELMOD3	3.03	4.01	7.32	ELMO/CED-12 domain containing 3
PIGP	10.29	16.26	7.22	phosphatidylinositol glycan anchor biosynthesis, class P
WARS2	13.48	18.82	7.15	tryptophanyl tRNA synthetase 2, mitochondrial
CCS	8.94	3.01	7.02	copper chaperone for superoxide dismutase
GPR174	1.48	4.96	6.92	G protein-coupled receptor 174
SESN2	2.45	6.26	6.81	sestrin 2
C2orf47	15.21	9.30	6.73	chromosome 2 open reading frame 47
RSPH3	1.26	1.42	6.70	radial spoke 3 homolog (Chlamydomonas)
CETN2	2.15	4.81	6.69	centrin, EF-hand protein, 2
SLC9A5	5.74	4.08	6.66	solute carrier family 9, subfamily A (NHE5, cation proton antiporter 5), member 5
SLC22A4	2.27	1.00	6.65	microRNA 3936
XAF1	1.06	2.46	6.63	XIAP associated factor 1
C9orf85	2.82	5.50	6.61	chromosome 9 open reading frame 85
TMEM88	1.75	1.00	6.58	transmembrane protein 88
KLHL22	2.89	3.97	6.56	kelch-like family member 22
WDFY1	7.15	3.38	6.55	WD repeat and FYVE domain containing 1
PLCB1	4.70	2.32	6.55	phospholipase C, beta 1 (phosphoinositide-specific)
PI4KAP2	4.59	4.70	6.53	phosphatidylinositol 4-kinase, catalytic, alpha pseudogene 2
NAPA-AS1	8.81	2.84	6.50	NAPA antisense RNA 1
TGIF1	4.83	8.98	6.42	TGFB-induced factor homeobox 1
C1orf192	2.58	4.40	6.42	succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa
IFITM3	9.69	6.43	6.41	interferon induced transmembrane protein 3
SCARNA5	3.22	3.81	6.40	small Cajal body-specific RNA 5
HSPA2	2.58	3.60	6.36	heat shock 70kDa protein 2
CAPS	2.08	6.94	6.33	calcyphosine
RBM4B	3.53	1.61	6.30	RNA binding motif protein 4
CDKN2D	3.15	2.40	6.30	cyclin-dependent kinase inhibitor 2D (p19, inhibits CDK4)
HIBCH	6.66	1.98	6.14	3-hydroxyisobutyryl-CoA hydrolase
SWI5	2.42	5.73	6.11	SWI5 recombination repair homolog (yeast)

Supplementary Table 15 The top 100 down-regulated genes at 2h GSK-J4 treated JJN3 cells as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene	2h	6h	24h	Gene name
ITGB8	0.01	0.01	0.01	integrin, beta 8
ENTPD1	0.05	0.07	0.00	chromosome 10 open reading frame 131
RPL12	0.09	1.22	0.94	small nucleolar RNA, H/ACA box 65
HNRNPF	0.09	0.33	0.41	heterogeneous nuclear ribonucleoprotein F
HIST1H2AH	0.09	0.56	0.42	histone cluster 1, H2ah
HIST1H3H	0.11	0.84	0.59	histone cluster 1, H3h
YWHAG	0.12	0.33	0.18	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide
RANBP9	0.12	0.32	1.57	RAN binding protein 9
RDH13	0.13	1.21	1.00	retinol dehydrogenase 13 (all-trans/9-cis)
EMC10	0.13	0.69	1.63	ER membrane protein complex subunit 10
CDK19	0.14	0.69	1.13	cyclin-dependent kinase 19
HIST1H4C	0.15	0.47	0.31	histone cluster 1, H4c
UBC	0.15	0.46	0.66	ubiquitin C
HIST1H2BD	0.15	0.50	0.34	histone cluster 1, H2bd
CDC42	0.16	3.20	0.51	cell division cycle 42
HMGN1	0.18	0.47	1.25	high mobility group nucleosome binding domain 1
RPS15	0.19	0.51	0.23	ribosomal protein S15
HIST1H2AM	0.19	0.64	0.73	histone cluster 1, H2am
HIST1H2BN	0.19	0.73	0.79	histone cluster 1, H2bn
HIST2H2BF	0.19	0.60	0.45	Fc fragment of IgG, high affinity Ia, receptor (CD64)
C19orf71	0.20	0.60	0.35	chromosome 19 open reading frame 71
HIST1H2AC	0.20	0.80	0.48	histone cluster 1, H2ac
RPPH1	0.20	0.99	1.41	ribonuclease P RNA component H1
TUBB	0.21	0.56	0.55	tubulin, beta class I
HIST1H2BJ	0.22	0.82	0.49	histone cluster 1, H2bj
MIR663A	0.22	1.02	0.78	microRNA 663a
RMRP	0.22	0.96	1.06	RNA component of mitochondrial RNA processing endoribonuclease
HIST1H4E	0.22	0.64	0.41	histone cluster 1, H4e
HIST1H3B	0.24	0.71	0.55	histone cluster 1, H3b
ECHDC2	0.24	0.62	0.78	enoyl CoA hydratase domain containing 2
HIST1H1B	0.24	0.68	0.46	histone cluster 1, H1b
SCARNA2	0.25	1.24	1.55	small Cajal body-specific RNA 2
RPLP1	0.26	0.86	0.86	ribosomal protein, large, P1
HIST1H4H	0.26	0.51	0.72	histone cluster 1, H4h
CPT1A	0.27	0.22	0.83	carnitine palmitoyltransferase 1A (liver)
HIST1H4B	0.28	1.23	1.00	histone cluster 1, H4b
HIST2H2AB	0.28	1.06	0.43	histone cluster 2, H2ab
HEXIM1	0.29	0.36	0.66	hexamethylene bis-acetamide inducible 1
HIST1H4I	0.29	0.69	0.84	histone cluster 1, H4i
IRX4	0.34	1.26	1.46	iroquois homeobox 4
ELL	0.35	5.52	4.96	elongation factor RNA polymerase II
ANKRD36BP1	0.35	0.27	0.88	ankyrin repeat domain 36B pseudogene 1
TTYH1	0.35	0.41	0.47	tweety homolog 1 (Drosophila)
TERC	0.36	2.20	3.01	telomerase RNA component
HK2	0.39	0.09	0.02	hexokinase 2
FAM65A	0.40	1.11	2.74	family with sequence similarity 65, member A
SMN2	0.43	1.47	0.31	survival of motor neuron 1, telomeric
AVPR2	0.43	0.51	0.60	arginine vasopressin receptor 2
HIST1H1E	0.44	1.07	0.72	histone cluster 1, H1e
PRR7-AS1	0.44	0.44	0.44	proline rich 7 (synaptic)
TG	0.45	0.17	2.05	Src-like-adaptor
GPR15	0.45	0.48	1.00	G protein-coupled receptor 15
TPT1-AS1	0.47	1.06	0.91	TPT1 antisense RNA 1
KNDC1	0.48	0.02	1.53	kinase non-catalytic C-lobe domain (KIND) containing 1
NDUFB1	0.50	0.81	0.84	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 1, 7kDa
SCAMP5	0.51	1.12	0.42	secretory carrier membrane protein 5
NOP14-AS1	0.51	0.49	0.88	NOP14 antisense RNA 1
ABHD8	0.52	0.48	0.44	abhydrolase domain containing 8
PRKRA	0.53	0.55	0.59	TTN antisense RNA 1
THOC1	0.54	0.81	1.03	THO complex 1
SLCO3A1	0.55	0.25	0.97	solute carrier organic anion transporter family, member 3A1
COL5A1	0.55	0.71	0.57	collagen, type V, alpha 1
HSPA1B	0.55	0.38	0.47	heat shock 70kDa protein 1B
LRRC48	0.55	0.80	0.93	leucine rich repeat containing 48

PRKRA	0.57	0.91	0.89	TTN antisense RNA 1
KCNV2	0.57	0.81	0.57	potassium channel, subfamily V, member 2
GP6	0.57	0.81	0.57	glycoprotein VI (platelet)
ENTPD1	0.58	0.84	0.44	chromosome 10 open reading frame 131
FKBP1AP1	0.58	0.52	0.45	zinc finger protein 552
ASF1A	0.59	1.22	0.93	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)
RDH13	0.65	0.76	0.60	retinol dehydrogenase 13 (all-trans/9-cis)
GPRIN1	0.68	1.40	5.43	G protein regulated inducer of neurite outgrowth 1
C19orf71	0.68	0.33	0.37	chromosome 19 open reading frame 71
SORT1	0.69	0.99	1.43	sortilin 1
JUND	0.70	3.97	1.24	jun D proto-oncogene
PRKRA	0.70	0.79	0.56	TTN antisense RNA 1
CPAMD8	0.72	0.75	0.66	C3 and PZP-like, alpha-2-macroglobulin domain containing 8
HEXIM2	0.72	0.97	0.63	hexamethylene bis-acetamide inducible 2
SMIM3	0.74	0.63	0.78	small integral membrane protein 3
RRS1	0.75	4.93	4.70	RRS1 ribosome biogenesis regulator homolog (S. cerevisiae)
FAM65A	0.76	0.54	0.47	family with sequence similarity 65, member A
TP53INP1	0.77	0.68	1.97	tumor protein p53 inducible nuclear protein 1
WARS2	0.78	1.29	0.47	tryptophanyl tRNA synthetase 2, mitochondrial
ENTPD1	0.79	0.73	0.78	chromosome 10 open reading frame 131
LEPROTL1	0.79	0.44	0.84	dynactin 6
EIF3C,EIF3CL	0.82	0.74	0.62	eukaryotic translation initiation factor 3, subunit C-like
BPTF	0.82	1.17	1.13	bromodomain PHD finger transcription factor
TG	0.82	0.73	0.73	Src-like-adaptor
GOLGA8F,GOLGA8G	0.83	0.89	0.74	golgin A8 family, member F
CPAMD8	0.84	0.40	1.02	C3 and PZP-like, alpha-2-macroglobulin domain containing 8
ARMC5	0.84	0.81	1.16	armadillo repeat containing 5
C1orf192	0.85	1.66	2.06	succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa
WTAP	0.86	1.04	0.81	Wilms tumor 1 associated protein
SLC22A23	0.87	1.10	0.83	solute carrier family 22, member 23
CRLF3	0.88	0.63	1.00	cytokine receptor-like factor 3
PCGF3	0.88	0.76	0.93	polycomb group ring finger 3
TEX2	0.89	1.01	0.88	testis expressed 2
ZBTB2	0.89	1.43	1.79	zinc finger and BTB domain containing 2
PROB1	0.90	1.35	0.75	proline-rich basic protein 1
ABHD15	0.91	0.21	0.30	abhydrolase domain containing 15

Supplementary Table 16 The top 100 down-regulated genes at 6h GSK-J4 treated JJN3 cells as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene	2h	6h	24h	Gene name
ITGB8	0.01	0.01	0.01	integrin, beta 8
KNDC1	0.48	0.02	1.53	kinase non-catalytic C-lobe domain (KIND) containing 1
ZNF785	3.32	0.03	0.84	zinc finger protein 785
LPCAT3	1.34	0.04	0.04	lysophosphatidylcholine acyltransferase 3
HIST1H2BM	1.55	0.05	1.11	histone cluster 1, H2bm
ALX3	1.98	0.05	0.84	ALX homeobox 3
RNF32	3.56	0.06	0.84	ring finger protein 32
ENTPD1	0.05	0.07	0.00	chromosome 10 open reading frame 131
HK2	0.39	0.09	0.02	hexokinase 2
SLFNL1	6.13	0.11	4.56	sex comb on midleg homolog 1 (Drosophila)
ISG15	1.16	0.13	0.86	ISG15 ubiquitin-like modifier
C2orf48	1.01	0.15	0.86	microRNA 4261
CAND2	3.30	0.16	0.16	cullin-associated and neddylation-dissociated 2 (putative)
CLDN20	3.29	0.16	2.26	T-cell lymphoma invasion and metastasis 2
TG	0.45	0.17	2.05	Src-like-adaptor
TMEM249	1.40	0.19	0.53	transmembrane protein 249
CDH9	2.66	0.19	0.19	cadherin 9, type 2 (T1-cadherin)
OLIG1	2.91	0.20	0.87	oligodendrocyte transcription factor 1
UFSP1	1.28	0.21	1.20	UFM1-specific peptidase 1 (non-functional)
ABHD15	0.91	0.21	0.30	abhydrolase domain containing 15
CPT1A	0.27	0.22	0.83	carnitine palmitoyltransferase 1A (liver)
NRN1L	1.01	0.22	0.22	neuritin 1-like
ERICH1	1.08	0.24	0.18	glutamate-rich 1
ZBTB8OS	1.06	0.24	0.44	zinc finger and BTB domain containing 8 opposite strand
MSRB1	1.62	0.24	0.42	methionine sulfoxide reductase B1
SLCO3A1	0.55	0.25	0.97	solute carrier organic anion transporter family, member 3A1
KDELR3	1.13	0.25	0.25	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
PARD6G-AS1	1.51	0.26	0.26	PARD6G antisense RNA 1
ANKRD36BP1	0.35	0.27	0.88	ankyrin repeat domain 36B pseudogene 1
C1RL-AS1	1.66	0.27	0.72	C1RL antisense RNA 1
NDUFB1	2.54	0.28	2.09	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 1, 7kDa
UBAP1L	1.01	0.28	0.28	ubiquitin associated protein 1-like
B3GNTL1	1.86	0.28	2.08	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like 1
FAM222A-AS1	2.94	0.29	2.07	family with sequence similarity 222, member A
MIR3656, TRAPPC4	1.10	0.30	1.27	microRNA 3656
ST6GALNAC6	1.07	0.30	0.45	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 6
RARS2	1.26	0.30	1.00	arginyl-tRNA synthetase 2, mitochondrial
AGPHD1	1.25	0.30	1.47	aminoglycoside phosphotransferase domain containing 1
NDUFC1	1.07	0.30	0.80	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 1, 6kDa
FAM86DP	1.10	0.31	0.22	family with sequence similarity 86, member D, pseudogene
RPL23AP64	1.82	0.31	0.84	ribosomal protein L23a pseudogene 64
FLJ35390	1.01	0.32	0.60	uncharacterized LOC255031
RANBP9	0.12	0.32	1.57	RAN binding protein 9
PARD6G	1.65	0.32	0.06	PARD6G antisense RNA 1
RAMP1	1.08	0.33	0.34	receptor (G protein-coupled) activity modifying protein 1
ZNF354C	1.01	0.33	0.51	zinc finger protein 354C
MED31	1.47	0.33	0.61	thioredoxin domain containing 17
YWHAG	0.12	0.33	0.18	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide
HNRNPF	0.09	0.33	0.41	heterogeneous nuclear ribonucleoprotein F
C19orf71	0.68	0.33	0.37	chromosome 19 open reading frame 71
TPRKB	1.48	0.34	1.04	TP53RK binding protein
ZFP69B	1.01	0.34	1.51	ZFP69 zinc finger protein B
LDLR	1.19	0.35	0.48	low density lipoprotein receptor
HEXIM1	0.29	0.36	0.66	hexamethylene bis-acetamide inducible 1
YY2	1.08	0.36	0.89	YY2 transcription factor
CPAMD8	1.46	0.36	0.85	C3 and PZP-like, alpha-2-macroglobulin domain containing 8
SLC9A3R2	1.06	0.37	0.20	solute carrier family 9, subfamily A (NHE3, cation proton antiporter 3), member 3 regulator 2
CRNKL1	1.06	0.37	1.04	chromosome 20 open reading frame 26
L3HYPDH	1.16	0.37	0.73	L-3-hydroxyproline dehydratase (trans-)
ACTL10	1.01	0.37	0.40	actin-like 10
METTL24	1.64	0.38	0.90	methyltransferase like 24

YIPF1	1.03	0.38	1.49	Yip1 domain family, member 1
RPL32P3	1.01	0.38	0.73	small nucleolar RNA, H/ACA box 7B
HSPA1B	0.55	0.38	0.47	heat shock 70kDa protein 1B
ANKRD9	1.10	0.38	1.09	ankyrin repeat domain 9
RPL26L1	1.88	0.38	0.15	ribosomal protein L26-like 1
SEMA6D	1.14	0.38	0.69	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D
PRR7-AS1	2.60	0.38	0.86	proline rich 7 (synaptic)
SMYD5	1.30	0.39	0.91	SMYD family member 5
C19orf71	1.45	0.39	0.39	chromosome 19 open reading frame 71
GLA	1.53	0.39	0.21	galactosidase, alpha
ASAP1-IT1	1.01	0.39	1.41	ASAP1 intronic transcript 1 (non-protein coding)
RTN4RL1	3.12	0.40	0.90	reticulon 4 receptor-like 1
CLDN5	1.26	0.40	1.19	uncharacterized LOC150185
CPAMD8	0.84	0.40	1.02	C3 and PZP-like, alpha-2-macroglobulin domain containing 8
TTYH1	0.35	0.41	0.47	tweety homolog 1 (Drosophila)
MTFMT	1.22	0.41	0.61	mitochondrial methionyl-tRNA formyltransferase
RAB20	1.01	0.41	0.52	RAB20, member RAS oncogene family
PIN4	1.01	0.41	0.90	excision repair cross-complementing rodent repair deficiency, complementation group 6-like
PRKRA	1.49	0.42	0.85	TTN antisense RNA 1
METTL8	1.21	0.42	0.67	DDB1 and CUL4 associated factor 17
C10orf10	1.10	0.42	0.56	chromosome 10 open reading frame 10
BHLHA15	1.16	0.43	0.31	basic helix-loop-helix family, member a15
ZNF785	1.22	0.43	0.34	zinc finger protein 785
P2RX5, P2RX5-TAX1BP3, TAX1BP3	1.33	0.43	0.55	Tax1 (human T-cell leukemia virus type I) binding protein 3
HLA-DOB	1.01	0.44	0.91	major histocompatibility complex, class II, DO beta
PDSS1	1.07	0.44	0.86	prenyl (decaprenyl) diphosphate synthase, subunit 1
FKBPL	1.56	0.44	1.10	FK506 binding protein like
LEPROTL1	0.79	0.44	0.84	dynactin 6
TRAF3IP3	1.23	0.44	1.15	TRAF3 interacting protein 3
ATG12	1.14	0.44	0.88	autophagy related 12
GRAMD3	1.02	0.44	2.16	GRAM domain containing 3
PRR7-AS1	0.44	0.44	0.44	proline rich 7 (synaptic)
HDHD3	1.24	0.45	0.17	haloacid dehalogenase-like hydrolase domain containing 3
PITPNA-AS1	1.09	0.45	1.24	PITPNA antisense RNA 1
PMEPA1	1.13	0.45	0.74	prostate transmembrane protein, androgen induced 1
IFNGR1	1.33	0.45	0.72	interferon gamma receptor 1
IDE	2.45	0.46	0.58	insulin-degrading enzyme
CYB561D2	1.16	0.46	1.03	cytochrome b561 family, member D2
UBE2NL	1.01	0.46	0.56	ubiquitin-conjugating enzyme E2N-like
UBC	0.15	0.46	0.66	ubiquitin C

Supplementary Table 17 The top 100 down-regulated genes at 24h GSK-J4 treated JJN3 cells as compared to the DMSO control. The numeric values show the fold increase in expression over DMSO control.

Gene	2h	6h	24h	Gene name
PROM2	2.37	1.12	0.00	Kv channel interacting protein 3, calsenilin
ENTPD1	0.05	0.07	0.00	chromosome 10 open reading frame 131
ITGB8	0.01	0.01	0.01	integrin, beta 8
HK2	0.39	0.09	0.02	hexokinase 2
LPCAT3	1.34	0.04	0.04	lysophosphatidylcholine acyltransferase 3
MIR4767	1.34	0.76	0.04	microRNA 4767
PARD6G	1.65	0.32	0.06	PARD6G antisense RNA 1
GP6	1.22	0.76	0.07	glycoprotein VI (platelet)
HIST1H4D	1.93	2.61	0.10	histone cluster 1, H4d
HSD11B2	1.12	0.81	0.10	hydroxysteroid (11-beta) dehydrogenase 2
FAM58A	1.01	0.75	0.13	family with sequence similarity 58, member A
CAMTA1	1.27	0.54	0.13	calmodulin binding transcription activator 1
TTYH1	2.76	1.10	0.14	tweety homolog 1 (Drosophila)
RPL26L1	1.88	0.38	0.15	ribosomal protein L26-like 1
CAND2	3.30	0.16	0.16	cullin-associated and neddylation-dissociated 2 (putative)
CST3	1.72	2.03	0.16	cystatin C
TPPP	1.95	0.79	0.17	tubulin polymerization promoting protein
HDHD3	1.24	0.45	0.17	haloacid dehalogenase-like hydrolase domain containing 3
DCTN6	1.01	1.87	0.17	dynactin 6
ERICH1	1.08	0.24	0.18	glutamate-rich 1
KATNB1	1.21	0.86	0.18	katanin p80 (WD repeat containing) subunit B 1
YWHAG	0.12	0.33	0.18	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide
CDH9	2.66	0.19	0.19	cadherin 9, type 2 (T1-cadherin)
R3HCC1L	1.26	0.65	0.19	R3H domain and coiled-coil containing 1-like
MITD1	1.01	0.64	0.19	MIT, microtubule interacting and transport, domain containing 1
SLC9A3R2	1.06	0.37	0.20	solute carrier family 9, subfamily A (NHE3, cation proton antiporter 3), member 3 regulator 2
GLA	1.53	0.39	0.21	galactosidase, alpha
CTXN1	1.16	1.14	0.22	cortexin 1
FIS1	1.52	1.15	0.22	fission 1 (mitochondrial outer membrane) homolog (S. cerevisiae)
CNIH4	1.01	3.70	0.22	cornichon homolog 4 (Drosophila)
FAM86DP	1.10	0.31	0.22	family with sequence similarity 86, member D, pseudogene
NRN1L	1.01	0.22	0.22	neuritin 1-like
RPS15	0.19	0.51	0.23	ribosomal protein S15
TST	1.53	0.88	0.23	thiosulfate sulfurtransferase (rhodanese)
SIRT6	1.08	0.86	0.24	sirtuin 6
ZNF136	1.10	0.95	0.24	zinc finger protein 136
CENPV	1.05	0.53	0.24	centromere protein V
ID3	1.50	1.11	0.24	inhibitor of DNA binding 3, dominant negative helix-loop-helix protein
NDUFA8	1.21	1.47	0.24	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19kDa
RHPN1-AS1	1.10	0.56	0.25	RHPN1 antisense RNA 1 (head to head)
KDELR3	1.13	0.25	0.25	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
HTATIP2	1.36	1.35	0.26	HIV-1 Tat interactive protein 2, 30kDa
NAF1	1.24	1.34	0.26	nuclear assembly factor 1 ribonucleoprotein
DNPEP	1.63	1.79	0.26	aspartyl aminopeptidase
CD83	1.45	0.98	0.26	CD83 molecule
MCM3AP-AS1	1.57	0.67	0.26	minichromosome maintenance complex component 3 associated protein
COQ3	1.39	0.88	0.26	coenzyme Q3 homolog, methyltransferase (S. cerevisiae)
PARD6G-AS1	1.51	0.26	0.26	PARD6G antisense RNA 1
HIST1H2AB	1.12	1.16	0.27	histone cluster 1, H2ab
FADS2	1.41	1.33	0.27	fatty acid desaturase 2
RNASET2	1.79	1.40	0.27	ribonuclease T2
PLEKHG5,TNFRSF25	1.30	1.73	0.27	pleckstrin homology domain containing, family G (with RhoGef domain) member 5
HEXIM2	1.75	2.51	0.27	hexamethylene bis-acetamide inducible 2
LRP3	1.13	1.04	0.28	low density lipoprotein receptor-related protein 3
UBAP1L	1.01	0.28	0.28	ubiquitin associated protein 1-like
FLJ20021	1.25	3.08	0.28	uncharacterized LOC90024

DDC	1.14	0.93	0.28	uncharacterized LOC100129427
RASSF4	1.07	1.13	0.29	chromosome 10 open reading frame 10
FKBP4	1.15	1.04	0.29	FK506 binding protein 4, 59kDa
C18orf56	1.01	1.08	0.30	thymidylate synthetase
ABHD15	0.91	0.21	0.30	abhydrolase domain containing 15
KTN1-AS1	1.01	0.69	0.30	KTN1 antisense RNA 1
RASAL1	1.08	0.55	0.31	RAS protein activator like 1 (GAP1 like)
SMN2	0.43	1.47	0.31	survival of motor neuron 1, telomeric
BHLHA15	1.16	0.43	0.31	basic helix-loop-helix family, member a15
HIST1H4C	0.15	0.47	0.31	histone cluster 1, H4c
PIF1	1.19	1.27	0.32	PIF1 5'-to-3' DNA helicase homolog (S. cerevisiae)
B3GNT1	1.05	0.68	0.32	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 1
CCDC51	1.15	0.75	0.32	coiled-coil domain containing 51
HRSP12	1.05	0.76	0.32	heat-responsive protein 12
C5orf55	1.82	1.84	0.32	chromosome 5 open reading frame 55
TESC	2.36	1.10	0.32	tescalcin
LETM2	1.01	1.08	0.33	leucine zipper-EF-hand containing transmembrane protein 2
TCTN3	1.16	0.64	0.33	tectonic family member 3
DUS3L	1.04	0.76	0.33	dihydrouridine synthase 3-like (S. cerevisiae)
SPC25	1.19	1.18	0.33	nitric oxide synthase trafficker
ZNF785	1.22	0.43	0.34	zinc finger protein 785
HIST1H2BD	0.15	0.50	0.34	histone cluster 1, H2bd
RAMP1	1.08	0.33	0.34	receptor (G protein-coupled) activity modifying protein 1
PFKFB3	1.02	0.48	0.34	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
C19orf71	0.20	0.60	0.35	chromosome 19 open reading frame 71
MOB2	1.50	1.69	0.35	keratin associated protein 5-4
HDDC3	1.52	0.58	0.35	unc-45 homolog A (C. elegans)
THRA	1.47	1.00	0.35	thyroid hormone receptor, alpha
HPS6	1.07	0.96	0.36	Hermansky-Pudlak syndrome 6
HMGCS1	1.32	0.78	0.36	3-hydroxy-3-methylglutaryl-CoA synthase 1 (soluble)
TRIM24	1.16	0.60	0.36	tripartite motif containing 24
LSM1	1.38	0.97	0.36	LSM1 homolog, U6 small nuclear RNA associated (S. cerevisiae)
TMEM170A	3.38	2.14	0.36	transmembrane protein 170A
COL18A1	2.51	0.77	0.36	collagen, type XVIII, alpha 1
NAGLU	1.76	0.98	0.37	N-acetylglucosaminidase, alpha
COPG2,TSGA13	1.15	0.76	0.37	coatomer protein complex, subunit gamma 2
NEIL2	1.80	1.12	0.37	nei endonuclease VIII-like 2 (E. coli)
C19orf71	0.68	0.33	0.37	chromosome 19 open reading frame 71
MECR	1.15	0.75	0.38	mitochondrial trans-2-enoyl-CoA reductase
LRRN4	1.36	0.61	0.38	leucine rich repeat neuronal 4
DOC2A	1.11	0.73	0.38	INO80 complex subunit E
ACPT	1.59	0.94	0.38	acid phosphatase, testicular
TCEAL8	1.54	1.41	0.39	transcription elongation factor A (SII)-like 8
WDR83	1.56	2.04	0.39	WD repeat domain 83
SFXN2	1.42	0.98	0.39	sideroflexin 2
TNFRSF14	1.12	1.32	0.39	uncharacterized LOC100133445
RGS9BP	1.51	0.51	0.39	regulator of G protein signaling 9 binding protein

Supplementary Table 18 List of genes that fall under the GO annotation category of "intracellular signalling" that exhibit a gradual increase of expression over 24h upon JJN3 cell treatment with GSK-J4. Gene expression values are displayed in FPKM.

Gene Symbol	24h DMSO	2h GSK-J4	6h GSK-J4	24h GSK-J4	Gene name
HTR7	0	0	0	1.23053	5-hydroxytryptamine (serotonin) receptor 7, adenylate cyclase-coupled
AKAP7	0	0	0	1.65799	A kinase (PRKA) anchor protein 7
GAB1	2.16245	2.59629	5.92581	7.22288	GRB2-associated binding protein 1
NOTCH4	0	0	1.33176	2.75757	notch 4
RAB33B	2.04241	1.79817	3.42014	5.1162	RAB33B, member RAS oncogene family
RAB9A	5.58977	6.17063	8.58054	16.3362	RAB9A, member RAS oncogene family
TIFA	8.86653	6.88786	16.3032	20.904	TRAF-interacting protein with forkhead-associated domain
ASB18	0	0	0	1.04239	ankyrin repeat and SOCS box containing 18
CLNK	0	0	1.13303	0.847449	cytokine-dependent hematopoietic cell linker
EPGN	0	0	0	1.66683	epithelial mitogen homolog (mouse)
HRH3	0	0	0.337286	1.00909	histamine receptor H3
IL31RA	0	0	0	1.01704	interleukin 31 receptor A
NPR2	1.39168	1.41376	1.81258	4.26084	natriuretic peptide receptor B/guanylate cyclase B (atrionatriuretic peptide receptor B)
NF1	5.46786	4.80068	9.46166	16.3728	neurofibromin 1
OPRD1	0	0	0	1.59037	opioid receptor, delta 1
PLA2G4C	0	0	0	1.63686	phospholipase A2, group IVC (cytosolic, calcium-independent)
PLCL1	0.935146	0.949982	1.17447	3.31858	phospholipase C-like 1
PHLDB2, PLCXD2	1.60889	1.90684	1.34707	5.37369	pleckstrin homology-like domain, family B, member 2; phosphatidylinositol-specific phospholipase C, X domain containing 2
P2RY4	0	0	0	1.73969	pyrimidinergic receptor P2Y, G-protein coupled, 4
RHOD	0	0	0.897248	2.68439	ras homolog family member D
RHOT1	5.18664	4.91766	8.39589	12.9926	ras homolog family member T1
RGS9	0	0	0.573892	1.717	regulator of G-protein signaling 9
RCAN1	3.36166	3.64265	9.00826	11.2279	regulator of calcineurin 1
SFN	0	0	0	1.12288	stratifin
TNFRSF10B	3.58721	2.51442	6.49142	8.23061	tumor necrosis factor receptor superfamily, member 10b

Supplementary Table 19 List of genes that fall under the GO annotation category of "regulation of cell proliferation" that exhibit a gradual increase of expression over 24h upon JJN3 cell treatment with GSK-J4. Gene expression values are displayed in FPKM.

Gene Symbol	24h DMSO	2h GSK-J4	6h GSK-J4	24h GSK-J4	Gene name
BTG2	63.8563	68.709	83.6131	134.545	BTG family, member 2
CD28	1.38423	1.29917	3.66974	5.60625	CD28 molecule
NKX2-3	6.32389	5.6212	4.85364	15.8413	NK2 homeobox 3
NOTCH4	0	0	1.33176	2.75757	notch 4
SMAD4	14.6342	17.8928	30.4621	44.6761	SMAD family member 4
TBX19	0	0	0.957033	2.38605	T-box 19
BMP4	0	0	0.545981	1.63382	bone morphogenetic protein 4
CTH	3.31568	2.24563	8.62019	25.803	cystathionase (cystathionine gamma-lyase)
CTLA4	0	0	0	1.5033	cytotoxic T-lymphocyte-associated protein 4
DDR1	0	0	0	1.39567	discoidin domain receptor tyrosine kinase 1
EPGN	0	0	0	1.66683	epithelial mitogen homolog (mouse)
HRH3	0	0	0.337286	1.00909	histamine receptor H3
IL31RA	0	0	0	1.01704	interleukin 31 receptor A
NF1	5.46786	4.80068	9.46166	16.3728	neurofibromin 1
PDCD1LG2	0	0	0	1.12836	programmed cell death 1 ligand 2
SFN	0	0	0	1.12288	stratifin
TOB1	196.194	158.804	275.36	452.014	transducer of ERBB2, 1
WARS	43.7535	33.9989	45.1282	89.4603	tryptophanyl-tRNA synthetase
TNFSF13B	0	0	0	1.03653	tumor necrosis factor (ligand) superfamily, member 13b

Supplementary Table 20 List of genes that fall under the GO annotation category of "regulation of apoptosis" that exhibit an early increase of expression followed by a return to basal levels upon JJN3 cell treatment with GSK-J4. Gene expression values are displayed in FPKM.

Gene Symbol	24h DMSO	2h GSK-J4	6h GSK-J4	24h GSK-J4	Gene name
ANGPTL4	0	1.27764	0	0	angiopoietin-like 4
BAG3	6.32577	13.9233	6.00265	5.28199	BCL2-associated athanogene 3
CARD9	0	1.57236	0.518355	0	caspase recruitment domain family, member 9
CD70	87.9953	118.325	111.583	36.8975	CD70 molecule
CIAPIN1	14.9482	21.0464	17.5662	5.25546	cytokine induced apoptosis inhibitor 1
COL18A1	7.47898	20.2603	5.56623	2.08163	collagen, type XVIII, alpha 1
CRADD	1.51784	4.11179	5.93068	0	CASP2 and RIPK1 domain containing adaptor with death domain
DEDD2	7.74348	23.9919	11.6211	11.8119	death effector domain containing 2
FOXC1	0.922359	3.90414	2.57425	0.770165	forkhead box C1
HSPA1B	27.1297	101.929	26.3208	29.1255	heat shock 70kDa protein 1B
ING4	1.36301	4.15415	1.52169	1.13809	inhibitor of growth family, member 4
P2RX4	1.78424	6.7769	1.91549	2.62408	purinergic receptor P2X, ligand-gated ion channel, 4
PHLDA3	1.13911	2.31445	1.27168	0	pleckstrin homology-like domain, family A, member 3
STRADB	0	1.69282	0	0	STE20-related kinase adaptor beta
TGFB3	0	1.53087	1.40195	0	transforming growth factor, beta 3
TNFRSF4	0	1.12359	0	0	tumor necrosis factor receptor superfamily, member 4
UCN	0	2.45132	0	0	urocortin

Supplementary Table 21 Genes that are constitutively upregulated upon JJN3 treatment with GSK-J4 and that have an MTF1 binding site around their transcription start site.

Gene ID	Distance to TSS	Binding Sequence	Gene Name
MT1E	-35	TGCGCCCGGCCC	metallothionein 1E
MT1E	37	GGGCCGGGTGCA	metallothionein 1E
MT1F	-97	TGCGCCCGGCCC	metallothionein 1F
MT1F	-8	TGCGCCCGGCCC	metallothionein 1F
MT1G	-53	TGCGCCCGGCCC	metallothionein 1G
MT1H	-72	GGGTCGGGTGCA	metallothionein 1H
MT1M	-16	TGCGCCCGGCCC	metallothionein 1M
MT1X	-122	TGCACCCGGCCC	metallothionein 1X
MT1X	-53	TGCGCCCGGCCC	metallothionein 1X
MT2A	-135	TGCGCCCGGCCC	metallothionein 2A
RGS16	-112	TGCACCCGGCCC	regulator of G-protein signaling 16

Supplementary Table 22 List of genes that fall under the GO annotation category of "oxidation reduction" that exhibit a gradual decrease as compared to control over 24 hours upon JJN3 cell treatment with GSK-J4. Gene expression values are displayed in FPKM.

Gene Symbol	24h DMSO	2h GSK-J4	6h GSK-J4	24h GSK-J4	Gene name
AKR7A2, AKR7A3	22.4511	17.1553	22.6568	6.72481	aldo-keto reductase family 7, members A2 and A3 (aflatoxin aldehyde reductase)
BLVRA	1.70233	0.576387	0.950184	0	biliverdin reductase A
CBR1	71.7477	71.4283	48.8594	26.3599	carbonyl reductase 1
CRYZ	7.98712	5.07241	4.05305	1.40934	crystallin, zeta (quinone reductase)
CYB561	11.2217	6.01472	6.90668	4.04295	cytochrome b-561
CYC1	337.036	332.328	302.499	163.647	cytochrome c-1
CYP2U1	3.29134	3.12065	1.46975	1.0993	cytochrome P450, family 2, subfamily U, polypeptide 1
DHCR7	37.0844	33.1466	24.0932	14.7209	7-dehydrocholesterol reductase
DUS3L	23.5287	24.4337	17.7114	7.20372	dihydrouridine synthase 3-like (S. cerevisiae)
DUS4L	5.50283	3.02843	4.90589	1.95637	dihydrouridine synthase 4-like (S. cerevisiae)
EHHADH	2.05241	1.88259	0.458192	0.342686	enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase
ENOX2	5.28427	3.35344	4.2966	1.94795	ecto-NOX disulfide-thiol exchanger 2
FADS2	10.6704	15.4853	14.4648	2.12136	fatty acid desaturase 2
FASN	117.969	144.209	98.0989	57.8088	fatty acid synthase
FDFT1	40.8895	33.0176	27.2133	19.041	farnesyl-diphosphate farnesyltransferase 1
FDX1L	23.3737	18.3897	19.5704	7.80679	ferredoxin 1-like
GPX1	155.761	152.093	129.668	76.6108	glutathione peroxidase 1
HSD11B2	17.847	20.1136	14.32	0.931009	hydroxysteroid (11-beta) dehydrogenase 2
HTATIP2	7.03857	9.95336	9.83863	1.06701	HIV-1 Tat interactive protein 2, 30kDa
IMPDH1	29.8878	36.2438	19.0663	12.478	IMP (inosine 5'-monophosphate) dehydrogenase 1
LDHA	202.811	178.601	178.42	78.5267	lactate dehydrogenase A
LEPREL2	6.14918	4.16449	3.77564	0.513453	leprecan-like 2
ME3	1.46931	0.806373	0.820152	0	malic enzyme 3, NADP(+)-dependent, mitochondrial
MECR	12.7457	14.8483	9.36155	4.15718	mitochondrial trans-2-enoyl-CoA reductase
MICAL2	2.83081	1.64327	1.58012	0.675346	microtubule associated monoxygenase, calponin and LIM domain containing 2
MOXD1	5.2781	5.00438	4.71387	1.76287	monooxygenase, DBH-like 1
NDUFA12	111.772	57.328	56.9347	50.6786	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 12
NDUFA13	439.484	305.078	299.284	150.456	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 13
NDUFA8	35.006	42.6737	52.1065	7.79463	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19kDa
NDUFAB1	122.189	76.2186	93.3323	53.6985	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa
NDUFAF1	10.5661	8.4884	12.4511	3.92119	NADH dehydrogenase (ubiquinone) complex I, assembly factor 1
NDUFS8	81.9957	69.8346	83.2165	29.0462	NADH dehydrogenase (ubiquinone) Fe-S protein 8, 23kDa (NADH-coenzyme Q reductase)
PAM	5.28612	2.2884	4.07043	1.03964	peptidylglycine alpha-amidating monooxygenase
PDHB	28.3174	26.0138	21.4415	10.817	pyruvate dehydrogenase (lipoamide) beta
PECR	3.58157	2.7288	2.99879	0.747648	peroxisomal trans-2-enoyl-CoA reductase
PGD	127.795	101.756	77.0532	59.3498	phosphogluconate dehydrogenase
PHYH	4.25264	2.16001	1.18682	0	phytanoyl-CoA 2-hydroxylase
PYCRL	43.1357	40.6092	24.6772	20.0337	pyrroline-5-carboxylate reductase-like
RDH10	12.7413	7.28069	10.4458	4.987	retinol dehydrogenase 10 (all-trans)
RDH11	15.2011	10.5094	13.0813	6.8753	retinol dehydrogenase 11 (all-trans/9-cis/11-cis)
RTN4IP1	2.46068	1.45817	1.37352	0	reticulon 4 interacting protein 1
SCO2	16.0539	9.86809	10.5849	3.16639	SCO2 cytochrome c oxidase assembly protein
SNCA	2.48494	0.336581	1.10965	0	synuclein, alpha (non A4 component of amyloid precursor)
SPR	14.2898	13.3068	5.98229	5.96595	sepiapterin reductase (7,8-dihydrobiopterin:NADP+ oxidoreductase)
SQRDL	12.1128	7.31901	7.28253	2.33213	sulfide quinone reductase-like (yeast)
SRXN1	27.0929	24.2608	16.1311	12.0653	sulfiredoxin 1
STEAP3	10.2091	8.86379	7.52194	4.23412	STEAP family member 3, metalloreductase
TYW1B	6.42795	3.49002	4.86799	2.68526	tRNA-yW synthesizing protein 1 homolog B (S. cerevisiae)

Supplementary Table 23 Genes that are gradually downregulated over 24 hours of JJN3 treatment with GSK-J4 and that have a MYC binding site around their transcription start site.

Gene ID	Distance to TSS	Binding Sequence	Gene Name
AACS	-239	TACACGTG	acetoacetyl-CoA synthetase
AACS	-230	CACGTGTA	acetoacetyl-CoA synthetase
ABHD8	14	GGCACGTG	abhydrolase domain containing 8
ABHD8	23	CACGTGGC	abhydrolase domain containing 8
AMN	-24	GTCACGTG	amnion associated transmembrane protein
AMN	-15	CACGTGGG	amnion associated transmembrane protein
ANAPC1	-291	TCCACGTG	anaphase promoting complex subunit 1
ANAPC1	-282	CACGTGGG	anaphase promoting complex subunit 1
ATPAF2	-156	AACACGTG	ATP synthase mitochondrial F1 complex assembly factor 2
ATPAF2	-147	CACGTGGC	ATP synthase mitochondrial F1 complex assembly factor 2
BHLHB9	-41	GTCACGTG	basic helix-loop-helix domain containing, class B, 9
BHLHB9	-32	CACGTGGG	basic helix-loop-helix domain containing, class B, 9
C12orf29	-94	CGCACGTG	chromosome 12 open reading frame 29
C12orf29	-85	CACGTGGG	chromosome 12 open reading frame 29
C2orf76	-165	CCCACGTG	chromosome 2 open reading frame 76
C2orf76	-156	CACGTGGG	chromosome 2 open reading frame 76
CBR1	-122	GCCACGTG	carbonyl reductase 1
CBR1	-113	CACGTGGG	carbonyl reductase 1
CHSY3	-54	ACCGCGTG	chondroitin sulfate synthase 3
CLCN5	-24	GTCACGTG	chloride channel, voltage-sensitive 5
CLCN5	-15	CACGTGAC	chloride channel, voltage-sensitive 5
CYB561	-116	CCCACGTG	cytochrome b561
CYB561	-107	CACGTGCG	cytochrome b561
CYB561	-53	GGCACGTG	cytochrome b561
CYB561	-44	CACGTGCG	cytochrome b561
CYTH3	-74	GTCACGTG	cytohesin 3
CYTH3	-65	CACGTGAC	cytohesin 3
DCBLD1	-286	ACCACGCG	discoidin, CUB and LCCL domain containing 1
DCBLD1	-277	CACGCGGC	discoidin, CUB and LCCL domain containing 1
DCBLD1	-225	CACGTGTA	discoidin, CUB and LCCL domain containing 1
DHCR7	-252	CCCACGTG	7-dehydrocholesterol reductase
DHCR7	-243	CACGTGGC	7-dehydrocholesterol reductase
DIXDC1	-48	GCCGCGTG	DIX domain containing 1
EBP	-80	GCCACGCG	emopamil binding protein (sterol isomerase)
ECE2	-216	CCCACGTG	endothelin converting enzyme 2
ECE2	-84	TCCACGTG	endothelin converting enzyme 2
ECE2	-75	CACGTGTT	endothelin converting enzyme 2
ESF1	-23	CCCACGTG	ESF1, nucleolar pre-rRNA processing protein, homolog (S. cerevisiae)
EXOSC1	-62	GACACGTG	exosome component 1
EXOSC1	-53	CACGTGTC	exosome component 1
FAM118B	-57	ACCACGTG	family with sequence similarity 118, member B
FAM118B	-48	CACGTGGA	family with sequence similarity 118, member B
FAM118B	-5	GTCACGTG	family with sequence similarity 118, member B
FAM118B	4	CACGTGGT	family with sequence similarity 118, member B
FAM118B	25	GTCACGTG	family with sequence similarity 118, member B
FAM118B	34	CACGTGGG	family with sequence similarity 118, member B
FAM13A	-40	GTCACGTG	family with sequence similarity 13, member A
FAM13A	-31	CACGTGAC	family with sequence similarity 13, member A
FAM173A	-235	GGCACGTG	family with sequence similarity 173, member A
FAM173A	-226	CACGTGCG	family with sequence similarity 173, member A
FAM98A	-218	CGCGTGGT	family with sequence similarity 98, member A
FBL	39	TCCACGTG	fibrillarin
FBL	48	CACGTGCG	fibrillarin
FDFT1	-298	GACACGTG	farnesyl-diphosphate farnesyltransferase 1
FDFT1	-289	CACGTGGG	farnesyl-diphosphate farnesyltransferase 1
FGFR4	42	GCCGCGTG	fibroblast growth factor receptor 4
FLVCR2	-19	GTCACGTG	feline leukemia virus subgroup C cellular receptor family, member 2
FLVCR2	-10	CACGTGTG	feline leukemia virus subgroup C cellular receptor family, member 2
GALNS	-157	GTCACGTG	galactosamine (N-acetyl)-6-sulfatase
GALNS	-148	CACGTGAC	galactosamine (N-acetyl)-6-sulfatase
GALNS	-129	CACGCGGC	galactosamine (N-acetyl)-6-sulfatase
GALNS	-15	GTCACGTG	galactosamine (N-acetyl)-6-sulfatase
GALNS	-6	CACGTGGC	galactosamine (N-acetyl)-6-sulfatase
GEMIN5	-104	CGCGTGGT	gem (nuclear organelle) associated protein 5
GPAM	-13	GCCACGCG	glycerol-3-phosphate acyltransferase, mitochondrial
GPN3	-238	AGCACGTG	GPN-loop GTPase 3

GPN3	-229	CACGTGGC	GPN-loop GTPase 3
GPX1	-21	GCCACGTG	glutathione peroxidase 1
GPX1	-12	CACGTGAC	glutathione peroxidase 1
HEATR2	-110	CCCACGTG	HEAT repeat containing 2
HEATR2	-101	CACGTGCG	HEAT repeat containing 2
HRSP12	-82	TCCACGTG	heat-responsive protein 12
HRSP12	-73	CACGTGGC	heat-responsive protein 12
HSF2	-233	TCCACGTG	heat shock transcription factor 2
HSF2	8	GTCACGTG	heat shock transcription factor 2
HSF2	17	CACGTGCG	heat shock transcription factor 2
HSP90AB1	-244	CGCGTGGT	heat shock protein 90kDa alpha (cytosolic), class B member 1
HSPH1	11	GGCACGTG	heat shock 105kDa/110kDa protein 1
HSPH1	20	CACGTGGA	heat shock 105kDa/110kDa protein 1
HTATIP2	-59	CGCACGTG	HIV-1 Tat interactive protein 2, 30kDa
HTATIP2	-50	CACGTGCT	HIV-1 Tat interactive protein 2, 30kDa
IQCG	-11	CCCACGTG	IQ motif containing G
ISM2	30	GCCGCGTG	isthmin 2
KCNH4	-120	ACCACGTG	potassium voltage-gated channel, subfamily H (eag-related), member 4
KCNH4	-111	CACGTGTG	potassium voltage-gated channel, subfamily H (eag-related), member 4
KIAA1715	-70	GTCACGTG	KIAA1715
KIAA1715	-61	CACGTGAC	KIAA1715
KIAA2013	-254	AACACGTG	KIAA2013
KIAA2013	-245	CACGTGTT	KIAA2013
LDHA	-13	CGCACGTG	lactate dehydrogenase A
LDHA	-4	CACGTGGA	lactate dehydrogenase A
LIN52	0	GTCACGTG	lin-52 DREAM MuvB core complex component
LIN52	9	CACGTGAC	lin-52 DREAM MuvB core complex component
LTV1	-7	TGCACGTG	LTV1 ribosome biogenesis factor
LTV1	2	CACGTGTG	LTV1 ribosome biogenesis factor
LYRM4	-100	GCCGCGTG	LYR motif containing 4
MACROD1	-107	CGCACGTG	MACRO domain containing 1
MACROD1	-98	CACGTGTG	MACRO domain containing 1
MAPRE2	-36	GCCACGTG	microtubule-associated protein, RP/EB family, member 2
MAPRE2	-27	CACGTGAC	microtubule-associated protein, RP/EB family, member 2
ME3	-163	ACCACGTG	malic enzyme 3, NADP(+)-dependent, mitochondrial
ME3	-154	CACGTGCA	malic enzyme 3, NADP(+)-dependent, mitochondrial
MED25	-116	CACGCGGT	mediator complex subunit 25
METTL25	-177	ACCACGCG	methyltransferase like 25
MICAL2	-124	CCCACGTG	microtubule associated monooxygenase, calponin and LIM domain containing 2
MICAL2	-115	CACGTGCG	microtubule associated monooxygenase, calponin and LIM domain containing 2
MIR4767	-92	CCCACGTG	microRNA 4767
MIR4767	-83	CACGTGCT	microRNA 4767
MIR4767	-23	GCCACGCG	microRNA 4767
MLX	-12	GTCACGTG	MLX, MAX dimerization protein
MLX	-3	CACGTGGG	MLX, MAX dimerization protein
MNT	-291	GCCACGTG	MAX network transcriptional repressor
MNT	-282	CACGTGTG	MAX network transcriptional repressor
MRPL36	-53	GCCACGTG	mitochondrial ribosomal protein L36
MRPL36	-44	CACGTGAC	mitochondrial ribosomal protein L36
MRPL52	-19	CGCGTGGC	mitochondrial ribosomal protein L52
MRPS18C	-79	CACGTGAC	mitochondrial ribosomal protein S18C
MRPS33	-94	CACGTGCA	mitochondrial ribosomal protein S33
MRPS9	-269	GCCACGTG	mitochondrial ribosomal protein S9
MRPS9	-260	CACGTGGA	mitochondrial ribosomal protein S9
MTX3	-26	CACGTGCC	metaxin 3
NDUFB2-AS1	-243	GCCACGTG	NDUFB2 antisense RNA 1
NDUFB2-AS1	-234	CACGTGAC	NDUFB2 antisense RNA 1
NDUFS8	-283	GCCGCGTG	NADH dehydrogenase (ubiquinone) Fe-S protein 8, 23kDa (NADH-coenzyme Q reductase)
NOL6	-212	GCCACGCG	nucleolar protein 6 (RNA-associated)
NOL6	30	CCCACGTG	nucleolar protein 6 (RNA-associated)
NOL6	39	CACGTGGG	nucleolar protein 6 (RNA-associated)
OSGEP	-102	CCCACGTG	O-sialoglycoprotein endopeptidase
OSGEP	-93	CACGTGGG	O-sialoglycoprotein endopeptidase
OSGEP	-81	GTCACGTG	O-sialoglycoprotein endopeptidase
OSGEP	-72	CACGTGAT	O-sialoglycoprotein endopeptidase
OSGEP	-20	ACCACGTG	O-sialoglycoprotein endopeptidase
OSGEP	-11	CACGTGAC	O-sialoglycoprotein endopeptidase
PACSIN3	-68	GCCGCGTG	protein kinase C and casein kinase substrate in neurons 3

PAGE1	1	CCCACGTG	P antigen family, member 1 (prostate associated)
PAGE1	10	CACGTGGA	P antigen family, member 1 (prostate associated)
PDHB	-218	GGCACGTG	pyruvate dehydrogenase (lipoamide) beta
PDHB	-209	CACGTGTG	pyruvate dehydrogenase (lipoamide) beta
PECR	-26	GTCACGTG	peroxisomal trans-2-enoyl-CoA reductase
PECR	-17	CACGTGGG	peroxisomal trans-2-enoyl-CoA reductase
PEPD	-271	CCCACGTG	peptidase D
PEPD	-262	CACGTGGT	peptidase D
PEX11G	-155	CACACGTG	peroxisomal biogenesis factor 11 gamma
PEX11G	-146	CACGTGTC	peroxisomal biogenesis factor 11 gamma
PFDN5	-190	CCCACGTG	prefoldin subunit 5
PFDN5	-181	CACGTGGA	prefoldin subunit 5
PFKFB3	-55	CCCACGTG	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
PFKFB3	-46	CACGTGGA	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
PIF1	-240	ACCACGCG	PIF1 5'-to-3' DNA helicase
PIGW	-36	CGCGTGGT	phosphatidylinositol glycan anchor biosynthesis, class W
PLA2G12A	-136	ACCACGTG	phospholipase A2, group XIIA
PLA2G12A	-127	CACGTGGG	phospholipase A2, group XIIA
PNO1	22	CGCACGTG	partner of NOB1 homolog (S. cerevisiae)
PNO1	31	CACGTGTT	partner of NOB1 homolog (S. cerevisiae)
POLR3G	16	GCCGCGTG	polymerase (RNA) III (DNA directed) polypeptide G (32kD)
PPID	-99	CACGCGGT	peptidylprolyl isomerase D
PRELID1	-20	GTCACGTG	PRELI domain containing 1
PRELID1	-11	CACGTGGC	PRELI domain containing 1
PRPF39	-28	CCCACGTG	pre-mRNA processing factor 39
PRPF39	-19	CACGTGAC	pre-mRNA processing factor 39
PSMD9	-74	CGCACGTG	proteasome (prosome, macropain) 26S subunit, non-ATPase, 9
PSMD9	-65	CACGTGTT	proteasome (prosome, macropain) 26S subunit, non-ATPase, 9
PUS7	-7	CGCACGTG	pseudouridylate synthase 7 (putative)
PUS7	2	CACGTGTG	pseudouridylate synthase 7 (putative)
PUSL1	-238	ACCGCGTG	pseudouridylate synthase-like 1
PUSL1	-229	CGCGTGGT	pseudouridylate synthase-like 1
RGS19	-251	TCCACGTG	regulator of G-protein signaling 19
RGS19	-242	CACGTGGA	regulator of G-protein signaling 19
RHOH	18	CGCACGTG	ras homolog family member H
RHOH	27	CACGTGCA	ras homolog family member H
RIOK1	32	CGCACGTG	RIO kinase 1
RIOK1	41	CACGTGGT	RIO kinase 1
RPL23AP7	-258	CACGTGCG	ribosomal protein L23a pseudogene 7
SETD6	-30	ACCACGTG	SET domain containing 6
SETD6	-21	CACGTGGT	SET domain containing 6
SGTB	-223	CGCGTGGT	small glutamine-rich tetratricopeptide repeat (TPR)-containing, beta
SGTB	-120	CCCACGTG	small glutamine-rich tetratricopeptide repeat (TPR)-containing, beta
SGTB	-111	CACGTGCG	small glutamine-rich tetratricopeptide repeat (TPR)-containing, beta
SHB	-33	CACGCGGC	Src homology 2 domain containing adaptor protein B
SLC16A6	-245	TCCACGTG	solute carrier family 16, member 6
SLC16A6	-236	CACGTGGG	solute carrier family 16, member 6
SLC38A5	-212	GACACGTG	solute carrier family 38, member 5
SLC38A5	-203	CACGTGGC	solute carrier family 38, member 5
SMNDC1	-59	GTCACGTG	survival motor neuron domain containing 1
SRFBP1	40	CACGTGCA	serum response factor binding protein 1
SYP	-108	CGCGTGGC	synaptophysin
TADA1	-44	ATCACGTG	transcriptional adaptor 1
TADA1	-35	CACGTGAC	transcriptional adaptor 1
TCEAL3	-16	CACGTGAC	transcription elongation factor A (SII)-like 3
TERT	-186	CCCACGTG	telomerase reverse transcriptase
TERT	-177	CACGTGGC	telomerase reverse transcriptase
TERT	22	CGCACGTG	telomerase reverse transcriptase
TERT	31	CACGTGGG	telomerase reverse transcriptase
TMEM132E	34	TGCACGTG	transmembrane protein 132E
TMEM132E	43	CACGTGCA	transmembrane protein 132E
TMEM168	-270	CGCGTGGC	transmembrane protein 168
TRAP1	-68	GCCACGCG	TNF receptor-associated protein 1
TRIP6	-211	GCCACGTG	thyroid hormone receptor interactor 6
TRIP6	-202	CACGTGCA	thyroid hormone receptor interactor 6
TRIP6	-29	GCCACGTG	thyroid hormone receptor interactor 6
TRIP6	-20	CACGTGAC	thyroid hormone receptor interactor 6
TUBA1C	-54	GACACGTG	tubulin, alpha 1c
TUBA1C	-45	CACGTGGC	tubulin, alpha 1c
UBXN8	39	CACGTGGG	UBX domain protein 8
USP2	-104	TCCACGTG	ubiquitin specific peptidase 2

USP2	-95	CACGTGGT	ubiquitin specific peptidase 2
UTP11L	-234	GCCACGTG	UTP11-like, U3 small nucleolar ribonucleoprotein (yeast)
UTP11L	-225	CACGTGTT	UTP11-like, U3 small nucleolar ribonucleoprotein (yeast)
UTP20	-34	TCCACGTG	UTP20, small subunit (SSU) processome component, homolog (yeast)
UTP20	-25	CACGTGAC	UTP20, small subunit (SSU) processome component, homolog (yeast)
UTP20	27	CGCACGTG	UTP20, small subunit (SSU) processome component, homolog (yeast)
ZBTB49	-119	AGCACGTG	zinc finger and BTB domain containing 49
ZBTB49	-110	CACGTGGA	zinc finger and BTB domain containing 49
ZNF890P	1	GACACGTG	zinc finger protein 890, pseudogene
ZNF890P	10	CACGTGTG	zinc finger protein 890, pseudogene

Supplementary Table 24 Genes that are upregulated at 2 hours after GSK-J4 treatment and also have their H3K4me3 marks differentially enriched as compared to DMSO control, and have a Jun binding site around their transcription start site.

Gene ID	Distance to TSS	Binding Sequence	Gene Name
ARG2	-115	GTGACGTCATCG	arginase 2
ARG2	-106	GGGTGACGTCAT	arginase 2
ARG2	-7	GTGACGTCACGG	arginase 2
ARG2	2	CGGTGACGTCAC	arginase 2
ASPHD1	-270	GTGACGTCAGGG	aspartate beta-hydroxylase domain containing 1
ASPHD1	-261	GAGTGACGTCAG	aspartate beta-hydroxylase domain containing 1
BACH2	-83	GTGACGTCAGCG	BTB and CNC homology 1, basic leucine zipper transcription factor 2
BACH2	-74	GGGTGACGTCAG	BTB and CNC homology 1, basic leucine zipper transcription factor 2
FAM122C	-64	GTGACGTCATTG	family with sequence similarity 122C
FAM122C	-55	AAGTGACGTCAT	family with sequence similarity 122C
FRAT1	-40	GTGACGTCACCC	frequently rearranged in advanced T-cell lymphomas 1
FRAT1	-31	ACGTGACGTCAC	frequently rearranged in advanced T-cell lymphomas 1
ING4	-262	GTGACGTCATAC	inhibitor of growth family, member 4
ING4	-253	CAGTGACGTCAT	inhibitor of growth family, member 4
ISL2	-62	TGTTGACGTCAT	ISL LIM homeobox 2
MED20	-76	AGATGATGTCAG	mediator complex subunit 20
MSI2	-192	GTGACGTCACCG	musashi RNA-binding protein 2
MSI2	-183	GCGTGACGTCAC	musashi RNA-binding protein 2
PTPRU	-43	CTGACGTCAGCG	protein tyrosine phosphatase, receptor type, U
RAB1A	-148	ATGACATCACTA	RAB1A, member RAS oncogene family
RAB1A	-139	TGATGACATCAC	RAB1A, member RAS oncogene family
SKIL	-25	GTGACGTCAGAG	SKI-like proto-oncogene
SKIL	-16	CGGTGACGTCAG	SKI-like proto-oncogene
SPAG4	-121	GTGACGTCACGG	sperm associated antigen 4
SPAG4	-112	AGGTGACGTCAC	sperm associated antigen 4
SPAG4	-58	GTGACGTCAGCA	sperm associated antigen 4
SPAG4	-49	CTGTGACGTCAG	sperm associated antigen 4
TM2D1	-81	CGATGATGTCAG	TM2 domain containing 1
ZNF200	-26	GTGACGTCACAG	zinc finger protein 200
ZNF200	-17	TTGTGACGTCAC	zinc finger protein 200

Supplementary Table 25 Genes that are upregulated at 6 hours after GSK-J4 treatment and also have their H3K4me3 marks differentially enriched as compared to DMSO control, and have an ATF4 binding site around their transcription start site.

Gene ID	Distance to TSS	Binding Sequence	Gene Name
ASF1A	-271	CTGATGTAAA	anti-silencing function 1A histone chaperone
ATF3	-62	CTTGCATCAC	activating transcription factor 3
C9orf91	-137	CTGATGAAAC	chromosome 9 open reading frame 91
CHAC1	-168	GTTGCATCAC	ChaC, cation transport regulator homolog 1 (E. coli)
DGCR6	-39	ATGACGCAAG	DiGeorge syndrome critical region gene 6
FGL2	-78	GTGATGCAAA	fibrinogen-like 2
NCR1	-97	CTGATGAAAG	natural cytotoxicity triggering receptor 1
NR3C1	-86	GTTGCATCGT	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)
PCGF5	-209	ATGATGTAAA	polycomb group ring finger 5
PTEN	-267	CTTTCATCAT	phosphatase and tensin homolog
PTK2B	-154	ATGATGTAAT	protein tyrosine kinase 2 beta
RPL9	-23	ATGACGCAAT	ribosomal protein L9
TMEM67	-241	CTGATGTAAC	transmembrane protein 67
VPS4B	-2	ATGACGAAAT	vacuolar protein sorting 4 homolog B (S. cerevisiae)
ZBED3	-166	ATTGCGTCAT	zinc finger, BED-type containing 3

Supplementary Table 26 Genes that are upregulated at 24 hours after GSK-J4 treatment and also have their H3K4me3 marks differentially enriched as compared to DMSO control, and are annotated by the GO category "regulation of transcription".

Gene Symbol	24h DMSO	2h GSK-J4	6h GSK-J4	24h GSK-J4	Gene name
AFF1	14.7175	17.4893	22.2473	22.7682	AF4/FMR2 family, member 1
BCL7A	4.34161	8.38007	3.87748	8.70068	B-cell CLL/lymphoma 7A
BRF2	8.32674	9.86866	10.6901	13.2103	BRF2, subunit of RNA polymerase III transcription initiation factor, BRF1-like
BUD31	25.8321	35.7391	35.3757	71.456	BUD31 homolog (S. cerevisiae)
CEBPB	10.9972	13.5656	28.0617	44.6011	CCAAT/enhancer binding protein (C/EBP), beta
CEBPG	10.4925	11.9838	14.5257	23.1295	CCAAT/enhancer binding protein (C/EBP), gamma
CRTC1	1.77689	3.819	3.47183	5.19332	CREB regulated transcription coactivator 1
CITED4	0	1.39619	1.53433	4.59042	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 4
DNMT3B	4.6257	8.721	9.59987	8.46979	DNA (cytosine-5-)-methyltransferase 3 beta
DDIT3	2.15192	2.49275	10.5715	23.3558	DNA-damage-inducible transcript 3
EAF2	7.31704	4.33599	4.08429	12.2194	ELL associated factor 2
GATAD1	1.14225	2.5386	1.48772	3.81514	GATA zinc finger domain containing 1
JAK2	0.589253	0.897903	1.31566	1.72208	Janus kinase 2
JDP2	0	0.283358	1.6352	1.04828	Jun dimerization protein 2
KLF11	2.84611	5.23187	2.26951	5.77154	Kruppel-like factor 11
LANCL2	0	2.80974	2.81924	3.01236	LanC lantibiotic synthetase component C-like 2 (bacterial)
MXD3	4.2607	4.47349	5.60345	8.90499	MAX dimerization protein 3
MDM2	9.27024	11.6496	10.5791	18.5774	Mdm2, p53 E3 ubiquitin protein ligase homolog (mouse)
PSIP1	5.66435	7.51139	11.255	25.7748	PC4 and SFRS1 interacting protein 1
PHF1	0.753411	2.80659	7.22544	4.80289	PHD finger protein 1
PHF20	3.48458	8.10566	6.43372	6.0432	PHD finger protein 20
POU6F1	0	0.451061	0.371747	1.66838	POU class 6 homeobox 1
SCAND1	19.2354	28.949	27.4389	33.9075	SCAN domain containing 1
SERTAD3	2.48434	5.46805	6.24028	8.29756	SERTA domain containing 3
SETD8	12.5273	12.7784	16.4717	19.6199	SET domain containing (lysine methyltransferase) 8
SIN3B	12.9237	19.2786	15.3713	22.7357	SIN3 transcription regulator homolog B (yeast)
SMAD4	14.6342	17.8928	30.4621	44.6761	SMAD family member 4
TOX2	13.5193	12.6369	25.5811	28.9503	TOX high mobility group box family member 2
TSC22D2	9.12279	15.0426	14.696	15.5847	TSC22 domain family, member 2
ATF3	0.870299	0.829187	8.03133	3.33722	activating transcription factor 3
ATF4	164.599	181.851	272.327	296.528	activating transcription factor 4 (tax-responsive enhancer element B67)
ACVR1B	0.343234	1.05773	1.93747	1.44908	activin A receptor, type IB
BHLHE40	19.5634	17.0091	22.1353	32.6707	basic helix-loop-helix family, member e40
CREG1	32.0291	22.5258	27.9635	48.6881	cellular repressor of E1A-stimulated genes 1
CHD2	8.08126	13.216	32.4605	24.4319	chromodomain helicase DNA binding protein 2
C14orf169	7.28768	11.8901	9.24522	12.1703	chromosome 14 open reading frame 169
CCDC101	4.7309	18.6898	30.8085	9.21731	coiled-coil domain containing 101
DMTF1	6.447	5.85953	8.93436	13.3663	cyclin D binding myb-like transcription factor 1
CCNL1	25.9651	20.3062	20.3598	27.8753	cyclin L1
CDK7	2.27665	3.46916	4.4478	4.75248	cyclin-dependent kinase 7
CDK8	4.7768	6.79361	11.1987	8.77492	cyclin-dependent kinase 8
CDKN1C	4.8083	9.73966	13.383	9.71373	cyclin-dependent kinase inhibitor 1C (p57, Kip2)
DVL3	20.0293	23.4038	26.5502	31.1706	dishevelled, dsh homolog 3 (Drosophila)
DR1	19.9258	26.9892	26.9716	48.6659	down-regulator of transcription 1, TBP-binding (negative cofactor 2)
ELL	29.2159	44.7303	33.1036	34.759	elongation factor RNA polymerase II
EPC1	4.61944	3.38916	5.47697	9.64483	enhancer of polycomb homolog 1 (Drosophila)
FAM120B	3.94449	6.96473	6.48629	8.64583	family with sequence similarity 120B
FOXD1	0	2.9448	7.68589	4.84098	forkhead box D1
FOXK1	15.9462	20.6297	19.7039	26.5162	forkhead box K1
FOXM1	24.7453	33.9991	33.6634	48.578	forkhead box M1
FOXN3	9.91886	9.64454	16.1449	15.9945	forkhead box N3
GTF2H2	0	0	1.43341	2.14425	general transcription factor IIH, polypeptide 2, 44kDa
GDF7	0	0	0	0.697605	growth differentiation factor 7
HDAC9	8.4519	5.33603	13.2721	17.9948	histone deacetylase 9
IKBKB	13.3773	14.8509	15.4996	21.0383	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta
INSR	9.67784	11.5439	13.5399	15.875	insulin receptor
ITGB3BP	7.41649	8.16225	12.4195	21.675	integrin beta 3 binding protein (beta3-endonexin)

IRAK1BP1	0	2.12257	1.74944	19.6274	interleukin-1 receptor-associated kinase 1 binding protein 1
JUND	51.1597	56.5303	58.1155	87.6845	jun D proto-oncogene
JUN	5.27773	8.7733	14.462	14.4225	jun proto-oncogene
L3MBTL4	2.6543	3.29561	2.46933	6.64897	l(3)mbt-like 4 (Drosophila)
KDM4C	2.68159	4.44629	7.16865	8.68126	lysine (K)-specific demethylase 4C
MAML2	4.58505	3.00816	7.99788	11.4855	mastermind-like 2 (Drosophila)
MED13	17.9497	12.6011	14.1684	19.0091	mediator complex subunit 13
MED8	11.729	18.3572	26.5	30.8945	mediator complex subunit 8
MAP3K13	1.34902	5.20593	2.16332	7.32249	mitogen-activated protein kinase kinase kinase 13
MALT1	15.9919	20.8189	19.7711	25.7411	mucosa associated lymphoid tissue lymphoma translocation gene 1
MEF2A	7.67842	9.58146	10.5705	12.036	myocyte enhancer factor 2A
NR4A3	0	0.637364	0.700442	1.38112	nuclear receptor subfamily 4, group A, member 3
NFX1	16.7035	24.1851	19.9148	28.2614	nuclear transcription factor, X-box binding 1
PCGF5	0.876117	6.72011	3.9399	13.7879	polycomb group ring finger 5
RUNX1	1.76348	2.62467	0.99497	3.47214	runt-related transcription factor 1
ZNF532	0.956647	1.53872	2.00246	3.39488	zinc finger protein 532
SLC30A9	1.46855	3.81249	6.5578	9.40109	solute carrier family 30 (zinc transporter), member 9
SRA1	7.04646	14.4783	17.1129	15.8087	steroid receptor RNA activator 1
SUPT3H	0.742895	1.00629	2.48831	3.33686	suppressor of Ty 3 homolog (S. cerevisiae)
TIGD3	0	0.287394	0.473743	0.708673	tigger transposable element derived 3
TCF19	32.028	38.691	35.6799	49.0225	transcription factor 19
TLE4	5.45549	5.76869	9.74779	9.5932	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)
TRIB3	24.6653	26.5161	47.1767	58.467	tribbles homolog 3 (Drosophila)
TSG101	12.2334	20.8116	12.9387	21.3582	tumor susceptibility gene 101
VEGFA	23.3415	34.9816	46.7251	44.8787	vascular endothelial growth factor A
VEZF1	6.08348	5.72222	7.73472	12.1348	vascular endothelial zinc finger 1
WNT1	0	0	0	1.20095	wingless-type MMTV integration site family, member 1
ZBTB2	12.9803	17.5816	11.5927	9.53784	zinc finger and BTB domain containing 2
ZFHX3	17.1503	9.59732	10.9058	40.3781	zinc finger homeobox 3
ZNF10	0	1.20709	1.98978	3.57182	zinc finger protein 10
ZNF138	1.35309	2.06188	3.02121	3.73992	zinc finger protein 138
ZNF214	0	0	0.338926	1.014	zinc finger protein 214
ZNF217	29.0244	36.984	41.7959	49.5802	zinc finger protein 217
ZNF226	5.06463	6.76241	2.82681	18.0173	zinc finger protein 226
ZNF233	0.584339	2.57346	0.978746	1.4646	zinc finger protein 233
ZNF235	0.501144	2.88487	3.07706	2.09226	zinc finger protein 235
ZNF263	4.85715	8.95411	7.49439	8.1861	zinc finger protein 263
ZNF276	9.4997	15.9619	15.3398	17.281	zinc finger protein 276
ZNF34	3.4516	1.16879	7.21656	6.4891	zinc finger protein 34
ZNF362	2.57818	2.26987	2.30257	4.73609	zinc finger protein 362
ZNF395	4.23205	9.80388	9.70335	7.06762	zinc finger protein 395
ZNF397	4.19422	5.81342	4.10691	7.50462	zinc finger protein 397
ZNF408	5.21398	10.0893	10.3946	9.3296	zinc finger protein 408
ZNF429	0	0.7045	1.16131	0.579067	zinc finger protein 429
ZNF501	0	0	1.15803	0.866144	zinc finger protein 501
ZNF510	2.41935	2.86736	3.03852	5.05037	zinc finger protein 510
ZNF518A	1.39242	1.88569	4.30487	4.15241	zinc finger protein 518A
ZNF594	0	2.60334	3.21853	4.27965	zinc finger protein 594
ZNF596	0	0	0	0.935496	zinc finger protein 596
ZNF623	6.39736	7.35204	7.13978	12.2915	zinc finger protein 623
ZNF641	1.07814	1.61816	1.02578	2.40067	zinc finger protein 641
ZNF7	12.4221	33.9482	10.6048	34.2087	zinc finger protein 7
ZNF768	19.1319	35.7516	36.3884	49.1085	zinc finger protein 768
ZNF772	2.05785	1.8914	4.59478	4.17317	zinc finger protein 772
ZNF774	0	2.47894	1.02158	1.52818	zinc finger protein 774
ZNF823	2.731	4.39268	2.28662	7.98129	zinc finger protein 823
ZNF837	0	0	0	1.44602	zinc finger protein 837
ZFPM1	4.92737	4.08916	8.98738	12.027	zinc finger protein, FOG family member 1

Supplementary Table 27 Genes that are upregulated at 24 hours after GSK-J4 treatment and also have their H3K4me3 marks differentially enriched as compared to DMSO control, and are annotated by the GO category "regulation of apoptosis".

Gene Symbol	24h DMSO	2h GSK-J4	6h GSK-J4	24h GSK-J4	Gene name
AKAP13	8.82959	7.76593	14.3427	16.4689	A kinase (PRKA) anchor protein 13
BCL2	14.747	31.9886	41.776	43.4192	B-cell CLL/lymphoma 2
BCL2L1	4.44682	9.46497	8.87151	11.6697	BCL2-like 1
CEBPB	10.9972	13.5656	28.0617	44.6011	CCAAT/enhancer binding protein (C/EBP), beta
CEBPG	10.4925	11.9838	14.5257	23.1295	CCAAT/enhancer binding protein (C/EBP), gamma
CD38	0	2.35939	2.59283	1.93931	CD38 molecule
DDIT3	2.15192	2.49275	10.5715	23.3558	DNA-damage-inducible transcript 3
JAK2	0.589253	0.897903	1.31566	1.72208	Janus kinase 2
RAB27A	8.61971	12.6305	9.52502	14.3896	RAB27A, member RAS oncogene family
SCAND1	19.2354	28.949	27.4389	33.9075	SCAN domain containing 1
TM2D1	4.28392	7.09821	7.97114	8.34685	TM2 domain containing 1
TRAIP	9.16615	13.811	9.63205	16.1485	TRAF interacting protein
ACVR1B	0.343234	1.05773	1.93747	1.44908	activin A receptor, type IB
ASNS	5.81243	13.157	29.9796	30.7771	asparagine synthetase (glutamine-hydrolyzing)
CAPN10	4.54386	7.70121	7.18282	7.58369	calpain 10
CTSB	5.43767	11.5591	7.70501	9.08094	cathepsin B
CLN3	15.1944	26.6742	26.6313	27.999	ceroid-lipofuscinosis, neuronal 3
CUL4A	12.1136	20.5544	16.4925	25.2921	cullin 4A
CDKN1A	6.29429	8.2912	10.1002	10.8599	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
DAPK2	0	2.92596	1.03354	1.03072	death-associated protein kinase 2
DDAH2	13.2573	14.3654	12.5801	27.6744	dimethylarginine dimethylaminohydrolase 2
FURIN	5.64707	8.72081	6.76457	9.17061	furin (paired basic amino acid cleaving enzyme)
HERPUD1	40.9777	42.454	76.2802	97.3904	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1
IKBKB	13.3773	14.8509	15.4996	21.0383	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta
IGF2R	13.3105	12.6659	16.1762	20.6809	insulin-like growth factor 2 receptor
ITGB3BP	7.41649	8.16225	12.4195	21.675	integrin beta 3 binding protein (beta3-endonexin)
JUN	5.27773	8.7733	14.462	14.4225	jun proto-oncogene
MBD4	13.25	7.40035	16.3548	20.8522	methyl-CpG binding domain protein 4
MALT1	15.9919	20.8189	19.7711	25.7411	mucosa associated lymphoid tissue lymphoma translocation gene 1
NOD1	1.74376	3.07047	3.11471	6.11533	nucleotide-binding oligomerization domain containing 1
NUDT2	13.7811	24.2929	17.5961	23.0139	nudix (nucleoside diphosphate linked moiety X)-type motif 2
PMAIP1	12.0875	16.3724	25.5426	43.2557	phorbol-12-myristate-13-acetate-induced protein 1
PTEN	12.3277	14.5162	23.1287	26.6708	phosphatase and tensin homolog
PLCG2	10.6703	15.948	17.252	16.8977	phospholipase C, gamma 2 (phosphatidylinositol-specific)
PLEKHF1	1.94799	3.9578	0.543673	4.87969	pleckstrin homology domain containing, family F (with FYVE domain) member 1
PSENEN	19.3763	60.145	79.3142	76.0321	presenilin enhancer 2 homolog (C. elegans)
PSMG2	29.2159	24.7329	37.1459	48.79	proteasome (prosome, macropain) assembly chaperone 2
RRM2B	1.58944	2.90339	3.72642	4.37406	ribonucleotide reductase M2 B (TP53 inducible)
STK3	1.14038	3.16547	4.7393	6.37307	serine/threonine kinase 3
ROCK1P1	0	0	0.370136	0.553687	Rho-associated, coiled-coil containing protein kinase 1 pseudogene 1
SOS2	3.80916	3.07583	6.05158	8.07391	son of sevenless homolog 2 (Drosophila)
SFN	0	0	0	1.12288	stratifin
SOCS3	2.36547	4.806	6.2718	9.382	suppressor of cytokine signaling 3
TP53INP1	0	23.0083	6.32119	104.015	tumor protein p53 inducible nuclear protein 1
TP53I3	0	0.357151	2.22413	4.78988	tumor protein p53 inducible protein 3
VEGFA	23.3415	34.9816	46.7251	44.8787	vascular endothelial growth factor A
ZMAT3	1.53315	1.61542	1.71185	2.98714	zinc finger, matrin-type 3