Verification and analysis of methylation profile assay data from RA CD4+ T cells by Bisulphite Sequencing

Submitted in accordance with the requirements for the degree of

Masters of Research (MRes)

University of Leeds

Faculty of Medicine, Dentistry and Health Graduate School

June 2018

Joshua Hugh Rogers

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement

©2018 The University of Leeds

1 Introduction

1.1 Defining epigenetics

1.2 Epigenetic modifications

1.2.1 Histone modification

1.2.2 DNA methylation

1.3 Environmental factors and epigenetic changes

1.3.1 Gestational effects

1.3.2 Postnatal effects

1.3.3 Transgenerational effects

1.3.4 Diet, lifespan and ageing

1.4 Epigenetics in disease

1.4.1 Epigenetics in cancer

1.4.2 Epigenetics in autoimmune conditions

1.4.2.1 Systemic lupus erythematosus

1.4.2.2 Sjogren's syndrome

1.4.2.3 Systemic sclerosis

1.4.2.4 Rheumatoid arthritis

1.5 Evaluation

1.5.1 Choosing a candidate cell type

1.5.2 Aims and Objectives

1.5.3 Hypothesis

2 Materials & methods

2.1 Ethics and IACON

2.2 Isolation by negative selection of CD4+ T-cells from PBMCs

2.3 CD4+ purity check

2.4 DNA extraction

2.5 Bisulphite conversion treatment

2.6 Gene amplification

2.7 Sequencing

2.8 Software analysis

2.8.1 Sequencing analysis

2.8.2 Initial gene selection and primer design

3 Results

3.1 DNA methylation data

3.1.1 TNF signalling cascade analysis

3.2 Primer design

3.2.1 PCR optimisation

3.3 DNA sequencing

4 Discussion

4.1 Candidate gene selection

4.1.1 Demonstrating TNF's relevance in RA pathology

4.1.2 NFkB activation of the TNF-dependant cytokine cascade in RA

4.1.2.1 Interleukin 1

4.1.2.2 Interleukin 6

4.1.2.3 Granulocyte-macrophage colony-stimulating factor

4.1.3 Other potential candidates

4.2 Candidate gene region selection

4.3 Confirming the candidate gene's veracity

4.4 Variance of TNF methylation

4.5 TNF cascade

4.5.1 TNF receptors

4.5.2 Downstream effects of TNF-TNFR signalling

4.5.2.1 TNF cascade in HC

4.5.2.2 P38

4.5.2.3 c-Jun N-terminal kinases

4.5.2.4 NFkB

4.5.2.5 Non-canonical activation of NFkB

4.5.3 TNF cascade-related methylation aberrations in RA CD4+ T cells

4.5.3.1 Inhibition of NFkB with DAXX and TRAF1

4.5.3.2 CD30

4.5.3.3 LIGHT/LTα/HVEM role in T-cell activation

4.5.3.4 RANK induction via NIK
4.5.3.5 cFOS activation via the TCR
4.5.3.6 NFkB signal modulation

4.6 TNF gene coexpression network
4.6.1 LNPEP
4.6.2 RPS18
4.6.3 GPSM3
4.6.4 DAXX
4.6.5 TLR9
4.6.6 ZNF217
4.6.7 TMC6
4.6.8 PTPRCAP
4.6.9 CORO1A
4.6.10 ABI3

4.7 The broader cause of hypomethylation and the MHC region
4.7.1 The link between coexpression and hypomethylation
4.8 Conclusion

5 Limitations and areas of further study
5.1 Limitations of methodology
5.2 Future steps

6 References

7 Appendix
7.1 Ethical approval
7.2 Raw data and ranking methodology
7.3 Designed primers and products
7.4 Gene definitions

1. Introduction

1.1 Defining epigenetics

The study of epigenetics in some form has existed since the 17th century in the form of the study of epigenesis which defined as "the process by which plants, animals and fungi develop from a seed, spore or egg through a sequence of steps in which cells differentiate and organs form" (1). Epigenesis was the school of thought that opposed preformatism, which was largely refuted in favour of epigenesis in 1762. In 1942, from epigenesis, the word "epigenetic" was coined, and used as an adjective to describe the differentiation of totipotent stem cells in embryonic development. However the precise definition of what epigenetics studied, epigenetic traits, wasn't agreed upon until 2008. "Epigenetic trait" was defined there as a "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (2).

According to this definition, epigenetics covers a large amount of modifications. These modifications includes alterations via: covalent modifications, RNA transcripts, mRNA, microRNA and sRNA. Covalent modification of histones and DNA play a central role in epigenetic inheritance and as a result, the word "epigenetics" is sometimes used as a synonym for these processes. It should be noted that this is misleading as chromatin remodelling is not always inherited, and not all epigenetic inheritance involves chromatin remodelling (3). Due to the importance of covalent modifications in epigenetics and relevance to this thesis they will be the focus in this section.

1.2 Epigenetic modifications

1.2.1 Histone modifications

Histone proteins are responsible for organising and packaging DNA. They do this by forming histone-DNA structural octamers called 'nucleosomes'. These nucleosomes form a larger complex called 'chromatin'.   The histone complex is made up of core histones called H2A, H2B, H3, H4, and a linker histone called H1.

The principle mechanism through which histones can bind to the DNA is because histones have a net positive charge and DNA has a net negative charge – they thus attract one another with ionic forces. As a result histone modifications typically exert their effects through modification of the histone's charge to better attract or repel the DNA strand. The aforementioned modifications occur primarily on their N-terminal tails, and in their globular domains (4, 5). These modifications include methylation, acetylation, citrullination, phosphorylation, ubiquitination, SUMOylation, and ADP-ribosylation.

1.2.2 DNA Methylation

DNA methylation is the process where the cytosine in a cytosine-phosphate-guanine motif (CpG) gets turned into 5-methylcytosine (5mC). The process is carried out by DNA methyltransferases (DNMTs). Three DNMTs (DNMT1, DNMT3a and DNMT3b) are required for establishment and maintenance of DNA methylation patterns. DNMT1 is responsible for the maintenance of established patterns of DNA methylation, while DNMT3a and 3b mediate establishment of new DNA methylation patterns. The overall effect of DNA methylation is gene suppression, however there are a few examples of DNA methylation upregulating expression. DNA methylation achieves transcriptional suppression by directly preventing transcription factor binding and leads to changes in chromatin structure that restrict access of transcription factors to the gene promoter (6, 7) –

therefore downregulating transcription and gene expression.

Due to their susceptibility to methylation, CpGs are highly mutagenic and through evolutionary processes genomes are consequently CpG deficient (7, 8). However CpGs do occur in high density in regions across the genome, these high density CpG regions are called CpG islands (CGIs). CGIs have been defined as being regions less than 500bp with a GC percentage greater than 60% and an observed:expected ratio of CpGs greater than 0.6.

In humans approximately 70% of known gene promoters are associated with a CGI, making this the most common promoter type in the vertebrate genome (8, 9). Nearly all of the housekeeping genes, as well as a proportion of tissue-specific genes and developmental regulator genes (10,11). More recent studies uncovered a large class of CGIs that are remote from annotated transcription start sites (TSSs), but still show evidence for promoter function (12,13). These findings emphasise the strong correlation between CGIs and transcription initiation.

Methylated CpGs can also become demethylated. The process of DNA demethylation can occur either actively or passively via a cyclic enzymatic cascade.

Passive DNA demethylation occurs during DNA replication, the DNA methylome maintenance machinery, DNMT1/ UHRF1, is responsible for copying the CpG methylation pattern by methylating the unmodified cytosine in the nascent DNA strand (14,15). Ergo, passive loss of 5mC can therefore be achieved through multiple cycles of DNA replication in the presence of dysfunctional methylome maintenance machinery. This demethylation process is termed ‘replication-dependent passive DNA demethylation’, an example of such is the erasure of 5mC from the maternal genome during mouse preimplantation development (16,17).

Active DNA demethylation is the process where 5mC can also be rapidly demethylated independently of DNA replication. Such change involves active DNA demethylation mechanisms, which require one or more enzymatic reactions to modify or remove 5mC - these reactions are catalysed by TET enzymes (18). The step-wise cytosine modification pathway that includes cytosine methylation by DNMTs and step-by-step oxidation of 5mC by TET proteins. DNMTs use S-adenosylmethionine (SAM) as a methyl donor to catalyse methylation at the 5-position of cytosine, yielding S-adenosylhomocysteine (SAH). As Fe^{2+}/a-ketoglutarate (a-KG)-dependent dioxygenases, TET proteins then use a base-flipping mechanism to flip the target base out of the DNA into the catalytic site (19-21). The active site, Fe^{2+} is bound by conserved amino acid residues in TETs. TET enzymes use molecular oxygen as a substrate to catalyse oxidative decarboxylation of a-KG, yielding CO2, enzyme-bound succinate, and a reactive highvalent Fe^{4+}-oxo intermediate. The TET-bound Fe^{4+}-oxo intermediate reacts with 5mC/ 5hmC/5fC to generate 5hmC/5fC/5caC. In mammals, base excision repair of 5fC/5caC completes the DNA demethylation process (18).

1.3 Environmental factors and epigenetic changes

This section will address how epigenetic changes can be brought about by diet, chemical pollutants, temperature changes and other external stresses.

For instance, animal models have shown that nutrition and exposure to environmental factors during development can cause region-specific changes in the epigenome (22).

For example, in humans, there is growing awareness that nutrition and exposure to toxic components may have long-term phenotypic effects (23). As is discussed below, human studies on the gestational effects of nutrition have thus far identified DNA methylation changes that are relatively small (24-26). However the relevance of these minor changes, and whether they are a consequence or cause of the observed developmental and metabolic aberrations is still something that needs to be addressed by future studies.

1.3.1 Gestational effects.

The degree of impact that environmental factors have on the mammalian epigenome seems to depend on what developmental stage the organism is at. Several lines of evidence suggest that the gestational period is particularly susceptible to epigenetic insult, and that the environment can have different effects on the embryo and placenta.

The early conceptional and embryonic stages of development seem particularly sensitive to nutritional influence upon epigenetic modifications. Studies on the ‘Dutch hunger winter’, at the end of the Second World War, indicate that famine exposure in the early conceptional period led to adverse metabolic and mental phenotypes in the following generation. Initial studies reported a minor, but significant, decrease in DNA methylation at a differentially methylated region (DMR) in the imprinted insulin-like growth factor 2 (IGF2) gene (27). Further research (24-26) identified small alterations in DNA methylation at other imprinted genes such as insulin (INS), maternally expressed gene 3 (MEG3) and guanine nucleotide binding protein a-stimulating (GNAS). Importantly, epigenetic alterations were also found at some loci that are involved in growth and implicated in metabolic disease, such as ATP-binding cassette A1 (ABCA1), leptin (LEP) and interleukin 10 (IL10). It should also be noted that famine exposure later during gestation had no significant effects on DNA methylation (24).

Another example of nutrition-induced epigenetic changes can be seen in the increased incidence of type 2 diabetes in rats when exposed to sub-optimal nutrition during early development. This was associated with the epigenetic repression, in pancreatic islets, of hepatocyte nuclear factor 4a (Hnf4a), which is involved in regulating β-cell differentiation and glucose homeostasis (28).

1.3.2 Postnatal effects

Environmental conditions can affect the epigenome postnatally as well, for instance one epidemiological study identified a possible link between particulate air pollution and genome-wide DNA methylation in blood cells (29). Genome-wide methylation was also altered in the blood of individuals that had been exposed to the carcinogen, benzene.

Epidemiological studies on skin and other tissues identified locus-specific DNA methylation changes associated with various environmental factors, including chronic exposure to sunlight, asbestos and tobacco smoke, consumption of alcohol and use of hair-dye (30-32).

However a possible confounding factor in these studies is that environmental stress can alter the relative abundance of different cell types in the blood. Such shifts could lead to altered levels of measured DNA methylation for specific genes. In addition, DNA methylation itself is associated with cell fate determination in haematopoiesis, and as a result alterations to this methylation pattern could affect the cell populations that constitute peripheral blood (33). To avoid such confounding effects, ideally individual cell types and tissues should be explored as well, if experimental design allows it.

In these and similar studies, it is also unknown whether the observed epigenetic changes are directly mediated by the toxic compound under study. Because not all tissues of an adult organism are equally exposed to chemical pollutants, and because sensitivity depends on the tissue type, epigenetic alterations in adults are probably tissue-specific in nature. However most human studies so far have only considered peripheral blood.

The reason for the difference in sensitivity to epigenetic alterations between postnatal could be linked to the level of cell pluripotency. Several studies have shown that the epigenomes of fully differentiated cells seem to be relatively stable in comparison to the epigenome of mouse embryonic stem cells which are particularly sensitive to in vitro culture conditions (34-37). Also, culture of mammalian pre-implantation embryos can readily affect the establishment and maintenance of DNA methylation, particularly at imprinted genes (34-37).

This increased sensitivity towards extrinsic factors inducing epigenetic changes is thought to be to do with the fact that early embryonic cells have higher expressions of genes involved with the regulation of the epigenome – DNA methyltransferases (DNMT). As a result, the increased amount of these regulatory machineries in early embryos, stem cells and germ cells is a likely cause of the susceptibility of these pluripotent cell types to environmental signals (34-37). Taken together it can be seen that the degree to which a cell is epigenetically modified by external cues depends upon how epigenetically malleable the cells at that time are.

1.3.3 Transgenerational inheritance

There has been research conducted in order to answer to question of whether environmentally-induced epigenetic alterations can be passed down from one progeny to the next (38). During gametogenesis, the epigenome is reprogrammed genome-wide in mammals. This remodelling processes make the gonadal cells particularly sensitive to environmental factors; this is so even in adults, and could explain why potential transgenerational effects are observed (39-42).

A truly epigenetic transgenerational inheritance would mean that the parent organism already had the epigenetic change and transmitted the change to its progeny, without the progeny being

exposed to the causative factor. Therefore true transgeneration epigenetic marks are classed as such if they transmit to the F3 generation. This is because the developing germ cells that give rise to the F2 generation are already present, and could thus have been exposed to the same extrinsic factors, during the embryonic development of the F1 generation (43-44).

However, only in exceptional cases are DNA methylation-associated phenotypes inherited from one generation to the next. The genome-wide DNA demethylation and chromatin remodelling that occurs in the germ cells of the developing embryo typically prevent this from happening (39,45). However, there are sequences in the genome that are relatively resistant to the genome-wide DNA methylation reprogramming and it only takes one of the mating pair to possess these inheritable epigenetic marks (46). For example, a 2011 study found that the effects of stress can be transmitted to the next generation, even when unstressed animals were mated with stressed animals (47).

1.3.4 Diet, lifespan and ageing

Dietary components have been linked to changes in DNA methylation such as: folate, vitamin B6, vitamin B12, betaine, methionine and choline. These nutrients can all affect the pathways of 'one-carbon metabolism' that dictate the levels of the methyl-donor available for DNA methylation and histone methylation - S-adenosylmethionine (SAM). Although the availability of methyl donors might be expected to induce a genome-wide effect on the epigenome, only region-specific effects have been reported in several studies (48,49).

Additionally, nutrient availability is also key in epigenetically altered lifespan (50).

Two key factors that appear to affect lifespan via epigenetic modification are the insulin signalling pathway and calorie restriction. The insulin and insulin-like growth factor 1 (IGF1) signalling pathway is well understood in regards to ageing. Nutrient sensing by insulin/IGF1 engages a cascade which results in inactivating and promoting the cytoplasmic sequestration of the transcription factor DAF16. Under calorie restriction, DAF16 is instead sequestered in the nucleus in which it can activate genes that promote longevity (51,52).

Various lines of evidence support this pathway, as it was shown that mutations in the age-1 gene lengthen lifespan (53). Subsequently, another study showed that mutations in DAF2 increase worm lifespan two-fold (54). It was further shown that DAF16 mutations suppressed longevity phenotypes of AGE1 and DAF2, suggesting that they functioned in the same pathway (55,56). Additionally, it was found that a high glucose diet decreased rate of dauer formation, the stasis-like state used to survive harsh conditions, shortened lifespan of DAF2 knockdowns, decreased DAF-16 activity, and shortened lifespan of wild type worms. In contrast, disabling DAF2 halted larval development in the dauer formation that shows increased lifespan, inhibited reproductive maturity, and delayed onset of senescence (54,57).

Not only have epigenetics been shown to effect lifespan but ageing itself has been shown to effect epigenetics. For instance, ageing has found to be linked to genome-wide loss of histones. The evidence for this is linked to cell division, specifically, replicatively-aged yeast were shown to have a genome-wide nucleosome loss of approximately 50%, in comparison to wild-type, which in turn leads to transcriptional deregulation (58). In addition, long-living strains of yeast are found to overexpress histones (59). In addition, ageing has been associated with dysfunctional histone activation and repression; as highlighted by a 2010 study which shows that levels of H3 histone methylation on lysine residue 27 (H3K27me3) are related to lifespan (60), these effects are also transgenerational (61).

Another genetic consequence of ageing is an altered transcriptional 'program', something that has been repeatedly highlighted in senescent human cells. These cell lines remodel the epigenetic mosaic to induce cellular defence and inflammatory genes, many of which the senescent cell secretes (62). These alterations to the cell are part of the senescence-associated secretory phenotype (SASP). SASP is thought to up-regulate immune clearance of pre-malignant senescent cells and in turn aid in tumour suppression. Chronic SASP on the other hand, is thought to promote tissue ageing (63).

In support of these data, a strong correlation was demonstrated between the loss of H3K27me3 and the up-regulation of SASP genes in senescent cell lines (64). Additionally, inhibition of the H3K4 methyltransferase, MLL1, was shown to indirectly reduce the SASP (65). Ergo, these studies indicate that transcriptional reprogramming in senescent cells occurs through the altered activity of certain epigenetic machinery.

The SASP can be seen in haematopoietic stem cells (HSCs) from aged mice as they show an up-regulated inflammatory response and down-regulated chromatin remodelling, transcriptional repression, and DNA repair in comparison to younger murine controls (66). An in-depth analysis revealed that genes related to cell proliferation, cell adhesion, and ribosomal protein (RP) genes were up-regulated while those related to cell cycle, DNA replication and DNA repair were down-regulated (67).

In addition to histone modifications, age-induced genome-wide alterations to DNA methylation is demonstrable in mammalian models (68). Data supports the model that ageing progressively increases genome-wide heterochromatin deregulation. The DNA hypomethylation occurs primarily in repetitive regions of the genome that correlate with the presence of conserved heterochromatin, while hypermethylation primarily occurs at promoter CpGs.

In murine models, age-induced hypomethylation in liver is apparent as the models age from 6-24 months (69). A study of livers from long-lived murine strains showed lower basal levels of DNMT1 regulated by growth hormone deficiency in comparison to normal murine models. The differential expression of DNMT is thought to cause these DNA methylation differences and impact lifespan by modifying lifespan-related genes (70).

It has been demonstrated that DNA methylation can serve as an effective predictor of age in non-cancerous tissues. A study identified that DNA methylation at 353 CpGs, termed 'clock' CpGs, and no other epigenetic modifications was an accurate predictor of age (71, 72). A study conducted in parallel to these, investigated the rate of DNA methylation change by building a predictive model of how ageing affected the methylome in individuals aged 19–101 years old. A smaller set of 71 CpGs were identified that have a very high accuracy of age-prediction and occur near age-related genes. The smaller number of CpGs identified increased the diagnostic value of 'clock' CpGs (73).

Taken together it can be seen that ageing is correlated with progressive loss of epigenetic control through histone loss, and DNA hypomethylation. Also it shows that lifespan is related to the levels of epigenetic dysfunction. Furthermore that epigenetic differences in senescent cells can affect cancer-related and inflammatory-related pathways.

1.4 Epigenetics in disease

1.4.1 Epigenetics in Cancer

The epigenome of normal cell undergoes vast changes after they turn into cancerous cells (74, 75). These epigenetic changes, dubbed “epimutations”, along with widespread genetic mutations, play a decisive part in the initiation and progression of cancer (76).

The epimutations result in a genome-wide dysregulation of gene expression leading to the development and progression of cancer. This is possible as epimutations can result in silencing of tumour suppressor genes and activating oncogenes independently to and in concert with deleterious genetic mutations. Epimutations therefore serve as a second 'hit' that further drives the cell towards a cancerous phenotype (77). As epimutations can be passed on through cell division, they are selected for in a rapidly growing cancer cell population and therefore would confer a growth advantage to tumour cells further exacerbating their uncontrolled growth.

DNA methylation undergoes profound changes during cancer initiation and progression. These changes were the first epimutations identified in cancer (78). A cancerous epigenome is consists of both genome-wide hypomethylation and site-specific CGI hypermethylation (76). It is thought that they contribute to cancer initiation due to some changes occurring very early on in cancer development (79).

The genome-wide DNA hypomethylation plays a significant role in tumour formation and occurs at various genomic regions including CpG-poor promoters, repetitive elements, introns and gene deserts (80). Hypomethylation in these regions all contribute towards genomic instability (76,81).

DNA hypomethylation and genomic instability is implicated in a wide variety of human cancers (81, 82). In contrast, CGI hypermethylation silences tumour suppressor genes and in doing so contributes to tumorigenesis. Various tumour suppressor genes have been shown to be silenced by hypermethylation in cancer p16, MLH1 and BRCA1 (74, 76, 83). These genes are involved in DNA repair, cell cycle, cell adhesion, apoptosis and angiogenesis – all of which are fundamental to

cancer initiation and progression.

It is interesting to note that regions that are hypermethylated in cancer are often pre-marked with H3K27me3 mark in embryonic stem cells (84) suggesting a link between the regulation of development and tumourigenesis. This observation partially explains the theory that there is organised methylation of a subset of CpG islands in tumours (84).

Whilst it is known that DNA methylation and histone modifications work independently and in concert to alter gene expression which provides a selective advantage for tumorigenesis; the mechanism by which these epigenetic abnormalities are acquired during cancer development is still not fully understood. It is postulated that it is unlikely that the epimutations found in cancer cell have occurred in a random fashion over several cellular divisions. It is more likely that the epimutation accumulations seen are the result of initial disturbances, in progenitor cells, to the epigenetic modification machinery. Such initial insults could predispose affected cell towards tumorigenesis – this model has been dubbed the “cancer stem cell model” (79).

In summation, can be seen that attaining a better understanding of how specific genomic regions are targeted for DNA hypermethylation in cancer would likely lead to additional therapeutic targets.

1.4.2 Epigenetics in autoimmune disorders

During the adaptive immune response, T cells become mobilised when they encounter antigen presenting cells (APCs) such as B cells, dendritic cells (DCs), macrophages or Langerhans cells. APCs present the T cell with an antigen which if complementary with the T cell's MHC receptor mature the T cell. The MHC-antigen complex results in cytokine production by the T cell which stimulate T cell growth. Some T cells become T helper cell, also called CD4+ T cells, which attract and direct macrophages, neutrophils and lymphocytes to the area of inflammation. Other T cells become cytotoxic T cells, also known as CD8+ T cells. Another subset of T cells form in the maturation process, memory T cells, which stay in an inactive state, for potentially a long period of time, until they encounter the antigen again. Perturbations to the adaptive immune response can

see it being directed at the host's body.

These perturbations to immune system that result an auto-reactive condition share a lot of similarities. Autoimmune conditions typically differ in regards to whether the immune response is localised or systematic, and what antigens the immune response targets. SLE, SSc, Sjorgens, and RA all result in joint damage and systematic inflammation. However RA is the only whose pathology primarily revolves around joint damage and degradation.

These immune system perturbations have been associated with aberrations to epigenetic mechanisms. These epigenetic mechanisms can that influence these perturbations can be seen in the pathogenesis of four of the most prevalent autoimmune diseases: Systemic lupus erythematosus (SLE), Rheumatoid arthritis (RA), systemic sclerosis (SSc), and Sjogren's syndrome (SS). Although RA will be the autoimmune system that will be focused on in this thesis, the contributions of aberrant DNA methylations to these disease pathologies will be examined (85)

1.4.2.1 Systemic lupus erythematosus

In regards to epigenetic modification, systemic lupus erythematosus (SLE) is one of the most well-studied autoimmune diseases. SLE is a chronic autoimmune disease that is characterized by the production of autoantibodies that target a variety of nuclear antigens - consequently the disease can target many organ systems.

Evidence suggests a strong interaction between the environment and the epigenome in the development of SLE. For example SLE CD4+ T cells that were treated with DNA methylation inhibitors were converted to autoreactivity. In addition, the introduction of synthetically demethylated CD4+ T cells into murine models resulted in the onset of a lupus-like syndrome (86,87). Pharmacological evidence is also supportive of this as two medications known to induce drug-induced lupus were found to inhibit methylation and produce murine lupus. It is thought that these drugs caused this by either by blocking the activity of the maintenance of DNA

methyltransferase DNMT1 or the ERK signalling pathway (88).

Some biological features of SLE actually contribute towards DNA hypomethylation and disease pathogenesis. There is strong evidence that apoptotic DNA is important in inducing autoimmunity. For example DNA taken from stimulated lymphocytes that is placed into murine models is capable of causing an autoimmune disease which resembles human SLE. However the autoimmune development seemed to be dependent upon the DNA's level of methylation as only the hypomethylated apoptotic DNA induced an autoimmune response (89). Though this effect didn't seem dependent on the DNA being 'apoptotic', as demethylation of necrotic or normal DNA via drug treatment, when injected, caused an autoimmune response similar to that seen with apoptotic DNA (89). These reactions are thought to be mediated by signalling of an innate immune-response receptor - Toll-like receptor 9 (TLR9) (90).

Peripheral blood mononuclear cells (PBMCs) of SLE patients have also been shown to possess changes in DNA methylation upon exposure to moderate to high levels of ultraviolet radiation, independent of the transcriptional activity of DNMT1. This would likely suggest an epigenetic component to the photosensitive rash seen in SLE patients (91).

Dysregulation of ERK signalling is known to be involved in SLE pathogenesis. A study found that murine models with an inducible ERK signalling block, a pathway known to be upstream of the maintenance DNA methyltransferase DNMT1. Upon induction, this murine model demonstrated hypomethylation and upregulation of CD11a, as well as autoantibody production and the production of interferon-regulated genes (92). Furthermore, data on oxidative stress in SLE also reflect this as increased oxidative stress results in decreasing signalling through the ERK pathway, previously identified as key to DNMT1-mediated DNA methylation (93)

Many a gene's level of methylation has been tied to SLE pathogenesis and progression. For example in CD4+ T cells, TGAL (CD11a) is involved in both cellular adhesion and co-stimulation. CD11a has been found to be overexpressed in SLE, this is likely due to the upstream promoter of

CD11a being demethylated; the degree of this demethylation is also correlated with disease activity (94).

IL10 is a cytokine pivotal to the development of both human and murine lupus. It is a well-described lupus-susceptibility gene overexpressed in lupus CD4+ T cells and functions to regulate many important other via several signalling pathways. It is significantly hypomethylated in SLE patients compared to disease-free controls. (95)

Genes encoded on the X chromosome are investigated as they possess relevance to SLE pathology. Specifically given that SLE is predominant in females and there is an apparent gene-dose effect (96). An example of a relevant SLE-associated T cell gene located on the X chromosome is CD40L, which is a type II transmembrane protein. Furthermore, elevations of soluble CD40L are found in human SLE patients and several lupus mouse models (97, 98).

This CD40L gene hypomethylation is at least partly drives its overexpression. This has been shown via treatment of human T cells with DNA methylation inhibitors or ERK pathway inhibitors which in turn hypomethylate the CD40L gene promoter sequences, lead to overexpression of CD40L, and were also found to be sufficient to cause T cell auto-reactivity in vitro (99). It can be seen that female SLE patients have roughly double the expression of CD40L compared to their male counterparts, considering females have twice as many X chromosomes as men then it could be a possible explanation for this apparent gene-dose effect (100).

A genome-wide methylation survey found that a SLE-afflicted twin in monozygotic twins possessed lower DNA methylation content than the healthy twin. This was found to be associated with significant reduction in expression levels of both DNMT1 and DNMT3b. Additionally, 49 genes were found to be differentially methylated in the SLE-afflicted twin, which had multiple functional ontologies, such as: defence response, cell activation, immune response, cell proliferation, and cytokine production (101).

A study analysed 27,000 CpG sites in CD4+ T cells of female SLE patients were compared to HCs. 337 CpG sites were found to be differential methylated, of this 337 approximately 70% were hypomethylated, and several methylation associations with SLE were found. For example, Matrix metalloproteinase 9 (MMP9), which has also found to be overexpressed in Sjogren's syndrome and RA, was hypomethylated (102). PDGFRA was also found to be hypomethylated, and had been previously identified as a target of autoantibodies in active SLE patients (103). A gene pathway analysis was conducted and it identified pathways and ontologies which were significantly overrepresented amongst the differentially methylated genes, the most prominent of which included folate biosynthesis, which is involved in production of the methyl-donor needed for DNA methylation. Multiple transcription factors were also overrepresented, including RUNX3, which interacts with the promoter region of the previously mentioned, CD11a. The last observation drawn from this study was that there was a significant correlation between the DNA methylation pattern of several genes and clinical disease activity (104). The same research group hypothesised that alterations in DNA methylation were a cause of human SLE rather than a response to the presence of disease.

They used a more comprehensive platform that would analyse over 485,000 CpGs and used it to examine the effects of differential methylation naïve CD4+ T cells from SLE patients. This identified 86 differentially methylated CpG sites and again approximately 70% of the CpGs were hypomethylated. These data were further verified as several genes were similarly differentially methylated as they were in the previous study with total CD4+ T cells. Furthermore, they found hypomethylation without evidence of overexpression of interferon-regulated genes, suggesting strongly that naïve CD4+ T cells in SLE are primed for a rapid type-I interferon regulated gene expression upon stimulation. Their data demonstrated that the majority of hypomethylated genes in naïve CD4+ T cells in SLE are interferon-regulated. Additionally they identified hypermethylation of two genes known to be SLE genetic susceptibility loci (CD247 and IL21R). This would suggest

that epigenetics is an alternate mechanism that may disrupt these genes in human SLE. These findings highlight the ability of epigenetic mechanisms to modulate the genetic risk of SLE (105, 106). However it is not clear how this genome-wide hypomethylation phenotype develops initially, but multiple studies have identified novel mechanisms that may play a role (105, 106).

Taken together it can be seen that epigenetic differences is not only linked to but is thought to cause SLE pathology - this is achieved via disruptions to the methylation machineries. However how the genome-wide methylation initially occurs isn't known. The literature also highlighted naïve CD4+ T cells as an appropriate candidate cell type for methylation analysis.

1.4.2.2 Sjogren's syndrome

Epigenetic mechanisms have been long been known to be involved in Sjogren's syndrome. Similar to other autoimmune diseases, DNA methylation in SS circulating immune cells is generally reduced. Similar to SLE, the CD70 gene has been demonstrated as hypomethylated and overexpressed in SS CD4+ T cells (107). Furthermore, one would expect that the population of regulatory T cells (Tregs) in SS would be reduced with disease progression. This is seen to be the case, as CD4+FOXP3+ T cells are reduced, due to hypermethylation and reduced expression of the FOXP3 gene (108)

A genome-wide DNA methylation study was conducted in naïve CD4+CD45RA+ T cells in comparison to HCs. 553 CpGs were found to be differentially methylated. 64% of these CpGs were hypomethylated. Several genes of interest were differentially methylated such as lymphotoxin α (LTα), which is a particularly interesting finding, as LTα is overexpressed in both salivary gland tissue and sera of SS patients. Furthermore, it is upregulated in the salivary glands of SS murine models, and if LTα is knocked out in these same murine models they don't develop SS (109).

Several genes involved in type I interferon pathway were also found to be differentially methylated, including STAT1, IFITM1, IFI44L, and USP18 – these findings were similar to those in the above SLE study. A gene in particular which had significance was the RUNX1 gene which is particularly important in T cell development, was found to be hypermethylated among SS patients. In addition, The T-cell receptor ζ-chain gene CD247 was hypomethylated, which highlights the importance of T cell signalling in SS (110).

Again, like SLE, SS also appears to be correlated with epigenetically changed genes which control and effect methylation levels. Furthermore containing differentially methylated genes which upregulate the immune system seems to be characteristic of SS.

1.4.2.3 Systemic sclerosis

T cell dysfunction is closely related to the development of SSc, just like SLE. This is particularly so in early disease, where the autoreactive T cells transfer signals to surrounding fibroblasts leading to collagen deposits and fibrosis (111). Also like SLE, it appears that hypomethylation of CD4+ T cells, is also driven, at least in part, by the downregulation of DNMTs. Therefore contributing significantly to the overexpression of various genes important to disease pathology (111).

The CD40L gene is found to be hypomethylated in SSc. In addition to its role as an adaptive immunity costimulatory molecule, a study has shown that CD40L plays a major role in the fibrosis of SSc (112,113). Furthermore, CD40 is found to be overexpressed in fibroblasts of SSc patients (114,115), and that blockade of the CD40L/CD40 interaction can reduce fibrosis in SSc murine models (116). Also similar to other autoimmune diseases, CD70 is hypomethylated and its transcript overexpressed in CD4+ T cells from SSc patients (117, 118).

In contrast to immune cells, the expression of DNMT1 is significantly increased in dermal

fibroblasts of SSc patients (119). Fibroblasts play a key role in SSc pathogenesis, particularly in organ and skin sclerosis where collagen is overproduced as well as other extracellular matrix components.

A genome-wide methylation assay was conducted examining large-scale differential methylation patterns in dermal fibroblasts taken from patients with different subtypes of SSc (120) Only 203 CpG sites were consistently methylated among the two SSc subtypes. Furthermore, the majority of these differentially methylated CpG sites were hypomethylated. This is despite the aforementioned studies showing an increase in DNMT1 expression (121)

Among the CpG sites that were shared, several were located on collagen genes, such as COL4A2 and COL23A1, or collagen breakdown enzymes. Among the SSc subsets, there was further hypomethylation of *COL6A3,* COL1A1*,* and COL12A1 in one subset and COL16A1, COL8A1, and COL29A1I in another. In addition, the transcription factors RUNX1 and RUNX2, which have also been associated with eroded cartilage in osteoarthritis and RUNX3, a T cell-associated transcription factor, were all hypomethylated in both SSc subtypes (121, 122).

Again, similar mechanisms to the other auto-immune conditions seem to appear. SSc is associated with epigenetic changes which effect the epigenetic methylation machineries, and the activity of immune cells. In addition, SSc seems to also possess disease-specific epigenetic changes, in the example of SSc this is likely highlighted by the changes to the 'COL' genes which facilitate collagen breakdown.

1.4.2.4 Rheumatoid arthritis

RA is perhaps best described as the prototype chronic inflammatory disease with features of autoimmunity (123).Even so, the first RA methylation studies showed that patients' T cells show a remarkably similar phenotype to SLE, characterised by genome-wide hypomethylation (123,124). Subsequent studies confirmed this, but did not find evidence for association between methylation levels and disease activity or DNMTs and disease activity; furthermore, it was found that the best therapeutic agents, TNF inhibitors, did not alter the methylome in PBMCs from RA. Indicating that changing the levels of TNF, a prime effector of RA pathology, had no effect on the differential methylation found in PBMCs from RA patients (125).

The development of larger DNA methylation assays have allowed for more unbiased evaluations of DNA methylation levels from a genome-wide perspective. Despite this, most studies have to date looked only at mixed cell populations such as PBMCs. So considering the specificity of methylation marks to particular cell types, it is therefore unsurprising that these data have not yet revealed significant differences between RA patients and controls.

DNA methylations pattern similar to that found in SLE have been found in candidate gene studies, wherein female RA patients have demethylation of the immune costimulator CD40L, which is not found in male RA patients, perhaps offering insight into the female predominance of the disease. However more study is needed in this area to deduce anything concrete (126).

A recent study did provide insights into the interaction of the RA PBMC epigenome and the underlying genomic code. This study identified ten DNA methylation phenotypes that appeared to mediate the underlying genetic risk for RA, with nine being within the major histone compatibility (MHC) complex, previously significantly implicated in RA in multiple genetic polymorphism studies (127).

A much stronger differential methylation has been found in synovium and synoviocytes from RA. Early candidate gene studies identified hypomethylation of elements that drive the expression of

MAP kinase, which in turn creates an aggressive cartilage-breakdown phenotype in RA fibroblast-like synoviocytes (FLS) (128,129).

It was also found that the maintenance methyltransferase, DNMT1, is downregulated in RA synovial fibroblasts (130). The subsequent findings supported similar findings in other autoimmune conditions regarding DNMT1 downregulation. They showed that there was a repression of machinery responsible for producing the methyl donor required for DNA methylation (131). Further studies indicated that there was chronic immune stimulation by IL1, TNF, or toll-like receptor ligands in various combinations in vitro which in turn mediated effective suppression of DNA methyltransferases (131-133).

Taken together, it can be seen that RA, like all the other auto-immune conditions listed here have epigenetic changes to their methylation machineries, they also possess epigenetic changes which encourage T cell activity. Furthermore, each autoimmune disease appears to possess epigenetic changes to genes which are unique to their given pathology and likely contribute towards their given pathological nuances.

1.5 Evaluation

1.5.1 Choosing a candidate cell-type

From this it can be seen that there is precedence of epigenetic aberration being involved in RA, however there do appear to be gaps in the published epigenetic RA studies. For instance, as previously mentioned, previous studies investigating epigenetics in RA have typically looked at PBMCs and not individual populations within the PBMC population. As a result investigating individual epigenetics of pathologically relevant cell types within the PBMC population would be a viable study model.

In such a study individual cell types should be focused on, as attempting to look at the entire PBMC population's epigenetic mosaic is disadvantageous as it assumes that the cell types within this population will all develop the same epigenetic aberration. Furthermore, as previously mentioned, the proportions of different cell subtypes can be better or worse represented in PBMC blood in RA patients. Not only might there be differences between RA patients, but the epigenetic data would favour the cells with the greater populations and would poorly represent the populations fewer in number - even if they have significant epigenetic aberration. As a result, for more reliable epigenetic data, a relevant cell type needs to be selected and isolated from PBMCs for an epigenetic investigation into RA.

PBMCs contain T cells, B cells, NK cells and monocytes. All of these have a role in RA pathology and are all potentially suitable for an epigenetic study. All of these are potentially suitable as, in addition to the several studies outlined above investigating epigenetic aberrations in RA T cells, literature suggests that WBCs in RA patients have increased telomere shortening in comparison to HC. Though literature does indicates that this shortening does appear independent of disease severity or duration. This telomere shortening is also found in early RA patients, specifically within

T cells. This telomere shortening is an indicator of cellular age (132). As previously mentioned, ageing is a key factor in the formation of epigenetic abnormalities. It is therefore likely that WBCs who endure advanced ageing develop abnormal epigenetic mosaics.

The candidate cell type should therefore be able to be found in PBMC samples, an individual cell type with a sufficiently large population to be analysed and relevant to the RA pathology. Whilst all PBMC cell types have a role within RA pathology, of these cell populations CD4+ T cells arguably has a more important role due to their role in attracting and directing immune response elements during inflammation. This role is a critical one within the setting of a chronic autoimmune disease such as RA as they are involved in maintaining the chronic inflammatory response. As a result CD4+ T cells has been appropriately selected as the candidate cell type for this epigenetic study.

1.5.2 Aims and Objectives

The aim of this study is to verify the data found in the previous Illumina 450k methylation array which suggests that RA CD4+ T cells are differentially and significantly methylated in comparison to healthy controls.

This objectives of this study will be to:

- Identify an RA-relevant candidate gene from the Illumina 450k methylation array data
- Bisulphite sequence the candidate gene in CD4+ T cells - in both RA and HC patients.
- Analyse the sequenced candidate gene, in RA and HC, to derive methylation levels.
- Determine if the candidate gene is differentially and significantly methylated in RA CD4+ T cells in comparison to HC CD4+ T cells.

1.5.3 Hypothesis

H1: My hypothesis is that the candidate gene identified in CD4+ T cells will be significantly differentially methylated in RA patients in comparison to HC ($p \leq 0.01$).

H0: My null hypothesis is that the candidate gene identified in CD4+ T cells are not significantly differentially methylated in RA in comparison to HC ($p > 0.01$).

This hypothesis will be tested by identifying a candidate gene from the data produced by a previously performed Illumina 450k methylation profile assay; which was done on isolated PBMC RA and HC CD4+ T cells. In order to identify a candidate gene with this data, statistical tests and ranking parameters will be employed to shortlist significantly hyper or hypomethylated genes. The most hyper or hypomethylated significant genes will then be reviewed by literature analysis to determine which may possess the greatest influence in RA pathology. Once a candidate gene has been chosen, an appropriate region within the gene shall be chosen to analyse which region would likely impart an effect on its transcription if differentially methylated – likely the promoter region. This area would be analysed by bisulphite sequencing and data from which would determine if the CpGs within are similarly methylated to what was found within the methylation profile data. If verified that the candidate gene region is differentially methylated, then it will be investigated how this might impact the CD4+ T cell and its role in RA pathology via a literature review of the gene's pathway and its various interactions it facilitates.

2. Materials and Methods

2.1 Ethics and IACON

The inflammatory arthritis disease continuum longitudinal study (IACON) obtains immunological data from the analysis of histological, serum and blood samples, and therapeutic response data and correlates it with clinical information from inflammatory arthritis patients. The cohorts have previously been used in identification of non-responder subtypes and analysing the pathophysiology of remission in RA patients. This study was ethically approved (see appendix 7.1) and conducted under the IACON study number 09/H1307/98.

2.2 Isolation by negative selection of CD4 T-cells from PBMCs

The CD4 T-cells were isolated with the Human CD4+ T cell enrichment kit (STEMCELL, UK). The cells were diluted if the numbers were greater than 4.25×10^8. The 'cocktail' solution (100 µL) was added to the sample and suspension was incubated for 10 min at room temperature (RT). Next 50 µL of the magnetic bead solution was added and was left to incubate for 5 minutes at RT. Sample volume was increased up to 5ml with the recommended medium (PBS + 2%FBS + 1mM EDTA). The tube was placed in the EasySep™ magnet at RT for 5 min. Without shaking, in one motion, the magnet and tube were inverted and the solution was transferred into a clean 15ml falcon tube. After centrifugation at 5,000 RCF for 5 min at 4°C the supernatant was discarded, and the cell pellet resuspended in 200µL of PBS.

2.3 CD4+ purity check

The purity of the isolated CD4 T-cells was then checked for each sample. Purity was assessed using flow cytometry method. First, 10 µL of the isolated cell suspension was placed in a flow tube. Then 2 µL of the antibody mix composed of: CD3 FITC, CD4 V500 and CD8 A700 antibodies were added (BD Biosciences, USA). This solution was incubated for 30 min at RT. Finally 100 µL of cell fix buffer was added. Stained cells were acquired using a flow cytometer (BD FACSCanto, BD Biosciences, USA). The CD4 T-cells subset was considered pure at the level of 90% and DNA was subsequently extracted for future analysis. .

2.4 DNA extraction

The DNA from CD4 T-cell's was extracted with the QIAamp DNA blood mini kit (QIAamp DNA blood mini kit, QIAGEN, UK). First, 20 µl of QIAGEN proteinase K, at a concentration of 100 µg/mL, was added to an Eppendorf tube followed by 200µL of cell suspension. 200 µL buffer AL was added and vortexed before incubating the samples at 56 °C for 10 min. After incubation, 200 µL of 100% ethanol was added, vortexed and briefly centrifuged to pull all the liquid to the bottom of the tube. The solution was placed in a QIAGEN MINI spin column and was centrifuged at 6,000g for 1 min; then the filtrate was discarded. The spin column was placed in a clean collection tube before adding 500 µL of the AW1 buffer. The column was centrifuged again at 6,000g for 1 min, filtrate was discarded and, column placed in clean collection tube followed by 500 µL of the AW2 buffer added before centrifugation at 20,000g for 3 min. Additional centrifugation for 1 min at 20,000g was performed. The column was placed in a clean 1.5ml Eppendorf tube and incubated with 200µl of elution buffer at RT for 5 min, then centrifuged at 6,000g for 1 min. The DNA concentration and quality was measured using The Nano Drop (ND-1000, Labtech International, UK) to determine whether samples were of a sufficient quantity.

2.5 Bisulphite conversion treatment

In order to observe the methylated CpG site the extracted DNA was bisulphite converted. The EZ DNA methylation-Gold^{TM} kit, Zymo Research, USA kit was used. First the conversion reagent () was prepared of 900μl of nuclease-free water, 300μl M-Dilution Buffer, and 50μl M-dissolving Buffer in a CT conversion reagent tube. The 130 μL of conversion reagent was added to 20 μL of DNA sample in a PCR tube. Usually between 200 and 500 ng of DNA was converted. The sample was placed in a thermal cycler (TC-512 Gradient Thermal Cycler, Fisher Scientific, UK) , with a 10 min cycle at 98 ° C, then 64 ° C for 2.5 hours. Converted DNA was then cleaned up using column method and 600 μL of M binding buffer. The sample on the column was centrifuged at 10,000g for 30 sec before adding 100 μL of wash buffer and centrifuging again at 10,000g for 30 sec. The desulphonation buffer was then added, incubated for 20 mins at RT and subsequently washed 2 times with 200 μL of washing buffer and centrifuged at 10,000g for 30 sec. After that, the column was placed in a 1.5 ml microcentrifuge tube, and 10 μL of elution buffer was added, it was centrifuged at 10,000g for 30 sec resulting in 10 μL of bisulphite converted DNA.

2.6 Gene amplification

A single polymerase chain reaction (PCR) had a total volume of 20 μL which consisted of: 10 μL of Master Mix, 0.1 μL total volume of forward and reverse primers, 0.8 μL of 25mM $MgCl_2$, 7.1 μL of nuclease-free water, and 2 μL of DNA sample (HotStarTaq DNA Polymerase, QIAGEN, UK). The primers were at an initial concentration of 500nM (Thermofisher Scientific, USA).

The PCR program contained 40 cycles of denaturation, annealing and elongation can be seen below in Table 1.

Table 1: The optimised parameters of the thermal cycler for PCR.

	Temperature (° C)	Time
Initial denaturation	94	15 min
Denaturation	94	10 seconds
Annealing	Primer specific	30 seconds
Elongation	72	60 seconds
Final maintenance	15	∞

Amplified PCR product was detected on a 1.5% agarose gel in the presence of 10% ethidium bromide alongside hyperladder 1kb (Bioline, USA) to determine the correct size of the product. The gel was run for 1 hour at 100V. The gel was then viewed using a Gel Doc™ XR+ Gel Documentation System (BioRad, UK).

2.7 Sequencing

The PCR product was then sequenced. First 2.5 µL of PCR product was added to a PCR plate with 1 µL of ExoProStar enzyme (Illustra™ ExoProStar™, GE Healthcare, UK). The solution was then placed in the thermocycler for 1 cycle of: 15 min at 37°C, 15 minutes at 80°C, then held at 15°C. The 3.5 µL of nuclease-free water was added to each well and was mixed with 0,25 µL of "Ready Reaction Mix" and 3.5 µL of sequencing buffer, 4.25 µL of nuclease-free water, and 1 µL of 1.6µM forward or reverse primer depending on whether the forward or reverse sequence is being sequenced, respectively. The plate was then run on the thermocycler for 28 cycles of: 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 min and was then held at 4°C.

In preparation for sequencing the sample was precipitated with ethanol. To each sample 1 µl of '3M sodium acetate pH 4.6' and 25 µL 95% ethanol were added and incubated for 30 min at RT. After incubation the samples were centrifuged for 30 min at 2254g.

To dry the subsequent pellet the plate was inverted on a tissue and was centrifuged at 185g for 1 min. Then 70 µL of 70% ethanol was added and centrifuged at 1650g for 15 min at 4°C. The pellet was again inverted on a tissue and was centrifuged at 185g for 1 min. The pellet was then dried for 1 minute at 95 ° C on a thermocycler. After drying the pellets were re-suspended in 20 µL of HiDi formamide. The samples were then ready for sequencing on the 3130xl Genetic Analyzer (ABI Prism, USA).

2.8 Software Analysis

2.8.1 Sequencing analysis

The raw data was then analysed by the Sequencing Analysis software 5.3.1 (Applied Biosciences, USA). This software produced a graph of signal strength of the nucleotides at all the positions on the DNA strand it was reading from 5' to 3'. The resultant DNA sequences were then aligned with the expected bisulphite sequence by using the Basic Local Alignment Search Tool (NCBI, USA) which showed where the 8 CpGs fell on the read DNA sequence. From the graph of the relative nucleotides the T and C signal peak heights were measured and what percentage was C signal was derived, the greater the C percentage the more methylated the CpG. From here, using this methodology, HC and RA patient DNA was sequenced multiple times and had their CpGs read and methylation level calculated.

2.8.2 Initial gene selection and primer design

A candidate gene is selected from the methylation profile data based upon how differentially methylated it is - both in regards to significance of the methylation and also the number of CpGs that are significantly differentially methylated. In addition, the gene must have some clinical relevance to the pathology in question – in this instance, RA.

Once a candidate gene has been selected it must have primers designed for it. However bisulphite sequencing restricts the use of PCR products longer than ~350bp, so multiple primer pairs are needed in order to properly sequence all the CpGs in a given gene. The candidate gene had its sequence found from the UCSC human genome database (www.genome.ucsc.edu). This sequence was then placed through a Zymo online bisulphite converter tool at (www.zymoresearch.eu) which converted the sequence into its bisulphite equivalent. The bisulphite sequence then had primers designed around fitting as many CpGs in a PCR product. The primers had to be between 25-30 bp long and have a similar Tm value.

Once a set of primers have been successfully designed they had to have their PCR conditions optimised. This was done by using temperature gradients on the thermocycler and also having an MgCl2 concentration gradient. By doing so the optimum temperature and MgCl2 concentration were found for each primer set. The primers were then sequenced on HeLa DNA, the sequenced product was then checked for quality and if it aligned with the desired sequence.

3. RESULTS

3.1 DNA Methylation data

DNA Methylation profiling was performed on CD4+ T-cell DNA extracted from 4 healthy controls (HC) and 10 RA patients using an Ilumina 450K Array. The data was analysed for difference in DNA methylation levels for over 450,000 CpGs across the whole genome initially by myself. I first filtered the raw data by significance ($p \leq 1x10^{-2}$), then I divided the data by chromosome. From here I analysed which significant CpGs were within 1 nucleosome of each other (250bp). I now could look at which genes had the most amount of clustered highly significant differentially methylated genes and rank them. This revealed 55 genes with 374 CpGs significantly differentially methylated, 162 CpGs were hypermethylated and 211 CpGs were $p \leq 10^{-3}$. In order to highlight clusters of these CpGs they were ranked by distance from the next CpG. 100bp, 250bp and 500bp were used as distances ranked highest to lowest respectively. CpGs separated by ≤100bp were highlighted green, between 100bp and 250bp highlighted yellow and between 250bp and 500bp were highlighted red. Then the CpGs with their 'cluster rankings' were lined up with their respective genes and those genes were ranked on a table by number of greens, number of consecutive greens, number of colours, number of consecutive colours. It was decided to rank the overall list by number of consecutive greens as this was thought to more likely to carry over a transcriptional effect. The list was then divided by 'Less methylated' and 'More methylated'. This wording was chosen carefully as some genes on the list had CpGs that were both hypo and hypermethylated, so if there were more hypo than hyper it would be classed as less methylated - and vice versa. Fortunately, the genes were often greatly or completely polarised toward methylation or hypomethylation, however I felt the distinction was important. It should be noted that there weren't any cases of having equal amounts of both methylated and unmethylated CpGs within a gene. The results for this process can be found in appendix 7.2.

In order to verify my top ranked genes I got a PhD student in the group, Miss Rujiraporn Pitaksalee to perform some computational analysis using R-programming software. She showed 648 genes with significantly differentially methylated CpGs ($p \leq 1x10^{-3}$). Of these 648 genes, 354 were hypermethylated and 294 were hypomethylated. They were ranked on how significantly and differentially methylated they were in comparison to HC. Her results were more inclusive than mine however the top hypomethylated genes were highly similar with B3GALT4, TNF, DAXX and CUTA being in the top 5 hypomethylated genes for both analyses. The top 15 hypermethylated and hypomethylated from the PhD student's analysis can be seen in Table 3. My ranking method and results can be found for reference in appendix 7.2.

Table 2: Showing the study number, age, gender and DAS28CRP scores of the 9 RA patients used in the verification study. DAS28CRP is a measure of RA disease activity with a score of <2.3 indicates remission, 2.3-2.7 means a low disease activity 2.7-4.1 means a moderate disease activity, and 4.1< is a high disease activity score.

IACON	Age	Gender	DAS28CRP
147	74	F	3.29
384	55	F	5.11
444	33	M	2.7
510	65	M	3.39
520	40	F	5.72
612	62	M	4.56
762	44	F	2.81
879	62	F	3.32
1190	51	M	4.08

Literature analysis showed that 3 genes in the top 10 ranked hypomethylated genes and 1 hypermethylated gene were directly or indirectly involved in the TNF pathway, including TNF itself (Table 3).

Table 3: showing in descending order the most significant 15 hypo and hypermethylated genes in RA CD4+ T cells. Those genes in red are found to be associated directly or indirectly with the TNF pathway.

Hypermethylated	Hypomethylated
Genes	Genes
HMX2	B3GALT4
ABAT	**TNF**
HIC1	ABI3
HPN	**DAXX**
BAIAP2	CUTA
OCSTAMP	**KSR1**
NR4A2	GNGT2
BAIAP2-AS1	MEOX1
RNF39	IFT140
SFRP2	COL11A2
HOXB-AS3	SNORD32A
RNF32	GPSM3
GAREM	ZC3H12D
BAALC	PBX2
CELF4	MAZ

According to literature searches, the other top 28 genes have various functions, some seemingly unrelated to RA pathology although the gene OCSTAMP could also be related to RA as it is involved in osteoclast differentiation. Two genes, ABAT and CUTA are involved in the metabolism of neurotransmitters (134,135). A number of genes are found to be associated with cancer; HIC1 is a gene that is found to be hypermethylated in cancer cells (136). HPN is associated with prostate cancer, and MAZ is found to be deleted in colorectal cancer (137). B3GALT4 and BAALC are found to be associated with lipid raft formation (138). A lot of the genes found in this list are those that are associated with proper embryonic development: HMX2, RNF32, HOXB-AS3, CELF4, MEOX1 (139).

All 16 top ranked hyper and hypomethylated genes were run through a gene network analysis program (Genemania).The software analysis generated a diagram showing the various types of interactions these genes may have with each other (Figure 1). The definitions of the genes mentioned can be found in the appendix 7.4.

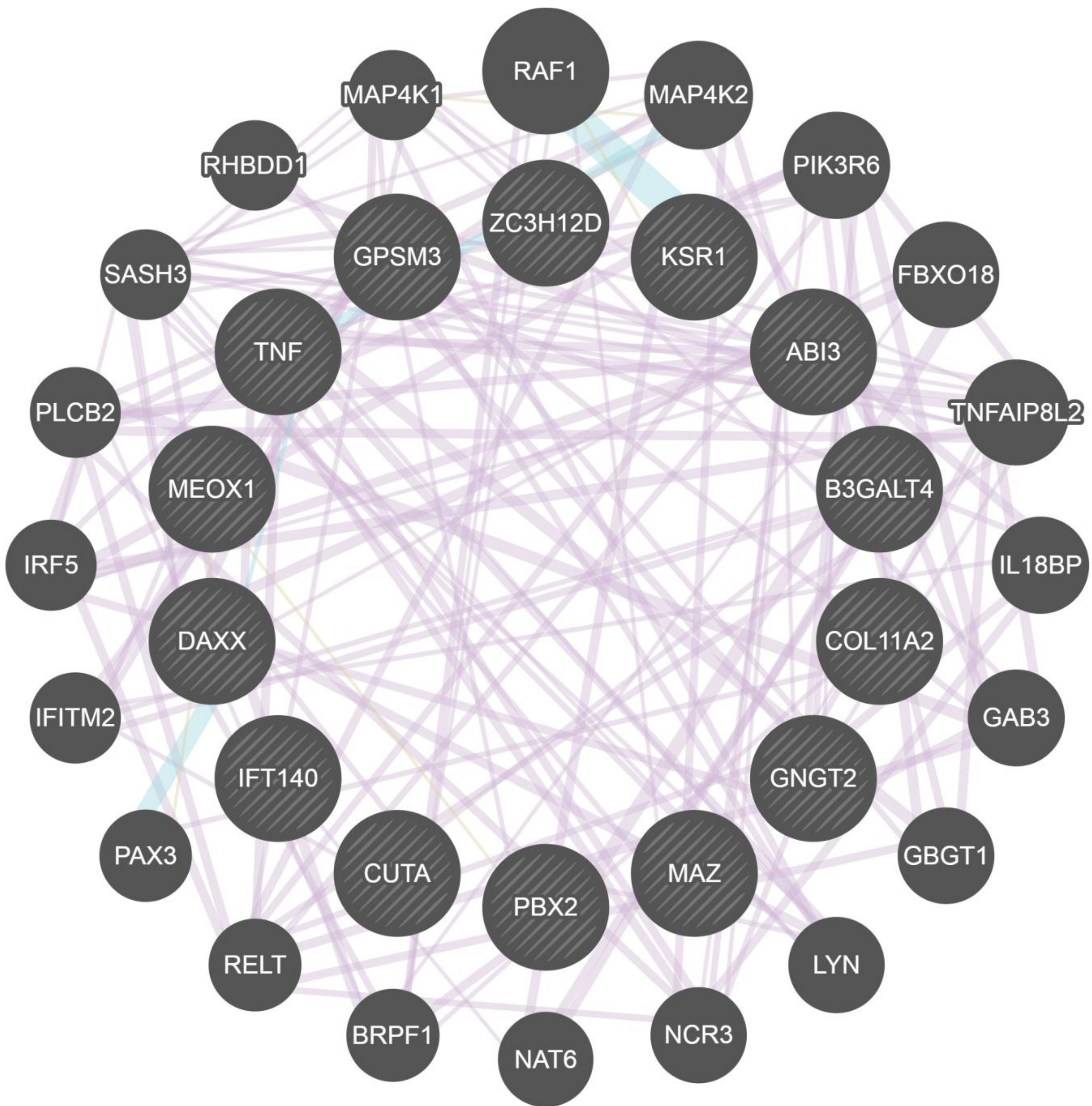


Figure 1. Gene network of the top 14 hypomethylated genes from the methylation profiling assay (circles with the grey stripes). The blue lines indicate that the genes share a pathway and the purple indicate coexpression between the connected genes.

Out of the hyper and hypomethylated genes, only the hypomethlyated gene network showed presence of any significant associations (Figure 1).The hypermethylated gene network showed very limited links (data not shown).

The 14 genes show that the only hypomethylated genes that are related by a shared pathway is DAXX and TNF. Also a lot of the 14 genes are coexpressed with one another and some that aren't directly coexpressed share coexpression with a secondary gene which isn't necessarily hypomethylated.

The software program also searches which functions the list of genes is likely involved in; then it attaches a Q-value for significance. This analysis suggested that the list of the top 15 differentially methylated genes in early RA is likely to be involved in positive regulation of JUN kinase activity, MAP kinase activity and the regulation of stress-activated protein kinase signalling cascade.

3.1.1 TNF signalling cascade analysis

The gene network map suggested that the TNF gene is involved in several associations and in the significant functions highlighted (regulation of MAP kinase activity and the regulation of stress-activated protein kinase signalling cascade). In addition, TNF is placed highly on the ranked list of significantly differentially methylated genes (3rd place) and so are others genes involved in its signalling cascade. As a result this project was focused on the TNF cascade. I reconstructed the TNF-alpha signalling cascade as shown in Figure 2, and identified around 65 proteins that are involved step-by-step in the cascade from TNF itself to its membrane receptor, the kinases and signalling molecules to the transcription factors and genes whose transcriptional regulation is

directly related to TNF-signalling.

I then investigated DNA methylation pattern individually for each gene. As can be seen from Figure 2, 14 of the proteins/genes (TNF, DAXX, CD30, CD160, LIGHT, NIK, MAP3K14-AS1, BCL2, UACA, TRAF1, TRAF5, TRAF6, ZAP70, KSR1) directly or indirectly involved in the TNF cascade, were differentially methylated however, with lower p-values ($p \leq 0.01$) compared to the TNF gene ($p \leq 0.001$), hence not at the top of the original list. To clarify genes stated to be indirectly involved in the TNF cascade are genes whose translated proteins share and contribute to the same downstream effects as the TNF cascade - such as NFkB activation. These NFkB-activating parallel pathways include TCR signalling, TNFRSF14 signalling, RANK signalling and CD30 signalling.

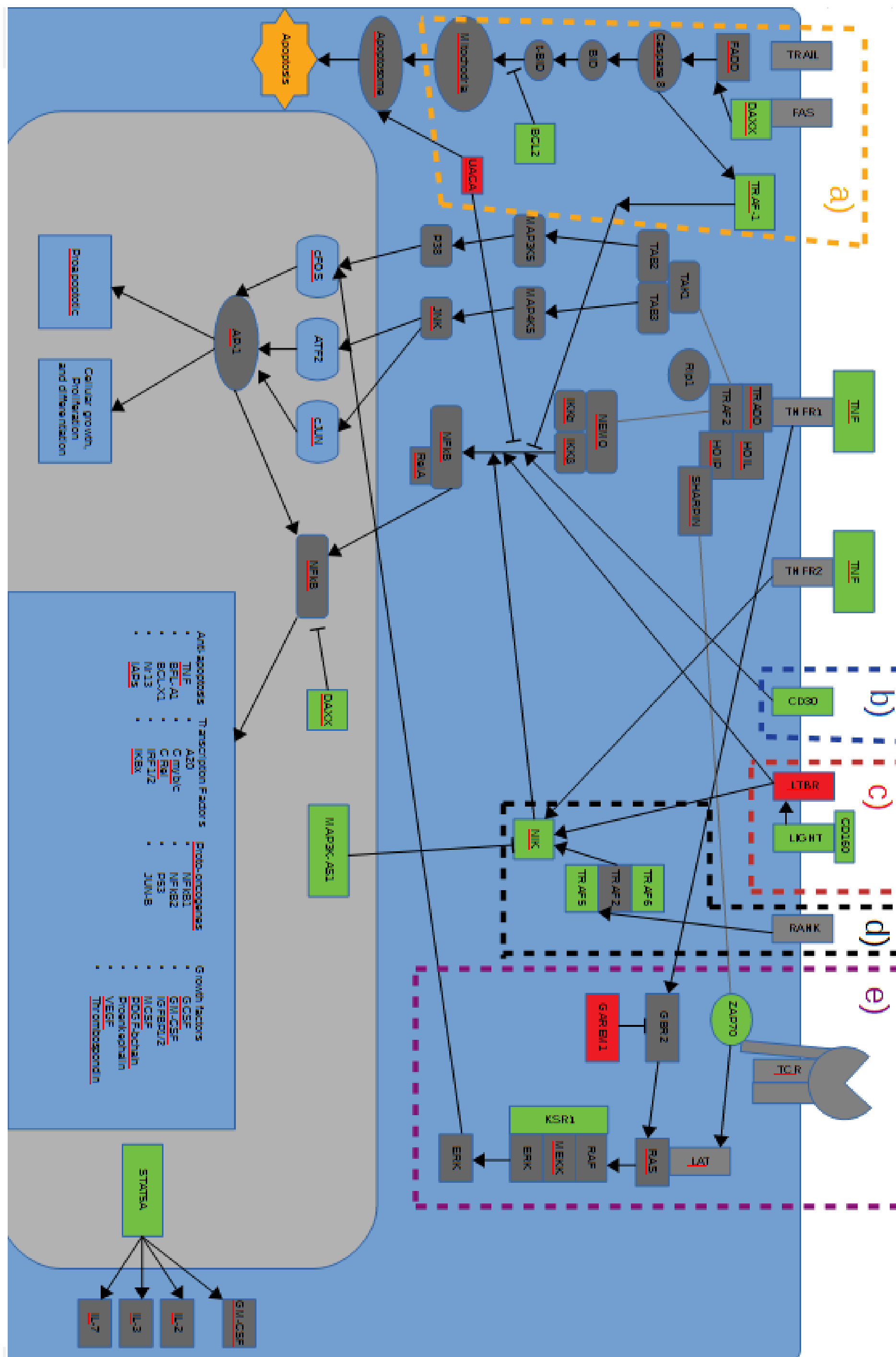


Figure 2. The TNF pathway 13 hypomethylated genes and 3 hypermethylated genes. In total 16 differentially methylated genesare involved in the TNF cascade in RA.a) showing the caspase apoptotic cascade, b) showing CD30 NFkB activation, c) TNFRSF14 stimulation of NIK via LTBR, d) RANK induction of NIK via a TRAF complex, and e) TCR induction of ERK2.

As can be seen from Figure 2, some of these genes are key elements of the TNF cascade. In the TNFR1-related apoptotic cascade, 4 genes are differently methylated; 4 genes are differently methylated in the TNF cascade; 1 genes that effect transcription; 9 genes involved in parallel pathways and are also differently methylated. This indicates that the epigenetic mosaic doesn't just effect one aspect of one signalling pathway, but multiple aspects and multiple parallel signalling pathways. The shared function of the majority of these pathways is NFkB activation.

Considering the TNF gene significance of methylation and relevance in RA it was therefore chosen as a candidate gene for direct bisulphite sequencing to confirm the differential DNA methylation data between HC and RA patients. The gene structure is described in Figure 3a. This gene does not have a defined CpG island but has several CpG rich regions spread through-out the gene. The CpG interrogated by the array are highlighted on the Figure 3b and the differential methylation levels observed between HC and RA patients are also depicted in Figure 3c.

3.2 Primer design

The TNF gene was found to have 14 CpGs that were significantly differentially methylated in RA compared to HC (Figure 3c). These CpGs were dispersed through the gene and as a result multiple PCR products were needed in order to be able to sequence these regions to validate the difference in methylation levels. As a result 6 primer pairs were designed to amplify 6 regions, hence 6 PCR products, based on the bisulphite converted DNA. All 6 products are shown in relation to the TNF gene in Figure 3a.

Various factors were taken into account when designing primers. The primary factor to be kept in

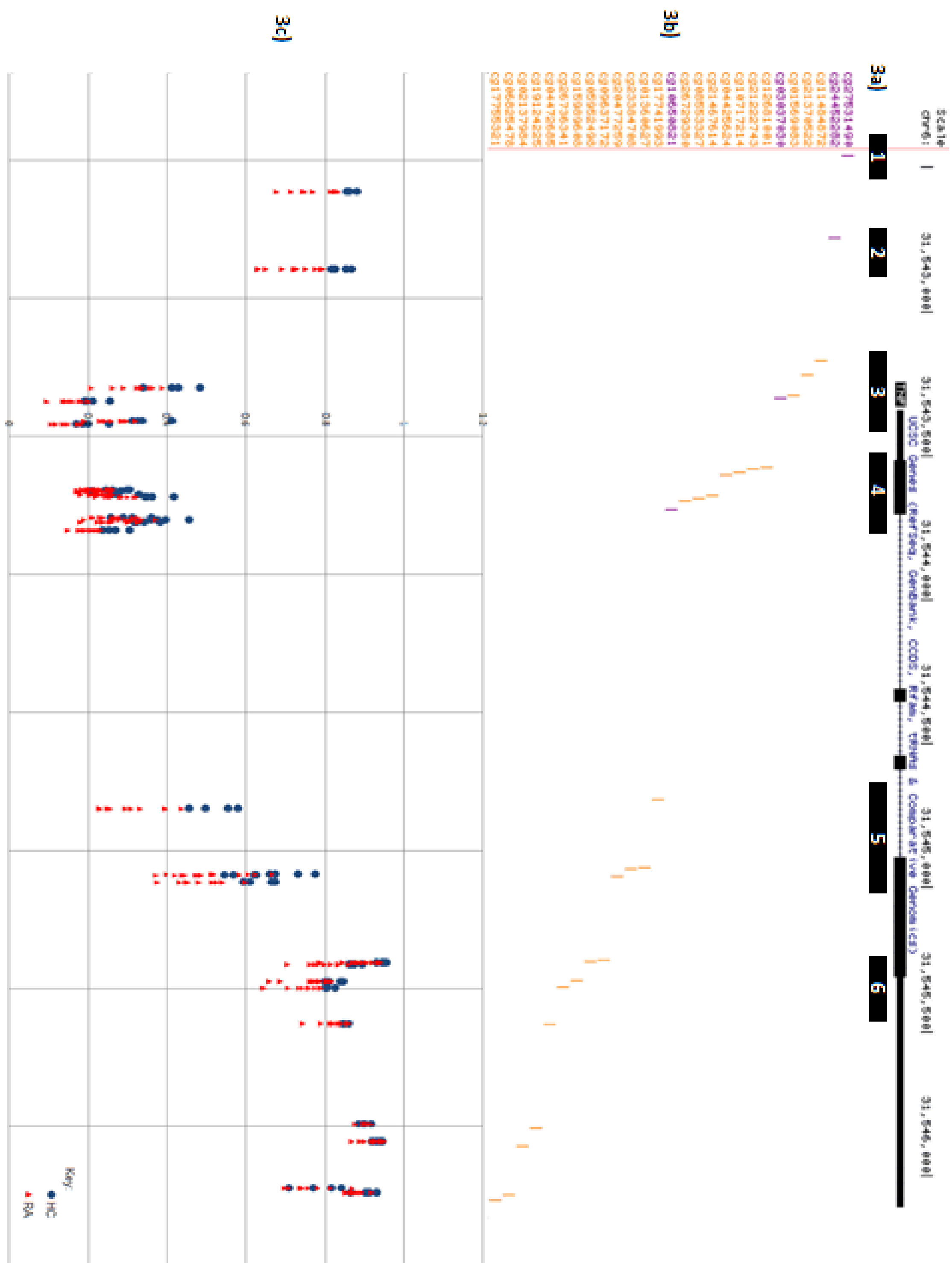


Figure 3: Diagram showing the TNF gene (a) in relation to where the CpGs are on the gene (b) and their corresponding methylation level according to the methylation profile array data (c).

mind is the melting temperature (Tm) of the primers. Typically the ideal range for primers is 52ºC to 58ºC; primers which have Tms exceeding 65ºC have a tendency towards secondary annealing during the extension phase of the PCR reaction. The Tm of a primer is a function of its length and its nucleotide composition therefore the longer the primer the higher the Tm and the higher the GC ratio the higher the Tm. Considering that, the ideal primer length is 18-30bp, primers based on bisulphite converted DNA tend to be long in order to compensate for the relative lack of Cs in bisulphite converted DNA.

A second factor to be considered is the distance from desired CpGs as the first 25bps of the sequence is usually of lower quality, hence CpGs to be tested should be ~25bps or more away from primers (i.e. “buffer zone”).
The final consideration is the desired product length with the recommended range being 150-300bps. When taking into account the primers and the ‘buffer zone’ required, the region that can be sequenced is approximately 40-190bp.

Following these criteria, the TNF gene was divided up into 6 suitable regions, each associated with a primer pair. Each resulting PCR product was numbered 1 to 6 from the 5’ to 3’ of the gene (Figure 3a); and the primer pairs associated with each products shared the same numbering system.

The process of primer design is illustrated on Figure 4. First a target DNA sequence to sequence is identified which contains desired CpGs; the bisulphite converted DNA sequence is then established in silico (using Zymo bisuphite conversion software) and from this converted sequence, primers are designed. As previously mentioned both primers need to be 25bp or longer and 25bp or more away from the desired CpGs. In addition the Tms of the forward and reverse primers must be within 1ºC of each other to ensure proper melting and annealing during PCR amplification. To verify the Tm of my primers I used Primer express software. Once the primers are positioned they should produce their respective products as shown on Figure 4.

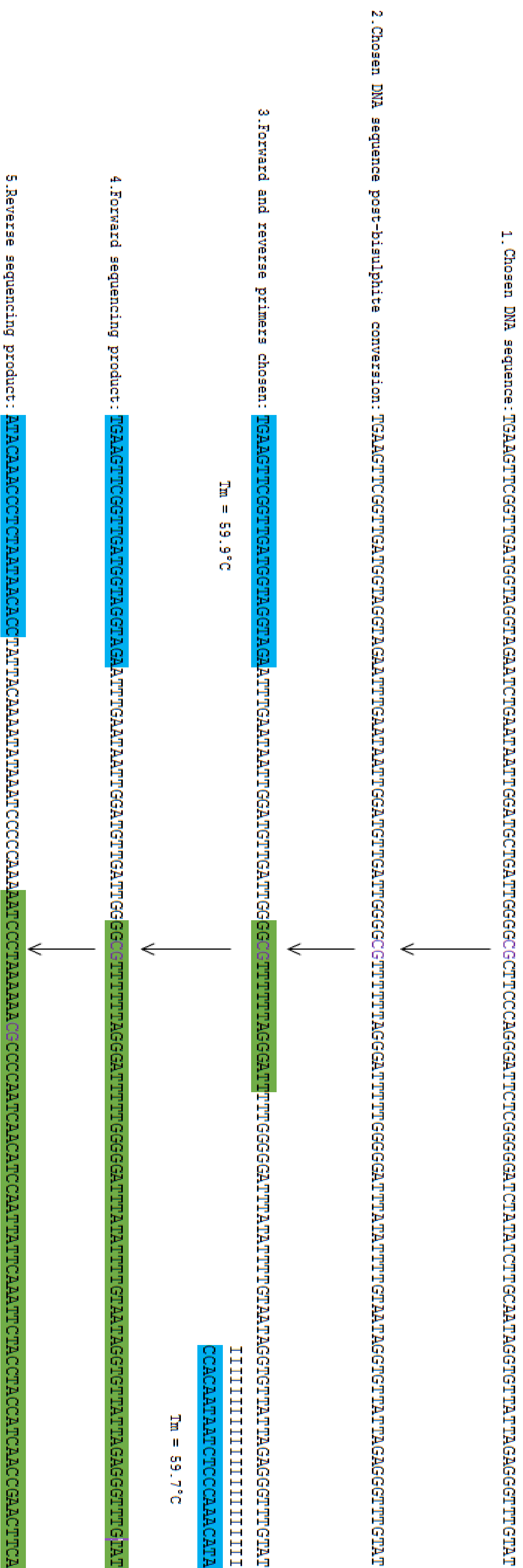


Figure 4: Flow chart showing how primers are designed

Following the above procedure for designing primers and PCR products, I identified 6 suitable regions in the TNF gene, for each of which I designed primers for a PCR product. The sequence of each primer pair sets is provided in appendix 7.3.

3.2.1 PCR optimisation

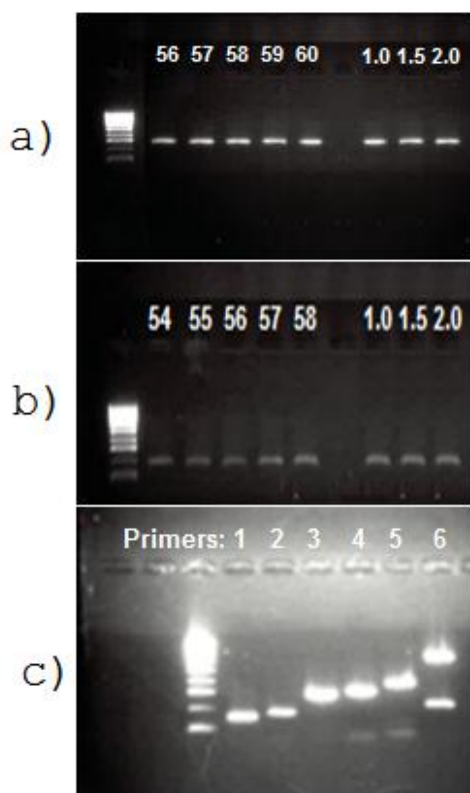
To optimise the amplification of my 6 PCR products, I proceeded to use a standard temperature gradient and in addition used a variable concentration gradient of MgCl2.

As an example, Figure 5a shows that primer pair 3 generated the same amount of PCR products across the temperature range from 56°C to 60°C. It was also stable across the MgCl2 concentration range of 1.0-2.0mM. This demonstrated a good, stable and efficient amplification of this PCR product 3.

In contrast, primer-set 2 (Figure 5b) was put through a similar PCR optimisation with a temperature gradient, ranging from 54°C to 58°C. An increase in the amount of PCR product is apparent as the temperature increases. This demonstrates how a 1-2 degree difference in temperature may affect annealing (despite the Tm being predicted to be 56°C.) and how the efficiency of the PCR reaction can be reduced. All 6 products were then optimised and final conditions are described in Figure 5d.

Finally, in Figure 5c each PCR products (1, 2, 3, 4, 5 and 6) were run alongside each other at their optimum conditions. It can be noted that even after optimisation products 4, 5 and 6 seem to produce secondary products of lower length. However for products 4 and 5, these secondary bands were

lower than the 100 bp ladder implying that they are likely to be primer dimers. For product 6 the secondary PCR product is much longer and likely to be the secondary amplification of an unknown sequence. It is unlikely that I would be able to attain a good sequence for this particular PCR product.



d)

Primer pair	Optimum annealing temperature (°C)	Optimum MgCl2 concentration
Primer 1	60	1.0mM
Primer 2	58	1.0mM
Primer 3	58	1.0mM
Primer 4	61	1.5mM
Primer 5	58	1.0mM
Primer 6	59	1.0mM

Figure 5: Diagram showing PCR product optimisation. 5a shows primer 3 temperature and MgCl2 gradient, 5b shows primer 2's temperature and MgCl2 gradient, 5c shows all 6 final optimised PCR products, and 5d shows all 6 primer pairs and their optimal PCR temperature and MgCl2 concentrations.

3.3 DNA Sequencing

I proceeded to the sequencing of all 6 products using HeLa DNA first to optimize the technique. After optimisation of the sequencing reaction and the acquisition of sequences for all 6 products, I aligned my sequences and compared them against the theoretical product expected from HeLa DNA. All products aligned to the expected sequence however, the quality of the alignment was not equivalent for all products, the worse being the longest one and product 6 due to contaminating product.

It was found that product 3 appeared to provide the most robust sequence. Based on alignment and location of the product in relation to the promoter region of the TNF gene, as well as the amount of DNA needed per reaction being greatly demanding, it was decided to focus the sequencing in HC and RA patients on product 3 only. This product contains 8 CpG sites of which 4 were directly interrogated on the Illumina array, and 1 (CpG1) of which was found to be significantly hypomethylated by the illumine array. Though it should be noted that all 4 would be considered significantly methylated in the initial experiment if $p \leq 3x10^{-2}$ was used instead of $p \leq 1x10^{-2}$. So whilst demonstrating a degree of significance the 3 other probes didn't meet the critical limit imposed on this analysis.

Figure 6a illustrates the beginning of a good quality sequence shortly after the primer, while there is a reduction of quality as the read of the sequence reaches the middle of the product. This is what is expected and is why the targeted CpGs are positioned 25bp or more away from the primers. This sequence also shows single peaks at position where a single C, A, T or G is expected while where a CpG site is located, two peaks (2 bases) can be read: T and C this is illustrating the difference in C-methylation being bisulphite converted in some cells (when methylated) but not in others (when unmethylated).

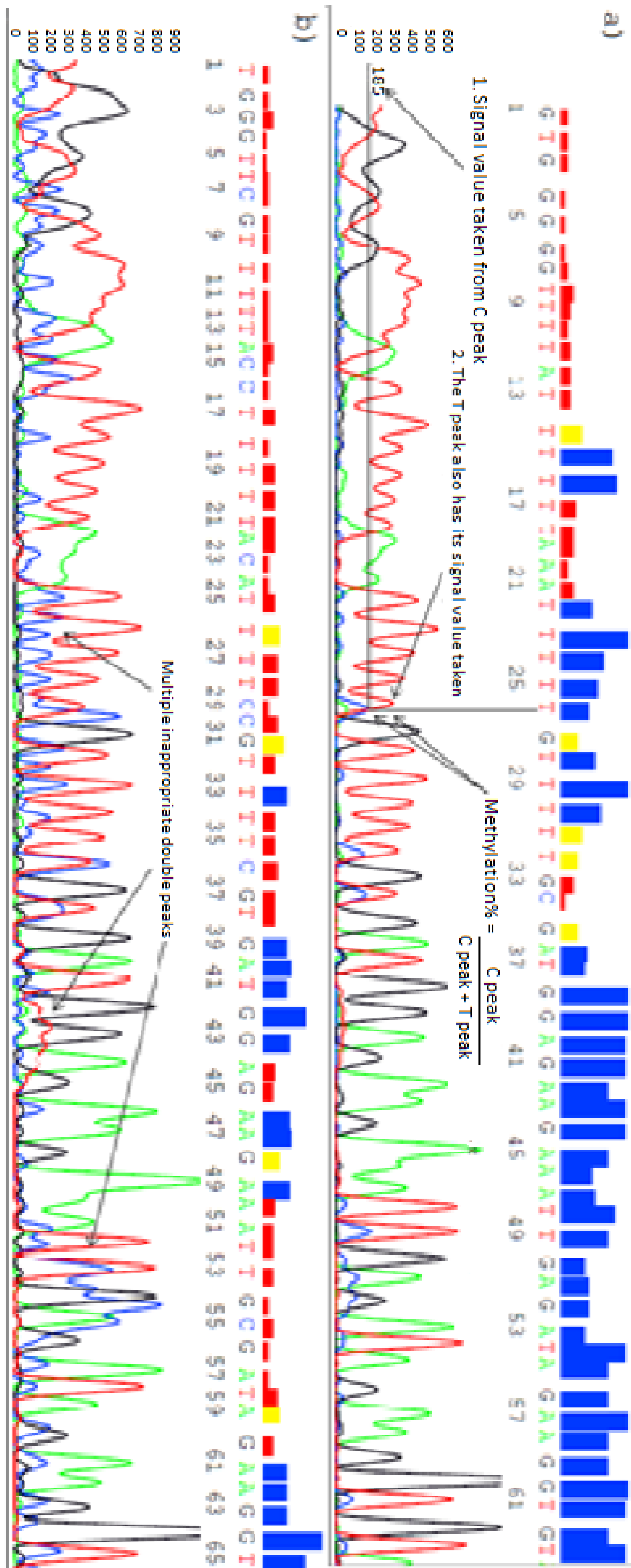


Figure 6: a) Showing how the CpG methylation levels were calculated, and b) demonstrating what a poor sequenced DNA product looks like.

Figure 6b shows the reads resulting of sequencing product 6, which was contaminated with a secondary product. As can be seen it is not showing single peaks for most of the sequence like it does for product 3 (as seen in Figure 6a). This is because the two products possess different nucleotide sequences at any given position and both are being sequenced.

The levels of methylation in the 8 CpGs contained in PCR product 3 were analysed by sequencing the region in HeLa cells, 8 HCs and 9 RA patients (RA demographics above in Table 2). Their relative C and T signals were measured as the average of sequencing results obtained at each position – each sample having being repeated a minimum of 6 times. The percentage of the total signal made up for cytosine at each of the CpG position in the TNF gene product 3 is shown in Figure 7.

In HeLa cells DNA of each of the 8 CpG was fully methylated in all individually repeated sequencing reactions (n=7). For HC (n=9, $\bar{x}$ =10.7 reactions per sample) DNA methylation of CpGs 1 to 7 on average shows variable levels of methylation in the range of 43%-51%, with CpG 3 being the most methylated, then CpGs 8 show methylation decreases to 35.95% on average.

The observed trend is similar but different in RA patients (n=9, $\bar{x}$ = 8.78 reactions per sample), where CpGs 1 to 6 are in the range of 9%-18%, then dropping down to almost 0% methylation for CpG 7 and a slight rise to 1.55% for CpG 8. Like HC, RA demonstrates a peak at CpG 3 and a general decline in methylation from CpG 3 to CpG 8.

Assuming the data in Figure 7 is representative of HC and RA populations and that RA patients once had HC methylation levels prior to RA onset then it can be deduced that CpG 7 is the most sensitive to demethylation and CpG 4 is the most resistant, based upon reduction in CpG methylation from HC to RA.

From the results in Figure 7 and Table 4, it can be seen that the CpGs 1 to 7 are significantly ($p<0.001$) hypomethylated in RA patients in comparison to HC. CpG 8 is conclusively hypomethylated and does demonstrates some degree of significance however it is orders less so ($p= 0.01368$) than the other CpGs. This difference in CpG 8 is likely due to possessing 1 fewer HC to plug into the significance test and possessing a larger variance – both of these factors are likely caused by CpG 8 being at the end of the sequence and thus possessing lower quality sequence data than the other CpGs. As a result this confirms the data obtained in the Illumina array suggesting hypomethylation in RA, for the 1 probe (CpG 1) located in product 3 of the TNF gene (Figure 8). The 4 probes are the CpGs 1, 5, 7 and 8, found below in Figure 7. For further reference, CpGs 1, 5, 7 and 8 seen in Figure 7 are cg11484872, cg21370522, cg01569083 and cg03037030, respectively, as seen in Figure 3 and Figure 8.

Sample	Replication repeats (Fwd,Rev)	CpG1	CpG2	CpG3	CpG4	CpG5	CpG6	CpG7	CpG8
HC2	9,4								
HC3	6,4								
HC4	10,4								
HC5	6,3								
HC6	6,4								
HC9	6,4								
HC11	5,4								
AVG		43.68%	50.37%	50.95%	44.00%	45.74%	48.35%	43.42%	35.95%
RA147	2,4								
RA384	4,5								
RA444	3,5								
RA510	2,4								
RA520	6,4								
RA612	3,4								
RA762	6,4								
RA879	5,6								
RA1190	5,7								
AVG		10.48%	14.72%	17.83%	11.39%	11.30%	9.85%	0.56%	1.55%

Figure 7: Diagram showing the pie chart averages of each RA and HC patient's 8 CpG methylation levels, and the percentage average of the RA and HC's CpG methylation levels.The blue section indicates the proportion of cells which are methylated at the CpG shown. Orange indicates the proportion of cells which are not methylated are the CpG shown.

Table 4. P-values demonstrating the significance of the RA and HC methylation at each individual CpG outlined in Fig 7. P-values were derived from a Welch's T test.

TNF CpG	CpG1	CpG2	CpG3	CpG4	CpG5	CpG6	CpG7	CpG8
P-Value	<0.0001	<0.0001	0.0003	0.0002	<0.0001	0.0002	0.0006	0.01329

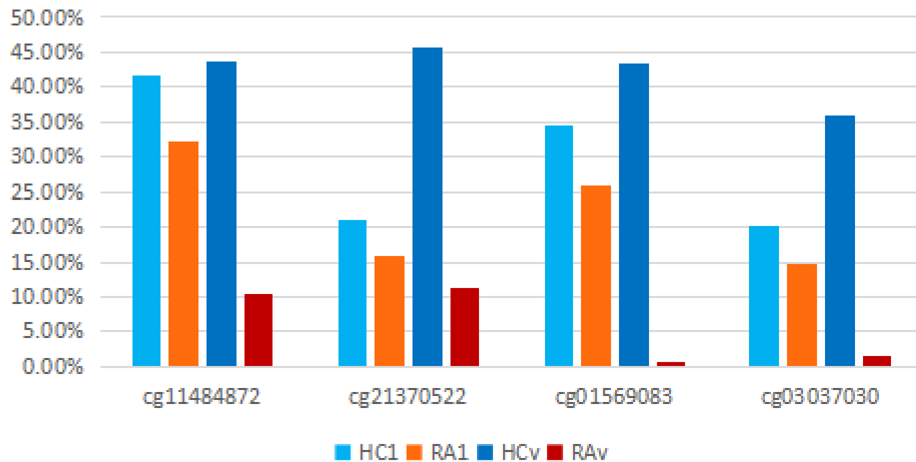


Figure 8. The methylation levels of HC and RA CD4+ T cells in the initial Illumina methylation profile data (HC1 and RA1) in comparison with the methlyation data (HCv and RAv) of the RA and HC CD4+ T cells in this verification experiment at the 4 CpGs picked up by the original Illumina array. Cg11484872, cg21370522, cg01569083 and cg03037030 (aka CpG1, CpG 5 CpG 7 and CpG 8, respectively. For reference cg11484872 was the probe used to confirm the veracity of the Illumina data.

4 Discussion

4.1 Candidate Gene Selection

For a gene to be considered as a candidate it will have to be highly ranked in regards to its degree of methylation or demethylation, it will have to have some relevance to the pathology of RA, and it should preferably be relevant to the role of CD4+ T cells.

The TNF gene has been found to be significantly hypomethylated in a methylation profile assay (Figure 3 and Figure 7). Furthermore, as will be explained below, it can be seen the TNF is responsible for initiating various cascades which cause inflammation and drive the pathogenesis of RA. Additionally, considering the relevance of the inflammatory pathway to CD4+ T cells it can be seen that the TNF gene is also relevant here too. Due to it satisfying all the outlined criteria the TNF gene has been chosen as a candidate.

4.1.1. Demonstrating TNF's Relevance in RA pathology

TNF was first elucidated as being involved with RA when a research group found that it and various other cytokines were being continuously produced in the synovium when usually cytokine production should be a transient occurrence. This dysregulation was further investigated by extracting cells from the synovium, culturing them and over the course of a week investigating any cytokine dysregulation. By using neutralising antibodies specific for certain cytokines they were able to deduce a pro-inflammatory hierarchy; they found that antibodies targeting TNF inhibited the production of all the other cytokines. It was this discovery that demonstrate that TNF was the apex pro-inflammatory cytokine and that TNF was involved in RA (140-145).

Out of this the first anti-TNF drug was produced – infliximab. Infliximab produced significant clinical effects in treating RA; importantly that its effects were long-lasting. This conclusively tied the pathogenesis to TNF. Clinical trials demonstrated that clinical improvements could be made better by taking methotrexate (MTX) as well. Subsequent studies into TNF elucidated its pathways (140-145).

4.1.2. NFkB activation of the TNF-dependant cytokine cascade in RA

TNF has been identified as the prime initiator of the cytokine cascade that is associated with in the inflammation-response seen in the pathology of RA. However a 2002 study indicated that although neutralising TNF reduces joint-damage in RA, TNF itself does not directly cause the joint damage; instead the cytokines and their pathways that it commands.

TNF produces these cytokines by activating NFkB activation via the TNF cascade (147). NFkB activation results in the initiation of the TNF-dependent cytokine cascade, which is critical in the formation of the inflammation and subsequent damages seen in the pathology of RA. The TNF-dependent cascade elucidated via anti-TNF drug development is complex and involves several cytokines namely: TNF, IL1, IL6 and GM-CSF (148). Specifically, TNF induces an intracellular cascade, via interacting with TNFR1, which typically culminates in the activation of NFkB – NFkB then stimulates the production of IL1, IL6 and GM-CSF. Here we will discuss here how IL1, IL6 and GM-CSF, the damaging elements of this cascade, carry out their effects and further RA pathogenesis.

4.1.2.1 Interleukin 1

IL1 is a family of 11 cytokines, 7 of which are pro-inflammatory, 3 are receptor antagonists and 1 which is anti-inflammatory. Although it is distinctly different from TNF there is a lot of overlap; both TNF and IL1 stimulate IL6, IL2 receptors, colony stimulating factors and acute phase proteins as a result both are effective pro-inflammatory cytokines. IL1 is stimulated via NFkB transcription which can be initiated by TNF. There is a body of evidence to suggest that IL1 contributes to the pathogenesis of RA. For example two members of the IL1 family, IL1α and IL1β, when activated, act as powerful pro-inflammatory cytokines. They trigger localised vasodilatation and attract monocytes and neutrophils to sites of tissue damage and stress.

Importantly, these cytokines are prime inducers of matrix enzymes and thus serve as powerful mediators of tissue damage by changing both cartilage and bone homeostasis. Uncontrolled IL1 activation is a primary factor of some inflammatory diseases and has been directly linked to RA pathology. This has been demonstrated in vivo where IL1β overexpression in rabbit's knee joints results in arthritis with clinical and histological symptoms of RA (149).

4.1.2.2 Interleukin 6

IL6 is responsible for stimulating B cells to produce immunoglobulins (Ig). In RA, antibodies towards citrullinated peptides and IgM and IgG rheumatoid factors are typically increased. IL6 does this by stimulating B cells to differentiate into plasma cells which in turn produce said Igs. The significance of this B cell activity in RA pathology is evident due to clinical improvement seen when B cells are depleted. IL6 is also involved in T cell proliferation and differentiation, specifically into TH17 cells. These TH17 cell produce IL17 which in turn can stimulate further IL6 production, and increase neutrophil migration by stimulating IL8 (147, 148). This process in human models is usually done in combination with IL1 β and IL23. Taking IL6's significant involvement in the

development in the adaptive immune response it can be seen how it may be involved in the pathogenesis of RA.

A hallmark of RA is the presence of chronic inflammation. There is evidence to suggest that IL-6 is involved in the shift from acute inflammation to chronic. IL6 is directly involved in activation of neutrophils which in turn migrate to tissues and release reactive oxidative species (ROS) and proteolytic enzymes which results in tissue damage – in RA this would be seen as joint damage. Neutrophils release soluble IL6R when they've reached the inflammation site; this causes recruitment of leukocytes via activation of adjacent epithelial cells which in turn release pro-inflammatory chemokines.

There is a correlation between matrix metalloprotinases' (MMPs) expression and articular cartilage damage in RA; this is due to overexpression of MMPs from cells lining the synovium. In early RA patients there is a correlation between levels of C-reactive protein (CRP), IL6 and pro-MMP3; indicating a relationship between MMP activity and IL6 (147,148).

Another manner through which IL6 can contribute towards joint damage in RA is through its interaction with vascular endothelial growth factor (VEGF). VEGF is a key mediator and promoter of the migration and proliferation of endothelial cells; it also increases local vascular permeability and inflammation. IL6 increases the levels of VEGF in the presence of sIL6R; VEGF levels are also positively correlated with RA disease activity.

The narrowing of joint-space and erosions are key characteristics of joint damage in RA. Osteoclasts are the cells that play a primary role in the erosions in RA, and IL6 has been shown to increase osteoclast recruitment by interacting with haematopoietic stem cells. A study demonstrated that IL6 deficient murine models with antigen-induced arthritis (AIA) had fewer osteoclasts at the bone erosion sites and a reduced severity of arthritis in comparison to wild-type (147,148).

The most convincing piece of evidence that IL6 is involved in RA pathology is that it was found that

neutralising IL6 results in clinical improvement. The drug, tocilizumab, which is an anti-IL-6R mAb, was given to RA patients who had a non-responder to one or more disease-modifying rheumatic drug (DMARD) demonstrated significant clinical improvement of disease activity within 1 week of monotherapy; this improvement held up until week 8. Later, larger trials demonstrated that these improvements could be sustained with repeat dosages. In addition it was found that combination therapy with MTX improved efficacy of the treatment, this was demonstrated through both symptom amelioration and reduction in joint damage (147,148).

4.1.2.3 Granulocyte-macrophage colony-stimulating factor

GM-CSF is a haemopoeitic growth factor but also modifies the functions of myeloid cells. For instance, GM-CSF can polarise macrophages into M1-like macrophages which in turn produce a variety of cytokines such as TNF, IL6, IL12p70, and IL23. On the other hand, neutralisation of GM-CSF results in the favouring of the formation M2-like macrophages which instead produce anti-inflammatory cytokines (150)

When you then take into account that pro-inflammatory cytokines such as TNF, IL1 and IL23 have been found to stimulate the production of GM-CSF and that GM-CSF is found in high levels in the synovium of RA patients. It is plausible to suggest that GM-CSF could propagate and sustain the pro-inflammatory cytokine cascade through polarisation of macrophages. (151)

In fact there is evidence to support the position that not only is GM-CSF involved in RA pathology but in some circumstances it can be the prime pro-inflammatory cytokine. As it was found that severe inflammatory arthritis can be caused by GM-CSF overexpression without TNF overexpression. This was further highlighted in RA patients who didn’t respond to anti-TNF drugs but did respond well to anti-GM-CSF drugs (MOR103 and Namilumab) (152).

Taken together it can be seen how TNF is capable of causing damage in RA without being directly involved in the damage process via NFkB activation. In addition, if NFkB is TNFs means of

implementing damage than other pathways which also stimulate NFkB should be considered.

4.1.3 Other potential gene candidates

Other considerations aside from TNF, are B3GALT4 and ABI3. Though B3GALT4 is the highest ranked gene in terms of differential methylation it doesn't appear to have any published specific relevance in RA. Furthermore the same can be said for ABI3, though ranked highly and possess the highest number of coexpressed genes (data not shown) there doesn't appear to be any published data linking it the RA pathology. Based on this, TNF is the superior choice for a candidate gene.

4.2 Candidate Gene Region Selection

Due to the limitations of bisulphite sequencing, the TNF gene which is 4.7kb long and possessed 24 significant CpGs (see appendix 7.2). This region which encompassed all 24 CpGs was divided into 6 regions for individual analysis. Out of the 6 regions only one was to be chosen as the candidate region. Based upon the sequencing data of the 6 primer products, products 4, 5 and 6 proved to be poor candidates and as a result they were omitted from further analysis. One of the main reasons thought to be behind these yielding poor sequencing data is the presence of repeat nucleotide sequences (as seen in Figure 7). These repeated thymine nucleotide sequences, when exceeding 11 repeats disrupt the sequencing data by introducing contractions or extensions. The further above 11 repeats these sequences were, the more contractions or extensions will be introduced (153).

This appears to be the reason for why products 4 and 5 are poor candidates as they have very long nucleotide repeats (up to 15 nucleotides). However product 6 had the same amount of repeats as product 2, which was found to give a good sequence. Therefore it is unlikely product 6's poor sequence data is due to a nucleotide repeat sequence. The answer may lie in Figure 6c. It can be seen that primer product 6 has a very strong secondary band occurring beneath the

desired product's band. It is likely this strong secondary product is due to the primers binding elsewhere on the DNA yielding a shorter product. It should also be noted that the product sequences containing nucleotide repeats of 10 or more (2, 4 and 5) have faint secondary products occurring below the desired products. Based upon their position and their length they look like primer dimers.

Primer products 1, 2 and 3 all provided very good sequencing data in regards to alignment and coverage with their expected products. Out of these three, primer product 3 was chosen as it contained more significant CpGs than 1 and 2, and also it is located in the promoter region of the gene (see Figure 3); implying that any epigenetic modifications in this region would have a larger effect on transcription than the other primer products designed, and thus would be more desirable to analyse as it would be more likely to contribute towards RA pathology. So, in summation, the promoter region of the TNF gene was chosen to be sequenced to confirm the veracity of the methylation profile data.

4.3 Confirming the Candidate Gene's Veracity

The methylation profiling assay interrogated 4 CpGs in the candidate region and showed 1 CpG significantly hypomethylated ($p \leq 1x10^{-2}$) in the TNF gene candidate region. However, the bisulphite-DNA Sanger sequencing conducted for confirmation showed that 7 out of 8 CpGs were significantly hypomethylated. This discrepancy isn't due to differences in methylation levels, but instead to differences in sensitivities, the methylation profiling assay was only able to analyse 4 CpGs maximum in the candidate region. The CpGs that were shared in this region in both methodologies both demonstrated significant hypomethylation in comparison to HC.

We can be confident that the results we have are veracious as the CpGs examined are significantly hypomethylated in comparison to HC data. Furthermore the results draw the same conclusions as those found previously in the data acquired by using a different methodology. The point of using a different methodology for confirmation is to rule out data-influencing factors which may be present in one methodology.

Now it has been asserted that the results we have are reliable, it should be discussed their relevance and meaning in the pathology of RA. However, inferences from this data should be carefully drawn as the confirmation of methylation level from the first assay doesn't confirm the entire dataset of the methylation profiling assay. It only confirms that the TNF gene is significantly hypomethylated in comparison to HC in the 1 CpG in the region analysed. That said, the region that the 7 hypomethylated CpGs are located in is the promoter region of the TNF gene, therefore the hypomethylation of the region is likely to impart a larger effect upon its ability to be transcribed in comparison to other CpG regions.

4.4 Variance of TNF methylation

As can be seen in Figure 7 there is a natural variation to the degree of methylation of CpGs in the TNF promoter region – both in HC and RA. However across all CpGs RA methylation is lower than HC. Variance is a common occurrence in nature and so to see some variance in both HC and RA populations is to be expected. Furthermore, such variance could aid in determining whether there is a healthy reference range of TNF methylation. Whilst such a determination would likely require a larger sample size and a better consideration of potentially confounding factors, such as disease severity, age and drug treatments. It can be seen from the pie charts in Figure 7 that the methylation level of the CpGs in RA patients rarely rises above 25%, furthermore the HC CpG methylation levels rarely drop below 25%. From this it can be said, crudely at least, that a TNF promoter CpG methylation level below 25% in the CD4+ T cell population is characteristic of RA. Now whether this 25% is the threshold which can predispose a healthy person to RA can't be asserted, but it should be an area of future study.

4.5 The TNF cascade

4.5.1 The TNF receptors

The members of the TNF ligand family demonstrate their functions by interacting with their complementary membrane receptors, these receptors make up the TNF receptor (TNFR) family (154). Two receptors, TNF receptor type 1 (TNFR1) and TNF receptor type (TNFR2) bind to the soluble TNF (sTNF), membrane-bound TNF (mTNF) and Lymphotoxin α (LTα). (155) TNFR1 is ubiquitously expressed in most tissues, whereas expression of TNFR2 is highly regulated and is typically found in cells of the immune system. In the majority of cells, TNFR1 appears to be the prime mediator of TNF-signalling, however in lymphocytes TNFR2 seems to play a major role (155); though due to the fact that TNFR2 can only be fully activated by mTNF its efficacy is often overlooked. The cause for TNFR2 only being able to be activated by mTNF is not fully understood but the different binding affinities of the individual ligand-receptor complexes may be a reason (155).

As well as being mediators of the TNF pathway TNFR1 and TNFR2 also have the ability to be inhibitors as their extracellular domains can be proteolytically cleaved, yielding soluble receptor fragments which have the capacity to neutralise TNF molecules (154). That said, the binding affinity of the soluble receptor fragments to TNF molecules is not as strong in comparison to their membrane-integrated forms.

TNF possesses a dichotomous functionality. On one hand it has the capacity to induce apoptosis but on the other it can also stimulate cell proliferation and survival pathways. One way the TNF ligand can induce different functions is dependent upon which of the TNF receptors it binds to. For instance due to the fact that TNFR2 doesn't possess a death domain and TNFR1 does then the choice for TNF-induced cell survival can be potentially made at this point in the pathway.

The other point is when TNFR1 is stimulated, as it can stimulate two pathways. TNFR1 activation can either stimulate cell death, or inflammation and cell survival. It is able to achieve this functional versatility by being able to interact with several complexes (Complex I, Complex IIa, Complex IIb and Complex IIc). The mechanism behind this live-or-die decision isn't well understood. It is thought that the levels of c-FLIPL and RIPK1 are vital components to this decision. c-FLIPL acts as an inhibitor of caspase-mediated apoptosis by binding to Caspase-8 and blocking its function. RIPK1 is thought to be ubiquitinated and tied to complex I. However, if the RIPK1 has its ubiquitin removed, it gets trafficked to and interacts with apoptotic machinery (156, 157).

One example of this dichotomous function of TNF in vivo can be seen in liver cells, as liver regeneration after partial hepatectomy in TNFR1-deficient models, hepatocyte DNA synthesis is impaired, indicating that TNF signalling through TNFR1 is involved in liver regeneration (158). In contrast to this, in models of acute hepatotoxicity TNF acting via TNFR1 appears as a key player in liver destruction (159). This destructive capability is largely due to its powerful pro-inflammatory and immune-stimulatory functions. As a result TNF is a key mediator of progression in many autoimmune diseases. Such examples of TNF-involved autoimmune diseases are rheumatoid arthritis and Crohn's disease. TNF involvement is demonstrated in these diseases as significant clinical improvements can be achieved when patients are treated with TNF-neutralising agents (160,161).

4.5.2 Downstream effects of TNF-TNFR signalling

The downstream non-apoptotic cascade after TNFR1 activation involves the activation of P38 and JNK by MAP3K5 and MAP4K5 kinase pathways. The other part involves the phosphorylation of the IKK complex, specifically the IKKβ subunit. This results in the dissociation of the NFkB-IkBα complex and allows NFkB to translocate into the nucleus. The TNF-TNFR2 interaction activates NFkB-Inducing Kinase (NIK) that causes in the same dissociation of the NFkB-IkBα complex and

subsequent NFkB translocation. The canonical and non-canonical TNF pathways are outlined in Figure 9.

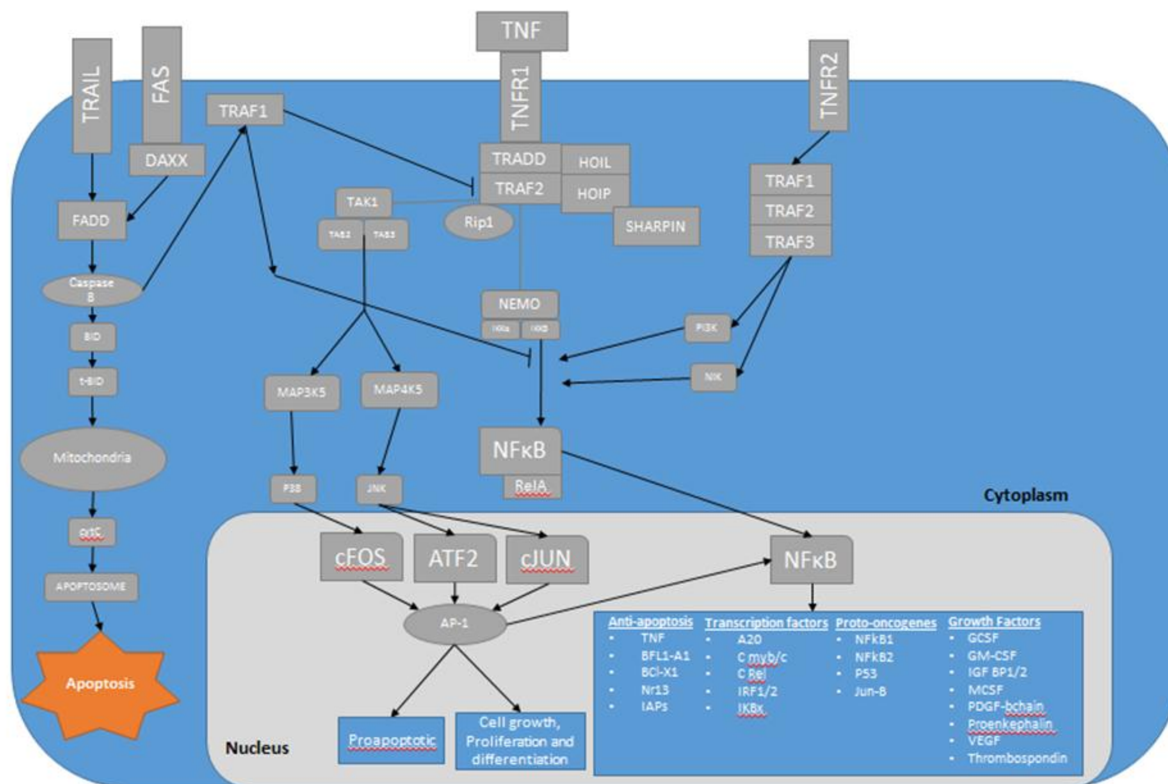


Figure 9. TNF pathway showing how TNF can either initiate pro-apoptotic or proliferative signalling via TNFR signalling molecules.

4.5.2.1 P38

P38 is known to mediate the response to cellular stress factors such as osmotic shock and UV light. To date, five isoforms of p38 MAPK have been identified. The p38 family can be divided into two groups, p38α/β/β2 and p38γ/δ, based on their ability to respond to different stimuli (162). P38 contributes to the process of leukocyte recruitment by activation of several processes: selectin-mediated rolling, integrin activation, diapedesis of leukocytes towards an injury site and activation of the effector functions in leukocytes.

Selectin-mediated rolling is the first step in leukocyte extravasion; this process is requires the cell to have adhesion that allows rolling along vessel walls. This temporary adhesion is provided by a family of adhesion molecules called ‘selectins’ (L-, E- and P selectins). L-selectins are the typically expressed selectin on the surface of leukocytes, and E- and P-selectins are usually expressed upon the surfaces of endothelial cells and also have the role of aiding leukocytic adhesion (163,164).

After rolling, and their destination reached, a stronger adhesion force is required between the endothelium and leukocytes. This firmer adhesion is produced by the binding of β1 and β2 integrins on the leukocytes to intracellular adhesion molecule 1 (ICAM1) and vascular cell adhesion molecule 1 (VCAM1). Once firmly binded to the vessel wall, the leukocyte migrates throughout the intracellular junctions between the endothelial cells following a gradient of chemokines to reach the site of injury. The type of chemokines produced by an injury site can attract specific subsets of leukocytes. For instance in COPD, the cytokines produced attracts an influx of neutrophils, whilst in asthma the cytokines attract a predominantly macrophage population (162).

Several studies have demonstrated that a p38 MAPK-specific inhibitor can initiate NFkB-dependent transcription (165, 167). However, phosphorylation of the NFkB subunits, nuclear translocation and DNA binding of NFkB were not affected by this p38 inhibitor. This suggests that NFkB and p38 MAPK are activated by separate effector pathways, which might converge further downstream in the cell nucleus (166, 167).

4.5.2.2 c-Jun N-terminal kinases

The mammalian c-Jun N-terminal kinase (JNK) was initially called stress activated protein kinase (SAPK) because they were activated by a variety of environmental stresses (168,169). Later, the JNK pathway was also discovered to be activated by TNF, IL1 and other growth factors. JNK is involved in several physiological and pathological processes. Specific stimuli trigger the activation of MAP3Ks, which then phosphorylate and activate the MAP2K isoforms MKK4 and MKK7, which in turn phosphorylate and activate JNK (149). JNK was discovered to phosphorylate c-Jun at the NH2-terminal Ser63 and 73 residues, and thus termed JNK. In addition to c-Jun, JNK can phosphorylate transcription factors such as JunB, JunD, c-fos, ATF2 and ATF3 (168,169). These transcription factors along with c-Jun, make up the Activator Protein-1 transcription factor (AP-1), which regulates the expression of several stress-responsive genes. The JNK class of enzymes comprises of three main types: JNK1, JNK2, and JNK3 (169, 170). The first two are ubiquitous, whereas the third is restricted to the brain, heart and testis. Differential splicing and exon usage results in multiple isoforms of JNK1, 2 and 3 genes. The alternative forms of each JNK1, 2, and 3 appear to differ in their ability of bind and phosphorylate different substrate proteins. Targeted gene disruption of each JNK has also defined differential functions for JNK1, JNK2 and JNK3 in many different cell types. JNK1 or JNK2 deletion resulted in defective T cell activation and differentiation (170,172).

JNK, by phosphorylation, modifies the activity of several proteins that are located at the mitochondria or perform roles in the nucleus. JNK is capable of activating and inhibiting various downstream molecules. Through JNK, TNF can regulate several cellular functions including, differentiation, cell growth, survival and apoptosis through the activation and inhibition of other small effector molecules.

4.5.2.3 NFkB pathway

The canonical nuclear factor kappa-light-chain-enhancer of activated B cells (NFkB) pathway has been defined primarily in response to TNF and IL1 signalling, NFkB activation is also widely implicated in inflammatory diseases such as RA, inflammatory bowel disease, asthma, and chronic obstructive pulmonary disease (173-177). Consequently much attention has become focused on the development of anti-inflammatory drugs targeting NFkB (178).

NFκB activity at sites of inflammation is associated with activation of the TNF canonical pathway and Rel-containing complexes. There have been several studies to show proinflammatory cytokine and chemokine production by disease tissue is NFkB-dependent; for example, using fibroblast-like synoviocytes from RA patients (179,180).

There has also been correlation of NFkB activation with inflammatory disease in animal models of arthritis (181). That said, it is clear from genetic experiments in mice that NFkB activation is not necessarily proinflammatory and has complex roles in the inflammatory response. The role of RelA as a critical component of the canonical pathway has been demonstrated with the development of RelA and IKKβ knockout mice (182, 183).

A 2001 study showed the involvement of NFkB in both the onset and resolution of acute inflammation in a single model system using pharmacological inhibitors (180). Studies have confirmed the expected role of NFkB in proinflammatory gene-induction during the onset of inflammation but also showed a putative role for NFkB in the expression of anti-inflammatory genes and in the induction of leukocyte apoptosis during the resolution of inflammation. Also these studies suggest an anti-inflammatory role for NFkB that limits the bactericidal and tumouricidal function of macrophages (180-184).

Gene knockout studies have also shown that NFkB subunit proteins can have both pro- and anti-inflammatory roles. For example, homodimers of the p50 subunit of NFkB, which lack

transactivation domains, have been shown to repress expression of NFkB target genes and inhibit inflammation (185). An anti-inflammatory role of NFkB was also demonstrated in inflammatory bowel disease in which "$p50^{-/-}p65^{+/-}$" murine models were more susceptible to H. hepaticus-induced colitis (186-187). As a result study data suggests that NF-kB can have an anti-inflammatory role, as well as an inflammatory role, by directly inhibiting expression of proinflammatory genes and by manipulating the expression or activity of anti-inflammatory cytokines such as IL10 (188). These multiple roles can be seen in T cells as NFkB can help T cells avoid glucocorticoid-mediated apoptosis during positive selection. Whilst on the other hand, NFkB has the opposite role in mature peripheral T cells, promoting apoptosis by increasing FasL expression, which may be linked to the termination of T cell responses (189).

4.5.2.4 Non-canonical activation of NFkB

The alternative NFκB pathway is characterised by the inducible phosphorylation of NFkB2 by IKKα, leading to activation of RelB/NFkB2 heterodimers. The upstream kinase that activates IKKα in this pathway has been identified as an NFkB-inducing kinase (NIK).

Recent studies suggest that IKKα has evolved distinct, but possibly complementary, roles in adaptive immunity and inflammation. IKKα functions to promote the resolution of inflammation by switching off the canonical NFkB pathway (176), and regulates the development of adaptive immunity through the alternative NFkB pathway (189,190). Although inflammation is classically considered to prime the adaptive response, the resolution of inflammation is also required to avoid tissue injury while supporting the development of the immunological memory. Cross-talk between the alternative and canonical NFkB pathways may regulate the transition from acute inflammation to antigen-specific immune responses that drive autoimmune diseases such as RA and multiple sclerosis. It is thought that, inhibition of IKKα may represent a therapeutic target to prevent autoimmune inflammation without compromising the innate immunity (191).

4.5.3 TNF cascade-related methylation aberrations in RA CD4+ T cells.

Now the key elements of the TNF cascade have been introduced, the aberrations found in the methylation levels of elements directly and indirectly involved in the TNF cascade should be analysed. This is a relevant area to investigate as any significant differences in methylation levels of downstream effectors likely highlight differences in the regulation of the TNF pathway and/or NFkB activation. It should be noted that data indicating hypomethylated genes alone doesn't indicate a tangible increase in the translation of a gene, it merely shows that the gene is primed and more ready for transcription. As a result the following analysis of the TNF cascade is simply in order to highlight aspects and elements which could be better or worse primed for transcription (Figure 2).

4.5.3.1 Inhibition of NFkB with DAXX and TRAF1

The apoptotic pathway is available to TNF due to TNFR1 having a death domain (TRADD), which in turn can interact with DAXX and in turn initiate a caspase cascade. However in this scenario it is possible that DAXX's up-regulated inhibits the TNF cascade, in conjunction with TRAF1, and not necessarily for the 'purpose' of initiating apoptosis (192).

As seen in (Figure 2a) FAS-associated DAXX initiates the caspase cascade by interaction with FADD-interaction. Now, FADD can then initiate the activation of caspase 8 who in turn can activate tBID or can also enzymatically interact with TRAF1. Caspase 8 cleavage of TRAF1 results in a product which inhibits the formation of IKKb dissociation and subsequent NFkB activation.

A 1998 study showed that NfkB activation blocks the activation of caspase 8 and its subsequent caspase cascade. The study noted that TRAF1, TRAF2, c-IAP1 and c-IAP2 were all required to fully block TNF-induced apoptosis. However the methylation data only indicates that TRAF1 is differentially methylated and not the others highlighted in this study. That said even if all were

hypomethylation, it wouldn't prove they are all being translated. (193).

Regardless, only TRAF1 being primed for transcription highlights either the greater complexities of transcriptional regulation or that TRAF1 can inhibit via another process. Whether caspase 8 interacts with TRAF1 or BID likely depends on their availability and as a result, an upregulated TRAF1 inhibiting caspase 8 via competitive inhibition could very well be a means to ensure this.

However, caspase 8's interaction with TRAF1 can still contribute towards a pro-apoptotic signal, likely through NfkB deactivation, for instance a 2003 study demonstrated in carcinoma cells that TRAF1 cleavage is an indicator of chemo-induced apoptotic signalling. However this pathogenesis isn't known for chronic NFkB activation as found in RA. So therefore TRAF1 cleavage wouldn't be as inducing of apoptosis in RA due to relatively higher level of NFkB activation (194). Furthermore, it should also be noted that DAXX is also able to inhibit NFkB function directly if it is located in the nucleus, further supporting the idea of it possessing functions outside of simple apoptotic initiation (194),

More evidence suggests the apoptotic cascade is downregulated due to the presence of differentially methylated genes which could act to inhibit the apoptotic caspase cascade. These include: hypomethylated BLC-2 which is the inhibitor of t-BID, and inhibits the release of cytochrome C from mitochondria; and further downstream a stimulator of apoptosomes is hypermethylated, UACA. In summary the methylation data suggests that TRADD-initiation of DAXX and caspase 8 could be in order to lessen the activity of NFkB and not necessarily for the purpose of apoptotic initiation.

Whether or not apoptosis or proliferation is being favoured can be demonstrated in this study fairly easily. It can reasonably be assumed that RA CD4+ T cells are selecting proliferation over

apoptosis due to the presence of a stable, epigenetically distinct cell population.

4.5.3.2 CD30

CD30 is a cell-surface receptor whose role is involved with lymphocyte activation and survival; CD30 does this through its ability to induce the transcription factor NF-kB (Figure 2b). In contrast, CD30 has also been implicated in the stimulation of apoptotic cell death of lymphocytes (195). It has been found that CD30 signal transduction renders cells sensitive to apoptosis induced by the TNFR1. However, this sensitisation is dependent on the TRAF-binding sites in the cytoplasmic domain of CD30. One of the proteins that binds to these sites is TRAF2. Study data suggests a model in which CD30 limits its own ability to transduce cell survival signals through reductions of levels of TRAF2 (195). In addition it was found that reduction in levels of intracellular TRAF2 and its associated proteins also increased the sensitivity of the cell to apoptosis to death-causing complexes such as the TRADD-TNFR1 complex. Additionally the study found that, the expression of a dominant-negative form of TRAF2 was found to potentiate TNFR1-mediated death (195). A potential mechanism has been elucidated, through which CD30, as well as other TRAF-binding members of the TNFR superfamily, can negatively regulate cell survival. Furthermore CD30 signal transduction and concomitant TRAF2 degradation reduces the ability of TNF to induce NFkB signalling (195).

As a result hypomethylation might indicate overexpression of CD30. Although CD30 would likely be involved in promoting cell survival excessive stimulation of CD30 would result in promotion of TNFR1-mediated apoptosis due to TRAF2 depletion. However, as mentioned above, the molecules involved in the initiation of the TNFR1-TRADD apoptotic pathway might well be used for NFkB suppression, as opposed to cell death initiation.

4.5.3.3 LIGHT/LTα/HVEM role in T-cell activation

Tumour necrosis factor super family member 14 (LIGHT) and LTα have been identified as herpes virus entry mediator (HVEM) ligands, based on their ability to bind to HVEM. As well as HVEM, LIGHT was additionally identified as a ligand for LTβR (196). LIGHT expression is more restricted and tightly regulated than HVEM with LIGHT mRNA expressed predominantly in the spleen and brain with relatively low levels of expression in peripheral blood leukocytes, amongst other sites (196-197). LIGHT is upregulated on T cells upon activation and thus is reciprocally expressed compared to HVEM which is high on unstimulated T cells, declines after activation and rises several days later (196, 198).

The expression of LIGHT on T cells in murine models leads to over-activation and expansion of activated T cells, resulting in autoimmune diseases and severe inflammation of the gut (200) which correlates with accumulation and activation of DCs (199-201). Murine models deficient in LIGHT have a defect in CD8+ T-cell activation and potentially in thymic selection (202-204).

LIGHT is directly and indirectly capable of stimulating NFkB and JNK (Figure 2c) which would achieve the same function as the TNF-dependant cytokine – stimulating cell survival, cytokine production, and proliferation (205-209). Furthermore, not only is LIGHT found to be hypomethylated, but so is TRAF5, which is an important associated factor in signal propagation. Due to LIGHT being tightly regulated and being specifically upregulated with T cell activation; it is likely that hypomethylation primes the cell to be better able to be transcribed upon T cell activation.

4.5.3.4 RANK induction via NIK

The receptor activator of NFkB's (RANK) cytoplasmic domain interacts with TRAF2, TRAF5, and TRAF6 (Figure 2d). RANK's overexpression also is found to activate NFkB and JNK pathways. A study demonstrated that TRAF2 and TRAF5 bind to different TRAF binding motifs than TRAF6 which interacts with a specific motif of RANK. Furthermore, it was found that the TRAF6-binding region, but not the TRAF2 or TRAF5-binding region, is necessary and sufficient for RANK-induced NFkB activation. It was found that RANK activates NFkB via NIK, as a kinase mutant of NIK inhibited RANK-induced NF-kB activation. In regards to JNK activation however, it was determined that TRAF2 was more important than TRAF5 or TRAF6 in JNK activation (210).

RANK-induced propagation of an NFkB and JNK signal is likely to be present in RA CD4+ T cells, as TRAF5, TRAF6 and NIK are all involved in this process and all found to be hypomethylated. Furthermore, highlighting the importance of TRAF2 in NFkB activation, TRAF2 has been found to be an important component in the outlined signalling mechanism.

4.5.3.5 cFOS activation via the TCR

TNF has been shown to activate Zap70 (Figure 2e). Zap70 has in turn been shown to activate linker for activation of T-cells (LAT). T-cell-receptor (TCR)-mediated LAT phosphorylation is crucial for the membrane recruitment of signalling complexes required for T-cell activation. Although phosphorylation of LAT is needed for recruitment and activation of signalling proteins, the exact mechanism associated with this event is not clear (211).

The signalling pathway mentioned essentially involves LAT forming a complex with RAS which is released when activated. RAS can then interact with the KSR1-RAF-MEKK-ERK complex. Which releases Erk and in turn stimulates the transcription factor c-FOS – which is a proto-oncogene. There does appear to be a route by which TNF can induce c-Fos activation and contribute to cell

survival other than by inducing NFkB, P38 and JNK. That said the c-Fos activation caused by Erk signalling is thought to be quite weak. So any pro-oncogenic effect caused by this mechanism would likely be marginal (212). As previously mentioned, ERK signalling could instead be involved in the regulation of methylation maintenance machineries, DNMT1, and continue to maintain and contribute towards the aberrant epigenetic mosaic.

4.5.3.6 NFkB Signal Modulation

Overall the differentially methylated pathways appears to be primed to stimulate NFkB and consequently support cell survival and proliferation of the CD4+T cell.

On the other hand there seems to be epigenetic modifications that would dampen NFkB activation; DAXX-mediated NFkB inhibition can occur either by the caspase 8 cleavage of TRAF1 (Figure 2a) or by direct inhibition by DAXX in the nucleus (Figure 2). The former would seem to target the canonical TNF cascade directly whilst the latter would target generic NFkB activation.

Further evidence for NFkB inhibition lies in the fact that there is a dampening of genes responsible for NIK activation. For instance the LTBR gene is hypermethylated, potentially inhibiting LTa or LIGHT stimulation of NIK (Figure 2c), and also the antisense RNA version of NIK, MAP3K-AS1 is hypomethylated, thus potentially downregulating NIK from the point of transcription (Figure 2).

4.6 TNF Gene coexpression network

Another aspect to analyse is whether the hypomethylated TNF belongs to a gene coexpression network and if any members of said network are also hypomethylated. A gene co-expression network is a phenomenon where a set of genes are expressed together, across a range of conditions. Usually this is to do with these genes being expressed by the same transcriptional regulator, functionally related, or are members of the same protein complex or pathway. Gene co-expression networks are stated to be a powerful approach to accelerate the elucidation of molecular mechanisms underlying important biological processes (213).

Considering that TNF can stimulate the production of more TNF, and that TNF is prevalent within RA pathology, it can be seen that genes that are coexpressed with TNF that are also hypomethylated would be sensitive to transcription in this pathology. As a result it is important to consider what effects any genes found to be both coexpressed with TNF and hypomethylated would likely have on the cell. TNF was found to be coexpressed with 10 other genes in the top 50 ranked hypomethylated genes.

4.6.1 LNPEP

LNPEP is an enzyme that is responsible for cleaving peptide hormones. It has also been associated with inflammatory conditions, specifically with psoriasis (214). Interestingly, TNF inhibitors are capable of inducing symptoms of psoriasis (215). Whether or not TNF inhibitors have an effect on LNPEP expression isn’t known. However studies show that both TNF and LNPEP are involved in inflammatory conditions (214, 215), and are also coexpressed together.

4.6.2 RPS18

RPS18 is a subunit of the 40s ribosomal complex which is the largest component of several transcription initiation complexes. The orthologue in e.coli of RPS18 is involved in the initiation of translation, specifically the initial fMET-tRNA binding. Therefore it is likely that RPS18 is also involved in translation initiation and thus is perhaps is a rate limiting step to this process if its expression was limited. It is seen that RPS18 is upregulated at the protein level in cancer, this is consistent with RPS18's role as cancer cells tends to have greater transcriptional requirements than healthy cells. In CD4+ T cells in RA it has been found that RPS18 is hypomethylated, indicating that this cell type has been primed to favour RNA translation (216)

4.6.3 GPSM3

GPSM3 is a regulator of G protein-coupled receptor signalling, which has its expression restricted to leukocytes and lymphoid organs. It has been demonstrated that SNPs in GPSM3's transcriptional start site is inversely correlated with RA. That is that SNPs that reduce the ability of GPSM3 to be transcribed provide a protective affect against RA (217) Therefore it is consistent with expectations to see that GPSM3 is hypomethylated, and thus primed for transcription, in CD4+ T cells from RA patients.

4.6.4 DAXX

DAXX is a protein responsible for initiating a caspase-mediated cascade which can result in apoptosis of a cell. Furthermore DAXX-induced apoptosis can be activated by TNFR1. Therefore one might expect that DAXX would be downregulated in order to promote cell survival in CD4+ T cells from RA. However it was found that DAXX was significantly hypomethylated in this cell type.

Upregulating an apoptotic gene may seem contradictory in considering what has been found so far. However it has been found that silencing DAXX can actually sensitise cells to several apoptotic mechanisms (241). Furthermore, several studies have found that DAXX may have a more complex role and may possess anti-apoptotic functions (242). Therefore, due to DAXX being multi-faceted it is difficult to state what is expected with only methylation data available.

4.6.5 TLR9

TLR9 is an important receptor expressed in immune system cells including dendritic cells, macrophages, natural killer cells, and other antigen presenting cells.TLR9 is also an important factor in autoimmune diseases, so much so that there is active research into synthetic TLR9 agonists and antagonists that could help regulate autoimmune inflammation (218, 219). It has also seen to be overexpressed in RA monocytes (220) and that it has a role to play in the pathogenesis of RA (221). As a result it is consistent and expected to see it being hypomethylated in CD4+ T cells from RA patients (222).

4.6.6 ZNF217

ZNF217 is a zinc finger protein is an oncogenic factor involved in various human cancers. It's amplification and increased copy number has been found to be linked to poor clinical prognosis. ZNF217 is so correlated to some cancer pathogeneses that it is being used as a potential biomarker and therapeutic target (223). ZNF217 has been found to be able to attenuate apoptotic signals resulting from telomere dysfunction and may promote neoplastic transformation and later stages of malignancy (223). TNF has been associated with breast cancer progression (224). ZNF217 whilst it is part of a HDAC and initial research thought it be involved in transcriptional repression. Further research has shown that it regulates specific genes and possess a more complex role than generically repressing transcription and is thought to contribute to tumorigenesis by dysregulating gene expression programs (225-228). From this it can be seen that if ZNF217 is being hypomethylated in CD4+ T cells it is indicative of the cell being primed

towards cell survival and proliferation - this strengthens the argument that the RA CD4+ T cells are primed for survival instead of apoptosis.

4.6.7 TMC6

TMC6 has been implicated in the rare skin disorder epidermodysplasia verruciformis, specifically when it possesses inactivating mutations. Although the exact mechanism and significance of this phenomenon remain to be determined, it seems probable that changes in nuclear and nucleolar zinc concentration could strongly influence function of the cell (229).

In this study it was found that TMC6 plays a role in the immune response; stimulating antiviral factors such as chemokines, cytokines and growth factors. In particular a study showed that the proinflammatory cytokine, IL6, was expressed less in knockout experiments. Specifically the effects caused are thought to be mediated by JNK and ELK1 activation (229). In addition to TMC6's outlined proinflammatory capacity, zinc influx has been stated to be an ionic signalling molecule after T cell activation and is essential for innate and adaptive immune response in another study (230). As a result it is understandable that TMC6, a proinflammatory, T cell activator, would be hypomethylated in RA CD4+ T cells.

4.6.8 PTPRCAP

PTPRCAP is a phosphoprotein that is specifically associated with CD45, a surface marker on lymphocytes, and has been shown in a study to be a key regulator of both T and B cell activation (231). Aberrations to CD45 have been shown to be related to autoimmune conditions (232). In addition CD45 deficiency has been linked to SCID (233); it would therefore appear that CD45 presence is correlated to immune activity. Ergo, in CD4+ T cells in RA one might expect to see PTPRCAP, CD45's associated protein, upregulated. This is what is seen as PTPRCAP is hypomethylated in this cell population from this pathology.

4.6.9 CORO1A

CORO1A was found to be a TNF-coexpressed gene that is hypomethylated. In a recent genome-wide longevity study, CORO1A expression levels were found to be negatively associated both with age at the time of blood sample and the survival time after blood draw. It is also thought to be involved in mitochondrial-mediated cell death (234), like that initiated by TNFR1 (235) and T cell-mediated immunity. It mediates T cell activity through its association with CD45 and its regulation of intracellular Ca^{2+} signalling (236). This is the second CD45-related molecule in this CGN, likely highlighting CD45 as relevant in RA CD4+ T cells.

It should be noted that CORO1B and also KCNN4 and KCNC4 encode proteins that are involved in Ca^{2+} signalling – an important component in regulation of T cell activation (237,238). Though not coexpressed with TNF - CORO1B and KCNN4 are also hypomethylated. As a result the expectations are met that T cells, whom are likely to be more active in a condition such as RA, would have pathways primed that are responsible for the regulation of T cell activation.

4.6.10 ABI3

The ABI protein family are primarily known to be involved in regulation of the actin cytoskeleton through the formation of a WAVE complex. The family contains three members ABI1, ABI2 and ABI3. ABI3 in particular was a later addition to this family and was done so simply on the grounds that it possessed amino acid similarity towards the other ABI members. Whilst all three proteins share a very similar domain structure, that includes a WAVE-binding (WAB) domain; in comparison to ABI1 and ABI2 little is known about the function of ABI3. This is due to the fact that the ABI3-WAVE complex appears to be functionally different to those ABI-WAVE complexes formed with

other ABI family members. The specific differences observed that the ABI3-WAVE2 complex is unable to form lamellipodial protrusions in response to fibronectin stimulation. In addition, expression of Abi3, but not Abi1 or Abi2, reduced the formation of invadopodia in the v-src- cancer cell line, and increased invadopodia formation when cAbl was inhibited. It is through these aspects that it can be seen that ABI3 is distinct from other ABI family members (239).

Whilst invasiveness of cells, from the perspective of cancer cells, is seen as a negative thing. In the immune system, immune cells need the ability to be able to invade regions of the body to help conduct a successful immune response. Invadopodia formation grants a cell the ability to invade very difficult places such as the basement membranes that surround tissues. In the context of autoimmune conditions, invadapodia-formation is associated with disease progression, as seen in RA (240).

As a result one might expect that CD4+ T cells from RA would have ABI3 suppressed – promoting cell invasiveness. However it is seen that ABI3 is actually hypomethylated and primed for transcription. Now obviously ABI3 could have its transcription regulated downstream, nor might it even be pathologically important for CD4+ T cells to possess the ability for aggressive invasion.

It should be noted that ABI3 is the highest ranked gene in regards to most coexpressed hypomethylated genes. As a result it might be that any effect ABI3 could have would be done so through the other genes it is co-expressed with – such as TNF.

It should also be noted that the CD4+T cells potential lack of invasive ability could be down to selection bias. As the CD4+ T cells were taken from peripheral blood, one naturally shouldn't expect to find a large proportion of cells which are very invasive as by definition they would have or be invading tissue and not be in peripheral blood. It should be therefore considered that the CD4+ T cells analysed could be a less-invasive subgroup. Though this can't be said for certain

without further investigation and analysis.

Taken together it can be seen that the hypomethylated genes in TNF's CGN are proinflammatory, prosurvival, involve molecules which interact with the T cell signalling regulator - CD45, and members whose transcription positively correlate with RA pathology. This is consistent with what is seen in the published literature and is expected in an autoimmune condition, where the CD4+ T cells are active and interact with TNF. It should also be noted that 4 out of the 10 genes listed here, TNF, DAXX, LNPEP and GPSM3 are all located in the MHC region. Considering this, 1 of the 4 is the candidate gene being analysed, and that the MHC region has been implicated in RA pathology (247). A closer look at the MHC region is warranted.

4.7 The broader cause of hypomethylation and the MHC region

The data from the top 50 hypo- and hypermethylated genes showed via genemania software analysis that no significant functional ontologies were found in the hypermethylated group but several were in the hypomethylated list. Specifically, it was found that genes involved with MHC class I antigen processing, regulation of JNK and positive regulation of NFkB were significantly represented ($p \leq 0.05$) and therefore the hypomethylated gene network, if transcribed, was significantly likely to be involved in promoting said processes. It should be noted that the functional ontologies highlighted are not mutually exclusive for instance TNF and DAXX are involved in both NFkB and JNK regulation and whilst not in MHC I antigen processing have locis in the MHC region.

In regards to MHC antigen processing, it has been long known to be associated with RA pathology. Furthermore it is the only region of the genome that has been consistently shown to be associated with RA (247). The MHC region is situated on chromosome 6 and extends over 3.6 Mb. It is divided into three regions, classes I, II and III. The class I region, located at the telomeric end of the MHC, contains HLA-A, -B and –C, and extends over 2000 kb. In the HLA class II region are

the HLA-DR, -DQ and -DP loci. The class III region is located between the class I and II regions. The MHC is a highly gene dense region containing around 220 genes, many of which have immunoregulatory functions (248)

The genes said to be connected to MHC class I by the gene network analysis are TAP1, HLA-E and LNPEP. Through literature analysis, some genes on the hypomethylated list have also been related to the MHC class II region, these genes are: PSMB8, TAP 1, PSMB9, Col11A2, HSD17B8, RPS18, B3GALT4, TAPBP, DAXX have all been classed as MHC Class II region genes and all of which are significantly hypomethylated in CD4+ T cells in RA. There are also hypomethylated genes found in the Class III HLA region these are TNF, DDH2, LSM2, STK19, TNXB, PBX2 and GPSM3 (249) This data strongly suggests that, at least at the DNA level, the MHC regions are significantly hypomethylated in CD4+ T cells in RA. To further demonstrate the MHC region's significance, out of the top 5 ranked Hypomethylated genes 3 are located in an MHC region as are 5 in the top 15. Whilst MHC region hypomethylation doesn't account for all the epigenetic aberration seen it is clear that it would explain the large portion of highly hypomethylated, seemingly unrelated genes.

MHC hypomethylation already has been shown to be involved in modulating genetic risk in several autoimmune conditions including RA (250, 251). As a result what has been found in this study is consistent with what has been found in previous published literature. The hypomethylation of MHC region appears particularly relevant in RA CD4+ T cells. This is to do with the fact that the MHC is fundamental to the CD4+ T cell's role of antigen recognition and subsequent cytokine signalling to other immune cell populations. Epigenetic changes to this region could adversely change the regulation of the MHC and therefore the cell's behaviour towards antigens, and as a result could contribute towards RA pathology (248-252). However the exact cause of why in particular the MHC region is affected and why in particular the TNF gene is hypomethylated is not something answerable from this study data. That said through literature review answers might lie in the link between coexpression and hypomethylation.

4.7.1 The link between coexpression and hypomethylation

In this study I have highlighted that TNF is hypomethylated, and has a hypomethylated coexpression gene network, hypomethylated TNF cascade and peripheral pathway members and is located in a hypomethylated DNA (MHC) region. Considering this overlap, in this section I will discuss how coexpression networks, gene loci and gene functions are inter-related in regards to hypomethylation.

Coexpression has been linked to colocalisation. That is, genes which are coexpressed often are in close proximity to each other. Therefore if a region, such as the MHC region, becomes hypomethylated it should be expected that coexpression networks located there become affected by said epigenetic changes. (242).

Also considering that these coexpressed gene networks often share functions and/or are part of a pathway, it can be seen how hypomethylation of a coexpression gene network can result in pathway members, even in peripheral but functionally-related pathways, demonstrating hypomethylation (243).

This link between coexpression and hypomethylation goes further; this is because transcription factors, which coexpressed genes networks share, also mediate DNA methylation maintenance (244,245). Ergo aberrations or conditions which negatively affect the DNA methylation maintenance machinery interacting with particular TFs can in theory predispose coexpressed gene networks, which share said transcription factors, to methylation loss (246).

This matches with what we see in this study as regional hypomethylation being the sole cause would only explain 4 of the 10 hypomethylated coexpressed genes. So it is clear there is more to it than the genes simply being in the same region. It is likely that certain TF or TFs, that this CGN

share, have been affected and any regional correlation of CGN genes is due to colocalisation of certain coexpressed genes in said network. Though of course this explanation for the hypomethylation of TNF’s CGN doesn’t explain the many other hypomethylated genes in the MHC region and the rest of the genome. It is probable that the regulation of many other TFs has been affected also and is the cause of the epigenetic mosaic seen here - however there needs to be further study to investigate this.

I postulate that the MHC region appears particularly central in this data set due to its functional importance in CD4+ T cells and its relevance in RA pathology. It is probable that this region would be transcriptionally active in RA CD4+ T cells. So it can be seen that either the region is more prone to wear and tear, due to increased use, and possesses a lower ability to repair itself due to dysfunctional DNMT machinery or RA CD4+ T cells are ‘purposefully’ modifying the MHC region to better interact and respond to signals in RA pathology. However none of these postulations can be ascertained without thorough further investigation.

The epigenetic mosaic observed in RA CD4+ T cells in this study can be generally characterised by global and MHC-region hypomethylation and priming of genes which facilitate NFkB interaction. Whether said epigenetic mosaic is a cause or consequence of RA is not something that's answerable just using the data found in this study. Though it is likely that ageing, environmental stressors and genetic and/or epigenetic predispositions are factors that leads to progressive genome-wide hypomethylation in CD4+ T cells, and that the hypomethylation of genes in CD4+ T cells, such as TNF, may contribute towards RA pathology.

4.8 Conclusion

In conclusion, it has been confirmed by a different methodology that the TNF promoter region analysed is significantly hypomethylated in CD4+ T cells in RA patient's peripheral blood in comparison to HC. The veracity of the CpG highlighted by the earlier methylation array assay has been confirmed and has highlighted a further 6 more CpGs that are significantly hypomethylated. It has discussed that the many of the highly-ranked hypomethylated genes happen to be in the MHC region on chromosome 6, indicating a degree of regional DNA hypomethylation; it has also been discussed that the genes TNF is coexpressed with genes that promote T cell activation, survival and proliferation; in addition various primed pathways in the TNF cascade have been discussed which seem to favour cell survival, proliferation, and inflammation propagation. All of which is consistent with what is seen in RA pathology. This verification study has been successful and demonstrates that CD4+ T cells are epigenetically distinct in RA in comparison to HC.
I therefore accept my hypothesis and reject my null hypothesis.

5.0 Limitations and areas of further study

5.1 Limitations of methodology

Using a different methodology to observe the same phenomenon not only is useful to confirm the veracity of the first data set through adding additional data to the data pool, but by using a different methodology you show that the methodology specific to producing the first data set wasn't the cause of the phenomenon. However there are various limitations to using the bisulphite DNA Sanger sequencing methodology.

A major limiting factor pertains to Sanger sequencing and it being unable to read DNA products over the length of 1,000bp. However this problem is compounded when the DNA undergoes bisulphite conversion, as converted DNA is unstable above 350bp. As a result not only is the length that can be read limited, but the length of bisulphite-converted DNA that can be amplified is limited. As a result this methodology is unable to verify all the data found all at once – in fact even with one gene this is found to be difficult.

As a result the short amplified products that result from amplifying bisulphite-converted DNA require more primers in order to cover the same sequence in comparison to normal DNA. In addition due to more primers there is more area on the DNA that isn't sequenced very well due to proximity to said primers. Getting the primers to be both of an appropriate length and possess high enough Tms for successful amplification and sequencing is challenging due to the lack of Cs in the DNA. This issue was resolved by finding G-rich areas and or placing an undesired CpG in the first third of a primer – both of which ameliorated both primer length and Tm problems.

Not only is primer design an issue but sequencing bisulphite-converted DNA is difficult due to its inability to read long nucleotide repeats. Specifically with bisulphite-converted DNA poly-T repeats

are common, however only repeats longer than 8 start to affect sequenced product with every additional repeat exponentially increasing the risk of a nucleotide expansion or contraction. Expansions and contractions occur due to the polymerase slipping during it reading the DNA strand. This adversely affects the read sequence as it doesn't always happen consistently. As a result you end up with multiple peaks occurring downstream as the sequence has different amounts of repeated nucleotide that have been added. Multiple peaks at a given point are a problem for a couple of reasons, firstly the quantification of methylation level become interfered with and less accurate if another base is being read at the locus of CpG; also in order to make sure that the correct CpGs are being measured at the right location the read sequence gets aligned with the expected sequence used to design the primer, this becomes a problem if the fewer or additional nucleotides result in the sequence not matching the expected sequence. A further problem with poly-T repeats is that the longer the repeat the more likely the polymerase will terminate reading the DNA strand, this results in variable length products and thus a degrading signal following the poly-T region, this reduces the accuracy of the CpG methylation quantification (153,253).

The Sanger sequencing bisulphite conversion methodology is restricted to observing DNA methylation. However but also the methodology also is unable to differentiate between modified cytosine variants. This is due to the bisulphite conversion process keeping both variants as a cytosine and as a result 5mC and 5hmC simply appear as a methylated cytosine in the sequencing data.

Modified cytosine variants occur because 5mC can go through sequential oxidation by the ten-eleven translocation enzyme (TET). TET family enzymes turns 5mC into 5-hydroxymethylcytosine (5hmC) and then to 5-carboxylcytosine (5caC) and 5-formylcytosine (5fC), the latter two are thymine substrates and are eventually turned back into an unmodified cytosine. Whilst 5hmC is an

intermediate in this process it is considered to be a stable epigenetic mark with it making up 1%-10% of all modified cytosines (254). 5hmC is known to possess distinct epigenetic function as studies have shown with genome-wide sequencing maps of various mammalian cells and tissues that 5hmC is a marker for gene expression with it being enriched in enhancers, promoters and gene bodies, also changes in 5hmC correlate to gene expression levels (255).
As a result knowing which modified CpG variants are present is important to know if one wishes to fully understand the effect of these modifications on a candidate's DNA and gene's expression.

Furthermore, the spontaneous deamination of 5mC to thymine isn't something that this methodology can pick up on, as these appear simply appear as any other demethylated CpG – a thymine. This spontaneous occurrence could be important in understanding the predisposing factors which aid the formation of the observed epigenetic mosaic in RA CD4+ T cells, however bisulphite sequencing can't detect this as they'd simply appear as unmethylated CpGs.

In addition, this methodology doesn't quantify actual gene expression of the gene being sequenced; as a result whether or not specific cytosine modifications are having any contribution to altered gene expression in a specific disease pathology cannot be determined quantitatively with this methodology.

There are limitations to the sequencing software used as it isn't designed for viewing multiple peaks at a given position. This is shown through the presence of the phred quality score (Q). This Q value is defined as a property that is logarithmically related to base-calling error probabilities, that is, the higher the score, the better confidence that can be had in the nucleotide shown at that position. Now considering that multiple peaks at a given point negatively affects the phred quality score it can be seen that software's means of deducing base-calling error and thus quality at a

CpG location won't be very reliable. Furthermore the software package has no DNA methylation- related features which would have assisted in data processing. All the CpGs had to be measured individually which was time consuming and labourious. As a result, if this experiment were repeated, efforts would be made to find software which facilitates sequenced bisulphite-converted DNA data processing.

An option available, if repeated again, would be to use pyrosequencing instead of Sanger sequencing. The method would allow the use of software packages that deduces the proportion Cs and Ts on a CpG at a given position without the need to do it manually. However this method doesn't address limiting factors associated with bisulphite conversion or designing primers for amplification, nor does it allow longer DNA products to be used in sequencing.

Another issue is the RA DNA that was used, generally had low DNA concentration in comparison to HC DNA concentrations. As a result the bisulphite conversion process had to be modified so that the conversion reagent was instead proportional to the concentration of DNA instead of DNA volume. That is, a sample with a low concentration of DNA needed more volume to be viable so more bisulphite conversion reagent needed to be used to maintain the same concentration needed. Despite doing this some of the RA patient's DNA concentrations were omitted for being too low. This low DNA concentration is likely to do with the age of the sample, fortunately despite age of a sample influencing DNA yield it doesn't seem to influence the DNA's integrity. However in future it would be better to use fresher DNA samples in order to get better DNA yields.

It should also be noted that the patient samples were not matched with the ones used in the original Illumina methylation array. This was done out of practicality, as samples were selected based upon their DNA yield. Due to RA patient's blood samples being inherently poor, the quality

of samples selected was important and was prioritised over any benefits drawn from matching samples. Furthermore if the effect is found in RA it should be expected to be seen in other RA patients and if it isn't then it would be down to chance or some idiosyncrasy of the patients in the first cohort. As a result using other RA patients to verify the first data was a stronger choice for verification as not only does it increase the pool of data confirming the phenomenon but it likely rules out any confounding idiosyncratic biases in the first study as the cause of the observed epigenetic aberration.

5.2 Future steps

Now that the promoter region of the TNF gene has been found to be hypomethylated in CD4+ T cells in RA patients the next step would be to repeat the experiment but with a larger cohort of patients. This would allow better understanding of the range of the hypomethylation of the TNF promoter in RA patients in comparison to HC. This is important as it would help to establish if there is a healthy range of methylation levels in the TNF promoter region.

Once the epigenetic mosaic occurring in the TNF gene has been better quantified, it would then be of interest to see at which point in the pathogenesis of RA that TNF gene undergoes significant hypomethylation. It would also be relevant to assess whether the level of TNF promoter methylation is associated with disease severity. This would show whether this is something that is acquired gradually by CD4+ T cells through the disease course or if significant hypomethylations exist in the TNF gene in pre-RA - such as that explained in the outlined cancer stem cell model. These points are relevant as if there is a gradual change to methylation levels that correlate with RA onset and disease severity then one could use this as an RA diagnostic screening tool. This might be advantageous over other diagnostic tools, as it doesn't require patients to have symptoms and it could also include both sero-positive and sero-negative RA patients, which is a major limitation of present diagnostic blood tests for RA.

A different approach would be to see if there are transcription differences in RA CD4+ T cells in comparison to HC. This would confirm whether or not the methylation differences elucidated actually carry over any effects on to the transcription of the respective genes.

This is important to analyse because genes being significantly hypo or hypermethylated doesn't guarantee that they will have their transcription significantly affected. A possible methodology of analysing the RA CD4+ T cell transcriptome would be to use gene expression profiling.

Translation is important to also look at; simply because a gene is being transcribed doesn't mean those mRNA transcript products are being translated into a given protein. This is due to posttranscriptional modifications such as splicing which can result in different protein variant being translated. Furthermore, it should also be considered that translational interference can occur, usually via antisense mRNA hybridising with the transcripted mRNA preventing it being translated - resulting in gene suppression. Translation analysis can be conducted by looking at the proteome of the RA CD4+ T cells. A key methodology used to analyse this would be separation of proteins using gel electrophoresis and subsequent identification of proteins by mass spectrometry.

Another avenue that could be approached would be whether or not TNF is actually being transcribed and translated in CD4+ T cells. This could be achieved by looking at the levels of TNF mRNA in CD4+ T cells in RA and HC patients. If TNF is being successfully translated and more than in HC, and/or at different time points in the pathogenesis of RA, then this could be pathogenically relevant. Considering that TNF encourages inflammation and disease progression. If found to be true, this could highlight CD4+ T cells being a relevant target for therapy. It should also be noted, that even if a gene, such as TNF, is not found to have any transcription or translation differences it wouldn't take away its potential as a diagnostic marker for RA as it's methylation differences are still significant and between RA and HC CD4+ T cells.

Fundamentally, analysing the methylome, transcriptome and proteome would give a comprehensive gene-activity profile and a better overall picture of what is happening in RA CD4+ T cells. Acquiring this data would allow more meaningful conclusions to be made on the role of individual genes or groups of genes in RA CD4+ T cells and the role that CD4+ T cells play in the pathology of RA.

A further aspect to consider is that CD4+ T cells have subpopulations. Specifically naïve CD4+ T cells can differentiate further into: Th1, Th2, Th9, Th17, Th22, Tfh or Tregs. These cell populations secrete different cytokines and are responsible for different elements of the inflammation process. Further investigation into one or more of these sub-populations is needed to see if their epigenetic mosaic differs from each other and in comparison with HC. Considering that the more totipotent a cell, the more sensitive it is to epigenetic change, the CD4+ population that should be analysed should be naïve CD4+ T cells.

6.0 References

1. Harvey W. English Anatomical Exercitations. 1653.
2. Berger S, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes & Development. 2009;23(7):781-783.
3. Jenuwein T, Laible G, Dorn R, Reuter G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cellular and Molecular Life Sciences (CMLS). 1998;54(1):80-93.
4. Mersfelder E, Parthun M. The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. Nucleic Acids Research. 2006;34(9):2653-2662
5. Tropberger P, Schneider R. Scratching the (lateral) surface of chromatin regulation by histone modifications. Nature Structural & Molecular Biology. 2013;20(6):657-661.
6. Lim D, Maher E. DNA methylation: a form of epigenetic control of gene expression. The Obstetrician & Gynaecologist. 2010;12(1):37-42.
7. Coulondre C, Miller J, Farabaugh P, Gilbert W. Molecular basis of base substitution hotspots in Escherichia coli. Nature. 1978;274(5673):775-780.
8. Bird A. DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Research. 1980;8(7):1499-1504.
9. Saxonov S, Berg P, Brutlag D. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proceedings of the National Academy of Sciences. 2006;103(5):1412-1417.
10. Larsen F, Gundersen G, Lopez R, Prydz H. CpG islands as gene markers in the human genome. Genomics. 1992;13(4):1095-1107.
11. Zhu J, He F, Hu S, Yu J. On the nature of human housekeeping genes. Trends in Genetics. 2008;24(10):481-484.
12. Illingworth R, Gruenewald-Schneider U, Webb S, Kerr A, James K, Turner D et al. Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. PLoS Genetics. 2010;6(9):e1001134.
13. Maunakea A, Nagarajan R, Bilenky M, Ballinger T, D'Souza C, Fouse S et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature. 2010;466(7303):253-257.
14. Bostick M, Kim J, Esteve P, Clark A, Pradhan S, Jacobsen S. UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. Science. 2007;317(5845):1760-1764.
15. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo T et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature. 2007;450(7171):908-912.
16. Wu H, Zhang Y. Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. Cell. 2014;156(1-2):45-68.
17. Law J, Jacobsen S. Establishing, maintaining and modifying DNA methylation patterns in plants and animals.

Nature Reviews Genetics. 2010;11(3):204-220.

18. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013;502(7472):472–9

19. Hashimoto, H., Pais, J.E., Zhang, X., Saleh, L., Fu, Z.-Q., Dai, N., Correˆ a, I.R.,Zheng, Y., and Cheng, X. Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. Nature.2013;506(7488):391-5

20. Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P., and Xu, Y. Crystal Structure of TET2-DNA Complex: Insight into TET Mediated 5mC Oxidation. Cell. 2013; 155(7): 1545-55.

21. Wu, H., and Zhang, H. Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. Cell. 2014; 156(0):45-68.

22. Rosenfeld CS. Animal Models of Transgenerational Epigenetic Effects. Transgenerational Epigenetics. 2014;:123–45

23. Gluckman PD, Hanson MA, Buklijas T, Low FM, Beedle AS. Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. Nature Reviews Endocrinology. 2009Feb;5(7):401–8.

24. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proceedings of the National Academy of Sciences. 2008;105(44):17046–9.

25. Tobi EW, Lumey L, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Human Molecular Genetics. 2009Apr;18(21):4046–53.

26. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, et al. Season of Conception in Rural Gambia Affects DNA Methylation at Putative Human Metastable Epialleles. PLoS Genetics. 2010;6(12).

27. Heijmans, B., Tobi, E., Stein, A., Putter, H., Blauw, G., Susser, E., Slagboom, P., and Lumey, L. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A. 2008; 105(44): 17046-49.

28. Sandovici I, Smith NH, Nitert MD, Ackers-Johnson M, Uribe-Lewis S, Ito Y, et al. Maternal diet and aging alter the epigenetic control of a promoter-enhancer interaction at the Hnf4a gene in rat pancreatic islets. Proceedings of the National Academy of Sciences. 2011Aug;108(13):5449–54.

29. Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, et al. Rapid DNA Methylation Changes after Exposure to Traffic Particles. American Journal of Respiratory and Critical Care Medicine. 2009;179(7):572–8.

30. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and Environmental Exposures Alter Tissue-Specific DNA Methylation Dependent upon CpG Island Context. PLoS Genetics.

2009;5(8).

31. Langevin SM, Houseman EA, Christensen BC, Wiencke JK, Nelson HH, Karagas MR, et al. The influence of aging, environmental exposures and local sequence features on the variation of DNA methylation in blood. Epigenetics. 2011;6(7):908–19.

32. Grönniger E, Weber B, Heil O, Peters N, Stäb F, Wenck H, et al. Aging and Chronic Sun Exposure Cause Distinct Epigenetic Changes in Human Skin. PLoS Genetics. 2010;6(5).

33. Calvanese V, Fernández AF, Urdinguio RG, Suárez-Alvarez B, Mangas C, Pérez-García V, et al. A promoter DNA demethylation landscape of human hematopoietic differentiation. Nucleic Acids Research. 2011Oct;40(1):116–31.

34. Khosla S, Dean W, Brown D, Reik W, Feil R. Culture of Preimplantation Mouse Embryos Affects Fetal Development and the Expression of Imprinted Genes1. Biology of Reproduction. 2001Jan;64(3):918–26.

35. Dean w, Bowden /l, Aitchison A, Klose J, Moore T, Meneses JJ, et al. Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. Development. 1998Jun;125(12):2273–82.

36. Doherty AS, Mann MR, Tremblay KD, Bartolomei MS, Schultz RM. Differential Effects of Culture on Imprinted H19 Expression in the Preimplantation Mouse Embryo1. Biology of Reproduction. 2000Jan;62(6):1526–35.

37. Young LE, Fernandes K, McEvoy TG, Butterwith SC, Gutierrez CG, Carolan C, et al. Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nat Gen. 2001Feb;27(2):153–4.

38. Daxinger L, Whitelaw E. Transgenerational epigenetic inheritance: More questions than answers. Genome Research. 2010Jan;20(12):1623–8.

39. Kota SK, Feil R. Epigenetic Transitions in Germ Cell Development and Meiosis. Developmental Cell. 2010;19(5):675–86.

40. Paszkowski J, Grossniklaus U. Selected aspects of transgenerational epigenetic inheritance and resetting in plants. Current Opinion in Plant Biology. 2011;14(2):195–203.

41. Jullien PE, Berger F. DNA methylation reprogramming during plant sexual reproduction? Trends in Genetics. 2010;26(9):394–9.

42. Teixeira FK, Colot V. Repeat elements and the Arabidopsis DNA methylation landscape. Heredity. 2010Dec;105(1):14–23.

43. Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, et al. Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits. PLoS Genetics. 2009;5(6).

44. Richards EJ. Natural epigenetic variation in plant species: a view from the field. Current Opinion in Plant Biology. 2011;14(2):204–9.

45. Sasaki H, Matsui Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nature Reviews Genetics. 2008;2008(2):129–40.

46. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter JR, et al. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis. 2003;35(2):88–93.

47. Seong K, Shimizu H, Li D, Nakamura R, Ishii S. This study describes an epigenetic, environmentally triggered phenotype in fruit flies that can be transgenerationally transmitted. Cell. 2011;145:1049–61.

48. Wolff GL, Kodell RL, Moore SR, Cooney CA. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J. 1999;12:949–57.

49. Waterland RA, Jirtle RL. Transposable Elements: Targets for Early Nutritional Effects on Epigenetic Gene Regulation. Molecular and Cellular Biology. 2003Jan;23(15):5293–300

50. Bordone L, Guarente L. Calorie restriction, SIRT1 and metabolism: understanding longevity. Nature Reviews Molecular Cell Biology. 2005;6(4):298–305.

51. Kenyon C. The Plasticity of Aging: Insights from Long-Lived Mutants. Cell. 2005;120(4):449–60.

52. Kenyon CJ. The genetics of ageing. Nature. 2010;467(7315):622–.

53. Friedman DB, Johnson TE. A mutation in the age-1 gene in Caenorhabditis elegans lengthens life and reduces hermaphrodite fertility. Genetics. 1988;118:75–86.

54. Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R. A C. elegans mutant that lives twice as long as wild type. Nature. 1993;366(6454):461–4.

55. Lin K. daf-16: An HNF-3/forkhead Family Member That Can Function to Double the Life-Span of Caenorhabditis elegans. Science. 1997;278(5341):1319–22.

56. Ogg S, Paradis S, Gottlieb S, Patterson GI, Lee L, Tissenbaum HA, et al. The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. Nature. 1997;389(6654):994–9.

57. Kimura KD. daf-2, an Insulin Receptor-Like Gene That Regulates Longevity and Diapause in Caenorhabditis elegans. Science. 1997;277(5328):942–6.

58. Hu Z, Chen K, Xia z, Chavez M, Pal S, Seol JH, et al. Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. Genes Dev . 2014;28:396–408.

59. Feser J, Truong D, Das C, Carson JJ, Kieft J, Harkness T, et al. Elevated histone expression promotes life span extension. Mol Cell . 2010; 39:724–35.

60. Greer EL, Maures TJ, Hauswirth AG, Green EM, Leeman DS, Maro GS, et al. Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in C. elegans. Nature. 2010;466(7304):383–7.

61. Maures TJ, Greer EL, Hauswirth AG, Brunet A. The H3K27 demethylase UTX-1 regulates C. elegans lifespan in a germline-independent, insulin-dependent manner. Aging Cell. 2011;10(6):980–90

62. Coppé J-P, Patil CK, Rodier F, Sun Y, Muñoz DP, Goldstein J, et al. Senescence-Associated Secretory Phenotypes Reveal Cell-Nonautonomous Functions of Oncogenic RAS and the p53 Tumor Suppressor. PLoS Biology. 2008Feb;6(12).

63. Adams PD. Healing and Hurting: Molecular Mechanisms, Functions, and Pathologies of Cellular Senescence. Molecular Cell. 2009;36(1):2–14.

64. Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, et al. Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. Genes & Development. 2013Sep;27(16):1787–99.

65. Capell BC, Drake AM, Zhu J, Shah PP, Dou Z, Dorsey J, et al. MLL1 is essential for the senescence-associated secretory phenotype. Genes & Development. 2016Jan;30(3):321–36.

66. Chambers SM, Shaw CA, Gatza C, Fisk CJ, Donehower LA, Goodell MA. Aging Hematopoietic Stem Cells Decline in Function and Exhibit Epigenetic Dysregulation. PLoS Biology. 2007;5(8).

67. Sun D, Luo M, Jeong M, Rodriguez B, Xia Z, Hannah R, et al. Epigenomic Profiling of Young and Aged HSCs Reveals Concerted Changes during Aging that Reinforce Self-Renewal. Cell Stem Cell. 2014;14(5):673–88.

68. Cruickshanks HA, Mcbryan T, Nelson DM, Vanderkraats ND, Shah PP, Tuyn JV, et al. Senescent cells harbour features of the cancer epigenome. Nature Cell Biology. 2013;15(12):1495–506.

69. Singhal RP, Mays-Hoopes LL, Eichhorn GL. DNA methylation in aging of mice. Mechanisms of Ageing and Development. 1987;41(3):199–210.

70. Armstrong VL, Rakoczy S, Rojanathammanee L, Brown-Borg HM. Expression of DNA Methyltransferases Is Influenced by Growth Hormone in the Long-Living Ames Dwarf Mouse In Vivo and In Vitro. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2013Jul;69(8):923–33.

71. Horvath S. DNA methylation age of human tissues and cell types. Genome Biology. 2013;14(10).

72. Horvath S. Erratum to: DNA methylation age of human tissues and cell types. Genome Biology. 2015;16(1).

73. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. Molecular Cell. 2013;49(2):359–67.

74. Jones PA, Baylin SB. The Epigenomics of Cancer. Cell. 2007;128(4):683–92.

75. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. Nature Genetics. 2006;38(5):540–9.

76. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Cell. 2007Feb23;128(4):683–92.

77. Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Gen. 1999;21(2):163–

78. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature. 1983;301(5895):89–92.

79. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. Nature Reviews Genetics. 2006;7(1):21–33.

80. Rodriguez J, Frigola J, Vendrell E, Risques R-A, Fraga MF, Morales C, et al. Chromosomal Instability Correlates with Genome-wide DNA Demethylation in Human Primary Colorectal Cancers. Cancer Research. 2006Jan;66(17):8462–9468.

81. Howard G, Eiges R, Gaudet F, Jaenisch R, Eden A. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. Oncogene. 2007Sep;27(3):404–8.

82. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2007;1775(1):138–62.

83. Baylin SB. DNA Methylation and Gene Silencing in Cancer. ChemInform. 2006;37(22).

84. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, et al. Epigenetic stem cell signature in cancer. Nat Gen. 2007;29:157–8.

85. Jeffries, M.A, Sawalha, A.H. Autoimmune disease in the epigenetic era: how has epigenetics changed our understanding of disease and how can we expect the field to evolve?. Expert Rev Clin Immunol. 2015;11(1): 45–58.

86. Quddus J, Johnson KJ, Gavalchin J, Amento EP, Chrisp CE, Yung RL, et al. Treating activated CD4 T cells with

either of two distinct DNA methyltransferase inhibitors, 5-azacytidine or procainamide, is sufficient to cause a lupus-like disease in syngeneic mice. Journal of Clinical Investigation. 1993Jan;92(1):38–53.

87. Richardson B. Effect of an inhibitor of DNA methylation on T cells. II. 5-azacytidine induces self-reactivity in antigen-specific T4 cells. Human Immunology. 1986;17(4):456–70.

88. Gorelik G, Fang JY, Wu A, Sawalha AH, Richardson B. Impaired T Cell Protein Kinase Cδ Activation Decreases ERK Pathway Signaling in Idiopathic and Hydralazine-Induced Lupus. The Journal of Immunology. 2007Feb;179(8):5553–63.

89. Wen ZK, Xu W, Xu L, Cao QH, Wang Y, Chu YW, et al. DNA hypomethylation is crucial for apoptotic DNA to induce systemic lupus erythematosus-like autoimmune disease in SLE-non-susceptible mice. Rheumatology. 2007May;46(12):1796–803.

90. Yasuda K, Richez C, Uccellini MB, Richards RJ, Bonegio RG, Akira S, et al. Requirement for DNA CpG Content in TLR9-Dependent Dendritic Cell Activation Induced by DNA-Containing Immune Complexes. The Journal of Immunology. 2009;183(5):3109–17.

91. Wang G, Zhang M, Li X, Zhang H, Chen W, Kan M, et al. Ultraviolet B exposure of peripheral blood mononuclear cells of patients with systemic lupus erythematosus inhibits DNA methylation. Lupus. 2009;18(12):1037–44.

92. Sawalha AH, Jeffries M, Webb R, Lu Q, Gorelik G, Ray D, et al. Defective T-cell ERK signaling induces interferon-regulated gene expression and overexpression of methylation-sensitive genes similar to lupus patients. Genes & Immunity. 2008;9(4):368–78.

93. Perl A. Oxidative stress in the pathology and treatment of systemic lupus erythematosus. Nature reviews Rheumatology. 2013;9(11):674–686

94. Lu Q, Kaplan M, Ray D, Ray D, Zacharek S, Gutsch D, et al. Demethylation ofITGAL (CD11a) regulatory sequences in systemic lupus erythematosus. Arthritis & Rheumatism. 2002;46(5):1282–91

95. Gateva, V, Sandling, J.K, Hom , G. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus . Nature genetics. 2009;11(1): 1228–1233.

96. Scofield RH, Bruner GR, Namjou B, et al. Klinefelter's syndrome (47, XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome. Arthritis and rheumatism. 2008;58(8):2511–2517

97. Ciferska H, Horak P, Hermanova Z, et al. The levels of sCD30 and of sCD40L in a group of patients with systemic lupus erythematodes and their diagnostic value. Clinical rheumatology.

98. Goules A, Tzioufas AG, Manousakis MN, Kirou KA, Crow MK, Routsias JG. Elevated levels of soluble CD40 ligand (sCD40L) in serum of patients with systemic autoimmune diseases. Journal of autoimmunity. 2006;26(3):165–171.

99. Zhou Y, Yuan J, Pan Y, et al. T cell CD40LG gene expression and the production of IgG by autologous B cells in systemic lupus erythematosus. Clinical immunology. 2009;132(3):362–370.

100. Lu Q, Wu A, Tesmer L, Ray D, Yousif N, Richardson B. Demethylation of CD40LG on the inactive X in T cells from women with lupus. Journal of immunology. 2007;179(9):6352–6358.

101. Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JI, Rodriguez-Ubreva J, et al. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Research. 2009;20(2):170–9.

102. Perez P, Goicovich E, Alliende C, et al. Differential expression of matrix metalloproteinases in labial salivary glands of patients with primary Sjogren's syndrome. Arthritis and rheumatism. 2000;43(12):2807–2817.

103. Kurasawa K, Arai S, Owada T, Maezawa R, Kumano K, Fukuda T. Autoantibodies against platelet-derived growth factor receptor alpha in patients with systemic lupus erythematosus. Modern rheumatology / the Japan Rheumatism Association. 2010;20(5):458–465

104. Jeffries MA, Dozmorov M, Tang Y, Merrill JT, Wren JD, Sawalha AH. Genome-wide DNA methylation patterns in CD4+ T cells from patients with systemic lupus erythematosus. Epigenetics : official journal of the DNA Methylation Society. 2011;6(5):593–601.

105. Jeffries M, Dozmorov M, Tang Y, Merrill JT, Wren JD, Sawalha AH. Genome-wide DNA methylation patterns in CD4 T cells from patients with systemic lupus erythematosus. Epigenetics. 2011;6(5):593–601.

106. Coit P, Jeffries M, Altorok N, Dozmorov MG, Koelsch KA, Wren JD, et al. Genome-wide DNA methylation study suggests epigenetic accessibility and transcriptional poising of interferon-regulated genes in naïve CD4 T cells from lupus patients. Journal of Autoimmunity. 2013;43:78–84.

107. Yin H, Zhao M, Wu X, et al. Hypomethylation and overexpression of CD70 (TNFSF7) in CD4+ T cells of patients with primary Sjogren's syndrome. Journal of dermatological science. 2010;59(3):198–203.

108. Yu X, Liang G, Yin H, Ngalamika O, Li F, Zhao M, et al. DNA hypermethylation leads to lower FOXP3 expression in CD4 T cells of patients with primary Sjögrens syndrome. Clinical Immunology. 2013;148(2):254–7.

109. Shen L, Suresh L, Wu J, Xuan J, Li H, Zhang C, et al. A Role for Lymphotoxin in Primary Sjogrens Disease. The Journal of Immunology. 2010;185(10):6355–63.

110. Thabet Y, Le Dantec C, Ghedira I, et al. Epigenetic dysregulation in salivary glands from patients with primary Sjogren's syndrome may be ascribed to infiltrating B cells. Journal of autoimmunity. 2013;41:175–181.

111. Hügle T, Oreilly S, Simpson R, Kraaij MD, Bigley V, Collin M, et al. Tumor necrosis factor-costimulated T lymphocytes from patients with systemic sclerosis trigger collagen production in fibroblasts. Arthritis &

Rheumatism. 2013;65(2):481–91.

112. Lei W, Luo Y, Lei W, Luo Y, Yan K, Zhao S, et al. Abnormal DNA methylation in CD4 T cells from patients with systemic lupus erythematosus, systemic sclerosis, and dermatomyositis. Scandinavian Journal of Rheumatology. 2009;38(5):369–74.

113. Lian X, Xiao R, Hu X, Kanekura T, Jiang H, Li Y, et al. DNA demethylation of CD40L in CD4 T cells from women with systemic sclerosis: A possible explanation for female susceptibility. Arthritis & Rheumatism. 2012;64(7):2338–45.

114. Kawai M, Masuda A, Kuwana M. A CD40-CD154 interaction in tissue fibrosis. Arthritis & Rheumatism. 2008;58(11):3562–73.

115. Fukasawa C, Kawaguchi Y, Harigai M, Sugiura T, Takagi K, Kawamoto M, et al. Increased CD40 expression in skin fibroblasts from patients with systemic sclerosis (SSc): role of CD40-CD154 in the phenotype of SSc fibroblasts. European Journal of Immunology. 2003;33(10):2792–800.

116. Allanore Y. Increased plasma soluble CD40 ligand concentrations in systemic sclerosis and association with pulmonary arterial hypertension and digital ulcers. Annals of the Rheumatic Diseases. 2004;64(3):481–3.

117. Komura K, Fujimoto M, Hasegawa M, Sato S. Blockade of CD40–CD40 Ligand Interactions Attenuates Skin Fibrosis and Autoimmunity in the Tight-skin Mouse. Clinical Immunology. 2007;123.

118. Jiang H, Xiao R, Lian X, Kanekura T, Luo Y, Yin Y, et al. Demethylation of TNFSF7 contributes to CD70 overexpression in CD4 T cells from patients with systemic sclerosis. Clinical Immunology. 2012;143(1):39–44.

119. Wang Y, Fan P-S, Kahaleh B. Association between enhanced type I collagen expression and epigenetic repression of theFLI1 gene in scleroderma fibroblasts. Arthritis & Rheumatism. 2006;54(7):2271–9.

120. Altorok N, Tsou P-S, Coit P, Khanna D, Sawalha AH. Genome-wide DNA methylation analysis in dermal fibroblasts from patients with diffuse and limited systemic sclerosis reveals common and subset-specific DNA methylation aberrancies. Annals of the Rheumatic Diseases. 2014Aug;74(8):1612–20.

121. Jeffries MA, Donica M, Baker LW, Stevenson ME, Annan AC, Humphrey MB, et al. Genome-Wide DNA Methylation Study Identifies Significant Epigenomic Changes in Osteoarthritic Cartilage. Arthritis & Rheumatology. 2014;66(10):2804–15.

122. Richardson B, Scheinbart L, Strahler J, Gross L, Hanash S, Johnson M. Evidence for impaired t cell dna methylation in systemic lupus erythematosus and rheumatoid arthritis. Arthritis & Rheumatism. 1990;33(11):1665–73.

123. Corvetta A, Bitta RD, Luchetti MM, Pomponio G. 5-Methylcytosine content of DNA in blood, synovial

mononuclear cells and synovial tissue from patients affected by autoimmune rheumatic diseases. Journal of Chromatography B: Biomedical Sciences and Applications. 1991;566(2):481–91.

124. Liu C-C, Fang T-J, Ou T-T, Wu C-C, Li R-N, Lin Y-C, et al. Global DNA methylation, DNMT1, and MBD2 in patients with rheumatoid arthritis. Immunology Letters. 2011;135(1-2):96–9.

125. Liao J, Liang G, Xie S, Zhao H, Zuo X, Li F, et al. CD40L demethylation in CD4 T cells from women with rheumatoid arthritis. Clinical Immunology. 2012;145(1):13–8.

126. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nature Biotechnology. 2013;31(2):142–7.

127. Neidhart M, Rethage J, Kuchen S, Künzler P, Crowl RM, Billingham ME, et al. Retrotransposable L1 elements expressed in rheumatoid arthritis synovial tissue: Association with genomic DNA hypomethylation and influence on gene expression. Arthritis & Rheumatism. 2000;43(12):2634–47.

128. Kuchen S, Seemayer CA, Rethage J, Knoch RV, Kuenzler P, Michel BA, et al. The L1 Retroelement-related p40 Protein Induces p38δ MAP Kinase. Autoimmunity. 2004;37(1):57–65.

129. Karouzakis E, Gay RE, Michel BA, Gay S, Neidhart M. DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis & Rheumatism. 2009;60(12):3613–22.

130. Karouzakis E, Gay RE, Gay S, Neidhart M. Increased recycling of polyamines is associated with global DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis & Rheumatism. 2012;64(6):1809–17.

131. Nakano K, Boyle DL, Firestein GS. Regulation of DNA methylation in rheumatoid arthritis synoviocytes. Journal of immunology. 2013;190(3):1297–1303

132. Steer SE, Williams FMK, Kato B, et alReduced telomere length in rheumatoid arthritis is independent of disease activity and duration Annals of the Rheumatic Diseases 2007;66:476-480.

133. Locksley RM, Killeen N, Lenardo MJ. The TNF and TNF Receptor Superfamilies. Cell. 2001;104(4):487–501.

134. Besse, A,.Wu, P., Brun,i F.,Donti, T., Graham, B.,Craigen, W.,McFarland, R., Moretti, P.,Lalani, S.,.Scott,K.,Taylor, R.,Bonnen, P. The GABA Transaminase, ABAT, Is Essential for Mitochondrial Nucleoside Metabolism. Cell. 2015; 21(3): 417-27.

135. Hou. P.,Liu, G., Zhao,Y.,Shi, Z., Zheng, Q., Bu, G., Xu, H.,Zhang,Y.The role of copper and the copper-related protein CUTA in mediating APP processing and Aβ generation. Neurobiolageing. 2015; 36(3):1310-15

136. Szczepny, A., Carey, K., McKenzie, L. Jayasekara, W., Rossello, F., Gonzalez-Rajal, A., McCaw, A., Popovski, D., Wang, D., Sadler, A.,_Mahar, A., Russell, P., Wright, G., McCloy, R., Garama, D., Gough, D., Baylin, S., Burgess, A., Cain J., and Watkins, N.The tumor suppressor Hic1 maintains chromosomal stability independent

of Tp53. Oncogene. 2018; 37: 1939-48.

137. Kim, H., Han, J., Chang, I., Kim, W., Myung, S. Variants in the HEPSIN gene are associated with susceptibility to prostate cancer. Prostate Cancer Prostatic Dis. 2012; 15:353-5

138. Schengrund, C. Gangliosides: glycosphingolipids essential for normal neural development and function. Cell. Trends in Biochemical Sciences. 2015; 40(7):397-406

139. Wang W., Grimmer J.F., Van De Water T.R., Lufkin T. Hmx2 and Hmx3 homeobox genes direct development of the murine inner ear and hypothalamus and can be functionally replaced by Drosophila Hmx. Dev Cell. 2004;7(3):439-53.

140. Grell M, Wajant H, Zimmermann G, Scheurich P. The type 1 receptor (CD120a) is the high-affinity receptor for soluble tumor necrosis factor. Proceedings of the National Academy of Sciences. 1998;95(2):570–5.

141. Grell M, Douni E, Wajant H, Löhden M, Clauss M, Maxeiner B, et al. The transmembrane form of tumor necrosis factor is the prime activating ligand of the 80 kDa tumor necrosis factor receptor. Cell. 1995;83(5):793–802

142. Wallach D, Engelmann H, Nophar Y, Aderka D, Kemper O, Hornik V, et al. Soluble and cell surface receptors for tumor necrosis factor. Agents Actions Suppl. 1991;35:51–7.

143. Russo C, Polosa R. TNF-α as a promising therapeutic target in chronic asthma: a lesson from rheumatoid arthritis. Clinical Science. 2005Jan;109(2):135–42.

144. Bystrom J, Clanchy F, Taher TE, Mangat P, Jawad AS, Williams RO, et al. TNFα in the regulation of Treg and Th17 cells in rheumatoid arthritis and other autoimmune inflammatory diseases. Cytokine. 2018;101:4–13.

145. Yamada Y, Kirillova I, Peschon JJ, Fausto N. Initiation of liver growth by tumor necrosis factor: Deficient liver regeneration in mice lacking type I tumor necrosis factor receptor. Proceedings of the National Academy of Sciences. 1997;94(4):1441–6.

146. Bradham CA, Plümpe J, Manns MP, Brenner DA, Trautwein C. Mechanisms of hepatic toxicity. I. TNF-induced liver injury. Am J Physiol . 275(3):G387–G392.

147. Darnay BG, Ni J, Moore PA, Aggarwal BB. Activation of NF-κB by RANK Requires Tumor Necrosis Factor Receptor-associated Factor (TRAF) 6 and NF-κB-inducing Kinase. Journal of Biological Chemistry. 1999;274(12):7724–31.

148. Paz PE, Wang S, Clarke H, Lu X, Stokoe D, Abo A. Mapping the Zap-70 phosphorylation sites on LAT (linker for activation of T cells) required for recruitment and activation of signalling proteins in T cells. Biochemical Journal. 2001Jan;356(2):461.

149. Tu Y-C, Huang D-Y, Shiah S-G, Wang J-S, Lin W-W. Regulation of c-Fos Gene Expression by NF-κB: A p65 Homodimer Binding Site in Mouse Embryonic Fibroblasts but Not Human HEK293 Cells. PLoS ONE. 2013;8(12).

150. Shiomi, A., Usui, T. & Mimori, T. GM-CSF as a therapeutic target in autoimmune diseases. Inflamm Regener. 2016; 36(8).

151. Zhang, Y., He, M., Wang, Y., and Liao, A. Modulators of the Balance between M1 and M2 Macrophages during Pregnancy. Front Immunol. 2017; 8: 120

152. Taylor, P.C., Saurigny, D., Vencovsky, J. Takeuchi, T., Nakamura, T., Matsievskaia, G., Hunt, B., Wagner, T. and Souberbielle, B. . Efficacy and safety of namilumab, a human monoclonal antibody against granulocyte-macrophage colony-stimulating factor (GM-CSF) ligand in patients with rheumatoid arthritis (RA) with either an inadequate response to background methotrexate therapy or an inadequate response or intolerance to an anti-TNF (tumour necrosis factor) biologic therapy: a randomized, controlled trial. Arthritis Res Ther. 2019; 21: 101.

153. Clarke LA, Rebelo CS, Gonçalves J, Boavida MG, Jordan P. PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences. Mol Pathol. 2001;54(5):351–353.

154. Deng M, Loughran PA, Zhang L, Scott MJ, Billiar TR. Shedding of the tumor necrosis factor (TNF) receptor from the surface of hepatocytes during sepsis limits inflammation through cGMP signaling. Sci Signal. 2015;8(361):ra11.

155. Chu WM. Tumor necrosis factor. Cancer Lett. 2013;328(2):222–225.

156. Dillon. P., Oberst, A.,Weinlich, R., Mak, T., Wallach, D., and Green, D. Survival Function of the FADD-CASPASE-8-cFLIPL Complex. Cell. 2012;1(5):401-7.

157. Zhang, X., Zhang, H., Xu, C. , Xu, C., Li, X., Li, M., Wu, X., Pu, W., Zhou, B., Wang, H., Li, D., Ding, Q., Ying, H., Wang, H. & Zhang, H.,et al. Ubiquitination of RIPK1 suppresses programmed cell death by regulating RIPK1 kinase activation during embryogenesis. Nat Commun.2019; 10: 4158

158. Yamada, Y., Kirillova, I., Peschon, J. and Fausto N. Initiation of liver growth by tumor necrosis factor: Deficient liver regeneration in mice lacking type I tumor necrosis factor receptor. Proceedings of the National Academy of Sciences. 1997; 94 (4): 1441-1446

159. Tarrats, N., Moles, A., Morales, A., García-Ruiz, C., Fernández-Checa, J.C., and Marí, M.Critical role of tumor necrosis factor receptor 1, but not 2, in hepatic stellate cell proliferation, extracellular matrix remodeling, and liver fibrogenesis..Hepatology. 2011: 54(1):319-27.

160. Taylor PC, Peters AM, Paleolog E, Chapman PT, Elliott MJ, Mccloskey R, et al. Reduction of chemokine levels and leukocyte traffic to joints by tumor necrosis factor α blockade in patients with rheumatoid arthritis. Arthritis & Rheumatism. 2000;43(1):38–47.

161. Blam M. Integrating anti–tumor necrosis factor therapy in inflammatory bowel disease: current and future

perspectives. The American Journal of Gastroenterology. 2001;96(7):1977–97.

162. Risco A, Cuenda A. New Insights into the p38γ and p38δ MAPK Pathways. J Signal Transduct. 2012; 2012:520289.

163. McEver RP. Selectins: initiators of leucocyte adhesion and signalling at the vascular wall. Cardiovasc Res. 2015;107(3): 331–339.

164. Nick JA, Avdi NJ, Young SK, Lehman LA, Mcdonald PP, Frasch SC, et al. Selective activation and functional significance of p38α mitogen-activated protein kinase in lipopolysaccharide-stimulated neutrophils. Journal of Clinical Investigation. 1999;103(6):851–8

165. Read MA, Whitley MZ, Gupta S, Pierce JW, Best J, Davis RJ, et al. Tumor Necrosis Factor α-Induced E-selectin Expression Is Activated by the Nuclear Factor-κB and c-JUN N-terminal Kinase/p38 Mitogen-activated Protein Kinase Pathways. Journal of Biological Chemistry. 1997;272(5):2753–61.Herlaar, E, Brown, Z. P38 MAPK signalling cascades in inflammatory disease. Molecular Medicine Today. 1999;Volume 5(10): 439-447

166. Bergmann M, Hart L, Lindsay M, Barnes PJ, Newton R. IκBα Degradation and Nuclear Factor-κB DNA Binding Are Insufficient for Interleukin-1β and Tumor Necrosis Factor-α-induced κB-dependent Transcription. Journal of Biological Chemistry. 1998;273(12):6607–10 .

167. Wesselborg S, Bauer MKA, Vogt M, Schmitz ML, Schulze-Osthoff K. Activation of Transcription Factor NF-κB and p38 Mitogen-activated Protein Kinase Is Mediated by Distinct and Separate Stress Effector Pathways. Journal of Biological Chemistry. 1997Sep;272(19):12422–9.

168. Cui J, Zhang M, Zhang Y-Q, Xu Z-H. JNK pathway: diseases and therapeutic potential. Acta Pharmacologica Sinica. 2007;28(5):601–8.

169. Johnson GL, Nakamura K. The c-jun kinase/stress-activated pathway: Regulation, function and role in human disease. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2007;1773(8):1341–8.

170. Weston CR, Davis RJ. The JNK signal transduction pathway. Current Opinion in Cell Biology. 2007;19(2):142–9.Archelos, J, Hans-peter, H. The role of adhesion molecules in multiple sclerosis: biology, pathogenesis and therapeutic implications.Molecular Medicine Today. 1997;Volume 3(7): 310-321.

171. Barr RK, Bogoyevitch MA. The c-Jun N-terminal protein kinase family of mitogen-activated protein kinases (JNK MAPKs). The International Journal of Biochemistry & Cell Biology. 2001;33(11):1047–63

172. Bogoyevitch MA. The isoform-specific functions of the c-Jun N-terminal Kinases (JNKs): differences revealed by gene targeting. BioEssays. 2006;28(9):923–34Holgate ST. Cytokine and anti-cytokine therapy for the treatment of asthma and allergic disease. Cytokine. 2004;28(4-5):152–7.

173. Holgate ST. Cytokine and anti-cytokine therapy for the treatment of asthma and allergic disease. Cytokine.

2004;28(4-5):152–7.

174. Chung K. Cytokines as Targets in Chronic Obstructive Pulmonary Disease. Current Drug Targets. 2006Jan;7(6):675–81.

175. Williams RO, Paleolog E, Feldmann M. Cytokine inhibitors in rheumatoid arthritis and other autoimmune diseases. Current Opinion in Pharmacology. 2007;7(4):412–7.

176. Lawrence T. The nuclear factor NF-kappaB pathway in inflammation. Cold Spring Harb Perspect Biol. 2009;1(6):a001651.

177. Tak PP, Firestein GS. NF-κB: a key role in inflammatory diseases. Journal of Clinical Investigation. 2001Jan;107(1):7–11.

178. Karin M, Yamamoto Y, Wang QM. The IKK NF-κB system: a treasure trove for drug development. Nature Reviews Drug Discovery. 2004;3(1):17–26.

179. Aupperle KR, Bennett BL, Boyle DL, Tak PP, Manning AM, Firestein GS. NF-kappa B regulation by I kappa B kinase in primary fibroblast-like synoviocytes. J Immunol. 1990;163(1):427–33.

180. Aupperle K, Bennett B, Han Z, Boyle D, Manning A, Firestein G. NF-kappa B regulation by I kappa B kinase-2 in rheumatoid arthritis synoviocytes. Firestein. 2001Feb15;166(4):2705–11.

181. Miagkov AV, Kovalenko DV, Brown CE, Didsbury JR, Cogswell JP, Stimpson SA, et al. NF- B activation provides the potential link between inflammation and hyperplasia in the arthritic joint. Proceedings of the National Academy of Sciences. 1998Oct;95(23):13859–64

182. Beg AA, Baltimore D. An Essential Role for NF-kappa B in Preventing TNF-alpha -Induced Cell Death. Science. 1996Jan;274(5288):782–4.

183. Li Q, Van Antwerp D, Mercurio F, Lee KF, Verma IM. Severe liver degeneration in mice lacking the IkappaB kinase 2 gene. Science. 1999Apr;284(5412):321–5.

184. Lawrence T, Gilroy DW, Colville-Nash PR, Willoughby DA. Possible new role for NF-kappaB in the resolution of inflammation. Nat Med. 2001Dec;7(12):1291–7

185. Bohuslav J, Kravchenko VV, Parry GC, Erlich JH, Gerondakis S, Mackman N, et al. Regulation of an essential innate immune response by the p50 subunit of NF-kappaB. Journal of Clinical Investigation. 1998Jan;102(9):1645–52.

186. Erdman SE, Fox JG, Dangler CA, Feldman D, Horwitz BH. Cutting Edge: Typhlocolitis in NF- B-Deficient Mice. The Journal of Immunology. 2001Jan;166(3):1443–7.

187. Tomczak MF, Erdman SE, Poutahidis T, Rogers AB, Holcombe H, Plank B, et al. NF- B Is Required Within the Innate Immune System to Inhibit Microflora-Induced Colitis and Expression of IL-12 p40. The Journal of Immunology. 2003;171(3):1484–92.

188. Tomczak MF, Erdman SE, Davidson A, Wang YY, Nambiar PR, Rogers AB, et al. Inhibition of Helicobacter hepaticus-Induced Colitis by IL-10 Requires the p50/p105 Subunit of NF- B. The Journal of Immunology. 2006Feb;177(10):7332–9.

189. Lin B, Williams-Skipp C, Tao Y, Schleicher MS, Cano LL, Duke RC, et al. NF-κB functions as both a proapoptotic and antiapoptotic regulatory factor within a single cell type. Cell Death & Differentiation. 1999;6(6):570–82.

190. Li Q, Lu Q, Bottero V, Estepa G, Morrison L, Mercurio F, et al. Enhanced NF- B activation and cellular function in macrophages lacking I B kinase 1 (IKK1). Proceedings of the National Academy of Sciences. 2005;102(35):12425–30.

191. Jacque E, Tchenio T, Piton G, Romeo P-H, Baud V. RelA repression of RelB activity induces selective gene activation downstream of TNF receptors. Proceedings of the National Academy of Sciences. 2005;102(41):14635–40.

192. Khelifi AF, Dalcontres MS, Salomoni P. Daxx is required for stress-induced cell death and JNK activation. Cell Death and Differentiation. 2005;12(7):724–33

193. Tak H, Jung Y, Joe C. Daxx promotes TNFα-induced apoptosis by inhibiting NF-kB activity. Cellular and Molecular Biology. 2005;46(10):192.

194. Park J, Lee JH, Muhnho L , et al. Inhibition of NF-Î°B Acetylation and its Transcriptional Activity by Daxx. Journal of molecular biology. 2007;368: 388-97.

195. Duckett CS, Thompson CB. CD30-dependent degradation of TRAF2: implications for negative regulation of TRAF signaling and the control of cell survival. Genes & Development. 1997Jan;11(21):2810–21.

196. Mauri DN, Ebner R, Montgomery RI, Kochel KD, Cheung TC, Yu G-L, et al. LIGHT, a New Member of the TNF Superfamily, and Lymphotoxin α Are Ligands for Herpesvirus Entry Mediator. Immunity. 1998;8(1):21–30.

197. Zhai Y, Guo R, Hsu TL, Yu GL, Ni J, Kwon BS, et al. LIGHT, a novel ligand for lymphotoxin beta receptor and TR2/HVEM induces apoptosis and suppresses in vivo tumor formation via gene transfer. Journal of Clinical Investigation. 1998;102(6):1142–51.

198. Morel Y, Truneh A, Sweet RW, Olive D, Costello RT. The TNF Superfamily Members LIGHT and CD154 (CD40 Ligand) Costimulate Induction of Dendritic

199. Tamada K, Shimozaki K, Chapoval AI, Zhai Y, Su J, Chen S-F, et al. LIGHT, a TNF-Like Molecule,

Costimulates T Cell Proliferation and Is Required for Dendritic Cell-Mediated Allogeneic T Cell Response. The Journal of Immunology. 2000;164(8):4105–10.

200. Sedy JR, Gavrieli M, Potter KG, Hurchla MA, Lindsley RC, Hildner K, et al. B and T lymphocyte attenuator regulates T cell activation through interaction with herpesvirus entry mediator. Nature Immunology. 2004;6(1):90–8.

201. Duhen T, Pasero C, Mallet FO, Barbarat B, Olive D, Costello RG. LIGHT costimulates CD40 triggering and induces immunoglobulin secretion; a novel key partner in T cell-dependent B cell terminal differentiation. European Journal of Immunology. 2004;34(12):3534–41.

202. Watts TH. Tnf/tnfr Family Members In Costimulation Of T Cell Responses. Annual Review of Immunology. 2005;23(1):23–68.

203. Wang J, Lo JC, Foster A, Yu P, Chen HM, Wang Y, et al. The regulation of T cell homeostasis and autoimmunity by T cell–derived LIGHT. Journal of Clinical Investigation. 2001;108(12):1771–80.

204. Wang J, Foster A, Chin R, Yu P, Sun Y, Wang Y, et al. The complementation of lymphotoxin deficiency with LIGHT, a newly discovered TNF family member, for the restoration of secondary lymphoid structure and function. European Journal of Immunology. 2002;32(7):1969.

205. Wang Y-G, Kim KD, Wang J, Yu P, Fu Y-X. Stimulating Lymphotoxin Receptor on the Dendritic Cells Is Critical for Their Homeostasis and Expansion. The Journal of Immunology. 2005Apr;175(10):6997–7002.

206. Wang J, Chun T, Lo JC, Wu Q, Wang Y, Foster A, et al. The critical role of LIGHT, a TNF family member, in T cell developmen. J Immunol. 2001;167(9):5099–105.

207. Wang J, Fu Y-X. LIGHT (a Cellular Ligand for Herpes Virus Entry Mediator and Lymphotoxin Receptor)-Mediated Thymocyte Deletion Is Dependent on the Interaction Between TCR and MHC/Self-Peptide. The Journal of Immunology. 2003;170(8):3986–93.

208. Scheu S, Alferink J, Pötzel T, Barchet W, Kalinke U, Pfeffer K. Targeted Disruption of LIGHT Causes Defects in Costimulatory T Cell Activation and Reveals Cooperation with Lymphotoxin β in Mesenteric Lymph Node Genesis. The Journal of Experimental Medicine. 2002Oct;195(12):1613–24.

209. Hsu H, Solovyev I, Colombero A, Elliott R, Kelley M, Boyle WJ. ATAR, a Novel Tumor Necrosis Factor Receptor Family Member, Signals through TRAF2 and TRAF5. Journal of Biological Chemistry. 1997;272(21):13471–4.

210. Marsters SA, Ayres TM, Skubatch M, Gray CL, Rothe M, Ashkenazi A. Herpesvirus Entry Mediator, a Member of the Tumor Necrosis Factor Receptor (TNFR) Family, Interacts with Members of the TNFR-associated Factor Family and Activates the Transcription Factors NF-κB and AP-1. Journal of Biological Chemistry.

1997;272(22):14029–32.

211. Kabashima K, Banks TA, Ansel KM, Lu TT, Ware CF, Cyster JG. Intrinsic Lymphotoxin-β Receptor Requirement for Homeostasis of Lymphoid Tissue Dendritic Cells. Immunity. 2005;22(4):439–50.

212. Xu G, Liu D, Okwor I, Wang Y, Korner H, Kung SKP, et al. LIGHT Is Critical for IL-12 Production by Dendritic Cells, Optimal CD4 Th1 Cell Response, and Resistance to Leishmania major. The Journal of Immunology. 2007Feb;179(10):6901–9.

213. Serin ER, Nijveen H, Hilhorst H.M.W, et al,. Learning from co-expression networks:possibilities and challenges. Front Plant Sci. 2016;7:444

214. Cheng H, Li Y, Zuo X-B, Tang H-Y, Tang X-F, Gao J-P, et al. Identification of a Missense Variant in LNPEP that Confers Psoriasis Risk. Journal of Investigative Dermatology. 2014;134(2):359–65.

215. Guerra ICA, Gisbert JP. Onset of psoriasis in patients with inflammatory bowel disease treated with anti-TNF agents. Expert Review of Gastroenterology & Hepatology. 2013;7(1):41–8.

216. Takemasa I. Potential biological insights revealed by an integrated assessment of proteomic and transcriptomic data in human colorectal cancer. International Journal of Oncology. 2011;

217. Gall BJ, Wilson A, Schroer AB, Gross JD, Stoilov P, Setola V, et al. Genetic variations in GPSM3 associated with protection from rheumatoid arthritis affect its transcript abundance. Genes & Immunity. 2016;17(2):139–47.

218. Christensen SR, Shupe J, Nickerson K, Kashgarian M, Flavell RA, Shlomchik MJ. Toll-like Receptor 7 and TLR9 Dictate Autoantibody Specificity and Have Opposing Inflammatory and Regulatory Roles in a Murine Model of Lupus. Immunity. 2006;25(3):417–28.

219. Peng S. Corrigendum: Increased Toll-Like Receptors Activity and TLR Ligands in Patients with Autoimmune Thyroid Diseases. Frontiers in Immunology. 2017;8.

220. Lacerte P, Brunet A, Egarnes B, Duchêne B, Brown JP, Gosselin J. Overexpression of TLR2 and TLR9 on monocyte subsets of active rheumatoid arthritis patients contributes to enhance responsiveness to TLR agonists. Arthritis Research & Therapy. 2016;18(1).

221. Huang Q, Pope RM. ROLE OF TOLL LIKE RECEPTORS IN RHEUMATOID ARTHRITIS. Curr Rheumatol Rep. 2009;11(5):357–64.

222. Littlepage LE, Adler AS, Kouros-Mehr H, Huang G, Chou J, Krig SR, et al. The Transcription Factor ZNF217 Is a Prognostic Biomarker and Therapeutic Target during Breast Cancer Progression. Cancer Discovery. 2012Oct;2(7):638–51.

223. Huang G, Krig S, Kowbel D, Xu H, Hyun B, Volik S, et al. ZNF217 suppresses cell death associated with

chemotherapy and telomere dysfunction. Human Molecular Genetics. 2005Mar;14(21):3219–25.

224. Yu M, Zhou X, Niu L, Lin G, Huang J, Zhou W, et al. Targeting Transmembrane TNF-α Suppresses Breast Cancer Growth. Cancer Research. 2013;73(13):4061–74

225. Banck MS, Li S, Nishio H, Wang C, Beutler AS, Walsh MJ. The ZNF217 oncogene is a candidate organizer of repressive histone modifiers. Epigenetics. 2009;4(2):100–6.

226. Cohen PA, Donini CF, Nguyen NT, Lincet H, Vendrell JA. The dark side of ZNF217, a key regulator of tumorigenesis with powerful biomarker value. Oncotarget. 2015;6(39):41566–81.

227. Lee D-F, Walsh MJ, Aguiló F. ZNF217/ZFP217 Meets Chromatin and RNA. Trends in Biochemical Sciences. 2016;41(12):986–8.

228. Lazarczyk M, Pons C, Mendoza J-A, Cassonnet P, Jacob Y, Favre M. Regulation of cellular zinc balance as a potential mechanism of EVER-mediated protection against pathogenesis by cutaneous oncogenic human papillomaviruses. The Journal of Cell Biology. 2007;179(7):i21–i21.

229. Yu M, Lee W-W, Tomar D, Pryshchep S, Czesnikiewicz-Guzik M, Lamar DL, et al. Regulation of T cell receptor signaling by activation-induced zinc influx. The Journal of Experimental Medicine. 2011;208(4):775–85.

230. Leitenberg D, Falahati R, Lu DD, Takeda A. CD45-associated protein promotes the response of primary CD4 T cells to low-potency T-cell receptor (TCR) stimulation and facilitates CD45 association with CD3/TCR and Lck. Immunology. 2007;121(4):545–54.

231. Schwinzer R, Witte T, Hundrieser J, Ehlers S, Momot T, Hunzelmann N, et al. Enhanced frequency of a PTPRC (CD45) exon A mutation (77C→G) in systemic sclerosis. Genes & Immunity. 2003;4(2):168–9.

232. Buckley RH. Molecular defects in human severe combined immunodeficiency and approaches to immune reconstitution. Annu Rev Immunol. 2004;22:625–55.

233. Kerber RA, O'Brien E, Cawthon RM. Gene expression profiles associated with aging and mortality in humans. Aging Cell. 2009;8(3):239–50.

234. Okumura M, Kung C, Wong S, Rodgers M, Thomas ML. Definition of Family of Coronin-Related Proteins Conserved Between Humans and Mice: Close Genetic Linkage Between Coronin-2 and CD45-Associated Protein. DNA and Cell Biology. 1998;17(9):779–87.

235. Mueller P, Massner J, Jayachandran R, Combaluzier B, Albrecht I, Gatfield J, et al. Regulation of T cell survival through coronin-1–mediated generation of inositol-1,4,5-trisphosphate and calcium mobilization after T cell receptor triggering. Nature Immunology. 2008;9(4):424–31.

236. Fracchia KM, Pai CY, Walsh CM. Modulation of T Cell Metabolism and Function through Calcium Signaling.

Frontiers in Immunology. 2013;4.

237. Rodríguez-Cortez VC, Pino-Molina LD, Rodríguez-Ubreva J, Ciudad L, Gómez-Cabrero D, Company C, et al. Monozygotic twins discordant for common variable immunodeficiency reveal impaired DNA demethylation during naïve-to-memory B-cell transition. Nature Communications. 2015;6:7335.

238. Sekino, S., Kashiwagi, Y., Kanazawa, H. et al. The NESH/Abi-3-based WAVE2 complex is functionally distinct from the Abi-1-based WAVE2 complex. Cell Commun Signal. 2015: 13 (41).

239. Lauzier A, Charbonneau M, Harper K, Jilaveanu-Pelmus M, Dubois CM. Formation of invadopodia-like structures by synovial cells promotes cartilage breakdown in collagen-induced arthritis: Involvement of the protein tyrosine kinase Src. Arthritis & Rheumatism. 2011;63(6):1591–602.

240. Chen L-Y, Chen JD. Daxx Silencing Sensitizes Cells to Multiple Apoptotic Pathways. Molecular and Cellular Biology. 2003;23(20):7108–21.

241. Salomoni P, Guernah I, Pandolfi PP. The PML-nuclear body associated protein Daxx regulates the cellular response to CD40. Cell Death and Differentiation. 2005;13(4):672–5.

242. Lian S., Liu T., Jing S., Yuan H., Zhang Z. and Cheng L. Intrachromosomal colocalization strengthens co-expression, co-modification and evolutionary conservation of neighboring genes. BMC Genomics. 2018;19(1):455

243. Michalak, P. Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. Genomics 2008; 91(3): 243-48.

244. Zhu H., Wang G. and Qian J. Transcription factors as readers and effectors of DNA methylation. Nat Rev Genet. 2016;17(9):551-65.

245. Marchal C. and Miotto, B. Emerging Concept in DNAMethylation: Role of Transcription Factors in Shaping DNA Methylation Patterns. J Cell Physiol. 2015; 230: 743-51.

246. Hervouet, E., Peixoto, P., Delage-Mourroux, R., Boyer-Guittaut, M. and Cartron, P. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. Clin Epigenet. 2018; 10(17).

247. Holmadahl R, Association of MHC and rheumatoid arthritis: Why is rheumatoid arthritis associated with the MHC genetic region? An introduction. Arthritis Res. 2000;2(3): 203-204

248. Maltby VE, Lea RA, Sanders KA, White N, Benton MC, Scott RJ, et al. Differential methylation at MHC in CD4 T cells is associated with multiple sclerosis independently of HLA-DRB1. Clinical Epigenetics. 2017;9(1).

249. Sambrook JG, Figueroa F, Beck S. A genome-wide survey of Major Histocompatibility Complex (MHC) genes and their paralogues in zebrafish. BMC Genomics. 2005;6:152

250. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nature Biotechnology. 2013;31(2):142–7.

251. Steenbergen, H., Luijk, R., Shoemaker, R., Heijmans, B., Huizinga, T., Mil, A. Differential methylation within the major histocompatibility complex region in rheumatoid arthritis: a replication study, *Rheumatology*, 2014; 53(12): 2317–18.

252. Anaparti, V., Agarwal, P., Smolik, I., Mookherjee, N., Elgabalawy, H.,Whole Blood Targeted Bisulfite Sequencing Validates Differential Methylation in C6ORF10 gene of Patients with Rheumatoid Arthritis. J Rheumatol. 2019: 47(1)

253. Bebenek K., Abbotts J., Wilson S. and Kunkel T.Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. J Biol Chem. 1993; 268(14):10324-34

254. Hahn M., Szabó P. and Pfeifer G. 5-Hydroxymethylcytosine: a stable or transient DNA modification?. *Genomics*. 2014;104(5):314-23.

255. Li W, Zhang X, Lu X, You L, Song Y, Luo Z, Zhang J, Nie J, Zheng W, Xu D, Wang Y, Dong Y, Yu S, Hong J, Shi J, Hao H, Luo F, Hua L, Wang P, Qian X, Yuan F, Wei L, Cui M, Zhang T, Liao Q, Dai M, Liu Z, Chen G, Meckel K, Adhikari S, Jia G, Bissonnette MB, Zhang X, Zhao Y, Zhang W, He C, Liu J. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res*. 2017;27(10):1243-57.

7. Appendix

7.1 Ethical Approval

National Research Ethics Service

Leeds (West) Research Ethics Committee
First Floor
Millside
Mill Pond Lane
Leeds
LS6 4EP

Tel: 0113 305 0116

24 August 2010

Professor Paul Emery
LEEDS TEACHING HOSPITALS NHS TRUST
Department of Rheumatology. Second Floor.
Chapel Allerton Hospital, Chapeltown Road
Leeds
LS7 4SA

Dear Professor Emery

Study title: **Inflammatory arthritis disease continuum longitudinal study**
REC reference: **09/H1307/98**
Protocol number: **RR09/9134**
Amendment number: **1**
Amendment date: **05 August 2010**

The above amendment was reviewed at the meeting of the Sub-Committee held on 17 August 2010.

Ethical opinion

Favourable Opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the provision that the following points be adhered to:

1) The word 'research' should be added to the invitation letter, after the phrase "potentially suitable for the ICON Disease continuum".

2) The invitation letter should come from (and be signed by) the research team – not the Consultant in charge of the patient's care.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Sample request form		
Protocol	3.0	02 August 2010
Letter of invitation to participant	1.0	02 August 2010
Participant Information Sheet: Information Sheet and Consent Form	3.0	02 August 2010

This Research Ethics Committee is an advisory committee to the Yorkshire and The Humber Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

Participant Information Sheet: Information Sheet and Consent Form	3.0	02 August 2010
Protocol	3.0	02 August 2010
Notice of Substantial Amendment (non-CTIMPs)		05 August 2010

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

09/H1307/98:	**Please quote this number on all correspondence**

Yours sincerely

Claire Kelly
Committee Assistant Co-ordinator

E-mail: Claire.kelly@leedspft.nhs.uk

Copy to: *Mrs Rachel de Souza*

7.2 Raw data and ranking methodology

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
388238	cg27531490	chr6	31542459	TNF	NA	3.3965764	0.8106	5.30E-03	0.856	0.853	0.878	0.859	0.824	0.814	0.748	0.827	0.742	0.676	0.767	0.825	0.714	0.833
388239	cg24452282	chr6	31542740	TNF	NA	4.4319505	0.9742	8.18E-04	0.813	0.824	0.851	0.864	0.789	0.774	0.749	0.721	0.689	0.649	0.63	0.793	0.723	0.726
388240	cg11484872	chr6	31543169	TNF	NA	2.6651857	0.59655	2.06E-02	0.339	0.429	0.484	0.413	0.363	0.365	0.389	0.357	0.334	0.263	0.338	0.324	0.29	0.209
388241	cg21370522	chr6	31543219	TNF	NA	2.4856151	0.5139	2.87E-02	0.19	0.21	0.253	0.194	0.201	0.185	0.178	0.201	0.161	0.095	0.145	0.137	0.143	0.156
388242	cg01569083	chr6	31543289	TNF	NA	3.5380812	0.61185	4.09E-03	0.326	0.336	0.412	0.312	0.32	0.297	0.289	0.283	0.259	0.19	0.232	0.232	0.242	0.242
388243	cg03037030	chr6	31543300	TNF	NA	2.7816249	0.53725	1.66E-02	0.186	0.198	0.252	0.169	0.186	0.179	0.173	0.174	0.137	0.108	0.12	0.151	0.144	0.116
388244	cg12681001	chr6	31543540	TNF	NA	3.2671512	0.41995	6.74E-03	0.261	0.297	0.304	0.245	0.257	0.264	0.249	0.221	0.23	0.186	0.206	0.197	0.215	0.203
388245	cg21222743	chr6	31543545	TNF	NA	1.7887543	0.2974	9.89E-02	0.21	0.254	0.281	0.204	0.24	0.259	0.229	0.186	0.216	0.171	0.174	0.177	0.196	0.175
388246	cg10717214	chr6	31543557	TNF	NA	3.3279019	0.42605	6.02E-03	0.258	0.272	0.327	0.253	0.263	0.252	0.241	0.224	0.229	0.181	0.203	0.204	0.22	0.207
388247	cg04425624	chr6	31543565	TNF	NA	4.3539862	0.5988	9.38E-04	0.343	0.362	0.417	0.348	0.303	0.326	0.317	0.288	0.285	0.22	0.282	0.259	0.255	0.248
388248	cg21467614	chr6	31543638	TNF	NA	1.8450568	0.3291	8.98E-02	0.289	0.312	0.359	0.256	0.315	0.328	0.281	0.262	0.273	0.208	0.233	0.24	0.215	0.23
388249	cg08553327	chr6	31543647	TNF	NA	3.2393231	0.6673	7.10E-03	0.36	0.397	0.457	0.376	0.308	0.337	0.371	0.337	0.316	0.189	0.296	0.265	0.261	0.278
388250	cg26729380	chr6	31543655	TNF	NA	2.7197805	0.5344	1.86E-02	0.322	0.341	0.383	0.315	0.304	0.329	0.324	0.299	0.253	0.181	0.261	0.239	0.223	0.236
388251	cg10650821	chr6	31543686	TNF	NA	3.8391181	0.54395	2.36E-03	0.252	0.267	0.304	0.235	0.233	0.22	0.229	0.211	0.205	0.148	0.181	0.193	0.175	0.19
388252	cg17741993	chr6	31544694	TNF	N_Shelf	4.6915091	1.30195	5.22E-04	0.497	0.456	0.579	0.555	0.396	0.296	0.437	0.396	0.31	0.228	0.257	0.331	0.249	0.23
388253	cg01360627	chr6	31544931	TNF	N_Shelf	3.6922141	0.9617	3.08E-03	0.671	0.658	0.773	0.729	0.608	0.626	0.667	0.608	0.518	0.4	0.517	0.585	0.523	0.513
388254	cg23384708	chr6	31544934	TNF	N_Shelf	3.9834911	0.71445	1.81E-03	0.545	0.568	0.624	0.621	0.485	0.488	0.585	0.485	0.447	0.373	0.439	0.479	0.475	0.421
388255	cg20477259	chr6	31544960	TNF	N_Shelf	4.0694105	0.84745	1.56E-03	0.593	0.61	0.671	0.662	0.516	0.528	0.598	0.541	0.454	0.375	0.449	0.541	0.48	0.434
388256	cg09637172	chr6	31545252	TNF	N_Shelf	3.1702333	1.0786	8.07E-03	0.955	0.93	0.946	0.944	0.933	0.919	0.87	0.905	0.879	0.786	0.844	0.94	0.894	0.885
388257	cg05952498	chr6	31545257	TNF	N_Shelf	3.7710335	0.797	2.67E-03	0.863	0.861	0.893	0.871	0.815	0.835	0.799	0.832	0.773	0.704	0.762	0.861	0.782	0.799
388258	cg15989608	chr6	31545321	TNF	N_Shelf	2.3415074	0.50325	3.73E-02	0.807	0.799	0.838	0.844	0.803	0.807	0.786	0.773	0.762	0.657	0.686	0.816	0.782	0.769
388259	cg26736341	chr6	31545342	TNF	N_Shelf	2.9013877	0.58675	1.33E-02	0.802	0.801	0.824	0.801	0.775	0.791	0.762	0.757	0.707	0.641	0.646	0.786	0.742	0.734
388260	cg04472685	chr6	31545473	TNF	N_Shelf	1.5812575	0.2688	1.40E-01	0.843	0.858	0.847	0.843	0.855	0.839	0.819	0.82	0.811	0.742	0.789	0.86	0.832	0.841
388261	cg19124225	chr6	31545836	TNF	N_Shelf	-0.424156	-0.0521	6.79E-01	0.916	0.9	0.895	0.882	0.906	0.904	0.898	0.909	0.901	0.876	0.897	0.911	0.896	0.918
388262	cg02137984	chr6	31545898	TNF	N_Shelf	0.8564277	0.2341	4.09E-01	0.942	0.929	0.918	0.937	0.944	0.926	0.95	0.918	0.898	0.867	0.887	0.94	0.928	0.924
388263	cg06825478	chr6	31546067	TNF	N_Shelf	0.7160598	0.19955	4.88E-01	0.706	0.769	0.839	0.815	0.782	0.813	0.869	0.743	0.735	0.697	0.741	0.756	0.74	0.709
388264	cg17755321	chr6	31546085	TNF	N_Shelf	2.1899906	0.44775	4.90E-02	0.863	0.902	0.911	0.928	0.873	0.882	0.917	0.865	0.87	0.851	0.85	0.871	0.893	0.852
388265	cg02581828	chr6	31546769	TNF	N_Shore	3.8086325	0.21425	2.49E-03	0.63	0.628	0.633	0.652	0.594	0.606	0.59	0.601	0.57	0.599	0.609	0.616	0.631	0.587
388266	cg10894274	chr6	31546773	TNF	N_Shore	3.4012896	0.24745	5.26E-03	0.664	0.692	0.695	0.669	0.646	0.654	0.61	0.652	0.622	0.614	0.664	0.653	0.635	0.667
388267	cg00710269	chr6	31546827	TNF	N_Shore	3.8902475	0.3096	2.15E-03	0.661	0.66	0.68	0.681	0.613	0.625	0.58	0.603	0.655	0.641	0.622	0.656	0.603	0.623

Figure 10. Raw data from initial methylation array of the TNF gene's CpGs. Row A is CpG probe name, B is chromosome position, C is CpG coordinate, D is gene name, E is CpG feature, F is T-statistic, G is difference in mean (negative and positive values indicates RA is hypermethylated and hypomethylated, respectively), H is P value, Rows I-L are HC methylation levels, M-V are RA methylation levels.

33	Probe_ID	Chromosome	CpG_coordir	Difference :	Gene	Feature	T statistic	Difference in mean	p value	C05N	C06N	C07N	C08N	R02N-	R03N-	R07N-	R08N-	R01N+	R04N+	R05N+	R06N+	R09N+	R10N+
34	cg2753149	chr6	31542459	281	TNF	NA	3.396576413	0.8106	5.30E-03	0.856	0.853	0.878	0.859	0.824	0.814	0.748	0.827	0.742	0.676	0.767	0.825	0.714	0.833
35	cg2445228	chr6	31542740	549	TNF	NA	4.431950502	0.9742	8.18E-04	0.813	0.824	0.851	0.864	0.789	0.774	0.749	0.721	0.689	0.649	0.63	0.793	0.723	0.726
36	cg0156908	chr6	31543289	251	TNF	NA	3.538081194	0.61185	4.09E-03	0.326	0.336	0.412	0.312	0.32	0.297	0.289	0.283	0.259	0.19	0.232	0.232	0.242	0.242
37	cg1268100	chr6	31543540	17	TNF	NA	3.267151223	0.41995	6.74E-03	0.261	0.297	0.304	0.245	0.257	0.264	0.249	0.221	0.23	0.186	0.206	0.197	0.215	0.203
38	cg1071721	chr6	31543557	8	TNF	NA	3.327901902	0.42605	6.02E-03	0.258	0.272	0.327	0.253	0.263	0.252	0.241	0.224	0.229	0.181	0.203	0.204	0.22	0.207
39	cg0442562	chr6	31543565	82	TNF	NA	4.35398624	0.5988	9.38E-04	0.343	0.362	0.417	0.348	0.303	0.326	0.317	0.288	0.285	0.22	0.282	0.259	0.255	0.248
40	cg0855332	chr6	31543647	39	TNF	NA	3.23932314	0.6673	7.10E-03	0.36	0.397	0.457	0.376	0.308	0.337	0.371	0.337	0.316	0.189	0.296	0.265	0.261	0.278
41	cg1065082	chr6	31543686	1008	TNF	NA	3.83911814	0.54395	2.36E-03	0.252	0.267	0.304	0.235	0.233	0.22	0.229	0.211	0.205	0.148	0.181	0.193	0.175	0.19
42	cg1774199	chr6	31544694	237	TNF	N_Shelf	4.691509088	1.30195	5.22E-04	0.497	0.456	0.579	0.555	0.396	0.296	0.437	0.396	0.31	0.228	0.257	0.331	0.249	0.23
43	cg0136062	chr6	31544931	3	TNF	N_Shelf	3.692214056	0.9617	3.08E-03	0.671	0.658	0.773	0.729	0.608	0.626	0.667	0.608	0.518	0.4	0.517	0.585	0.523	0.513
44	cg2338470	chr6	31544934	26	TNF	N_Shelf	3.983491072	0.71445	1.81E-03	0.545	0.568	0.624	0.621	0.485	0.488	0.585	0.485	0.447	0.373	0.439	0.479	0.475	0.421
45	cg2047725	chr6	31544960	292	TNF	N_Shelf	4.069410539	0.84745	1.56E-03	0.593	0.61	0.671	0.662	0.516	0.528	0.598	0.541	0.454	0.375	0.449	0.541	0.48	0.434
46	cg0963717	chr6	31545252	5	TNF	N_Shelf	3.170233259	1.0786	8.07E-03	0.955	0.93	0.946	0.944	0.933	0.919	0.87	0.905	0.879	0.786	0.844	0.94	0.894	0.885
47	cg0595249	chr6	31545257	1512	TNF	N_Shelf	3.771033511	0.797	2.67E-03	0.863	0.861	0.893	0.871	0.815	0.835	0.799	0.832	0.773	0.704	0.762	0.861	0.782	0.799
48	cg0258182	chr6	31546769	4	TNF	N_Shore	3.808632481	0.21425	2.49E-03	0.63	0.628	0.633	0.652	0.594	0.606	0.59	0.601	0.57	0.599	0.609	0.616	0.631	0.587
49	cg1089427	chr6	31546773	54	TNF	N_Shore	3.401289621	0.24745	5.26E-03	0.664	0.692	0.695	0.669	0.646	0.654	0.61	0.652	0.622	0.614	0.664	0.653	0.635	0.667
50	cg0071026	chr6	31546827		TNF	N_Shore	3.890247533	0.3096	2.15E-03	0.661	0.66	0.68	0.681	0.613	0.625	0.58	0.603	0.655	0.641	0.622	0.656	0.603	0.623

Figure 11. TNF gene CpGs after filtering by ($p \leq 1x10^{-2}$). A difference column has been added to observe the bp distance between sig

CpGs. Green, yellow and red cells n 500bp respectively.

Gene	Chromosome	Largest green cluster	Number of greens	Largest coloured cluster	Number of coloured	Less or More Methylated	encodes Codes for	Notes
B3GALT4	6	4	10	13	15	Less Methylated	Beta-1,3-Galactosyltransferase 4	Its family catalyzes the transfer of galactose. Galactose is involved in T-cell activation.
TNF	6	4	9	5	13	Less Methylated	Tumor Necrosis Factor	TNF-α Cytokine
DAXX	6	4	6	8	10	Less Methylated	Death-Domain Associated Protein	Down-regulates basal and activated transcription. Transcription corepressor known to repress transcriptional potential of several sumoylated transcription factors. Aids in Apoptosis
ABAT	16	4	6	7	7	More Methylated	4-Aminobutyrate Aminotransferase	Responsible for catabolism of gamma-aminobutyric acid (GABA), an important, mostly inhibitory neurotransmitter in the central nervous system, into succinic semialdehyde
MIR21	17	3	5	3	8	Less Methylated	MicroRNA 21	miR-21 is one of the most frequently upregulated miRNAs in solid tumours, and its high levels were first described in B cell lymphomas
NEAT1	11	3	5	3	6	Less Methylated	Nuclear Enriched Abundant Transcript 1	this novel noncoding RNA appears to have an important structural role in the nuclear paraspeckles. Paraspeckles contribute to transcriptional regualtion.
HOXB-AS3	17	4	5	5	6	More Methylated	HOXB Cluster Antisense RNA 3	Regulates HOX-B3. Increased expression of HOXB3results is associated with a distinct biologic subset of acute myeloid leukemia.
CUTA	6	5	5	6	8	Less Methylated	Acetylcholinesterase-associated protein	May form part of a complex of membrane proteins attached to acetylcholinesterase (AChE)
SNORD32A	19	5	5	5	5	Less Methylated	Small Nucleolar RNA, C/D Box 32A	SNORDS are small nucleolar RNAs which usually functions in guiding the modification of other non-coding RNAs.
RNF39	6	5	5	5	5	More Methylated	RING finger protein 39	Plays a role in an early phase of synaptic plasticity
OCSTAMP	20	2	4	5	5	More Methylated	Osteoclast Stimulatory Transmembrane Protein	Probable cell surface receptor that plays a role in cellular fusion and cell differentiation. Involved in osteoclast bone resorption
ABI3	17	4	4	4	4	Less Methylated	ABI Family Member 3	The encoded protein inhibits ectopic metastasis of tumor cells as well as cell migration. This may be accomplished through interaction with p21-activated kinase. Alternative splicing results in multiple transcript variants
HORMAD2	22	4	4	5	5	More Methylated	HORMA domain-containing protein 2	Essential for synapsis surveillance during meiotic prophase via the recruitment of ATR activity
HLA-H	6	4	4	4	4	More Methylated	MHC Class I Antigen H	Involved in the presentation of foreign antigens to the immune system.
KSR1	17	3	3	7	7	Less Methylated	Kinase Suppressor Of Ras 1	Among its related pathways are Negative regulation of MAPK pathway and Interleukin-3, 5 and GM-CSF signaling.

Figure 12. Top 15 ranked genes based on proximity of significant differentially methylated CpGs. Red genes are hypermethylated and green are hypomethylated. They are ranked in descending order of significant 'green' CpGs.

ABAT	B3GALT4
HOXB-AS3	TNF
RNF39	DAXX
OCSTAMP	CUTA
HORMAD2	SNORD32A
HLA-H	MIR21
BAIAP2-AS1	NEAT1
CALCOCO1	ABI3
CLSTN1	KSR1
GLMP	MEOX1
HEYL	CD160
HCG9	RFPL2
LOC81691	TAP1
MIR7-3HG	TRIM69
PAX1	HLA-C

Figure 13. Top 15 ranked hypomethylated genes (green), and top 15 ranked hypermethylated genes (red). Listed from top to bottom in order of significance. Taken from the data set shown in Figure 12.

7.3 Designed primers and products

The DNA sequences below show the designed primers and products for the TNF gene. The DNA is orientated 5' to 3', the primers are highlighted yellow and the desired significant CpG to be sequenced is highlighted pink.

Product 1

ATAGGAGATTTTTGGGGAGATGTGATTATAGTAATGGGTAGGAGAATGTT
TAGGGTTATGGAAGTCGAGTATGGGGATTTTTTTTTAACGAAGATAGGGT
TATGTAGAGGGTTTTAGGGAGTGAAAGAGTTTTTAGGATTTTTAGGTATG

150bp Product

Fwd: 25bp Tm:60.0

Rev: 32bp Tm:59.7

Product 2

TTTGTTTTAGTGGGGTTTGTGAATTTTCGGGGGTGATTTTATTTTTCGGG
GTTGTTTTAGGTTTGTTTTTGTTATTTTTATTTAGTTTTTTTTGAGGTTT
TAAGTTTGTTATTAAGTTTTTAGTTTTTTTTTTTCGTAGGGATTTAAATA
TAGGTTTTAGGATTTAAT

166Bp Product

Fwd: 22bp Tm:55.9

Rev: 30bp Tm:56.0

Product 3

GGGGAGTGTGAGGGGTATTTTTGATGTTT
GTGTGTTTTTAATTTTTTAAATTTTCGTTTTCGCGATGGAGAAGAAATCG
AGATAGAAGGTGTAGGGTTTATTATCGTTTTTTTTAGATGAGTTTATGGG
TTTTTTTATTAAGGAAGTTTTTCGTTGGTTGAATGATTTTTTTTTCGTTT
TTTTTTCGTTTTAGGGATATATAAAGGTAGTTGTTGGTATATTTAGTTAG
TAGACGTTTTTTTAGTAAGGATAGTAGAGGATTAGTTAAGAGGGAGAG

273 bp Product

Fwd: 25bp Tm:58.0

Rev: 30bp Tm:58.2

Product 4

ATATT**ATG**AGTATTGAAAGTATGATTCGGG
ACGTGGAGTTGGTCGAGGAGGCGTTTTTTAAGAAGATAGGGGGGTTTTAG
GGTTTTAGGCGGTGTTTGTTTTTTAGTTTTTTTTTTTTTTTGATCGTGGT
AGGCGTTATTACGTTTTTTTGTTTGTTGTATTTTGGAGTGATCGGTTTTT
AGAGGGAAGAGGTGAGTGTTTGGTTAGTTTTTATTTATTTTTTTATTTAA
GGGGAAATGGAGACGTAAGAGAGGGAGAGAGATGGATGGG

270 bp product

Fwd: 27bp Tm:58.8

Rev: 20bp Tm:58.9

Product 5

TGAAGTTCGGTTGATGGTAGGTAGAATTTGGAGATAATGTGAGAAGGATT
CGTTGAGTTTAAGGGAAGGGTGGAGGAATAGTATAGGTTTTAGTGGGATA
TTTAGAACGTTATGGTTAGGTGGGATGTGGGATGATAGATAGAGAGGATA
GGAATCGGATGTGGGGTGGGTAGAGTTCGAGGGTTAGGATGTGGAGAGTG
AATCGATATGGTTATATTGATTTTTTTTTTTTTTTTTTTTTTTTTTTAGT
AAATTTTTAAGTTGAGGGGTAGTTTTAGTGGTTGAATCGTCGGGTTAATG
TTTTTTTGGTTAATGGCGTGGAGTTGAGAGATAATTAGTTGGTGGTGTTA
TTAGAGGGTTTGTATTT

370bp product

Fwd: 25bp Tm: 59.9

Rev: 30bp Tm: 59.7

Product 6

GAGGGGTTTTTTAGTTGGAGAAGGGTGATCGATTTAGCGTTGAGATTA
ATCGGTTCGATTATTTCGATTTTGTCGAGTTTGGGTAGGTTTATTTTG
GGATTATTGTTTTG**TGA**GGAGGACGAATATTTAATTTTTTTAAACGTTT
TTTTTGTTTTAATTTTTTTATTATTTTTTTTTTTAGATATTTTTAATTT
TTTTTGGTTTAAAAAGAGAATTGGGGGTTT

224 Bp product

Fwd: 24bp Tm:59.4

Rev: 30bp Tm:59.3

7.4 Gene definitions

Gene Acronym	Gene Name
A20	Tumor necrosis factor, alpha-induced protein 3
ABAT	4-Aminobutyrate Aminotransferase
ABI3	ABI Family Member 3
AP-1	Activator protein 1
ATF2	Activating transcription factor 2
B3GALT4	Beta-1,3-galactosyltransferase 4
BAALC	Brain and Acute Leukemia, Cytoplasmic
BAIAP2	Brain-specific angiogenesis inhibitor 1-associated protein 2
BAIAP2-AS1	Brain-specific angiogenesis inhibitor 1-associated protein 2 - Antisense 1
BCL-X1	B-cell lymphoma-extra large
BCL2	B-cell lymphoma 2
BFL-A1	BCL2-related protein A1
BID	BH3 interacting-domain death agonist
BRCA1	Breast Cancer 1
CRel	REL proto-oncogene
C myb/c	Cancers Myelocytomatosis
CASP 8	CASPASE 8
CD11a	**Integrin, alpha L**
CD160	Cluster of Differentiation 160
CD247	Cluster of Differentiation 247
CD30	Cluster of Differentiation 30
CD40	Cluster of Differentiation 40 Ligand
CD40L	Cluster of Differentiation 40 Ligand
CD70	Cluster of Differentiation 70
CELF4	CUG-BP- And ETR-3-Like Factor 4
COL11A2	Collagen alpha-2 (XI) chain
COL12A	Collagen alpha-1 (XII) chain
COL16A1	Collagen alpha-1(XVI) chain
COL1A1	Collagen alpha-1 (I) Chain
COL23A1	Collagen alpha-1 (XXIII) Chain
COL29A1	Collagen alpha-1 (XXIX) Chain
COL6A3	Collagen alpha-3(VI) chain
COL8A1	Collagen alpha-1 (VIII) Chain
CORO1A	Coronin-1A
CUTA	Acetylcholinesterase-Associated Protein
CXCL12	stromal cell-derived factor 1
DAXX	Death-associated protein 6
DR3	Death receptor 3.
ERK	Mitogen-activated protein kinase
FADD	Fas-associated protein with death domain,
FAS	TNF receptor superfamily 6
GAREM1	GRB2 associated regulator of MAPK1 subtype 1
GBR2	Growth factor receptor-bound protein 2
GCSF	Granulocyte-colony stimulating factor
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GNGT2	guanine nucleotide-binding protein, gamma-transducing activity polypeptide 2
GPSM3	G Protein Signaling Modulator 3
HIC1	Hypermethylated in cancer 1

Gene Acronym	Gene Name
RNF39	RING finger protein 39
RPS18	**40S ribosomal protein S18**
RUNX1	Runt-related transcription factor 1
RUNX2	Runt-related transcription factor 2
RUNX3	Runt-related transcription factor 3
SFRP2	Secreted frizzled-related protein 2
SHARPIN	SHANK Associated RH Domain Interactor
SNORD32A	Small nucleolar RNA, C/D box 32A
STAT1	Signal transducer and activator of transcription 1
TAB2	TGF-Beta Activated Kinase 1/MAP3K7 Binding Protein 2
TAB3	TGF-Beta Activated Kinase 1/MAP3K7 Binding Protein 3
TAK1	MAP3K7
TBID	BH3 interacting-domain death agonist
TCR	T cell receptor
TGFβ	Transforming Growth Factor Beta 1
TLR9	Toll-like receptor 9
TMC6	Transmembrane Channel Like 6
TNF	Tumour Necrosis Factor
TNFR1	Tumour Necrosis Factor Receptor 1
TNFR2	Tumour Necrosis Factor Receptor 2
TRADD	Tumor necrosis factor receptor type 1-associated DEATH domain protein
TRAF1	TNF Receptor Associated Factor 1
TRAF2	TNF Receptor Associated Factor 2
TRAF5	TNF Receptor Associated Factor 5
TRAF6	TNF Receptor Associated Factor 6
TRAIL	TNF-related apoptosis-inducing ligand
THBS1	Thrombospondin
UACA	Uveal autoantigen with coiled-coil domains and ankyrin repeats
USP18	Ubiquitin specific peptidase 18
VEGF	Vascular endothelial growth factor,
ZAP70	Zeta Chain Of T Cell Receptor Associated Protein Kinase 70
ZC3H12D	zinc finger CCCH-type containing 12B
ZNF217	Zinc finger protein 217
cFOS	FBJ murine osteosarcoma viral oncogene homolog

Gene Acronym	Gene Name
HOXB-AS3	HOXB Cluster Antisense RNA 1
HPN	Hepsin
HVEM	Herpesvirus entry mediator
IAPs	inhibitor of apoptosis family of proteins
IFI44L	Interferon-induce protein 44-like
IFITM1	Interferon-induced transmembrane protein 1
IFT140	Intraflagellar transport 140 homolog
IGFBP1/2	Insulin-like growth factor-binding protein 1/2
IKKa	Inhibitor of nuclear factor kappa-B kinase subunit alpha
IKKb	Inhibitor of nuclear factor kappa-B kinase subunit beta
IL1	Interleukin 1
IL10	Interleukin 10
IL13	Interleukin 13
IL21R	Interleukin 21R
IRF1/2	Interferon regulatory factor 1/ 2
JNK	c-Jun N-terminal kinases
JUN-B	JunB Proto-Oncogene, AP-1 Transcription Factor Subunit
KSR1	Kinase suppressor of Ras 1
LAT	Linker For Activation Of T Cells
LIGHT	TNF superfamily member 14,
LNPEP	Leucyl/cystinyl aminopeptidase
LTBR	**Lymphotoxin beta receptor**
MAP3K14-AS1	NF-kappa-B-inducing kinase – antisense 1
MAZ	Myc-associated zinc finger protein
MEKK	Mitogen-activated protein kinase kinase kinase
MEOX1	homeobox protein MOX-1
MLH1	MutL homolog 1, colon cancer, nonpolyposis type 2
MLL1	Histone-lysine N-methyltransferase 2A
NEMO	NF-kappa-B essential modulator
NFkB	nuclear factor-kappa-B
NFkB1	Nuclear factor NF-kappa-B p105 subunit
NFkB2	See NFkB definition
NIK	NF-kappa-B-inducing kinase
NR4A2	nuclear receptor subfamily 4 group A
Nr13	Anti-apoptotic protein NR13
OCSTAMP	osteoclast stimulatory transmembrane protein
P16	cyclin-dependent kinase inhibitor 2A
P38	Mitogen-activated protein kinase 14
P53	Tumor protein p53
PBX2	Pre-B-cell leukemia transcription factor 2
PDGF-bchain	Platelet-derived growth factor subunit B
PDGFRA	platelet-derived growth factor receptor A
PTPRCAP	Protein tyrosine phosphatase receptor type C-associated protein
Proenkephalin	proenkephalin A
RAF	MAP kinase kinase kinase
RANK	TNF ligand superfamily member 11
RAS	Rat sarcoma viral oncogene homolog
RELA	P65