The study of the prevalence of cartilage age-CpGs in articular joint tissues and the consequence on target-gene expression, homeostasis and pathogenesis.

Raul Cristian Fulea

Thesis submitted to Newcastle University for the degree of Doctor of Philosophy

Biosciences Institute

March 2023

Abstract

Osteoarthritis is a debilitating musculoskeletal disease having age as one of the biggest risk factors with DNA methylation being reported as a risk factor. DNA methylation is an epigenetic mechanism of gene regulation occurring at cytosine-guanine dinucleotides (CpG). Our group performed a genome-wide DNA methylation analysis of 179 human knee cartilage samples with 716-CpGs showing age-related DNA methylation changes (age-CpGs). The impact on gene expression is unknown, although DNA methylation is usually associated with repression of gene expression. The aims were to assess age-related DNA methylation changes in musculoskeletal tissues and the impact on gene expression, determine DNA methylation impact on transcription factor binding and the cellular consequences of these changes.

In total 23 age-CpGs were identified in cartilage, synovium and fat pad tissues using bisulphite pyrosequencing with 17 being across multiple tissues. These were predominantly located in promoter regions with 19 in total, with four in enhancer regions, where methylation has been observed to have a repressive activity on gene transcription. DNA methylation was correlated with repressed target gene expression at 27 CpG sites across musculoskeletal tissues. The promoter region of *FHL2* contains highly correlated age-CpGs in musculoskeletal tissues and publicly available datasets and the analysis further focused on the *FHL2* region using the electrophoretic mobility shift assay. DNA methylation increased the binding of multiple transcription factors such as Sp1 and Sp3 by up to 2.3-fold. The impact of CRISPR-Cas9 mediated knockdown and a plasmid overexpression vector of *FHL2* was studied by RNA-sequencing in SW1353 chondrosarcoma cells, and it was observed that this impacted several cellular pathways such as cellular proliferation, cellular metabolism and extracellular matrix synthesis.

In conclusion, the current study indicates the importance of age-CpG sites in gene expression regulation during ageing and the study of how these changes could impact cellular functionality.

Acknowledgements

I would like to thank my supervisor Dr. Louise Reynard and Prof. David Young, for their patience and continuous support during the project as well as their advice and encouragement along the way. I would also like to thank the Centre for Integrative Research into Musculoskeletal Ageing (CIMA) for the MRes/PhD studentship opportunity.

Special thanks to Marjolein Burgers, my colleague that greatly helped during this project especially during the difficult times of restricted laboratory access as well as support in ideas and discussion across the project. I would like to express my gratitude to Dr. Jamie Soul that helped me with guidance on the analysis of the RNA-sequencing data from Chapter 6 and to the members of the Skeletal Research Group that offered their support and help during the project.

I would also like to thank my partner Andreea Lungu, that helped me greatly with her support and understanding during the project period as well as for believing in me.

Table of contents

List of figures

List of tables

Chapter 1: Introduction

1.1 Ageing

Ageing is defined as a natural time-dependent process which affects living organisms. Marked by an decrease in functionality and an associated increase towards the predisposition to disease with an increased risk of mortality (Schumacher *et al.*, 2008). Ageing is a process where its cause is still in debate as some theories suggest that the continuous accumulation of cellular damage leads to loss of function, senescence, disease and ageing (Kirkwood, 2005; López-Otín *et al.*, 2013), whereas others support the theory of cell hyperfunction, as the process of developmental growth that drives the organism to ageing after reaching adulthood (Kirkwood and Austad, 2000); this implying that ageing causes cellular damage (e.g. DNA damage) and not the other way (Blagosklonny and Hall, 2009; Blagosklonny, 2010).

1.2 Cellular damage and link with ageing

At the organ and tissue level the changes associated with ageing need to be distinguished from the changes associated with the development of pathological conditions. Key changes with ageing are being marked as decrease and sometimes loss of the flexibility of compensatory mechanisms and the ability of the individual's system to respond to disease development (Flint and Tadi, 2021). Ageing is associated with many physiological, functional and metabolic changes, these including: altered liver metabolism, gastrointestinal malabsorption (Bhutto and Morley, 2008), kidney function reduction (Denic *et al.*, 2016), overall cerebral atrophy (Wyss-Coray, 2016), respiratory capacity reduction, decreased immune system defence capacity (Sharma and Goodwin, 2006), decrease in muscle strength (Keller and Engelhardt, 2014) and general altered hormonal homeostasis(van den Beld *et al.*, 2018). Age is considered as an important factor in the development of several illnesses including: sarcopenia (Franceschi *et al.*, 2018) , osteoarthritis (Loeser, 2017), neurodegenerative (Hou *et al.*, 2019), cancer (Berben *et al.*, 2021) and many more. These changes can impact individuals in many ways through the change in healthspan by increasing the predisposition to age-associated disease development as well as changing the overall treatment strategies for some of these conditions (Li *et al.*,

2021). The molecular changes associated with ageing seem to share pathways in common with the development of age-related diseases where their development rate seem to be mostly influenced by genetic and environmental factors (Franceschi *et al.*, 2018). The decline in cellular functionality in older individuals can be difficult to diagnose where an age related disease is not observed to be in development due to biological compensatory mechanisms, that buffer the detrimental changes up to a point of their failure this being the basis for the remodelling theory of ageing (Collino *et al.*, 2013).

1.3 Ageing as mechanism of further organismal growth

The cell hyperfunction theory is powered by the mechanisms of growth and reproduction such as the Insulin growth factor 1 and the target of the rapamycin kinase pathway that act in infancy and these are supposed to continue in adulthood, these driving a demand for growth that leads to oxidative stress and damage (Blagosklonny and Hall, 2009; Leontieva *et al.*, 2012). Rapamycin has been found to extend the lifespan in mice (Leontieva *et al.*, 2012; Bitto *et al.*, 2016; Swindell, 2017), however there are no conclusive results in humans yet. The theories around ageing also discuss a point where there was no evolutionary pressure to select for improved fitness in an aged individual (Kirkwood, 2005). Factors that could give certain advantages in younger individuals such as increased elimination of senescent cells or a more aggressive immune system could give rise to a better fit to survive young individual, however these characteristics would also present issues in an aged individual where these could give rise to autoimmune disorders or cancer development. Those genes that confer disadvantages later in life are not selected against because the fit younglings already had the chance to reproduce, and if these genes confer an advantage towards that they may have outcompeted most of the others that do not possessed this gene and its advantages.(Williams, 1957; Kowald *et al.*, 2020).

1.4 Hallmarks of ageing

Based on the some of these associated changes with ageing certain hallmarks were determined. These hallmarks are defined by the following criteria: the presence of the factors in normal ageing, the accentuation of the factors leads to accelerated ageing and that the repression of these factors decelerates ageing (López-Otín *et al.*, 2013). The hallmarks are characteristics that are present in mammalians, during normal ageing. The hallmarks of ageing: from epigenetic alterations to cellular senescence, these define the ageing phenotype as a decline in function with time with an increase towards disease development. (López-Otín *et al.*, 2013)

1.4.1 Primary hallmarks

A key contributor to ageing is the connection with the accumulation of DNA damage from both intrinsic and extrinsic factors. The sources for cellular damage are diverse, from voluntary factors such as tobacco smoking to diet, and from unavoidable factors such as stochastic effects as oxidative stress accumulation and pollution (Rattan, 2006; Soares *et al.*, 2014). The targets of this damage are diverse and cover all cellular components from cellular membranes, proteins and DNA. DNA damage affects many cellular functions such as: cell repair mechanisms, cell metabolism functions and genome stability (Lombard *et al.*, 2005). These functions when impaired have drastic consequences, especially if an essential gene is hit, an example being when DNA repair mechanisms are compromised this leading to accelerated ageing as in Werner syndrome (Chen and Oshima, 2002; Hoeijmakers, 2009; Gregg *et al.*, 2012). The accumulation of DNA damage is hypothesized to be a consequence of the decline in the cellular maintenance machinery of the DNA repair system (Maynard *et al.*, 2015) this leading to genomic instability with its characterized association with diseases that also show a higher association frequency with age such as cardiovascular diseases (Botto *et al.*, 2001; De Majo *et al.*, 2021), cancers (Negrini *et al.*, 2010) and dementias (Myung *et al.*, 2008).

Telomere shortening and damage are regarded as one of the main contributors for cellular senescence and their link with ageing (Jurk *et al.*, 2014). This issue with telomere shortening is intrinsic as the DNA replication machinery (e.g. DNA polymerase) is unable to fully replicate chromosomal DNA endings resulting in shortening at each cellular division (Harley *et al.*, 1990), this resulting in progressively shorter chromosomes until a point of critical length where cell cycle arrest is activated (d'Adda di Fagagna *et al.*, 2003). Still, such short telomeres retain sufficient length and protein telomere binding sites in order to avoid chromosomal fusions with the consequence of still signalling the critical length issue as DNA damage (Cesare and Karlseder, 2012). The activation of cell cycle arrest might also drive cells towards cellular apoptosis (d'Adda di Fagagna, 2008) or autophagy (Cheon *et al.*, 2019). Additionally, factors such as chronic inflammation are observed to contribute greatly towards accelerated telomere attrition (Jurk *et al.*, 2014). All these changes manifest similarly to cellular senescence, where an accompanied change in chromosomal architecture is seen together with the transcriptomic and proteomic changes (Gorgoulis *et al.*, 2019).

Further to the presented changes with age is the loss of proteostasis, a reduction in protein quality control mechanisms for protein homeostasis and protein chaperoning that maintains the normal functionality of cellular processes (Santra *et al.*, 2019). This mechanism shares similar causes to genomic instability where the proteins accumulate diverse forms of damage that further impact their ability in protein synthesis, degradation as well as their function in DNA repair mechanisms, gene expression and the many more other functions that proteins take part (Feser *et al.*, 2010; de Graff *et al.*, 2016). The loss of proteostasis is additionally marked by the accumulation of unfolded and misfolded proteins. In healthy situations the cellular machinery deals with these through the endoplasmic reticulum and mitochondrial unfolded proteins response mechanism (Hetz *et al.*, 2020; Metcalf *et al.*, 2020), these dealing and eliminating any abnormal proteins in order to ensure optimal cellular functionality. When these systems are overwhelmed, the cellular response is to trigger cellular apoptosis pathways resulting in the elimination of such compromised cell. The impact of this elimination is the maintenance of cellular fitness within the organism, however with ageing these mechanisms are less responsive and adaptive this possibly being the main cause for the observed

accumulation (Wodrich *et al.*, 2022). Unresponsive and defective cellular protein homeostasis mechanisms are linked to the development of cancers, this being an important hallmark towards the development of such conditions (Urra *et al.*, 2016), immunological hyperactivation that facilitates the activation of many proinflammatory pathways through NF-κB activation (Wu *et al.*, 2004; Schmitz *et al.*, 2018) and neurodegeneration in conditions such as Alzheimer (Uddin *et al.*, 2020) and Parkinson (Costa *et al.*, 2020) disease.

1.4.2 Antagonistic hallmarks

With ageing a decline in mitochondrial function and quality has been observed in relatively healthy conditions (RH, 2019). Evidence is being found how such changes can be linked with other changes as observed with ageing such as cellular senescence, inflammageing and stem cell depletion (Sun *et al.*, 2016). Accumulation of mitochondrial DNA damage is thought to be one key factor responsible for their decrease in quality (Rossignol *et al.*, 2003) and function similar to what is observed for nuclear genomic DNA, however since there are many more copies of mitochondrial DNA this effect is not always observed as severe (Rossignol *et al.*, 2003; Zhao and Sumberaz, 2020). The accumulation of such mitochondrial DNA damage has been modelled in mice (Kujoth *et al.*, 2005) and interestingly its main impact seems to be observed as increased oxidative stress with the accompanied increase in reactive oxygen species these resulting in an avalanche of damage across the cell (Hämäläinen *et al.*, 2015) with this also being linked with functionally handicapped stem cells and their exhaustion (Chen *et al.*, 2009). A decrease in mitochondrial function is linked with changes in the cellular metabolic profile with some evidence that these link with some changes towards cellular senescence (Habiballa *et al.*, 2019) and additionally the installation of cellular senescence introduces further mitochondrial metabolic changes with increased reactive oxygen species levels (Passos *et al.*, 2007). Together these changes are still poorly understood in regards to their connection and their link to overall ageing, however their presence and impact are thought to impact overall longevity (Sun *et al.*, 2016). The mitochondrial decline is also linked with a decline in protein

homeostasis where the mitochondrial unfolded protein response mechanism is impacted (Lima *et al.*, 2022). This mechanism being a key pathway of adapting mitochondrial function to stress such as increased load of damaged, misfolded or unfolded proteins that would be functionally inferior (Zhao *et al.*, 2002; Gu *et al.*, 2021). The consequence of such accumulation of abnormal proteins leads to the decline of stem cell function (e.g. haematopoietic stem cells) with a reduction in their regenerative capacity and viability (Mohrin *et al.*, 2015).

Cellular senescence is marked by cellular growth arrest with the development of distinct morphological, metabolic and chromatin architectural changes and is accompanied by the secretion of the senescence- associated secretory phenotype (SASP) factors (Salama *et al.*, 2014; McHugh and Gil, 2018). For young healthy individuals senescence is a powerful tool for tumorigenesis suppression (Stoczynska-Fidelus *et al.*, 2022), however with older individuals there is an accumulation of senescent cells across body this being also declared a marker of ageing (van Deursen, 2014). Cellular senescence can be induced by a multitude of factors from oxidative stress, genomic damage, activation of oncogenes and telomere shortening (d'Adda di Fagagna *et al.*, 2003; Rodier *et al.*, 2009; Correia-Melo and Passos, 2015) with the DNA damage response being a central initiation pathway for this. The accumulation of senescent cells in a multitude of tissues has been observed with ageing (Hewitt *et al.*, 2012; Biran *et al.*, 2017) and these have been thought to also be contributors towards the installation of chronic inflammation and its link with inflammageing (Jurk *et al.*, 2014). The elimination of senescent cells using mouse models has been observed to contribute towards increased healthspan by increasing metabolic functionality and to delay the onset of some age-associated diseases (Baker *et al.*, 2011; Xu *et al.*, 2015; Hernandez-Gonzalez *et al.*, 2021). Unfortunately, there are a number of disadvantages to eliminating cellular senescence. The elimination of senescent cells first would require immediate cellular replacement where these cells provide structural integrity to tissues and their replacement through stem or progenitor cells, alternatively there are suppressive drugs that target SASP and therefore help to mitigate the main insult of senescence (Acosta *et al.*, 2013; Habiballa *et al.*, 2019).

1.4.3 Integrative hallmarks

Another marker for ageing also linked with senescence is that multipotent stem cells population depletion occurs during ageing with the consequence of this being declined stem cell functionality, the decreased tissue repair capability leading to tissue deterioration (McHugh and Gil, 2018). Such a decline is being observed in a multitude of tissues from the hematopoietic stem cells to even those with a relatively low cellular turn-over. Hematopoietic stem cells have age-associated decreased functionality that leads to the decreased functionality and number of naive B, T cells as well as natural killer cells (Kollman *et al.*, 2001; Geiger and Van Zant, 2002). In tissues such as the nervous system where low cellular proliferation and low turn-over is observed, the neural stem cells show increasingly limited neurogenesis capacity, decreased proliferation and increase in senescent markers (Molofsky *et al.*, 2006).

Ageing manifests also through the increase in ectopic intracellular communication pathways, such as those associated with the low-level of inflammation termed “inflammageing”. This condition leads to serious chronic complications that is thought the be one of the most important risk factor for disease development risk (Franceschi and Campisi, 2014). This condition of low grade inflammation is brought up by the increased presence of the inflammatory factors IL-1 and IL-6, these acting as prognostic markers for the development of many diseases including diabetes of both types (Tsalamandris *et al.*, 2019) and cardiovascular disease (Alfaddagh *et al.*, 2020). Inflammageing is also seen as a link between the altered intracellular communication and cellular senescence (Coppé *et al.*, 2010).

Further to the impact of these changes there is the economic burden associated with managing and the treatment of these ailments and with the report of the increasing ageing population such figures are only predicted to increase in cost in the future (Yang *et al.*, 2021b). A solution to mitigate this impact is through improvements in diagnostic technologies as well as recommending earlier treatment of early form age-related diseases reducing the impact these will have later on in life. It is estimated that through the introduction of such measures the

economic gain for countries like the USA will be around $38 trillion per year of life expectancy extension (Scott *et al.*, 2021).

1.5 Epigenetics alterations are a primary hallmark of ageing

Epigenetics refers to heritable changes in gene expression and/or phenotype that occur without changes to the underlying DNA sequence (Hamilton, 2011). The field of epigenetics encompasses chromatin interaction mechanisms (i.e. euchromatin-heterochromatin boundary), DNA methylation, histone post-translational modifications and non-coding RNAs including microRNAs (Kim *et al.*, 2011) (Figure 1. 1). Epigenetic patterns are dynamic during development and cell differentiation, yet these can also be stable allowing cell identify to be maintained (Bird, 2007). Epigenetic patterns are mitotically and meiotically inherited but can undergo continuous modulation in response to both the internal and external environment, including diet, physical activity and smoking. (Wu and Morris, 2001; Dupont *et al.*, 2009; Breton *et al.*, 2021).

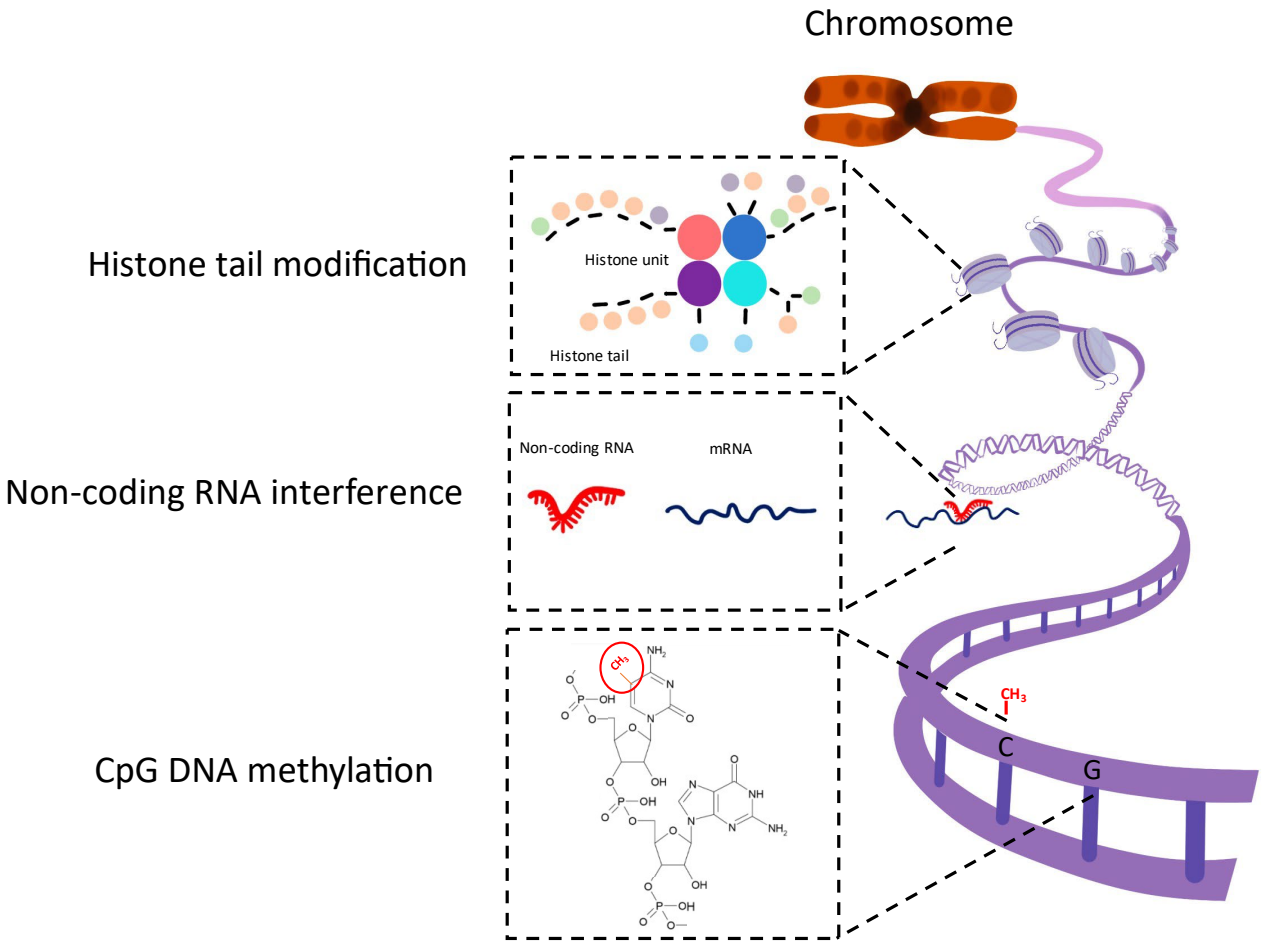


Figure 1. 1. Diagram of the mechanism of epigenetic regulation. Histone tail modification, non-coding RNA interference and CpG DNA methylation changes are three recognised mechanisms of epigenetic regulation.

The epigenome shows a progressive loss in its signature and stability with ageing, this being marked by fundamental changes in chromosomal architecture, gene expression pattern and genome stability (Brunet and Berger, 2014; Saul and Kosinsky, 2021). Ageing is also closely linked with heterochromatin shifts and decays where endogenous genetical elements like retrotransposons begin to show activation and expression leading to transposition. In young

healthy individuals these genetic elements are silenced however with ageing the silencing mechanisms disappears with the changes in chromatin accessibility leading to their activation. This results in random insertions of these sequences in the genome leading to further genomic instability, abnormal gene expression and abnormal protein structures (Pal and Tyler, 2016; Gorbunova *et al.*, 2021). Several studies have observed a decrease in histone methylation associated with ageing this being found for multiple histone marks such as H3K9me3 (Djeghloul *et al.*, 2016) ,H3K27me3(Bracken *et al.*, 2007; Shah *et al.*, 2013) H3K9me3 is most abundant at gene promoters and serves as marker for such regulatory regions where its absence could lead to gene silencing or altered transcription (Cano-Rodriguez *et al.*, 2016). H3K27me3 is a chromatin mark strongly linked with the heterochromatin signature and the age associated decrease in this mark supports the observed heterochromatin decay signatures across the genome (Yi and Kim, 2020).

During ageing there is observable change in the nuclear architecture of the chromatin where a reorganization between euchromatin and heterochromatin correlates with age (Adams, 2007) giving rise to the "heterochromatin loss model of ageing" (Tsurumi and Li, 2012). Cell models of ageing present enlarged nuclei due to increased nuclear decondensation with further consequences of disrupted transcription (Zhang *et al.*, 2015). In senescent cells the nuclear structural changes observed are those of increase in heterochromatin with the formation of senescence associated heterochromatin foci at genes associated with cellular proliferation (Narita *et al.*, 2003),although there is also an overall global loss of heterochromatin followed by a loss of gene silencing (Tsurumi and Li, 2012; Lee *et al.*, 2020).

1.6 DNA methylation changes as a hallmark of ageing.

1.6.1 Regulation of gene expression DNA methylation

In mammals, the majority of DNA methylation involves the addition of a methyl group on the 5' carbon of a cytidine base usually preceding a guanine base, forming a CpG unit. The mechanism is accomplished by the DNA methyl transferase (DNMT) family of proteins, which use the methyl group donor S-adenosyl L methionine as a substrate (Lyko, 2018). The enzymes are essential in development (e.g. global K.O. is embryonically lethal (Li *et al.*, 1992)), cell differentiation and in normal cell function (e.g. genome stability) (Robertson, 2005; Santos *et*

al., 2005). DNMT1 is the enzyme that has a maintenance responsibility where it imprints the existing DNA methylation pattern onto the newly synthesised DNA strand during mitosis and DNMT3a and DNMT3b are *de novo* imprinters that methylate new CpG sites in response to various factors. Global knock out mice for DNMT1 are embryonically lethal (Li *et al.*, 1992) and therefore only conditional knock outs can be used to study the temporal absence of this enzyme (Jorgensen *et al.*, 2018). Knock out of DNMT1 in human embryonic stem cells results in rapid cell death (Liao *et al.*, 2015; Pathania *et al.*, 2015)

The compound 5-Aza 2'deoxycytidine, an analog of cytosine, is used in the treatment of acute myeloid leukemia and certain types of breast cancer (Momparler, 2005; Mirza *et al.*, 2010) and recently it has been used to improve visual function in aged mice (Chen *et al.*, 2020). The mechanism through which it acts is by blocking CpG methylation through the binding to DNA methyl transferase 1, also marking this for ubiquitination (Patel *et al.*, 2010). The result is believed to be a passive global loss of DNA methylation at CpG sites, loss increasing with each cell replication (Ramos *et al.*, 2015), however it still has efficacy in non-dividing cells such as neurons (Yang *et al.*, 2017). The treatment also causes transcriptomic changes where genes are observed to either have increased expression or decreased (Yang *et al.*, 2012), a mechanism possibly dependent on the availability of cell type transcription factors. The compound is used to investigate whether certain genes of interest respond to this change in DNA methylation and what impact this has on their expression.

1.6.2 Effect of DNA methylation on transcription factor binding.

Transcription factors are fundamental biological regulators of gene transcription being involved heavily in the initiation of the process at the right time and location. The binding of transcription factors occurs at gene promoters and enhancers together forming complexes of multi-protein factors forming the transcription machinery (Grove and Walhout, 2008). Transcription factor binding is determined and influenced by the underlying DNA sequence, the DNA methylation status of the underlying sequence and the competition for the site from other transcription factors. These together influencing the affinity of the factors that bind certain DNA regions (Islam *et al.*, 2021). Transcription factors can be divided into four categories, these being activators, coactivators, chromatin remodellers and repressors (Li *et al.*, 2015; Reiter *et*

al., 2017). Transcription activators directly impact the transcription of a particular gene through their binding of DNA at the beginning of the transcription process, these factors can also serve as a base for the formation of the transcription complex that advances the transcription operation (Spitz and Furlong, 2012; Lambert *et al.*, 2018). Transcription coactivators are factors without a DNA binding domain and therefore rely on other factors to bind the DNA sequence first in order to perform their function however, coactivators participate in the formation and stabilization of the bigger transcription complex machinery (Frietze and Farnham, 2011; Reiter *et al.*, 2017). Chromatin remodelling transcription factors influence the accessibility of chromatin through their interaction with the DNA sequence and nucleosome (i.e. interaction with histone marks) resulting in distinct chromatin conformations that influences the binding of the other types of transcription factors (Rippe *et al.*, 2007; Li *et al.*, 2015). Transcription repressors on the other hand have a negative impact with transcription, where their binding can stop the binding of activators or can stop the formation of the transcription complex and therefore reduce gene transcription through mechanisms such a steric hindrance (Lee and Young, 2013; Domcke *et al.*, 2015). Another property of transcription factors is that they can be either tissue/cell type specific tending to drive differentiation programs (e.g. MyoD in myogenic differentiation) (Ishibashi *et al.*, 2005; Shintaku *et al.*, 2016) and others are ubiquitously expressed across a range of cell types (e.g. SP1) (O'Connor *et al.*, 2016) this mechanism greatly influencing the diversity and complexity of the overall transcription machinery and its action. The system permits the change of the proportion or population of transcription factors in response to the changes in DNA methylation which then can alter the expression level of genes and therefore of the cellular signalling (Moore *et al.*, 2013; Héberlé and Bardet, 2019a).

1.6.3 Effect of DNA methylation of gene promoters

A gene promoter is a specialized region present at both coding and non-coding genes serving as a starting point for transcription machinery assembly and initiation. This region enables a high degree of control over the ability of gene transcription through the integration and support of regulatory machinery through distal enhancers, histone marks and the associated regulatory transcription factors (Haberle and Stark, 2018). The classical initiation of transcription is starting at a transcription start site (TSS) that is integrated in the gene's promoter region, this

serves as binding platform for the transcription machinery composed of RNA polymerase II and the associated transcription factors (Hampsey, 1998). Although promoters have sufficient capabilities to drive gene transcription this is only observed to be at a very low level this being also influenced by factors such as chromatin accessibility, histone marks and DNA methylation levels (Kadonaga, 2012; Moore *et al.*, 2012; Fan *et al.*, 2021; Hsu *et al.*, 2021) and the associated gene distal enhancers (Ishibashi and Taguchi, 2021).

Promoter hypermethylation has been declared as being a key mechanism in the development of certain cancers where aberrant methylation of tumour suppressor gene promoters results in gene silencing (Ma *et al.*, 2014; Molnár *et al.*, 2018; Dhar *et al.*, 2021). Interestingly, the correlation between promoter DNA methylation levels and gene expression has been linked with tissue specificity with a main target impact on genes for transcription factors (Moarii *et al.*, 2015). And additionally any transcriptional changes were also correlated with DNA methylation based on the type of CpG island density located in the promoters (Messerschmidt *et al.*, 2014).

1.6.4 Effect of DNA methylation on transcriptional enhancers

Enhancers are usually short regulatory regions of DNA that can be situated almost anywhere in the genome, from inside introns of genes to gene deserts (i.e. regions of DNA lacking protein-coding regions). These regulatory regions activate or enhance expression of their target gene(s) from a distance (Bulger and Groudine, 2011).

Their presence amplifies a target gene's transcript level and allows fine tuning of this transcript in different tissues (Guerrero *et al.*, 2010; Kolovos *et al.*, 2012; Schaffner, 2015), more so that can be achieved by a promoter alone (Magnusson *et al.*, 2015). The mechanism through which they achieved this is the recruitment of transcription factors at its site and of RNA polymerase II (Koch *et al.*, 2011). This recruitment causes a DNA loop to form (see figure below) that may increase the concentration of transcription machinery at the location of the enhancer and its target causing an increase in transcription levels (Pennacchio *et al.*, 2013). The RNA polymerase also transcribes a part of the enhancer to produce non-coding enhancerRNA (eRNA), which is thought to be a key component for an enhancer's activity, this eRNA when present increases the mRNA synthesis of the target gene by several fold (Orom *et al.*, 2010; Pennacchio *et al.*,

2013). Enhancers are identified conventionally by their open chromatin state (a sensitive region for DNAse I) (He *et al.*, 2010), characteristic histone tail modifications (e.g. H3K27ac, H3K4me1, H3k4me2) (Zentner *et al.*, 2011) as well as the presence of bound RNA polymerase II (De Santa *et al.*, 2010).

An enhancer's activity may be required to be only active during a certain stage in development in a certain region in the body (Sakabe *et al.*, 2012). Such a specific cell type, temporal and spatial control is achieved through a tight feedback system (Calo and Wysocka, 2013). In this system, enhancers that are no longer needed are shut down (i.e. are poised) by either acquiring DNA methylation at CpG sites by *de novo* methylation by DNMT3a or DNMT3b and new specific histone markers (e.g. H3K4me3).This changes the chromatin structure to a closed conformation, making it inaccessible to the transcription machinery and therefore it results in transcriptional repression (Sakabe *et al.*, 2012; Jang *et al.*, 2017) . In contrast, when some enhancers are methylated this induces acetylation at the nearby histone at the H3K27 position increasing the gene expression of the target gene by recruiting additional transcription factors that can present a methyl-CpG binding domain (Charlet *et al.*, 2016; Zhu *et al.*, 2016).

Other locations that have been observed to be impacted by DNA methylation are a gene's first intron where several inverse proportional links have been established between the occurrence of DNA methylation at the gene's first intron and its impact of transcription. Such an observation is also linked with the findings that these regions are usually populated by gene enhancers (Anastasiadi *et al.*, 2018; Dhar *et al.*, 2021). Another is the gene body where DNA methylation changes were observed to correlate with an increase in expression (Yang *et al.*, 2014) (Molnár *et al.*, 2018)

1.6.5 CpG Methylation changes and ageing

Age-linked tissue specific CpG methylations are hypothesised to predispose an individual towards the development of certain disease (Bjornsson *et al.*, 2004; van Otterdijk *et al.*, 2013) possibly by decreasing the expression of anabolic genes. Anabolic genes control the synthesis of components from the extracellular matrix (e.g. collagen synthesis), promote muscle growth (*MyoD*) these promoting tissue growth (Aigner *et al.*, 2001).

A study in monozygotic twins revealed that young twins had a very similar methylome, while in older twins a more drastic difference in the methylome pattern was observed. This study provides a good model to exemplify the influence of the environment even on genetically identical beings (Fraga *et al.*, 2005).

In the decade more attention been focused on the implications of epigenetic mechanisms during ageing and how these can influence the aging process (Sierra *et al.*, 2015; Pal and Tyler, 2016). Interestingly, a phenomenon observed during ageing is that the global methylation level of CpG sites decreases with age (Vanyushin *et al.*, 1973; Fraga *et al.*, 2007; Jintaridth and Mutirangura, 2010; Heyn *et al.*, 2012; Zampieri *et al.*, 2015) while the methylation level of certain CpG sites seem to increase with age showing tissue specificity (Day *et al.*, 2013; Hannum *et al.*, 2013). This lead to the hypothesis that ageing results in global CpG hypomethylation which could impact on gene expression and could permit ectopic expression of silenced genes in cells and tissues (Zs-Nagy *et al.*, 1988). No reports of any global hypomethylation changes associated with ageing have been made (Lister *et al.*, 2013; Raddatz *et al.*, 2013), however an enrichment of hypermethylation or hypomethylation at certain CpG sites is a characteristic observed across multiple tissues with the ageing process (Raddatz *et al.*, 2013; Slieker *et al.*, 2016; Slieker *et al.*, 2018). The CpG sites that change their methylation levels with age are called age-CpG sites, these being the focus of many studies that have found different diseases exhibiting a methylation pattern in only one tissue or many different tissues, such as blood, kidney, lungs etc., suggesting that these changes can occur across multiple tissues linking them in the disease pathogenesis. Such studies identified a great deal of age-CpGs however the functional link between ageing still presents a fundamental point of research in order to gain better understanding of the phenomenon.

As DNA methylation of promoter and enhancer regions can be correlated with reduced gene expression, the implications of this are of high importance towards disease susceptibility and progression (Alvarez-Garcia *et al.*, 2016b). If an enhancer is *de novo* methylated by the cellular machinery as a consequence of the advancement in age, this could change the functionality of tissues causing the observed increased disease predisposition observed in the elderly (Divo *et al.*, 2014).

1.6.6 The development of DNA methylation based epigenetic clocks

As the technology started to permit the study of thousands of CpGs in multiple samples and as the accuracy of these improved, a number of investigators observed a correlation between the methylation level of a CpG and the chronological age of a person (Day *et al.*, 2013; Hannum *et al.*, 2013; Horvath, 2013). Based on the data sets generated from methylation array datasheets, an epigenetic age predictor was established (Bocklandt *et al.*, 2011; Hannum *et al.*, 2013; Horvath, 2013). This can be used to investigate the epigenetic age of a tissue to predict whether there is a risk for an individual for epigenetically-linked disease development (Teschendorff *et al.*, 2010; Bocklandt *et al.*, 2011). Studying the patterns of DNA methylation from embryonal development to adulthood and beyond it was possible to generate a tool called the DNA methylation age/epigenetic clock, a clock formed by the methylation levels of 353 CpG sites that occur across different tissues (Horvath, 2013). The DNA methylation age in Horvath's epigenetic clock is the measure of the epigenetic maintenance system, where the cells from the embryonic stage show natively an age of ~0 with a logarithmic increase of this age in comparison to chronological age (Bocklandt *et al.*, 2011; Horvath, 2013) (Figure 1. 2). Another such clock has been devised based on whole blood samples of 656 individuals between 19-101 years of age. Even though the clock has been trained on blood samples, it seems that it can be used with some degree of success in other tissues as well (Hannum *et al.*, 2013).

Recently an analysis of epigenetic age changes observed from childhood to old age using the two epigenetic clock systems (Hannum *et al.*, 2013; Horvath, 2013) was performed. This study used whole blood sample as a longitudinal study, and it revealed that the Hannum epigenetic clock performed better than the Horvath clock. This is mainly due to the fact that the Hannum clock was trained on blood samples only, whereas the Horvath clock is a multi-tissue trained clock and it shows only very few CpGs in common (Marioni *et al.*, 2019). Interestingly, the study revealed the epigenetic clock seems to slow down in the older population (Marioni *et al.*, 2019), which is contrary to what it was suggested previously (Benayoun *et al.*, 2015).

These epigenetic clocks also permit the study of the impact of diseases or ageing on the epigenetic system of a particular tissue or cell type, and has been used to observe that in certain diseases such as obesity and Werner syndrome there is an accelerated epigenetic age of tissues of blood and liver respectively (Horvath *et al.*, 2014; Maierhofer *et al.*, 2017).

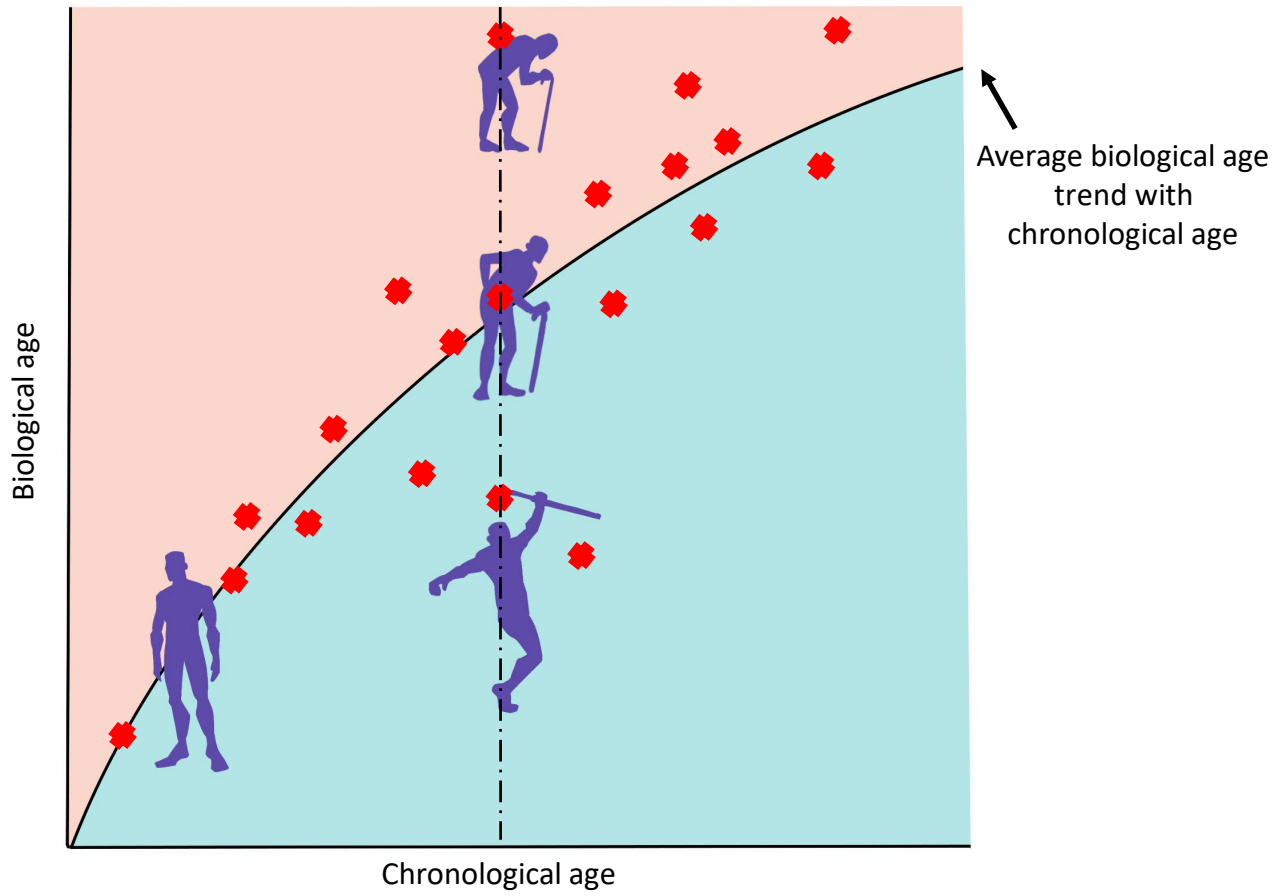


Figure 1. 2. Diagram representation of the impact of the relationship between biological age and chronological age and how this can impact an individual's healthspan.

Gene expression link with epigenetics and changes with age

The dynamic nature of gene expression has long been observed (Alberts et al., 2002; Ramoni et al., 2002) together with the different patterns of gene expression of different tissues to give rise

to the different phenotype. Although this has been studied extensively during embryogenesis (Kocabas et al., 2006; Li et al., 2006; Yi et al., 2010) only more recently the study has been taken up to study the impact of chronological age on the gene expression patterns in different tissues (Glass et al., 2013; Harris et al., 2017). The skin, adipose tissue and blood of 856 of female twins with ages ranging from 39 to 85 was studied for changes in gene expression. The findings were that the skin tissue shows the most dramatic changes across the age interval although all the tissues showed changes. Genes implicated in fatty acid metabolism, mitochondrial activity and splicing were changing their expression levels and most interesting the expression changes showed tissue specificity (Glass et al., 2013).

Having more un-biased approaches to study gene expression changes, with the use of whole transcriptome arrays, provides a more valuable and profound information about the inter-connection of tissues and their changes with age. With these tools and samples from approx. 200 individuals of nine paired tissues (heart, lung, adipose, muscle, nerve, skin, thyroid, blood, artery) it was again confirmed how tissues show unique changes in their transcriptome and more interestingly, that there seems to be a sort of synchronization between tissues. These synchronous changes, termed co-ageing of tissues, describe the similar gene expression changes with age observed in two or more tissues. Moreover, there was an up-regulation of genes linked with immune-mediated inflammation and inflammageing in tissues such as arteries and blood. This observation could provide insight in the mechanism of ageing as a factor towards disease vulnerability (Yang et al., 2015).

In a twin study of 336 monozygotic twins (Viñuela et al., 2019), 137 genes were identified to change with age which were also previously identified in other studies (Glass et al., 2013; Yang et al., 2015), however 42 genes showed inconsistency between the twins implicating environmental factors. These changes were believed to be the consequence of age-methylation changes and they were present in genes previously associated with different diseases (IRS1 and type 2 diabetes) (Viñuela et al., 2019)

1.7 Osteoarthritis is an age-related disease of the synovial joints

1.7.1 Osteoarthritis

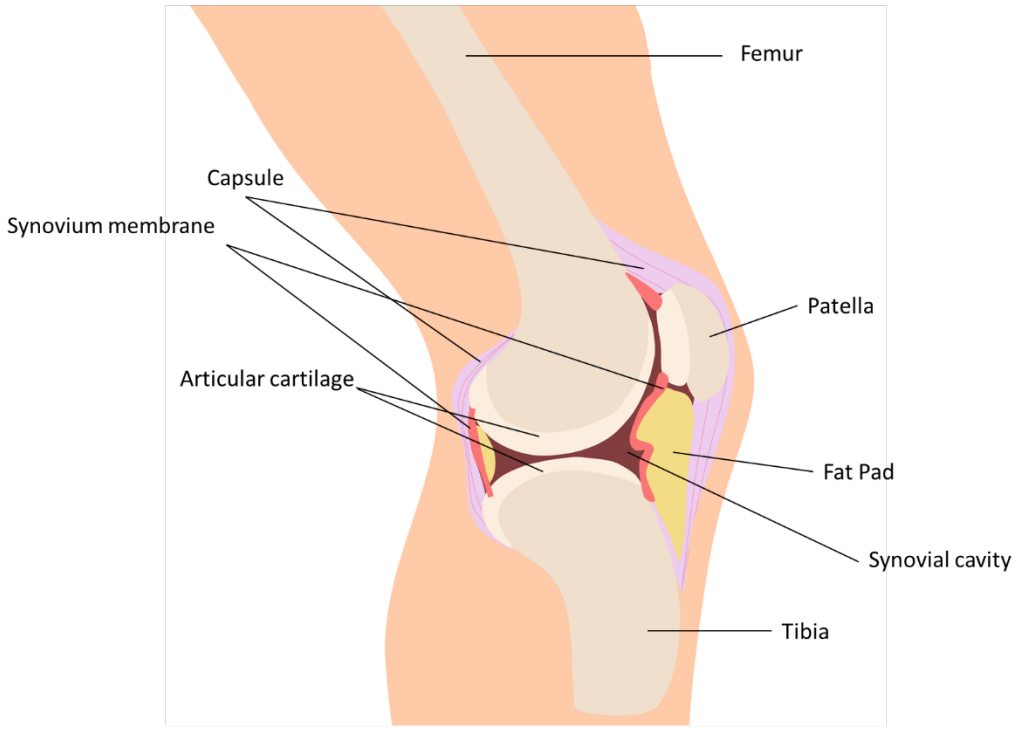
Osteoarthritis (OA) is the most common degenerating disease of the synovial joint affecting multiple joints such as the knees, hips, hands, foot, neck and spine. OA of the knee is the most frequently diagnosed form of OA, with somewhat lower incidence for hip, hand and vertebral column (van Saase *et al.*, 1989; Helmick *et al.*, 2008; Lawrence *et al.*, 2008; Fernandes *et al.*, 2017). It is marked by the loss of the articular cartilage and the inflammation of the synovial membrane which leads to joint rigidity, joint movement limitation, and a painful condition where individuals become immobilized, creating disability (Musumeci *et al.*, 2015a). It affects >8.75 million people in the UK (UK, 2018) with over 17 million across the globe in 2010, this being ranked the 11th highest contributor to global disability (Cross *et al.*, 2014). OA is a growing problem to the health authorities across the entire planet where it is estimated to affect more than 50% of the individuals aged over 65 (Musumeci *et al.*, 2015b). OA also presents itself as an economic burden more so than other musculoskeletal diseases (e.g. rheumatoid arthritis) (Bitton, 2009; Cross *et al.*, 2014) with an estimate indirect cost of around £3.2 billion /annum for the UK government that come from the hospitalization, treatment (e.g. NSAIDs), surgery and disability of the patients (Chen *et al.*, 2012).

OA has traditionally been considered a disease of the articular cartilage(Martel-Pelletier and Pelletier, 2010), a smooth, elastic, avascular and aneural tissue that covers the ends of the bone in the joint. It is specialised to support and distribute the load across the articular joint by providing a low friction for smooth joint movement (Pearle *et al.*, 2005). Due to this tissue having only one type of resident cells, the chondrocytes that synthesize and repair the cartilage matrix this made them an attractive start point to investigate their role in the pathogenesis of OA (Archer and Francis-West, 2003; Akkiraju and Nohe, 2015). This view was challenged by new studies that showed the implications of the other tissues in the joint and now OA is considered a whole joint disease (Ashraf and Walsh, 2008; Das and Farooqi, 2008; Egloff *et al.*, 2012; Loeser *et al.*, 2012). It presents itself as a disease where both the genetic makeup of an individual (e.g. inheritance of risk genes) and environmental factors (e.g. injury, body weight) show a high implication in the disease's development and progression (Yucesoy *et al.*, 2015). As

a whole joint disease, it means that researchers need to also investigate the other tissues in the articular joint (e.g. synovium membrane, fat pad, ligaments, subchondral bone) that might be modified during or before OA development (Man and Mologhianu, 2014) (Figure 1. 3). Some of the tissues that are affected in OA are discussed below.

A)

Anatomy of the knee joint schematic



B)

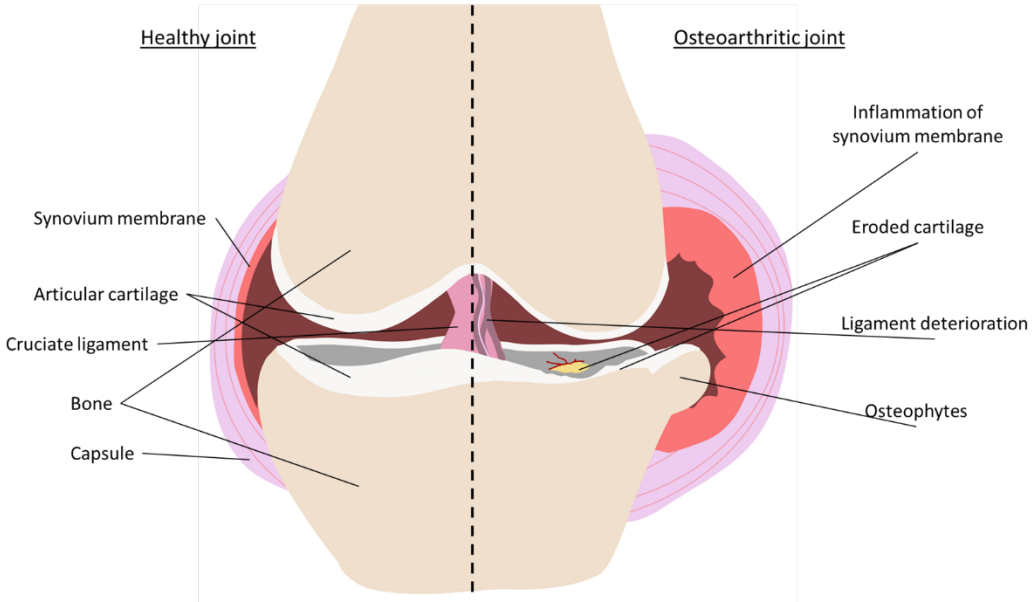


Figure 1. 3. Diagram representation of the tissues present in the knee articular joint (A) and their change with osteoarthritis development (B).

1.7.2 Cartilage

The cartilage is composed out of an extracellular matrix of collagen (mainly type II), aggrecan, hyaluronan and other proteoglycan, and a cellular part with chondrocytes being the only resident cells. The chondrocytes are the cells that synthesize the articular cartilage matrix, they are also the ones that repair the tissue but also, they are responsible for degrading the cartilage (figure 3) matrix through the action of matrix degrading enzymes (e.g. collagenases) (Archer and Francis-West, 2003; Sophia Fox *et al.*, 2009). An understanding that came with the study of these cells was that in healthy individuals there is a balance between the destruction of the cartilage and its repair, however in OA cartilage the balance is shifted towards degradation (Li *et al.*, 2013b). The activity of chondrocytes can be modified by the presence of pro-inflammatory cytokines (e.g. IL-1, IL-6) and growth factors (e.g. Transforming growth factor- β) that impact the anabolic and catabolic pathways (Stabellini *et al.*, 2003; Goldring *et al.*, 2008; Fortier *et al.*, 2011; Kapoor *et al.*, 2011). The source for these mediators in a normal healthy joint was thought only to appear after trauma, as the damage to the articular cartilage stimulates the chondrocytes to produce reactive oxygen species, upregulate their anabolic activity and pro-inflammatory cytokine production, which also brought an upregulation of cartilage degrading enzymes such as matrix metalloproteinases (MMPs) (e.g. MMP-13) (Ding *et al.*, 2010; Goodwin *et al.*, 2010; Goldring and Otero, 2011). These results propose the formation of a vicious cycle that ends with the total destruction of the cartilage from the affected articular joint.

1.7.3 Synovium

The synovium is a specialized connective tissue membrane with the role of sealing the synovial cavity from the surrounding tissue. The synoviocytes are the main cell type in this tissue which are responsible for the maintenance of the synovial fluid, a viscous fluid in the articular joint composed mainly of lubricin and hyaluronic acid. The synovial fluid serves as lubrication for the

cartilage to prevent friction at the articular surface and as nutrient reserve for chondrocytes. Because the cartilage is an avascular tissue, chondrocytes have to rely on the diffusion of nutrients from the synovial fluid to maintain their normal function (Sellam and Berenbaum, 2010; Scanzello and Goldring, 2012). In OA patients, the synovial membrane becomes inflamed (i.e. synovitis), and is infiltrated with mononuclear cells, being associated with both early stage OA patients and with late stage OA patients (Myers *et al.*, 1990; Oehler *et al.*, 2002; Benito *et al.*, 2005). Synoviocytes produces MMPs as well (Yuan *et al.*, 2004) together with pro-inflammatory cytokines and adipokines (Lee *et al.*, 2009; Sellam and Berenbaum, 2010), these being responsible for inducing the increased neo-vascular invasion of the synovial membrane (Ashraf and Walsh, 2008; Wenham and Conaghan, 2010) which contributes towards the OA pathogenesis (Scanzello and Goldring, 2012; Mathiessen and Conaghan, 2017). Other factors such as obesity further aggravate the synovitis and have a higher associated cartilage degeneration leading to a more severe form of OA (Kanthawang *et al.*, 2021)

1.7.4 Ligament

The ligament is a strong and elastic tissue that connects the muscles to the bone permitting the transmission of the muscle's strength to the bone, allowing joint movement. The main resident cells in this tissue are the tenocytes which respond to the load of the tissue to better adapt the tendon for supporting load (Andarawis-Puri *et al.*, 2015).

The tendon tissue undergoes changes caused by different stimuli such as physical exercise or certain diseases, changes that can influence its capacity to function normally (Bordoni and Varacallo, 2018). During the normal function of the tendon it is continuously modelled to make it better adapted to support load, this action is made possible by the secretion of MMPs from tenocytes with subsequent repair by the same cells (Magnusson *et al.*, 2010). Injury to the collateral ligaments of the knee joint has been found to be a great risk factor for OA development, this being acknowledged by several studies (Fleming *et al.*, 2005; Blalock *et al.*, 2015). This is also being confirmed through animal models where through ligament damage the

joint becomes unstable increasing the risk of abnormal bear on the articular joint with this possibly damaging it (Lorenz and Grassel, 2014; Blalock *et al.*, 2015).

1.7.5 Fat Pad

The fat pad is an intracapsular adipose tissue located under the patella in the knee joint (Saddik *et al.*, 2004; Ioan-Facsinay and Kloppenburg, 2013) with the possible function in reducing joint mechanical overloading and shock absorbance (Han *et al.*, 2014). Recently attention has been given to the fat pad tissue in knee OA, as potentially a site that can release pro-inflammatory signals in the articular joint, contributing to the OA phenotype (Distel *et al.*, 2009; Ioan-Facsinay and Kloppenburg, 2013). The role of the fat pad in the knee has been investigated and it was found that the bigger the surface area of the intra patellar fat pad the more protective against OA it seemed to be, possibly due to an increased shock absorption capacity of the tissue (Han *et al.*, 2014). The fat pad during OA pathogenesis suffers from increased vascularization, inflammation and fibrosis compared to non-OA patients (Favero *et al.*, 2017) it also becomes a site where IL-1 and IL-6 are released, key contributors to the inflammatory phenotype (Distel *et al.*, 2009; Chuckpaiwong *et al.*, 2010).

1.7.6 Bone

The bone tissue is special type of connective tissue, being strongly mineralized a property which confers its shock absorption capacity as well as inelasticity. Four types of cells are resident in this tissue, these being: osteoclasts, osteoblasts, osteocytes and bone lining cells (Buckwalter *et al.*, 1996). The tissue is under continuous remodelling by its resident cells working in concert to ensure homeostasis and repairs resulting in a strong material that can absorb shock from movement (Florencio-Silva *et al.*, 2015; ElSayed and Varacallo, 2018).

The current view based on recent evidence suggests that the increase in thickness of the subchondral bone found in OA patients is thought to be due to the increase in the bone turnover and the activation of the secondary point of ossification (Burr and Radin, 2003; Muir *et al.*, 2006). The main reasons hypothesized to cause this are linked with the incapacity of the

articular cartilage to absorb impact shocks, the subchondral bone suffering micro lesions. These micro lesions initiate the reactivation of the secondary point of ossification which causes the bone remodelling mechanisms to make an advancement of the calcified tissue towards the cartilage zone. This all leads to the slow replacement of the articular cartilage from the subchondral bone front. The cross talk between the cartilage and bone is also implicated in this vicious mechanism of degradation as the chondrocytes in the deep zone of the cartilage secrete mediators that stimulate osteoclast activation, which is thought to initiate the bone remodelling (Bellido *et al.*, 2010; Goldring, 2012; Henrotin *et al.*, 2012). Another OA associated change is the microstructural thickening of the trabecular bone which could lead to a higher tensile stress of the cartilage-bone juncture, this participating further in the deterioration of both tissues involved (Li *et al.*, 2013a); additionally in obese patients the prevalence of type I collagen homotrimers leads to further architectural changes of the bone leading to more severe phenotypes (Philp *et al.*, 2017).

1.7.7 Risk factors and treatments for OA development

The main risk factors for OA are age, sex, obesity, prior joint injury and genetics (Heidari, 2011; Reynard and Loughlin, 2012). Females show a higher predisposition towards OA development with 13% of females compared to ~10% of males in adults over 60 years of age (Zhang and Jordan, 2010). Prior joint injury as in the case of athletes has been shown to be highly correlated with a high risk of OA development (Buckwalter, 2003; Brown *et al.*, 2006). Over 90 independent OA associated genetic risk loci have been discovered that increase the susceptibility of OA development. Several of these loci increase risk of OA development at a specific joint site or in a specific sex, and for the majority of these loci, the target gene and mechanism through which genetic risk is acting remains unknown. Ageing is described to be the highest risk factor for OA development, however it is not as simple to interpret OA as a consequence of articular joint ageing but rather as ageing being a predisposition factor for OA through its mechanisms (Shane Anderson and Loeser, 2010).

Obesity as a risk factor

https://www.nature.com/articles/s41598-020-60587-1

OA is currently managed with a treatment of its symptoms (e.g. pain) and not its cause. Use of NSAIDs has been shown to help with early OA management however as the disease continuous to progress patients can reach a stage where these become ineffective (Cho *et al.*, 2015). Therefore the use of stronger painkillers (e.g. opioids) proceeds this which can have strong adverse effects (e.g. addiction) (Bedson *et al.*, 2016) or sometimes a controversial procedure is performed, total joint replacement(Edwards *et al.*, 2018). Several other treatments have been proposed (e.g. moderate and supervised physical exercise) which are used in the clinical environment some with more success than others however no big advancement has been made that can be applied to all patients (Dziedzic and Allen, 2018).

1.7.8 Cartilage DNA methylation differ between people with and without OA.

Over the last years, the role of DNA methylation changes between osteoarthritic and non-osteoarthritic cartilage has been studied at both the gene-specific and genome-wide level (Reynard, 2017). Several groups have used the Illumina Human Methylation Bead Chip 450K array to assesses methylation in cartilage at ~485,000 CpG sites across the genome (Rushton *et al.*, 2014; Aref-Eshghi *et al.*, 2015). These studies have identified thousands of CpG sites that are differentially methylated between OA-affected and unaffected cartilage, the majority located in putative chondrocyte enhancer regions.

1.7.9 Cartilage investigation for age-related methylation changes

Prior to the start of this PhD project, a meta-analysis of 179 osteoarthritic (OA) and non-OA cartilage samples from patients ranging from 50 to 95 years was performed to identify age-related DNA methylation changes in cartilage. A total of 716 CpG sites were identified that showed a significant correlation between the methylation levels and chronological age (termed age-CpGs), with over 90% of them showing hypermethylation with ageing, whilst some of the age-CpGs found in this study had been reported to correlate with age in non-musculoskeletal tissues (e.g. blood, liver, lung) (Day *et al.*, 2013; Hannum *et al.*, 2013) other CpGs had not been identified previously, suggesting that they may be specific to OA cartilage. Using data from the

ENCODE and ROADMAP project several potential targets were identified of these age-CpG regions together with their potential functionality from the chromatin state in different cell types. A great majority of the age-CpGs were mapped to gene promoters (~70%) but approximately 9% were located within regions with enhancer chromatin state in the E049 in vitro differentiated chondrocyte cells.

Chapter 2: Methods

2.1 Tissue preparation

2.1.1 Patient tissue sample collection.

Tissue samples of cartilage (CN), synovium (SYN), fat pad (FP) and blood (BL) were received from patients undergoing total knee replacement surgery due to end-stage knee OA, these being received after informed consent and with ethical approval from the Newcastle and North Tyneside Ethics Committee (REC reference number 14/NE/1212). The patient samples received were then stored at 4 °C with tissue being collected within 48 hours from the surgery. Additionally, samples of cartilage and synovium were received from UHN OA Tissue Biobank, Toronto, Canada (REB: 14-7592-AE) under request ID: OBR-029 DNA Methylation, UHN/Newcastle University. A summary of characteristics of the samples used for DNA methylation and gene expression listed in Tables Table 2. 1 and Table 2. 2 (Full table of individual sample details Table 2. 10).

Table 2. 1. Summary of patient data of samples used for bisulphite pyrosequencing.

	Cartilage	Synovium	Fat Pad	Blood
Age mean (range)	67 (48-87)	69 (46-88)	69 (46-85)	71(45-85)
%female	62.5	59	55	57
Sample size	72	22	22	7

Table 2. 2. Summary of patient details of the samples used for qRT-PCR analysis.

	Cartilage	Synovium	Fat Pad
Age mean (range)	67 (48-87)	66(45-87)	68(45-86)
%female	60.4	61.5	55
Sample size	48	39	40

2.1.2 Tissue grinding.

Tissue samples of cartilage, synovium and fat pad were ground using a RETSCH MM400 grinder and RETSCH MM200 grinder using the following settings: 1.5ml stainless steel (RETSCH, Cat# 10099031) grinding jar with a stainless steel 10mm ø ball size (RETSCH, Cat# 053680063) cooled in liquid N_2 using a frequency of 25-30 oscillations/sec for 90 sec per cycle with in between cycle cooling in liquid N_2 with three cycles for cartilage samples and five cycles for synovium and fat pad samples. Ground tissue was then transferred to a storage container and stored at -80 °C.

2.2 Nucleic acid extraction.

2.2.1 Cartilage DNA extraction.

DNA was extracted from ~200 mg of ground knee cartilage (for the samples received from UHN OA Tissue Biobank, Toronto ~100 mg of cartilage) using the E.Z.N.A.® Tissue DNA Kit (Omega Bio-tek, Cat#D3396-01) using the following protocol. Per sample, on dry ice the sample was transferred to a bijou tube and one ml of TL buffer and 60 µL of OB protease were added. The samples were then vortexed for 10 seconds and incubated for two hours at 55℃ with occasional vortexing. After the incubation, the samples were then centrifuged at 16200xG for five minutes and the eluent was transferred to a new bijou tube, to which 1040 µL of BL buffer

was added and then mixed by vortexing. The sample was then incubated for 10 minutes at 70℃ after which 1040 µL of 100% Ethanol was added and mixed in by vortexing. The solution was transferred to a HiBind minicolumn in portions of 1ml and centrifuged at 16200xG for 1 min until all the samples solution was through the column. Following this, the columns were washed according to the manufacturer's instructions. To the columns containing the bound DNA, 100 µL of provided elution buffer (50 µL for the Toronto samples) pre-heated at 70°C was added to the membrane with subsequent incubation at 70 °C for five minutes followed by centrifugation at 16200xG for one minute. The elution was then transferred again to the column and incubated for five minutes at room temperature followed by centrifugation at 16200xG for one minute. Another elution was performed using fresh elution buffer by adding 100 µL (50 µL for Toronto samples) pre-heated at 70°C to the column incubating for five minutes at room temperature and centrifuging at 16200xG for one minute. The two eluents were then combined and were quantified on an 8-sample spectrophotometer ND-8000 (Labtech) and ran on a 1% (Tris-Acetate-EDTA) TAE agarose gel to check for quality and integrity (Figure 2. 1 A) with then storage at -20°C.

2.2.2 Cartilage RNA extraction.

RNA was extracted from knee cartilage using a modified Phenol chloroform method. Approximately one gram of ground cartilage sample (for the samples received from UHN OA Tissue Biobank, Toronto 300-400 mg of sample was used) was transferred to a 50 ml universal tube being kept on dry ice to which four ml of Qiazol Lysis Reagent (Qiagen, Cat#79306) were added followed by vortexing for 20 seconds and passing several times through a syringe with a 21 G needle to aid homogenization. After incubating for 15 minutes at room temperature (RT) the samples were split equally into four 1.5 ml Eppendorf tubes and centrifuged at 13000xG for three minutes at 4℃ to remove the debris. The supernatant was collected and transferred to new 1.5 ml Eppendorf tubes to which 200 µL of chloroform was added per one ml of Qiazol solution. The solution was shaken vigorously for 30 seconds and again incubated at RT for 10 min. This was followed by centrifugation at 13000xG for 15 min at 4℃ and subsequent storage at 4℃ overnight. The following day the samples were centrifuged again at 13000xG for 15 min

at 4℃ and the top aqueous layer was collected and transferred to a new 1.5 ml Eppendorf tube. To this, 250 µL of 100% EtOH per 500 µL of aqueous top layer volume, was added and then mixed by inverting several times. The solution was then added to a Qiagen RNeasy Mini Kit (Cat# 74104) spin column and centrifuged at 10300xG for 30 sec at RT, discarding the filtrate. This procedure was repeated until all of the solution from the same cartilage sample was passed through the column. To the column, 700 µL of prepared RW1 buffer were added and after centrifugation and removal of the filtrate, 500 µL of RPE was added to the column. The column was centrifuged at 10300xG for 30 sec at RT, and after the filtrate was discarded another 500 µL of RPE was added and centrifuged as before. The column was then centrifuged for 1 minute at 10300xG for 1 min to dry. To the column membrane, 30 µL of provided elution buffer (20 µL for the Toronto samples) was added and after incubating at RT for 5 min and centrifuging at 10000xG for 1 min and elution was collected and re-added to the membrane with the incubation and centrifugation step being repeated. The RNA samples were quantified on an 8-sample spectrophotometer ND-8000 (Labtech) and ran on a 1% TAE agarose gel at 65V to check for quality and integrity (Figure 2. 1 B) and then stored at -80°C.

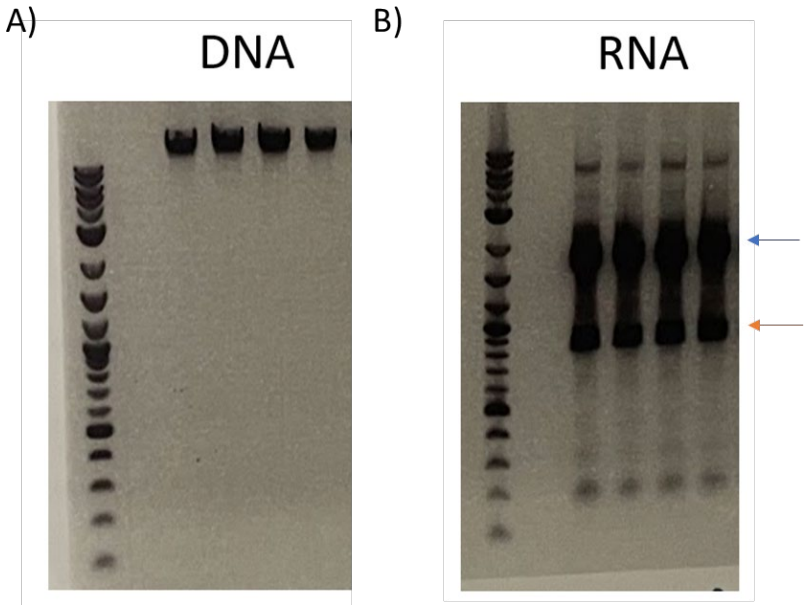


Figure 2. 1. DNA (A) and RNA (B) samples extracted from tissues samples showing good integrity with no degradation. Blue arrow indicates 28S band and orange arrow indicates 18S band.

2.2.3 Blood DNA extraction.

DNA was extracted from whole blood samples collected in PaxGene RNA tubes using the following protocol. One ml of blood was thawed and mixed with eight ml of red blood cell lysis buffer solution (0.14 M NH4Cl. 0.017 M Tris at pH 7.65), and incubated at 37℃ for 15 min. The solution was then centrifuged at 2500xg for five min and the supernatant was disposed. Red blood cell lysis was added until a volume of 4.5 ml was reached, and this was mixed and incubated further at 37℃ for 10 min. The solution was centrifuged again at 2500xG, and the supernatant was removed. Four ml of saline solution (0.15 M NaCl) was added to the pellet, vortexed and centrifuged at 2500xg with the removal of the supernatant. Next, 750 µL of nucleic lysis buffer (10 mM Tris pH 8 , 2mM EDTA pH 8, 400mM NaCl) was added to the pellet and after vortexing, 50 µL of 10% SDS and 125 µL of Pronase E solution (3 mg/ml Pronase E from *Streptomyces griseus* (Sigma, Cat# P5147-100MG), 1% SDS, 2mM EDTA pH 8) was added and the solution was incubated at 37℃ overnight. Following this, 250 µL of ~6 M NaCl was added and after vigorous shaking for 15 sec the solution was centrifuged 2800xg for 18 min. The supernatant was transferred to a new tube, 2 volumes of absolute ethanol were added and after mixing the samples was centrifuged at 17000g for 15 min. The supernatant was removed, and the pellet was washed in 70% ethanol and centrifuged at 17000xg for 5 min, removing the supernatant. The pellet was air dried for 3 min resuspended in 50 µL of TE buffer (10 mM Tris pH 8, 1mM EDTA pH 8) and stored at -20°C.

2.2.4 Synovium and Fat Pad DNA & RNA extraction.

DNA and RNA were extracted from ~400 mg of synovium and fat pad samples using E.Z.N.A.® DNA/RNA Isolation Kit (Omega Bio-tek, Cat#R6731-01) using the following optimised protocol. Each sample was incubated with 2 ml TRK lysis buffer (Omega Bio-tek, Cat#PR021) + 60 µL of 2-

mercapto-ethanol (Sigma-Aldrich). The samples were then vortexed for 20 seconds and passed through a 21 G needle using a 5 ml syringe. The lysate was then centrifuged at 17000xG for five minutes at RT, after which the protocol was followed per the instructions of the manufacturer. DNA was eluted in 100 µL of elution buffer pre-heated at 70°C with the RNA being eluted in 50 µL DEPC water pre-heated to 70°C. The DNA and RNA samples were quantified on an 8-sample spectrophotometer ND-8000 (Labtech) and ran on a 1% TAE agarose gel to check for quality and integrity. The DNA samples were stored at -20°C and RNA at -80°C.

If following the gel visualisation, the RNA samples showed DNA contamination, the samples were then DNase treated using DNA-free Kit DNase Treatment and Removal (Thermofisher Scientific, Cat#AM1906) using the Routine DNase treatment protocol from the manufacturer. The samples were then run on a 1% TAE agarose gel to check the removal of the DNA contamination and stored appropriately.

2.2.5 Nucleic acid extraction from cell lines

Nucleic acids were extracted from frozen cell pellets containing approximately $1x10^6$ cells/pellet using the E.Z.N.A.® DNA/RNA Isolation Kit (R6731-01, Omega Bio-tek). The extraction was carried out as instructed by the manufacturer with a modification to the elution step in order to improve yields. For the RNA, 50 µL of DEPC water pre-heated at 70°C was added to the column membrane containing the bound RNA and was incubated at RT for five minutes being followed by centrifugation at 13000xG for one minute. The eluent containing the RNA was added again to the membrane and incubated for a further five minutes at RT, after which it was centrifuged at 17000xG for one minute to collect the eluent. The RNA samples were stored at -80°C. For the DNA, 100 µL the provided elution buffer was pre-heated at 70°C and the added to the column membrane containing the bound DNA following with the incubation at 70°C for five minutes and centrifugation at 13000xG for one minute. The eluent was again added to the same membrane and a further five-minute incubation at 70°C and centrifugation at 13000xG for one minute was performed to collect the eluent. The DNA samples were stored at -20°C. Nucleic

acid samples were then run on a 1% TAE agarose gel to check for integrity and quantified on an 8-sample spectrophotometer ND-8000 (Labtech).

2.2.6 Nucleic acid extraction from primary human chondrocytes

Approximately $1x10^6$ H.A.C. cultured cells at passage two were collected for nucleic acid extraction. This was then split into 1/3 for DNA extraction and 2/3 for RNA extraction. DNA was extracted using the E.Z.N.A.® Tissue DNA Kit (Omega Bio-tek, Cat#D3396-01) using the following protocol. The cell pellet was resuspended in 200 µL of PBS at 4°C and 25 µL of OB proteinase solution and 220 µL of BL buffer was added followed by vortexing and incubation at 70°C for 10 minutes. To this 220 µL of 100% ethanol was added, mixed by vortexing and the solution was transferred to a HiBind DNA Mini Column and centrifuged at 13000xG for one minute. Following this, the manufacturer's instructions were followed for the washing of the membrane bound DNA. To the column membrane 100 µL of elution buffer heated at 70°C were added and the column was incubated at 70°C for five minutes. This was followed by centrifugation at 13000xG for one minute with the filtrate containing the DNA sample. The DNA samples were analysed for quantity and quality using an 8-sample spectrophotometer ND-8000 (Labtech) and by gel electrophoresis. RNA samples were store at -20°C until further use.

RNA was extracted using a phenol chloroform methodology as follows. The pellet for RNA was resuspended in 250 µL of Qiazol Lysis Reagent (Qiagen, Cat#79306) and incubated at room temperature for five minutes, this being followed by the addition of 50 µL of chloroform (Sigma Aldrich) mixed by vortexing and incubation for two minutes at RT. Following the incubation, the solution was centrifuged at 16200xg for 15 min at 4 °C with the separated aqueous phase being transferred to a new 1.5 Eppendorf tube and 125 µL of isopropanol and 1 µL GlycoBlue Coprecipitant (15 mg/mL, Invitrogen) were added, mixed by vortexing and incubated at RT for 10 minutes. The solution was then centrifuged at 16200xg for 15 min at 4 °C, discarding the supernatant and to the RNA pellet 250 µL of 75% Ethanol (v/v) were added to wash it. This was followed by centrifugation 16200xg for 15 min at 4 °C and the removal of the supernatant with the air pellet being left to air dry for five minutes. The air-dried pellet was resuspended in 15 µL DEPC water and the RNA content and quality was checked using an 8-sample

spectrophotometer ND-8000 (Labtech) and by gel electrophoresis. RNA samples were store at -80°C until further use.

2.3 Tissue cell culture of cell lines and human primary cells.

2.3.1 Tissue cell culture cell lines.

The SW1353 human chondrosarcoma cell line (Donor: Caucasian 72 years old female; Company: ATCC) and TC28a2 (Donor: 15 years old, female;(Goldring *et al.*, 1994)) were cultured in DMEM/F12 prepared medium with incubation at 37°C and 5% CO_2 in T25, T75 or T175 flasks (ThermoFisher Scientific).

- DMEM-F12 prepared media:

Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM-F12) media (Gibco) supplied with L-glutamine (Sigma Aldrich) (1%), Foetal Bovine Serum (10%) (ThermoFisher Scientific), Penicillin-Streptomycin (Sigma Aldrich) (100 U/ml) and Nystatin (1U/ml) (Sigma Aldrich).

When the confluency reached ~90%, the cells had their media removed and were washed in sterile PBS pH 7.2. The PBS was removed, and 0.05% Trypsin-EDTA solution was added followed by incubation at 37°C 5% CO_2 for five minutes after which the flask containing the cells was inspected under the microscope to check for cellular detachment. Once the cells were detached, 3x the volume of trypsin-EDTA solution was added of culture media and the solution was mixed and transferred to a suitable centrifugation tube to be centrifuged at 500xG for 5 minutes at RT. The supernatant was removed, and fresh culture media was added followed by re-suspension of the cell pellet and counting on a Haemocytometer to determine cell number. Upon requirement a part of the cells were added to new tissue culture flask to continue the culture.

Cell pellets were collected for experiments by centrifuging at 500xG for 5 minutes the appropriate volume to obtain $1x10^6$ cells followed by a (Phosphate buffer saline) PBS wash and final centrifugation 500xG for 5 minutes. The supernatant was removed, and the cell pellet was frozen at -80°C until further use. Frozen aliquots for long term liquid nitrogen storage were

collected by resuspending a cell pellet of $1x10^6$ cells in one millilitre of FBS containing 5% DMSO and transferred to 1.2 ml cryovials (Corning, Cat# 430487) followed by incubation at -80°C inside a Mr. Frosty Freezing container (ThermoFisher Scientific, Cat# 5100-0001) for up to one week and subsequent transfer to liquid N_2 storage.

2.3.2 Human articular chondrocytes extraction and cell culture.

Human articular cartilage tissue samples were collected, after informed consent, from patients with end-stage OA that were undergoing arthroplasty surgery at the Freeman Hospital, Newcastle upon Tyne. Ethical approval for this was obtained from the Newcastle and North Tyneside Ethics Committee (REC reference number 14/NE/1212). The cartilage tissue was removed from the joint using a scalpel, cut into small pieces (roughly 2-3 mm^2) and articular chondrocytes isolated through enzymatic digestion using 4 ml/g (1mg/ml) hyaluronidase (Sigma Aldrich Cat#H3506) in PBS of for 15 mins at 37°C, 4ml/g cartilage (2.5 mg/ml) of trypsin (Sigma Aldrich Cat# T7409) in PBS (Sigma Aldrich, T7409) for 30 min at 37°C and final digestion with 3 ml/g (2 mg/ml) collagenase type 2 (Sigma Aldrich Cat# C0130) overnight at 35°C. The following morning the cells were passed through a 100 µm Cell Strainer and centrifuged at 500xG for five minutes. The cells were resuspended in DMEM H.A.C. media and seeded in 25 cm^2 or 75cm^2 plastic tissue culture flasks at 37°C 5% CO_2 and were let to recuperate for two weeks before beginning treatment with stimuli, this being considered the first passage.

- DMEM-F12 H.A.C. media:

Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM-F12) media (Gibco) supplied with L-glutamine (Sigma Aldrich) (1%) Foetal Bovine Serum heat inactivated (10%) (ThermoFisher Scientific, Cat# 10270106), Penicillin-Streptomycin (Sigma Aldrich) (100 U/ml) and Nystatin (1U/ml) (Sigma Aldrich).

2.3.3 FHL2 overexpression cell line generation

The generation of a line overexpressing *FHL2* was done using the FHL2-GFP plasmid vector (Cat#: RG220395, Origene) (Figure 2. 6) containing a *FHL2* human mRNA transcript variant fused to a green fluorescent protein. The control plasmid was made by enzymatic digestion removal of the *FHL2* ORF with 20 units of AsiS I (New England Biolabs, Cat#R0630S) and 20 units of Mlu I (New England Biolabs, Cat#R3198S) restriction enzymes was ligated with T4 ligase as described in section 2.7.2. with a replacement sequence with a multiple cloning site with the following sequence:

5' CGCCGGCGCGCCAGATCTCAAGCTTAACTAGTTAGCGGACCGA 3'

The sequence is found in the original plasmid (pCMV6-AC-GFP) (Figure 2. 7) used to generate the FHL2-GFP vector. The transfection was done by seeding $1x10^5$ SW1353 cells and transfecting using 1 µg of each respective plasmid linearized with 30 units of Pci I (New England Bioblabs, Cat# R0655S, digested and purified as described in section 2.9.2) and 3 µL of FugeneHD. This was followed by the selection of the cells that have acquired the plasmid using G418 1mg/ml for two weeks in order to generate a stably overexpressing line. After the selection period the antibiotic treatment was maintained at 0.5 mg/ml G418 while the cells were in culture.

The presence of the construct was determined using qRTPCR, fluorescent microscopy together with western blotting using an anti-FHL2 (Abcam, Cat#ab12327) antibody to determine the presence of a ~56 kDa band FHL2-GFP fusion protein. The qRTPCR primers used are listed in Table 2. 6.

2.3.4 Reintroduction of FHL2 construct in FHL2 knock down SW1353 cell line.

Using the SW1353 Cas9 cells that presented the highest level of knock down of FHL2, a reintroduction of FHL2 was done using the plasmid containing the construct for FHL2-GFP and these were selected, as described previously (section 2.3.3). The presence of the construct was determined using qRT-PCR, fluorescent microscopy and western blotting.

2.4 Cell proliferation assays

2.4.1 BrdU cell labelling proliferation assay.

SW1353 cells were seeded at a density of $1x10^4$ cells/well in 8 well glass chamber sliders (Cat#154534, LabTek) for 24 hours, after which these were incubated with the cell proliferation labelling reagent bromodeoxyuridine (BrdU), fluorodeoxyuridine (FdU) (Cat# RPN201, Amersham Biosciences) reagents at the concentration recommended by the manufacturer for 3 hours at 37°C 5% CO_2. The cells were washed in PBS and then fixation was performed using a solution of 95% Ethanol and 5% acetic acid for 10 min at room temperature. Following this, the cells were washed in PBS and incubated with 4M hydrochloric acid for 10 minutes at RT. After this step, the cells were washed and incubated with 0.1 M Borate buffer pH 8.5 for 20 minutes at RT. Following this, the method described in the immunocytochemistry section 2.8.4 (from the permeabilisation step) was followed. The primary antibody used to detect BrdU positive cells was Anti-BrdU antibody [BU1/75 (ICR1)] (Cat# ab6326, Abcam) diluted to 5 µg/ml in PBS and the incubation was done overnight at 4°C, this being followed by incubation with donkey anti-rat Alexa Fluor488 secondary antibody at a concentration of 1µg/mL (Invitrogen, Cat# #A-21208) in 10% Goat serum in PBS (Vector Laboratories, Cat# S-1000) for two hours at RT and mounted in VECTASHIELD® Antifade Mounting Medium with DAPI (Vector Laboratories, Cat#H-1200-10). BrdU positive cells were counted as a percentage of total DAPI positive cells.

2.4.2 WST-1 cell proliferation assay.

The WST-1 assay is a colorimetric cell proliferation assay based on the cleavage of the tetrazolium salt WST-1 molecule by mitochondrial dehydrogenase enzymes generating a coloured product (formazan) with an absorption maximum at 435 nm.

Cells were seeded at a density of $5x10^3$ cells/100µL/ well in a 96 well plate for 24 hours at 37°C 5%CO_2. After this period, 10µL/well of the WST-1 reagent (Roche, Cat# 5015944001) was added and further incubated for two hours at 37°C. The absorbance was read at 435 nm on a

Varioskan LUX Multimode Microplate Reader (Thermofisher Scientific). The experiment was performed in triplicate with four technical repeats.

2.4.3 Cellular treatment and stimulation

2.4.4 5-Aa 2'deoxycytidine treatment.

Approximately 2.5x10^5 SW1353 or TC28a2 cells were seeded in a T25 flask with the treatment being done in the presence or absence of 5-Aza 2'deoxycytidine (Sigma-Aldrich, Cat# A3656-5MG) at 0.25 & 0.5 µM for SW1353 and 5 & 10 µM for TC28a2 with an equivalent dosage of DMSO solvent for the control for 72 hours of culture, with a media change at every 24 hours. After this, the cells were collected and nucleic acids were extracted as described previously (section 2.2.5). The experiment was performed in biological triplicate.

For the HACs (patient details available in Table 2. 11) approximately 4x10^5 cell were seeded in T25 flask with the 5-Aza treatment done at 5 µM or the equivalent of Dimethyl sulfoxide (DMSO) solvent for the control for a period of two weeks with fresh treatment being added every two days and media fully refreshed every four days. After this, the cells were collected and nucleic acids were extracted as described previously (2.2.5).

2.4.5 IL-1 stimulation of H.A.C.s.

RNA was available from H.A.C.s (patient details available in Table 2. 11) treated with or without 0.1ng/ml IL-1 alpha for two consecutive passage approximately 22 day by Dr. Reynard and William Thompson These samples were reverse transcribed as described in section 2.7.1 and subsequently used for qRT-PCR to determine whether the treatment impact *FHL2* expression together with the expression of proliferation of cell cycle genes such as *MKI67, TPX2, CCND1* and *CDKN1A*. The list of the primers used can be found in Table 2. 6.

2.4.6 IL1 stimulation of SW1353 cells, SW1353 FHL2 K.D. and SW1353 FHL2-GFP

Control SW1353 cells or SW1353 with *FHL2* knock down or *FHL2-GFP* overexpression were seeded at a density of 1x10^4 cells/well in 8 well glass chamber sliders (Cat#154534, LabTek) and left to settle for 4 hours, after which IL1 alpha at 0.5 ng/ml was added for 24 hours at 37°C 5%CO_2.

2.4.7 Native FHL2 nuclear translocation after IL-1 stimulation

Native SW1353 cells were seeded at a density of $1x10^4$ cells/well in an eight-well glass chamber slide (LabTek, Cat#154534) and incubated for 24 hours, after which these were treated with IM-1 alpha 0.5ng/ml for 1 hour, 40 min, 20 min or 0 min at 37°C 5% CO_2. Following this, the cells were fixed, and the immunocytochemistry protocol was afterwards followed (section 2.8.4). DAPI was used to visualise the nuclear signal and FHL2 with anti-FHL2 (ab12327) and respective fluorophore tagged secondary antibody. The cells that showed positive nuclear staining for FHL2 were then counted and reported to the total number of cells as given by DAPI staining for each different time point. This gave a percentage measure of the cells that show nuclear translocation of FHL2 after IL1 stimulation at a given time period. The experiment was done in biological triplicate with four images being analysed per well per replicate.

2.5 CRISPR-Cas9 and derivatives

2.5.1 crRNA design for CRISPR-Cas9

The region surrounding the found age-CpGs were subjected to Cas9 deletion using a pair system of sgRNAs located on both sides of the age-CpGs. When designing single guide RNA (sgRNA) nucleotides for CRISPR-Cas9 work, the first part is to design crisprRNA (crRNA) sequences. The crRNA is a RNA nucleotide sequence of around 17-20 bases that is complementary to the target DNA region. Additionally the tracrRNA nucleotide is also needed as this is a scaffold sequence to bind to the Cas9 nuclease.

The design of crRNA for CRISPR targeting was realized using four software webtools:

http://www.crisprscan.org/

https://crispr.cos.uni-heidelberg.de/

https://www.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE

http://chopchop.cbu.uib.no/

The crRNAs were chosen based on their proximity to the age-CpG region (distance of ±200 bp away from the age-CpGs) and based on the best given off-target and on target score for the region of interest (Table 2. 8).

2.5.2 sgRNA duplex formation

The crRNA was ordered from IDT and 20 nmol lyophilised crRNA was resuspended in 200 µL Nuclease-Free Duplex buffer (IDT) giving a concentration of 100 µM and was stored at -80°C until further use. The crRNAs and tracrRNA (Integrated DNA Technologies (IDT), Cat# 1072533) were combine in an equimolar ratio of 200 picomoles, followed by annealing together by heating to 95°C for five minutes and slow cooling to RT to form the sgRNA duplex. At this stage the formed complex sgRNA contains both the designed crRNA sequence to target the region of interest as well as the tracrRNA sequence to interact with the Cas9 nuclease combine in a duplex formation.

2.5.3 Transfection of sgRNAs into SW1353 cells

SW1353 cells stabling expressing the Cas9 nuclease enzyme generated using the lentiCRISPR-v2 (Addgene 52961, Sanjana et al 2014) were a kind gift of the Young lab (Barter *et al.*, 2021), and cells were maintained in SW1353 with 0.5ug/ml puromycin selection. Newcastle University For performing the genomic region deletions, $1x10^5$/well SW1353-Cas9 cells were seeded in 1 ml prepared DMEM-F12 medium in a six well plate followed by incubation for 24 hours at 37°C and 5% CO_2. Transfection reaction per well was prepared using 200 picomoles of prepared sgRNA duplex with 4 µL of DharmaFECT 1 Transfection Reagent (Dharmacon, Cat#T-2001-01) and neat DMEM-F12 medium to make 200 µL reaction volume, this being followed by incubation at RT for 20 minutes and addition to the respective plate well. The cells were maintained in culture

until confluency after which they were transferred to a T25 flask. Once confluent, the cells were then collected for liquid N_2 storage as well as for protein, DNA and RNA extraction.

2.5.4 Determination of deletion efficiency.

Primers to determine the deletion efficiency through qPCR were designed as following the following publication (Li *et al.*, 2019) and also the advice of Prof. Qilai Huang. Primers were designed located at a distance of 200-500 bp away from respective deletion. Controls where a random sgRNA pair with a known outside targets were used together with primers located inside the deletion and at 200-500 bp away from the deletion (Table 2. 6). The determination was done by using the following equation.

$$Editing\ eff.\% = 1 - 2^{-\Delta Ct(test\ amplicon) - \Delta Ct(control\ amplicon)}$$

2.5.5 Generation of CRISPR-Cas9 knock down variants of FHL2 in SW1353 cells.

In order to generate a *FHL2* knock down cell line, crRNAs were designed to target the third common exon (common exon between all known isoforms of this protein) or the promoter of *FHL2* (Table 2. 8). The crRNAs were complexed and were transfected into the Cas9 SW1353 cell line as described in section 2.5.3. Confirmation of the knock down was performed using western blotting. Three different *FHL2* deletion variants were produced, one by targeting the promoter of *FHL2,* another by transfecting just one gRNA targeting the third isoform common exon producing an approximately 40% reduction in FHL2 levels and a third one where both gRNAs were transfected simultaneously both targeting the common third exon for the isoforms, this producing a reduction in FHL2 levels of >90% compared to the control.

2.5.6 dCas9-DNMT3a and dCas9-TET1cd SW1353 cell line generation.

To attempt methylation and demethylation of age-CpG regions, SW1353 cells stably expressing the catalytically dead Cas9 protein fused to either the catalytic domain of the DNA methyltransferase enzyme DNMT3a or catalytic domain of the TET1 dioxygenase were generated. Fuw-dCas9DNMT3a (Addgene #84476, Figure 2. 4) (Liu *et al.*, 2016) and Fuw-dCas9TET1cd (Addgene #84475, Figure 2. 5) (Liu *et al.*, 2016) plasmids were supplied by Addgene as bacterial stabs , these being then streaked on LB agar plates with 100 µg/ml Ampicillin (Sigma Aldrich, Cat# A9518) and grown overnight at 37°C. Clones were then picked and grown further in 3 ml of LB broth supplied with 100 µg/ml Ampicillin and miniprepped using the following protocol. The culture was transferred to 1.5 ml centrifuge tubes and was centrifuged at 10000xG to pellet the bacterial cells, disposing of the supernatant. To this 100 µL of P1 solution (Qiagen) was added to the tubes containing the pellet after which it was vortexed until the pellet was resuspended. This was followed by the addition of 200 µL of P2 solution (Qiagen) and incubation at room temperature for 3 min. After the 3 min, 150 µL of P3 solution (Qiagen) was added and the tubes were centrifuged at 17000xG for 5 min. The supernatant was transferred to clean tubes, 1 ml of 100% EtOH was added followed by incubation at -80°C for 15 min. After the incubation, the samples were centrifuged at 14000xG for 5 min and the supernatant was removed. The tubes were allowed to air-dry for 5 min after which 80 µL of DPEC H2O was added to resuspend the plasmid DNA. Glycerol stocks were made for the positive clones by collecting approximately 200 µL of culture and mixing with 200 µL of glycerol with this being stored at -80°C. After confirmation of the correct sequence clones, the cultures were grown up by adding 30 µL from the respective glycerol stocks in 100 ml of LB broth with 100 µg/ml Ampicillin and cultured at 200 RPM and 37°C. The plasmid was extracted from this culture using the PureYield™ Plasmid Midiprep System (Promega, Cat# A2492) according to the manufacturer's instructions and eluting in 400µL of DEPC water. Plasmids were linearised by digestion with 30 units Pvu I HF (Promega Cat# R6321) at 37°C for 16 hours and purified using a QIAquick Gel Extraction kit (Qiagen, Cat# 28704) following the manufacturer's instructions with this being eluted in 30 µL of DEPC water. The purified product was then stored at -20°C.

SW1353 chondrosarcoma cells were seeded at a density of $1x10^5$ cells/ well in a six well plate in prepared DMEM/F12 medium. After 24 hours, the cells were transfected with 3 µL /well FuGENE® HD Transfection Reagent (Promega, Cat# E2311) and one µg /well of Fuw-dCas9DNMT3a (Addgene #84476) (Liu *et al.*, 2016) or Fuw-dCas9TET1cd (Addgene #84475)

(Liu *et al.*, 2016) constructs in either circular form (undigested) or linearised version. In order to generate stable construct expressing cell lines, after 48 hours from the transfection, the cells were treated with 200 µg/ml Zeocin antibiotic until confluency (approximately three weeks). The variants generated were screened using qRT-PCR with the primer sequences available in Table 2. 6, western blotting and immunocytochemistry with antibodies details available in Table 2. 10.

2.5.7 dCas9-DNMT3a and dCas9TET1cd cell line validation.

A literature search revealed that certain sgRNAs had been successfully used to induce DNA methylation changes at specific regions using the same or similar constructs to those we generated. As a positive control to test the functionality of these cell lines the following crRNA were used: for the dCas9-DNMT3a construct 5'-CTGTGTGCTAAACCTCCCGT (Pflueger *et al.*, 2018a) and for the dCas9-TET1cd 5'-GGCGTCCATTCTGGCCGTGC (Choudhury *et al.*, 2016).

In a six well plate $1x10^5$ SW1353-Cas9 cells were seeded in 1 ml prepared DMEM-F12 medium per well in a six well plate followed by incubation for 24 hours at 37°C and 5% CO_2. Transfection reaction per well was prepared using 200 picomoles of prepared sgRNA duplex, that was formed by annealing a fluorescently tagged ATTO 550 tracrRNA (IDT, Cat# 1075927) with the respective crRNA (Table 2. 8) with 4 µL of DharmaFECT 1 Transfection Reagent (Dharmacon, Cat#T-2001-01) and additive free DMEM-F12 medium to make 200 µL reaction volume, this being followed by incubation at RT for 20 minutes and addition to the respective plate well. The control cells were transfected with tracrRNA only as a negative control. Additionally, after 24

hours some cells were transfected with gRNAs as a second time. Following this, the cells were collected after they were in culture for 72 hours as the control (tracrRNA only), transfected once cells, and twice transfected group, with the nucleic acids being extracted as indicated in section 2.2.5. DNA methylation changes by bisulphite pyrosequencing was analysed as described in section 2.6, using the pyrosequencing primers listed in Table 2. 6. The primers for the bisulphite pyrosequencing assays were designed based on the indicated target regions of the gRNAs reported in the literature (Choudhury *et al.*, 2016; Pflueger *et al.*, 2018a) (section Table 2. 8).

2.6 Bisulphite pyrosequencing.

2.6.1 Bisulphite PCR primer design

Primers for the amplification and pyrosequencing of the age-CpGs were designed based on the DNA sequence at approximately ±100bp DNA from the age-CpGs under investigation. The DNA sequences were extracted from UCSC genome browser (https://genome.ucsc.edu/, using the Human Assembly: Feb.2009 (GRCh37/hg19)) and were used to design pyrosequencing assays using the PyroMark assay design SW 2.0 software (QIAGEN), choosing primer sets that had a primer set score >60, primer length between 18- 30 and max CG of ~50% (Table 2. 6). For each pair of primers one oligomer sequence in either forward or reverse direction had a biotin tag that allows the binding to sepharose beads and the subsequent purification of tagged products and one sequencing primer to be used to sequence the amplified region on the Qiagen Pyromark Q24 pyrosequencer machine. Primer sequences were ordered through IDT and these were resuspended at 100 mM concentration in DEPC water and further diluted to 10 mM for the PCR.

2.6.2 Generation of calibration standard for pyrosequencing methylation assays.

To confirm whether the DNA methylation readings were accurate and no amplification bias is observed synthetic DNA oligomers as gBlocks gene fragments (IDT) of the regions encompassing the age-CpGs were designed and ordered through IDT as two bisulphite converted DNA versions: for 0% methylation the normal sequence, for 100% the cytosines in

the CpGs were replaced with thymine this representing fully methylated DNA; additionally short sequence adapters were added at 5’ and 3’ ends to decrease sequence complexity to permit the synthesis. The list of the sequences is available in Table 2. 3.

The gBlocks at 250 ng were resuspended in 100 µL DEPC water (2.5 ng/ µL) and incubated at 50°C for 20 minutes followed by slow cooling to RT. This was further diluted to 25 pg/µL and used to produce a gradient starting from 0-100% with 10% increments subsequently used for PCR amplification. From the results of this, a function was fitted between the expected vs. observed points (either linear or quadratic) and the raw results from all the pyrosequencing was adjusted using the function for each individual assay (Figure 2. 2).

Table 2. 3. Gblock sequence used for the calibration standard for the age-CpGs pyrosequencing assays. Blue highlight marks the position of primer binding and

Region CpGs & nearby gene name	Methylation status	gBlock sequence 5'-3'
cg14566959 (*PCDHGA1*)	Unmethylated	TTGTGTTAGAAGAGATGGATAAAGGTTTTTTTGTGGGTAATATTTTTAAGGATTTGGGGTTGGATTTTTGGAAG TTGGTGAAGTATGGAGTTTGTATTGTTTTTAGAGGTAGGATGTAGTTTTTTGTTTTGAATTTGTGTAGTGGTAGT TTGATTATTGTGGGTAGGATAGATTGGGAGGAGTTTTGTGTTTAGAGTTTGTGGTGTTTGATAAATATTAATAT TTTGGTTGAGGATAAAGGAAAATTTTTTGGGGGTAGAAATAGAAATAATTGATATTAATGATAATAATTTAAAAT TTTAGGTTGAAGATTTAGAAGTAAAAATTAATGAAATTGTGGTTTTTGGAGTATGTTATTTATTTTTAGAAGTTG
	Methylated	TTGTGTTAGAAGAGATGGATAAAGGTTTTTTTGTGGGTAATATTTTTAAGGATTTGGGGTTGGATTTTTGGAAG TTGGTGAAGTATGGAGTTTGTATTGTTTTTAGAGGTAGGATGTAGTTTTTTGTTTTGAATTTGTGTAGTGGTAGT TTGATTATTGTGGGTAGGATAGATTGGGAGGAGTTTTGCGTTTAGAGTTCGCGGTGTTTGATAAATATTAATAT TTTGGTTGAGGATAAAGGAAAATTTTTTGGGGGTAGAAATAGAAATAATTGATATTAATGATAATAATTTAAAAT TTTAGGTTGAAGATTTAGAAGTAAAAATTAATGAAATTGTGGTTTTTGGAGTATGTTATTTATTTTTAGAAGTTG
cg23500537 (*PCDHB1*)	Unmethylated	TGGGCATGGGGGGTGCTTTTCATCCCCGCTTCCTACTGCAATAATTTTTGTAATTTTTTTTGTAGTTTTAAGTAAGG AGATTTGTGTTGTTAGGAGTGTGGTGTAGTTATATATTTAAGGTTGATAGGGTGGGTATTTTGTTAAGTTTTGT GTGTTGTTGTTTATAATATTTTTTTTAGTTTTGTTTGTATTTGAAG
	Methylated	TGGGCATGGGGGGTGCTTTTCATCCCCGCTTCCTACTGCAATAATTTTTGTAATTTTTTTTGTAGTTTTAAGTAAGG AGATTTGTGTTGTTAGGAGTGTGGTGTAGTTATATATTTAAGGTTGATAGGGTGGGTATTTTGTTAAGTTTTGC GCGTTGTTCGTTTTTTATAATATTTTTTTTAGTTTTGTTTGTATTTGAAG
cg22331349, cg09547119 (*ZNF577*)	Unmethylated	GGGCGGGCGGGTGTTTAGGGGATGTTTGTTGGGAGTAGTTATTTTTAATTTTATTGTTGTAGAGTAGGTGGAG TTTTTTTTTTGTGTTTTAAGATAGGTAGGTTTTGATGATGAAAAGTAATTGAAAATGATTTATTTTATTTTTTTTT TAGTTTATGTGAATTGTTAGATTTTGGTTTTGTAATATTAGTTAGGGGTGCGCGCGCCGGGG
	Methylated	GGGCGGGCGGGTGTTTAGGGGATGTTTGTTGGGAGTAGTTATTTTTAATTTTATTGTTGTAGAGTAGGCGGAG TTTTTTTTTTCGCGTTTTAAGATAGGTAGGTTTTGATGATGAAAAGTAATTGAAAATGATTTATTTTATTTTTTTT TTAGTTTATGTGAATTGTTAGATTTTGGTTTTGTAATATTAGTTAGGGGTGCGCGCGCCGGGG
cg01644850, cg24214260 (*ZNF776*)	Unmethylated	CAGCCGGCGCCACACTTTTAGAAATTTGGGAAAAGTTTTTTTAAGAAGGAGTAAATATCTTCTTTTTCTTTAGA GGTTGAAGTTGTTATTTTTTTGTATGGGGTATAGTATTTTAGAATATAAGTTTGTAGAGGTTAAAGTAGTGGAT ATATTTTGAAGAGTTTTGTGGAGTTTTGGAAATTATATTATTTAGAGTGTAGAGTGTAAAATGGTGGTGGAGTT GAGTTAATGATAGTTTAGGATAGCAAC
	Methylated	CAGCCGGCGCCACACTTTTAGAAATTTGGGAAAAGTTTTTTTAAGAAGGAGTAAATATCTTCTTTTTCTTTAGA GGTCGAAGTCGTTATTTTTTTGTACGGGGTATAGTATTTTAGAATATAAGTTCGTAGAGGTTAAAGTAGTGGAT ATATTTCGAAGAGTTTCGTGGAGTTTTGGAAATTATATTATTTAGAGTGTAGAGCGTAAAACGGCGGCGGAGTT GAGTTAATGATAGTTTAGGATAGCAAC
cg16867657, cg24724428, cg21572722 (*ELOVL2*)	Unmethylated	GGGAGGGGTGTAGGGTAAGTGAGGTGGTGTTTTTTGTTTTTGTGGTTTTGTGTGTTTTTTTTTGGGTGATTGAT TTTGTTTTTGTGTTTGTGGTGTTTTTTGTTGGTTGGGTGGTGATTTGTAGGTTTAGTTGGTGTTGGTTTTGTGTG GTGGTTTAATGTTTATGGAGTTTTAGGAATATTTATTTGTTGTTTAGATTGGTAGTTGTTGTTGTGGGGAGAAGT AGTATTGTGTAGGGTGGGTATGTTGGTTTTGTTTATAGTTGGGTTTTGGTGGGTTTGAAGTATATATTAGGGGG AAATGGTTTTGT
	Methylated	GGGAGGGGTGTAGGGTAAGTGAGGTGGTGTTTTTTGTTTTTGTGGTTTTGTGTGTTTTTTTTTGGGTGATTGAT TTTGTTTTTGTGTTTGTGGTGTTTTTTGTTGGTTGGGTGGTGATTTGTAGGTTTAGTTGGTGTTGGTTTTGCGCG GCGGTTTAACGTTTACGGAGTTTTAGGAATATTTATTTGTTGTTTAGATTGGTAGTTGTTGTTGTGGGGAGAAG TAGTATTGTGTAGGGTGGGTATGTTGGTTTTGTTTATAGTTGGGTTTTGGTGGGTTTGAAGTATATATTAGGGG GAAATGGTTTTGT
cg06639320, cg22454769, cg24079702 (*FHL2*)	Unmethylated	TTTGCCAGGGCTCCTTTCTTCGTGTTTTTTGGGTTTTGGGAGTATAGTAGTTATTGGGAGTGTTGTTTTTGGTGT GGGTTTTTGGGGTGTGAGTTTTGGATGAGGTTTGGGTGTGGTGGTAGGGGTTTGTTTATGTTGGGATTTTTGTTT GGTTTTAGGAGTGGGAGATTGGAGAAGTTTTAGGATGTGTCGTGCCGGGGGAGGCGGAGGGAGGAG
	Methylated	TTTGCCAGGGCTCCTTTCTTCGTGTTTTTTGGGTTTTGGGAGTATAGTAGTTATCGGGAGCGTCGTTTTCGGCGT GGGTTTTCGGGCGCGAGTTTCGGACGAGGTTTGGGCGCGGTGGTAGGGGTTTGTTTACGTCGGGATTTTTGTT TGGTTTTAGGAGCGGGAGATTGGAGAAGTTTTAGGATGTGTCGTGCCGGGGGAGGCGGAGGGAGGAG

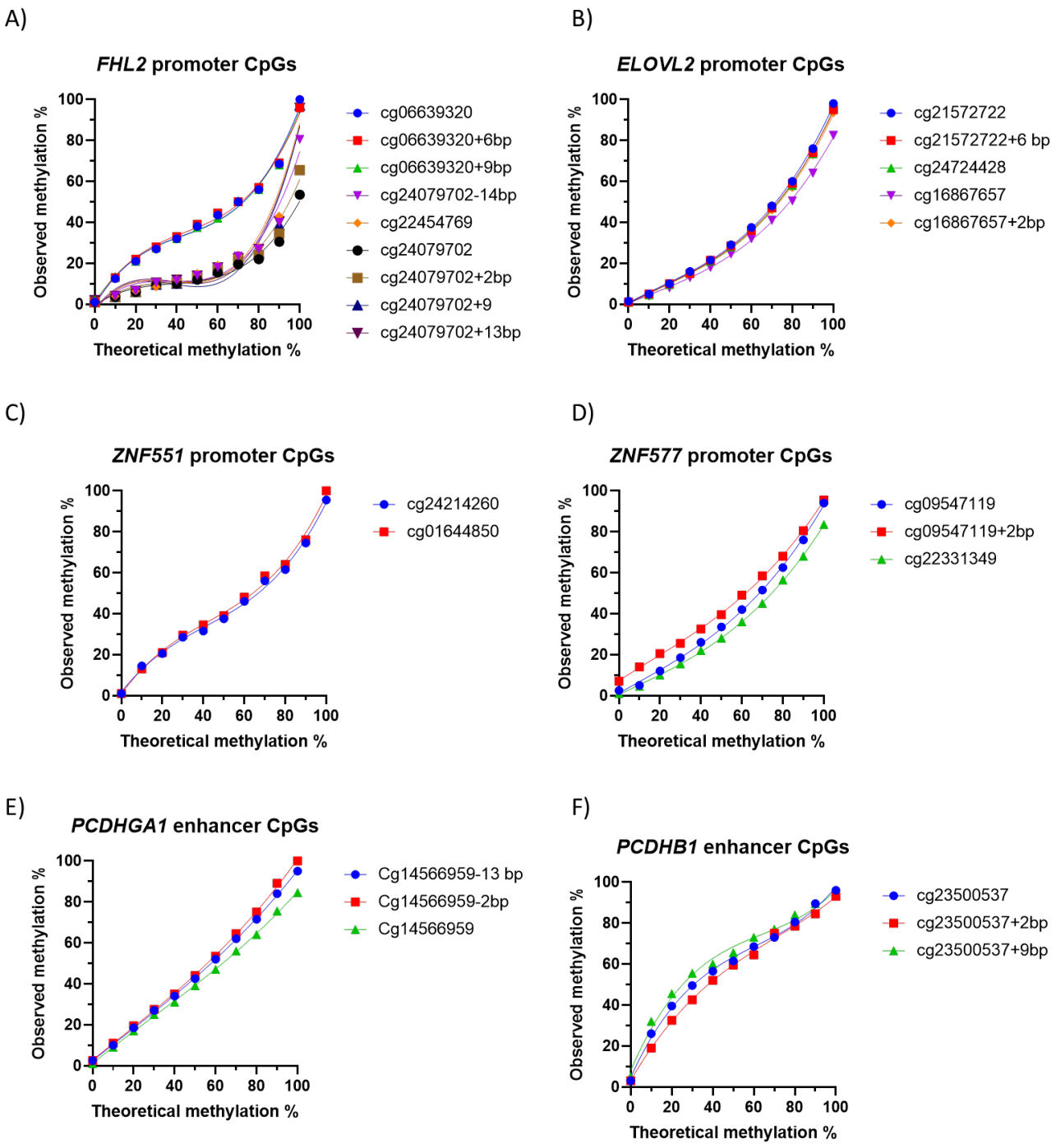


Figure 2. 2. Plot of the results of the response between the expected theoretical methylation (from the gradient standards) and response from the pyrosequencer machine. A cubic polynomial equation was fit to the data points in order to adjust the raw values to better correspond to the expected values of the standard.

Table 2. 4. Summary of the polynomial cubic equations for the calibration of DNA methylation from bisulphite pyrosequencing data for each CpG analysed. Adj.meth% stands for the adjusted methylation from the equation, OB stands for observed methylation that is the raw methylation values from the bisulphite pyrosequencing.

Region	CpG	R^2	Equation
ELOVL2 promoter	cg21572722	0.9994	Adj. meth% = -2.190 + 2.3511*OB - 0.02120*OB^2 + 0.000080*OB^3
	cg21572722+6 bp	0.9996	Adj. meth% = -1.744 + 2.3777*OB - 0.02174*OB^2 + 0.000084*OB^3
	cg24724428	0.9991	Adj. meth% = -1.101 + 2.3613*OB - 0.02131*OB^2+ 0.000080*OB^3
	cg16867657	0.9992	Adj. meth% = -1.162 + 2.733*OB - 0.02915*OB^2 + 0.000132*OB^3
	cg16867657+2bp	0.9994	Adj. meth% = -1.228 + 2.3915*OB - 0.02192*OB^2 + 0.000085*OB^3
FHL2 promoter	cg06639320	0.9904	Adj. meth% = -1.90 + 0.719*OB + 0.02492*OB^2 - 0.000220*OB^3
	cg06639320+6bp	0.9944	Adj. meth% = -1.38 + 0.537*OB + 0.02961*OB^2 - 0.000253*OB^3
	cg06639320+9bp	0.9918	Adj. meth% = -3.17 + 0.773 *OB + 0.02478*OB^2 - 0.000224*OB^3
	cg24079702+13bp	0.9419	Adj. meth% = -11.42 +5.378*OB -0.0894*OB^2+0.000474*OB^3
	cg24079702+9	0.9363	Adj. meth%= -7.87 +5.521*OB -0.0973*OB^2+0.000535*OB^3
	cg24079702+2bp	0.9664	Adj. meth%= -7.07 +5.144*OB -0.0793*OB^2+0.000392*OB^3
	cg24079702	0.9759	Adj. meth%= -7.82 +5.431*OB-0.0783*OB^2+0.000269*OB^3
	cg22454769	0.9477	Adj. meth%= -6.76 +4.835*OB -0.07636*OB^2+0.00039*OB^3
	cg24079702-14bp	0.9634	Adj. meth%= -6.76 +4.835*OB -0.07636*OB^2+0.00039*OB^3
PCDHGA1 enhancer	cg14566959-13 bp	0.9997	Adj. meth% = -2.04 +*OB-0.01169*OB^2+0.00003*OB^3
	cg14566959-2bp	0.9996	Adj. meth% = -13.98+1.905*OB-0.00838*OB^2+0.00001*OB^3
	cg14566959	0.9995	Adj. meth% = -1.281 +2.3049*OB-0.01959*OB^2+0.000078*OB^3
PCDHB1 enhancer	cg23500537	0.9972	Adj. meth% = 1.71 - 0.406*OB + 0.02705*OB^2 - 0.000126*OB^3
	cg23500537+2bp	0.9987	Adj. meth% = -0.44 + 0.280*OB+ 0.01170*OB^2 - 0.000032*OB^3
	cg23500537+9bp	0.9919	Adj. meth% = 1.89 - 0.417*OB + 0.02090*OB^2- 0.000057*OB^3
ZNF577 promoter	cg09547119	0.9989	Adj. meth% = -2.04 +1.922*OB-0.01169*OB^2+0.00003*OB^3
	cg09547119+2bp	0.9996	Adj. meth% = -13.98+1.905*OB-0.00838*OB^2+0.00001*OB^3
	cg22331349	0.9997	Adj. meth% = -1.281 +2.3049*OB-0.01959*OB^2+0.000078*OB^3
ZNF551 promoter	cg24214260	0.9963	Adj. meth%=-2.878+0.9158*OB+0.01612*OB^2+-0.00015*OB^3
	cg01644850	0.9957	Adj. meth%=-1.917+0.8203*OB+0.01588*OB^2+-0.00014*OB^3

2.6.3 Bisulphite conversion.

Approximately 750 ng of cartilage, blood, fat pad and synovium DNA was bisulphite converted using the EZ DNA Methylation kit (Zymo Research, Cat# D5001) as per manufactures instructions. Bisulphite conversion was performed for 16 hours at 50°C on a Veriti Fast Thermal Cycler (Applied Biosystems), this being followed by the washing sequences from the manufacturer's instructions. The bisulphite converted DNA was eluted in 50 µL of the provided elution buffer giving a final concentration of ~15ng/ul and then stored at -20°C.

2.6.4 PCR amplification.

For each sample and region, the PCR was run in duplicate. The PCR amplification was done using the following master mix: 0.4 µL Titanium Taq Polymerase (Takara, Cat# 639208), 2 µL of

Titanium Taq Polymerase buffer (Takara), 0.4 µL of 10mM Forward primer, 0.4 µL of 10mMReverse primer, 0.4 µL 10 mM dNTPs, 15.4 µL of DPEC H2O 31 and 1 µL of template DNA, in a total volume of 20 µL. PCR Cycle conditions: Initial denaturation 95°C for 60 s followed by 40x cycles of denaturation 95°C for 15 s, annealing X°C for 60 s, extension 68°C for 60 s and a final step at 68°C for 7 min. The annealing temperature was determined experimentally using a temperature gradient between 55-70°C program. The reaction's success was confirmed by running it on a 3% agarose TAE gel at 90 V and identifying the right sized product corresponding to each assay as well as the absence of any band in the negative control.

2.6.5 Qiagen Pyromark Q24 pyrosequencing

DNA methylation quantification was performed on a Qiagen Pyromark Q24 pyrosequencer by transferring 10 µL of the PCR product to a 24 well PCR plate (Starlab) and adding binding solution consisting of 40 µL Pyromark Binding Buffer (Qiagen), one µL Streptavidin Sepharose High Performance (GE Healthcare) and 29 µL of DEPC water. The plate was then placed on a an Orbis 700-235 Microplate Shaker (Cole-Parmer Instruments) for 10 minutes at room temperature. The solution was then processed at a PyroMark Q24 Vacuum Workstation (Qiagen), where the beads were collected using the vacuum filter probe followed by a five second wash in 75% ethanol, followed by five seconds in 0.2 M NaOH solution and a final wash for 10 seconds using 1x PyroMark Wash Buffer (Qiagen). The vacuum was switched off and then the filter probes were then placed on a 24 well plate containing 24.25 µL Pyromark Annealing Buffer (Qiagen) and 0.75 µL sequencing primer (10 µM) per each well and was gently shaken to release the beads in the solution. This was then incubated for two minutes at 80°C followed by slow cooling at RT for five minutes. Following this, the plate was placed in the pyrosequencer machine and the PyroMark Q24 dispensing cartridge was loaded with the required nucleotides, enzyme and substrate (PyroMark Gold Q96 Reagents, Qiagen) as instructed by the software as required for each of the corresponding assay. The result output is a thymine:cytosine ratio that represent the percentage of DNA methylation at the CpG site. A positive control consisting of bisulphite converted SW1353 DNA was run for every plate to confirm that the readings remain consistent with previous data. Any samples where the percentage methylation between PCR replicates varied by >5% were excluded from the analysis.

2.6.6 Generation of epigenetic clock using bisulphite pyrosequencing results.

An epigenetic clock was made for the cartilage, synovium+ fat pad tissues as well as one to contain all three tissues by using the pyrosequencing data from the CpG sites. This model was developed in order to understand whether a collaborative link can be established between age and multiple CpG DNA methylation level, where the DNA methylation levels of CpG that change with age all contribute and can be used to better determine a person's age using pyrosequencing data. The analysis was realized using Multiple-variate regression analysis in GraphPad Prism v9.3.1. At first the samples were divided into two groups where the first consisting of 2/3 of the samples was dedicated for model training and the second group consisted of 1/3 of the samples being left for testing the model. All the age-CpG were used in the beginning of the analysis, and this provided a correlation value marked by R^2 based on the contribution of each age-CpG. This value was then used to select only the top 10 age-CpGs and these were using to model the epigenetic clock. Another clock was also developed using only age-CpG sites contained in the promoter region of *FHL2* due to these having high correlation coefficients as well as being only from two pyrosequencing assays. After the models were developed these were tested against the reserved samples left for testing and the age error was determined from this.

2.7 Gene expression quantification and analysis

2.7.1 cDNA synthesis.

After the RNA was extracted as described in section 2.2.2 and 2.2.5 , cDNA was synthesised by reverse transcription of 400-1000 ng of RNA that presented >1.6 A260/280. A DNaseI treatment was first performed to remove any traces of DNA contamination using 0.5 µL DNase TURBO DNase and 0.5 µL 10X Reaction Buffer (Invitrogen) for 30 minutes at 37°C. After the incubation the DNaseI was inactivated by the addition of 0.9 µL 100 mM EDTA and incubation at 75°C for 10 minutes. For the synthesis of cDNA from RNA the SuperScript™ II First-Strand Synthesis System kit (Invitrogen, Cat # 18080051) was used as follows. Following the inactivation step, per reaction 1 µL 10mM dNTPs (Invitrogen) and 1 µL random primers (Invitrogen) at concentration of 3 µg/µL was added and incubated at 70°C for five minutes. The next step was the addition of 4 µL 5xReverse Transcriptase first strand buffer (Invitrogen), 4 µL 25 mM $MgCl_2$

(Invitrogen), 2 µL 0.1 Dithiothreitol (DTT) (Invitrogen) and 1 µL RNaseOUT (Invitrogen) and finally 1 µL of 3x Reverse Transcriptase SuperScript II enzyme was added and the solution was mixed by pipetting. The reaction was incubated in a a Veriti Fast Thermal Cycler (Applied Biosystems) with the following programme: 10 minutes at 25°C, 50 minutes at 42°C and 15 minutes at 70°C. The cDNA was diluted 1/20 in DEPC H_2O and stored at -20°C until further use.

2.7.2 Quantitative Real Time Polymerase Chain Reaction.

Gene expression levels were quantified using quantitative Real Time PCR (qRTPCR) method with the PCR primers being designed for the genes of interest using the Universal Probe Library tool (Roche) by selecting the homo sapiens target organism and the adding the target gene name. Assays with the highest scores were selected where the sequence of the primer oligomers and the number of the Roche Library probe were provided, and the primer oligomers were ordered through IDT. Primer oligomers were resuspended in DEPC water to give a 100 mM solution and further diluted to 10 mM for the PCR reaction. The master mix reaction components for the qRTPCR reaction were: 5 µL of TaqMan Fast Advanced Universal Master Mix (Applied Biosystems, Cat#4444555), 0.2 µL of 10 µM Forward primer, 0.2 µL of 10 µM Reverse primer, 0.1 µL of the Roche Library probe, 2 µL of DPEC H2O and 2. 5 µL of template cDNA, with the total volume of the reaction being 10 µL. Gene expression was quantified using qRT-PCR using a Quant Studio 3 (Applied Biosystems, by ThermoFisher Scientific) machine using the following cycling conditions; cycling conditions: 1 step at 95°C for 20s, followed by 45x cycles at 95°C for 1 s and 60°C for 20 s. Each sample was measured in a technical triplicate per gene on the same plate and only samples where the difference between the three technical repeats was ≤0.5 cycle threshold (Ct) units were included in the analysis.

Gene expression was normalised to the house keeping genes 18S, GAPDH and HRPT1 using the equation:

$$Gene\ expression = 2^{-\Delta Ct(Gene\ of\ interest) - \Delta Ct((GAPDH + HPRT1 + 18s)/3)}$$

Samples where the housekeeper average was ≥25 Ct units were excluded from the analysis.

The IDT prime time XL qPCR assay for housekeeper genes *GAPDH* (Hs.PT.5840035104, IDT), HPRT1 (Hs.PT.58v.45621572, IDT) and for *18s* (FP: CGAATGGCTCATTAAATCAGTTATGG, RP: TATTAGCTCTAGAATTACCACAGTTATCC, IDT probe: 6-FAM/ TCCTTTGGTCGCTCGCTCCTCTCCC/ ZEN/IBFQ), was used with the other genes assays being designed using Roche Universal Probe Library software (Roche). The qRT-PCR primers used are listed in Table 2. 6.

2.7.3 RNA-sequencing and bioinformatics analysis

RNA samples from the SW1353 cell lines that were FHL2 knock down (section 2.5.5) or overexpressing of FHL2 (section 2.3.3) and their controls were treated with DNase using the DNA-free™ DNA Removal Kit (Invitrogen, Cat#AM1906). Approximately 10 µg of RNA in a 50 µL volume reaction was treated with 1 µL DNase I and incubated at 37°C for 30 minutes. After this step, the RNA samples were purified as specified by the manufacturer's instructions using 5 µL resuspended DNase Inactivation Reagent (Invitrogen). The purified RNA samples were submitted to a quality check on an Agilent 4200 Tapestation where all samples had a RIN score = 10 with concentrations of >150 ng/µL.

For the library preparation around 1µg RNA per sample was processed using the Illumina® Stranded mRNA Prep Ligation kit then submitted for sequencing with the reads per sample at $20x10^6$ reads and with read length of 75 bp per sample on an Illumina NextSeq500 machine at the Genomics Core Facility (Newcastle University). After receiving the sequencing results, the fastq files were inspected using the MultiQC tool to determine the sequence depth and quality as well as for sequence nucleotide score to determine whether any trimming is required. No trimming was performed on the samples as they all showed a quality score of >30. Salmon was used for alignment using the human Gencode V38 index where two extra chromosomes were added to the index files, containing the sequence for the Cas9 and GFP gene. The alignment results showed that >92% of the read aligned to the Gencode V38 index. The data was then imported into RStudio using the TxImport library where gene counts and gene transcript per million (TPM) levels produced from the transcript levels generated through Salmon were

compiled for all the samples. Genes that showed a value ≥2 TPM in at least a condition within all the replicates were then selected for differential analysis using the DeSeq2 library against the respective control condition, followed by shrinking using the ashr library. The data was then annotated using the biomaRt library ensembl version 104 homo sapiens gene database. Based on the differential analysis, genes that showed a significant p adjusted value ($p<0.05$) and a log2FC>0.585 (FC>1.5) were considered upregulated and a log2FC<-0.599 (FC<0.66) were considered downregulated. Gene Ontology enrichment and KEGG enrichment analysis was then performed on the list of upregulated and downregulated genes. Graphical representation was designed in either RStudio version 1.4.1717 or GraphPad Prism version 9.31. KEGG pathview library was used for the analysis and visualisation of all the genes that had a significant p value and a TPM>2 in the respective condition comparison in order to test for pathway changes based on transcriptomic changes.

2.8 Protein analysis.

2.8.1 Protein extraction from cell lines

Cell pellets of approximately $1x10^6$ cells were collected and then resuspended in 50 µL self-prepared lysis buffer for whole cell protein with incubation on ice for 20 minutes. Following this the cells were centrifuged at 16200xg for 15 min at 4 °C and the supernatant was transferred to a new 1.5 ml Eppendorf tube. This lysate was stored at -80°C until further use.

- Lysis buffer whole cell protein:

Tris Base 50 mM, NaCl 150 mM, NaF 5 mM, EDTA 1 mM, Triton x-100 1%, Glycerol 10%, complete Mini EDTA-free Protease Inhibitor Cocktail (Roche, Cat#4693159001) 1 tablet.

2.8.2 Bradford assay for protein quantification.

The whole cell protein lysate was quantified using the Bradford assay (Sigma Aldrich, Cat#B6916) and a bovine serum albumin (BSA; Fisher Scientific, Cat#BP1605-100) protein standard from 40-400 µg/ml was prepared. Absorbance readings were taken at 595nm and

450nm on a Varioskan LUX Multimode Microplate Reader (Thermofisher) followed up by taking the ratio of the absorbance values as 595/450 and the generation of a standard linear regression equation from which the samples concentration was interpolated. All samples and standards were quantified in duplicate and the mean values used.

2.8.3 Western blotting.

The quantified whole cell protein lysated were then diluted to 1µg/µL using DEPC water and with 5x of self-prepared Laemmli buffer (appendix section); with each sample adjusted to contain the same amount of Laemmli buffer. The prepared sample (approximately 10-20 µg of protein per well) was then loaded on either self-made acrylamide gels of 8-12.5% acrylamide (appendix section) or precast NuPage Gradient 3-8% Tris-acetate gels (Invitrogen, Cat#EA0375BOX) concentration together with a HiMark pre-stained Protein Standard ladder (Invitrogen, Cat#LC5699) or Pageruler prestained protein ladder (Thermofisher Scientific Cat#26616). The gels were ran using running buffer (appendix section) at 120 V for 1.5-2 hours, after they were transferred onto a PVDF membrane (Fisher Scientific, Cat# 88518) using a semi-wet transfer technique performed at 90 mA for 1.5 hours using a self-made transfer buffer (appendix section). The membrane was blocked in 5% w/v milk for one hour and incubated with primary antibody overnight at 4°C according to the manufacturer's instructions. Following this, it was incubated with the secondary antibody in 5% BSA in TBS-Tween for one hour at RT. Visualisation was performed using the Immobilon Western Chemiluminescent HRP Substrate (Millipore, Cat# WBKLS0500) using an Azure Biosystems c600 Imager. After the blots were imaged, the membrane was re-probed with either an anti-beta-tubulin or anti-GAPDH antibody as a loading control. Primary and secondary antibody (Table 2. 9) incubations were performed as above. This was then visualised using the low-sensitivity ECL™ Western Blotting Reagent (Sigma, Cat# GERPN2106) as above. Imaging was also performed using infrared-fluorescent tag IRDye 800CW and IRDye 680CW secondary antibodies (Li-Cor, Cat#P/N 925-32211, Cat#P/N 925-32210) at 1:10000 dilution in 5% BSA in TBS-Tween for the respective primary antibody and then visualised on a Li-Cor Odyssey DLx machine.

- Laemelli buffer 5x:

Tris Base 250 mM (Sigma Aldrich), Glycerol 20% (Sigma Aldrich), Sodium dodecyl sulfate 10% (Sigma Aldrich), Bromophenol blue 0.02% (Sigma Aldrich), 2-mercaptoethanol 0.8% v/v (Sigma Aldrich).

- Western blot gel:

Gel: 8-12.5% Bis:Acrylamide (Sigma Aldrich), Sodium dodecyl sulfate 0.1% (Sigma Aldrich), Tris 0.38% (Sigma Aldrich), 0.1% Ammonium Persulphate (Sigma Aldrich), 0.17% TEMED (Sigma Aldrich).

- Stacking gel:

4.5% Bis:Acrylamide (Sigma Aldrich), Sodium dodecyl sulfate 0.1% (Sigma Aldrich), Tris 0.13% (Sigma Aldrich), 0.1% Ammonium Persulphate (Sigma Aldrich), 0.17% TEMED (Sigma Aldrich).

- Running buffer:

25 mM Tris (Sigma Aldrich), 0.192 M glycine (Sigma Aldrich), 0.1% (w/v) Sodium dodecyl sulfate (Sigma Aldrich).

- Transfer buffer:

39mM glycine (Sigma Aldrich), 48mM tris base (Sigma Aldrich), 1.127mM Sodium dodecyl sulfate (Sigma Aldrich) and 20% methanol (Sigma Aldrich).

2.8.4 Immunocytochemistry.

The cells were seeded at a density of $1x10^4$ cells/well in an eight-well glass chamber slide (LabTek, Cat#154534) and incubated for 24 hours. The cells were fixed with 10% Neutral buffered formalin (Sigma, Cat#MKCG6196) for 10 minutes at RT, in PBS and permeabilized in

0.5% (in PBS) Triton X-100 (Sigma, Cat#SLBR3411V) for five minutes at RT and blocked in 10% (in PBS) goat-serum (Vector Laboratories, Cat# S-1000) for one hour at RT. After, they were incubated primary antibody at the dilution recommended by the manufacturer (see table below) in PBS for two hours at RT followed by secondary antibody at a dilution factor of 1:500 in 10% Goat serum (Vector Laboratories, Cat# S-1000) for two hours at RT and mounted in VECTASHIELD® Antifade Mounting Medium with DAPI (Vector Laboratories, Cat#H-1200-10).

Images were acquired using a Zeiss Axioplan 2 microscope using a CMOS 1300x1030 pixel detector camera and 10x (N.A. 0.3) objective and 20x (N.A. 0.5) objective and on a Nikon A1R confocal microscope using the 20x (N.A. 0.75) air objective, 40x (N.A. 1.3) oil objective and 60x (N.A. 1.4) oil objective. The configuration details of the channels used for signal acquisition for confocal microscopy are shown in Table 2. 5.

Table 2. 5. Configuration details of the channels used for signal acquisition for confocal microscopy.

Channel	Excitation λ (nm)	Emission λ (nm)	Pinhole size (μm)
DAPI	405.9	>450	24.27
Green (FITC)	488	>525	24.27
Red (TRITC)	561.5	>595	24.27
Cy5	640.2	>700	24.27

2.8.5 FHL2 cellular distribution and co-localisation of FHL2 with actin and mitochondria and analysis.

The intracellular distribution of FHL2 was assessed using immunofluorescent confocal microscopy together with the co-localization with actin and mitochondria.

SW1353 cells were seeded at $1x10^4$ cells/100µL/ well in CELLview™ Slides (Greiner Bio-One, Cat#543079) and were incubated for 24 hours at 37°C 5% CO_2. The cells were fixed in 10% formalin and the method described in the immunocytochemistry section was followed where they cells were incubated with anti-FHL2 primary antibody and either Anti-TOM20, Anti-beta-Actin or Phalloidin-AF488 stain (Table 2. 9), followed by respective fluorescently tagged secondary antibody. The slides were imaged in a Nikon A1R confocal microscope fitted with a photo-multiplier tube detection sensor at a resolution matrix of 4096x4096 pixels with a resolution of 0.11µm/pixel and a 40x oil objective N.A. = 1.35.

Co-localization analysis between FHL2 signal an either beta-actin or mitochondria was done in ImageJ where the coloc2 function was selected with the parameters of selection being a mask made by getting firstly the cytoplasmic borders from a brightfield image followed by selecting the signals for FHL2 and the other factor for co-localization. Pearson correlation and Manders coefficient were then used to determine the extent of co-localization of the two signals. The Pearson correlation looked at the total signal for both channels as pixel intensity and position to check for a correlation relationship. The Manders coefficient looked at the signal pixel position corresponding to either the beta-actin or mitochondrial signal and used this to determine whether there is any signal above a threshold found for FHL2.

2.8.6 Immunohistochemistry.

Knee tissue samples of mice aged seven days, four months and one year old, were formalin fixed and embedded in paraffin by Hua Lin (Young group). These were then deparaffinized in xylene and rehydrated in a series of decreasing ethanol solutions. A solution of 10mM Citrate buffer pH 6.0 was used for antigen retrieval with the samples being heated in a microwave oven for 6 minutes. The samples were afterwards incubated with a 0.3% v/v H_2O_2 in PBS for 30 min followed by incubation with 10% blocking serum and avidin in PBS for 1 hour at room temperature. The sections were then incubated with Anti-FHL2 (Abcam, Cat#ab12327) 1:100 dilution in PBS (10 µg/ml) at room temperature. The samples were washed in PBS and then

incubated with a secondary peroxidase conjugated anti-rabbit antibody for two hours at RT. Immunostaining was developed then using 3,3'-diaminobenzidine tetrahydrochloride (DAB) substrate kit (Vector Laboratories) and counterstained using Mayer's Haematoxylin for 2 min incubated at room temperature. The samples were then dehydrated in an increasing series of ethanol solutions followed by clearing with xylene and mounted in Permamount. Light microscopy was used to visualize the slides on a Zeiss AxioVert 2 using a 5x objective N.A. =0.15 and a 10x objective N.A. =0.32 with image acquisition using Zeiss Axiocam HRc camera module.

Slides containing samples from sequential sections where staining was performed with Haematoxylin and Eosin (H&E) and Safranin-O and Fast Green were also received in order to differentiate the different tissues present in the sections. This was received from Hua Lin.

2.9 Luciferase assay.

2.9.1 PCR amplification of region of interest

PCR primers were designed to the 500-1000bp region surrounding the age-CpGs of interest using Primer 3 software (http://primer3.ut.ee/) with restriction enzyme sites were added at the 5' and 3' ends of the construct for the enzymes SpeI, BamHI / HindIII. The promoter regions containing the age-CpG region for: FHL2, ELOVL2, ZNF577, and ZNF551 together with the transcription start site (TSS), were cloned into the pCpGL basic vector (Klug and Rehli, 2006). The putatitive enhancer regions containing age-CpGs PCDBH1 and PCDHGA1 were cloned upstream of a CpG free EF1 minimal promoter version of the pCpGL basic vector (produced by Dr. Louise Reynard). The primers used for the amplification are found in (Table 2. 6). Due to the sequence's increased density of cytosine and guanine within the regions surrounding FHL2 and ELOVL2 this presented a challenge for PCR amplification from genomic template DNA. Several pairs of primers were tried with no success and due to this issue, the gene synthesis service (IDT) was employed to produce synthetic double stranded DNA of the region of interest. To the sequences for synthesis (Table 2. 1) at the 5' and 3' ends restriction sites were added for the restriction enzymes SpeI and HindIII.

Table 2. 6. Summary details of the regions where the gene synthesis service was used. The blue highlighted text marks the position of restriction sites for SpeI and HindIII and in red the position of CpGs that were analysed by pyrosequencing.

Custom gene synthesis service IDT	
Region name	**Sequence 5'-3'**
FHL2	GGGGACTAGTCCGTAGTGGAACGAAGGATTCTTTATGGTCAT CTGTGGTGTTCATTTCGTAAAGAAGCATCTGCAGGGTAAGAA GGAAAAACGTCCACCTTGCAAATATATCCCAGGCACATGCCTC CTGAGAAGTGACCCCCTCCTCCCTCCGCCTCCCCCGGCACGTC CTGGGGCTTCTCCAGTCTCCCGCTCCTGGGACCAGGCAGAGA TCCCGGCGTGGGCAGACCCCTGCCACCGCGCCCAGGCCTCGT CCGAAACTCGCGCCCGAGAGCCCACGCCGGAGGCGACGCTCC CGATAACTACTGTGCTCCCAAGACCCGGAGGGCACGAAGAAA GGAGCCCTGGCAAACAAAGGGTACGGGCCGGGACCGCCGCA GCCCGGGGCGGGGGCACGGCAAAGCTTCCCC
ELOVL2	GGGGACTAGTATTCGCCTGCCTCAGCCTCCCAAAGTGCTGGG ATTACAGGAGGGAACCACCACGCCCGCCAACTTCCCATGCTT GAGGGAGAAATGGAAGAAAGTTCATGTAATACTCAGGCAAG TCCAATTTTTTCGACGTCTTTCACTTGGGCCACACACACAACTA AAGTAACTAGAAGCGCAGGCTCTAGGAGGCCACCGTTCTGTT CACAGTGAAGAGGGTGCGCTCACCGTTGGTCGTGTCCGCTGG AAGCCCCGCGTCAGGCCGGGAGCGGGACAGAGACTCTTGCTC AGGGCCGTTATCCGAACTGATCCGCTTCCCACCGCACCCCCAG AGAAACCCACCCAACCCCCTAAACCTAAGAAACCCAGACTGC GCAAACCTGCAGGAACAGAGCCATTTCCCCCTAATGTGTGCTT CAAACCCACCGAAGCCCAACTGTAAGCAAGACCAGCGTGCCC GCCCTGCACGATACTGCTTCTCCCCGCAGCAGCGGCTGCCGAT CTGGGCAGCGGGTGGGTATTCCTGGGGCTCCGTGGACGTTGA GCCGCCGCGCGAAACCGGCGCCGGCTGGACCTGCAAATCGCC GCCCGGCCGGCAGGGGACGCCGCGGACAAGCTTCCCC

The DNA regions of interest were cloned from a template DNA of pooled blood genomic (for PCDHGA1, PCDHB1, ZNF577 and ZNF551) DNA using the following mastermix: 0.3 µL Phire Hot Start II DNA Polymerase (ThermoFisher Scientific, Cat# F122S), 3 µL Phire Taq PCR buffer, 0.3 µL of 10 mM dNTP, 0.3 µL 10 µM Forward primer and 0.3 µL 10 µM Reverse primers, 9.8 DPEC H2O, 1 µL blood genomic DNA template, in a total reaction volume of 15 µL. The cycling

conditions used: A first step hot start at 98°C for 30s followed by 35x cycles of denaturation at 98°C for 5s, annealing at x°C for 10s, elongation at 72°C for 15 s and a final step at 72°C for 1 min. The optimal annealing temperature was determined experimentally using a temperature gradient program for each combination of primers. After the PCR, the product was purified using a QIAquick Gel Extraction kit (Qiagen, Cat# 28704) following the manufacturer's instructions with this being eluted in 30 µL of DEPC water. The purified product was then stored at -20°C.

After the PCR and gene synthesis service, the samples were purified using a QIAquick Gel Extraction kit (Qiagen, Cat# 28704) following the manufacturer's instructions and this was eluted in 30 µL of DEPC water.

2.9.2 Digestion of insert and plasmid.

The purified PCR products together with 10 µg of the empty pCpGL vectors were then digested with 20 units of each of the restriction enzymes (New England Biolabs) indicated for each primer in the above table, for 16 hours (overnight) at 37°C in a total volume of 40 µL. After this digestion, the PCR products were again purified using the QIAquick Gel Extraction kit and the digested plasmids were run on an 1% agarose TAE gel and gel purified using the same kit according to the manufacturer's instructions. The PCR construct and respective plasmid were combined in a molar ratio of 3 :1, 3x parts of PCR construct DNA (~90 ng) to 1x part digested plasmid (~200 ng) based on the DNA concentration and ligated overnight at room temperature using a T4 DNA ligase (New England Biolabs) following the manufacturer's instructions in a 20µL total reaction volume.

The 1µL of the ligated plasmid solution was transformed into 20µL GT115 *E. Coli* Chemically competent cells (Invivogen) following the manufacturer's instructions. This mixture was then plated on agar plates with Zeocin (Invivogen, Cat#ant-zn-05) at a concentration of 25µg/ml as

30µL and 90µL of the reaction and cultured overnight at 37°C. Clones were picked, grown up overnight at 37°C and 200 RPM in 5 ml LB containing Zeocin at (Invitrogen) (25µg/ml) then miniprepped as previously described in section 2.5.6.

2.9.3 Screening of clones and sequencing.

Miniprepped clones were screened for the insertion by restriction enzyme digestion using the respective restriction enzymes for each primer sequence (Table 2. 6). The clones that were positive for the insert, genomic DNA ~200 ng was sent for Sanger sequencing using the sequencing service (SourceBioscience Cambridge) with the sequencing primer: 5'-TAAATCTCTTTGTTCAGCTCTCTG.

2.9.4 In vitro DNA methylation.

In vitro methylation was carried out using the following master mix: 2 µL 20000 U/ml M.SssI (New England Biolabs, Cat#MO226M), 25 µL 3.2 mM S-adenosyl-methionine (SAM), 12 µL NEBuffer™ 2 and 10 µg of plasmid DNA, with DPEC H_2O being added to bring the total volume to 120 µL. A mock methylation was set as a control, following the same mastermix recipes as listed above but replacing the M.SssI with DPEC H_2O. The reaction was incubated at 37°C for 16 hours. After 16 hours, a further 2 µL 20000 U/ml M.SssI (New England Biolabs, Cat#MO226M), 4 µL 3.2 mM SAM, 2 µL NEBuffer™ 2 with DPEC H2O being added to bring the total volume to 140 µL were added and incubated for 4 hours at 37°C. After this, the reaction was inactivated by heating at 65°C for 20 min. The DNA was the purified using the Wizard SV gel and PCR clean-up system (Promega, Cat# A9281) following the manufacturer's instructions.

Methylation status was confirmed using a combination of methylation sensitive restriction enzymes, Hae II and Hpa II (New England Biolabs, Cat# R0107L and R0171) with the reaction being deemed positive only when no digestion occurred on the methylated plasmid after 2 hours of incubation at 37°C. After confirming the successful methylation status of the plasmids, the concentration of all DNA samples was adjusted to 50 ng/µL (see fig X as an example).

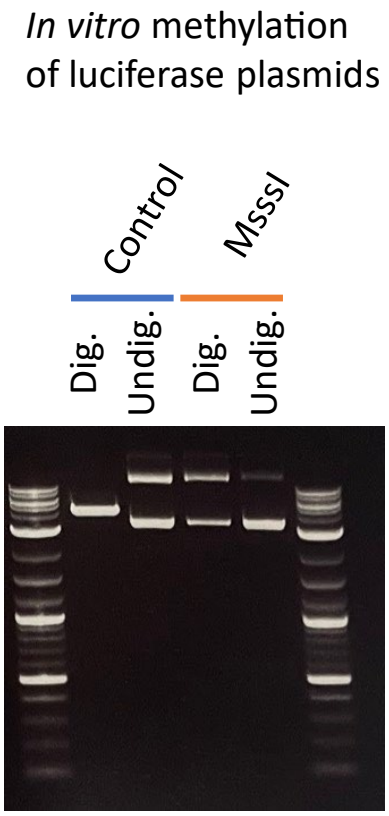


Figure 2. 3. Gel electrophoresis image showing the digestion with the methylation sensitive restriction enzymes Hae II and Hpa II of MsssI methylated and unmethylated (control) luciferase vectors.

2.9.5 Transfection into mammalian cell line and luciferase activity determination.

The SW1353 human chondrosarcoma cell line was seeded at a density of $5x10^3$ cells/ well, in a 96 well plate in DMEM/F12 medium. After 24 hours from the seeding, the cells were transfected with 100 ng/well of plasmid constructs and 6ng/well of *Renilla* luciferase reporter vector (Promega) using 0.3 µL/well FuGENE® HD Transfection Reagent (Promega, Cat# E2311). After an additional 24 hours, media was removed, and the cells were washed in sterile PBS then lysed with 30 µL 1x Passive Lysis Buffer (Promega) according to the manufacturer's instructions followed by storage overnight at -80°C. Following this, 10 µL of lysate was transferred to a luciferase plate and the luciferase activity was determined using the Dual-Luciferase® Reporter Assay System (Promega, Cat. # E1910) on a GloMax-Multi Detection System (Promega).

The following settings were used to each well 40µL of Luciferase Assay Reagent II was added (200 µL/sec speed) followed by 0.5 sec delay and acquisition of the signal with an integration time of 1 sec for the luciferase signal. This was then followed by the addition of 40µL of Stop & Glo® Reagent followed by the same parameters of acquisition for the renilla signal.

2.9.6 Normalisation and statistical analysis for luciferase assay data

The Firefly luciferase signal was normalized against the signal of the Renilla luciferase, after which the data were normalized against the empty vector respective to each plasmid. Six technical repeats and up to six biological repeats were performed for each *in* vitro methylated or unmethylated plasmid construct. Statistical analysis was performed using the Mann Whitney test to compare ranks due to the luciferase data being non-parametric (Kolmogorov-Smirnov test for normality).

2.10 Electrophoretic mobility shift assay (EMSA).

2.10.1 Probe design and preparation.

The region containing the age-CpGs present in the promoter of *FHL2* were selected for investigation using EMSA with the region being divided into two assays. The regions were designed as forward and reverse primers as fully methylated and unmethylated probes using the oligo design service from Eurofins Genomics (Germany) and ordered through the oligomer synthesis service from Eurofins Genomics (Germany), these being tagged at the 5'end with DY-682 fluorophore. Upon received the oligomers, these were resuspended in DEPC water according to the manufacturer's instructions, giving a stock solution of 100 pmol/µL. The probes were annealed by adding 10 µL of each respective forward and reverse oligomer per assay, 10 µL EMSA annealing buffer and 70 µL DEPC water and heating to 95°C for five minutes followed by cooling to RT. The annealed primer solution was further diluted to a concentration of 100 fmol/µL in DEPC water and stored at -20°C until further use.

2.10.2 Nuclear protein extraction from cell lines

Nuclear protein was extracted from SW1353 chondrosarcoma cells, TC28a2 immortal rib chondrocyte cells, U2OS osteosarcoma cells, MDA-MB-231 epithelial breast cancer cells using a self-made 1ml per pellet of hypotonic buffer for 15 minutes on ice followed by centrifugation and the retaining of the pellet. This was followed by the addition of 200 µL hypertonic buffer incubating for 30 min on ice. After centrifugation the supernatant was retained and quantified using the Bradford assay. The U2O2 and MDA-MB-231 nuclear extracts were received from Dr. Reynard. The optimal binding reaction contained 1× Binding Buffer, 2.5 mM DTT, 1 µg polydIdC, 200 fmol annealed probe oligonucleotide, 5 µg nuclear extract, and either 2.5 mM $MgCl_2$+NP-40 for P1 or 5 mM $MgCl_2$ for P2 variant.

2.10.3 Gel preparation and running conditions.

A 0.5x TBE (Tris-Borate-EDTA) polyacrylamide gel was prepared and once solidified was run for 30 min at 100V to reduce the traces of the polymerizing agents. The reactions were then carried out according to the manufacturer's instructions using the Li-Cor Biosciences ODYSSEY EMSA buffer kit (Cat#11502187, Li-cor Biosciences) and the gel was ran at 100 V and a max of

8mA for 4 hours at 4°C in 0.5x TBE buffer, followed by imaging on the Odyssey DLx Imaging System (LI-Cor).

- EMSA buffer:

Tris pH 7.6 (0.065M)-Borate (23 mM)-EDTA (1.3 mM) (Sigma Aldrich).

- Hypertonic buffer:

20 mM HEPES (Sigma Aldrich) pH 7.9, 420 mM NaCl (Sigma Aldrich), 20% glycerol (v/v) (Sigma Aldrich), 1 mM DTT (Sigma Aldrich), 10 mM NaF (Sigma Aldrich), 1 mM Na3VO4 (Sigma Aldrich), 1× complete protease inhibitor cocktail tablet (Roche) per 50 ml of buffer.

- Hypotonic buffer:

10 mM HEPES pH 7.6 (Sigma Aldrich), 1.5 mM MgCl2 (Sigma Aldrich), 10 mM KCl (Sigma Aldrich), 1 mM DTT (Sigma Aldrich), 10 mM NaF (Sigma Aldrich), 1 mM Na3VO4 (Sigma Aldrich), 0.1% Triton-X (v/v) (Sigma Aldrich), 1× complete protease inhibitor cocktail tablet per 50 ml solution (Roche).

- Annealing buffer:

100 mM Tris HCl pH 8 (Sigma Aldrich), 500 mM NaCl (Sigma Aldrich), 10 mM EDTA ph 8 (Sigma Aldrich).

2.10.4 Identification of and deign of competitors for Transcription factors that potential bind the FHL2 EMSA probes.

The following online webtools were used to obtain information about potential transcription factor binding that could occur at the site of the FHL2 age-CpGs.

Tfbind (http://tfbind.hgc.jp/),

Alggen (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3),

Jaspar (https://jaspar.genereg.net/),

MEME suite (https://meme-suite.org/meme/),

TFsitescan (http://www.ifti.org/Tfsitescan/),

LASAGNA (https://bio.tools/lasagna-search)

UCSC TFBS ChIPseq tracks (Raney *et al.*, 2014)

Based on the results from these a list was generated with the potential transcription factors that could bind each probe region based on consensus motif sequence matching together with the possible position this might occur at. From this list, the common transcription factors with at least two incidents between the results from the different tools were selected to be further studied. Competitor oligos were designed corresponding to the consensus motif sequences of the transcription factors identified by introducing the consensus competitor sequence at the identified position into a random non-competitive sequence that was determined experimentally to not show any form of competition with the protein complex bands identified for each region. A list of the competitor sequences can be found in Table 2. 7.

2.10.5 Antibody supershift assay.

Supershift EMSAs reactions were performed using 2 µg of EMSA-grade antibody (antibody list in Table 2. 9) with this being added to the binding reaction at the same time as the nuclear extract and incubated as recommended by the kit manufacturer. The impact of the supershift was assess by visual inspection.

2.10.6 Competitor design and competition assay.

Oligo competitors fully methylated or unmethylated were designed for the assay having the exact sequence as the probes with the exemption of the 5'fluorescent tag. These were annealed as previously described for the EMSA probes (section 2.10.1) and diluted accordingly as to give a concentration of 100 fmol/µL in DEPC water. These were incubated with the probe in the normal reaction conditions with increasing concentrations of competitor at: 50, 100, 250, 500, 1000 nM.

2.10.7 Densiometric analysis of EMSA image results.

After acquisition on the Licor machine, the resulting TIFF images are analysed in ImageJ as follows. The images were opened in ImageJ (or Fiji) and the colours were inverted and the image was rotated horizontally at 90 ° left. Then the rectangle drawing function was selected and a rectangle was drawn per lane of at least 5 pixels height and length that spawns all the visible bands, including the ones generated by the probes and the upper one from near the wells. Signal acquisition was made using the "Plot Profile" function the results were then exported to Excel and GraphPad Prism 9.3.1 for further analysis. Care was taken during acquisition exposure as not to have oversaturated pixels of the 8-bit sensor (255 max value). A background signal measurement was also performed to determine the background value that will be subtracted from the value of the bands of interest. The peak of each band was found by looking at the maximum signal value where the band was present based on the pixel distance from the upper band (near the wells). The values from this peak with ±5 pixels were summed up, representing the signal from that particular band. The background signal was then subtracted from the band signal, and afterwards the values were normalised to the band intensity found in the control unmethylated condition. Due to the variability in the distance travelled between the different gel replicates a method of standardisation was developed using the upper band near the well (representing the start as 0) and the band generated by the free probes (representing the maximum distance as 1). Therefore, each protein band that binds in between these reference bands will have a value between 0-1 with the scale being called Retention factor (Rf). Each result was replicated three times for the competition assay and twice for the mutation assay, consensus assay and supershift assay.

To determine the impact of competition on band intensity, the values for the controls (no competition) and that of the oligo competitors were analysed using the "Binding-Saturation – One site Total" function in GraphPad Prism 9.3.1 to perform affinity calculations per each band. This generated a disassociation constant value (Kd) that reports the concentration of competitor required to reach 50% saturation, with lower Kd values showing a greater affinity for the substrate. Further analysis was performed to determine the impact of the decrease of

signal with the increase in the concentration of the competitor or type of competitor using one-way ANOVA with multiple comparison against control and multiple test correction against hypothesis (Šidák correction) and two-way ANOVA with multiple comparisons against the mean of each row and multiple test correction against hypothesis (Šidák correction) respectively.

2.10.8 Tables

Table 2. 7. Primers

Use	Target CpGs & nearby gene	Region coordinates hg19	Primer orientation	Primer sequence		
Bisulphite pyrosequencing	cg22331349, cg09547119 (*ZNF577*)	chr19:52391092-52391663	FP	5'-[Btn]GGGAAGTTTGTTGGGAGTAGTTATTT		
			RP	5'-ATATTACAAAACCAAAATCTAACAATTCAC		
			Seq P	5'-TTCATCCTCAAAACCTACCTATC		
Bisulphite pyrosequencing	cg01644850 (*ZNF551*)	chr19:58192773-58193537	FP	5'-TTTTAGAAATTTGGGAAAAGT		

			RP	5’-[Btn]ATCCTAAACTATCATTAACTCAACTC		
			Seq P	5'-AGAGGTTAAAGTAGTGG		
Bisulphite pyrosequencing	cg06639320 (*FHL2*)	chr2:106015624-106016026	FP	5'-TGTTTTTAGGGTTTTGGGAGTATAGTAG		
			RP	5’-[Btn]ACACCTCCTAAAACTTCTCCAATCTCC		
			Seq P	5'-GTTTTGGGAGTATAGTAGT		
Bisulphite pyrosequencing	cg22454769, cg24079702 (*FHL2*)		FP	5’-[Btn]GTTTTTAGGGTTTTGGGAGTATAGTAGTT		
			RP	5'-ACACCTCCTAAAACTTCTCCAATCTCC		
			Seq P	5'-ACCACCCCCCCCAAACCT		
Bisulphite pyrosequencing	cg16867657, cg24724428,	chr6:11044815-11045424	FP	5’-[Btn]AGGGGAGTAGGGTAAGTGAG		

	cg21572722 (*ELOVL2*)		RP	5'-ACCATTTCCCCCTAATATATACTTCA		
			Seq P	5'-TAAAACTCCR TAAACRTTAAACCRCCRC		
Bisulphite pyrosequencing	cg23500537 (*PCDHB1* enhancer)	chr5:140419365-140420157	FP	5’-[Btn]GAAGTTGAGGAAGGTGTTGTG		
			RP	5’-CCCCTCTACAACCCCAAAC		
			Seq P	5’- ACACTCTACCAAATCCTA		
Bisulphite pyrosequencing	cg14566959 (*PCDHGA1* enhancer)	chr5:140772371-140772999	FP	5’-GGTAGGATAGATAGGGAGGAGTTTTG		
			RP	5’-[Btn]ACCCCAAAAAATTTTCCTTTATCCTC		
			Seq P	5’-GAGTTTTGAGTTTAGAGT		
Bisulphite pyrosequencing		chr8:27850326-27851223	FP	5' -TTG GGTTTGGGGGATTTAGAAT		

	cg07634191 (*SCARA5* enhancer)		RP	5' -[Btn]ACCTCCCCAACTACTCCAAAAACTTTAT		
			Seq P	5' -TGAGATGAGAATTGTAGTTGTT		
Bisulphite pyrosequencing	cg13144059 (*SPATS2L* enhancer*)*	chr2:201245124-201246184	FP	5' -AATTTAAAGTTGGGATTGTTATAAAGTG		
			RP	5' -[Btn]TACAAATTTAAAACCTTCATAAAAA		
			Seq P	5' -TGGGATTGTTATAAAGTGT		
Bisulphite pyrosequencing	cg06926735 (*TMEM189* enhancer)	chr2:48732802-48733705	FP	5' -[Btn]TGGAGTAGTGTTTGATGGGAGTTGT		
			RP	5' -AACTCTTATCCCATCAACTAAATTTATC		
			Seq P	5' -CCCATCAACTAAATTTATCA		

Bisulphite pyrosequencing	cg06217736 (*DUSP8* enhancer)	chr1:1579494-1580450	FP	5' -[Btn]GGGTTAGGTGAGGGTTAAGATTG		
			RP	5' -ACCCCAAACCTAACTTCATCT		
			Seq P	5' -CCCTAACTAAACAAATCCA		
Bisulphite pyrosequencing	cg07477282 (*SPG11* enhancer*)*	chr1:44956155-44957246	FP	5' -TTTTGGTATGTTTGAAGGTTTAGTTTAG		
			RP	5' -[Btn]AAAAAAATAATATTATTTTAAAATTTATAA		
			Seq P	5' -AGTTTAGAGTAGTAGGT		
Bisulphite pyrosequencing	cg18569335 (*NKIRAS2* enhancer)	chr1:44956154-44957239	FP	5' -TTTGGTATGTTTGAAGGTTTAGTTTAG		
			RP	5' -[Btn]TAACACCACCCCAAAACCTATA		
			Seq P	5' -AGTTTAGAGTAGTAGGT		

Use	Region Name	Region coordinates hg19	Primer orientation	Sequence	Restriction site	
Luciferase PCR assays	PCDHB1	chr5:140419365-140420157	FP	5' GGGGGGATCCATTGTCTCTTCCCTTGGGCT	BamHI	
			RP	5' GGGGACTAGTAGTTCAGGGGCTAGCGTTTA	SpeI	
Luciferase PCR assays	PCDHGA1	chr5:140772371-140772999	FP	5’ GGGGATCCCAAAACAACAATGGCCGCTC	BamHI	
			RP	5’ GGGGACTAGT CAGCCTCTTCCTCCCTGTC	SpeI	
Luciferase PCR assays	ZNF577	chr19:52391092-52391663	FP	5' GGGGACTAGTTAACCCCTCCGATTGTTGCT	SpeI	
			RP	5' GGGGAAGCTTGATGACCTGTCTCCACTGCT	HindIII	

Luciferase PCR assays	ZNF551	chr19:58192773-58193537	FP	5' GGGGACTAGTGCGTGATCCACCCACCTT	SpeI	
			RP	5' GGGGAAGCTTGGGCATTCGAACACAACAGT	HindIII	
Use	Assay name	Region coordinates hg19	Orientation	Sequence		
dCas9 validation BIS PCR	DNMT3a Pyro	chr9:38016355-38017662	FP	5’-[Btn]AGTAGTTTTGGTAGGGTGAGTAAGT		
			RP	5’-CTAAACCTCCCCTTAACCTCCAACC		
			Seq P	5’-ACTACCCCCTACTATCCCTT		
dCas9 validation BIS PCR	TET1cd Pyro	chr17:41278564-41279065	FP	5’-[Btn]GGGTTAGTTAGGGGTGGGGTTA		
			RP	5’-CCTCTCCCTCCACACTTC		

			Seq P	5’-CAATAACCAACTAAAAAACTCCTC		
Use	Name	Region coordinates hg19	Location relative to deleted region	Orientation	Sequence	UPL Probe #
CRISPR deletion check PCR primers	FHL2	chr2:106015842-106015916	Inside	FP	5' GAGCGGGAGACTGGAGAAG	52
				RP	5' ATGCCTCCTGAGAAGTGACC	
		chr2:106016380-106016441	Outside	FP	5' CATTCGGGGCAGCTCATA	65
				RP	5' AGCAATCACGGAGTTCTTCC	

CRISPR deletion check PCR primers	ELOVL2	chr6:110450055-11045066	Inside	FP	5' GGTTTGAAGCACACATTAGGG	21
				RP	5' AAACCCAGACTGCGCAAA	
		chr6:110457044-11045780	Outside	FP	5' TGTTTATTGGGGTTGAGAGCA	5
				RP	5' CAGGTGTGGGATGCTAAGTGT	
CRISPR deletion check PCR primers	ZNF577	chr19:52391266-52391343	Inside	FP	5' CCGTCCCTTTATGATGCAAT	32
				RP	5' CTACGGCAGTAGGGTTGAAAA	
		chr19:52392001-52392068	Outside	FP	5' AATTAACTGATGAATTCTAGCACCAA	32

				RP	5' AAAGGTTCTCTGAGGCTGGTT	
CRISPR deletion check PCR primers	ZNF551	chr19:58193182-58193302	Inside	FP	5' GAATACAAGTTCGCAGAGGTCA	86
				RP	5' GCTGTCATTGGCTCAACTCC	
		chr19:58193597-58193686	Outside	FP	5' GAGGGACTCGGCTCAGGT	3
				RP	5' AGCAGGCATCCCTCACTG	
CRISPR deletion check PCR primers	PCDHB1	chr5:140419776-140419849	Inside	FP	5' CACACATCCAAGGCTGACA	53
				RP	5' GAGGAAGGTGTTGTGGAAGG	

		chr5:140420416-140420483	Outside	FP	5' CCCAGTGCTGGTGTCTTAGG	27
				RP	5' GTTATGGCGGCCAATACCT	
CRISPR deletion check PCR primers	PCDHGA1	chr5:140772872-140772940	Inside	FP	5' GAACTCCCTCCAGAGCTACCA	7
				RP	5' TTGTCTCCAGTCTGCACGTC	
		chr5:140772365-140772458	Outside	FP	5' CTGCAGCAAAACAACAATGG	18
				RP	5' GATTTCCCACAGCGTTCC	
Use	Gene name	Region coordinates hg19	Orientation	Sequence	UPL Probe#	

qRTPCR primers	*FHL2*	chr2:105977803-105979761	FP	5' GTGCACCAACCCCATCAG	59	
			RP	5' CACCAGTGAGAGGGAGCACT		
qRTPCR primers	*ELOVL2*	chr2:106002556-106003266	FP	5' CCCTTCGGTTGTCTCATCTT	35	
			RP	5' GCATATCTTTCTTCATTGGCTTTT		
qRTPCR primers	*PCDHGA1*	chr5:140713745-140713809	FP	5' AAGCAGAAATTGAAGGTTAGGAAA	67	
			RP	5' AGTCTGAAGCACTTCTGTCCCTA		
qRTPCR primers	*PCDHB1*	chr5:140431131-140431204	FP	5’ GGGGATGCGACAACTATCC	59	
			RP	5’ TTAGCTACGTTGGCCACAAAC		
qRTPCR primers	*ZNF776*	chr19:58267849-58267915	FP	5’ AAACTGTGTGACTCTTAGGGAACA	8	
			RP	5’ CCAAACTAGGCCACAAATGC		
qRTPCR primers	*ZNF551*	chr19:58200467-58200534	FP	5’ TGCCATTTTTGTCAGGTCAG	16	
			RP	5’ ACAGAAGCACTGGGGATTGT		
qRTPCR primers	*P21 (CDKN1A)*		FP	5’ CACCTCACCTGCTCTGCTG	6	

		chr6:36652166-36652242	RP	5’ GAGGCACAAGGGTACAAGACA		
qRTPCR primers	*CCND1*	chr11:69462858-69465924	FP	5' TGTCCTACTACCGCCTCACA	55	
			RP	5' CAGGGCTTCGATCTGCTC		
qRTPCR primers	*PCDHB13*	chr5:140594563-140594631	FP	5' GGCAAAACCTTTAAGATCAATCC	2	
			RP	5' TTTTTCGAAATCGAGTTGTTTTT		
qRTPCR primers	*JUNB*	chr19:12902880-12902984	FP	5' ATCACGACGACGCCTACAC	87	
			RP	5' CTCCTGCTCCTCGGTGAC		
qRTPCR primers	*FOS*	chr14:75745481-75748937	FP	5' GGGGCAAGGTGGAACAGT	46	
			RP	5' TCTCCGCTTGGAGTGTATCA		
qRTPCR primers	*TGFB1*	chr19:41836976-41837042	FP	5' CGCTGCCCATCGTGTACTA	39	
			RP	5' CGCACGATCATGTTGGAC		
qRTPCR primers	*GFP*	N/A	FP	5' CGACGGCGGCTACTACAG	5	
			RP	5' GTGGATGGCGCTCTTGAA		
qRTPCR primers	*RUNX2*	chr6:45517050-45517123	FP	5' TCAAACCCACCTTTGTAGGC	43	
			RP	5' CAGGCAGGAAGCACAGGT		

Table 2. 8. EMSA competitors

Use	Sequence name	P1	P2		
EMSA mutation assay	Original probe	AGCCCACGCCGGAGGCGACGCTCC CGATAACT	AGGCCTCGTCCGAAACTCGCGCCCG AGAGCCC		
	C1	AAAAAACGCCGGAGGCGACGCTCC CGATAACT	AAAAAACGTCCGAAACTCGCGCCCG AGAGCCC		
	C2	AGCCAAAAAAGGAGGCGACGCTCC CGATAACT	AGGCAAAAAACGAAACTCGCGCCCG AGAGCCC		
	C3	AGCCCACGAAAAAAGCGACGCTCC CGATAACT	AGGCCTCGAAAAAAACTCGCGCCCG AGAGCCC		
	C4	AGCCCACGCCGGAAAAAACGCTCC CGATAACT	AGGCCTCGTCCGAAAAAAGCGCCCG AGAGCCC		
	C5	AGCCCACGCCGGAGGCAAAAAACC CGATAACT	AGGCCTCGTCCGAAACAAAAAACCG AGAGCCC		
	C6	AGCCCACGCCGGAGGCGACGAAAA AAATAACT	AGGCCTCGTCCGAAACTCGCAAAAA AGAGCCC		
	C7	AGCCCACGCCGGAGGCGACGCTCC AAAAAACT	AGGCCTCGTCCGAAACTCGCGCCCA AAAAACC		

	C8	AGCCCACGCCGGAGGCGACGCTCCCGAAAAAA	AGGCCTCGTCCGAAACTCGCGCCCGAAAAAAA		
			P1 region		
Use	Sequence name	Consensus	Sequence 5'-3'	Frequency of prediction	Base position
EMSA TF consensus competitor	Random non-competitive sequence		TCTCCATAAGGAGTCCGGTGTAGCGAAGGATC		
	Zfx	AGGCCTGG	AGGCCTGGAGGAGTCCGGTGTAGCGAAGGATC	2	1
	NF-κB	NGGGGAMTTTCCNN	TCTCCATATGGGGACTTTCCGTGCGAAGGATC	2	9
	TP53	CGGACATGTCCGGACATGTC	TCTCCATAAGGAGTCCGCGGACATGTCCGGAC	2	18
	Sp1+3	GGGGGCGGGGG	TCTCCATAAGGAGTCCGGGGGCGGGGGGGATC	2	16

	Myc	NNNCACGTGNNN	TCTCCATAAGGAATGCACGTGCTGGAAGGATC	2	13
	AP2	MKCCCSCNGGCG	TCTCCATAAGGAGTCAGCCCGCTGGCGGGATC	4	16
	ELK1	NNNACMGGAAGTNCNN	TCTCCATTCACCGGAAGTTCGAGCGAAGGATC	2	7
	E2F	NNGCGCGAAANTK	TCTCCATAAGGAGTTCGCGCGAAATTGGGATC	5	15
	p300	CGTCC	TCTCCACGTCCAGTCCGGTGTAGCGAAGGATC	2	7
			P2 region		
Use	Sequence name	Consensus	Sequence 5'-3'	Frequency of prediction	Base position
EMSA TF consensus competitor	Random non-competitive sequence		TCTCCATAAGGAGTCCGGTGTAGCGAAGGATC		

	E2F	NNGCGCGAAANTK	TCTCCATAAGGAGTCCGGTTCGCGCGAAATTG	2	20
	NF-κB	NGGGGAMTTTCCNN	TCTCCATAAGGATGGGGACTTTCCGTAGGATC	2	13
	Zfx	AGGCCTGG	TCTCCATAAGGAGTCCGAGGCCTGGAAGGATC	2	18
	CREB	NNGNTGACGYNN	TCTCCATAAGGTCGCTGACGTACCGAAGGATC	2	12
	AP2	MKCCCSCNGGCG	AGCCCGCTGGCGGTCCGGTGTAGCGAAGGATC	5	1
	STAT	TTCCCRKAA	TCTCCATAAGGAGTCCGGTGTTTCCCGTAATC	2	22
	Sp1+3	GGGGGCGGGGG	TCTCCAGGGGGCGGGGGGTGTAGCGAAGGATC	2	7
	cMYC	NNNCACGTGNNN	TATGCACGTGCTGTCCGGTGTAGCGAAGGATC	2	2
	ELK1	NNNACMGGAAGTNCNN	TCTCTTCACCGGAAGTTCGATAGCGAAGGATC	4	5

Table 2. 9. crRNA sequences

Age-CpG region Cas9 deletion crRNAs			
Name region	Region chromatin state	crRNA sequence	hg 19 coordinates
FHL2	Promoter	5' AAGAATCCTTCGTTCCACTA	chr2:106016004-106016023
		5' CGTGCCCCCGCCCCGGGCTG	chr2:106015640-106015659
ZNF577	Promoter	5' GGTTCCAATGCGGAGACCAT	chr19:52390903-52390922
		5' GGACCTGAGCAACAATCGGA	chr19:52391637-52391656
PCDHB1	Promoter	5' AAGAAACCGTTTACGTCTCA	chr5:140419553-140419572
		5' AGTAGGCAGTTAAGGTTCAC	chr5:140420063-140420082
ZNF551	Promoter	5' TGAGGTGCTGTGCCCCGTGC	chr19:58193162-58193181
		5' GCGCGGAAGTGTTGCTATCC	chr19:58193306-58193325
ELOVL2	Promoter	5' CTTCCAGCGGACACGACCAA	chr6:11045178-11045197
		5' CCACTCACCATGATCCGCAG	chr6:11044453-11044472
PCDHGA1	Enhancer	5' GACGGACAAAGGCTCCTTCG	chr5:140772494-140772513
		5' TTGTATCCAACACTGTCACG	chr5:140773067-140773086
PCDHB1	Enhancer	5' AAGAAACCGTTTACGTCTCA	chr5:140419553-140419572
		5' AGTAGGCAGTTAAGGTTCAC	chr5:140420063-140420082
FHL2 knockdown crRNAs			
Region name	Region of gene	crRNA sequence	hg 19 coordinates
Guide pair 1	Promoter	5' CGTGCCCCCGCCCCGGGCTG	chr2:105399183-105399202

		5' AAGAATCCTTCGTTCCACTA	chr2:105399547-105399566
Guide 2	Exon 3 (common)	5' CAAGAAGTACATCCTGCGGG	chr2:105386450-105386469
Guide pair 2	Exon 3 (common)	5' CAAGAAGTACATCCTGCGGG	chr2:105386450-105386469
		5' CCAACACCTGCGAGGAGTGT	chr2:105386388-105386407

Table 2. 10. Antibodies

Western blotting				
Antibody	Species	Manufacturer	Cat#	Dilution factor
Anti-Cas9	Mouse	Cell signalling technology	14697	500
Anti-GAPDH	Mouse	Merck Millipore	MAB374	20000
Anti-FHL2	Rabbit	Abcam	ab12327	2000
Anti-beta-tubulin	Mouse	Proteintech	66240-1-Ig	20000
Anti-CDKN1A	Rabbit	Proteintech	28248-1-AP	1000
Anti-EDIL3	Rabbit	Proteintech	12580-1-AP	1000
Anti-FOS	Rabbit	Cell signalling technology	4384T	1000
Anti-CELF2	Rabbit	Sigma Aldrich	HPA035813-25UL	2000

Anti-rabbit HRP	Goat	Agilent Dako	P0448	1000
Anti-mouse HRP	Goat	Agilent Dako	P0447	1000
Anti-rabbit IRDye 800CW	Goat	Li-Cor	P/N 925-32211	10000
Anti-mouse IRDye 680CW	Goat	Li-Cor	P/N 925-32210	10000
Immunocytochemistry				
Antibody	Species	Manufacturer	Cat#	Dilution factor
Anti-Cas9	Mouse	Cell signalling technology	14697	500
Anti-BrdU	Rat	Abcam	ab6326	200
Anti-FHL2	Rabbit	Abcam	ab12327	2000
Anti-TOM20	Mouse	Santa Cruz Biotechnology	sc-17764	1000
Anti-beta-actin	Mouse	Proteintech	66240-1-Ig	500
Alexa Fluor 488 Phalloidin	N/A	Invitrogen	A12379	40
Anti-rat Alexa Fluor™ 488	Goat	ThermoFisher Scientific	A-11006	200
Anti-rabbit Alexa Fluor™594	Goat	ThermoFisher Scientific	A11037	200
Anti-mouse Alexa Fluor™488	Goat	ThermoFisher Scientific	A11029	200
EMSA				
Name	Manufacturer	Cat#		
Anti-SP1(E-3)	Santa Cruz Biotechnology	sc-17824X		
Anti-SP3(G-7)	Santa Cruz Biotechnology	sc-365220X		
Anti- NF-κB	Abcam	ab16502		

Anti-STAT1	Proteintech	10144-2-AP		
Anti-STAT2	Cell Signaling Technology	4597s		
Anti-STAT3	Cell Signaling Technology	4904s		
Anti-STAT3	Proteintech	10253-2-AP		
Anti-TFAP2A (AP2)	Proteintech	13019-3-AP		

Table 2. 11. Patient details and information on tissues studied for DNA methylation and gene expression.

Patient details			DNA methylation analysis				Gene expression analysis			
Sample code	Sex	Age	CN	SYN	FP	WB	CN	SYN	FP	WB
606	F	84					✓			
3055	M	69			✓					
3060	M	76			✓					
3122	F	68			✓					
3297	F	79			✓				✓	
3335	M	74			✓					
3348	M	57			✓					
3360	M	64			✓					
3411	M	71		✓				✓		
3458	F	71			✓					
3504	M	87	✓				✓			
3555	M	71			✓					
3772	M	72							✓	
3912	F	74			✓				✓	
3921	M	81			✓					
4251	M	56			✓					
4257	M	66			✓				✓	
4285	F	65			✓				✓	
4338	F	79		✓						
4419	M	64					✓			
4445	F	77			✓					
4483	F	73	✓				✓			
4653	M	62			✓					
4659	F	62			✓				✓	

4660	F	70			✓				✓	
4661	F	79							✓	
4676	M	67							✓	
4695	M	63			✓					
4761	F	82			✓					
4762	M	56	✓				✓			
4763	F	62	✓							
4773	F	76	✓							
4803	M	55	✓							
4810	F	66	✓				✓			
4863	M	48	✓				✓			
4901	F	69	✓				✓			
4923	M	82	✓							
4929	M	58					✓			
4962	M	68	✓							
4965	F	64	✓				✓			
4966	F	87	✓				✓			
4969	F	58	✓				✓			
4983	M	75		✓						
4984	M	62		✓						
4994	F	61	✓				✓			
5003	M	64	✓							
5009	F	58	✓				✓			
5012	M	56	✓				✓			
5014	M	55			✓					
5019	M	50	✓	✓						
5023	M	75		✓						
5031	F	57	✓				✓			

5034	F	74	✓				✓			
5130	M	82	✓				✓			
5140	M	75	✓				✓			
5146	F	76	✓							
5171	F	69	✓				✓			
5173	M	77		✓						
5175	F	79	✓				✓			
5210	F	68	✓				✓			
5215	F	61	✓				✓			
5218	F	77	✓				✓			
5222	F	77	✓							
5225	F	79	✓							
5240	F	62		✓						
5254	F	62		✓						
5266	F	73	✓				✓			
5271	F	68		✓						
5275	F	53	✓				✓			
5281	F	59		✓						
5299	F	75		✓						
5400	F	62	✓							
5489	M	61			✓					
5564	F	73		✓						
5569	F	63		✓				✓		
5575	M	62		✓				✓		
5597	M	71							✓	
5603	F	79		✓						
5622	F	62		✓	✓			✓	✓	
5632	F	52		✓	✓			✓	✓	

5643	M	76		✓						
5644	M	76		✓				✓		
5648	F	66			✓				✓	
5649	M	79			✓				✓	
5650	F	74			✓				✓	
5652	F	82		✓	✓					
5665	M	61		✓						
5677	M	55		✓				✓		
5692	F	67		✓				✓	✓	
5694	F	82		✓				✓		
5700	F	81		✓				✓	✓	
5703	M	77							✓	
5712	F	67		✓				✓		
5713	F	65						✓		
5720	M	68	✓	✓	✓		✓	✓	✓	
5723	M	79		✓					✓	
5725	M	81		✓	✓					
5731	F	58		✓						
5740	F	88		✓	✓			✓		
5747	F	88		✓						
5748	M	67		✓						
5769	F	60	✓							
5770	F	52		✓						
5776	F	85		✓						
5796	F	74	✓				✓			
5803	M	54			✓				✓	
5809	M	79		✓	✓				✓	
5818	M	71			✓					

5846	F	53	✓				✓			
5847	M	70	✓							
5942	F	58	✓							
5980	F	75							✓	
5994	M	58	✓							
6001	F	64	✓							
6002	M	58	✓				✓			
6006	F	75	✓							
6008	M	68					✓			
6009	F	59	✓							
6052	M	76			✓					
6054	F	60	✓				✓			
6057	F	76	✓				✓			
6062	F	64			✓				✓	
6068	F	59		✓					✓	
6090	F	72	✓				✓			
6094	F	54			✓				✓	
6095	M	84					✓			
6102	F	70			✓				✓	
6104	F	53	✓				✓			
6111	M	81			✓				✓	
6114	F	56	✓							
6153	F	72	✓				✓			
6165	M	64	✓							
6175	M	80					✓			
6190	F	70	✓				✓			
6223	M	78	✓				✓			
6225	M	60	✓							

6228	M	78	✓				✓			
6268	M	86							✓	
6535	M	67						✓	✓	
6547	F	51	✓	✓	✓	✓		✓	✓	
6548	F	82	✓	✓	✓	✓		✓	✓	
6549	M	46	✓	✓	✓	✓		✓	✓	
6550	F	81	✓	✓	✓	✓		✓	✓	
6593	M	73	✓	✓	✓	✓		✓	✓	
6597	M	85	✓	✓	✓	✓		✓	✓	
6599	F	81	✓	✓	✓	✓		✓	✓	
6603	F	54						✓	✓	
6636	F	50						✓	✓	
6638	M	72						✓	✓	
6655	M	58						✓		
6810	F	52						✓		
6877	F	71						✓	✓	
6890	M	68						✓		
101k	M	60					✓	✓		
115k	M	62					✓	✓		
132k	M	71					✓	✓		
151K	M	84						✓		
33K	F	58						✓		
50k	F	69					✓	✓		
51K	F	67						✓		
56k	F	52					✓	✓		
95k	F	55					✓	✓		
96k	M	75					✓	✓		
SH007	F	68	✓							

SH040	F	76	✓							
SH041	M	63	✓							
SH043	M	68	✓							
SH045	F	78	✓							
SH046	M	63	✓							
SH048	M	63	✓							
SH049	M	75	✓							

Table 2. 12. Human articular chondrocyte donor details.

HAC 5-Aza treatment			HAC IL-1 stimulation		
Sample code	Sex	Age	Sample code	Sex	Age
CIMA-363	M	77	6626	F	66
CIMA-364	M	68	6898	F	73
CIMA-365	M	72	6960	F	63
CIMA-318	M	65	6968	M	64
CIMA-320	F	66	6969	F	65
CIMA-323	F	47	6973	M	61
			7032	F	62
			7035	F	61
			7038	F	52
			7061	F	73
			7212	F	53
			5419	F	67
			5345	M	66
			5410	M	68
			5378	M	58
			5418	M	60
			5417	F	73

2.10.9 Plasmids

Fuw-dCas9-Dnmt3a

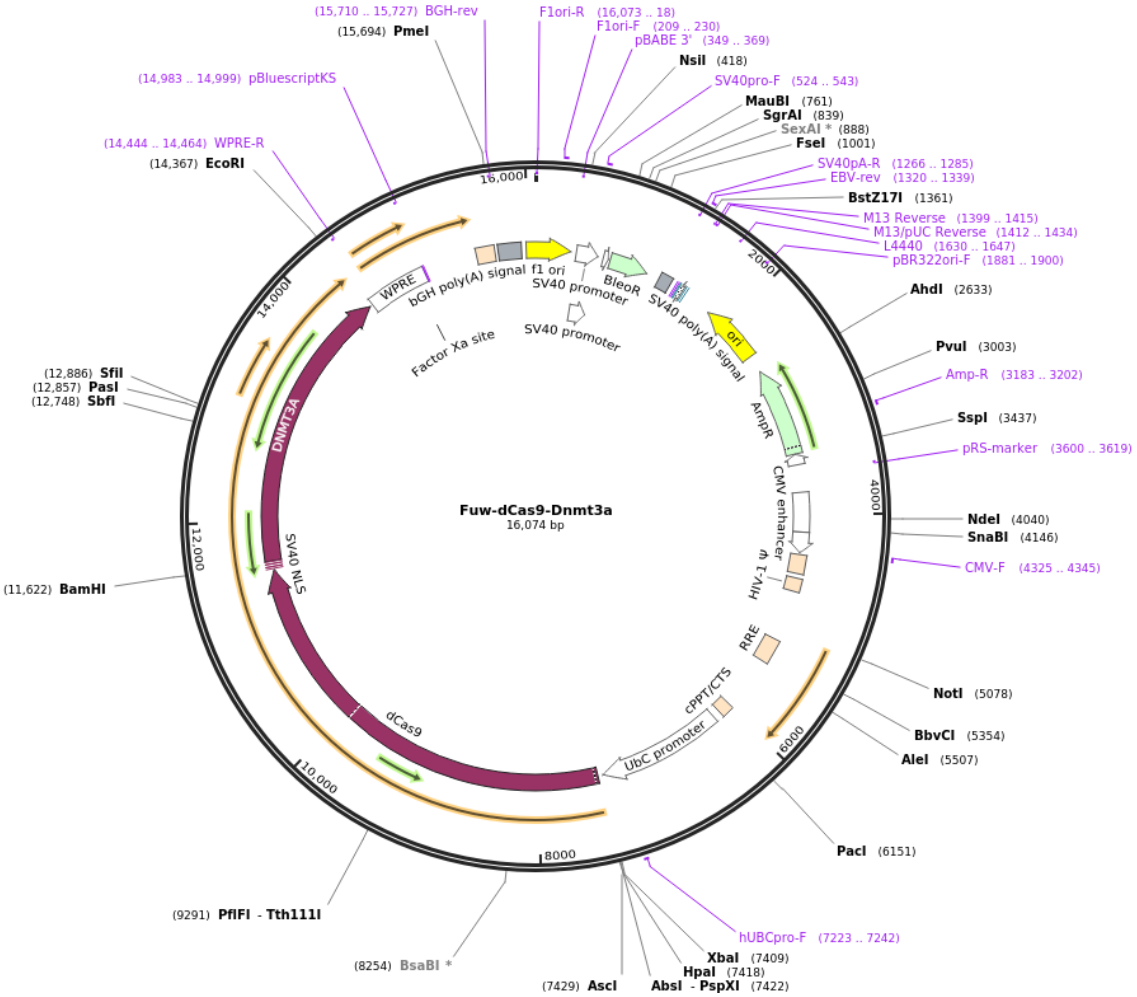


Figure 2. 4. Plasmid map of the dCas9-DNMT3a vector with the functional regions highlighted.

Fuw-dCas9-Tet1CD

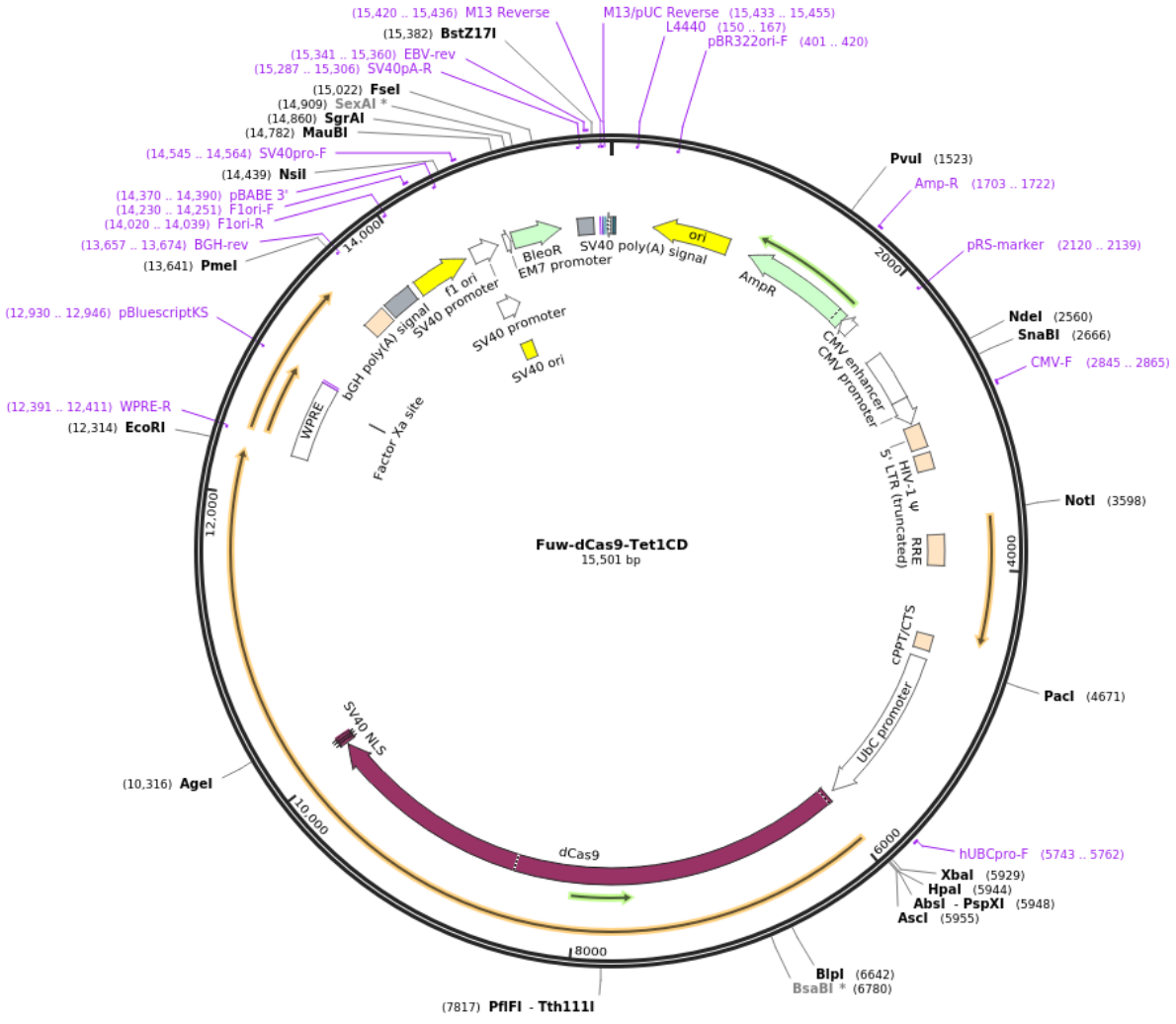


Figure 2. 5. Plasmid map of the dCas9-TET1cd vector with the functional regions highlighted.

FHL2-GFP RG220395 vector

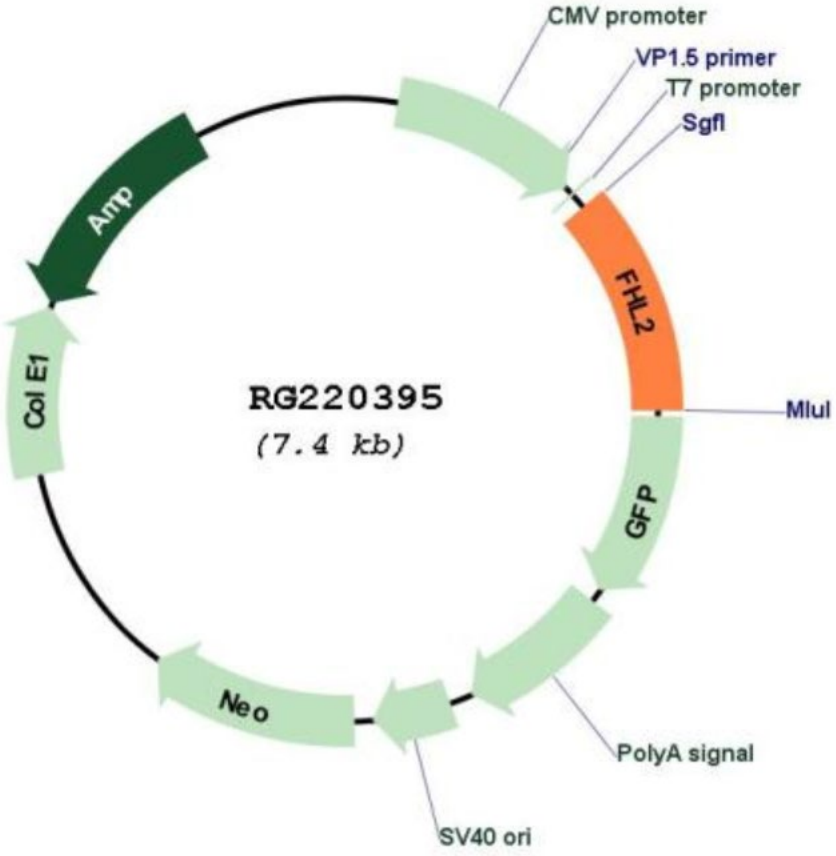


Figure 2. 6. Plasmid vector that contains the FHL2-GFP mRNA transcript under a mammalian G418 selection marker.

pCMV6-AC-GFP control plasmid

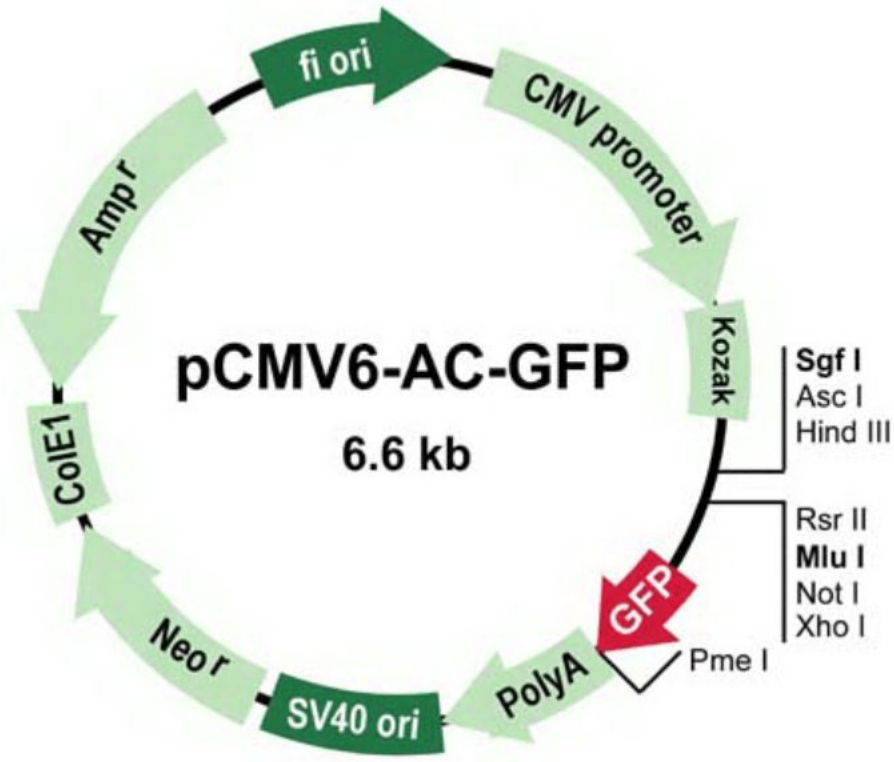


Figure 2. 7. Plasmid vector that contains the GFP mRNA transcript under a mammalian G418 selection marker.

pCpGL basic vector Luciferase

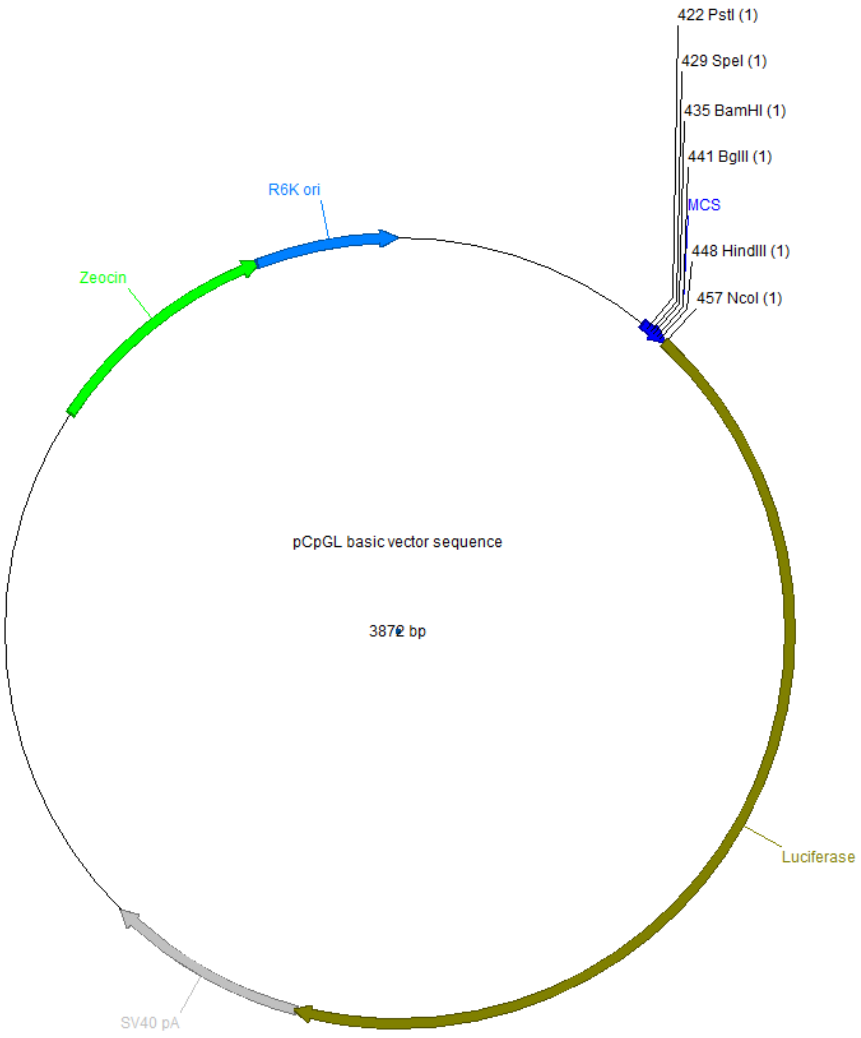


Figure 2. 8. Plasmid map of the pCpGL basic vector with the important functional regions highlighted.

pCpGL EF1 vector Luciferase

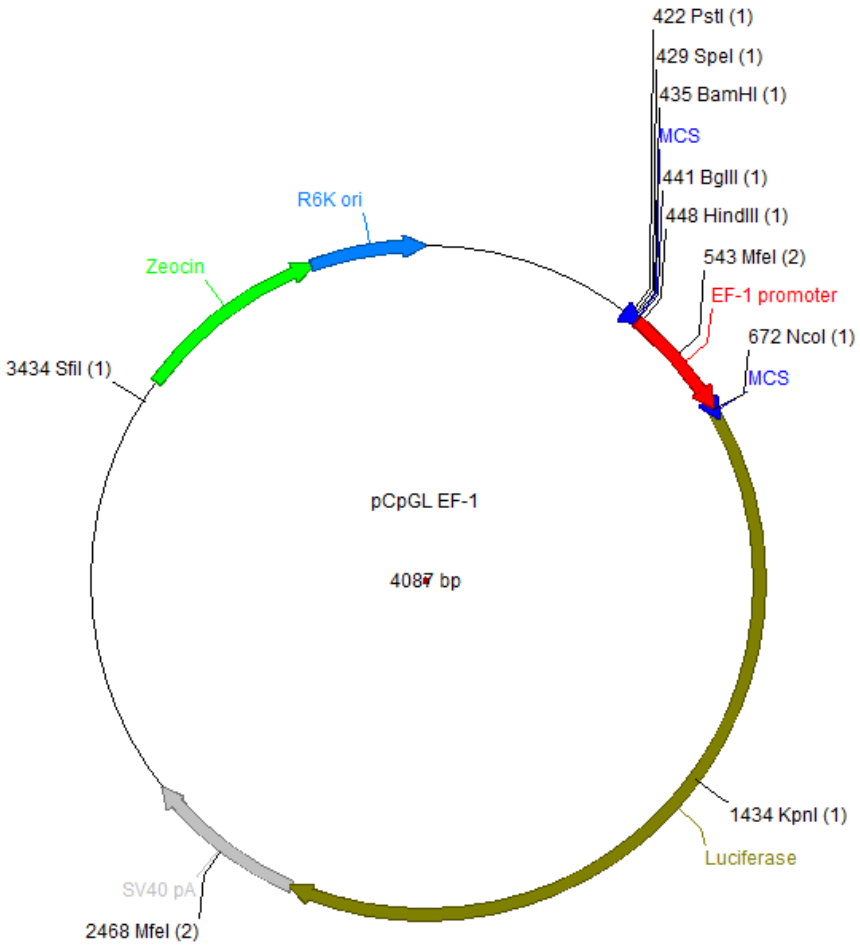


Figure 2. 9. Plasmid map of the pCpGL-EF1 vector with the important functional regions highlighted.

Chapter 3: Age-related CpG DNA methylation changes in the musculoskeletal system and non-musculoskeletal tissues

3.1 Introduction

DNA methylation in mammals is most commonly found as 5-methylcytosine that is the addition of a methyl group on the 5' carbon of a cytosine base preceding a guanine base, forming a CpG unit. These covalent modifications have been generally linked with a repressive activity in gene expression when the CpG sites in a promoter or enhancer are methylated (Day *et al.*, 2013) while methylation in the gene body was observed to correlate with an increase in expression (Yang *et al.*, 2014). The importance of these modifications is related to the adherence to the cell phenotype, but also to maintain genomic stability (Moore *et al.*, 2013). Mechanisms through which DNA methylation can induce gene repression is by inhibiting the binding of transcription factors and by recruiting methyl CpG binding proteins (e.g. Polycomb group proteins) that change the surrounding chromatin structure to a closed conformation as heterochromatin (Ringrose and Paro, 2004; Klose and Bird, 2006). These DNA methylation modifications at CpG sites have been observed to have tissue specificity (Day *et al.*, 2013) but also some have been highly associated with ageing (Johnson *et al.*, 2012; Day *et al.*, 2013; Hannum *et al.*, 2013).

During ageing, the global methylation levels of CpGs drops when measured by techniques such as reduced representation bisulphite sequencing (RRBS) or whole genome bisulphite sequencing (WGBS) (Jintaridth and Mutirangura, 2010; Zampieri *et al.*, 2015; Corso-Díaz *et al.*, 2020) while several investigators observed that the levels of some CpG sites increase with the chronological age of a person with some CpG sites showing tissue specific patterns (Day *et al.*, 2013; Hannum *et al.*, 2013; Horvath, 2013). The CpG sites whose DNA methylation levels correlate with age have been termed as age-CpGs and these have been identified in several different healthy tissues as well as diseased tissues e.g. multiple types of cancers (Jin and Liu, 2018). An example of such age-CpG sites such as those located in the gene promoters of *ELOVL2* (cg21572722, cg24724428, cg16867657) (Slieker *et al.*, 2018) and *FHL2* (cg06639320, cg22454769, cg24079702) (Habibe *et al.*, 2021) have been reported in several tissues and are almost regarded as non-tissue specific age-CpGs, being proposed as an ubiquitously changing age-CpG across the human body.

Based on the data sets generated from the Illumina 27K and 450K methylation array datasets, several groups have established epigenetic age predictor models using both tissue specific and multi-tissue age-CpGs (Bocklandt *et al.*, 2011; Hannum *et al.*, 2013; Horvath, 2013). For example, Horvath used 27K and 450K array-based methylation data from 8,000 samples from 51 healthy tissues and cell types from embryonal development to adulthood and old age to develop an 'aging clock' (Horvath 2013). This clock uses methylation levels at 353 CpGs present on both the 27K and 450K array to calculate the epigenetic age of an individual. The DNA methylation age in Horvath's epigenetic clock is the measure of the epigenetic maintenance system, where the cells from the embryonic stage show natively an age of ~0 with a logarithmic increase of this age in comparison to chronological age until about the age of 30-40 years this being followed by a more linear increase (Bocklandt *et al.*, 2011; Horvath, 2013). This tool permits the study of the impact of diseases on a particular tissue or cell type, and has been used to observe that in certain diseases such as obesity and Werner syndrome there is an accelerated epigenetic age of tissues (e.g. blood and liver respectively) (Horvath *et al.*, 2014; Maierhofer *et al.*, 2017; Nevalainen *et al.*, 2017). This accelerated epigenetic age has been correlated with an increased risk for disease development (e.g. autoimmune diseases or age-related diseases) (Szyf, 2008; Pal and Tyler, 2016; Jin and Liu, 2018; McCartney *et al.*, 2018), being termed a "biological doomsday clock" (Mendelson, 2018; Huang *et al.*, 2019). However, although several studies have identified age-CpGs or have generated DNA methylation -based epigenetic clocks, very few of these studies have assessed the consequences of these DNA modifications on gene expression levels of target genes (McKinney *et al.*, 2019)

Several groups have used the Illumina 450K array to quantify DNA methylation in healthy and OA cartilage ((den Hollander *et al.*, 2015; Rushton *et al.*, 2015; Alvarez-Garcia *et al.*, 2016a)) and have observed that certain loci are differentially methylated with the disease however further work is needed to understand mechanisms and how factors such as ageing influence disease onset . In order to investigate if epigenetic ageing was accelerated in OA cartilage, a meta-analysis of genome-wide DNA methylation levels from 179 osteoarthritic (OA) and non-OA cartilage samples from patients ranging from 50 to 95 years was performed. Although there were no differences in epigenetic ageing rate between OA and non-OA cartilage samples, linear regression of methylation against chronological age

identified in cartilage 716 age-CpG sites (Reynard, unpublished data). Whilst some of the age-CpGs found in this study have been reported to correlate with age in non-musculoskeletal tissues (e.g. blood; (Day *et al.*, 2013; Hannum *et al.*, 2013), analysis of data sets from five other studies encompassing multiple different tissue types suggested several age-CpGs were specific to cartilage. Over 90% of the cartilage age-CpGs were hypermethylated with age and are enriched in CpG islands. Furthermore, overlap of age-CpG location with chromatin state information for in vitro differentiated chondrocytes demonstrated that the majority of cartilage age-CpGs (~70%) were mapped to gene promoters and approximately 9% were located within enhancer regions.

3.2 Aims

1. Using published DNA methylation data generated in other synovial joint tissues to examine if cartilage age-CpGs are also present in other tissues

2. Validating cartilage age-CpGs by analysing additional knee cartilage DNA methylation datasets and quantifying DNA methylation at specific age-CpGs by bisulphite pyrosequencing in a new cohort of OA knee samples

3. Quantifying methylation at specific cartilage age-CpGs in additional synovial joint tissues from OA knee patients and comparing methylation ageing rates across CpGs and tissues.

4. Development of an epigenetic clock that works with bisulphite pyrosequencing data for musculoskeletal tissues.

3.3 Results

3.3.1 Investigation of cartilage age-CpGs in other Illumina 450k or EPIC DNA methylation datasets

The Infinium HumanMethylation450 BeadChip (Illumina 450k methylation array) is a platform for the study of over 450,000 CpG sites across the genome that uses probe-based technology to estimate DNA methylation levels at selected probe sites. This platform was used extensively in the literature to investigate the link between DNA methylation levels and other variables such as disease (Fortin *et al.*, 2014; Smith *et al.*, 2021) or ageing (Bacalini *et al.*, 2015). The Illumina DNA methylation 450k array was used on cartilage samples to determine the age-CpGs present in this tissue, where 716 CpG sites passed the significance threshold at an adjusted p value <0.05 for the linear regression analysis. A comparison was performed between the Illumina 450k data and data from the newer EPIC array as there are reports of poor correlation between the two arrays in different tissues (Logue *et al.*, 2017; Fernandez-Jimenez *et al.*, 2019). The newer Infinium MethylationEPIC array that gives data for over 800,000 CpG sites and the dataset was composed of 77 new cartilage samples with an age range between 46-95 (average age of 70) for the investigation. The analysis looked at these previously identified 716 age-CpGs and whether these still can be detected as age-CpGs and it was found that for the EPIC array only 638 of the 716 age-CpGs are present as probes on the array and from this only 473 CpGs (74%) were observed to have a significant change with age (Figure 3. 1). In addition to the comparison to the EPIC array another smaller dataset generated on the Illumina 450k with 23 new cartilage samples with an age range between 19-82 (average age = 47) was analysed for age-related changes and compared to the previously identified 716 age-CpG sites. For this analysis only 365 CpGs very identified (~51%) to be statistically significant age-CpGs (Figure 3. 1)

Further investigation looks at how the 716 cartilage age-CpGs can be detected in other datasets that have not necessarily investigated the age-related changes in DNA methylation through linear regression analysis. Publicly available Illumina 450k array datasets were obtained for several tissues to compare against the 716 cartilage CpGs to understand

whether these are specific changes with age for cartilage or these could be observed for other tissues such as those from the musculoskeletal system tissues or others outside this system. These datasets had DNA methylation information for the tissue part of the musculoskeletal system such as of bone (Lien *et al.*, 2018) and skeletal muscle (Zykovich *et al.*, 2014) as well as for tissues that are not part of the musculoskeletal system such as whole blood (Hannum *et al.*, 2013), liver (Kurokawa *et al.*, 2022), breast (Panjarian *et al.*, 2021b), dermis and epidermis (Vandiver *et al.*, 2015b).

Investigating the occurrence of the 716 cartilage age-CpGs in the bone tissue indicates that only 232 out of the 716 CpGs (~32%) displayed age-CpGs characteristics, this being followed by 70 CpGs (~9.8%) for the skeletal muscle (Figure 3. 2 Musculoskeletal tissues). For the investigation of the occurrence of the 716 cartilage age-CpGs in tissues of non-musculoskeletal origin, in the blood tissue 507 of the age-CpGs (70.8%) were detected, for the liver 340 CpGs (47.5%), the breast tissue had 226 CpGs (31.5%), the epidermis tissue with 279 CpGs(38.9%) and dermis tissue 517 CpGs(72.2%) (Figure 3. 2 Non-musculoskeletal tissues). A summary table for a sample of 48 from the 716 cartilage age-CpGs is shown in Table 3. 1, where CpGs such as cg22454769, cg24079702, cg02650266, cg03555227 were detected as age-CpGs in all the datasets regardless of the origin, whereas the CpGs of cg24214260, cg21653581, cg22448433, cg22643867 were only found in the cartilage dataset.

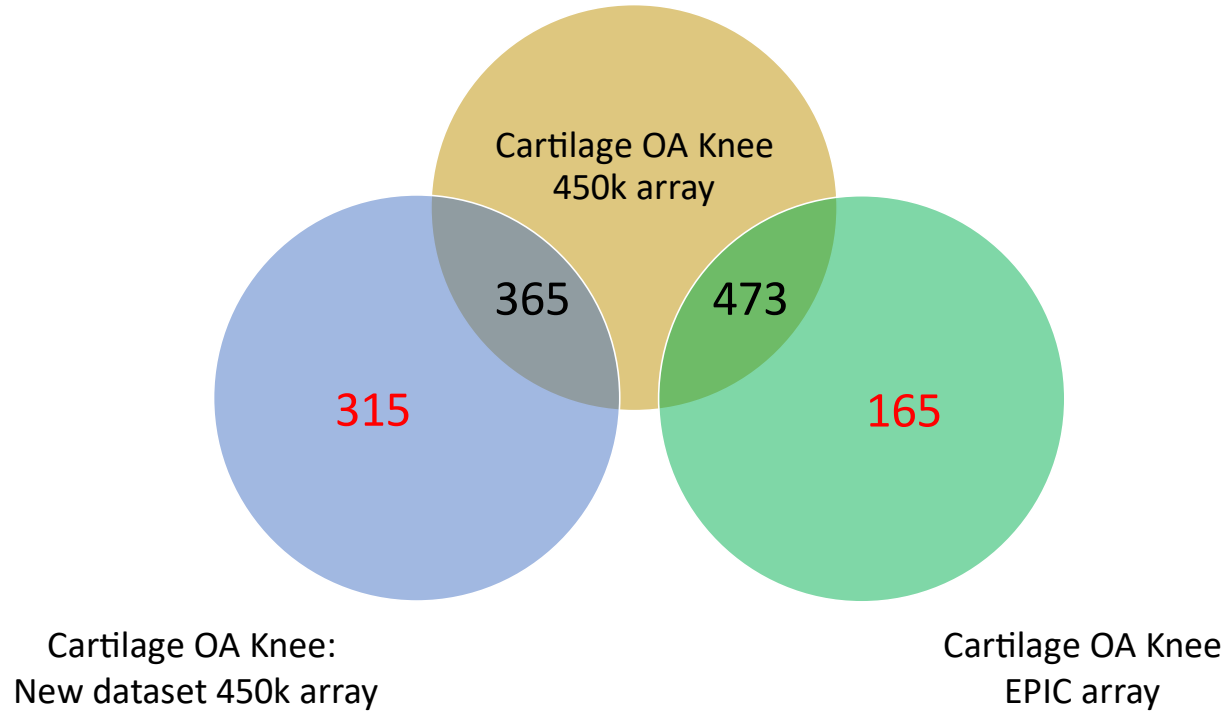


Figure 3. 1. Investigation of the reproducibility of the findings of the 716 age-CpGs on the Illumina EPIC array as well as a new cohort of cartilage samples on the same Illumina 450k array. Red highlighted text represents the number of CpGs that did not pass the significance threshold to be an age-CpG.

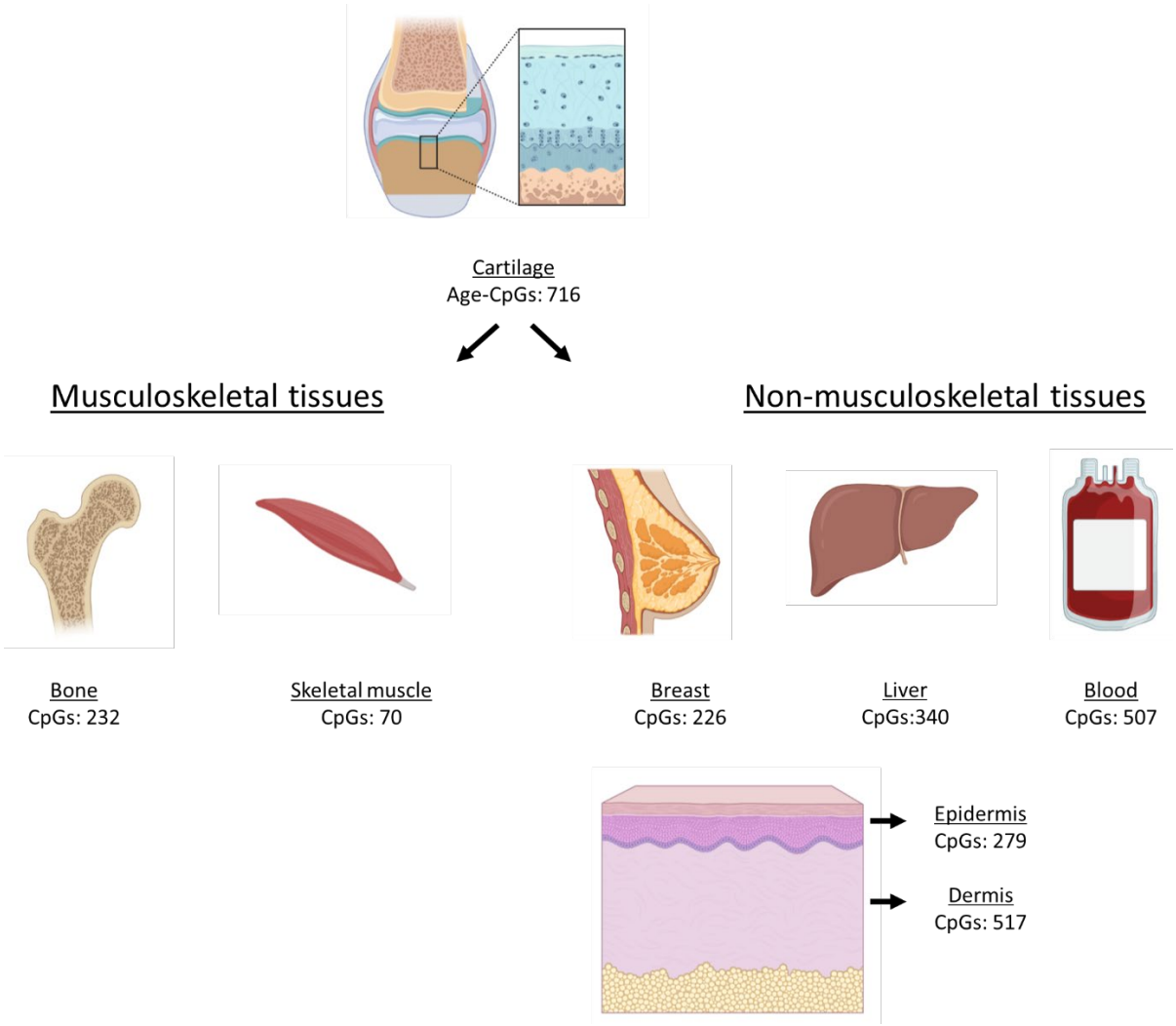


Figure 3. 2. Investigation of the 716 cartilage age-CpGs in other musculoskeletal and non-musculoskeletal publicly available Illumina 450k array datasets. The musculoskeletal tissues include bone and skeletal muscle, and non-musculoskeletal tissues include breast, liver, blood, epidermis and dermis.

Table 3. 1. Summary table of the occurrence of 48 of the 716 cartilage age-CpG sites in other Illumina 450k datasets for musculoskeletal tissues or non-musculoskeletal tissues. Each CpG site with hg19 coordinates is shown together with a counter of the occurrence in all of the other Illumina 450k datasets, occurrence in non-musculoskeletal tissue datasets or musculoskeletal tissues datasets.

		Occurrence in other Illumina 450k datasets						Occurrence in other Illumina 450k datasets		
CpG	hg19 coordinates	All	Non-musculoskeletal	Musculoskeletal		CpG	hg19 coordinates	All	Non-musculoskeletal	Musculoskeletal
cg22454769	chr2:106015768-106015769	7	5	2		cg07634191	chr8:27850178-27850178	2	2	0
cg24079702	chr2:106015772-106015773	7	5	2		cg13144059	chr2:201245077-201245077	2	2	0
cg02650266	chr4:147558239-147558239	7	5	2		cg19191984	chr4:120549793-120549793	2	2	0
cg03555227	chr5:170289071-170289071	7	5	2		cg19250101	chr19:54666216-54666216	2	2	0
cg06458239	chr19:58038574-58038574	7	5	2		cg19258882	chr12:56472805-56472805	2	1	1
cg06782035	chr5:16179135-16179135	7	5	2		cg19340278	chr10:97453976-97453976	2	2	0
cg07570470	chr8:142318841-142318841	7	5	2		cg19843036	chr4:159131517-159131517	2	1	1
cg08715791	chr11:66189297-66189297	7	5	2		cg19869037	chr6:116732515-116732515	2	1	1
cg21572722	chr6:11044894-11044895	6	4	2		cg06926735	chr20:48732667-48732667	1	1	0
cg06639320	chr2:106015740-106015741	6	4	2		cg20552468	chr17:79259699-79259699	1	1	0
cg23500537	chr5:140419820-140419820	6	4	2		cg20699200	chr4:6675675-6675675	1	1	0
cg00481951	chr3:187387651-187387651	6	5	1		cg20946277	chr16:77247012-77247012	1	1	0
cg00590036	chr6:158957434-158957434	6	5	1		cg21000165	chr13:53191308-53191308	1	1	0
cg00702638	chr3:44803294-44803294	6	4	2		cg16867657	chr6:11044877-11044878	0	0	0
cg00731185	chr7:155167703-155167703	6	5	1		cg24724428	chr6:11044888-11044889	0	0	0
cg00745389	chr7:32467436-32467436	6	5	1		cg24214260	chr19:58192221-58192223	0	0	0
cg14566959	chr5:140772681-140772681	5	4	1		cg21653581	chr3:167452757-167452757	0	0	0
cg22331349	chr19:52391351-52391351	5	4	1		cg22448433	chr13:76123478-76123478	0	0	0
cg09547119	chr19:52391367-52391367	5	3	2		cg22643867	chr2:131148878-131148878	0	0	0
cg06217736	chr11:1579331-1579331	4	4	0		cg23601994	chr7:134001916-134001916	0	0	0
cg07477282	chr15:44956108-44956108	3	3	0		cg24299306	chr12:10096152-10096152	0	0	0
cg18569335	chr17:40171970-40171970	3	2	1		cg24773522	chr15:75918163-75918163	0	0	0
cg01644850	chr19:58193232-58193232	3	3	0		cg24824725	chr10:97453974-97453974	0	0	0

3.3.2 Selection of cartilage age CpGs for further analysis

Based on the Illumina 450k Methylation array results correlating age and DNA methylation at ~450,000 CpG sites, 716 CpG sites were found to change with age by linear regression analysis. Due to the poor correlation that has been reported between the Illumina methylation platforms (Logue *et al.*, 2017; Fernandez-Jimenez *et al.*, 2019) as well as when compared to bisulphite pyrosequencing (Cheung *et al.*, 2020) for some CpG sites it was decided to validate some of the cartilage age-CpGs in a new cohort of 72 cartilage samples (average age= 67 with a range 48-87) using bisulphite pyrosequencing.

From the 716 age-CpGs, further filtering was performed where Dr. Reynard selected 65 of these sites based on their significance level and location in promoter or enhancer regions based on chromatin state data. Further selection was done by taking the CpGs that showed a difference of >10% variability in Beta level between the youngest and the oldest samples, this resulting in 53 CpGs remaining. In order to produce reliable bisulphite pyrosequencing assays the CpGs were then screened for any nearby single nucleotide polymorphisms (SNP) and only the CpGs that had no nearby SNPs within 500 bases distance from a CpG, this resulting in 31 CpGs remaining. Assays were then designed using the Qiagen Pyromark Q24 software where 15 CpGs had assays with acceptable scoring in order to proceed with the validation (Figure 3. 3). The CpG sites chosen for validation in a new cohort were the following: cg24079702, cg22454769, cg06639320, cg14566959, cg01644850, cg16867657, cg21572722, cg22331349, cg23500537, cg06926735, cg18569335, cg07634191, cg13144059, cg09547119, cg07477282. A summary table with the details for each selected CpG is presented for the chromosomal coordinates in hg19, chromatin state in E049 cells (Mesenchymal Stem Cell Derived Chondrocyte) and the nearby gene (Table 3. 2).

Plan for investigation

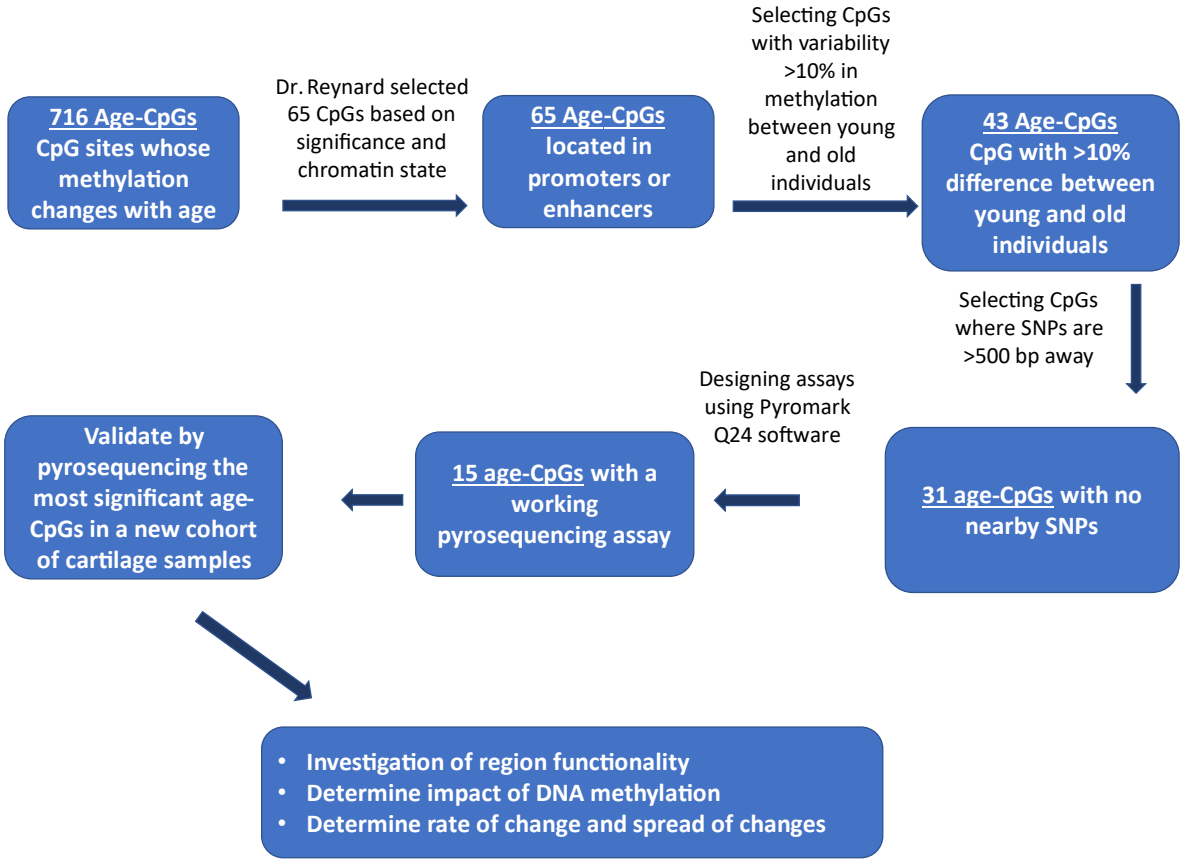


Figure 3. 3. Methodology for selecting the age-CpGs to investigate through bishulphite pyrosequencing. Initial screening was accomplished at first by the selection of the most significant CpGs present in either promoter or enhancer regions followed by filtering was on the range of change with age in DNA methylation, proximity to SNPS and feasibility to design pyrosequencing assay.

Table 3. 2. CpG list of selected age-CpGs for pyrosequencing validation. CpG hg19 coordinates linked to chromosome number and the nearest gene is shown for each CpGs site as well as the chromatin state found in E049 cell is presented.

CpGs	hg19 coordinates	Chromating state	Neighbouring gene
cg16867657	chr6:11044877-11044878	Promoter	*ELOVL2*
cg21572722	chr6:11044894-11044895	Promoter	*ELOVL2*
cg06639320	chr2:106015740-106015741	Promoter	*FHL2*
cg22454769	chr2:106015768-106015769	Promoter	*FHL2*
cg24079702	chr2:106015772-106015773	Promoter	*FHL2*
cg07477282	chr15:44956108-44956109	Enhancer	*SPG11*
cg18569335	chr17:40171970-40171971	Enhancer	*NKIRAS2*
cg01644850	chr19:58192232-58192233	Promoter	*ZNF551, ZNF776*
cg07634191	chr8:27850178-27850179	Enhancer	*NUGGC, SCARA5*
cg06926735	chr20:48732667-48732668	Enhancer	*TMEM189*
cg13144059	chr2:201245077-201245078	Enhancer	*SPATS2L*
cg14566959	chr5:140722681-140722682	Enhancer	*PCDHGA1*
cg22331349	chr19:52391351-52391352	Promoter	*ZNF577*
cg09547119	chr19:52391367-52391368	Promoter	*ZNF577*
cg23500537	chr5:140419820-140419820	Enhancer	PCDHB1

3.3.3 ELOVL2 promoter region

The first region under investigation is the promoter region of *ELOVL2* that contained two age-CpGs cg16867657, cg21572722 located in close proximity. The region shows a consistent promoter chromatin signature with high levels of the histone marks H3K27Ac and H3K4Me1 (Figure 3. 4). The *ELOVL2* (Elongation Of Very Long Chain Fatty Acids-Like 2) gene encodes for an transmembrane enzyme involved in the synthesis of very long polyunsaturated acids, these playing a role in retina and photoreceptor renewal of the eye (Chao and Skowronska-Krawczyk, 2020).

For the analysis of the Illumina 450k and EPIC data, the beta values were converted to M values to account for heteroscedasticity. The probe based array data is described to be non-linear towards either 0% or 100% DNA methylation this following a sigmoidal curve (heteroscedasticity) however, the values can be adjust to M values that show better linearity at these levels (Du *et al.*, 2010). Investigating publicly available Illumina 450k datasets to understand whether the CpG sites present at this region of interest could also be identified as age-CpGs in other datasets such as for tissues present in the musculoskeletal system (e.g. skeletal muscle, bone) or non-musculoskeletal tissues (e.g. blood, liver). In the

musculoskeletal tissues (Figure 3. 5), the CpG of cg16867657 is found to have the highest overall correlations between the DNA methylation changes and ageing. For this CpG site, for the bone tissue this had a R^2=0.388 and p value < 0.0001, followed by the skeletal muscle tissue with a correlation coefficient R^2=0.799 and p value< 0.0001 and with no available data for the tendon tissue at this CpG site (Table 3. 3). In the non-musculoskeletal tissues (Figure 3. 6), the highest correlation was observed for the dermis tissue with R^2=0.852 and p value< 0.0001 and the lowest for the breast epithelial tissues at R^2=0.2 and p value= 0.003 (Table 3. 3). The CpG site with the lowest average correlation with age for the investigated tissues was CpG cg21572722 where in musculoskeletal tissues (Figure 3. 5) the analysis results were for bone R^2=0.315 and p value<0.0001, for skeletal muscle R^2=0.48 and p value=<0.0001, for tendon R^2=N/A and p value=0.393 (Table 3. 3). Unfortunately, no R^2 correlation data was available for this tissue due to the owner of the dataset performing an un-paired t-test analysis for the age-related impact on DNA methylation, however this would still indicate that a significant change is observed for this tissue. In the non-musculoskeletal tissues (Figure 3. 6), the highest correlation was observed for the tissue of epidermis R^2=0.9 and p value< 0.0001 and the lowest correlation that was also non-significant was for the breast tissue with R^2=0.02 and p value=0.34 (Table 3. 3). Due to the fact that cg24724428 was located in between cg16867657 and cg21572722 this CpG was also analysed in the datasets and a summary for the age correlation is found in (Table 3. 3).

For the CpG sites found in the promoter of *ELOVL2,* the three age-CpGs were investigated by bisulphite pyrosequencing, with pyrosequencing assay having an additional two neighbouring CpGs that were also investigated, these being named according to the proximity to the nearest CpG probe coordinates. The CpG site cg24724428 was not found to be an age-CpG in the initial Illumina 450k array cartilage dataset tissues however this was included in the pyrosequencing assay due to its proximity to cg16867657 and cg21572722. In order to validate the findings from the Illumina array 72 new cartilage samples were analysed having an age range from 48-87 with an average age of 67, where approximately 62.5% were female samples.

Furthermore, a calibration curve was generated in order to adjust for any amplification bias that can result from PCR. A standard of synthetically made oligos of the region under analysis was ordered for 0% DNA methylation and 100% DNA methylation, with this being

used to make a serial gradient with increments of 10% to study the accuracy of the pyrosequencing results as well as to adjust for the PCR bias using a cubic equation adjustment of the raw values (Moskalev *et al.*, 2011) (Methods 2.8.2.).

The positive relationship between age and methylation levels of several age-CpGs previously observed on the array was replicated in the new cartilage cohort. For the cartilage tissue (Figure 3. 7), this CpG that had the highest correlation with age was cg24724428 (R^2=0.2815 and p value< 0.0001) followed by cg16867657 (R^2=0.2815 and p value< 0.0001), whereas the CpG of cg21572722 had a non-significant trend with ageing with R^2=0.06859 and p value=0.0607 (Table 3. 4). Further analysis investigates how other articular joint tissues could also have age-associated DNA methylation changes and 22 samples of synovium tissues as well as 22 samples of fat pad were investigated for the now four cartilage age-CpGs. In the synovium tissue (Figure 3. 7), the CpG with the highest correlation was observed to be cg24724428-9bp (R^2=0.703 and p value< 0.0001) this being followed by cg16867657-4bp (R^2=0.664 and p value< 0.0001) and the lowest correlation and passed significance was observed cg21572722 (R^2=0.613 and p value< 0.0001) (Table 3. 4). For the fat pad tissue (Figure 3. 7), the CpG site of cg24724428 had the highest correlation coefficient with a R^2=0.553 and p value< 0.0001 and the CpG site of cg21572722 (R^2=0.53 and p value= 0.0001) came second, while the CpG site of cg24724428-9bp had the lowest correlation coefficient at R^2=0.285 and p value=0.01 (Table 3. 4). Additionally, an investigation in blood tissue was also performed on seven samples and the results for linear regression on DNA methylation versus age indicated that all five CpGs were significantly changing with ageing. The highest correlation was observed for the CpG site of cg24724428-9bp with R^2=0.886 and p value= 0.001 and cg16867657 with R^2=0.851 and p value= 0.003, whereas the lowest was for the CpG site of cg21572722 that had R^2=0.722 and p value= 0.02 (Table 3. 4).

As for all the tissues studied there was an indication that almost all of the age-CpGs found in cartilage were also present in synovium, fat pad and blood (Table 3. 4) from this further investigation seeks to analyse whether there is any correlation to the changes between the tissues that could be linked. The analysis is based on the accumulated unmatched samples using one-way ANOVA as well as seven quadruple matched samples for cartilage, synovium, fat pad and whole blood using two-way ANOVA. With the unmatched samples (Figure 3.

8A), the one-way-ANOVA indicated a significant difference between almost all the tissues with exception of the synovium and fat pad tissue (p=0.45) that was present at most CpG sites within the region. The only exception was observed for cg24724428 where no significant differences were also observed for the tissues of synovium and blood (p=0.99) as well as for fat pad and blood (p=0.58) (Table 3. 5 A). Interestingly the same correlation relationship was observed for the matched samples (Figure 3. 8B), where no detectable differences were observed between synovium and fat pad for most of the CpGs where additionally for cg24724428 no significant difference was observed for the comparison between synovium and blood as well as fat pad and blood (Table 3. 5 B). Principal component analysis (Figure 3. 8A III) shows a separation on the first principal component between cartilage, blood and synovium+ fat pad, where the synovium and fat pad signature were undistinguishable on this PC1 with PC2 being able to produce a bigger separation at a smaller lever between all of the tissues. Matrix correlation analysis on the matched samples revealed that the highest correlations between the tissues is observed for the fat pad and cartilage tissues (R^2=0.857) (Figure 3. 8B III).

The intra-regional analysis between tissues (Figure 3. 9 A) and the mean methylation pattern observed across all the tissues would indicate that only for the cartilage tissues there is a difference in overall methylation levels (average = 34.5%), this showing a hypomethylated state when compared to the signature for the synovium (average = 72%), fat pad (average = 69%) and blood (average = 78.3%). Slope normalisation analysis reveals that the CpG site cg16867657 shows the highest degree of change per year in cartilage and cg16867657+2bp shows the highest degree of yearly change across the synovium, fat pad and blood tissues (Figure 3. 9B). Across the region the correlation coefficient R^2 is higher consistently in blood (R^2>0.72) this being followed by the fat pad tissue (R^2>0.61) (Figure 3. 9C).

Matrix correlation analysis (Figure 3. 10 A) identified the CpG site of cg21572722 to stand outside the cluster formed by its neighbouring CpGs cg21572722+6bp, cg24724428, cg16867657 and cg16867657+2bp. This relationship however is not observed for the synovium and fat pad tissues, where cg21572722 seems to be integrated into this correlation cluster. This observation is somewhat also replicated for the principal

component analysis (Figure 3. 10 B) for these tissues on the PC2 axis however on PC1 there is a higher degree of separation between all of the CpG sites for all the articular joint tissues.

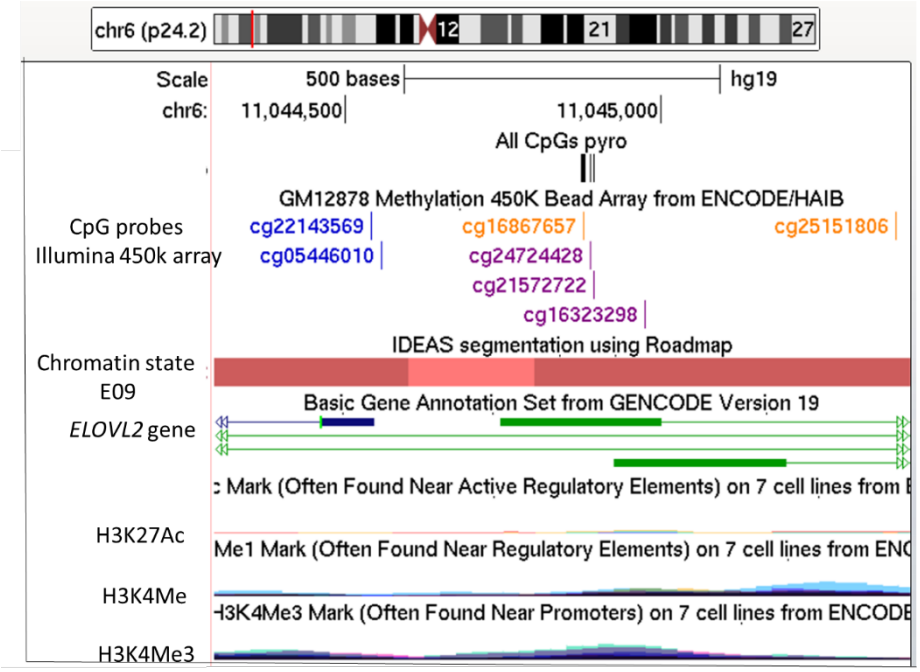


Figure 3. 4. UCSC genome browser view of the chromosomal location of the age-CpGs located in the promoter of ELOVL2 cg16867657, cg24724428, cg21572722. The track All CpGs pyro provides the relative location of the CpGs included in the pyrosequencing assay; the chromatin state information is presented for the E049 cell line. Information about the nearby genes/isoforms is presented below this followed by information regarding the histone mark intensity present at the region in 7 cell lines provided by the ENCODE project

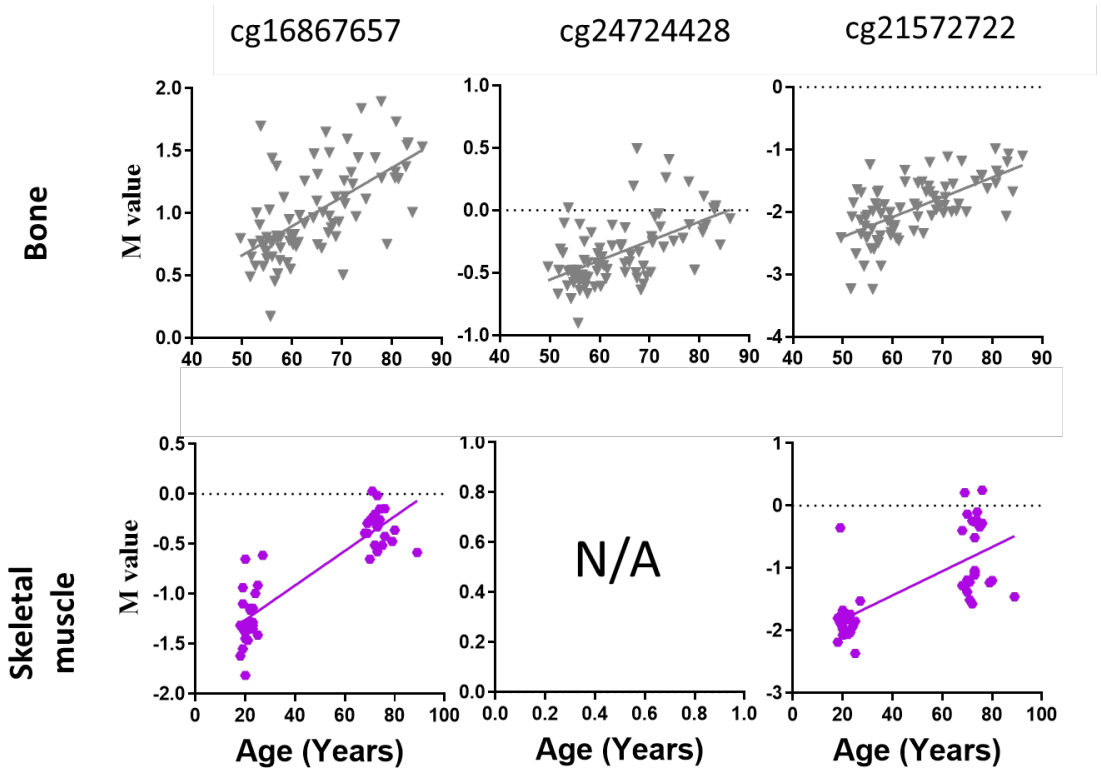


Figure 3. 5. Publicly available Illumina 450k datasets information for the musculoskeletal tissues of bone and skeletal muscle and linear regression between age and DNA methylation analysis for cg16867657, cg24724428 and cg21572722 CpGs.

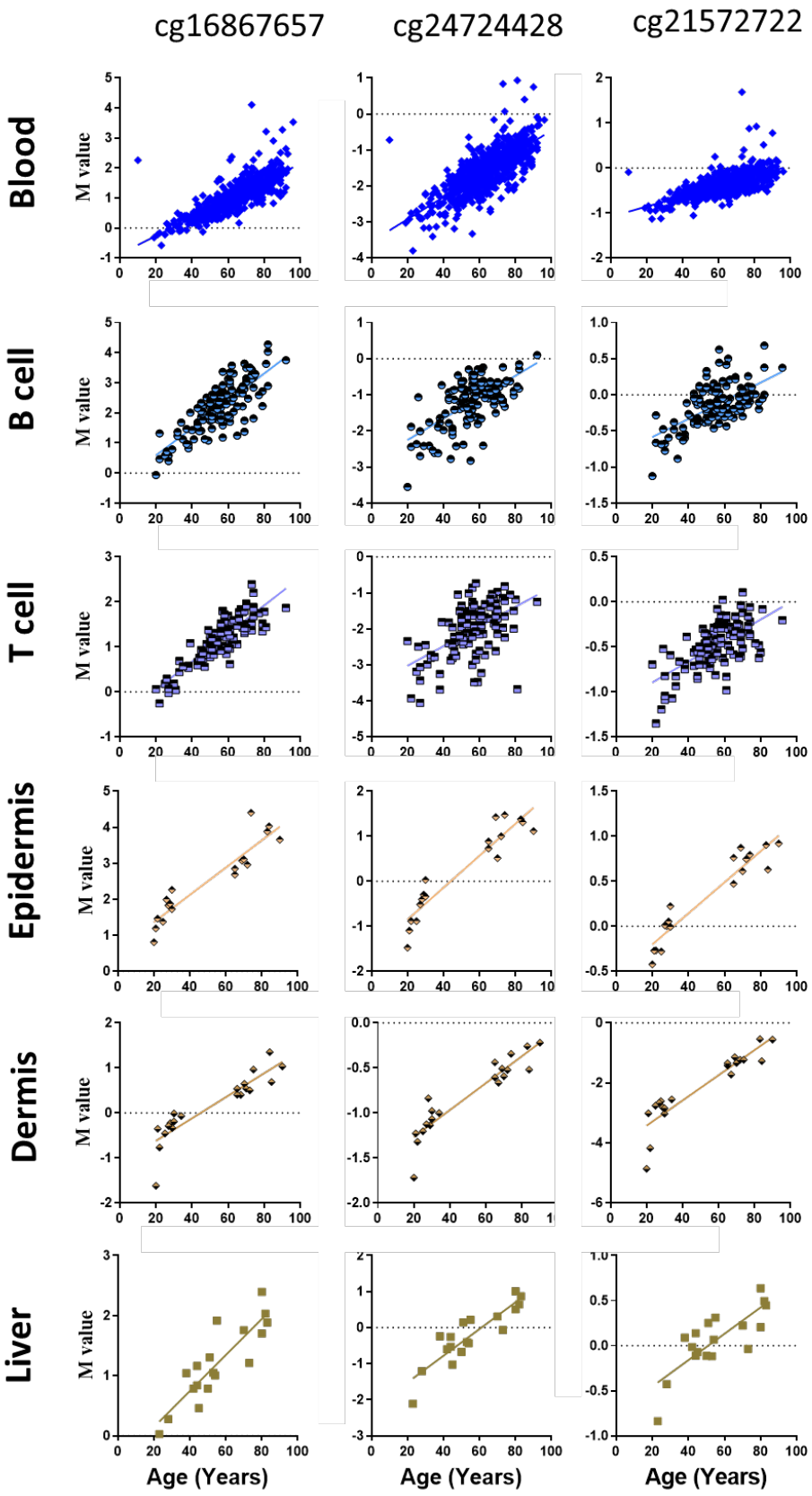


Figure 3. 6. Publicly available Illumina datasets information for the non-musculoskeletal tissues of blood, B cell, T cells, epidermis, dermis and liver together with linear regression

between age and DNA methylation analysis for cg16867657, cg24724428 and cg21572722 CpGs.

Table 3. 3. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the ELOVL2 region. N/A stands for "Not available" due data missing for the particular probe for the respective dataset.

		MSK tissues				non-MSK tissues						
		Cartilage	**Bone**	**Skeletal muscle**	**Tendon**	**Blood**	**B cell**	**T cell**	**Liver**	**Dermis**	**Epidermis**	**Breast**
ELVOL2 region	**Samples**	73	84	48	6	656	130	109	18	20	18	42
Age-CpG	**Reference**	Reynard	PMID: 29514638	PMID: 24304487	Riasat	GSE40279	GSE137634	GSE137634	GSE48325	GSE52980	GSE52980	GSE88883
cg16867657	**Slope**	0.003	0.004	0.003	N/A	0.004	0.005	0.005	0.004	0.004	0.003	0.002
	P value	0.002	< 0.0001	< 0.0001	N/A	<0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.003
	R square	0.128	0.388	0.799	N/A	0.698	0.621	0.753	0.759	0.834	0.852	0.200
cg24724428	**Slope**	0.000	0.004	0.003	N/A	0.004	0.004	0.003	0.002	0.004	0.006	0.002
	P value	0.663	<0.0001	<0.0001	0.255	<0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.193
	R square	0.003	0.425	0.799	N/A	0.528	0.387	0.278	0.636	0.926	0.895	0.042
cg21572722	**Slope**	0.002	0.003	0.003	N/A	0.002	0.002	0.002	0.006	0.002	0.003	0.001
	P value	0.001	<0.0001	<0.0001	0.393	<0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.345
	R square	0.156	0.315	0.480	N/A	0.453	0.377	0.397	0.800	0.879	0.902	0.022

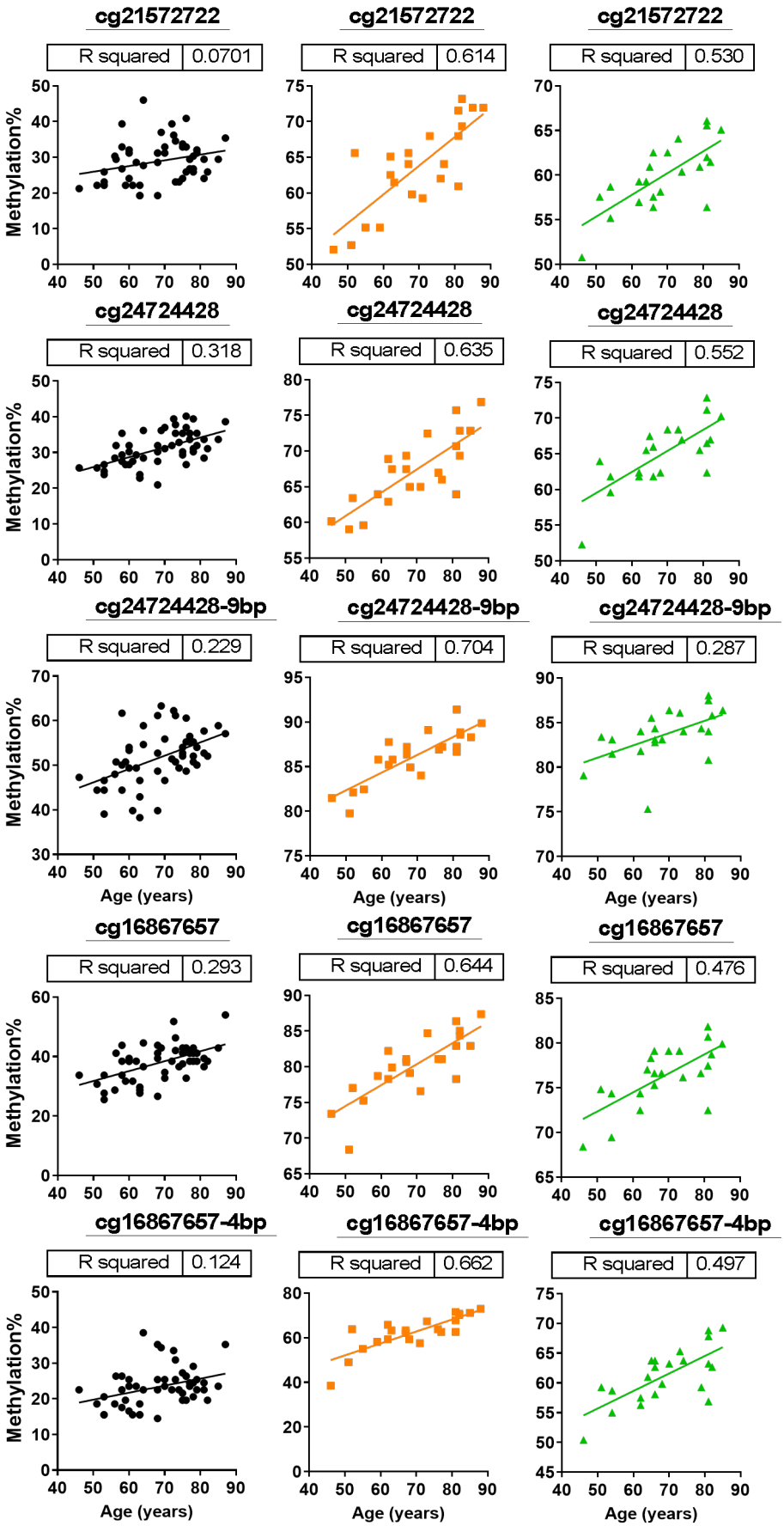

cg21572722
R squared 0.0701
cg21572722
R squared 0.614
cg21572722
R squared 0.530
cg24724428
R squared 0.318
cg24724428
R squared 0.635
cg24724428
R squared 0.552
cg24724428-9bp
R squared 0.229
cg24724428-9bp
R squared 0.704
cg24724428-9bp
R squared 0.287
cg16867657
R squared 0.293
cg16867657
R squared 0.644
cg16867657
R squared 0.476
cg16867657-4bp
R squared 0.124
cg16867657-4bp
R squared 0.662
cg16867657-4bp
R squared 0.497
Methylation%
Age (years)

Figure 3. 7. Linear regression analysis of age versus DNA methylation at the CpG sites located in the promoter of ELOVL2 for the tissues of cartilage, synovium and fat pad.

Table 3. 4. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the ELOVL2 region.

		Cartilage			Synovium			Fat Pad			Blood		
Region	CpG	Slope	R^2	p value	Slope	R^2	p value	Slope	R^2	p value	Slope	R^2	p value
ELOVL2	cg21572722	0.1585	0.06859	0.0607	0.4016	0.6127	< 0.0001	0.2436	0.5304	0.0001	0.4956	0.7221	0.0155
	cg24724428	0.2733	0.3121	< 0.0001	0.3271	0.6357	< 0.0001	0.2954	0.5526	< 0.0001	0.461	0.7285	0.0146
	cg24724428-9bp	0.2995	0.2285	0.0003	0.1987	0.703	< 0.0001	0.1373	0.2854	0.0104	0.1778	0.8857	0.0016
	cg16867657	0.3294	0.2815	< 0.0001	0.2944	0.645	< 0.0001	0.2128	0.4777	0.0004	0.2891	0.8513	0.0031
	cg16867657-4bp	0.1965	0.1233	0.0107	0.5354	0.6638	< 0.0001	0.2937	0.499	0.0002	0.5655	0.7488	0.0119

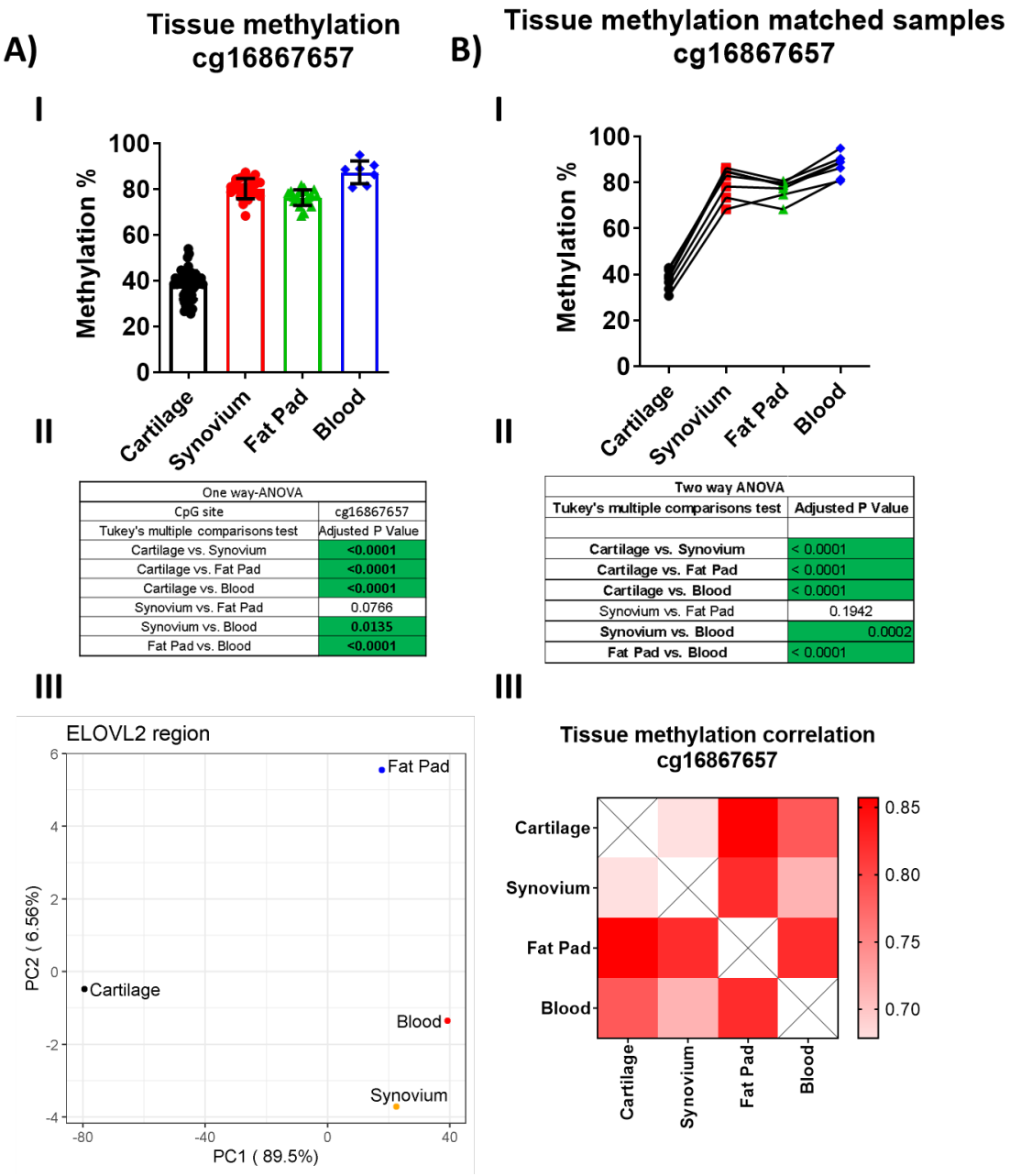


Figure 3. 8. Analysis of cg16867657 methylation level in different tissues. (A) Pooled samples from similar age groups I. The methylation level range for cg16867657 in the four tissues studied. Bar plot presents mean with S.D. II One-way ANOVA analysis table of the methylation levels found in the tissues studied. III Principal component analysis of the relationship between the tissues of cartilage, synovium, fat pad and blood (B) Quadruple matched samples for cartilage, synovium, fat pad and whole blood. I Methylation levels of cg16867657 as found in patient matched samples. II Two-way ANOVA paired analysis table for the matched samples. III Spearman correlation matrix analysis for the matched samples at cg16867657.

Table 3. 5. Table of the analysis of DNA methylation between tissues comparison A) for the unmatched tissues by the One-way ANOVA and B) for the matched samples using paired Two-way ANOVA.

A)

One-way ANOVA					
CpG site	cg21572722	cg24724428	cg24724428-9bp	cg16867657	cg16867657-4bp
Tukey's multiple comparisons test	Adjusted P Value				
Cartilage vs. Synovium	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
Cartilage vs. Fat Pad	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
Cartilage vs. Blood	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
Synovium vs. Fat Pad	0.1623	0.4569	0.3153	0.0766	>0.9999
Synovium vs. Blood	**0.0105**	0.9896	0.0521	**0.0135**	**<0.0001**
Fat Pad vs. Blood	**<0.0001**	0.5434	**0.0014**	**<0.0001**	**<0.0001**

B)

Two-way ANOVA					
CpG site	cg21572722	cg24724428	cg24724428-9bp	cg16867657	cg16867657-4bp
Tukey's multiple comparisons test	Adjusted P Value				
Cartilage vs. Synovium	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Cartilage vs. Fat Pad	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Cartilage vs. Blood	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Synovium vs. Fat Pad	0.5406	0.6945	0.3062	0.1942	0.9776
Synovium vs. Blood	**0.0065**	0.9974	**0.0002**	**0.0002**	**0.0096**
Fat Pad vs. Blood	**0.0004**	0.5797	**< 0.0001**	**< 0.0001**	**0.0223**

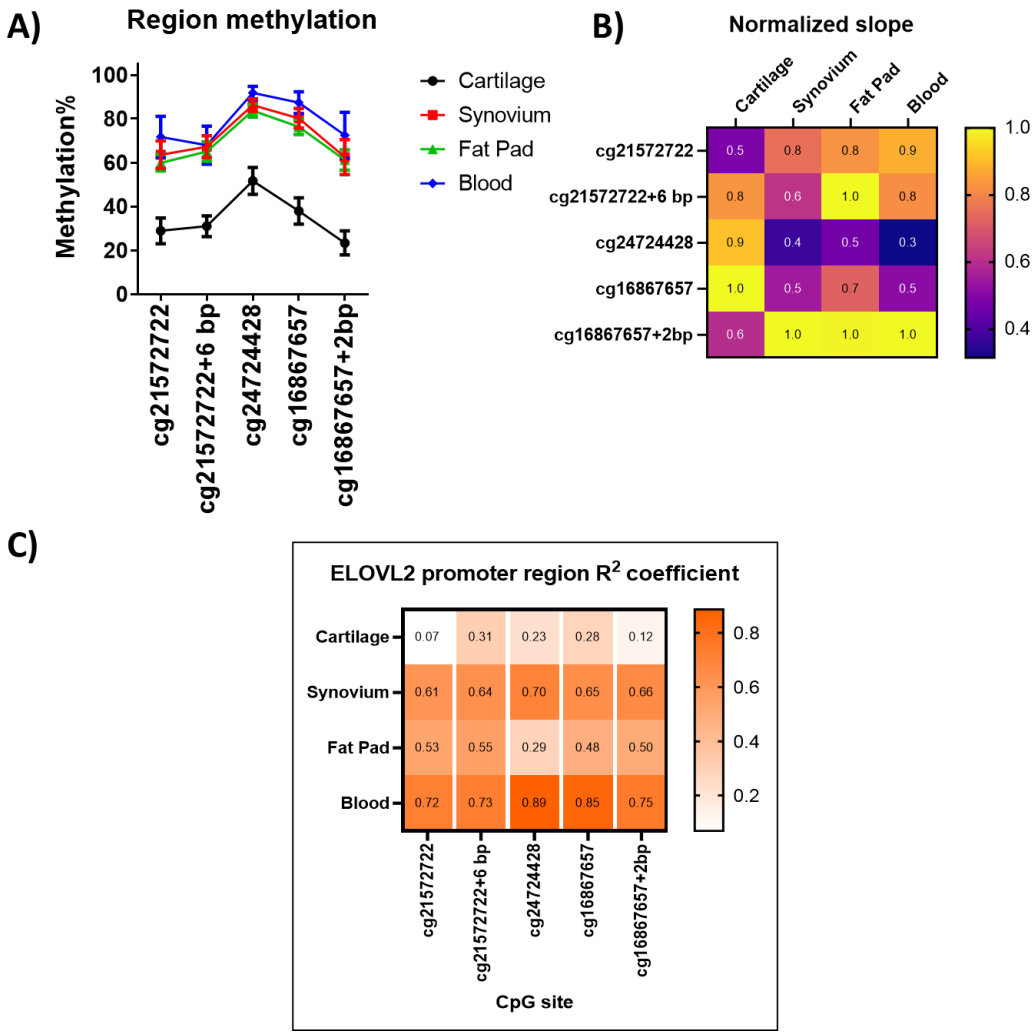


Figure 3. 9. Intra-regional analysis of the ELOVL2 promoter region that contains age-CpGs. A) Mean DNA methylation levels across five neighboring CpG sites with standard deviation across cartilage, synovium, fat pad and whole blood tissues. B) Normalized slope heatmap plot using the maximum slope value observed change per tissue. C) Heatmap plot of the correlation coefficient R^2 of the age-CpGs located in the ELOVL2 promoter across different tissues.

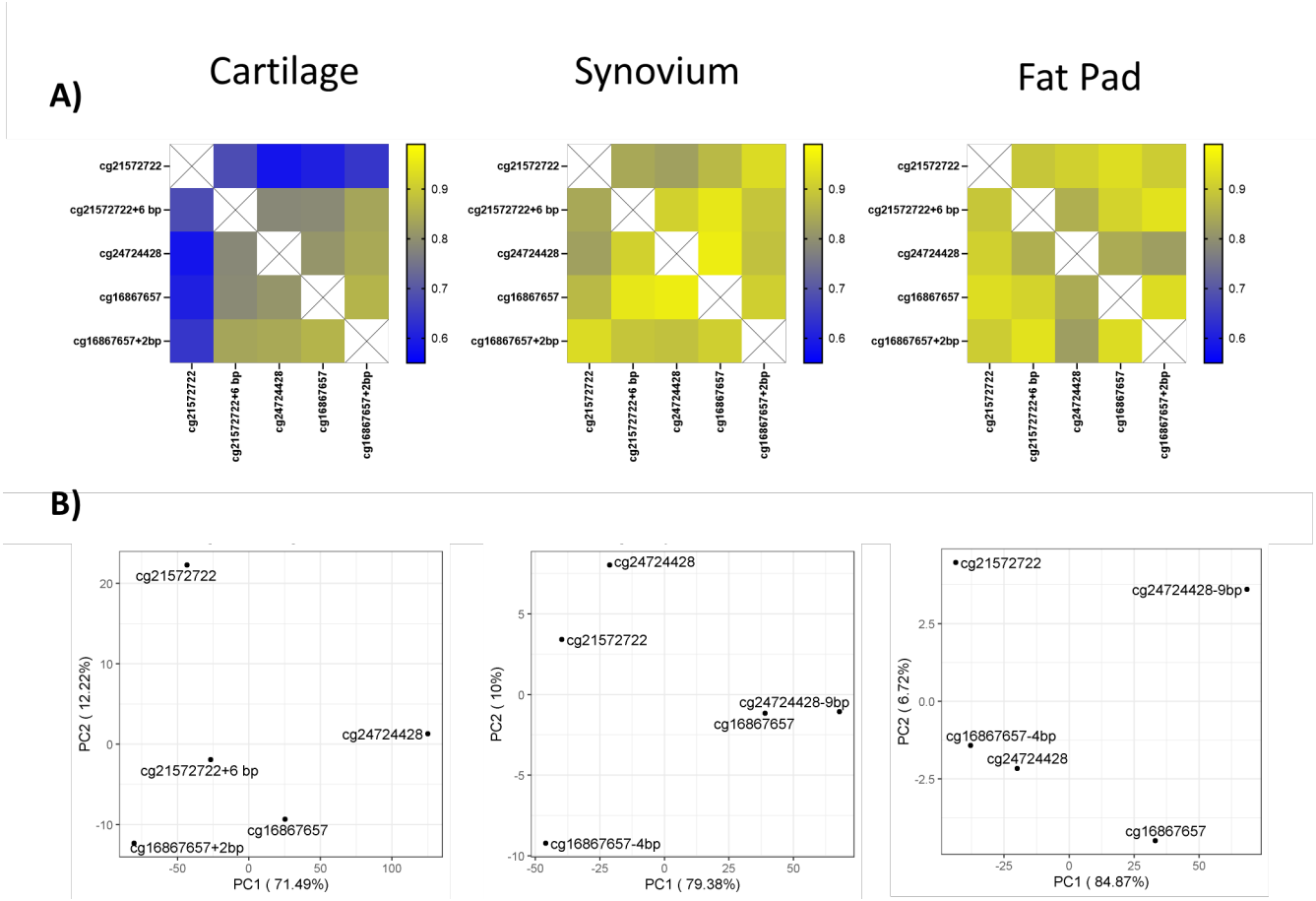


Figure 3. 10. Intra-regional analysis of the DNA methylation levels of age-CpGs across the ELOVL2 promoter region. A) Matrix correlation analysis for the articular tissues of cartilage, synovium and fat pad. B) Principal component analysis for cartilage, synovium and fat pad tissues.

3.3.4 FHL2 promoter region

Another investigated region is the promoter region of *FHL2* that contained three age-CpGs cg24079702, cg22454769, cg06639320 located at close proximity to each other at ~60bp. The region shows a consistent promoter chromatin signature in the E049 MSC-derived chondrocytes, with high levels of the histone marks H3K27Ac and H3K4Me1 across different cell types (Figure 3. 11).

The *FHL2* (Four-and-a-Half-LIM-only 2) gene encodes for a transcription co-factor protein a transcription co-activator (Morlon and Sassone-Corsi, 2003) involved stabilization of

transcription factor complexes (Tran *et al.*, 2016) and a complex stabilizer for focal adhesion (Park *et al.*, 2008).

Looking at the publicly available Illumina 450k musculoskeletal datasets (Figure 3. 12 and Table 3. 6) the highest correlation with age was observed for the CpG of cg22454769 where for bone tissue this had a R^2=0.266 and p value< 0.0001 followed by R^2=0.8517 and p value< 0.0001 for skeletal muscle tissue and for tendon R^2=N/A and p value=0.052. For tissues that are not part of the musculoskeletal system (Figure 3. 13), the highest correlation was observed for the dermis tissue with R^2=0.87 and p value< 0.0001 and the lowest correlation for the epidermis with R^2=0.23 and p value=0.04 (Table 3. 6). The CpG site with the lowest overall correlation was cg06639320 where this had the following analysis values for the musculoskeletal tissues (Figure 3. 12 and Table 3. 6) of bone (R^2=0.25, p value< 0.0001), skeletal muscle (R^2=0.54 and p value=< 0.0001) and tendon R^2=N/A and p value= 0.09. In non-musculoskeletal tissues (Figure 3. 13), the highest correlation was observed for the dermis tissue (R^2=0.83, p value< 0.0001) and the lowest for at a non-significant level for epidermis (R^2=0.15, p value= 0.11) (Table 3. 6).

Through the pyrosequencing assay an additional six neighbouring CpGs were also investigated, these being named according to the proximity to the nearest CpG probe coordinates. A calibration curve was generated as for the *ELOVL2* region CpGs and the results from this were used to adjust the raw methylation values.

The three cartilage age-CpGs under investigation all had significant age associated DNA methylation changes and moreover the additional six CpGs present in the assays also had significant age associated changes in the new cohort of cartilage samples. In the cartilage tissue (Figure 3. 14), the age-CpGs that had the highest R^2 correlation coefficient were cg06639320 (R^2=0.263 and p value <0.0001), cg24079702 (R^2=0.262 and p value <0.0001), cg22454769 (R^2=0.193 and p value= 0.0003) and the CpG with the lowest R^2 coefficient was cg06639320 (R^2=0.08 and p value= 0.02) (Table 3. 7). In other articular joint tissues such as in the synovium tissue (Figure 3. 15) all the CpGs investigated at this region were found to significantly change with age. The CpG sites with highest correlation R^2 were cg24079702+13bp (R^2=0.62 and p value<0.0001), cg24079702 (R^2=0.606 and p value<0.0001) and cg24079702+2bp (R^2=0.46 and p value= 0.0005) with the lowest correlation being observed for cg06639320+6bp (R^2=0.255 and p value= 0.016) (Table 3. 7).

For the fat pad tissue (Figure 3. 16), all nine CpG sites investigated were observed to change their DNA methylation levels with age at a significant level. The CpG sites with highest correlation coefficient R^2 were cg24079702+2bp (R^2=0.574 and p value<0.0001), cg22454769 (R^2=0.57 and p value<0.0001), cg06639320+9bp (R^2=0.555 and p value <0.0001) and the CpG with the lowest correlation with age was cg24079702+13bp (R^2=0.24 and p value=0.02) (Table 3. 7). In the blood tissue, high correlation values were observed for the CpG sites of cg24079702+9bp (R^2=0.94 and p value=0.0003), cg06639320+6bp (R^2=0.92 and p value =0.0007), cg24079702+13bp (R^2=0.85 and p value=0.0033) and the lowest correlating CpG site was that of cg24079702+2bp (R^2=0.6 and p value=0.04) (Table 3. 7).

All the tissues studied indicated that for all of the age-CpGs found in cartilage these were also present in synovium, fat pad and blood (Table 3. 7) similar to what was observed for the CpG sites located in the promoter of *ELOVL2.* Further analysis is done to determine whether there is any correlation to the changes between the tissues that could be linked. An analysis summary is presented for the cg06639320 age-CpG site (Figure 3. 17), however the same relationship is observed for all of the age-CpGs located in the *FHL2* promoter (Table 3. 8). To investigate the relationship the one-way-ANOVA was used as the analysis for the unpaired samples for the articular tissues and blood for the nine age-CpGs (Figure 3. 17 A) with this indicating that most tissues have significant differences in the DNA methylation levels observed with exception of the synovium and fat pad tissues where none of the CpGs showed significant different levels ($p<0.05$). Additionally, at cg24079702+13 a non-significant difference was found for cartilage vs synovium and cartilage vs fat pad tissues. For the matched samples (Figure 3. 17B) a paired two-way ANOVA was performed on the age-CpGs located in the promoter of *FHL2* where this presented a similar pattern of significance between the tissues as observed for the results of the unpaired samples. Principal component analysis on the unmatched samples (Figure 3. 17 A III) shows a separation on the first principal component between cartilage, blood and synovium+ fat pad where again the synovium and fat pad signature were undistinguishable on this PC1. Only PC2 was able to produce a bigger separation at a smaller level between all of the tissues. Matrix correlation of the tissue analysis on the matched samples revealed that the highest correlations between the tissues is observed for the fat pad and blood tissues ($R^2= 0.96$) (Figure 3. 17 BIII).

Observing the data for the all the age-CpGs across the *FHL2* promoter region for all the tissues investigated shows that the mean methylation pattern observed across all the tissues is mostly divided based on the tissue of origin (Figure 3. 18A). The cartilage tissue showed consistently lower methylation values (20-60%) with the exception of cg24079702+13bp that showed an average value of 15.5%, this is followed by the synovium and fat pad tissue where there was a higher methylation level across the CpGs (30-80%) with both tissues having very similar levels across the region and with cg24079702+13bp having values at 14.3% for synovium and 14.5% for the fat pad. Slope normalization was

performed in order to observe whether there are some CpG that are consistently observed to show a higher degree of change across the tissues and this analysis indicates that cg06639320 together with its proximal neighbours cg06639320+6bp and cg06639320+9bp consistently show the highest degree of age-related change across all the tissues (Figure 3. 18B). The correlation coefficient for the linear regression analysis of DNA methylation vs age, R^2 shows somewhat lower values for cartilage across the region (0.08-0.26) with the highest values being observed for the blood tissue (0.6-0.94) across all the CpGs this is followed by the synovium (0.26-0.62) and fat pad tissues (0.24-0.57) (Figure 3. 18C).

As the assay presents multiple CpGs, all with age-associated effects on DNA methylation the next analysis looks for any relationship that these CpGs have with each other. In order to investigate this relationship, the matrix region correlation analysis was performed (Figure 3. 19 A).This analysis generated a correlation coefficient for the relationship between one CpG site and its neighbours and when plotted on a heatmap graphic two clusters and an isolated CpG are observed. The first cluster containing cg06639320, cg06639320+6bp and cg06639320+9bp and the second cluster the CpGs cg24079702-14bp, cg22454769, cg24079702, cg24079702+2bp, cg24079702+9 but not the age-CpG cg24079702+13bp, this being outside of the two observed clusters. This relationship of clustering was strongly observed in the cartilage tissue and synovium however the pattern is more diffuse for the fat pad tissue. Principal component analysis of the age-CpGs also revealed a clustering effect where a small cluster can be observed as previously with cg06639320 and its close neighbours and a second bigger cluster with cg22454769, cg24079702 and their proximal neighbours again excluding cg24079702+13bp. This effect, as previously observed for the matrix correlation analysis, was also observed for the PCA plot in all the tissues investigated (Figure 3. 19B).

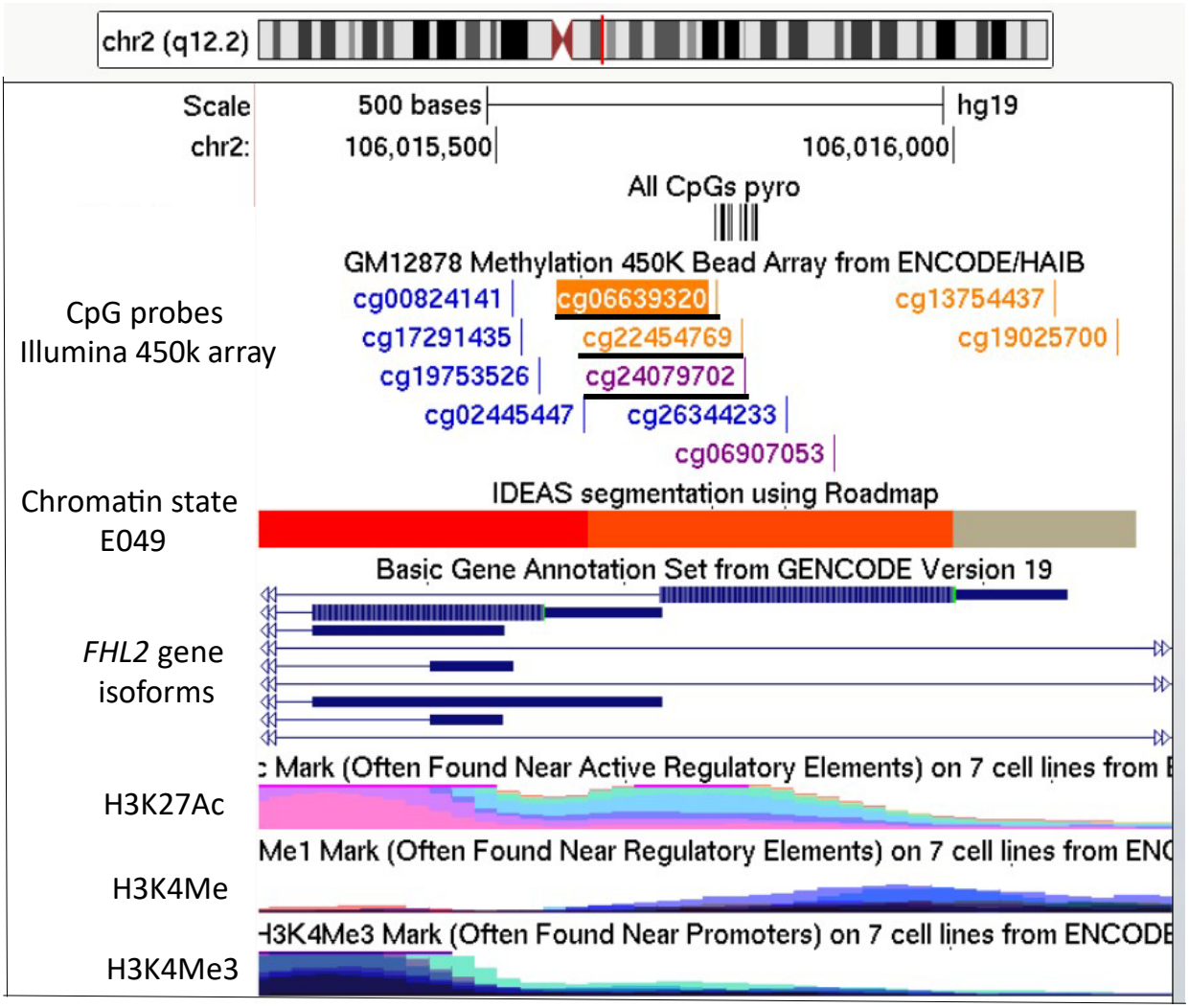


Figure 3. 11. UCSC genome browser view of the chromosomal location of the age-CpGs located in the promoter of FHL2 cg00639320, cg22454769, cg24079702. The track All CpGs pyro provides the relative location of the CpGs included in the pyrosequencing assay; the chromatin state information is presented for the E049 cell line. Information about the nearby genes/isoforms is presented below this followed by information regarding the histone mark intensity present at the region in 7 cell lines provided by the ENCODE project.

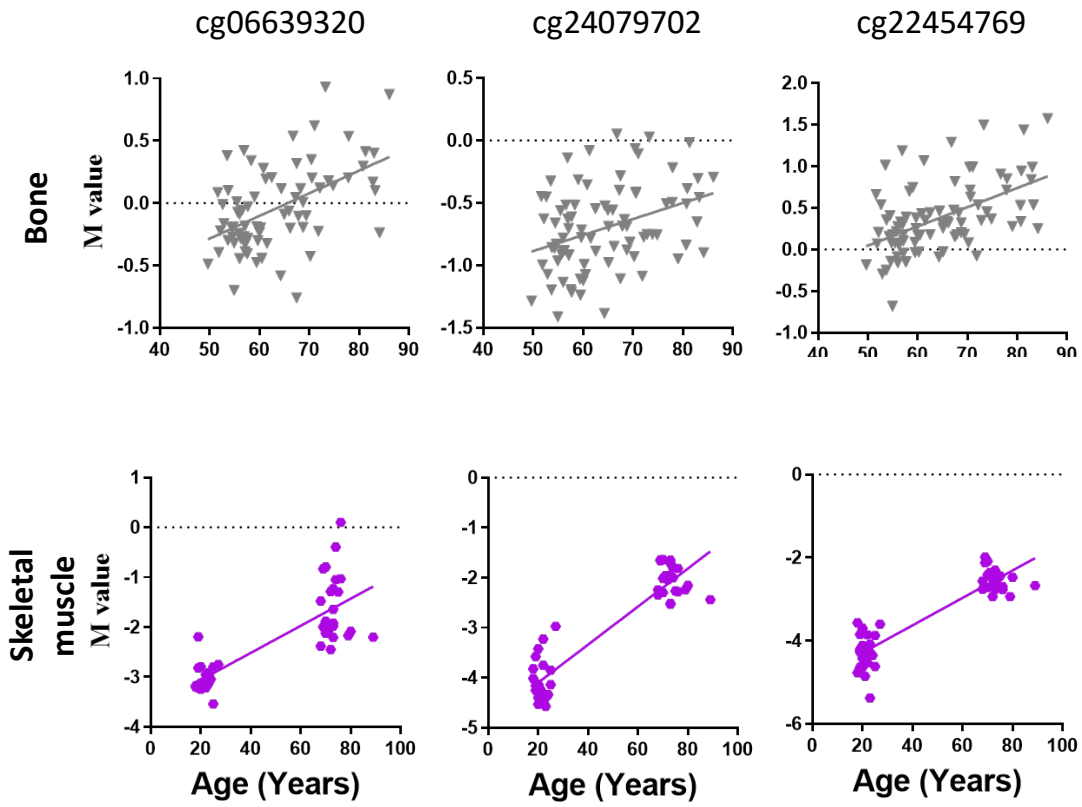


Figure 3. 12. Publicly available Illumina 450k datasets information for the musculoskeletal tissues of bone and skeletal muscle and linear regression between age and DNA methylation analysis for cg06639320, cg24079702 and cg22454769.

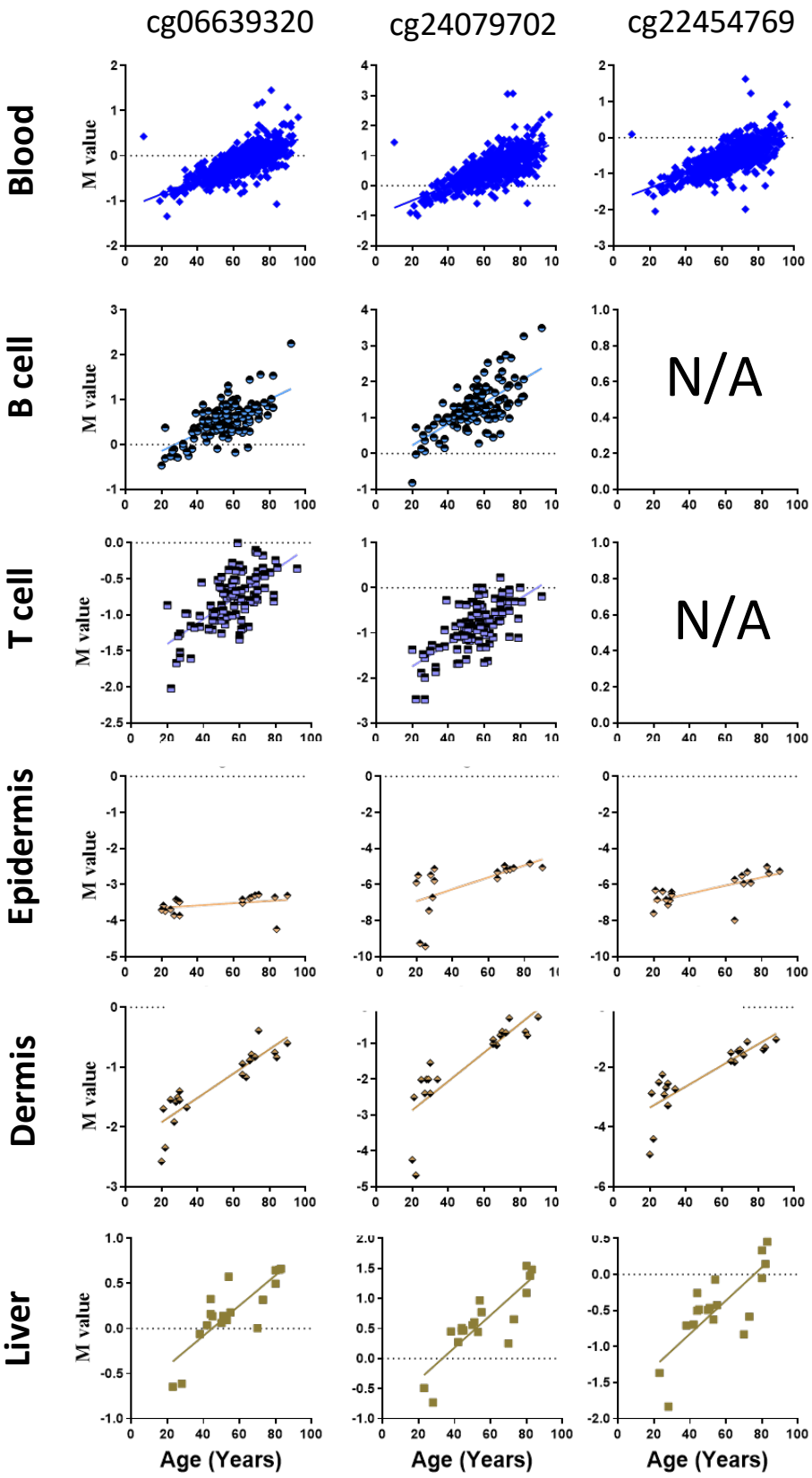

cg06639320
cg24079702
cg22454769
Blood
B cell
T cell
Epidermis
Dermis
Liver
M value
N/A
N/A
Age (Years)
Age (Years)
Age (Years)

Figure 3. 13. Publicly available Illumina datasets information for the non-musculoskeletal tissues of blood, B cell, T cells, epidermis, dermis and liver together with linear regression between age and DNA methylation analysis for cg06639320, cg24079702 and cg22454769.

Table 3. 6. Table summary of the linear regression analysis of the available methylation datasets with linear regression analysis for age vs. DNA methylation for musculoskeletal and non-musculoskeletal tissues for the FHL2 region. N/A stands for “Not available” due data missing for the particular probe for the respective dataset.

		MSK tissues				non-MSK tissues						
		Cartilage	**Bone**	**Skeletal muscle**	**Tendon**	**Blood**	**B cell**	**T cell**	**Liver**	**Dermis**	**Epidermis**	**Breast**
FHL2 region	**Samples**	73	84	48	6	656	130	109	18	20	18	42
Age-CpG	**Reference**	Reynard	PMID: 29514638	PMID: 24304487	Riasat	GSE40279	GSE137634	GSE137634	GSE48325	GSE52980	GSE52980	GSE88883
cg06639320	**Slope**	0.002	0.003	0.003	N/A	0.003	0.003	0.003	0.003	0.003	0	0.003
	P value	< 0.0001	< 0.0001	< 0.0001	0.088	<0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.111	0.001
	R square	0.238	0.251	0.544	N/A	0.527	0.451	0.438	0.688	0.828	0.151	0.25
cg24079702	**Slope**	0.002	0.002	0.002	N/A	0.003	N/A	N/A	0.004	0.003	0	0.003
	P value	0.001	0.001	< 0.0001	N/A	<0.0001	N/A	N/A	< 0.0001	< 0.0001	< 0.0001	0.003
	R square	0.16	0.131	0.844	N/A	0.474	N/A	N/A	0.613	0.872	0.639	0.201
cg22454769	**Slope**	0.004	0.004	0.003	N/A	0.004	0.004	0.004	0.004	0.005	0	0.004
	P value	< 0.0001	0	< 0.0001	0.053	<0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.043	0.001
	R square	0.235	0.266	0.852	N/A	0.524	0.462	0.441	0.706	0.872	0.232	0.256

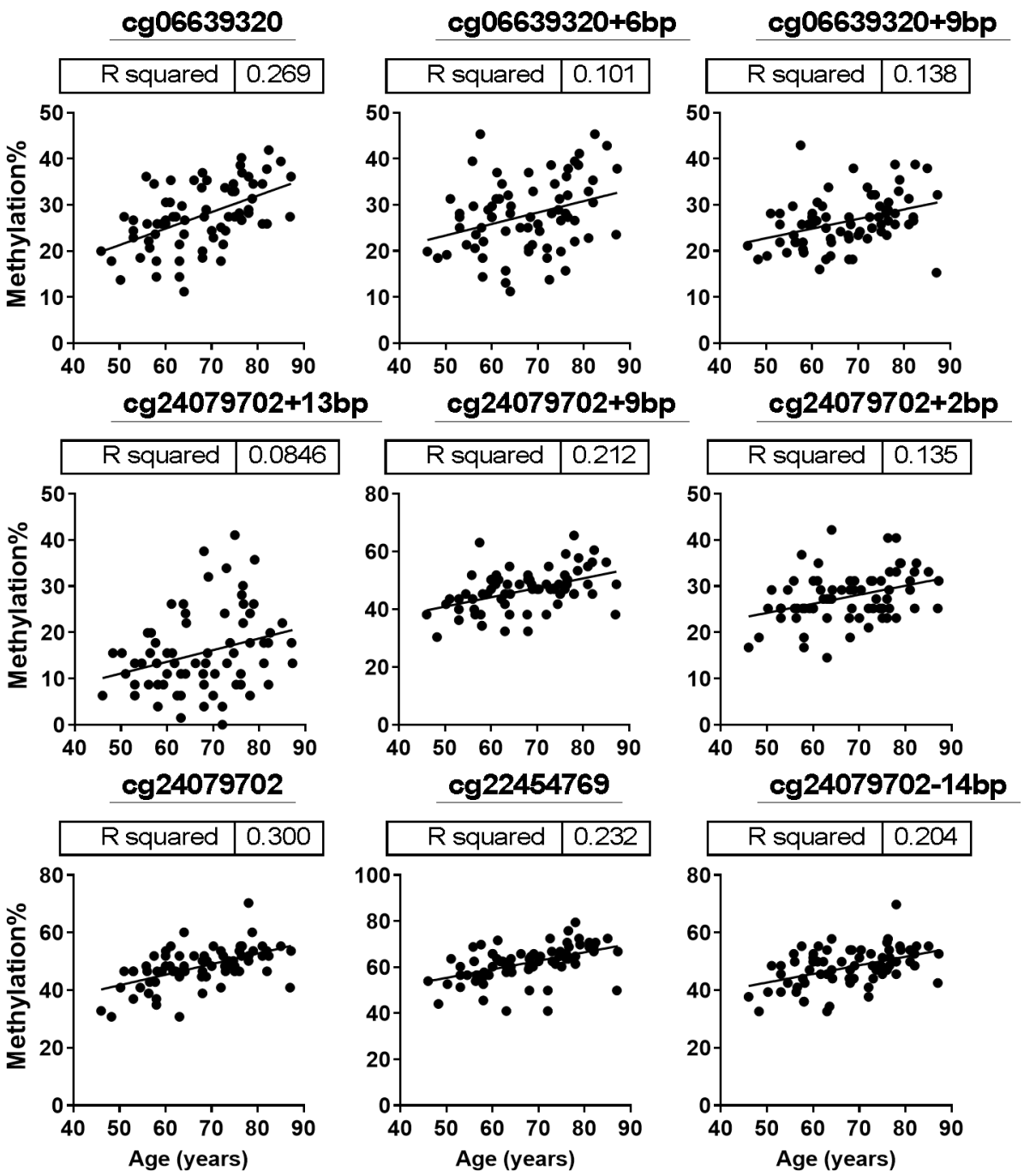


Figure 3. 14. Linear regression analysis between age and DNA methylation results for the nine CpGs located in the promoter of FHL2.

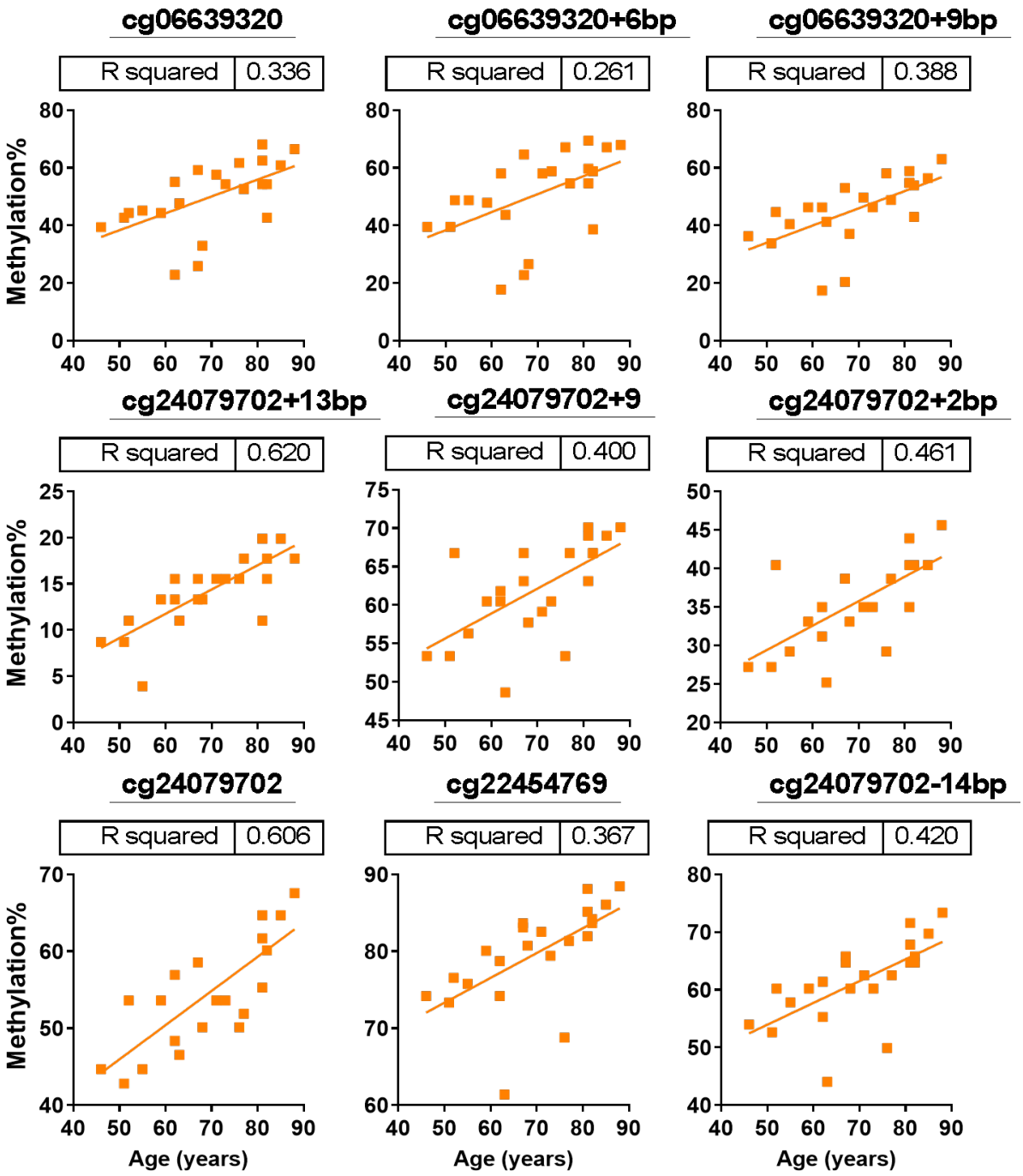


Figure 3. 15. Linear regression analysis between age and DNA methylation performed in the synovium tissue for the nine age-CpGs located in the promoter of FHL2.

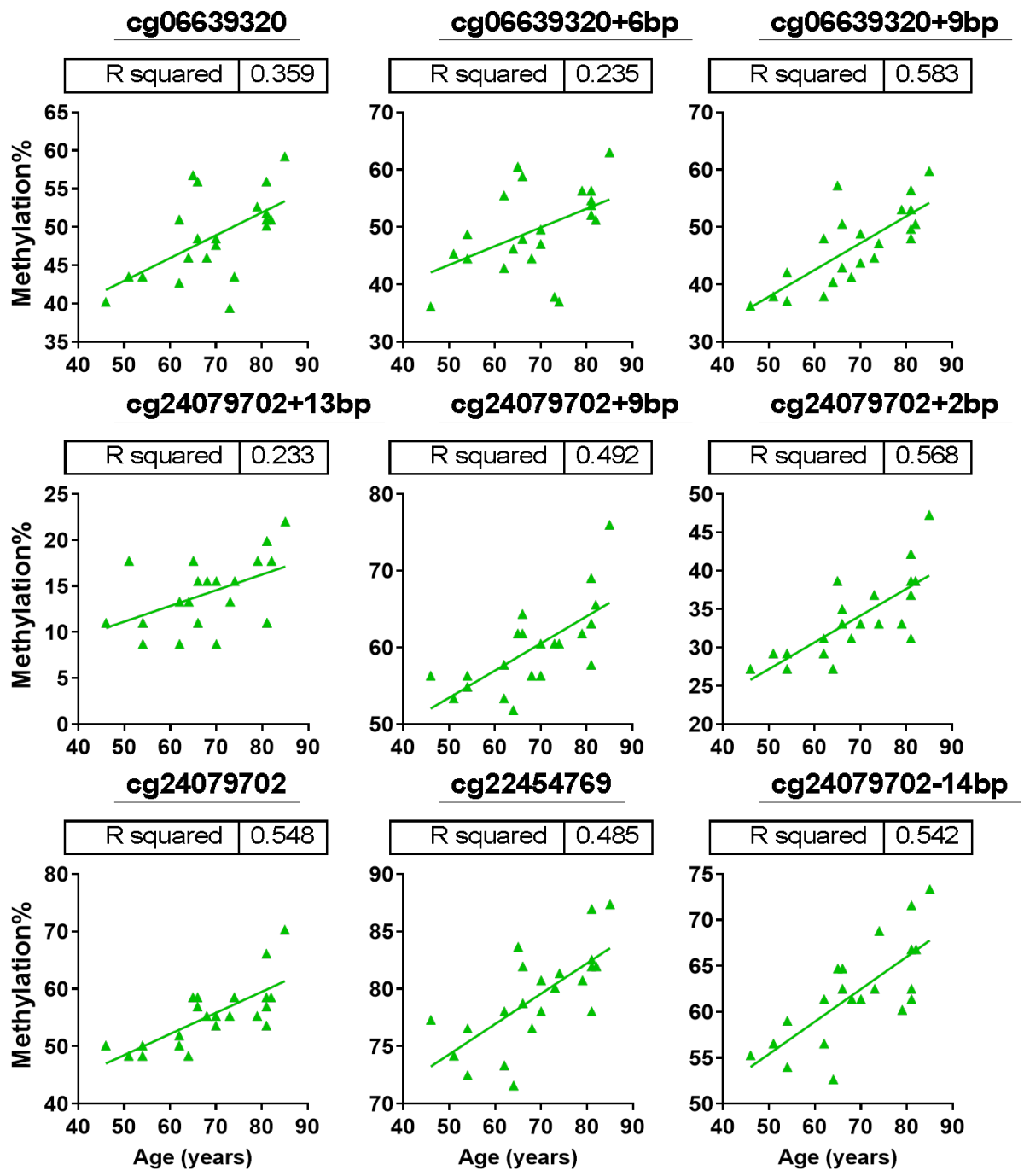


Figure 3. 16. Linear regression analysis between age and DNA methylation for the nine age-CpGs located in the promoter of FHL2 in the fat pad tissue.

Table 3. 7. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the FHL2 region.

		Cartilage			Synovium			Fat Pad			Blood		
Region	CpG	Slope	R^2	p value	Slope	R^2	p value	Slope	R^2	p value	Slope	R^2	p value
FHL2	cg06639320	0.4233	0.2634	< 0.0001	0.5925	0.3321	0.005	0.2968	0.3579	0.0033	0.9546	0.5969	0.0417
	cg06639320+6bp	0.261	0.0802	0.0223	0.6258	0.2553	0.0165	0.3218	0.2355	0.0221	0.9053	0.7487	0.0119
	cg06639320+9bp	0.2226	0.1145	0.0058	0.6073	0.3769	0.0024	0.4662	0.5737	< 0.0001	1.064	0.7656	0.0099
	cg24079702-14bp	0.3099	0.1813	0.0004	0.3779	0.4202	0.0011	0.3532	0.544	< 0.0001	0.4552	0.8467	0.0033
	cg22454769	0.3679	0.1913	0.0003	0.3228	0.3665	0.0028	0.2655	0.4912	0.0003	0.2234	0.8398	0.0037
	cg24079702	0.3703	0.2621	< 0.0001	0.4447	0.606	< 0.0001	0.3695	0.5549	< 0.0001	0.4478	0.9418	0.0003
	cg24079702+2bp	0.1972	0.1142	0.0059	0.3179	0.4608	0.0005	0.3492	0.5699	< 0.0001	0.52126	0.8473	0.0033
	cg24079702+9	0.3257	0.1745	0.0005	0.3243	0.3999	0.0016	0.3587	0.5069	0.0002	0.2583	0.9161	0.0007
	cg24079702+13bp	0.2791	0.08255	0.0203	0.2629	0.6198	< 0.0001	0.1792	0.2569	0.0161	0.4511	0.6836	0.0218

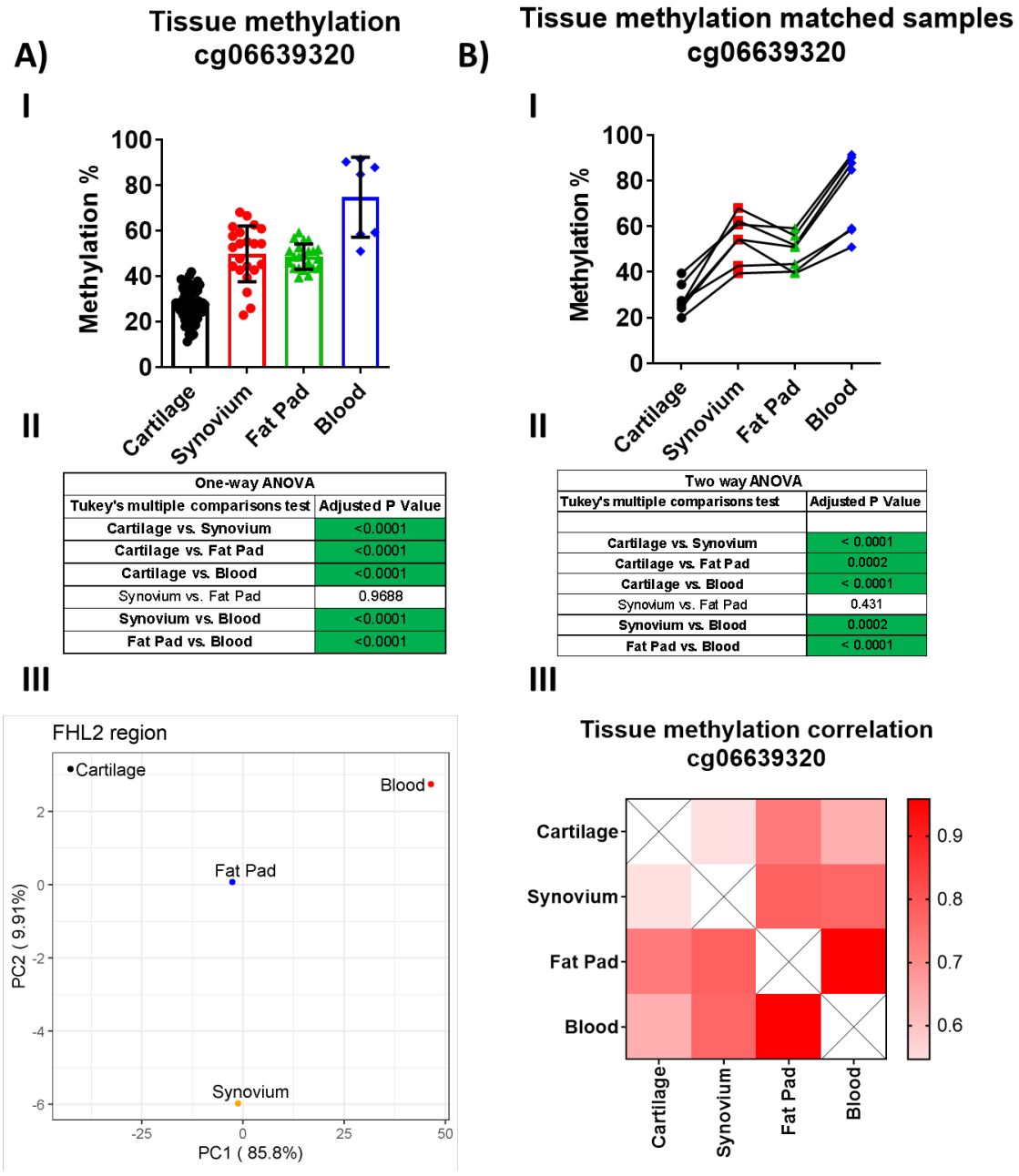


Figure 3. 17. Analysis of cg06639320 methylation level in different tissues. (A) Pooled samples from similar age groups I. The methylation level range for cg06639320 in the four tissues studied. Bar plot presents mean with S.D. II One-way ANOVA analysis table of the

methylation levels found in the tissues studied. III Principal component analysis of the relationship between the tissues of cartilage, synovium, fat pad and blood (B) Quadruple matched samples for cartilage, synovium, fat pad and whole blood. I Methylation levels of cg06639320 as found in patient matched samples. II Two-way ANOVA paired analysis table for the matched samples. III Spearman correlation matrix analysis for the matched samples at cg06639320 versus tissue of origin.

Table 3. 8. Table of the analysis of DNA methylation between tissues comparison A) for the unmatched tissues by the One-way ANOVA and B) for the matched samples using paired Two-way ANOVA.

FHL2 promoter region

A)

One-way ANOVA									
CpG site	cg06639320	cg06639320+6bp	cg06639320+9bp	cg24079702+13bp	cg24079702+9bp	cg24079702+2bp	cg24079702	cg22454769	cg24079702-14bp
Tukey's multiple comparisons test	**Adjusted P Value**								
Cartilage vs. Synovium	<0.0001	<0.0001	<0.0001	0.8986	<0.0001	<0.0001	0.0007	<0.0001	<0.0001
Cartilage vs. Fat Pad	<0.0001	<0.0001	<0.0001	0.9193	<0.0001	0.0001	0.0001	<0.0001	<0.0001
Cartilage vs. Blood	<0.0001	<0.0001	<0.0001	0.0004	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Synovium vs. Fat Pad	0.9688	0.9856	0.9768	>0.9999	0.7835	0.7	0.9807	0.9985	0.9823
Synovium vs. Blood	<0.0001	<0.0001	<0.0001	0.0003	<0.0001	<0.0001	<0.0001	0.0058	0.0001
Fat Pad vs. Blood	<0.0001	<0.0001	<0.0001	0.0004	<0.0001	<0.0001	0.0002	0.004	0.0004

B)

Two-way ANOVA									
CpG site	cg06639320	cg06639320+6bp	cg06639320+9bp	cg24079702+13bp	cg24079702+9bp	cg24079702+2bp	cg24079702	cg22454769	cg24079702-14bp
Tukey's multiple comparisons test	**Adjusted P Value**								
Cartilage vs. Synovium	< 0.0001	< 0.0001	< 0.0001	0.9421	< 0.0001	0.0035	0.0038	< 0.0001	< 0.0001
Cartilage vs. Fat Pad	< 0.0001	< 0.0001	< 0.0001	0.3029	< 0.0001	0.0007	0.0001	< 0.0001	< 0.0001
Cartilage vs. Blood	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Synovium vs. Fat Pad	0.4309	0.2378	0.9998	0.6082	0.9051	0.877	0.3506	0.9739	0.517
Synovium vs. Blood	0.0002	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Fat Pad vs. Blood	< 0.0001	< 0.0001	< 0.0001	0.0014	< 0.0001	< 0.0001	0.0002	< 0.0001	< 0.0001

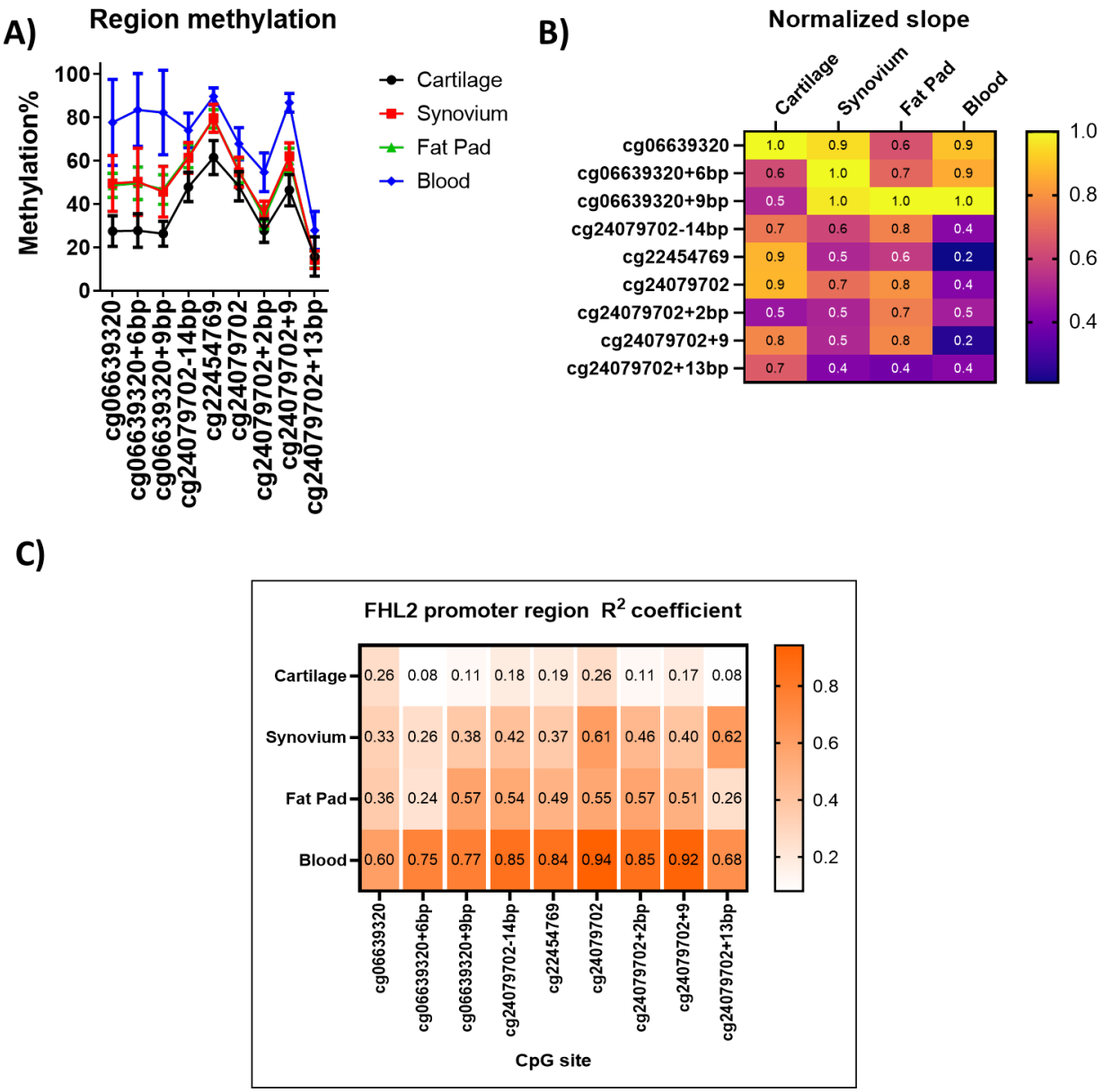


Figure 3. 18. Intra-regional analysis of the age-CpGs contained in the FHL2 promoter. A) Mean methylation with standard deviation across the region and representation in different tissues. B) Heatmap of normalized slope values across cartilage, synovium, fat pad and blood. C) Correlation coefficient (R^2) for the age-CpGs in the investigated tissues.

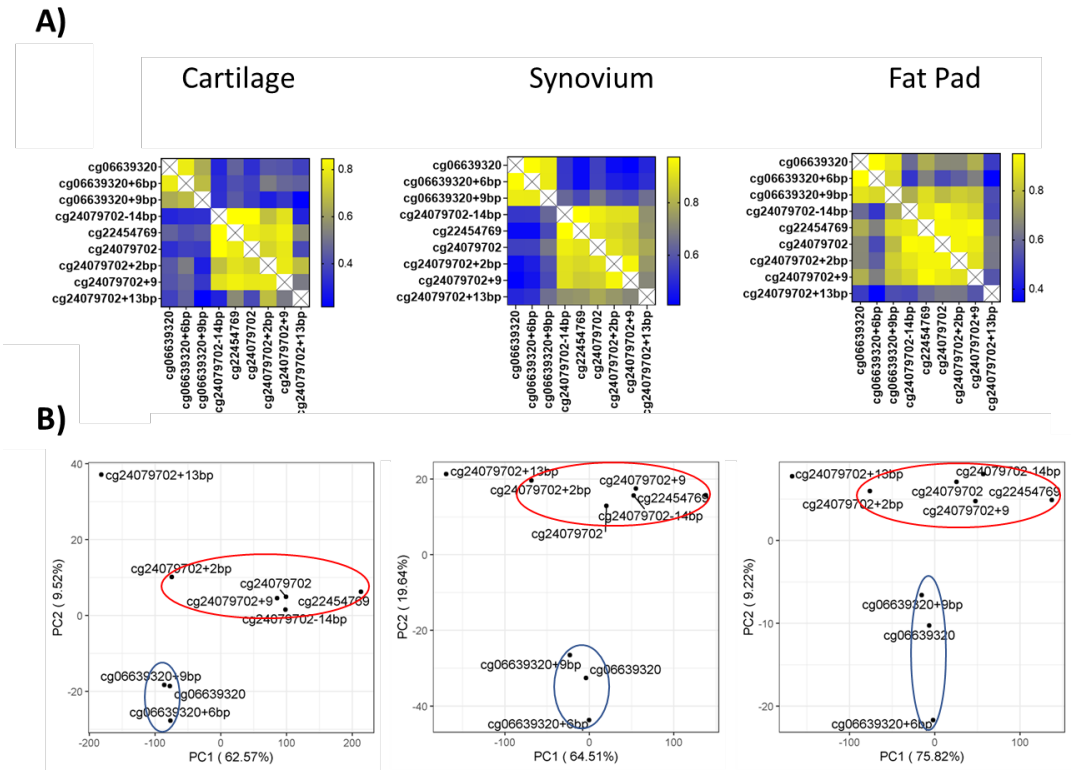


Figure 3. 19. Intra-regional correlation analysis of the DNA methylation levels of age-CpGs across the FHL2 promoter region. A) Matrix correlation analysis for the articular tissues of cartilage, synovium and fat pad. B) Principal component analysis for cartilage, synovium and fat pad tissue samples.

3.3.5 PCDHGA1 enhancer region

Another region under investigation is that of the age-CpGs cg14566959 located in a putative enhancer region in the intron of *PCDHGA1*. This region has low levels of the histone marks H3K27Ac, H3K4Me1 and H3K4Me3 and is located intron of a cluster of the *PCDHGA* gene family and potentially the exon of *PCDHGA1* (Figure 3. 20). The *PCDHGA* family (protocadherin gamma) cluster of genes likely play a role as cadherin-like cell adhesion proteins with possible roles in the control of specific cell-cell connections (Dilling *et al.*, 2017).

In the publicly available datasets, for the musculoskeletal tissues (Figure 3. 21 A) cg14566959 was observed for the bone tissue to have a non-significant correlation with age with R^2=0.042 and p value=0.061, however for the skeletal muscle tissue this CpG had a high significant (R^2=0.465 and p value<0.0001) and no data for this CpG was available for the tendon tissue (Table 3. 9). For the non-musculoskeletal tissues (Figure 3. 21 B), the highest

correlation was observed for the epidermis tissue R^2=0.743 and p value<0.0001 and the lowest for the breast tissues seen as a non-significant relation with age (R^2=0.048 and p value=0.1605) (Table 3. 9).

For the pyrosequencing analysis an additional two CpGs were also analysed as part of the assay. In the cartilage tissue (Figure 3. 22), for cg14566959 there was a significant relationship observed with R^2=0.196 and p value=0.0002, however none of the additional CpGs included in the assay had any significant correlation with age (Table 3. 10). Looking for the relationship between age and DNA methylation in the synovium tissue (Figure 3. 22) this indicated that again only cg14566959 had a significant correlation with age with R^2=0.233 and p value<0.0001 while the other CpGs had no significant change with age (Table 3. 10). Interestingly in the fat pad tissue (Figure 3. 22), no CpG site had a significant change with age with cg14566959 having R^2=0.136 and p value=0.09 (Table 3. 10). For the blood tissue none of the CpGs significantly changed with age with cg14566959 having R^2=0.2133 and p value=0.29 (Table 3. 10).

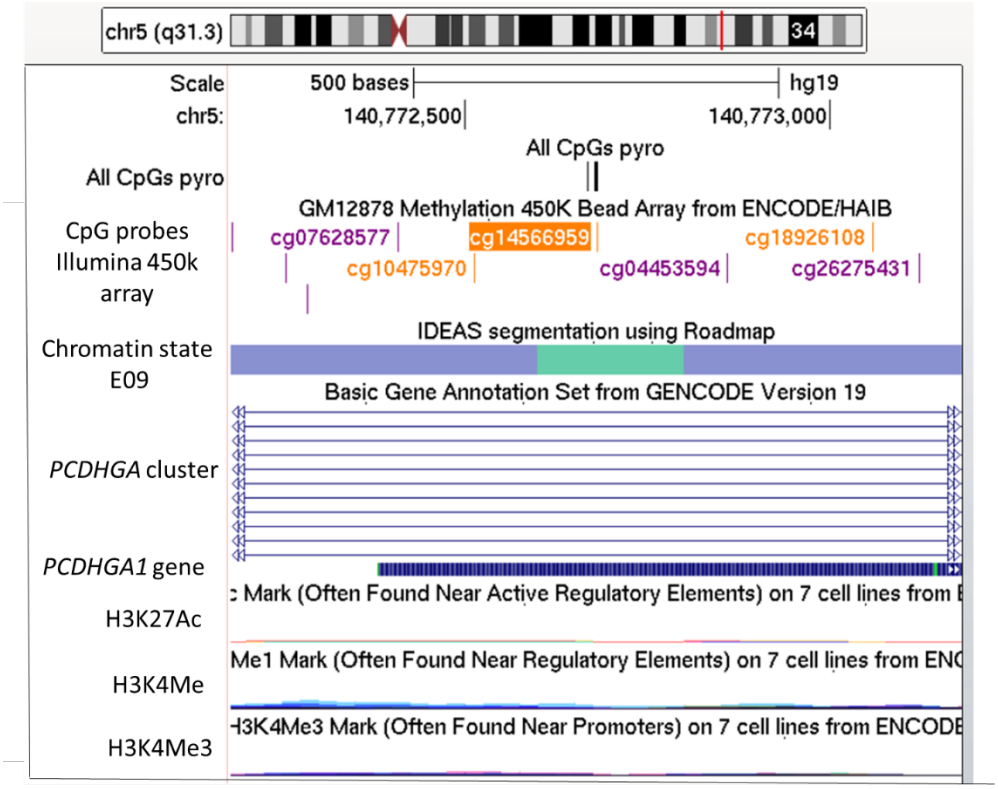


Figure 3. 20. UCSC genome browser view of the chromosomal location of the age-CpG located in the PCDHGA1 potential enhancer region cg14566959. The track All CpGs pyro provides the relative location of the CpGs included in the pyrosequencing assay; the

chromatin state information is presented for the E049 cell line. Information about the nearby genes/isoforms is presented and information regarding the histone mark intensity present at the region in 7 cell lines provided by the ENCODE project.

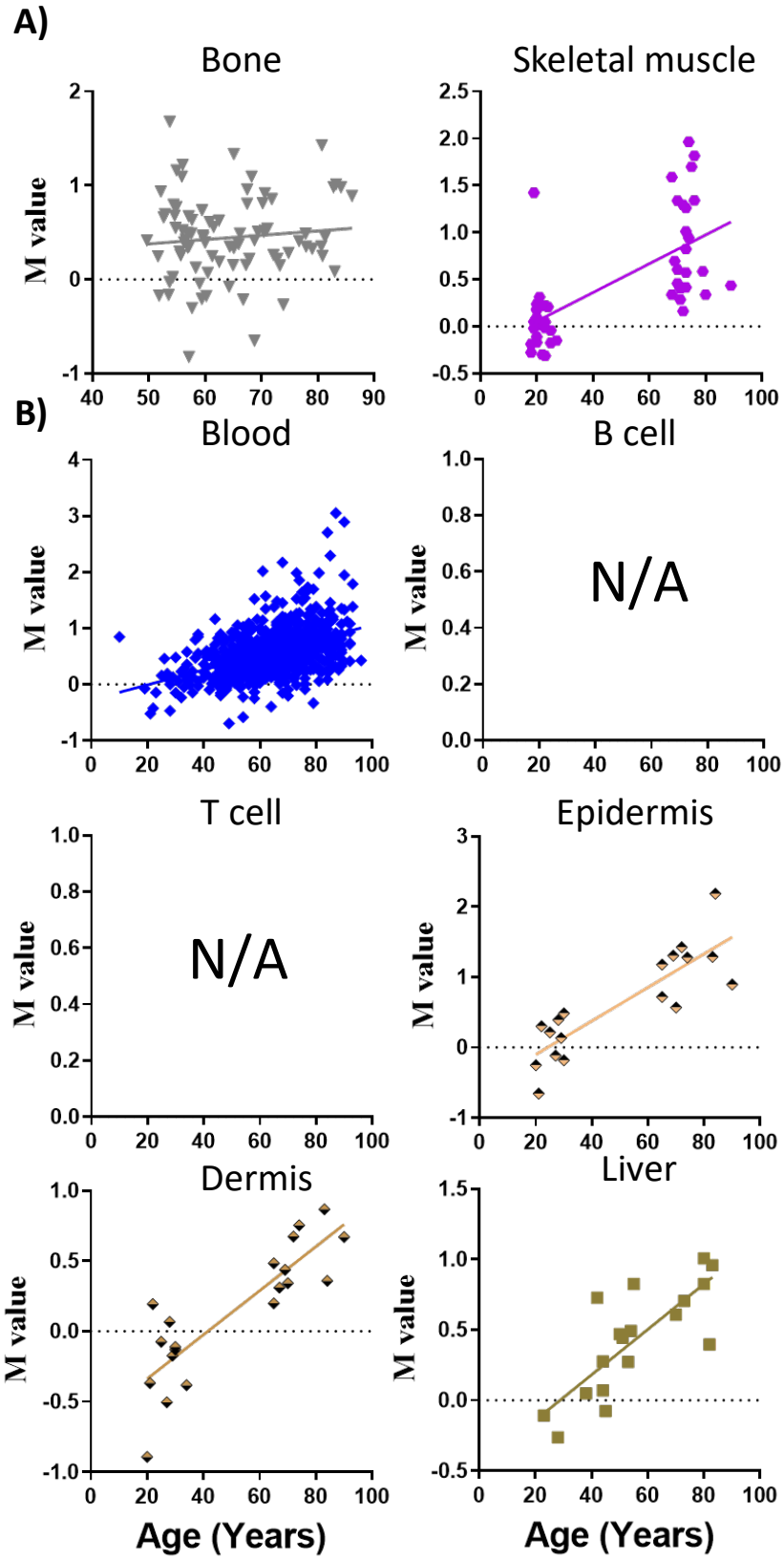


Figure 3. 21. Publicly available Illumina datasets information with linear regression analysis between age and DNA methylation for the musculoskeletal tissues of bone and skeletal

muscle (A) and non-musculoskeletal tissue of blood, B cell, T cell, dermis, epidermis and liver (B) for the cg14566959 CpG.

Table 3. 9. Table summary of the linear regression analysis of the available methylation datasets with linear regression analysis for age vs. DNA methylation for musculoskeletal and non-musculoskeletal tissues for the PCDHGA1 enhancer region. N/A stands for “Not available” due data missing for the particular probe for the respective dataset.

		MSK tissues				non-MSK tissues						
PCDHGA1 region		**Cartilage**	**Bone**	**Skeletal muscle**	**Tendon**	**Blood**	**B cell**	**T cell**	**Liver**	**Dermis**	**Epidermis**	**Breast**
Age-CpG	**Reference**	Reynard	PMID: 29514638	PMID: 24304487	Riasat	GSE40279	GSE137634	GSE137634	GSE48325	GSE52980	GSE52980	GSE88883
cg14566959	**Slope**	0.003	0.002	0.002	N/A	0.002	N/A	N/A	0.003	0.003	0.004	0.002
	P value	0.001	0.062	<0.0001	N/A	<0.0001	N/A	N/A	<0.0001	<0.0001	<0.0001	0.161
	R square	0.150	0.042	0.465	N/A	0.213	N/A	N/A	0.617	0.726	0.743	0.049

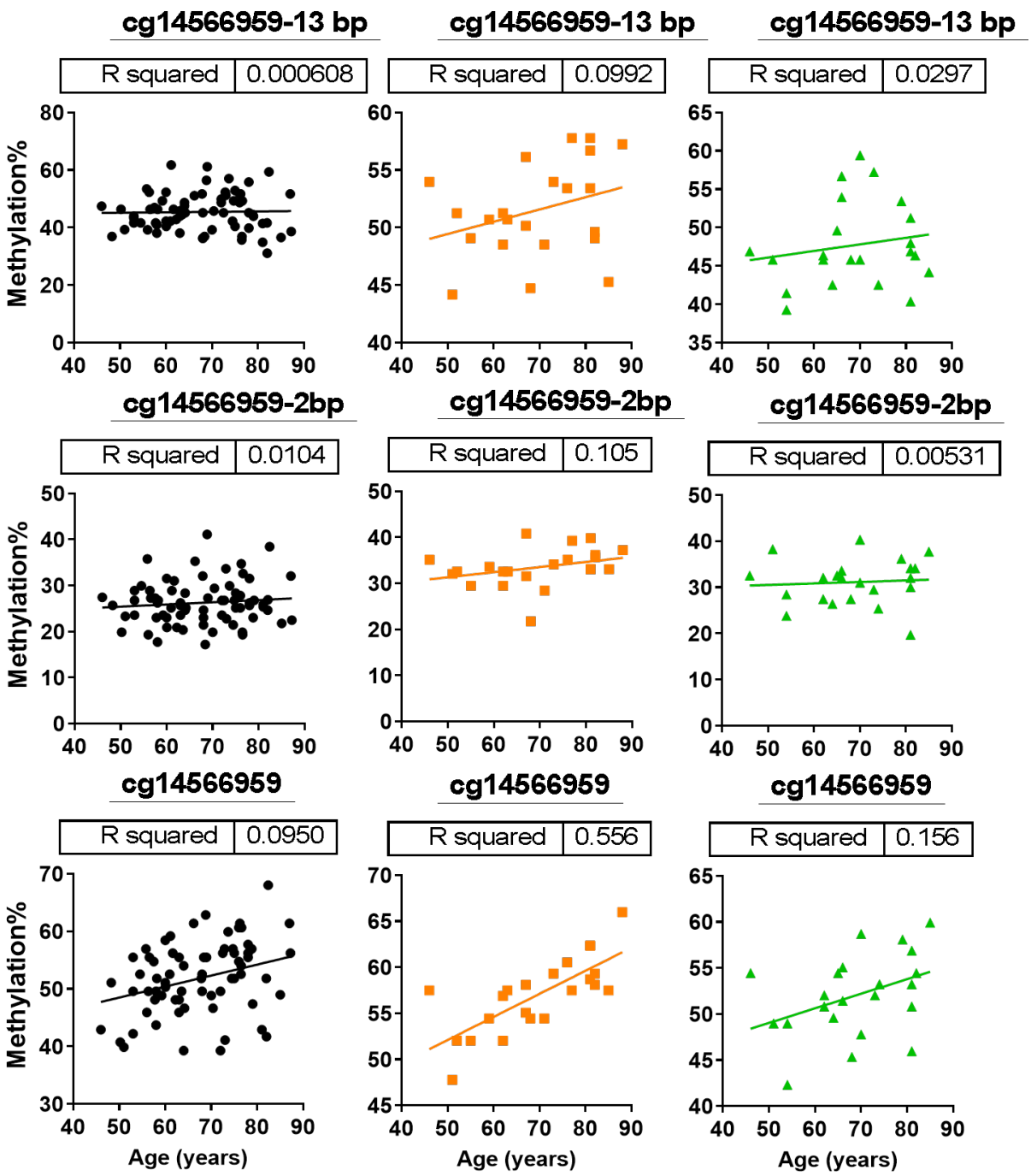


Figure 3. 22. Linear regression analysis of age versus DNA methylation at the CpG sites located in the enhancer region near PCDHGA1 for the tissues of cartilage, synovium and fat pad.

Table 3. 10. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the PCDHGA1 enhancer region.

		Cartilage			Synovium			Fat Pad			Blood		
Region	CpG	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value
PCDHGA1	cg14566959-13 bp	0.1118	0.02693	0.1915	0.1059	0.09802	0.156	0.06788	0.01866	0.5444	0.357	0.391	0.1331
	cg14566959-2bp	0.07928	0.02253	0.2327	0.1103	0.1042	0.1428	0.01732	0.001525	0.863	0.1711	0.04767	0.6381
	cg14566959	0.2738	0.1962	0.0002	0.2502	0.5562	< 0.0001	0.1475	0.1359	0.0914	0.2559	0.2133	0.2968

3.3.6 PCDBH1 enhancer region

The cg23500537 CpG is located in a putative enhancer region similar to what was found for cg14566959. The CpG site is located ~14kb upstream from the *PCDHB1* gene and ~27kb downstream from the *PCDHA* gene family cluster. Histone marks are enriched for this region where H3K27Ac, H3K4Me1 and in some instances H3K4Me3 is seen to be present. Nearby genes around this region are those in the PCDHB family (Figure 3. 23). The *PCDHB1* (protocadherin beta 1) gene is thought to be involved in neural functions such as cell density and dendritic arborization (Miyake *et al.*, 2011), however no known functions outside the brain tissue are known.

Analysing the CpG for age-related changes in DNA methylation in Illumina 450k datasets for musculoskeletal tissues (Figure 3. 24) indicated that for bone (R^2=0.257 and p value< 0.0001), skeletal muscle (R^2=0.6234 and p value< 0.0001) there is a significant change with age (Table 3. 11). In non-musculoskeletal tissues (Figure 3. 24) the highest and significant correlation with age was observed for dermis R^2=0.9409 and p value< 0.0001 and the lowest and still significant for the tissue of breast R^2=0.101 and p value=0.0402, whereas the liver had a non-significant change with age (R^2=0.1151 and p value=0.1479) (Table 3. 11).

The pyrosequencing assay had an additional two CpG site and this revealed that for the cartilage tissue (Figure 3. 25) the CpG of cg23500537 had a highly significant change with age (R^2=0.536 and p value=< 0.0001), with its neighbours cg23500537+2bp (R^2=0.587 and p value=< 0.0001) and cg23500537+9bp (R^2=0.311 and p value=< 0.0001) also significantly changing with age (Table 3. 12). For the investigation in the synovium tissue (Figure 3. 25) cg23500537 (R^2=0.305 and p value=0.0006) also showed a significant change with age alongside the neighbouring CpG sites of cg23500537+2bp (R^2=0.2761 and p value=0.0012) and cg23500537+9bp (R^2=0.3551 and p value=0.0002) (Table 3. 12). Additionally, all three CpG sites were also significantly changing with age for the fat pad tissue (Figure 3. 25) with cg23500537 (R^2=0.1447 and p value=0.0265), cg23500537+2bp (R^2=0.141 and p value=0.0286) and cg23500537+9bp (R^2=0.3762 and p value=0.0001) (Table 3. 12) passing the significance threshold of p<0.05. Interestingly for the blood tissue there was no significant change with age at any of the investigated sites with cg23500537 (R^2=0.37 and p value=0.147), cg23500537+2bp (R^2=0.324 and p value=0.182) and cg23500537+9bp (R^2=0.232 and p value=0.274) (Table 3. 12).

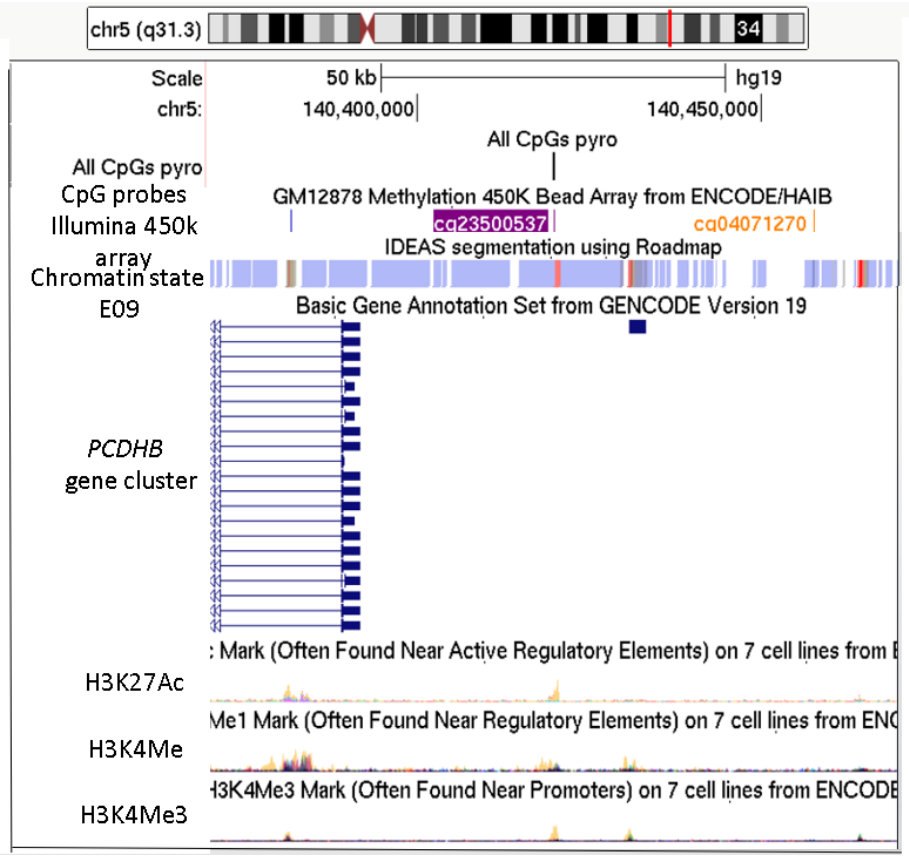


Figure 3. 23. UCSC genome browser view of the chromosomal location of the age-CpGs located in the putative enhancer of PCDHB1 cg23500537. The track All CpGs pyro provides the relative location of the CpGs included in the pyrosequencing assay; the chromatin state information is presented for the E049 cell line. Information about the nearby genes/isoforms is presented bellow this followed by information regarding the histone mark intensity present at the region in 7 cell lines provided by the ENCODE project.

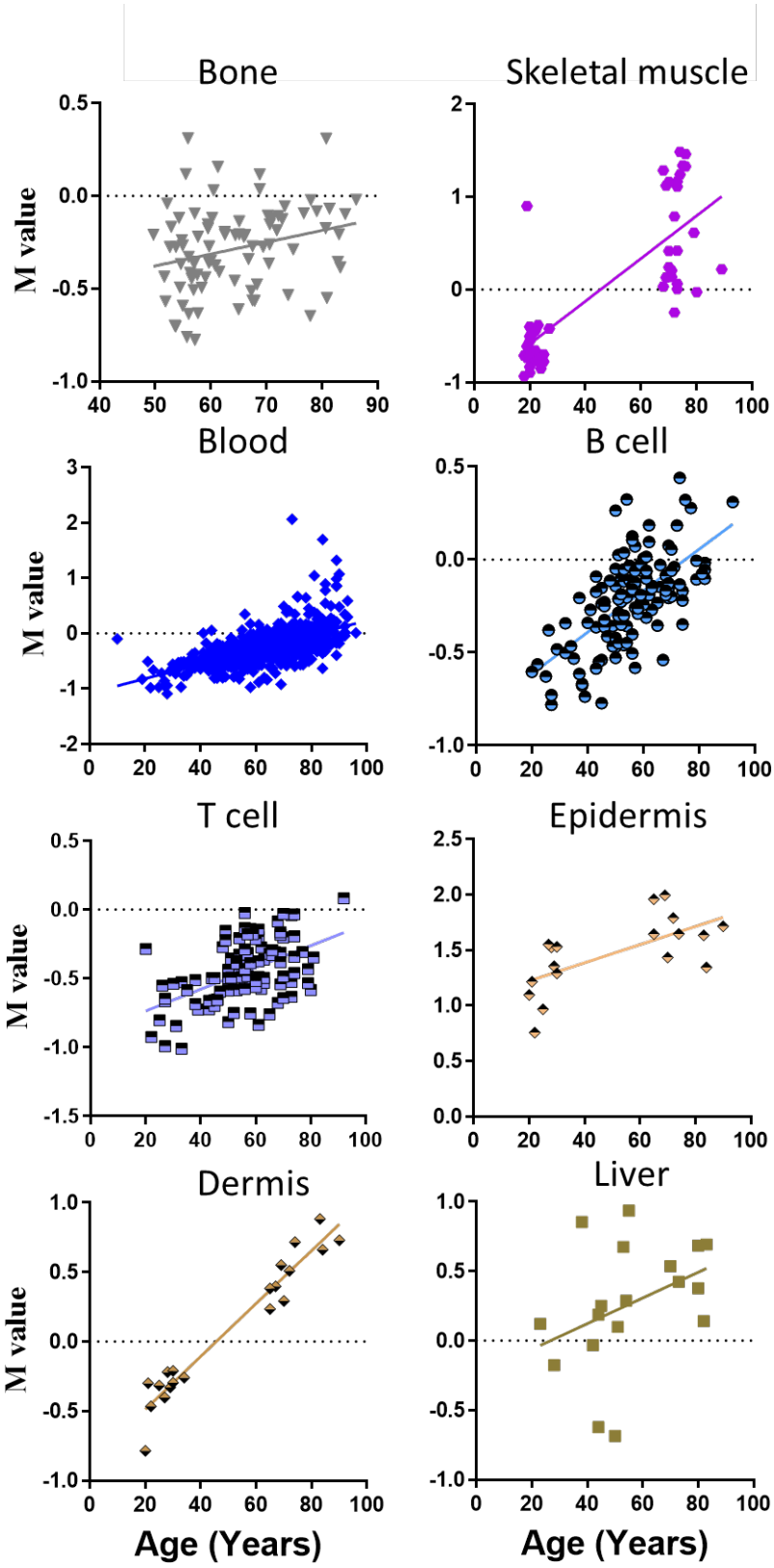


Figure 3. 24. Publicly available Illumina datasets information with linear regression analysis between age and DNA methylation for the musculoskeletal tissues of bone and skeletal

muscle and non-musculoskeletal tissue of blood, B cell, T cell, dermis, epidermis and liver for the cg23500537 CpG.

Table 3. 11. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the PCDHB1 region. N/A stands for "Not available" due data missing for the particular probe for the respective dataset.

		MSK tissues				non-MSK tissues						
PCDHB1 region		**Cartilage**	**Bone**	**Skeletal muscle**	**Tendon**	**Blood**	**B cell**	**T cell**	**Liver**	**Dermis**	**Epidermis**	**Breast**
Age-CpG	**Reference**	Reynard	PMID: 29514638	PMID: 24304487	Riasat	GSE40279	GSE137634	GSE137634	GSE48325	GSE52980	GSE52980	GSE88883
cg23500537	**Slope**	0.003	0.002	0.004	N/A	0.002	0.002	0.001	0.002	0.003	0.001	0.002
	P value	< 0.0001	< 0.0001	< 0.0001	N/A	<0.0001	< 0.0001	< 0.0001	0.148	< 0.0001	0.003	0.040
	R square	0.406	0.257	0.623	N/A	0.375	0.411	0.259	0.115	0.941	0.425	0.101

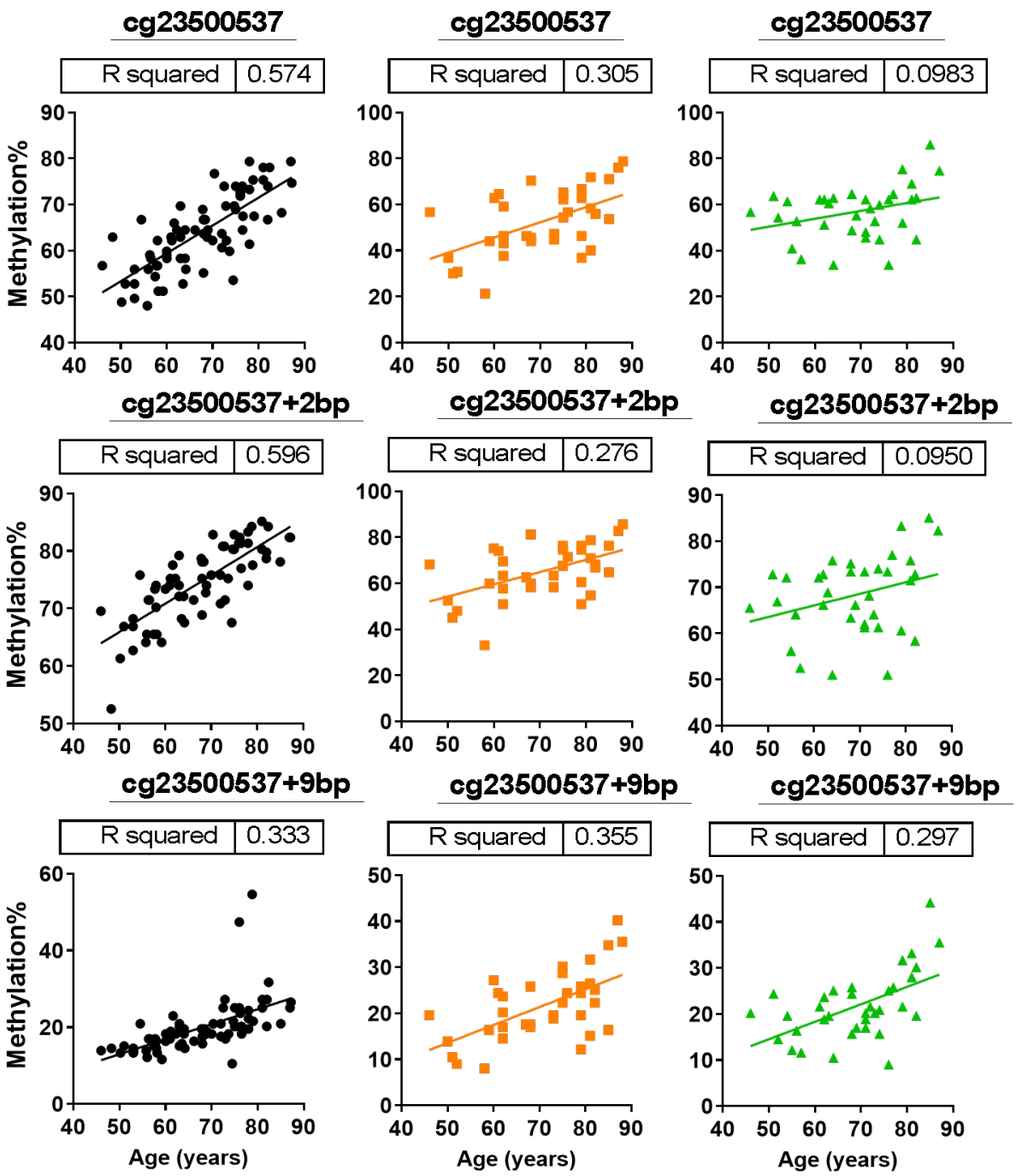


Figure 3. 25. Linear regression analysis of age versus DNA methylation at the CpG sites located in the putative enhancer region located near PCDHB1 for the tissues of cartilage, synovium and fat pad.

Table 3. 12.Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the PCDHB1 enhancer region.

		Cartilage			Synovium			Fat Pad			Blood		
Region	CpG	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value
PCHDB1	cg23500537	0.6226	0.5362	< 0.0001	0.6623	0.3051	0.0006	0.3437	0.1447	0.0265	0.6234	0.3701	0.1472
	cg23500537+2bp CpG	0.5322	0.5869	< 0.0001	0.5325	0.2761	0.0012	0.2515	0.141	0.0286	0.4979	0.3244	0.1819
	cg23500537+9bp CpG	0.418	0.3112	< 0.0001	0.3885	0.3551	0.0002	0.3791	0.3762	0.0001	0.3115	0.2319	0.2738

3.3.7 ZNF577 promoter region

In the promoter region of *ZNF577* there are two age-CpGs present, cg09547119 and cg22331349. The region can show an enrichment for H3K27Ac and H3K4Me3 histone marks (Figure 3. 26). The *ZNF577* (zinc finger 577) gene's function is unknown, however other members of the zinc finger family of proteins are involved in adipogenesis and muscle function control (Lorenzo *et al.*, 2020).

For the Illumina 450k datasets in the musculoskeletal tissues, the CpG cg09547119 (Figure 3. 27) had a significant change with age for bone (R^2=0.096 and p value=0.004) and for skeletal muscle R^2=0.446 and p value< 0.0001 however not for the tendon tissue (R^2=N/A and p value=0.145) (Table 3. 13). The other CpG cg22331349 (Figure 3. 27), in musculoskeletal tissue for bone R^2=0.068 and p value=0.017, for skeletal muscle R^2=0.446 and p value< 0.0001, for tendon R^2=N/A and p value=0.379 (Table 3. 13). In the non-musculoskeletal tissues, for cg09547119 (Figure 3. 28) the most significant change was observed for the tissue of dermis (R^2=0.8108 and p value< 0.0001) and in contrast the tissues of epidermis (R^2=0.093 and p value=0.218) and breast (R^2=0.0498 and p value=0.155) did not have any significant change (Table 3. 13). For cg22331349 (Figure 3. 28), the highest correlation was found for dermis (R^2=0.811 and p value< 0.0001) with the epidermis tissue showing a non-significant change with age (R^2=0.1984 and p value=0.064) (Table 3. 13).

The pyrosequencing assay also had an additional CpG not present as a probe this being named cg09547119+2bp. For the cartilage tissue (Figure 3. 29), only cg22331349 (R^2=0.083 and p value=0.04) was significantly observed to change with age with the CpG of cg09547119 (R^2=0.07547 and p value=0.0511) not passing the significance threshold (Table 3. 14). In contrast in the synovium tissue (Figure 3. 29) only cg09547119 (R^2=0.126 and p value=0.043) passed the significance threshold and not cg22331349 (R^2=0.1134 and p value=0.0554) (Table 3. 14). In the fat pad tissue (Figure 3. 29) the only CpG that was found to significantly changed with age was cg09547119 (R^2=0.159 and p value=0.048) (Table 3. 14). Finally, in the blood tissue no CpG contained in the *ZNF577* promoter region was observed to pass the significance threshold (Table 3. 14).

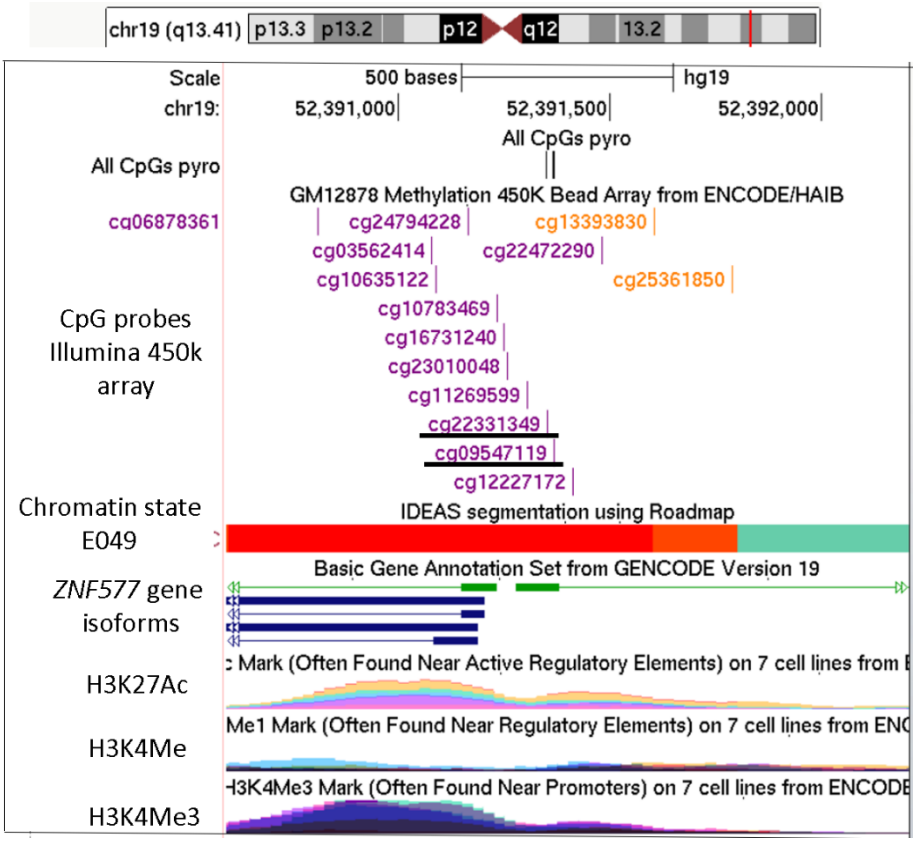


Figure 3. 26. UCSC genome browser view of the chromosomal location of the age-CpGs located in the promoter of ZNF577 cg22331349 and cg09547119. The track All CpGs pyro provides the relative location of the CpGs included in the pyrosequencing assay; the chromatin state information is presented for the E049 cell line. Information about the nearby genes/isoforms is presented bellow this followed by information regarding the histone mark intensity present at the region in 7 cell lines provided by the ENCODE project

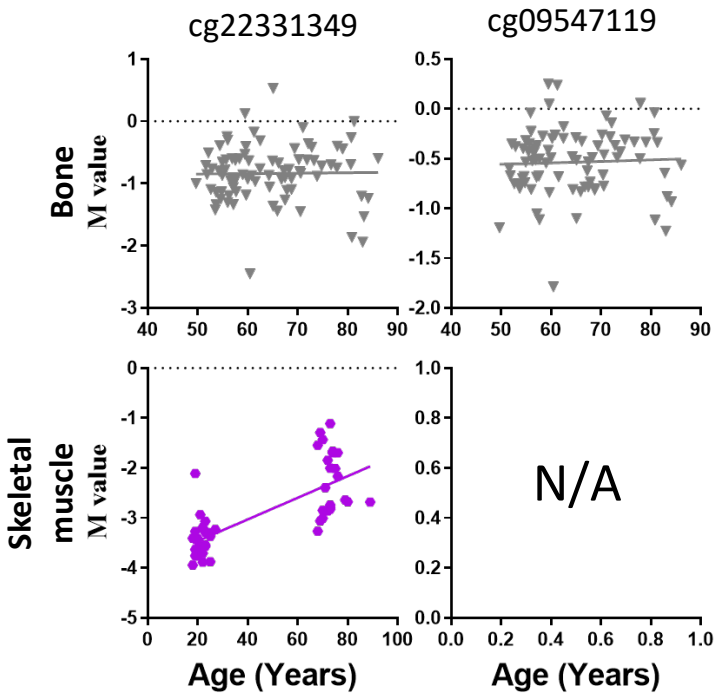


Figure 3. 27. Publicly available Illumina 450k datasets information for the musculoskeletal tissues of bone and skeletal muscle and linear regression between age and DNA methylation analysis for cg122331349 and cg09547119 CpGs.

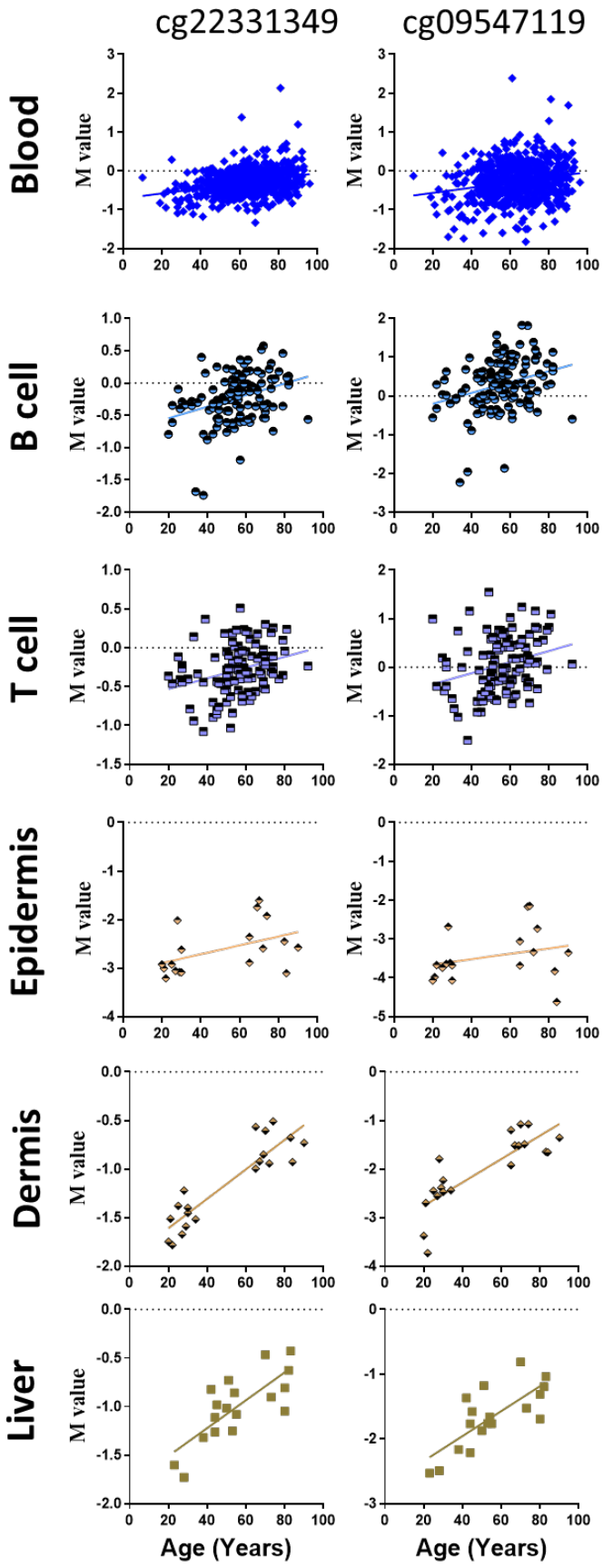

cg22331349
cg09547119
Blood
B cell
T cell
Epidermis
Dermis
Liver
M value
Age (Years)

Figure 3. 28. Publicly available Illumina datasets information for the non-musculoskeletal tissues of blood, b cell, t cells, epidermis, dermis and liver together with linear regression between age and DNA methylation analysis for cg122331349 and cg09547119 CpGs.

Table 3. 13. Table summary of the linear regression analysis of the available methylation datasets with linear regression analysis for age vs. DNA methylation for musculoskeletal and non-musculoskeletal tissues for the ZNF577 region. N/A stands for "Not available" due data missing for the particular probe for the respective dataset.

		MSK tissues				non-MSK tissues						
ZNF577 region		**Cartilage**	**Bone**	**Skeletal muscle**	**Tendon**	**Blood**	**B cell**	**T cell**	**Liver**	**Dermis**	**Epidermis**	**Breast**
Age-CpG	**Reference**	Reynard	PMID: 29514638	PMID: 24304487	Riasat	GSE40279	GSE137634	GSE137634	GSE48325	GSE52980	GSE52980	GSE88883
cg09547119	**Slope**	0.002	0.002	0.002	N/A	0.001	0.002	0.002	0.002	0.003	0.000	0.001
	P value	< 0.0001	0.004	< 0.0001	0.146	<0.0001	0.001	0.003	0.001	< 0.0001	0.218	0.155
	R square	0.212	0.096	0.446	N/A	0.044	0.089	0.079	0.541	0.718	0.093	0.050
cg22331349	**Slope**	0.002	0.001	0.002	N/A	0.001	0.001	0.001	0.002	0.002	0.001	0.002
	P value	< 0.0001	0.017	< 0.0001	0.379	<0.0001	< 0.0001	0.001	< 0.0001	< 0.0001	0.064	0.002
	R square	0.325	0.068	0.446	N/A	0.094	0.116	0.093	0.583	0.811	0.198	0.219

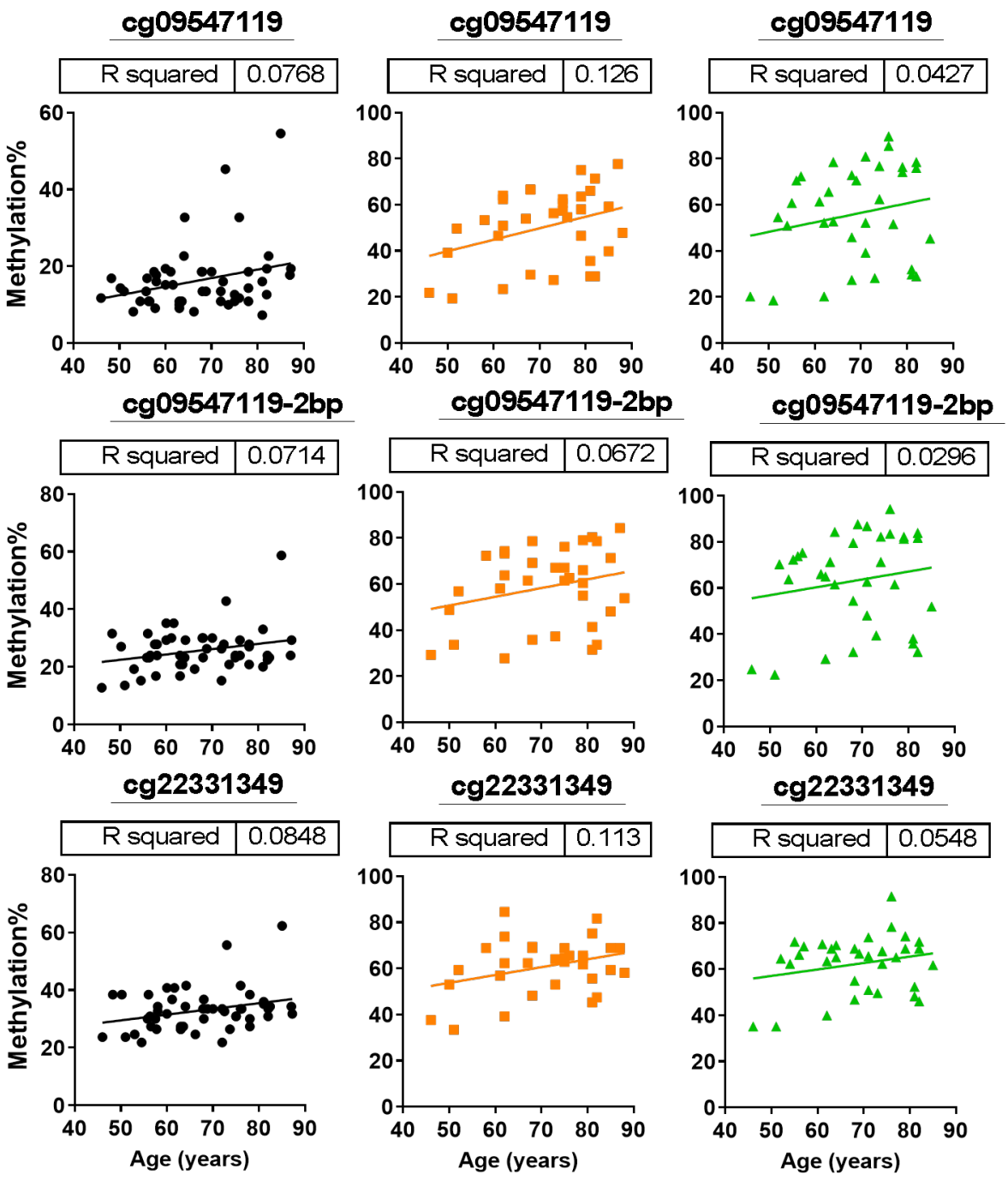


Figure 3. 29. Linear regression analysis of age versus DNA methylation at the CpG sites located in the promoter of ZNF577 for the tissues of cartilage, synovium and fat pad.

Table 3. 14. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the ZNF577 region.

		Cartilage			Synovium			Fat Pad			Blood		
Region	CpG	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value
ZNF577	cg09547119	0.2209	0.07547	0.0511	0.4992	0.1259	0.0427	0.5898	0.1594	0.048	0.3027	0.3139	0.1908
	cg09547119+2bp	0.1851	0.0709	0.0589	0.3788	0.06723	0.1451	0.4377	0.1221	0.0869	0.07947	0.02538	0.7329
	cg22331349	0.1972	0.08304	0.0403	0.3411	0.1134	0.0554	0.2039	0.06001	0.2379	0.03625	0.02568	0.7314

3.3.8 ZNF551 promoter region

The promoter region of *ZNF551* contains the age-CpGs of cg1644850. The region has some enrichment for H3K27Ac and H3K4Me3 histone marks at the site of the CpG sites and additionally the region could be a shared promoter site for the gene of *ZNF776* (Figure 3. 30). The *ZNF551* gene similarly to *ZNF577* gene (section 3.3.7.) is a member of the zinc finger family of proteins with the function of this gene being currently unknown.

For the analysis of the publicly available datasets the CpG of cg24214260 was also included in the analysis due to its close proximity to cg1644850 at 10 bp upstream. For these CpGs the following relationships were found in the datasets. In the musculoskeletal tissues (Figure 3. 31), the CpG cg24214260 had no significant change with age for bone (R^2=0.0281 and p value=0.127), with a non-significant change for tendon R^2=N/A and p value=0.301 and no data availability for skeletal muscle (Table 3. 15). For the age-CpG cg01644850 (Figure 3. 31) there was no significant change observed with age for the musculoskeletal tissues of bone (R^2=0.031 and p value=0.111) and tendon (R^2=N/A and p value=0.262) and again no available data for this CpG site for the skeletal muscle tissue (Table 3. 15). In the non-musculoskeletal tissues (Figure 3. 32), for cg24214260 the most significant change was observed for the tissue of dermis (R^2=0.68 and p value<0.0001), whereas the liver (R^2=0.001 and p value=0.89) and epidermis (R^2=0.0003 and p value=0.95) were both non-significantly changed with ageing (Table 3. 15). For cg1644850, the most significant change was observed for the dermis (R^2=0.8064 and p value< 0.0001) as well, with a non-significant change for the tissues of B cell (R^2=0.006 and p value=0.37) and epidermis (R^2=0.058 and p value=0.33) (Table 3. 15).

For the pyrosequencing assay this also contained information for cg24214260 alongside cg1644850. The results for cartilage tissue (Figure 3. 33) indicate that both CpGs investigated have a significant change with age with cg24214260 (R^2=0.50 and p value< 0.0001) and cg1644850 (R^2=0.5841 and p value< 0.0001) (Table 3. 16). For the synovium

tissue (Figure 3. 33), cg24214260 (R^2=0.105 and p value=0.142) had no significant change with age and in contrast cg1644850 (R^2=0.213 and p value=0.031) was found to be significant (Table 3. 16). A similar observation for the fat pad tissue (Figure 3. 33) where cg24214260 (R^2=0.1707 and p value=0.056) had no significant change with age whereas cg1644850 did have a significant change with R^2=0.375 and p value=0.0025 (Table 3. 16). In contrast to the findings from the synovium and fat pad tissue, in the blood tissue cg24214260 (R^2=0.8253 and p value=0.0046) was observed to significantly change with age and not cg1644850 R^2=0.07124 and p value=0.5629 (Table 3. 16).

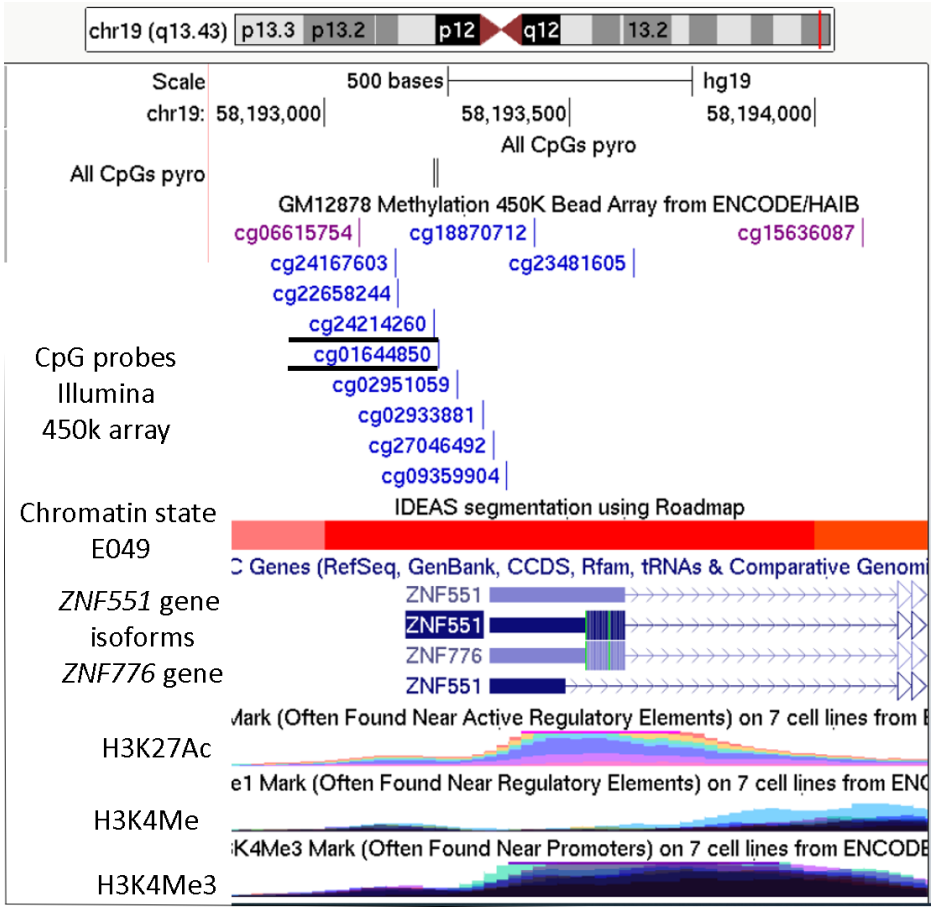


Figure 3. 30. UCSC genome browser view of the chromosomal location of the age-CpGs located in the promoter of ZNF551 cg24214260 and cg01644850. The track All CpGs pyro provides the relative location of the CpGs included in the pyrosequencing assay; the chromatin state information is presented for the E049 cell line. Information about the nearby genes/isoforms is presented bellow this followed by information regarding the histone mark intensity present at the region in 7 cell lines provided by the ENCODE project

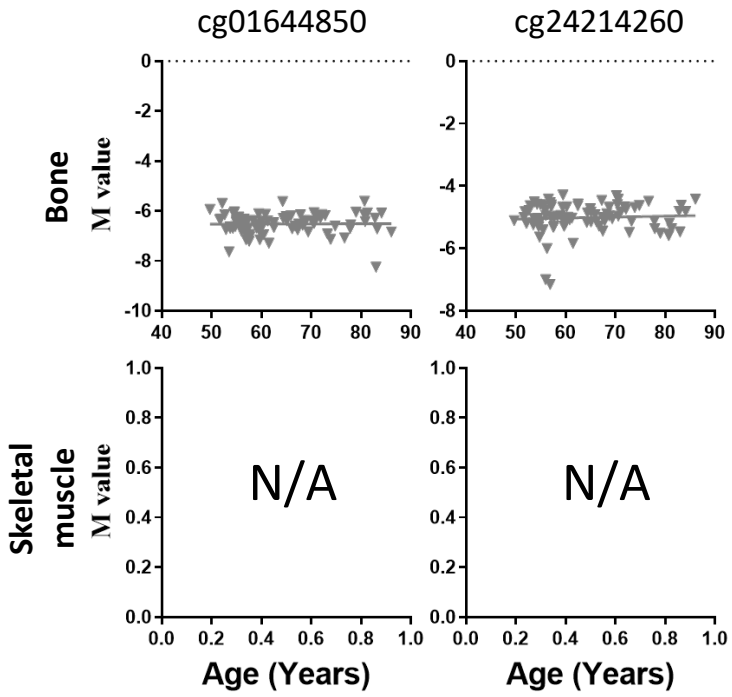


Figure 3. 31. Publicly available Illumina 450k datasets information for the musculoskeletal tissues of bone and skeletal muscle and linear regression between age and DNA methylation analysis for cg01644850 and cg24214260 CpGs.

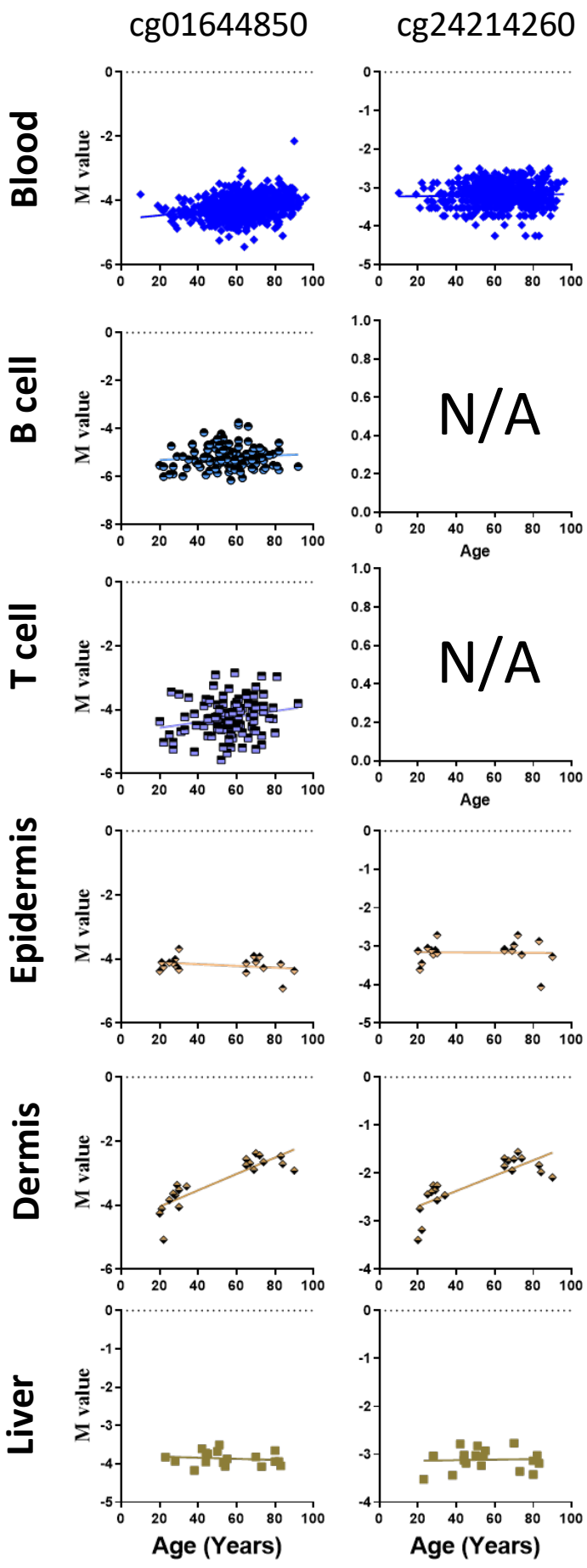

cg01644850
cg24214260
Blood
B cell
T cell
Epidermis
Dermis
Liver
M value
N/A
N/A
Age
Age (Years)
Age (Years)

Figure 3. 32. Publicly available Illumina datasets information for the non-musculoskeletal tissues of blood, B cell, T cells, epidermis, dermis and liver together with linear regression between age and DNA methylation analysis for cg01644850 and cg24214260 CpGs.

Table 3. 15. Table summary of the linear regression analysis of the available methylation datasets with linear regression analysis for age vs. DNA methylation for musculoskeletal and non-musculoskeletal tissues for the ZNF551 region. N/A stands for “Not available” due data missing for the particular probe for the respective dataset.

		MSK tissues				non-MSK tissues						
ZNF551 region		**Cartilage**	**Bone**	**Skeletal muscle**	**Tendon**	**Blood**	**B cell**	**T cell**	**Liver**	**Dermis**	**Epidermis**	**Breast**
Age-CpG	**Reference**	Reynard	PMID: 29514638	PMID: 24304487	Riasat	GSE40279	GSE137634	GSE137634	GSE48325	GSE52980	GSE52980	GSE88883
cg24214260	**Slope**	0.001226	0.000144	N/A	N/A	0.000212	N/A	N/A	2.65E-05	0.001544	0.00001203	0.001088
	P value	0.078	0.127	N/A	0.302	0.032	N/A	N/A	0.891	<0.0001	0.946	0.001
	R square	0.043	0.028	N/A	N/A	0.028	N/A	N/A	0.001	0.680	0.000	0.245
cg01644850	**Slope**	0.002	0.000	N/A	N/A	0.000	0.000	0.000	0.000	0.001	0.000	0.001
	P value	0.001	0.111	N/A	0.262	<0.0001	0.374	0.036	0.022	< 0.0001	0.334	<0.0001
	R square	0.135	0.031	N/A	N/A	0.071	0.006	0.040	0.558	0.806	0.058	0.377

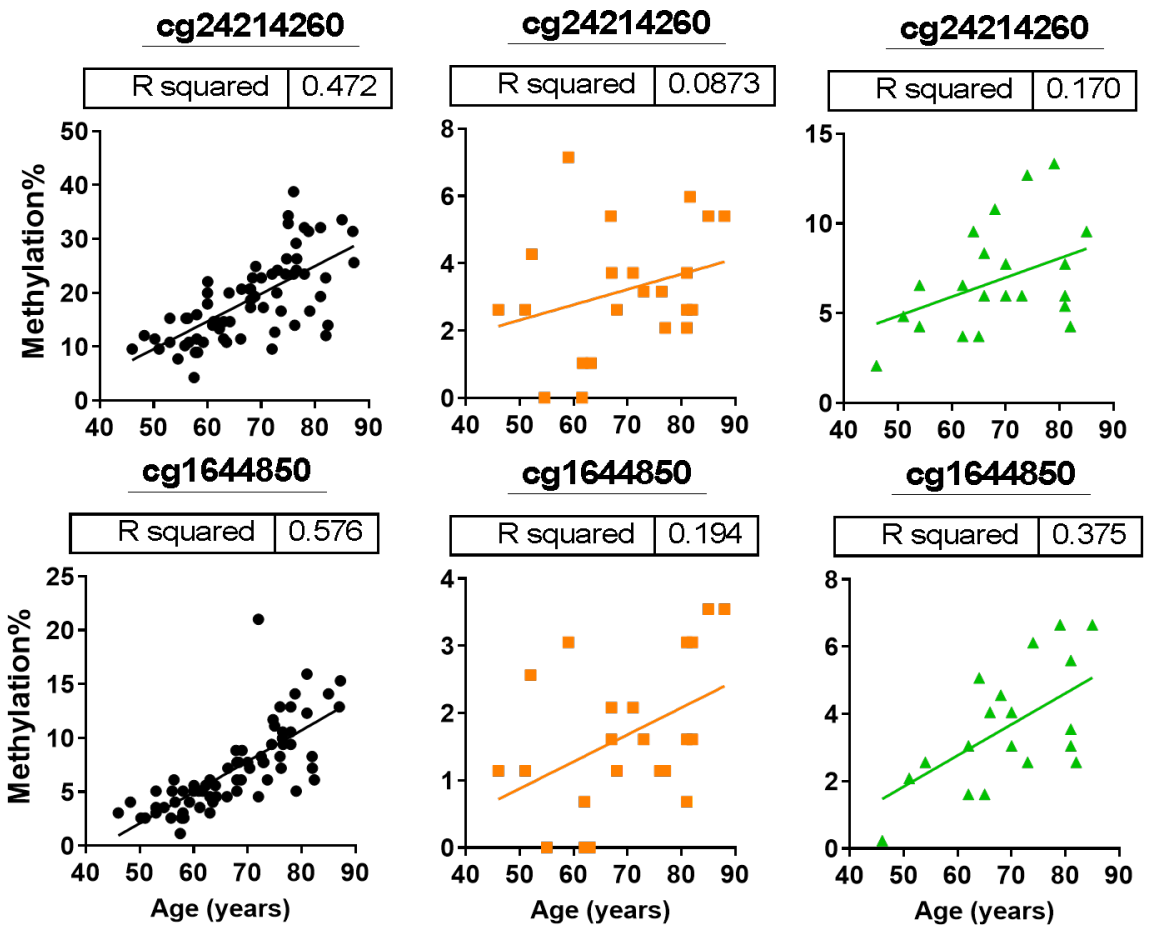


Figure 3. 33. Linear regression analysis of age versus DNA methylation at the CpG sites located in the promoter of ZNF551 for the tissues of cartilage, synovium and fat pad.

Table 3. 16. Summary table for the linear regression analysis between age and DNA methylation for the bisulphite pyrosequencing analysis in the articular joint tissues of cartilage, synovium and fat pad as well as for the blood tissue for the ZNF551 region.

		Cartilage			Synovium			Fat Pad			Blood		
Region	CpG	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value	Slope	r2	p value
ZNF551	cg24214260	0.5128	0.5009	< 0.0001	0.05036	0.1045	0.1422	0.1074	0.1707	0.056	0.02825	0.8253	0.0046
	cg1644850	0.2766	0.5841	< 0.0001	0.04258	0.2128	0.0307	0.09186	0.3745	0.0025	0.0044	0.07124	0.5629

3.3.9 Non validated regions

For the CpGs of cg07634191, cg13144059, cg06926735, cg18569335, cg07477282 there was no significant increase with age observed in DNA methylation levels for the new cartilage cohort by bisulphite pyrosequencing (Figure 3. 34, Table 3. 17). These CpG sites also showed weak age-CpGs properties when investigated on the EPIC array cohort and the additional Illumina 450k cohort (Figure 3. 34) with the closest CpG to the significance threshold being cg13144059 where it passed the significance threshold for the EPIC dataset (R^2=0.5671 and p value=<0.0001) and for new Illumina 450k dataset (R^2=0.1287 and p value=0.0143), however failing to show any significant change for bisulphite pyrosequencing (R^2=0.022 and p value=0.5096) (Table 3. 17). For the CpG site of cg06926735 there was no significant change with age detected for the EPIC dataset (R^2=0.1599 and p value=0.0587) or for the new Illumina 450k dataset (R^2=0.0653 and p value=0.0865), however significant changes with age are detected using bisulphite pyrosequencing (R^2=0.0287 and p value=0.6187) (Table 3. 17).

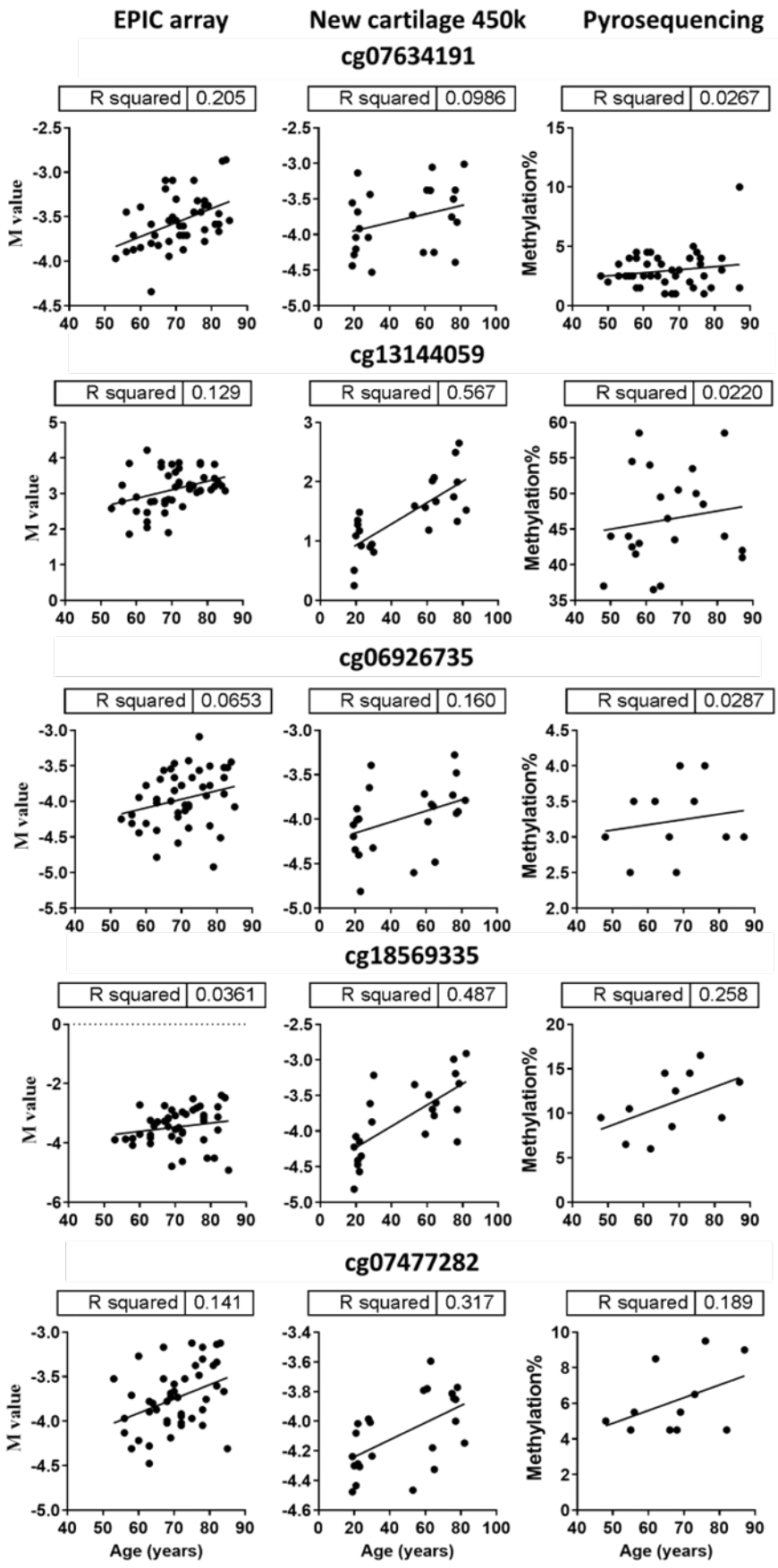

EPIC array
New cartilage 450k
Pyrosequencing
cg07634191
R squared 0.205
R squared 0.0986
R squared 0.0267
cg13144059
R squared 0.129
R squared 0.567
R squared 0.0220
cg06926735
R squared 0.0653
R squared 0.160
R squared 0.0287
cg18569335
R squared 0.0361
R squared 0.487
R squared 0.258
cg07477282
R squared 0.141
R squared 0.317
R squared 0.189
M value
Methylation%
Age (years)

Figure 3. 34. Linear regression analysis of age versus DNA methylation at the CpG sites that did not validate in the new cartilage cohort analysed by bisulphite pyrosequencing.

Table 3. 17. Summary table of the CpG sites that could not be validate in the new cohort using bisulphite pyrosequencing.

	CpG	cg07634191	cg13144059	cg06926735	cg18569335	cg07477282
Scripts	R^2	0.098	0.567	0.160	0.488	0.318
	p value	0.145	<0.0001	0.059	<0.0001	0.005
	r	0.314	0.753	0.400	0.698	0.564
EPIC	R^2	0.205	0.129	0.065	0.036	0.141
	p value	0.002	0.014	0.087	0.206	0.010
	r	0.453	0.359	0.256	0.190	0.375
Pyrosequencing	R^2	0.000	0.022	0.540	0.523	0.189
	p value	0.947	0.510	0.010	0.059	0.181
	r	0.010	0.148	0.735	0.723	0.435

3.3.10 Can pyrosequenced quantified methylation values be used to develop an ageing clock?

An epigenetic clock was developed using multiple-variate regression analysis data from bisulphite pyrosequencing data in order to model DNA methylation changes to chronological age. Two epigenetic clocks were developed with one starting with the data from all the age-CpGs followed by investigation of the highest contributors to the model by their R^2 coefficient and selection of the top 10 age-CpGs (Figure 3. 35 All); the second epigenetic clock was developed using the data from the age-CpGs contained in the promoter of *FHL2*, this being due to their high correlation and degree of change with chronological age (Figure 3. 35 FHL2). After selecting the best scoring CpGs, the epigenetic clock was trained on 2/3 of the available samples for cartilage (48 samples) synovium +fat pad (29) and all tissues (77) with 1/3 of the data being kept for testing the model generated. This generated clock for cartilage had a R^2=0.891 with an age error of 6 years (Figure 3. 35 A), for synovium + fat pad a R^2=0.922 with an age error of 12 years (Figure 3. 35 B) and a multi tissue articular joint epigenetic clock with a R^2=0.705 with an age error of 9.6 years (Figure 3. 35 C). The FHL2 model was generated using only one region that consistently showed a high correlation coefficient value that being the promoter of *FHL2* promoter age-CpGs (Figure 3. 35). For this version of the clock the cartilage had a R^2=0.605 with an age error of 9.5 years (Figure 3. 35A), for synovium + fat pad a R^2=0.855 with an age error of 11 years (Figure 3. 35B) and a multi tissue articular joint epigenetic clock with a R^2=0.583 with an age error of 16.5 years (Figure 3. 35C). This secondary clock has a lower overall R^2 values and bigger overall age error with the benefit of being only two assays that permit a quick determination of the samples age.

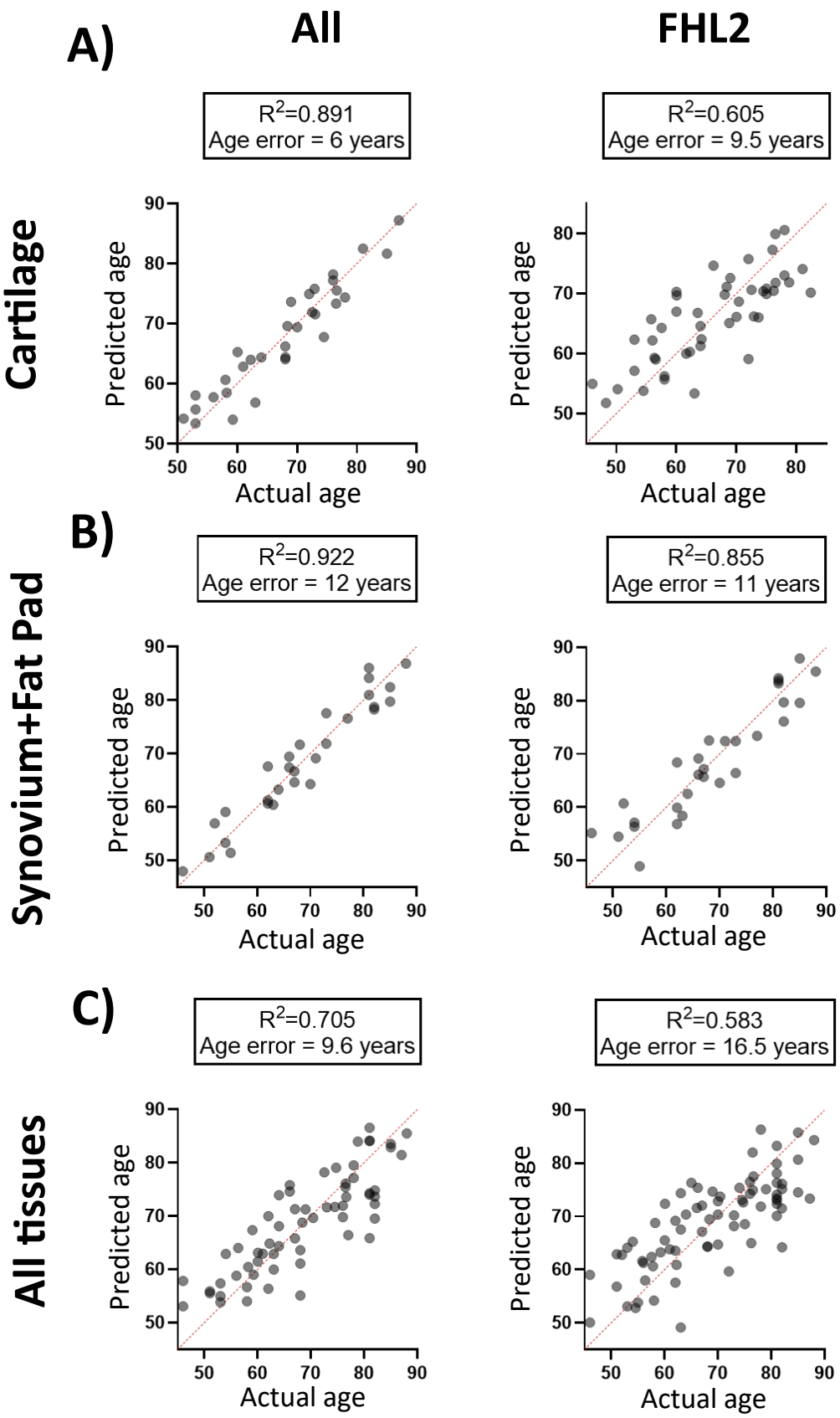

A)
All
FHL2
Cartilage
R2=0.891
Age error = 6 years
R2=0.605
Age error = 9.5 years
B)
Synovium+Fat Pad
R2=0.922
Age error = 12 years
R2=0.855
Age error = 11 years
C)
All tissues
R2=0.705
Age error = 9.6 years
R2=0.583
Age error = 16.5 years
Predicted age
Actual age
50
60
70
80
90

Figure 3. 35. Using multi-variate regression analysis an epigenetic clock was developed using bisulphite pyrosequencing data. Two epigenetic clocks were developed one using data from all the CpG assays and the second using the high scoring region of the CpGs from the FHL2 promoter. A) Cartilage specific epigenetic clock B) Synovium and Fat Pad epigenetic clock C) Epigenetic clock trained with samples from cartilage, synovium and fat pad samples.

3.4 Discussion

3.4.1 Validation of age-related DNA methylation changes using bisulphite pyrosequencing

Based on the Illumina 450k Methylation array results generated in our group using 179 knee cartilage samples from patients with end-stage OA, a total of 716 CpG sites were identified that show a correlation between DNA methylation levels and chronological age. Looking into the reproducibility of these findings using the Illumina EPIC array and a new cohort with the same Illumina 450k array, this revealed that only ~66% of the CpGs could be replicated using a different platform however this issue is also caused by the removal of ~ 11% probes from the EPIC array that were previously present on the Illumina 450k array. On the same platform with a new smaller cohort ~50% of the CpGs were replicated however a reason for the low number of CpGs that replicated the age-CpGs characteristics could be due to the lower number of samples that this dataset had. A recent publication from our group (Cheung *et al.*, 2020) has shown that the reliability of both the Illumina 450k array and of the Illumina EPIC array when compared to bisulphite sequencing, is highly dependent on the site in question, with false positives being reported at certain CpG sites. The investigation of other publicly available datasets was done to understand firstly whether the 716 cartilage age-CpG were specific to cartilage and to what extent these could be identified as age-CpGs in other tissues. The analysis revealed that several hundred CpGs seem to be found as age-CpGs in musculoskeletal tissues such as bone followed by skeletal muscle or other non-musculoskeletal tissues (i.e. dermis, blood, liver, epidermis and breast). Such results suggest that the age-associated DNA methylation changes are not rarely observed, and most are not unique to the cartilage tissue. Tissues examined include whole blood, liver, breast, etc, with the CpG sites of cg22454769, cg24079702 being identified in all tissues and in contrast the CpG sites of cg22448433, cg22643867 were observed to be only identified in the cartilage dataset this possibly making it unique for this tissue.

The results of the Illumina 450k array do not always corelate with the findings from another tool such as bisulphite pyrosequencing (Roessler *et al.*, 2012), with bisulphite pyrosequencing being regarded as a gold standard in measuring DNA methylation due to its

more reliable sequencing technology (Bassil *et al.*, 2013; Hop *et al.*, 2020). Due to this reason, it was decided to validate some of the age-CpGs identified in the cartilage Illumina 450k dataset. The selection of the CpGs to validate was decided based on the factors such as CpG location within chromatin state regions, overall range of DNA methylation change observed between the youngest and oldest subjects, and the successful design of bisulphite pyrosequencing assays to study these. Based on the starting number of 716 age-CpGs these have been reduced to a shorter list of 15 CpG site, selected for further validation in a new cohort of 72 OA knee cartilage samples. A calibration curve was generated in order to test the validity of the pyrosequencing results as well as to adjust using a cubic polynomial regression equation adjustment for any amplification bias that can result from PCR (Moskalev *et al.*, 2011). From this work 9/15 CpGs validated as being age-CpGs in the new cohort of cartilage samples. A possible reason or A difference below 5% is regarded as undetectable by pyrosequencing (Jing-bin *et al.*, 2014). Additionally, the bisulphite pyrosequencing assays were also able to provide additional data on neighbouring CpGs to the main age-CpG in analysis and through this additional CpG sites were identified to show age-related DNA methylation changes in cartilage. From the initial nine validated CpG sites an additional of 11 new CpGs sites were found to exhibit the properties of age-CpGs this giving a total of 20 age-CpGs for the cartilage tissue out of 41 analysed CpGs (Table 3. 18). An interesting observation is that most of the CpGs that did not validate by pyrosequencing in the new cohort were located in putative enhancer regions (e.g. cg18569335, cg07634191). A possible reason for this could be that the DNA methylation levels at these enhancers do not have a big range in variability with age, this limiting the detection by pyrosequencing. On the other hand, the variability is far greater with age at CpG sites located in promoter regions and therefore more age-CpGs located in promoter regions were validated in the new cohort. This possibly indicates the different epigenetic control that takes places at enhancers and promoters giving rise to different levels of DNA methylation being imprinted. Since for the extra CpG sites there are no corresponding Illumina 450k platform/EPIC probes this limits their measurement of DNA methylation to sequencing based technology like bisulphite pyrosequencing. Their measurement allows the investigation of a broader region and therefore can better give insight into the degree of age-associated changes across the region in question.

Table 3. 18. Summary table of CpGs that validated by bisulphite pyrosequencing and comparison to the Illumina 450k data and R2 value and p value for linear regression. A) Represents age-CpGs that are predicted to be located in a region with a promoter signature and B) Age-CpGs predicted to be located in a putative enhancer region. Cells under the p value row are marked with green when these are significant to a threshold of <0.05.

A)

Nearby gene	*FHL2*			*ZNF551*	*ELOVL2*			*ZNF577*		*TMEM189*
CpG	cg24079702	cg22454769	cg06639320	cg01644850	cg16867657	cg24724428	cg21572722	cg22331349	cg09547119	cg06926735
Chromatin state	Promoter			Promoter	Promoter			Promoter		Promoter
450k										
R square	0.14	0.15	0.24	0.16	0.13	0.01	0.16	0.33	0.18	0.08
P value	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.31	< 0.0001	< 0.0001	< 0.0001	0.04
Bis. Pyro.										
R square	0.26	0.19	0.26	0.58	0.28	0.23	0.07	0.08	0.08	0.54
P value	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.06	0.04	0.05	0.01

B)

Nearby gene	*NKIRAS2*	*SCARA5*	*SPATS2L*	*SPG11*	*PCDHB1*	*PCDHGA1*
CpG	cg18569335	cg07634191	cg13144059	cg07477282	cg23500537	cg14566959
Chromatin state	Enhancer	Enhancer	Enhancer	Enhancer	Enhancer	Enhancer
450k						
R square	0.19	0.12	0.07	0.15	0.41	0.23
P value	0.02	< 0.0001	0.02	< 0.0001	< 0.0001	< 0.0001
Bis. Pyro.						
R square	0.52	0	0.02	0.19	0.54	0.2
P value	0.06	0.95	0.51	0.18	< 0.0001	< 0.0001

3.4.2 DNA methylation changes at the promoter of ELOVL2

For the promoter region of *ELOVL2* there are a total of three probe CpGs cg16867657, cg24724428, cg21572722 that have been observed to be age-CpGs in tissues such as saliva, blood, kidney, brain. The CpGs of cg16867657, cg21572722 are part of the 716 cartilage age-CpGs and moreover they have been identified in the musculoskeletal datasets of bone, skeletal muscle and the non-musculoskeletal tissues of liver, breast, epidermis, etc. The analysis using bisulphite pyrosequencing confirmed that cg16867657 had age-related methylation changes yet not cg21572722. The analysis in the cartilage tissue identified

cg24724428 as also being an age-CpGs even though this was not part of the 716 cartilage age-CpGs. The CpG of cg24724428 is a neighbouring CpG of cg16867657, cg21572722 in close proximity to these CpGs and the identification of this using pyrosequencing suggests that there is at least a case of a false negative being given by the Illumina 450k platform. However, an interesting observation is that at cg24724428 there was no association with age in the cartilage age-CpGs from the initial Illumina 450k dataset and only because of the proximity of this site to cg21572722 it was measured using pyrosequencing and was found to show characteristics of an age-CpG. For the tested CpG sites the results from the Illumina 450k array findings did not always corroborate with the findings by bisulphite. The pyrosequencing assay also analysed two other CpG neighbouring the mentioned age-CpGs, both of which show age-CpG characteristics in cartilage, synovium and fat pad. The cartilage tissues show consistently lower levels of DNA methylation across all five CpGs being followed by the synovium, fat pad and blood all of which show very similar levels and no significant discernment can be made between these three tissues. For the CpGs sites located in the *ELOVL2* promoter for the matrix correlation analysis and principal component analysis in articular tissues indicate that all neighbouring CpGs are highly associated (with exception of cg21572722 in cartilage).

The relationship of the CpGs in the region to cluster could reveal different functionalities for the region or could therefore give insight into their role with ageing and how these changes in DNA methylation can reflect the functionality of the regions. The age-associated hypermethylation observed in this study for the promoter of *ELOVL2* age-CpGs in articular tissues has been reported in the literature for various tissues has also been proposed as a marker of cellular division (Bacalini *et al.*, 2017) , where DNA methylation at these sites was progressively detected to increases with serial passaging in fibroblast cells. These age-CpGs have also been proposed as ubiquitous age-related DNA methylation marker (Slieker *et al.*, 2018) however the link with gene expression and other associations are not currently known.

3.4.3 DNA methylation changes at the promoter of FHL2

The promoter region of *FHL2* contains a total of three probe age-CpGs cg06639320, cg22454769, cg24079702 that seem to have age-associated DNA methylation changes in

multiple tissues from the musculoskeletal tissues of cartilage, synovium, fat pad, bone and skeletal muscle and non-musculoskeletal tissues such as blood, liver, breast, dermis and epidermis. With the pyrosequencing analysis a further six proximal CpGs were identified where these are also showing age-associated DNA methylation changes in cartilage, synovium and fat pad tissue. Intra-regional analysis indicates the formation of two clusters with one cluster being formed by cg06639320, cg06639320+6bp and cg06639320+9bp and the second cluster being formed by cg24079702-14bp, cg22454769, cg24079702, cg24079702+2bp but not cg24079702+13bp with this relationship observed in all investigated tissues. Such a relationship between the CpGs could represent their functionality on how the region interacts with the transcription machinery. The age-CpGs present at this promoter (i.e. cg06639320, cg22454769, cg24079702) have been also identified in longitudinal studies to show age-associated DNA methylation changes (Florath *et al.*, 2014; Bacalini *et al.*, 2017) which compliments the data presented here, as this is part of a cross-sectional analysis. DNA methylation levels at cg06639320 has been proposed as part of a forensic age determination tool for blood samples (Pan *et al.*, 2020; Han *et al.*, 2022) to be used for age identification of unknown samples to its robust changes with age. Additionally, DNA methylation levels at cg06639320 in whole blood were highly associated with high blood pressure (Kazmi *et al.*, 2020), however it is unknown through what mechanism these could be associated. This makes the age-CpGs present at the promoter of *FHL2* as interesting targets for further analysis to determine the impact these DNA methylation changes have on cellular functionality as well as understanding the mechanisms through which this occurs.

3.4.4 DNA methylation changes at the enhancer of PCDHGA1

The CpG of cg14566959 is located in a putative enhancer region in the intron of the *PCDHGA1* gene. This has been observed to be an age-CpG in the tissue of cartilage, synovium skeletal muscle, blood, liver, dermis and epidermis however not in bone, fat pad and breast tissue. The pyrosequencing analysis also determined that the two neighbouring CpG sites of cg14566959-2bp and cg14566959-13bp did not have any age-associated DNA methylation changes in either tissue investigated. Due to the proximity to the cluster of

PCDHGA gene family it could therefore be that one or more of these could be the target/targets of the enhancer (Qin *et al.*, 2022), however it is well known in the field that enhancers do not necessarily target the nearest gene and could modulate genes located at much farther distances (Hariprakash and Ferrari, 2019). With this region being classed as a putative enhancer the current target for this region remains unknown. Future experiments will focus in determining whether this region has enhancer properties as shown by methods such as the luciferase assay (Yun and DasGupta, 2014; Wong *et al.*, 2021). As for the determination of the target gene or genes for this putative enhancer, methods such as through dCas-VP64 enhancer activation (Xu *et al.*, 2019) or CRISPR-Cas9 deletion of the enhancer region (Moorthy and Mitchell, 2016) followed by gene expression analysis of the surrounding genes could be used. Another possible method to determine the target for this enhancer region could be through Capture Hi-C (Tomás-Daza *et al.*, 2023), where the enhancer's interaction with the target promoter can be captured and sequenced.

3.4.5 DNA methylation changes at the enhancer of PCDHB1

The CpG of cg23500537 is also located in a potential enhancer region that is located nearby to the *PCDHB1* gene ~14kb away upstream. The CpG has been identified to be an age-CpGs in the musculoskeletal tissues of cartilage, synovium, fat pad, bone and skeletal muscle and in the non-musculoskeletal tissues of blood, B cell, T cell, dermis, epidermis and breast. Additionally, the two proximal CpG sites to cg23500537 (cg23500537+2bp and cg23500537+9bp) that also showed age associated changes in DNA methylation levels for cartilage, synovim and fat pad indicating the spread of the age-associated changes to these CpGs as well, however no significant correlation was detected in the blood samples. As cg23500537 was previously detected as an age-CpG in the whole blood 450k datasets (Hannum *et al.*, 2013) it could therefore be that the inability to significantly detect this in the current dataset by pyrosequencing could be due to the lower sample size used and future work for this would focus to increasing this. Again, with this region being another possible enhancer, future work will focus in testing this property followed by a possible investigation into the target of this region in order to understand the possible role this could have in the tissues and the consequences of the age-related DNA methylation changes.

Currently there are no reports of linking DNA methylation levels cg23500537 with any functional changes and this will be investigated in the next chapter.

3.4.6 DNA methylation changes at the promoter of ZNF577

The promoter region of *ZNF577* contained two CpGs as part of the 716 discovered age-CpGs, cg09547119 and cg22331349. Both CpGs are age-CpGs in bone and skeletal muscle as well as blood, B cell, T cell, liver and dermis. Interestingly none are significant age-CpGs in the epidermis tissues and only cg22331349 is significant in the breast tissue. For the cartilage tissues in the new cohort through pyrosequencing, only cg22331349 passed the significance threshold. Additionally, an extra CpG was analysed as part of the assay, cg09547119+2bp, with this CpG also showing a trend of change with age with a close to threshold significance. For the synovium and fat pad tissues interestingly only cg09547119 shows a significant relationship with age and the other CpGs are close to the threshold. None of the CpGs seems to be significantly changing with age in the blood tissues suggesting that possibly these CpGs could only slowly change with age in these tissues and therefore more samples are needed to detect this subtle change as this was observed in the bigger 450k datasets for blood. Currently not much is known about any other associations for the DNA methylation levels at the CpGs of cg09547119 and cg22331349 and primarily it is through the age-related changes discussed. Whether these changes impact the functionality of the *ZNF577* gene is currently unknown and this will the focus of the next investigations in Chapter 4. For the CpG of cg22331349 a report found that the methylation levels of the mother directly influence the methylation levels in their children with this also being correlated with an increased sensitisation to allergens in early life (Acevedo *et al.*, 2021) although the mechanism is unknown.

3.4.7 DNA methylation changes at the promoter of ZNF551

The CpG of cg1644850 is located in the promoter region of *ZNF511* and this has been observed as an age-CpGs in blood, t cell, liver, dermis and breast but not in the musculoskeletal tissues of bone and skeletal muscle. After the assessment of the CpG by pyrosequencing this was found to validate in the new samples and moreover another CpG that showed age-related changes was included which was found to also correspond to the

probe of cg24214260. In the initial 716 age-CpGs in cartilage cg24214260 was not identified as an age-CpG however in the other datasets this was determined to be an age-CpG in blood, dermis and breast and again not in any of the other musculoskeletal tissues. For the synovium and fat pad tissues only cg1644850 showed a significant change in DNA methylation with age with the CpG of cg24214260 only showing a non-significant trend of increasing with age. Interestingly in the blood tissue a reverse occurs where cg24214260 passed the significance threshold and cg1644850 did not. A possible reason could be that the two CpGs in this region could be under the regulation of different mechanisms in order to maintain their DNA methylation levels where this could differ by the type off tissue and therefore influence differently their functionality. No current knowledge links the DNA methylation changes observed with any functional changes at these sites and this will be investigated in the following chapter.

In the list of the age-CpGs selected by further validation through pyrosequencing there have been several CpGs that either did show weak change with age in the DNA methylation levels or did not change across the samples. These include cg07634191, cg13144059, cg06926735, cg18569335, cg07477282 where only cg06926735 was showing a relationship with age with very little variation between the youngest subject at 48 years to the older subject at 87 years with ~4% DNA methylation level difference. As previously mentioned, a difference below 5% is regarded as undetectable by pyrosequencing (Yan *et al.*, 2014) and therefore any methylation changes below this threshold are difficult to study by this technique. The rest in this list are not showing any significant changes with age to validate the findings from the initial Illumina 450k analysis. Due to this, these CpGs were not analysed further, and the focus was placed on those that validated in the initial study. A possible reason why the five CpG sites did not validate as age-CpGs could either be due to false positives being given by the platform (Zheng *et al.*, 2022) or the requirements of a bigger sample number to detect the relationship that they could have with age as the study could be under-powered for such detection.

Table 3. 19. Summary of the linear regression analysis for cartilage, synovium and fat pad tissues from bisulphite pyrosequencing data.

		Cartilage		Synovium		Fat Pad	
Region	**CpG**	**R^2**	**p value**	**R^2**	**p value**	**R^2**	**p value**
Promoter of *ELOVL2*	cg21572722	0.0686	0.0607	0.6127	< 0.0001	0.5304	0.0001
	cg24724428	0.3121	< 0.0001	0.6357	< 0.0001	0.5526	< 0.0001
	cg24724428-9bp	0.2285	0.0003	0.7030	< 0.0001	0.2854	0.0104
	cg16867657	0.2815	< 0.0001	0.6450	< 0.0001	0.4777	0.0004
	cg16867657-4bp	0.1233	0.0107	0.6638	< 0.0001	0.4990	0.0002
Promoter of *FHL2*	cg06639320	0.2634	< 0.0001	0.3321	0.0050	0.3579	0.0033
	cg06639320+6bp	0.0802	0.0223	0.2553	0.0165	0.2355	0.0221
	cg06639320+9bp	0.1145	0.0058	0.3769	0.0024	0.5737	< 0.0001
	cg24079702-14bp	0.1813	0.0004	0.4202	0.0011	0.5440	< 0.0001
	cg22454769	0.1913	0.0003	0.3665	0.0028	0.4912	0.0003
	cg24079702	0.2621	< 0.0001	0.6060	< 0.0001	0.5549	< 0.0001
	cg24079702+2bp	0.1142	0.0059	0.4608	0.0005	0.5699	< 0.0001
	cg24079702+9	0.1745	0.0005	0.3999	0.0016	0.5069	0.0002
	cg24079702+13bp	0.0826	0.0203	0.6198	< 0.0001	0.2569	0.0161
Enhancer *PCDHGA1*	cg14566959-13 bp	0.0269	0.1915	0.0980	0.1560	0.0187	0.5444
	cg14566959-2bp	0.0225	0.2327	0.1042	0.1428	0.0015	0.8630
	cg14566959	0.1962	0.0002	0.5562	< 0.0001	0.1359	0.0914
Enhancer *PCHDB1*	cg23500537	0.5362	< 0.0001	0.3051	0.0006	0.1447	0.0265
	cg23500537+2bp	0.5869	< 0.0001	0.2761	0.0012	0.1410	0.0286
	cg23500537+9bp	0.3112	< 0.0001	0.3551	0.0002	0.3762	0.0001
Promoter of *ZNF577*	cg09547119	0.0755	0.0511	0.1259	0.0427	0.1594	0.0480
	cg09547119+2bp	0.0709	0.0589	0.0672	0.1451	0.1221	0.0869
	cg22331349	0.0830	0.0403	0.1134	0.0554	0.0600	0.2379
Promoter of *ZNF551*	cg24214260	0.5009	< 0.0001	0.1045	0.1422	0.1707	0.0560
	cg1644850	0.5841	< 0.0001	0.2128	0.0307	0.3745	0.0025

3.4.8 Epigenetic clock development

With the introduction of the epigenetic DNA methylation clocks and with the increase of their usage as a tool in research following the development of the Horvath clock, many researchers have started to make their own epigenetic clocks in order to compensate for the disadvantages that multi tissue epigenetics clocks present (i.e. sex and ethnicity measurement error in the early Horvath clock and inconsistent accuracy of the clock for different tissues). The development of tissue specific clocks permits a more accurate analysis of the tissue in questions however these are still commonly done on data from Illumina 450k or Illumina EPIC array data and are difficult to introduce into a cheap and quick method of analysis and together with the poor translation to pyrosequencing of some of the CpG sites make it still difficult to use as a tool in research. Since the direct data from these arrays is used in the production of epigenetic clocks with no further validation by techniques such as bisulphite pyrosequencing, it is therefore difficult to ascertain whether the models are built on artefacts of the array or the actual physiological changes with ageing. An attempt here is made at developing a cartilage epigenetic clock and an articular joint clock is here attempted using multivariate regression analysis that is based on pyrosequencing data. After selecting the best scoring CpGs from the multivariate analysis by taking only the CpGs with the highest correlation coefficient in this analysis, the epigenetic clock was trained on 2/3 of the available samples for cartilage (48 samples) synovium +fat pad (29 samples) and all (77 samples) with 1/3 of the data being kept for testing the model generated to assess the age error given by each model. Another model was generated using only one region that consistently showed a high correlation coefficient value that being the promoter of *FHL2* promoter age-CpGs. This secondary clock has a lower overall R^2 values and bigger overall age error with the benefit of being only two assays that permit a quick determination of the samples age. Similar tools for study in other tissues were used to determine that conditions such as obesity cause an increased biological age as determined by an epigenetic clock (de Toro-Martín *et al.*, 2019) and treatments such as those with growth hormone have decrease the biological age of individuals (Bartke *et al.*, 2021). Such clocks for musculoskeletal tissues could be useful in the assessment of the biological age of

the tissue after certain treatments as well as to determine comorbidities that could influence the normal functionality.

Chapter 4: Do age-related DNA methylation changes correlated with changes in gene expression?

4.1 Introduction

The previous chapter confirmed that specific CpG sites undergo age-related DNA methylation changes in knee cartilage, with several of these acting as age-CpGs in additional musculoskeletal tissues. However, the effect of these age-related DNA methylation changes on gene expression and cellular phenotype is unknown. Associations have been observed between DNA methylation and gene expression with such a relationship previously believed to be mainly an inverse proportionally relationship, with increases in DNA methylation being associated with a decrease in target gene expression (Moore *et al.*, 2012). However, such a theory has been adjusted as with the introduction of more powerful techniques to quantify DNA methylation and interrogate singular CpG sites, there have been other associations formed, where the impact of DNA methylation can either be activatory (Rauluseviciute *et al.*, 2020) and repressive (Curradi *et al.*, 2002; Lopes *et al.*, 2008) towards gene expression depending mainly on the location of the DNA methylation changes relative to the gene . Increases in DNA methylation at the promoter or the first intron of genes has been linked with both repression and activation of gene transcription (Li *et al.*, 2017; Anastasiadi *et al.*, 2018; Nam *et al.*, 2020). Other links have been observed for enhancer regions where DNA methylation increases have been negatively linked with gene expression (Angeloni and Bogdanovic, 2019; Ordoñez *et al.*, 2019) with the activity of some enhancers being strongly repressed by DNA methylation due to the inhibition of specific transcription factor binding with allosteric interference or competition from methyl-dependent transcription factors (Clemens *et al.*, 2020). Due to the different sensitivity of transcription factors to changes in DNA methylation this could be a mechanism that is believed to participate in the fine tuning of gene expression as a response to environmental factors as well as disease (Héberlé and Bardet, 2019b). As DNA methylation plays a role in ageing (Johnson *et al.*, 2012) and disease state progression (McMahon *et al.*, 2017) the study of the consequences of methylation at particular sites is essential. Age-related DNA methylation changes have been reported for numerous tissues (Day *et al.*, 2013; Hannum *et al.*, 2013; Slieker *et al.*, 2018) yet the consequences on target gene expression has only been investigated in a few studies

that correlated the DNA methylation changes with associated target gene expression (Bacos *et al.*, 2016; Lien *et al.*, 2018) . Functional studies are further needed to understand this impact as causation with this being accomplished by the direct alteration of DNA methylation either at the global level or at the target site with the follow up on the directionality of gene expression changes. Whilst drugs such as decitabine (5-Aza-2-deoxycitidine) have been used in both the lab and the clinic to induce DNA demethylation in a non-specific manner (Christman, 2002; Ramos *et al.*, 2015), it is only within the last decade that targeted DNA methylation editing has become possible.

The development of genome editing tools such as TALEN or CRISPR-Cas9 systems permitted the study of more depth the consequences of DNA variants and epigenetic marks in on region functionality in both cell lines and animal models providing (Adli, 2018). The CRISPR-Cas9 system is a DNA editing construct composed of a Cas9 endonuclease from *Streptococcus pyogenes* and a short guide RNA sequence that serves to direct the Cas9 to the target region after which a double strand DNA break occurs near the target's PAM site (i.e. protospacer adjacent motif contained in the guide RNA sequence) (Jinek *et al.*, 2012; Redman *et al.*, 2016). This system used as a tool greatly improved the ability of genetic modification capabilities such as deletion, replacement or insertion of DNA sequences into the target's genome that results in transcriptomic changes such as gene inactivation (Knott and Doudna, 2018).The Cas9's catalytic domain for inducing double strand DNA breaks can be modified to nullify this activity yet still retaining the capacity for the Cas9 to be targeted anywhere in the genome (Qi *et al.*, 2013). This permitted the development of variants by fusing the catalytically inactive Cas9 (termed dCas9) to different functional proteins (e.g. DNMT3a, TET1CD) (Liu *et al.*, 2016; McDonald *et al.*, 2016) giving rise to molecular tools that permit targeted tuning of epigenetic mechanisms. DNA Methyltransferase 3 Alpha (DNMT3a) is one of two *de* novo DNA methyltransferases proteins that imprint the methylation pattern observed normally for X-inactivation and gene silencing of different cell types in humans (Zhang *et al.*, 2018). Since DNMT3a is a key enzyme in the normal methylation pathway this puts it as a good candidate to fuse to the dCas9 to achieve a construct that does targeted methylation (Smith and Meissner, 2013; Liu *et al.*, 2016). The produced fusion of dCas9-DNMT3A can catalyse targeted DNA methylation within 320 bp region surrounding the target PAM sequence (Liu *et al.*, 2016). The Ten-eleven

translocation dioxygenase 1 (TET1) is an enzyme that catalyses the conversion of 5mC to 5hydrocymC, thus playing a role in active CpG DNA demethylation pathways in mammalian cells (Wu and Zhang, 2014). The dCas9-TET1cd has been successfully used to reprogram fibroblasts to myoblasts with demethylation of the *MyoD* distal enhancer, in turn activating the expression of *MyoD (Liu et al., 2016)*. Although tools for targeted methylation/demethylation were only developed recently (Liu *et al.*, 2016), they are increasing being used to study the impact of DNA methylation at specific loci on transcription (O'Geen *et al.*, 2017; Stepper *et al.*, 2017; Pflueger *et al.*, 2018a).

Studies linking ageing with transcriptome changes or correlating age-related DNA methylation changes with gene expression have not yet been performed in musculoskeletal tissues and this is therefore the focus of this chapter. For the six regions that undergo age-associated DNA methylation changes discussed in the previous chapter, the effect of DNA methylation on their activity and the impact of this on their target's gene expression will be investigated in cell line as well as in tissues such cartilage, synovium, fat pad, bone and peripheral blood mononuclear cells (PBMC).

4.2 Aims

1. To understand the impact of DNA methylation on gene expression using cell line models.

2. Confirm that the regions containing age-CpGs can function as promoters or enhancers in a chondrocyte-like cell line model.

3. To examine if gene expression of the potential target genes of age-CpGs correlates with age in musculoskeletal tissues.

4. Investigate the relationship between gene expression and DNA methylation in human musculoskeletal tissues and peripheral blood mononuclear cells.

4.3 Results

4.3.1 Examining the effect of DNA methylation on gene expression in SW1353, TC28a2 and HAC cells using the DNA methylation inhibitor 5-Aza 2'deoxycytidine

The previous chapter confirmed that methylation of specific CpGs within the promoters of the *ELOVL2, FHL2*, *ZNF551* and *ZNF577* genes as well as two putative enhancers within the *PCDHB* gene and *PCDHGA* clusters on chromosome 5 undergo age related hypermethylation in cartilage. Based on the location of the age-CpGs, we hypothesis that they may regulate expression of these genes in a methylation dependent manner. To investigate the potential transcriptional consequence of DNA methylation at age-CpGs, the expression of these genes was quantified in cells exposed to the global demethylation compound 5-Aza 2'deoxycytidine (5-Aza, Decitabine). 5-Aza is a cytosine analogue that gets incorporated into DNA during replication with subsequent binding to DNMT1, inhibiting its activity to methylate further DNA as well as by inducing selective degradation of DNMT1 protein (Christman, 2002; Ghoshal *et al.*, 2005). This causes the cell's genome to passively lose methylation marks at the DNA replication step (Cheng *et al.*, 2004; Palii *et al.*, 2008). 5-Aza has been previously used to globally decrease DNA methylation levels in tissue culture environments and is used in clinical practice for myelodysplastic syndromes (Valdez *et al.*, 2015). Genes whose expression is regulated by the DNA methylation status of their promoter or enhancers will be expected to show changes in the expression profile after 5'AZA treatment.

In order to examine the effect of DNA methylation on gene expression, the SW1353 chondrosarcoma cell line, the TC28a2 immortalised rib chondrocyte cell line and primary human articular chondrocytes (HAC) (average age = 66 years, N=6) were treated with 5-AZA. 5-AZA can be cytotoxic to cells, inducing cell apoptosis (Stresemann and Lyko, 2008), and therefore the SW1353 and TC28a2 cells were exposed to two concentrations of 5-Aza for 72hrs where these were used to investigate cytotoxicity and impact on DNA methylation and gene expression. The concentration of 5-Aza chosen for the SW1353 treatment was

0.25 µM and 0.5 µM this being based on previous optimisations from within our group (Bui *et al.*, 2012) with the concentration of 5-Aza for TC28a2 at 5 µM and 10 µM being based on a previous publication (Alvarez-Garcia *et al.*, 2016b). For HAC cells the concentration of 5 µM 5-Aza was chosen based on a previous optimisation and publication from our group with the cells being treated for 14 days (Reynard *et al.*, 2014).

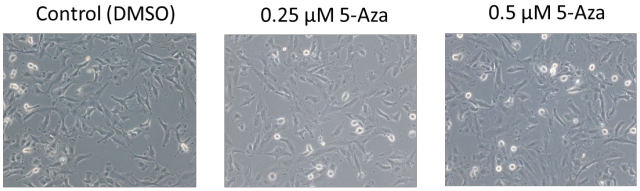
No evident morphological changes were observed during the treatment with either concentration for SW1353 cells, however a decrease in the recovered number of the cells was observed compared to the vehicle only control, with the cell number decreasing with 0.5 µM 5-Aza concentration in SW1353 (Figure 4. 1 A). To determine the impact the 5-Aza treatment had on DNA methylation levels, bisulphite pyrosequencing was used to investigate the DNA methylation levels at three age-CpG regions analysed (encompassing a total of nine CpGs (Figure 4. 1 B). Whole cell protein samples were also collected to determine whether the treatment at 0.25µM 5-Aza in SW1353 had an impact on DNA methyl transferase levels of DNMT1, DNMT3A and DNMT3B. As previously reported (Ghoshal *et al.*, 2005) DNMT1 protein levels are confirmed to be decreased following the treatment with no modification in the protein levels of DNMT3A and DNMT3B (Figure 4. 1 C).

Bisulphite pyrosequencing revealed the treatment with 5-AZA decreased methylation of the *FHL2* promoter CpG cg06639320 in SW1353 cells compared to control DMSO treated cells by ~ 49% for 0.25 µM 5-Aza and by ~32% for the 0.5 µM 5-Aza (Figure 4. 1 B) with this being accompanied by an increase in FHL2 gene expression in a dose dependent manner by up to FC= 2.04 (Figure 4. 2 A). *FHL2* promoter methylation levels were not quantified in 5-Aza-treated TC28a2 or HACs due to technical issues with the pyrosequencer and time constraints. However, *FHL2* expression was significantly increased in both these cell types in 5-Aza treated cells compared to untreated cells with expression in TC28a2 cells increasing by up to FC= 5.02x with 10 µM 5-Aza and by FC= 1.81x in HAC cells (Figure 4. 2 A, Table 4.1). For the ELOVL*2* gene, the treatment with 5-Aza significantly upregulated it's expression by up to FC= 2.41 in SW1353 cells and FC= 3.49 in TC28a2 cells where this treatment also upregulated *ELOVL2* expression in 5/6 HAC donors (FC = 1.81) (Figure 4. 2 ,Table 4. 1). The age-CpG cg01644850 is located within the promoter region of the Z*NF551* and *ZNF776*

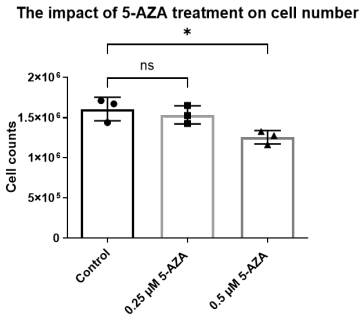
genes, which have overlapping first exons. ZNF551 expression was upregulated in both SW1353 and TC28a2 cells in a concentration dependent manner (Figure 4. 2 C) with up to a FC= 2.25x increase and FC= 2.34x increase respectively. In contrast ZNF*776* expression was reduced by ~18% (FC= 0.82) in SW1353 cells exposed to 0.25 µM 5-Aza compared to control cells, and no significant expression change in SW1353 cells with 0.5 µM 5-Aza, nor at either 5-Aza concentration in TC28a2 cells (Figure 4. 2D). For both of the genes of *ZNF551* and *ZNF776* there was no significant change in expression HACs after exposure to 5-Aza (Table 4. 1). The *ZNF577* promoter contains two age-CpGs as the cg09547119 and cg22331349 age-CpGs and the treatment with 5-Aza treatment has no significant effect on expression of *ZNF577* gene in either of the three cell types studied (Figure 4. 2, Table 4. 1). The next investigated regions were those found in putative enhancer regions with the age-CpGs of cg14566959 and cg23500537 respectively. These age-CpGs are located within enhancer regions that are nearby to the genes of *PCDHGA1* and *PCDBH1,* however neither gene was detectable in SW1353, TC28a2 cells or HAC cells (data not shown). A summary of 5-Aza induced gene expression changes is presented in Table 4.1

A)

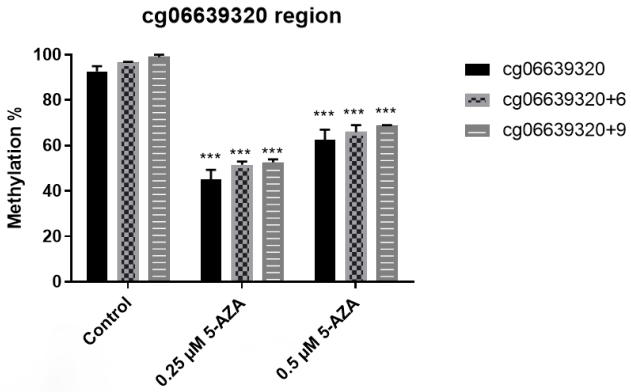
I.


Control (DMSO)
0.25 µM 5-Aza
0.5 µM 5-Aza

II.


The impact of 5-AZA treatment on cell number
*
ns
Cell counts
Control
0.25 µM 5-AZA
0.5 µM 5-AZA

B)


cg06639320 region
Methylation %
cg06639320
cg06639320+6
cg06639320+9
Control
0.25 µM 5-AZA
0.5 µM 5-AZA

C)

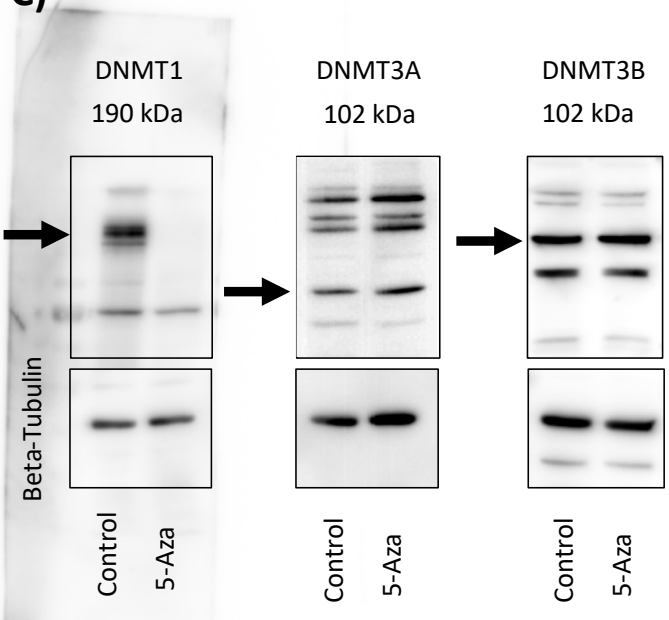

DNMT1
190 kDa
DNMT3A
102 kDa
DNMT3B
102 kDa
Beta-Tubulin
Control
5-Aza
Control
5-Aza
Control
5-Aza

Figure 4. 1. 5-Aza treatment of SW1353 concentration optimisation. A) I. Photomicrographs of the SW1353 treated with two concentrations of 5-Aza and control vehicle. II. Quantification of cell number after exposure to 5-Aza using manual counting on the haemocytometer. A dose dependent reduction was observed after the treatment with 5-Aza. Bars indicate S.D., statistical analysis was done using one-way ANOVA using Dunnette correction for multiple comparisons (n = 3), $p<0.05$ *. B) Bisulphite pyrosequencing results for the treatment of SW1353 with different concentrations of 5-Aza for 72 hours. DNA methylation changes presented per region for the CpG cg06639320 and its neighbouring CpGs. One-way ANOVA with Dunnette correction method ***$p<0.001$. C). Western blotting results for the impact of 5-Aza treatment (0.25 µM) in SW1353 on DNMT1, DNMT3A and DNMT3B proteins with Beta-tubulin being used as a loading control. The black arrow marks the expected sized of the protein.

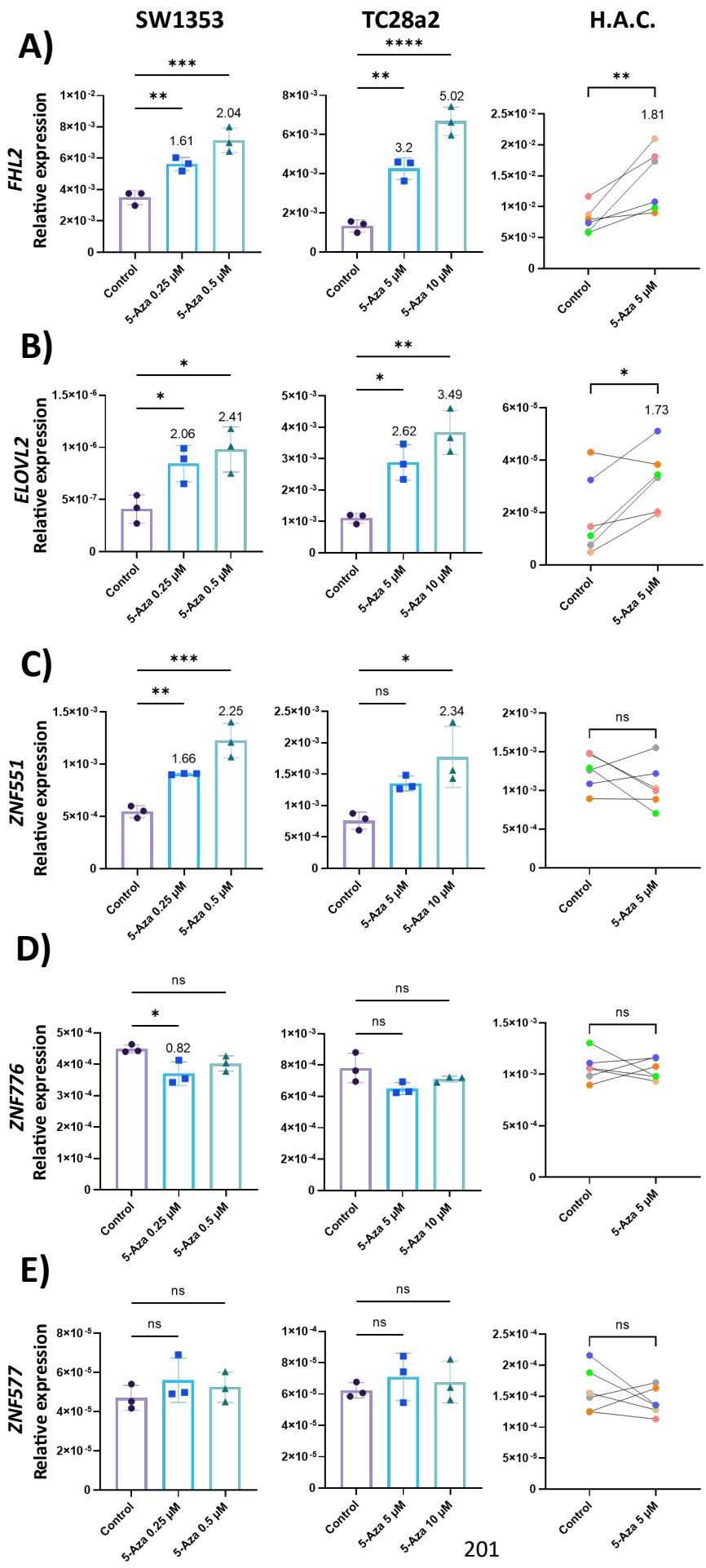

SW1353
TC28a2
H.A.C.
A)
FHL2
Relative expression
B)
ELOVL2
Relative expression
C)
ZNF551
Relative expression
D)
ZNF776
Relative expression
E)
ZNF577
Relative expression
Control
5-Aza 0.25 µM
5-Aza 0.5 µM
5-Aza 5 µM
5-Aza 10 µM
ns

Figure 4. 2. 5-Aza treatment impact on the target gene's expression for age-CpGs performed in SW1353 and TC28a2 cell lines and on HACs. Gene expression data for FHL2 (A), ELOVL2 (B), ZNF551 (C), ZNF776 (D), ZNF577 (E). For the cell lines SW1353 and TC28a2 (n=3) one-way ANOVA analysis was performed with the fold change value between treatment and control indicated above significant conditions. Relative expression is based on the expression of *GAPDH, HPRT1* and *18s* housekeeper genes; ns-nonsignificant, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. For the HAC samples (N=6) paired t-test analysis, ns-nonsignificant, *p<0.05, **p<0.01.

Table 4. 1. Summary of the impact of 5-Aza on fold change gene expression for several genes investigated as a heatmap table. Relative expression is based on the expression of *GAPDH, HPRT1* and *18s* housekeeper genes

Fold change		Gene						
Cell type	Group	*FHL2*	*ELOVL2*	*PCDHB1*	*PCDHB6*	*ZNF551*	*ZNF776*	*ZNF577*
SW1353	Ctl vs 0.25 µM Aza	1.61	2.06	N/A	N/A	1.66	0.82	1.19
	Ctl vs 0.5 µM Aza	2.04	2.41	N/A	N/A	2.25	0.9	1.11
TC28a2	Ctl vs 5 µM Aza	3.2	2.62	N/A	N/A	1.78	0.83	1.14
	Ctl vs 10 µM Aza	5.02	3.49	N/A	N/A	2.34	0.91	1.08
H.A.C.	Ctl vs 5 µM Aza	1.81	1.73	N/A	N/A	0.85	0.98	0.88

4.3.2 Investigating the effect of DNA methylation on promoter or enhancer activity using luciferase reporter assays

A key aim is the understanding of the impact that age-related DNA methylation changes have on gene expression activity. The above 5-Aza gene expression data suggests that several genes with promoter age-CpGs are regulated by DNA methylation but does not actually demonstrate that DNA methylation impacts the regulatory potential of the region. In order to investigate this further, the relationship between DNA methylation and regulatory activity was investigated for the four promoter and two enhancer regions containing age-CpG using a luciferase assay method. To study this phenomenon CpG free luciferase reporter assays were used to assess and quantify the impact of region DNA methylation in SW1353 in a tissue culture environment. The region of interest together with the transcription start site (~200bp) containing the age-CpGs of interest were cloned into a CpG free luciferase vector and *in vitro* DNA methylation was performed to assess what impact this could have on luciferase activity. For the age-CpG sites located in

promoter regions (Figure 4. 3 B), DNA methylation reduced the luciferase activity readings for the promoter regions of *FHL2* region by F.C. = 0.13 (-87%.) and of the *ELOVL2* region by a F.C. of 0.37 (-63%). For the age-CpG region located in the promoter of *ZNF551*, DNA methylation reduced significantly the luciferase activity of the region F.C. = 0.05 (-95%) and for the promoter region of ZNF*577* by F.C. = 0.02 (-98%). For the regions being marked as putative enhancer (Figure 4. 3 C), for the *PCDHGA1* enhancer region DNA methylation reduced the luciferase activity by a F.C. = 0.35 (-65%) with a similar relationship being observed for *PCDHB1* where a reduction in activity of F.C. = 0.19 (-81%) was observed. This indicates that DNA methylation of the age-CpG regions has the potential to greatly represses the transcriptional activity of the target gene.

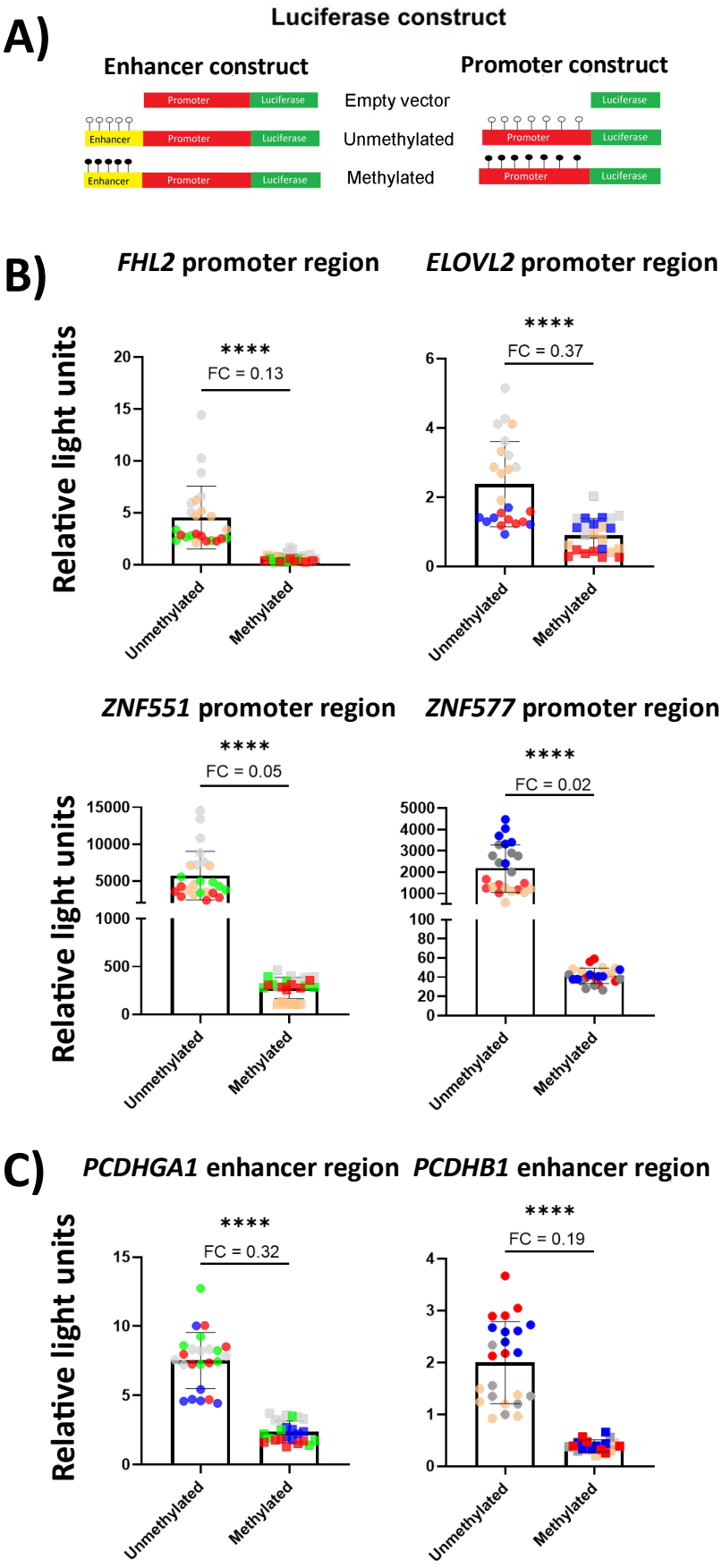

A)
Luciferase construct
Enhancer construct
Promoter construct
Promoter
Luciferase
Enhancer
Empty vector
Unmethylated
Methylated
B)
FHL2 promoter region
ELOVL2 promoter region
ZNF551 promoter region
ZNF577 promoter region
Relative light units
****
FC = 0.13
FC = 0.37
FC = 0.05
FC = 0.02
Unmethylated
Methylated
C)
PCDHGA1 enhancer region
PCDHB1 enhancer region
FC = 0.32
FC = 0.19

Figure 4. 3. Impact of DNA methylation on luciferase activity in SW1353. A) Schematic of the constructs used for the luciferase assay. B) Impact for the age-CpGs located in promoter regions and C) impact for the age-CpGs located in potential enhancer regions. Each experimental replicate was performed with six internal replicates (colour coded) for a total of 4 biological repeats for the promoter regions of FHL2, ELOVL2, ZNF551, ZNF577 and enhancer regions of PCDHGA1 and PCDHB1. Mann-Whitney U-test statistical analysis due to the data being non-parametric. $p<0.0001$ ****

4.3.3 CRISPR-Cas9 deletion of region containing age-CpGs in SW1353 cells

Following on the investigation of the role of the age-CpG regions, CRISPR-Cas9 was used to delete the region in SW1353 cells and to follow up on the impact this could have on the target's gene expression. The *FHL2* region had an average deletion % of 59.1, where the consequence on *FHL2* expression was a significant decrease of F.C. = 0.49. In contrast, for the *ELOVL2* region with an average deletion % of 49.1, the impact on *ELOVL2* expression was an increase by F.C. = 7.01. Deletion of the *ZNF551* region was not as successful as with the previous regions with an average deletion % of 31.8 and there was no significant expression change for either *ZNF551* gene F.C. = 0.78 (p value = 0.399) or for *ZNF776* F.C. = 0.86 (p value = 0.591). On the other hand, for the *ZNF577* region, the average deletion= 63.8% and this had a significant impact of the expression of *ZNF577* with a reduction of F.C. = 0.28 of the *ZNF577* gene (

Figure 4. 4).

For the enhancer regions the deletion efficiency was at an average deletion= 49.6% and 27.6% (

Figure 4. 4) for the age-CpG region of *PCDHB1* and *PCDHGA1* respectively, however both genes had no detectable expression in SW1353 cells for both the control and the deletion condition.

A)

	Deletion efficiency % (2-ΔΔCt method)					
Age-CpG region	***FHL2***	***ELOVL2***	***ZNF776***	***ZNF577***	***PCDHGA1***	***PCDHB1***
Repl. 1	52.1	43.5	38.5	65.3	31.5	48.8
Repl. 2	67.9	45.9	21	69	25.5	53.9
Repl. 3	57.3	57.9	35.9	57.3	25.9	46.1
Average	59.1	49.1	31.8	63.9	27.6	49.6

B)

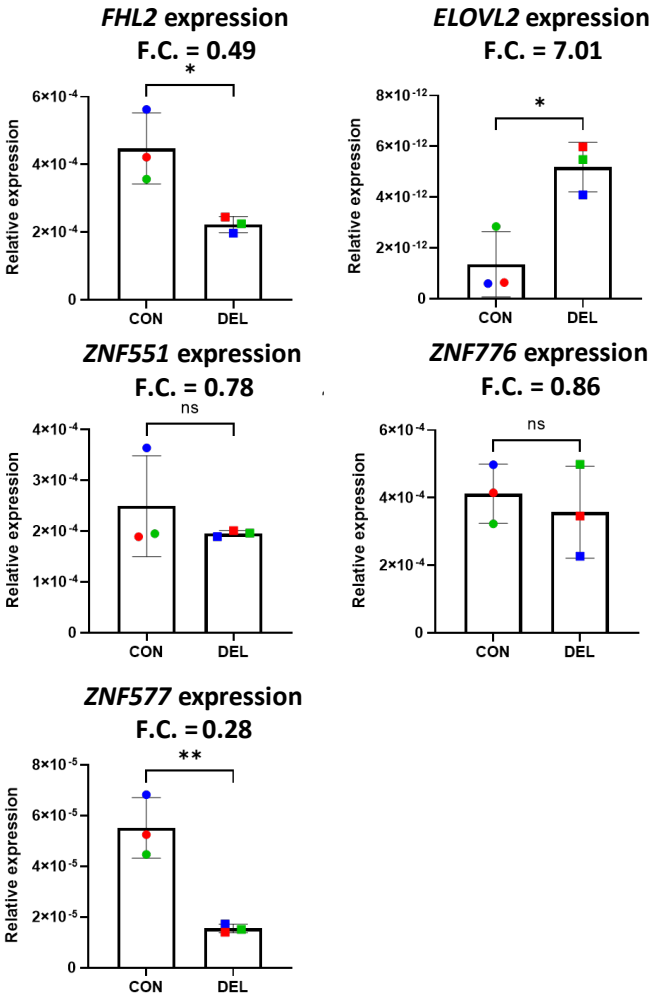


Figure 4. 4. CRISPR-Cas9 deletion of regions of FHL2, ELOVL2, ZNF551, ZNF776, ZNF577 and PCDHB1 containing age-CpGs with the impact on potential target gene expression. A) Deletion efficiency of age-CpG region deletion through the CRISPR-Cas9 system in SW1353 cells based on the qRT-PCR method. B) Deletion of age-CpG region impact on target gene expression in SW1353 cells. Relative expression is based on the expression of *GAPDH, HPRT1* and *18s* housekeeper genes. Paired t-test analysis, colours mark paired experimental sample (n=3), *p<0.05.

4.3.4 SW1353 dCas9-DNMT3a/TET1cd cell line production and validation

In order to further test whether this DNA methylation change could alter gene expression and whether this can be mimicked in a tissue culture environment an attempt was made to generate SW1353 dCas9-TET1cd and SW1353 dCas9-DNMT3A stable expression lines. The experimental objective was to alter DNA methylation levels at the target age-CpG sites and to observe the effect this will have on target gene expression. SW1353 cells have been transfected with plasmid vectors encoding for the dCas9-TET1cd and dCas9-DNMT3A constructs together with a Zeocin antibiotic resistance gene. After the transfection the cells were selection with Zeocin antibiotic for three weeks, after which whole cell protein samples were collected for protein analysis. Western blotting was performed with antibodies against the Cas9 to detect the dCas9 fusion proteins dCas9-DNMT3A and dCas9-TET1cd (Figure 4. 5 A). Several lines were tested for the constructs, where only the cell line marked dCas9-TET1cd v1 produced the right sized product at ~240 kDa corresponding to the dCas9TET1cd construct with this line being kept for further validation experiments. For the dCas9-DNMT3A construct, the cell lines marked dCas9-DNMT3Av1, v2 and v4 produced a product with a size of ~263kDa that corresponds to the expected size for the dCas9-DNMT3A construct, however dCas9-DNMT3Av2 was found to also produced a second band at ~ 160kDa and due to this reason, this was excluded from further experiments. The dCas9-DNMT3Av4 was selected for further validation due to this having a higher level of dCas9-DNMT3A according to the densiometric analysis (Figure 4. 5 B).

After confirmation of the correct molecular weight fusion protein, the cell lines were transfected with gRNAs that have been successfully used in the literature to modify CpG DNA methylation levels and produce an effect on gene expression. For the SW1353 dCas9-DNMT3a cell line, the validation gRNA targeted the site of an enhancer of *SHB* located in the first intron of the gene, which was reported to impact the expression of *SHB* when a similar system was used (Pflueger *et al.*, 2018b), and for the SW1353 dCas9-TET1cd cell line the gRNA sequence use for the validation experiment targeted the promoter region of *BRCA1* affecting the expression of the gene *BRCA1* with this system being previously used successfully for this application (Choudhury *et al.*, 2016). Several methods of transfection were attempted where the gRNAs were transfected either in two consecutive doses or just one dose and the transfection efficiency was monitored using a fluorescently labelled tracrRNA.

After 72 hours from the transfection, the cells were collected and genomic DNA was extracted for the determination of DNA methylation levels at the target CpG sites. Bisulphite pyrosequencing analysis indicated no detectable difference between control and the samples transfected with the gRNA in either configuration (Figure 4. 6).

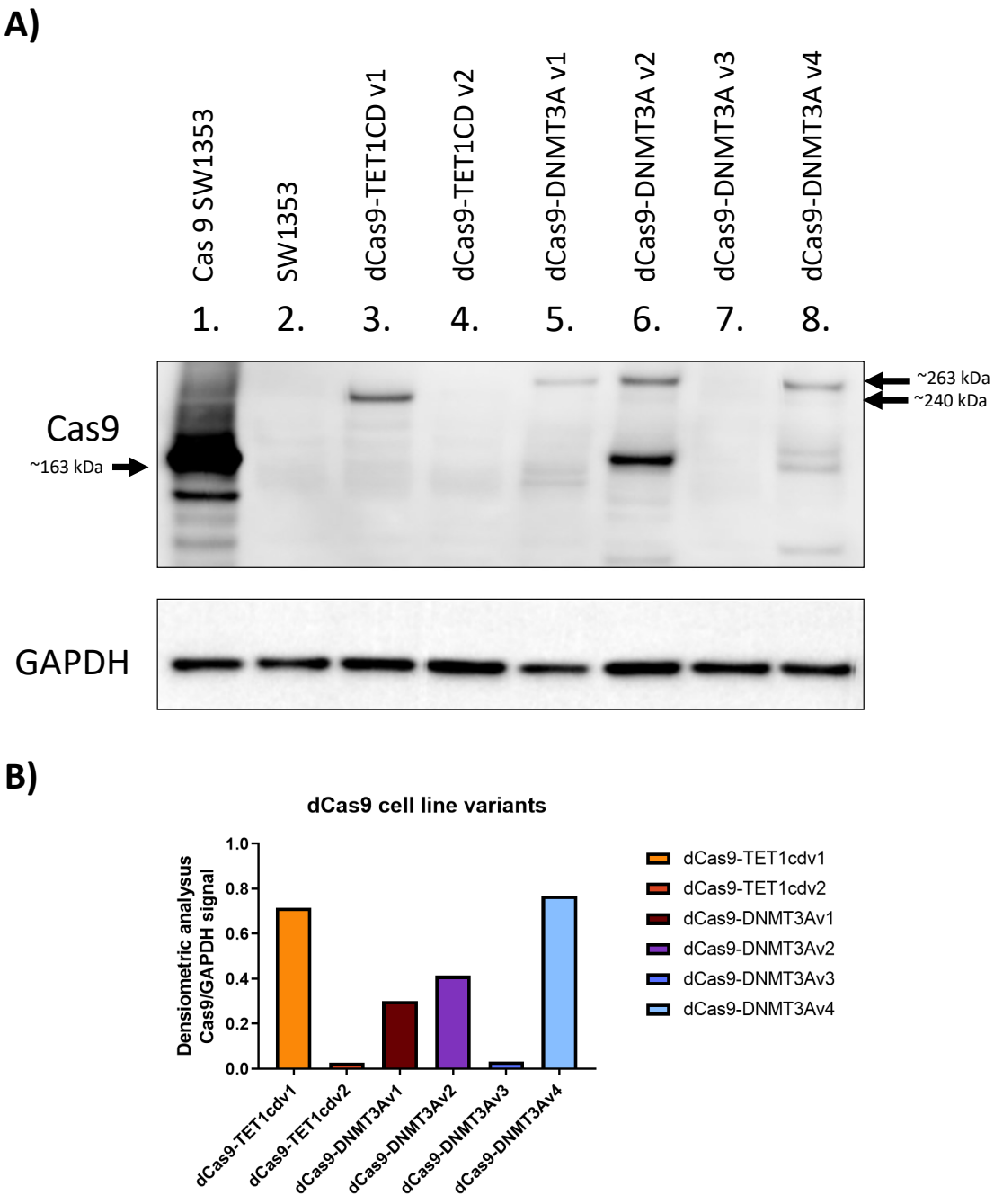


Figure 4. 5. Analysis of dCas9-TET1cd and dCas9-DNMT3A SW1353 cell lines. A) Western blotting result of the whole cell protein from SW1353 cell lines transfected with the constructs dCas9-DNMT3a and dCas9-TET1cd after selection with Zeocin 200µg/ml in the culture media for three weeks. The presence of the dCas9 constructs was assessed using anti-Cas9 primary antibody and the loading control for this experiment with anti-GAPDH primary antibody. Cas9 SW1353 cells were used as a positive control with plain SW1353 cells as a negative control. Cas9 protein detected at 163 kDa, dCas9-TET1CD detected at

~240kDa, dCas9-DNMT3A detected at 263 kDa.B) Densiometric analysis of Cas9 normalised against the loading control GAPDH for each respective sample.

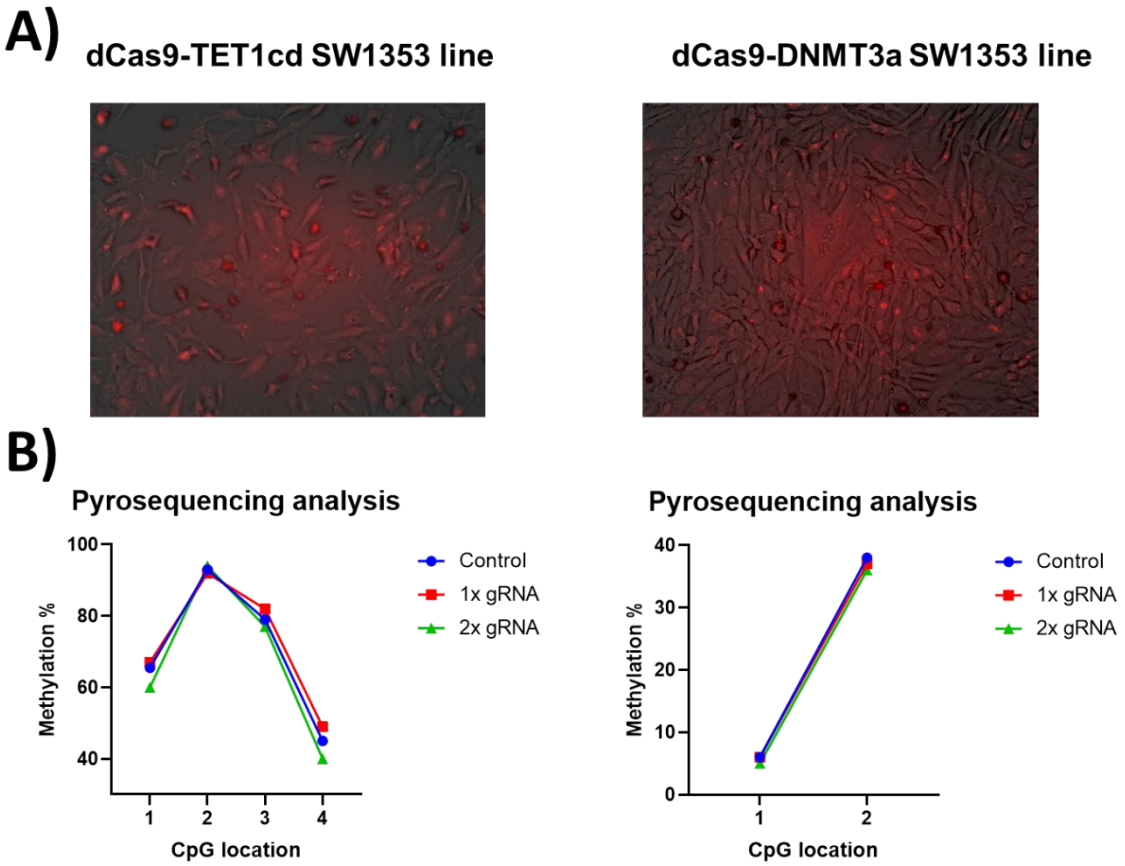


Figure 4. 6. Testing of the selected SW1353 cell lines expressing the dCas9-TET1cd constructs or the dCas9-DNMT3a. A) Fluorescent microscopy image of the cell lines after transfecting with a fluorescently tagged tracrRNA. B) Bisulphite pyrosequencing analysis results 48 hours after gRNA transfection, with indication of DNA methylation levels for control and target gRNA. Y-axis represents the determined % methylation levels at the CpGs investigated and their location on the X-axis.

4.3.5 Gene expression changes with age and correlation with DNA methylation

Age-related DNA methylation changes have been previously observed in Chapter 3 for age-CpG regions located in the promoter of *ELOVL2, FHL2, ZNF551* and *ZNF577* gene as well as in enhancer region located nearby the genes of *PCDHGA1* and *PCDHB1*. The next experimental aim is to understand whether the expression of the supposed target genes is also related with ageing and whether these changes would correlate with the DNA methylation changes at age-CpG sites. For the age-CpG sites located in promoter regions the nearest gene's expression was assessed, however for the age-CpGs located in potential enhancers several neighbouring genes had to be assessed due to enhancer potentially targeting multiple targets as well as not necessarily the nearest gene. This analysis was performed in the previously mentioned articular joint tissues of cartilage, synovium and fat pad as well as microarray data was obtained for bone (Lien *et al.*, 2018) and peripheral blood morphonuclear cells (PBMC) (Steegenga *et al.*, 2014) to understand the significance of age in the modulation of gene expression and to understand how gene expression correlates with DNA methylation levels. All the linear regression analysis and correlations with age as well as the analysis between gene expression and DNA methylation were performed for all the CpG sites investigated in Chapter 3, however only the highest correlating CpG per age-CpG regions is displayed in the figures below. The results from all the correlation analysis are summarized in (Table 4. 2).

The analysis for the *FHL2* gene in cartilage indicated that expression is significantly decreasing with age (p value = 0.047) as well as negatively correlating with DNA methylation levels at cg06639320+9bp (p value = 0.002). In the synovium tissue, expression also showed a negative correlation with age (p value = 0.034), with DNA methylation at cg22454769 significantly repressing *FHL2* expression (p value = 0.01). In the fat pad, bone and PBMC tissue, there was no significant correlation with age or DNA methylation for this gene (Figure 4. 7, Table 4. 2).

For the gene of *ELOVL2* in the cartilage tissue no analysis with age could be performed due to the undetectable levels of *ELOVL2* expression in the majority of patient samples analysed. In the synovium tissue, expression was significantly decreased with age (p value = 0.011), where a negative correlation between cg16867657 levels and *ELOVL2* expression was

observed (p value = 0.027). For the fat pad tissue, again expression significantly decreased with age (p value = 0.001) and moreover expression also was found to be inhibited by DNA methylation at the age-CpG cg21572722. In contrast, for the bone tissue *ELOVL2* expression increased with age (p value = 0.0006) and also positively correlated with cg16867657 levels (p value = 0.03). In the PBMC, *ELOVL2* expression showed no correlation with age and no correlation with DNA methylation was observed (Figure 4. 8, Table 4. 2).

For the age-CpG sites of cg24214260 and cg1644850 due to these being located in a shared promoter of *ZNF551* and *ZNF776* both genes were investigated for both their correlation in expression with age as well as with DNA methylation. The *ZNF551* gene expression in the cartilage synovium and bone tissues, there was no significant correlation with age or DNA methylation. Interestingly, in the fat pad tissue expression did not change with age, however expression was significantly repressed by DNA methylation at cg24214260 (p value = 0.04); with a similar relationship being found in PBMC, where *ZNF551* expression did not correlate with age but instead a significant decrease in expression was associated with cg01644850 levels (p value = 0.01). For the overall differences of *ZNF551* expression levels found in the articular tissues, all the tissues had a statistically significant differences in their expression levels (Figure 4. 9). For the other potential gene of *ZNF776,* in cartilage tissue the expression of *ZNF776* was positive correlated with age (p value = 0.002) and with DNA methylation levels at cg24214260 (p value = 0.01). For the tissue of synovium, fat pad and PBMC, there was no significant correlation found with age and DNA methylation for the *ZNF776*gene, however for the bone tissue a negative correlation with age was observed (p value = 0.02) yet no further correlation with DNA (Figure 4. 10 , Table 4. 2).

Investigations for the gene of *ZNF577* in the cartilage tissue, indicated that expression positively correlated with age (p value = 0.042), however no significant relationship could be established with DNA methylation. For the synovium, a significant correlation with age is observed (p value = 0.001) and moreover a significant negative correlation is also observed with DNA methylation at cg09547119 (p value = 0.001) with this being similar in the fat pad tissue where a significant negative correlation is seen with age (p value = 0.032) and with cg09547119-2bp (p value = 0.006). In the bone tissue, *ZNF577* expression had no significant link with age or with DNA methylation. Interestingly in the PBMC, expression showed no

significant correlation with age with however a negative correlation was observed for cg09547119 (p value = <0.0001) (Figure 4. 11, Table 4. 2).

For the gene of *PCDHGA1* no detectable expression levels were found for any of the articular tissues as well as for bone and PBMC and therefore no correlation with age or DNA methylation could be established (Table 4. 2).

Lastly, for the gene of *PCDHB1* in the cartilage, synovium, bone and PBMC there was no correlation with age or DNA methylation, however for the fat pad tissue there was a negative correlation with age (p value = 0.029) as well as a negative correlation with cg23500537+9bp (p value = 0.007) (Figure 4. 12, Table 4. 2).

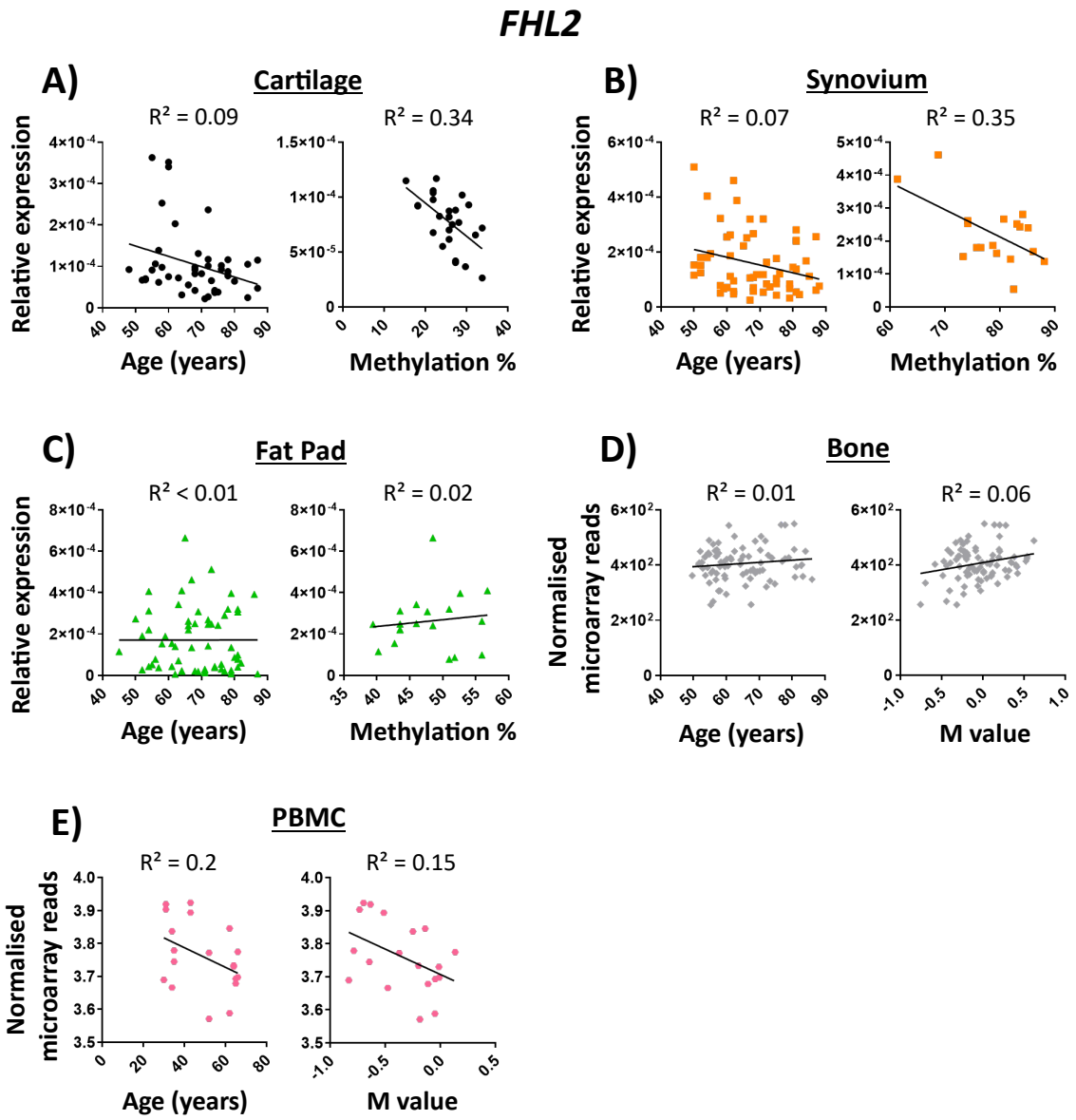


Figure 4. 7. Investigation of FHL2 gene expression changes with age and the link with DNA methylation. Linear regression analysis against age and matched DNA methylation level in multiple tissues: cartilage (A), synovium (B), fat pad (C), bone (D), PBMC (E).

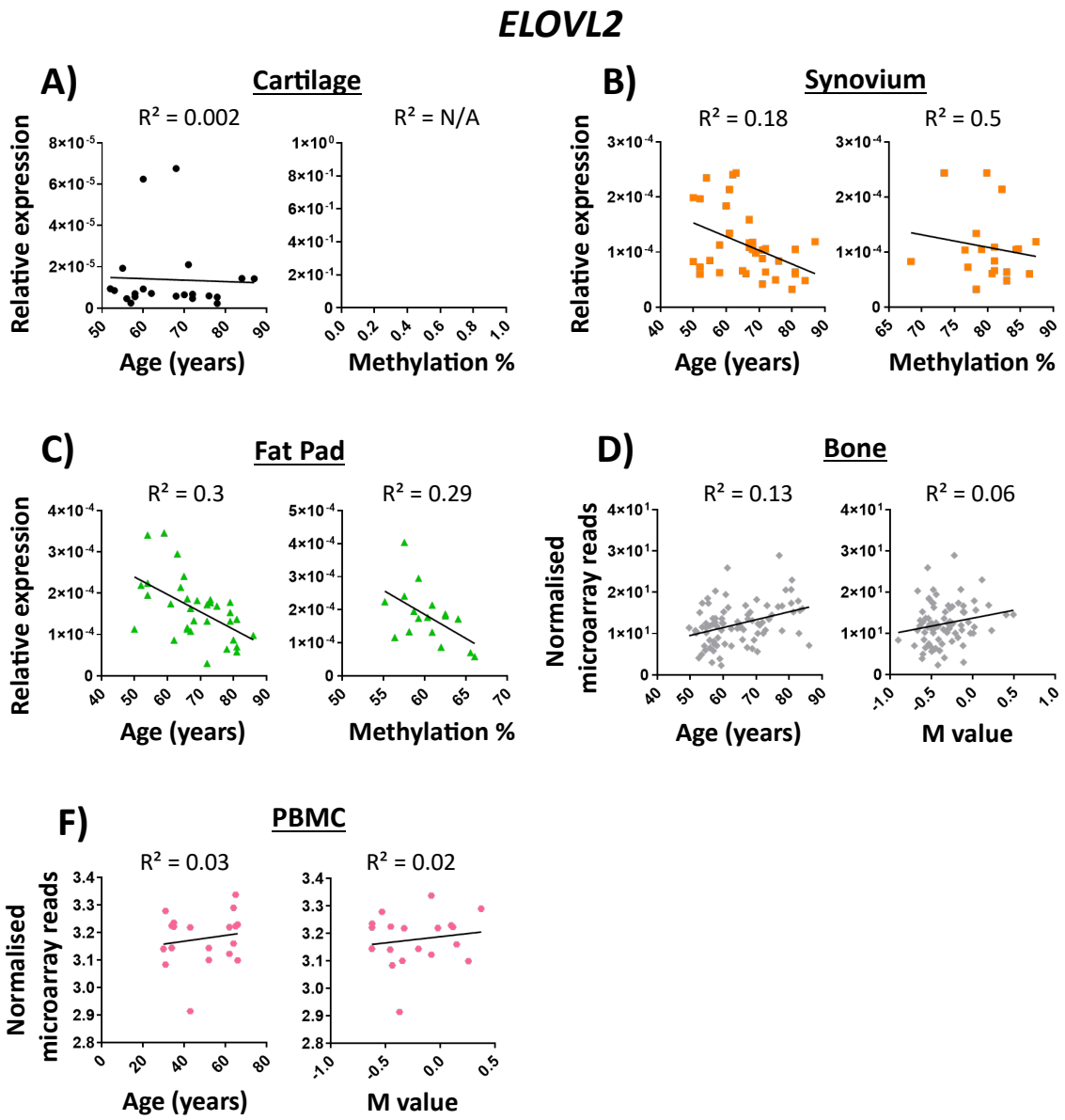


Figure 4. 8. Investigation of ELOVL2 gene expression changes with age and the link with DNA methylation. Linear regression analysis against age and matched DNA methylation level in multiple tissues: cartilage (A), synovium (B), fat pad (C), bone (D), PBMC (E).

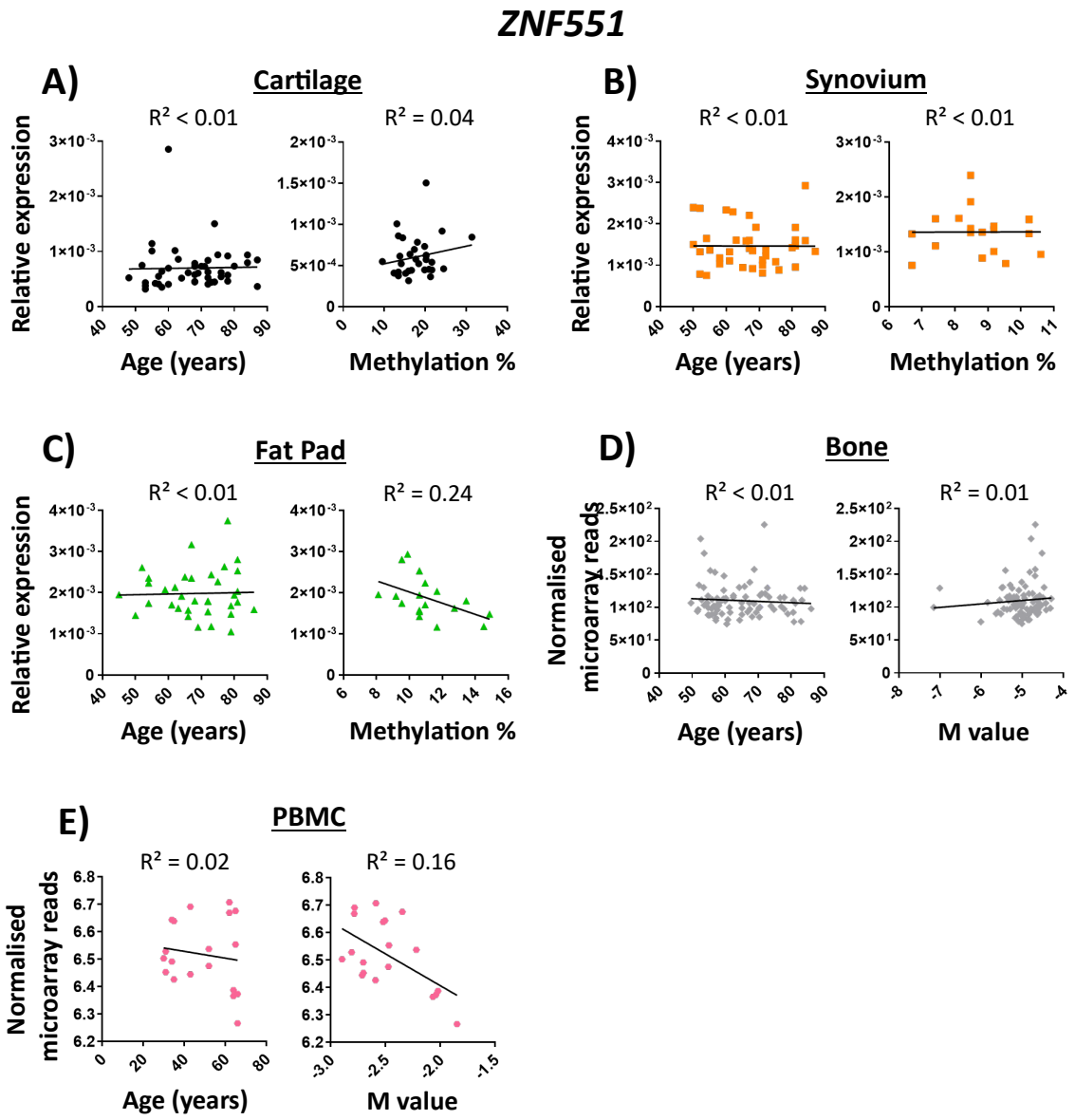


Figure 4. 9. Investigation of ZNF551 gene expression changes with age and the link with DNA methylation. Linear regression analysis against age and matched DNA methylation level in multiple tissues: cartilage (A), synovium (B), fat pad (C), bone (D), PBMC (E).

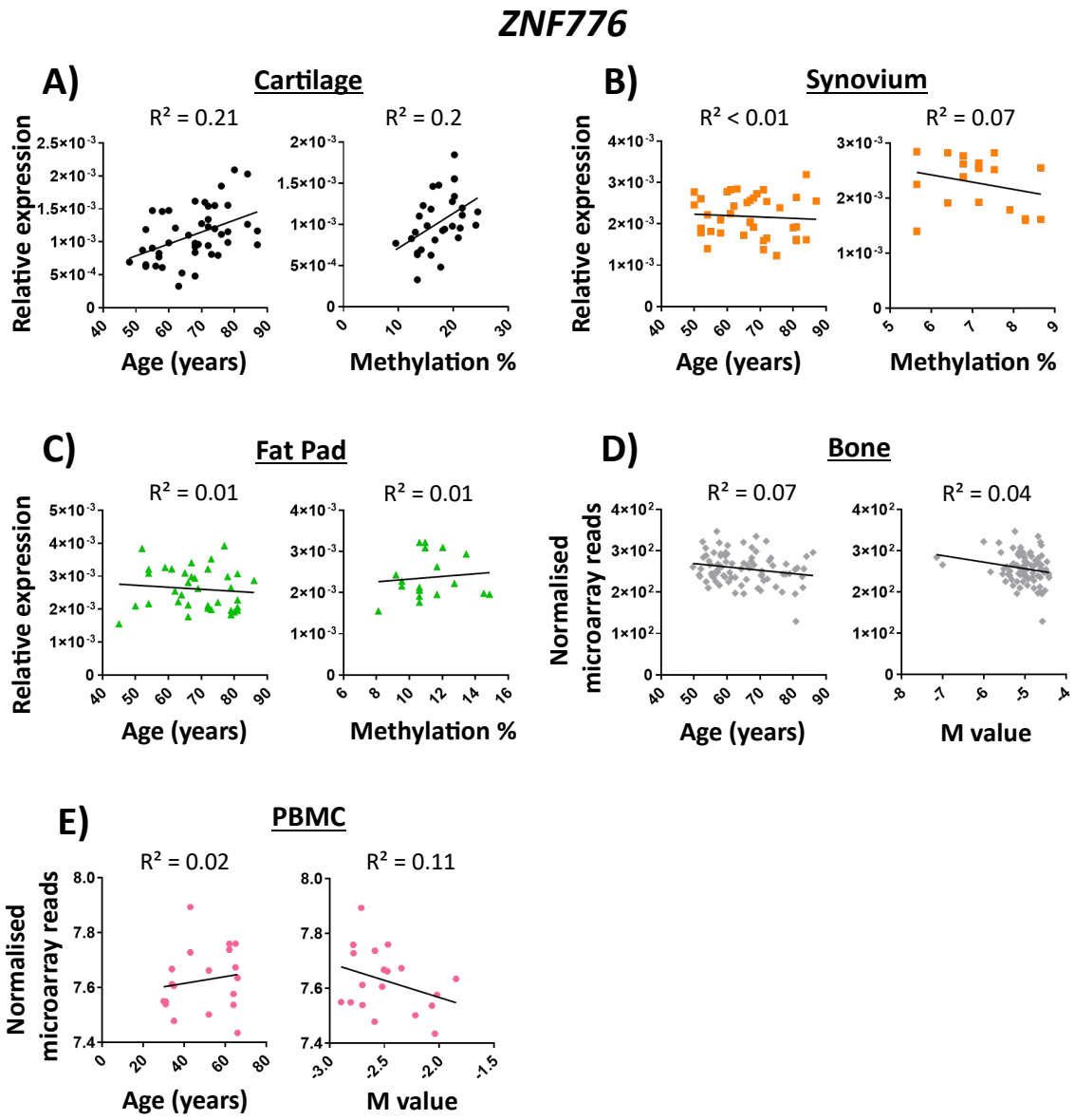


Figure 4. 10. Investigation of ZNF776 gene expression changes with age and the link with DNA methylation. Linear regression analysis against age and matched DNA methylation level in multiple tissues: cartilage (A), synovium (B), fat pad (C), bone (D), PBMC (E).

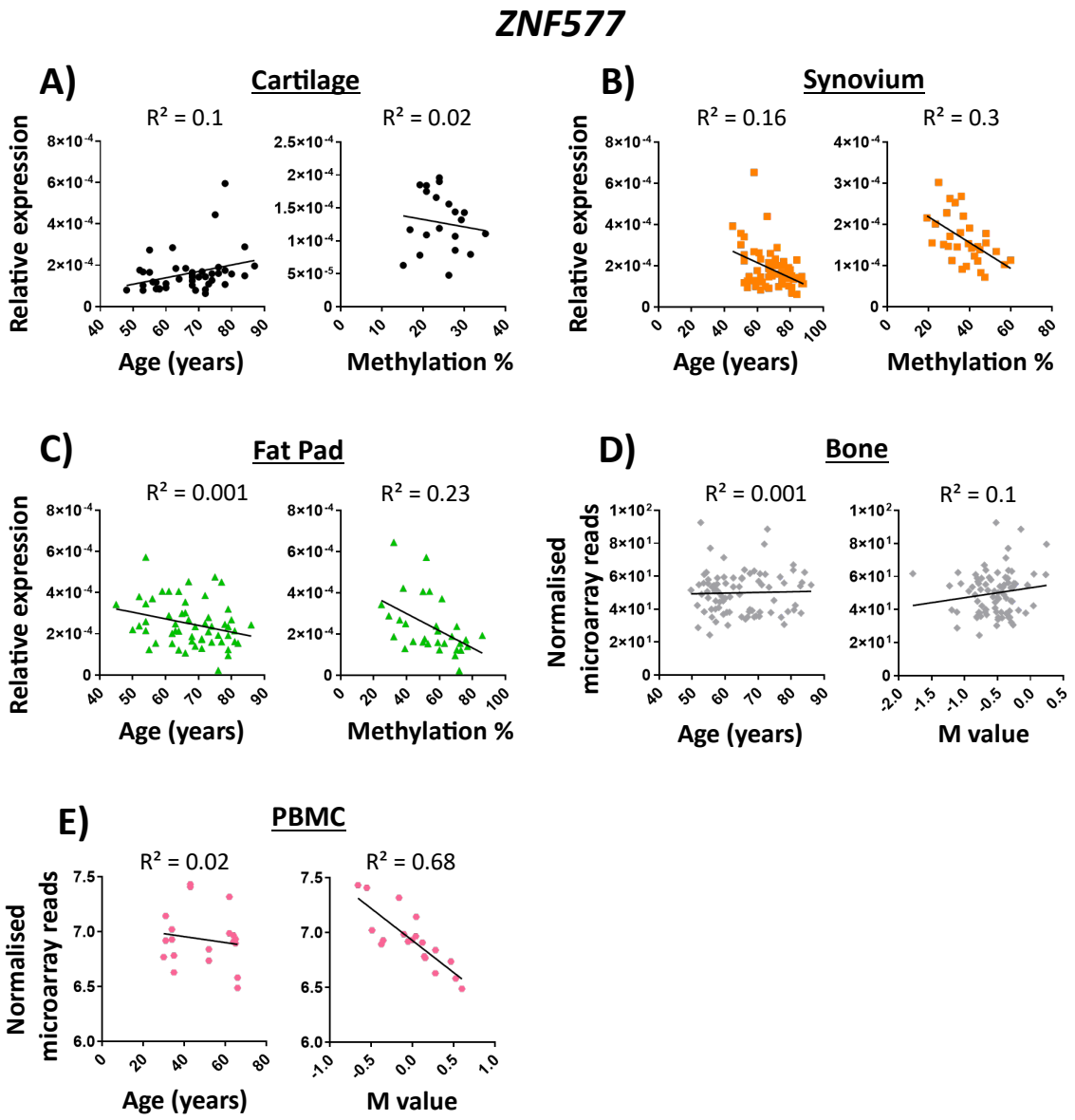


Figure 4. 11. Investigation of ZNF577 gene expression changes with age and the link with DNA methylation. Linear regression analysis against age and matched DNA methylation level in multiple tissues: cartilage (A), synovium (B), fat pad (C), bone (D), PBMC (E).

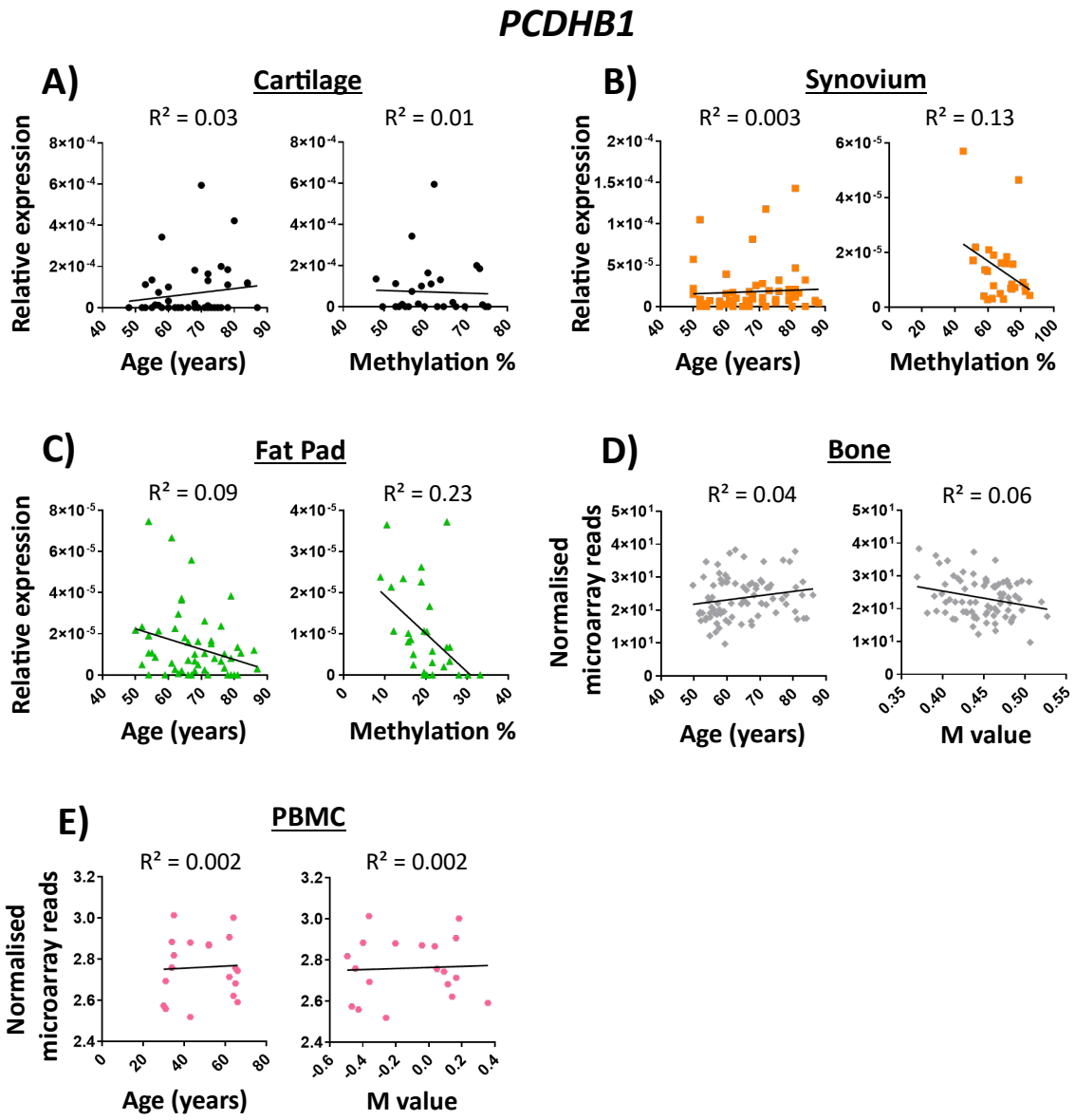


Figure 4. 12. Investigation of PCDHB1 gene expression changes with age and the link with DNA methylation. Linear regression analysis against age and matched DNA methylation level in multiple tissues: cartilage (A), synovium (B), fat pad (C), bone (D), PBMC (E).

Table 4. 2. Summary table for the linear regression and correlation analysis between gene expression and age or gene expression with CpG DNA methylation levels for the regions investigated. Significant relationships are highlighted in light green and bold text together with a mention of the type of trend observed of the correlation analysis. -ve= negative correlation relationship and +ve a positive correlation relationship between the variables analysed.

			Cartilage		Synovium		Fat Pad		Bone		PBMC	
Gene	Chromatin state region of age-CpGs	CpG	R^2 Age vs. Expression	Meth. vs. Expr.	R^2 Age vs. Expression	Meth. vs. Expr.	R^2 Age vs. Expression	Meth. vs. Expr.	R^2 Age vs. Expression	Meth. vs. Expr.	R^2 Age vs. Expression	Meth. vs. Expr.
ELOVL2	Promoter	cg21572722	0.002	N/A	**0.18 ; -ve**	0.12	**0.3 ; -ve**	**0.29; -ve**	**0.13 ; +ve**	**0.06; +ve**	0.03	0.02
		cg21572722+6 bp		N/A		**0.4; -ve**		0.14		N/A		N/A
		cg24724428		N/A		0.33		**0.29; -ve**		0.04		0.02
		cg16867657		N/A		**0.5; -ve**		0.19		N/A		N/A
		cg16867657+2bp		N/A		0.08		0.17		**0.05; +ve**		0.01
FHL2	Promoter	cg06639320	**0.09 ; -ve**	**0.21; -ve**	**0.08 ; -ve**	0.21	~0	0.02	0.02	**0.06; +ve**	0.17	0.2
		cg06639320+6bp		0.12		0.03		0.02		N/A		N/A
		cg06639320+9bp		**0.34; -ve**		0.01		0.00		N/A		N/A
		cg24079702-14bp		0.02		**0.32: -ve**		0.00		N/A		N/A
		cg22454769		0.02		**0.35; -ve**		0.01		0.02		**0.26: -ve**
		cg24079702		0.18		0.05		0.00		**0.05; +ve**		**0.29; -ve**
		cg24079702+2bp		0.10		0.13		0.01		N/A		N/A
		cg24079702+9		0.13		0.15		0.00		N/A		N/A
		cg24079702+13bp		0.08		0.00		0.01		N/A		N/A
PCDHB1	Enhancer	cg23500537	0.03	0.01	0.003	0.09	**0.09; -ve**	**0.3; -ve**	0.04	**0.06; -ve**	0.002	0.002
		cg23500537+2bp		0.01		0.13		**0.18; -ve**		N/A		N/A
		cg23500537+9bp		~0		0.07		**0.23; -ve**		N/A		N/A
PCDHGA1	Enhancer	cg14566959-13 bp	N/A	N/A	N/A	N/A	N/A	N/A	0.02	0.001	**0.4; +ve**	N/A
		cg14566959-2bp		N/A		N/A		N/A		N/A		N/A
		cg14566959		N/A		N/A		N/A		N/A		**0.43: +ve**
ZNF577	Promoter	cg09547119	**0.1 ; +ve**	0.01	**0.16 ; -ve**	**0.3; -ve**	**0.08; -ve**	**0.19: -ve**	0.001	**0.1; +ve**	0.024	**0.68: -ve**
		cg09547119-2bp		0.02		**0.21; -ve**		**0.23; -ve**		N/A		N/A
		cg22331349		0.00		**0.16; -ve**		0.01		**0.07; +ve**		**0.67; -ve**
ZNF551	Promoter	cg24214260	~0	0.04	~0	~0	~0	**0.24; -ve**	0.006	0.006	0.02	0.11
		cg1644850		0.01		~0		0.21		0.01		**0.16 ; -ve**
ZNF776	Promoter	cg24214260	**0.21 ; -ve**	0.20	0.005	0.03	0.01	0.01	**0.07; -ve**	~0	0.02	0.11
		cg1644850		0.15		0.07		0.01		0.04		0.006

4.4 Discussion

Based on the age-associated DNA methylation changes observed in Chapter 3, the work in this chapter focuses to investigate the impact of these changes as a consequence on target gene expression.

4.4.1 DNA methylation impact after 5-Aza treatment

The treatment with 5-Aza, indicated that most genes studied are consistently and significantly upregulated when treated with 5-Aza, that causes a decrease in global DNA methylation levels with the strongest effect being observed for the genes *FHL2* and *ELOVL2*. Moreover, these showed a dose dependent effect in the cell lines with increasingly amounts of 5-Aza resulting in higher expression. Other genes such as *ZNF577* showed no significant modulation by this treatment and genes such as *ZNF551* showed a significant change only in the SW1353 and TC28a2 cell lines. This experiment would suggest that DNA methylation could play a role in modulating their expression, however the limitation with the treatment is the fact that there is no targeted loss of DNA methylation, the phenomenon being global. The increase in gene expression could be due to the regions of interest that are promoters/enhancer becoming less methylated, however there could be other regulatory regions that could become active or inactive and these could contribute to the observed expression changes as well. Therefore, further work investigates the impact of DNA methylation at the age-CpG sites.

4.4.2 Impact of DNA methylation using the luciferase assay

The luciferase assay investigation looked at how the DNA methylation status could influence gene transcription by cloning the age-CpG regions in a CpG free luciferase vector followed by *in vitro* DNA methylation. All the regions assessed showed that they can be impacted by DNA methylation and consistently a reduction in luciferase activity was observed with an increase in DNA methylation, suggesting that it is repressive towards gene expression. The smallest, yet still significant, impacts of DNA methylation were observed for the promoter of *ELOVL2* region and *FHL2* region, whereas the highest changes were observed for the promoter regions of *ZNF577* and *ZNF551*. For the enhancer regions of *PCDHGA1* and *PCDHB1* there was a strong reduction in the luciferase activity with methylation suggesting

that DNA methylation could impact the functionality of these regions in an inhibitor manner. DNA methylation reduces luciferase activity and therefore is hypothesised to be inhibitory in these conditions however, the limit of the technique is the potential interactions between DNA methylation and histone marks that are absent in a luciferase vector making the conditions artificial (Kumar and James, 2015; Zhu *et al.*, 2016). Another limitation is with the enhancer construct where enhancers do not natively sit upstream of the target promoter at a distance of a few bases therefore this might force an interaction to occur even if it could not in normal conditions (Kyrchanova and Georgiev, 2021). What this technique does show is that there are transcription factors in the cell line studied that can interact with the regions studied here and that DNA methylation greatly reduces the ability of the region to support target gene expression.

4.4.3 CRISPR-Cas9 systems

The CRISPR-Cas9 system was used to delete the regions of the age-CpGs to further understand whether these regions are important to the target gene's functionality. The results indicated a reduction in gene expression in the expression of *FHL2, ZNF577* when the respective age-CpG regions were deleted and in contrast, when the age-CpG *ELOVL2* region was deleted expression levels increase for *ELOVL2* suggesting that this region could have some inhibitory functions in this cell line. No other transcriptomic changes were observed for the other regions for the supposed target regions.

To continue in the investigation for the impact of DNA methylation changes, epigenetic modulating tools were produced such as dCas9-DNMT3a and dCas9-TET1cd system in SW1353 cells. These systems could be used to alter the DNA methylation levels at the age-CpG sites and look for the expression consequences of this directly in the cells. After the transfection period, the DNA methylation levels were assessed for the respective sites where none of the transfection protocols were able to modify the methylation levels of their supposed target region CpG sites. Several unsuccessful attempts were made, however no modification of DNA methylation occurred. It is currently unknown why these systems did not work in the current experiments and future experiments could focus on the newer and stronger variants of the constructs produced such as the dCas9-SunTag DNMT3A (Pflueger *et al.*, 2018a) and dCas9- scFv- TET1cd (Morita *et al.*, 2018; Hanzawa *et al.*, 2020).

4.4.4 DNA methylation correlation with gene expression in human tissues

Further investigation was done on human tissue samples for knee cartilage, synovium and fat pad as well as publicly available datasets for bone and PBMC, to investigate the relationship between age and the expression of the supposed target genes of the age-CpG. Gene expression changes were observed to either not change with age, decrease and in some significant examples increase with age. As DNA methylation changes were found to correlate with age at certain CpGs sites and the supposed target gene of these regions can also show changes associated with ageing then naturally the next step is to look at the correlation between these two variables. Such an analysis indicated that not all age-CpGs located in the same region show the same relationship with the target gene expression, however some of them indeed showed a significant correlation between DNA methylation and the expression of the target gene. For the *FHL2* gene's expression this significantly decreased with age in cartilage and synovium, where these changes were associated significantly with the increase in DNA methylation at the promoter age-CpGs sites. *FHL2* expression has been previously observed to positively correlate with the age-CpGs cg06639320 and cg22454769 in the pancreas tissue (Bacos *et al.*, 2016), where this is in contrast with the current findings. Although no significant expression changes for *FHL2* with age were detected in the bone tissue, the expression levels correlated with DNA methylation at cg06639320 in a positive correlation similar to what is described for the pancreas tissue. These observations would suggest that the impact of DNA methylation is not always a repressive one where the evidence suggest this can enhance expression as well. Such an observation could be possible based on the different populations of transcription factors that could be tissue specific and facilitate the interaction with the age-CpGs for different results (T'Jonck *et al.*, 2018). *FHL2* function has been observed to be involved in the tissue healing pathways and inflammation, where reduced expression of *FHL2* has been associated with slowed healing (Wixler, 2019). Since the expression of this gene changes with age and DNA methylation levels in some tissues such pathways could impact tissue functionality in ageing and makes *FHL2* a key target to study to further understand the impact of ageing as well as how this can be mitigated.

For the *ELOVL2* gene, the expression negatively correlated with age in the synovium and fat pad tissue while positively in the bone tissue. Moreover, these changes were also correlated with the changes in DNA methylation at sites such as cg21572722 and cg16867657 suggesting that DNA methylation could play a role in facilitating these changes. *ELOVL2* expression was previously observed to negatively correlate with DNA methylation at CpG sites such as cg21572722 and cg16867657 in adipose tissue (Rönn *et al.*, 2015) these being similar observations as for the synovium and fat pad tissues. These results again would suggest that these tissues could have some transcription factors in common that respond to changes in DNA methylation and facilitate these expression changes where the absence of these produced no expression in the cartilage tissue although DNA methylation changes occurred. The *ELOVL2* gene is implicated in producing an enzyme that facilitates the elongation of unsaturated fatty acids (Chen *et al.*, 2020) and its function has been linked with lipid metabolism and age-related macular degeneration where it is believe that the age associated DNA methylation changes contribute to a decline in *ELOVL2* expression and results the disturbance of mitochondrial functions and increased stress on the endoplasmic reticulum (Li *et al.*, 2022).

The age-CpG sites of cg24214260 and cg1644850 are contained in the promoter of the *ZNF551* gene, a promoter that seems to be shared with the *ZNF776* gene as well. Due to this the investigation looked at both genes to observe if they show a link with the DNA methylation changes. For all the tissues studied the expression of *ZNF551* did not have a significant correlation with age however even though no changes between expression with age were observed, the expression levels did negatively correlate with DNA methylation at cg24214260 in fat pad and cg1644850 in PBMC. For *ZNF776* expression on the other hand, there was a significant and negative correlation with age in the cartilage and bone tissue however none of these changes also correlated with DNA methylation. Both of these genes encode for zinc fingered proteins, however not much is known about their functionality.

The *ZNF577* gene expression had a positive correlation with age that was not associated to any DNA methylation changes, however in the fat pad and synovium expression negatively correlated with age and moreover these also correlated with DNA methylation changes at CpGs such as cg09547119. Interestingly no correlation with age has been observed for the

bone and PBMC tissue yet expression was negatively correlated with cg09547119 and cg22331349 in both tissues. No other previous reports found any associations between this gene and age or DNA methylation, however a report observed that DNA methylation changes located at two CpG sites that were not currently studied were associated with adiposity and menopausal state (Lorenzo *et al.*, 2020).

The age-CpG cg23500537 is located in a putative enhancer region situated at ~14kb distance from the *PCDHB1* gene where this was supposed to be a possible target. For this gene no expression changes with age or methylation were detected for cartilage, synovium, pbmc however this had a negative correlation with age and with DNA methylation in the fat pad tissue; with the bone tissue also having a negative correlation with DNA methylation. This would suggest that the impact of DNA methylation at this enhancer reflects the observations from the luciferase activity where DNA methylation was repressive towards gene expression. Since enhancers can target multiple targets (Qin *et al.*, 2022) it could be possible that there are other genes to investigate and could correlate with the changes at the age-CpG sites.

Such gene expression changes indicate how these supposed target genes for the age-CpG region could be under regulation based on ageing and gives the hypothesis that perhaps the changes at DNA methylation levels could be intermediaries for these. Even though many of the age-associated changes in DNA methylation levels were found to occur in these tissues studied, it would seem that the expression changes are not so consistent with these observations. Given that the DNA methylation variability between the youngest and oldest sample is around a 30% difference this might not be high enough to exhibit a strong and detectable effect on gene expression for some cases.

The current findings indicate that some of the age-associated changes in DNA methylation could either be responsible for the changes in gene expression associated with age or that the expression changes associated with age can result in changes in DNA methylation at CpG sites. Due to the work with 5-Aza and the luciferase assays, DNA methylation changes were strongly linked gene repression, however due to the inability of making the dCas9-DNMT3a/dCas9-TET1cd system work it still would need further investigation on the degree of actual change when DNA methylation is altered.

Further investigation would look into the possible mechanism through which DNA methylation could impact gene expression such as transcription factor recruitment or binding inhibition in order to alter gene expression and explain the observation of the significant correlations. In addition to this further investigation into the possible consequences of the changes linked expression changes would be necessary to better understand what impact these could have on cellular functionality. For the more in depth analysis the region of *FHL2* was selected to be investigated for transcription factors that can bind the region and how can this change with DNA methylation changes as well as to investigate what would be the cellular consequences of the modulation of *FHL2* transcription as it is observed to occur with ageing and DNA methylation changes.

Chapter 5: How does DNA methylation at the *FHL2* age-CpGs influence transcription factor binding?

5.1 Introduction

In previous chapters, the *FHL2* promoter was observed to undergo hypermethylation with age in many tissues with a correlation being observed between *FHL2* promoter CpG methylation and gene expression in cartilage, synovium and bone. How this impacts transcription factor (TF) binding and therefore gene expression has not been studied. To understand what could cause the observed effect, an investigation of the identity of the transcription factors that bind to the promoter of *FHL2,* and the effect of DNA methylation on their binding was performed using Electrophoretic Mobility Shift Assay (EMSA) technique.

Transcription factors are fundamental biological regulators of gene expression that are involved in controlling the rate off transcription in a temporal and spatial manner. The binding of transcription factors occurs at regulatory regions including gene promoters (Vinson *et al.*, 2011) and enhancers (Azofeifa *et al.*, 2018) together forming complexes of multi-protein factors forming the transcription complex with RNA polymerases to initiate RNA synthesis (Zhang *et al.*, 2017). Transcription factor binding may be determined and influenced by several factors including: 1. the chromatin state accessibility and structure (i.e. these factors are not observed to bind heterochromatin), 2. The underlying primary DNA sequence, 3. the DNA methylation status of the region and 4. the competition for the site from other transcription factors (e.g. steric hindrance). The impact of DNA methylation influence on transcription factor binding is of particular interest to the subject of the thesis. DNA methylation was initially associated with repression of TF binding by allosteric interference with the binding site of the factors (Bird, 2002) however, modern high throughput binding assays have found that DNA methylation promotes binding and recruitment of certain specific TFs these being termed methyl-binding domain transcription factors (MBD-TF) (Du *et al.*, 2015; Héberlé and Bardet, 2019a). More recently, CpG methylation has been linked with the binding of 542 transcription factors using methyl-SELEX and bisulphite-SELEX, with the findings of this determining the binding is influenced 23% negatively, 34% positively and ~39% neutrally and the rest showing more complex binding influence based on the underlying sequence(Yin *et al.*, 2017). Such a system permits

the change of the proportion or population of transcription factors in response to the changes in DNA methylation which then can alter the expression level of genes and therefore of the cellular signalling (Moore *et al.*, 2013; Héberlé and Bardet, 2019a). Additionally, changes in DNA methylation are seen as essential in embryonal development acting both to influence and inhibit differentiation and therefore determine cell fate, as well as important for the maintenance of phenotype (Messerschmidt *et al.*, 2014).

To study the effect of DNA methylation on transcription factor binding to the *FHL2* promoter and potentially identify these TFs, the EMSA technique was used (Garner and Revzin, 1986). This technique allows the study of interactions between DNA and protein and their ability to form complexes that uses an acrylamide gel to run under non-denaturing conditions. Fluorophore labelled double stranded DNA is incubated with cellular nuclear extract containing nuclear proteins after which it is run though a gel electrophoresis setup. The operation principle is that unbound dsDNA moves at a greater rate through the acrylamide gel than DNA bound by protein complexes and therefore this generates a banding pattern based on the size and charges of the protein complexes formed. Previously this technique employed the use of radio-labelled ^{32}P and the subsequent detection using radiographic film, however further developments permitted the use of 5' tagged nucleic acid infrared fluorophore probes to be used instead, this providing better signal stability and shorter acquisition time and the ability of signal quantification with greater linearity (Hsieh *et al.*, 2016). There are several advantages of this technique over mass-spectrometry, including the ease of use with little optimisation and training, the robustness and replicability of the data generated as well as the ability to infer the binding position as well as running binding kinetics assays (Seo *et al.*, 2019). Whereas some disadvantages can include higher time expenditure, uneven band migration between some lanes, and the fact that multiple antibodies targeting the transcription factor of choice need to be used in order to identify them (Holden and Tacon, 2011). This technique is successfully employed in the literature to study such protein-DNA interactions (Reynard *et al.*, 2014; Reynard *et al.*, 2016) where it shows great reproducibility and validation by other techniques such as mass-spectrometry. The banding pattern observed then can be interrogated using the many online prediction tools that have been developed to infer the binding of certain transcription factors based on the underlying sequence of the region in question or based on experimental chromatin-

immuno precipitation followed by sequencing (ChIPseq) results (Lee and Huang, 2013). Further addition to determine the identity of the transcription factors involved in the complex formation can be achieved through the use of antibodies against the predicted transcription factor that will either change the migration properties of the complex or inhibit its formation (Hellman and Fried, 2007). The technique is usually used as a qualitative method for the assessment of the impact of certain conditions on transcription factor binding, however with the development of conjugated infrared-fluorophores, the integration of an eight-bit photodiode sensor and together with software capable of performing densiometric analysis it is therefore possible to generate quantitative data (Pagano *et al.*, 2011). This in turn opens further possibilities of application and under certain conditions it permits to determine characteristics such as affinity and binding kinetics, determination of the preferential site of transcription factor binding and moreover the identification of bands that could be omitted due to the human error aspect of optical inspection (Pagano *et al.*, 2011).

The focus of this chapter is to determine the effect of DNA methylation on the binding of transcription factors to the age-CpG region at the *FHL2* promoter and potentially the identity of the transcription factors involved.

5.2 Aims

1. Investigate the impact DNA methylation has on the binding of protein complexes at the *FHL2* promoter age-CpG sites.

2. Using competition assays determine the strength of binding differences between unmethylated and methylated DNA.

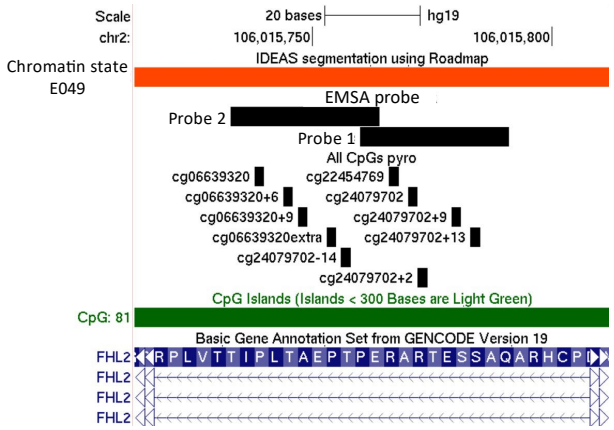
3. Determine the binding location and identity of the proteins that bind the region.

5.3 Results

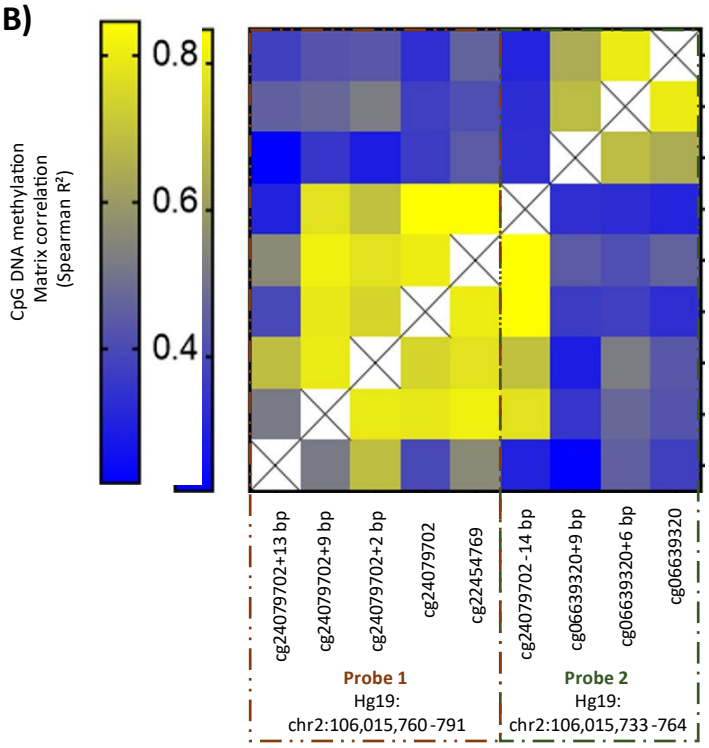
5.3.1 Design of EMSA probes

The age-CpG region contained in the promoter of *FHL2,* that was studied by pyrosequencing in chapter 3, spans approximately 60 bases and contains a total of nine CpGs that show age-related DNA methylation changes and an extra CpG that was outside the pyrosequencing assay range where any age-related DNA methylation changes are not known. Although a single probe could be designed to this region, there are some disadvantages when regions this size are used such as loss in resolution, increase in technical difficulty and increase in limitations for designing mutation assays and transcription factor consensus assays. Therefore, for the study through EMSA of the region in question it was decided to use two 32-length nucleotide probes as this permits enough resolution to assess the banding pattern, differentiate between complexes of similar size, infer the binding location of the different factors and investigate binding kinetics for individual factors. Two probes were therefore designed (Figure 5. 1), representing approximately the two clusters of CpGs whose DNA methylation level highly correlated with each other (see Chapter 3 figure 3.19). Probe 1 is the closest to the TSS of *FHL2* and this maps to hg19: chr2:106,015,760-106,015,791, including the CpGs of: cg24079702+13 bp, cg24079702+9 bp, cg24079702+2 bp, cg24079702, and cg22454769. The second probe maps to hg19: chr2:106,015,733-106,015,764 and contains the following CpGs: cg24079702-14 bp, cg06639320+14 bp (extra CpG), cg06639320+9 bp, cg06639320+6 bp, cg06639320, where no previous measurement through pyrosequencing being done for the extra CpG.

A)


Scale
20 bases
hg19
chr2:
106,015,750
106,015,800
Chromatin state
IDEAS segmentation using Roadmap
E049
EMSA probe
Probe 2
Probe 1
All CpGs pyro
cg06639320
cg22454769
cg06639320+6
cg24079702
cg06639320+9
cg24079702+9
cg06639320extra
cg24079702+13
cg24079702-14
cg24079702+2
CpG Islands (Islands < 300 Bases are Light Green)
CpG: 81
Basic Gene Annotation Set from GENCODE Version 19
FHL2
FHL2
FHL2
FHL2

B)


CpG DNA methylation
Matrix correlation
(Spearman R²)
0.8
0.6
0.4
cg24079702+13 bp
cg24079702+9 bp
cg24079702+2 bp
cg24079702
cg22454769
cg24079702 -14 bp
cg06639320+9 bp
cg06639320+6 bp
cg06639320
Probe 1
Hg19:
chr2:106,015,760 -791
Probe 2
Hg19:
chr2:106,015,733 -764

C)

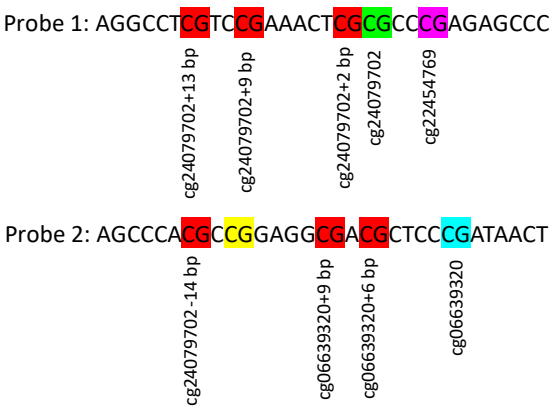

Probe 1: AGGCCTCGTCCGAAACTCGCGCCCGAGAGCCC
cg24079702+13 bp
cg24079702+9 bp
cg24079702+2 bp
cg24079702
cg22454769
Probe 2: AGCCCACGCCGGAGGCGACGCTCCCGATAACT
cg24079702 -14 bp
cg06639320+9 bp
cg06639320+6 bp
cg06639320

Figure 5. 1. EMSA probe map for the FHL2 age-CpG region. A) UCSC genome browser view of the region containing age-CpGs within the promoter of FHL2 showing chromosomal coordinates, chromatin state (in vitro differentiated chondrocytes E049), CpG island location, probe location and FHL2 isoforms. B C) DNA sequence of the designed probes, the turquoise highlighted CpG is cg06639320, green highlighted is cg24079702 and magenta highlighted is cg22454769. The yellow highlighted CpG was not analysed by pyrosequencing with the red highlighted CpGs do not have a corresponding probe on either the Illumina 450k or EPIC array, however these were included in the pyrosequencing assays and were found to also exhibit age-associated DNA methylation changes.

5.3.2 Optimisation of EMSA binding conditions

The conditions under which EMSAs are run need to be optimised to prevent artifact complex binding. Protein complex binding to these probes was initially examined using nuclear protein lysates from SW1353 chondrosarcoma cells. For both probes, the binding conditions were optimised using the reactive additives recommended by the manufacturer (Figure 5. 2 A), with the protein-DNA complexes present and the strength of these complexes varying with different additives. For both probes, one extra protein-DNA complex was only observed when $MgCl_2$ was added to the binding reaction (marked by the arrow with the "*" symbol for both P1 and P2 in Figure 5.2). Further optimisation was performed using combinations of additives (Figure 5. 2 B). The specificity of the protein-probe complexes formed were confirmed by the unlabelled probe sequence being added to the binding reaction at 50-fold excess to act as a competitor. Based on these optimisations, $MgCl_2$ was added to the probe 1 binding reaction, and $MgCl_2$+NP40 to probe 2 binding reactions. These conditions gave the strongest probe-protein complex formation (based on band intensity) and were thus used for all subsequent experiments.

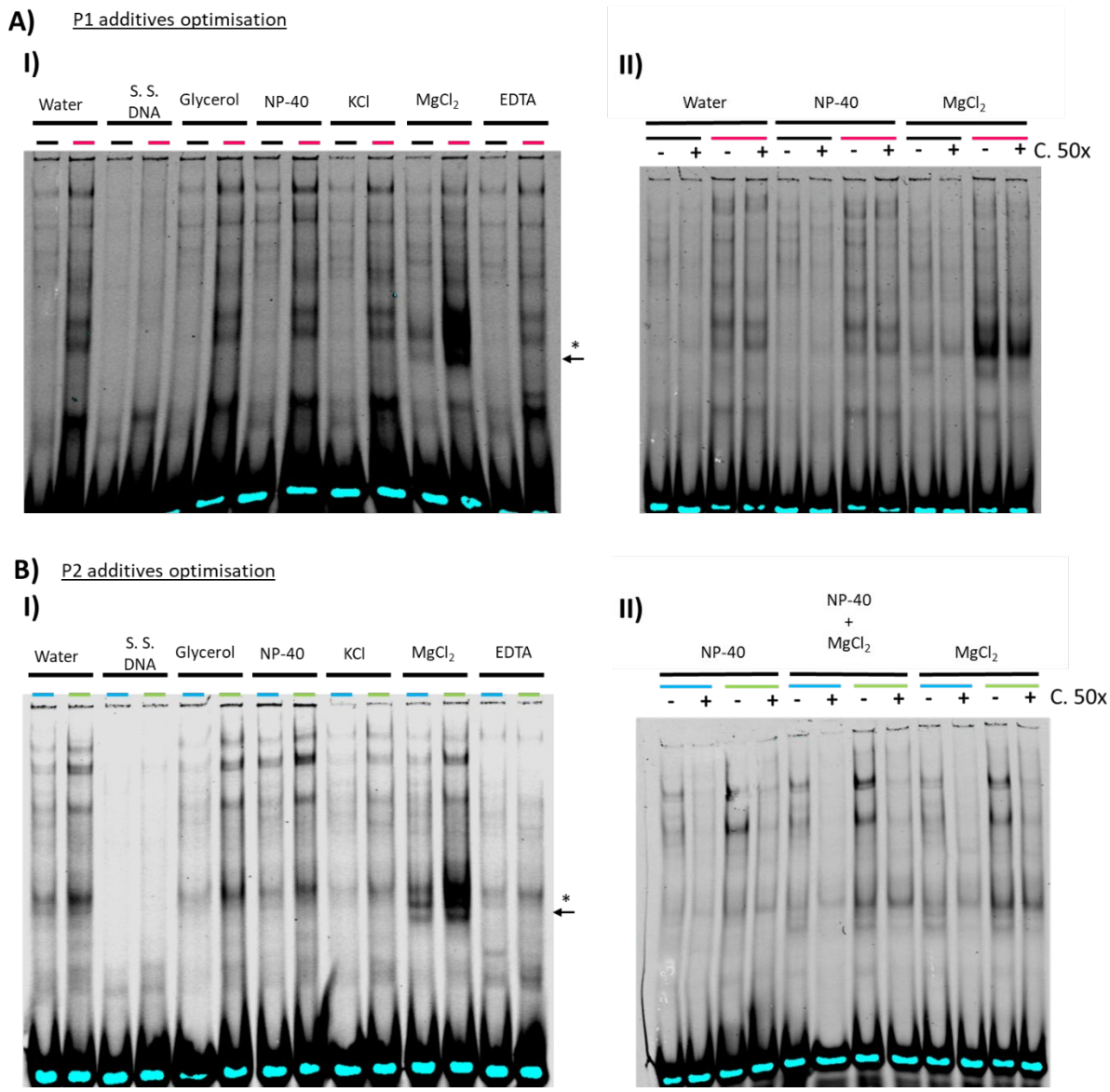


Figure 5. 2. EMSA additive optimisation. Optimisation of the EMSA binding reaction for A) probe 1 (P1) region and B) probe 2 (P2) region) with the addition of an unlabelled competitor at 50x concentration (C. 50x). A) For the P1 probe the black line indicates the unmethylated probe with the magenta line indicating the methylated version. B) For the P2 region, the blue line indicates the unmethylated probe with the olive green indicating the methylated version. Marked by the arrow with the "*" symbol for both P1 and P2 is an extra complex band that appeared only for the $MgCl_2$ additive; I) Indicates the results of the first optimisation and II) indicates the results of the second stage of optimisations for fine tuning the parameters.

5.3.3 Identification of protein complexes binding to P1

5.3.3.1 How many transcription factor complexes bind the P1 probe sequence?

To investigate what protein complexes bind the region of interest, cell nuclear protein extracts from multiple cell lines were incubated with the EMSA probe. Since the age-related DNA methylation changes occurs in other non-synovial joint tissues from around the body based on the literature, it might be expected that the same transcription factors bind to this region in multiple cell types. The initial EMSAs were thus performed using nuclear extracts from the human SW1353 chondrosarcoma cell line, the TC28a2 an immortalized rib chondrocyte line, the U2OS osteosarcoma line and MDA-MB-231 epithelial breast cancer line.

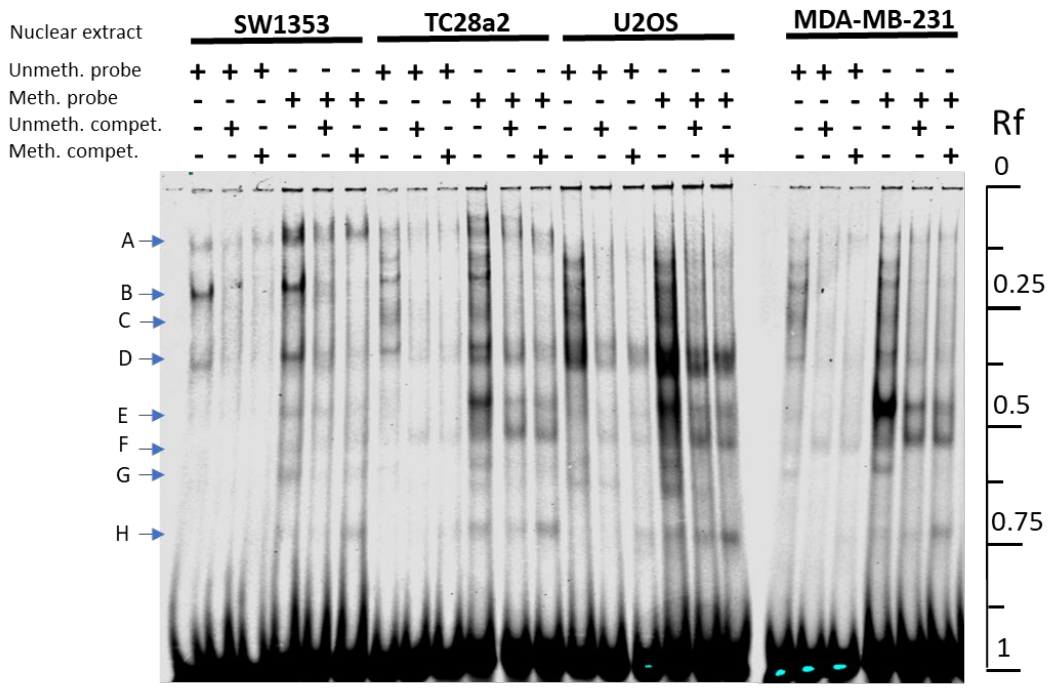
The migration distance of each band was standardised to make for a more comparable system where the migration can be more easily replicated between technical replicates as well as to determine the bands of equal migration between different cell line nuclear extracts regardless of technical artefacts. Each identified band has a migration coefficient assigned from 0 (no migration) to 1 (free probe migration) and letter nomenclature which provides also a more objective assessment based on migration characteristics. Through this migration standardisation the value is named retention factor (Rf) and because it is a ratio it does not have named units.

Based on the total bands present across all the cell nuclear extracts tested here (Figure 5. 3 A), there are a total of 14 different bands that have different corresponding Rf values, suggesting 14 different protein complexes bind this region *in vitro* (Figure 5. 3 B). Across the 4 cell line extracts, more bands were observed for the methylated version of the probe than unmethylated probe. In addition, for protein complex that bind both forms of the probe, the band signal is stronger for the methylated probe, suggesting these complexes bind more avidly when the sequence is methylated. There is variability between the cell lines in the number of bands detected for each probe, with a total of 5 bands were present in all 4 cell lines with the same probe type. A single band (termed band C) was detected with both the unmethylated and methylated probes in all cell lines, with four bands only detected with the methylated probe in all 4 lysates (bands D, E, F and G). The remaining nine bands were only

present in certain cell line nuclear extracts or with a certain version of the probe with or without methylation.

To investigate the affinity of the different protein complexes for the probe, competition EMSAs were performed, with a 50x of fold excess unlabelled unmethylated or methylated competitor added to the binding reaction. Competition showed an effect for most of the bands observed in the nuclear extracts studied (Figure 5. 3 A). Depending on the methylation status of the competitor, different effects could be observed for bands such as band A, B, and C. For the A band, the methylated competitor reduced the intensity of the band less than the methylated competitor, whereas bands B and C were fainter after addition of the methylated competitor than the unmethylated competitor, suggesting these proteins bind more strongly to the sequence is methylated. The other bands showed minimal or a non-discernible impact between the two types of competitors in any cell lines. The effect of the methylated and unmethylated competitor on protein binding is summarised in Table 5. 1.

A)



B)

Cell line		SW1353		TC28a2		U2OS		MDA-MB-231	
Letter assigned to band complex	Rf	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.
A-1	0.06	No	No	No	Yes	No	No	No	No
A	0.09	Yes	Yes	Yes	Yes	No	No	Yes	Yes
A+1	0.16	No	No	Yes	Yes	Yes	Yes	Yes	Yes
A+2	0.19	No	No	Yes	Yes	Yes	Yes	Yes	Yes
B	0.22	Yes	Yes	No	No	Yes	Yes	No	No
B+1	0.25	No	No	No	No	Yes	Yes	No	No
B+2	0.28	No	No	Yes	Yes	Yes	Yes	Yes	Yes
C	0.33	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
C+1	0.36	No	No	No	Yes	Yes	Yes	No	Yes
D	0.45	No	Yes	No	Yes	No	Yes	Yes	Yes
E	0.51	No	Yes	No	Yes	Yes	Yes	No	Yes
F	0.58	No	Yes	No	Yes	No	Yes	Yes	Yes
G	0.63	No	Yes	No	Yes	Yes	Yes	No	No
H	0.72	No	Yes	No	Yes	No	Yes	No	Yes

Figure 5. 3. P1 EMSA complexes in multiple cell lines. A) Representative EMSA results for the P1 region using nuclear extracts from SW1353 chondrosarcoma line, TC28a2 chondrocyte line, U2OS osteosarcoma line and MDA-MB-231 breast epithelial cell line and unmethylated and methylated probes. The impact of an unmethylated and methylated competitor is also tested for each probe and nuclear extract. The symbol "+" marks the addition of either probe and competitor type, where "-" marks the absence of these B) Summary of EMSA gel result banding pattern observation after visual inspection and densiometric analysis and Rf value assignment. Yes means the band was observed and no means the band is absent.

Table 5. 1. Summary of the effect of protein-probe binding upon addition of 50x fold excess methylated or unmethylated competitor at the P1 probe. Densiometric analysis each band in the EMSA gel was performed and used to determine the effect of the competitor on band intensity (and thus protein binding). “X” represents that the specific band was not observed to occur at the reported Rf value for a particular cell line with that probe, ”0 “ indicates that the addition of the competitor had no effect on band intensity, , “+” represents a less than a 50% reduction of upon addition of competitor (i.e. some competition occurred) , “++”represents >50% band intensity reduction with the competitor (i.e. strong competition occurred with the competitor).

	SW1353				TC28a2				U2OS				MDA-MB-231			
Probe	Unmethylated		Methylated		Unmethylated		Methylated		Unmethylated		Methylated		Unmethylated		Methylated	
	Competitor type															
Letter assigned	Unmethyl.	Methyl.	Unmethyl	Methyl.	Unmethyl	Methyl.	Unmethyl	Methyl.	Unmethyl	Methyl.	Unmethyl	Methyl.	Unmethyl	Methyl.	Unmethyl	Methyl.
A-1	X	X	X	X	X	X	0	++	X	X	X	X	X	X	X	X
A	++	+	++	+	++	++	++	+	X	X	X	X	++	++	++	++
A+1	X	X	X	X	++	++	++	++	++	++	++	++	++	++	++	++
A+2	X	X	X	X	++	++	++	++	++	++	++	++	++	++	++	++
B	+	++	+	++	X	X	X	X	++	++	++	++	X	X	X	X
B+1	X	X	X	X	X	X	X	X	++	++	++	++	X	X	X	X
B+2	X	X	X	X	++	++	++	++	++	++	++	++	+	+	+	+
C	+	++	+	++	++	++	++	++	+	+	+	+	+	+	+	+
C+1	X	X	X	X	X	X	++	+	+	+	+	+	X	X	+	+
D	X	X	++	++	X	X	+	++	X	X	+	+	+	+	+	+
E	X	X	++	++	X	X	+	+	+	+	+	+	X	X	+	+
F	X	X	++	++	X	X	++	++	X	X	++	++	++	++	++	++
O	X	X	++	++	X	X	++	++	+	++	+	++	X	X	X	X
P	X	X	++	++	X	X	++	0	X	X	+	0	X	X	+	0

5.3.3.2 Binding position determination of factors that bind the P1 region

Determining the binding position of the different protein factor bands can be accomplished by using oligomer competitors that replace a region of six successive nucleotides with the base adenine; this results in a sequence that has 26/32 bases that match the original sequence. By having a successive and systematic scan of the entire sequence while also keeping at least two nucleotides in common between each succession the binding position of the factors can be determined. Each mutated sequence gives rise to a total of eight combinations that spawn the entire region of interest. When these mutated competitors are used in a 50x fold excess with the normal reaction described previously, it should reveal where these protein factors bind by leaving the signal unaltered, as there is no competition from the mutated competitor.

The results of this experiment can be seen in Figure 5. 4, where both raw EMSA image (Figure 5. 4 B) and quantified results (Figure 5. 4 C) are shown. Interpretation of the binding pattern suggest a small difference between competitors for most complex bands present with exception for bands F and H. The signal for F was seen to increase at position 16-26 above the signal values of the control for both the unmethylated and methylated probe. The results suggest that for complex band H the binding region on the probe is at position 1-14 with these effects being observed stronger for the methylated probe (Table 5. 2).

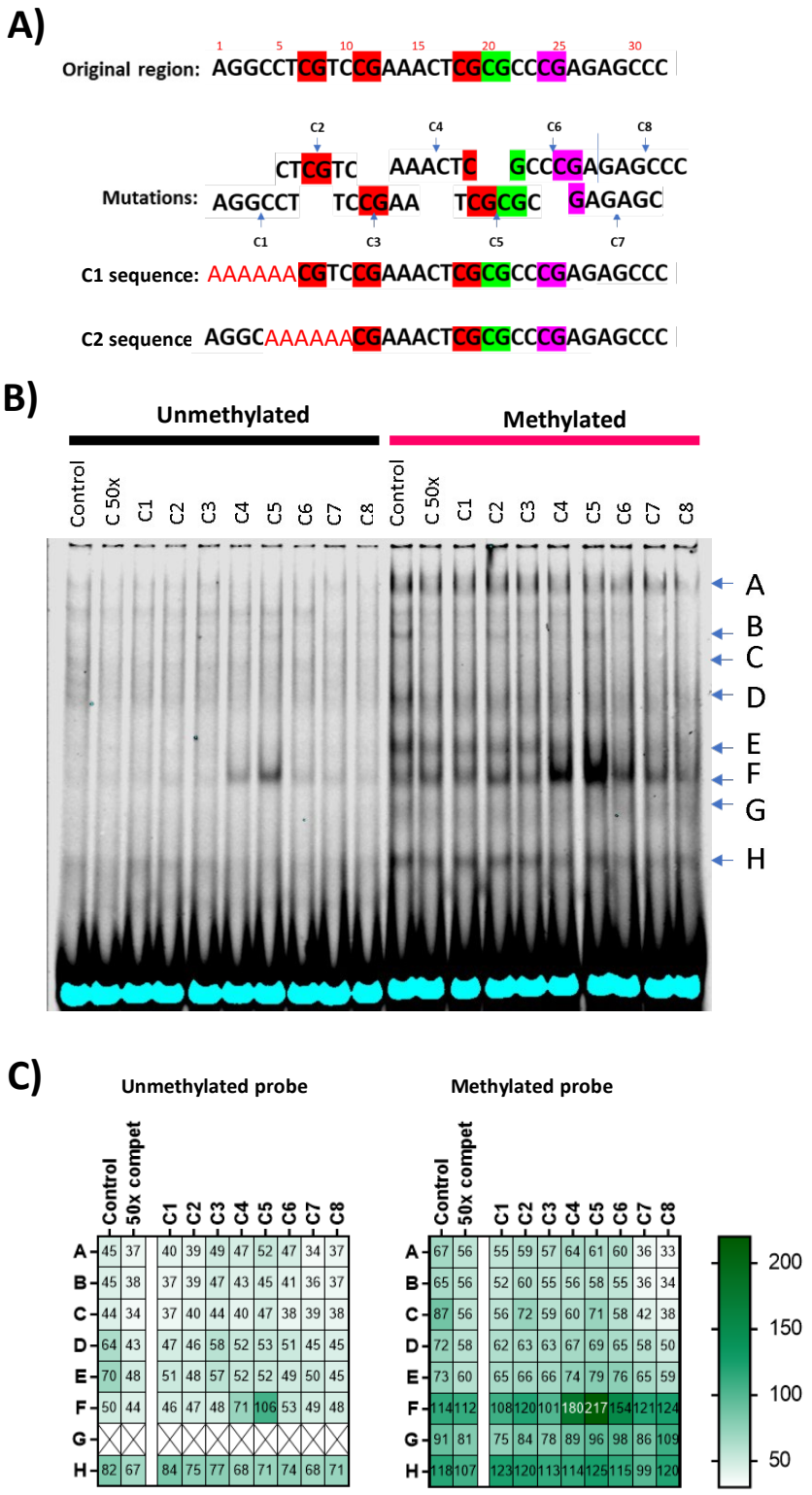


Figure 5. 4. P1 mutated competitor assay. P1 region determination of the binding location of the protein factor bands found in SW1353, a combination of eight competitors were used to determine the binding position. A) Schematic of the original probe sequence and the mutated combinations used for the identification of the binding position. Each mutated

competitor had six consecutive bases replaced with adenine, with each combination of competitor having an overlap of two bases. This gave rise to a total of eight combinations marked by C and a digit represented in the schematic B) Representative EMSA result using the different mutated competitors on the C) Heatmap of the densiometric quantification result with 50x mutated competitor to determine the preferred binding site for each identified band in SW1353 nuclear extract.

Table 5. 2. Summary of the findings from the mutated competitor assay on the P1 region. The position of binding site found is indicated for the complex bands studied as well as the determined binding sequence of the complex for each type of probe.

P1 Probe	Unmethylated probe		Methylated probe	
Band name	Position of binding	Determined binding sequence	Position of binding	Determined binding sequence
A	Unknown		Unknown	
B	Unknown		Unknown	
C	Unknown		Unknown	
D	Unknown		Unknown	
E	Unknown		Unknown	
F	16-26	TCGCGCCCGA	16-26	TCGCGCCCGA
G	X	X	Unknown	
H	1-14	AGGCCTCGTCCGAA	1-14	AGGCCTCGTCCGAA

5.3.3.3 Identification of transcription factors that may bind to probe 1.

In order to determine the possible protein factors that are binding the region of interest, the probe sequence was inputted into webtools Tfbind (Tsunoda and Takagi, 1999), Alggen (Messeguer *et al.*, 2002), JASPAR (Sandelin *et al.*, 2004), MEME (Bailey *et al.*, 2015), ChIP datasets from UCSC (Rosenbloom *et al.*, 2013), TFsite (Ghosh, 2000), LASAGNA (Lee and Huang, 2013). As there was little overlap in the factors the different tools identified, transcription factors that were identified by at least two tools were considered for further analysis, giving a total of 15 factors, only nine of which are actually expressed in SW1353 cells based on RNAseq data (Figure 5. 5 A). Transcription factors competitors were designed to mimic the binding position of the factors with the probe, and the template for this competitor was chosen based on its experimentally determined inability to exhibit any competition with the bands previously identified. None of the TF competitors was observed to have an impact on the binding pattern observed for the SW1353 nuclear extract (Figure 5. 5 B), and so the identity of the proteins binding to the probe 1 DNA sequence remains undetermined at this time.

A)

	Frequency of prediction	Probe position	TF Consensus sequence	Sequence 5'-3'
Original sequence				AGGCCTCGTCCGAAACTCGCGCCCGAGAGCCC
Factor				
Zfx	2	1-8	AGGCCTGG	AGGCCTGGTCCGAAACTCGCGCCCGAGAGCCC
NFKB	2	9-22	NGGGGAMTTTCCNN	AGGCCTCGTGGGGACTTTCCGTCCGAGAGCCC
TP53	2	18-32	CGGACATGTCCGGACATGTC	AGGCCTCGTCCGAAACTCGGACATGTCCGGAC
Sp1+3	2	16-26	GGGGGCGGGGG	GGCCTCGTCCGAAAGGGGGCGGGGGAGCCC
Myc	2	13-24	NNNCACGTGNNN	GGCCTCGTCCGATGCACGTGCTGGAGAGCCC
AP2	4	16-27	MKCCCSCNGGCG	AGGCCTCGTCCGAAAAGCCCGCTGGCGAGCCC
ELK1	2	7-22	NNNACMGGAAGTNCNN	AGGCCTTTCACCGGAAGTTCGACCGAGAGCCC
E2F	5	15-27	NNGCGCGAAANTK	AGGCCTCGTCCGAATCGCGCGAAATTGAGCCC
p300	2	7-11	CGTCC	AGGCCTCGTCCGAAACTCGCGCCCGAGAGCCC

B)

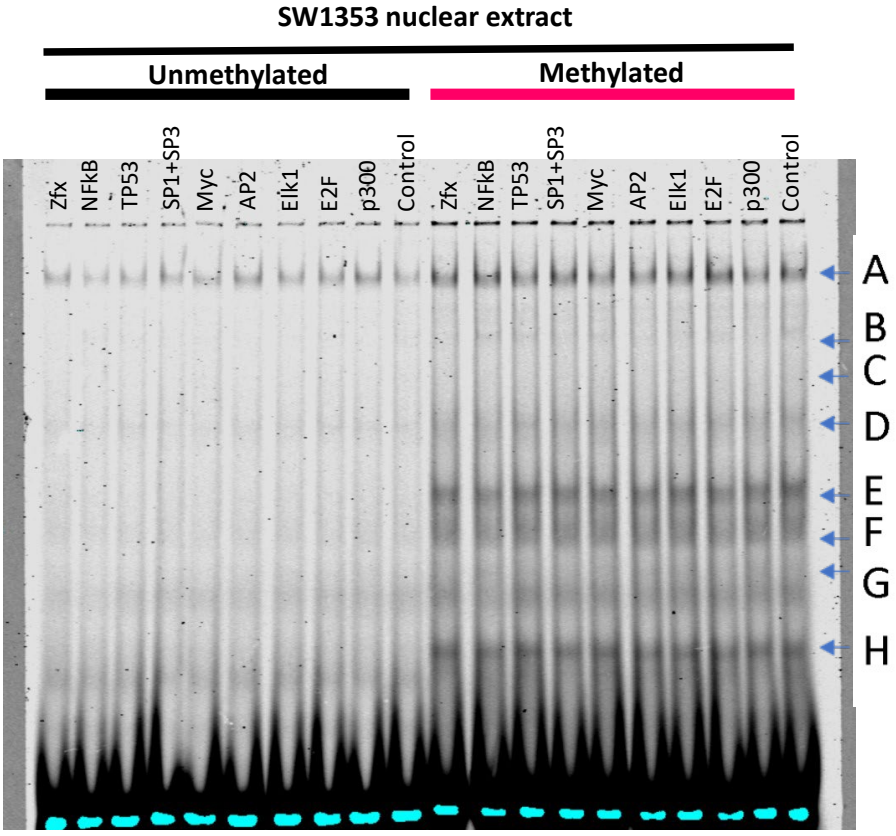


Figure 5. 5. P1 TF consensus competitor assay. Identification of and consensus competitor testing of transcription factors predicted to the P1 region. A) Table summarising the nine TF predicted by at least 2/7 tools the predicted TFs together with the frequency at which the

factor was identified, the binding position on the probe, consensus sequence, and binding sequence on the probe where the consensus sequence was added. Highlighting in green is for cg24079702, in magenta for cg22454769, and in red the other CpG sites present. Red underlined sequence corresponds to the binding position of the transcription factor on the original probe sequence. B) The representative EMSA result for the incubation with the 50x TF consensus competitors.

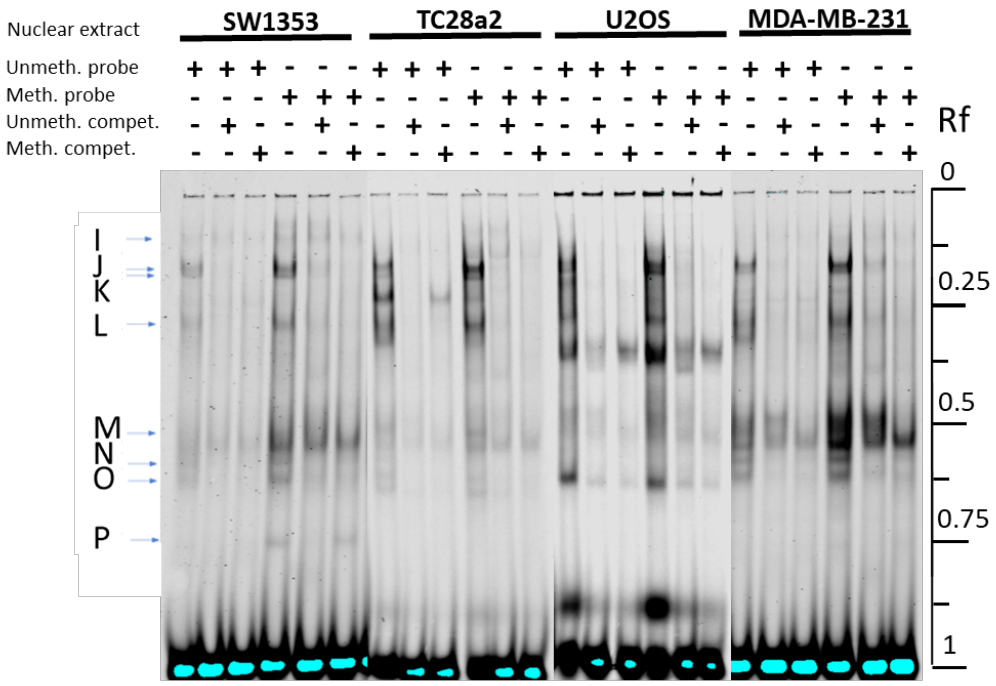
5.3.4 Identification of protein complexes binding P2 probe

5.3.3.4 How many transcription factor complexes bind the P2 probe sequence?

In order to identify protein complexes that bind to the probe 2 region of the FHL2 promoter, the probe was incubated with nuclear extract from the same four cell lines as for probe 1 (Figure 5. 6 A). A total of 20 different bands with unique RF values were identified across the different cell line-probe combinations. More bands are generally observed bands in number for the methylated version of the probe with these also showing higher signal strength (Figure 5. 6 A & B).

Four complex bands (termed J, K, L and O) were observed in all cell lines regardless of the methylation status of the probe, with the remaining bands being only present in only certain cell line nuclear extracts or with a certain version of the probe with or without methylation. Addition of 50-fold excess of an unlabelled competitor oligo with the same nucleotide sequence as the probe showed an effect for most of the bands observed in the nuclear extracts studied, with the competitor reducing these bands (Figure 5. 6). For several bands, different competition effects were observed depending on the methylation status of the competitor, such as bands J, K, L where the unmethylated competitor or was a weaker competitor than the methylated competitor. Whereas for bands such as band O, there was minimal difference between the unmethylated and methylated competitor. Some complex bands such as band I did not seem to be affected by competition from any version of the competitor (Table 5. 3).

A)



B)

Cell line		SW1353		TC28a2		U2OS		MDA-MB-231	
Letter assigned	Rf	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.
I	0.09	Yes	Yes	No	No	No	No	No	Yes
I+1	0.13	No	No	Yes	Yes	No	Yes	No	No
J	0.15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
K	0.17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
K+1	0.19	No	No	Yes	No	No	No	No	No
K+2	0.21	No	No	No	No	No	No	Yes	Yes
K+4	0.23	Yes	No	No	Yes	No	No	No	No
L	0.27	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
L+1	0.31	No	No	No	No	Yes	Yes	Yes	Yes
L+2	0.34	No	No	Yes	Yes	No	No	No	No
L+4	0.38	No	No	No	No	No	Yes	No	Yes
L+5	0.44	No	No	No	Yes	No	No	No	No
L+6	0.48	No	No	Yes	Yes	No	No	Yes	Yes
L+7	0.5	No	No	No	No	Yes	Yes	Yes	Yes
M	0.53	Yes	Yes	Yes	Yes	No	No	Yes	Yes
N	0.57	Yes	Yes	No	No	No	No	Yes	Yes
O	0.6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
O+1	0.63	No	No	Yes	Yes	No	No	No	No
P	0.74	Yes	Yes	No	No	No	No	No	No
P+1	0.88	No	No	No	No	Yes	Yes	No	No

Figure 5. 6. P2 EMSA complexes in multiple cell lines. Representative EMSA results for the P2 region using nuclear extracts from SW1353 chondrosarcoma line, TC28a2 chondrocyte line, U2OS osteosarcoma line and MDA-MB-231 breast epithelial cell line and unmethylated and methylated probes. The impact of an unmethylated and methylated competitor is also tested for each probe and nuclear extract. The symbol "+" marks the addition of either probe and competitor type, where "-" marks the absence of these B) Summary of EMSA gel result banding pattern observation after visual inspection and densiometric analysis and Rf value assignment. Yes means the band was observed and no means the band is absent

Table 5. 3. Summary of the effect of protein-probe binding upon addition of 50x fold excess methylated or unmethylated competitor at the P2 probe. Densiometric analysis each band in the EMSA gel was performed and used to determine the effect of the competitor on band intensity (and thus protein binding). “X” represents that the specific band was not observed to occur at the reported Rf value for a particular cell line with that probe, ”0 “ indicates that the addition of the competitor had no effect on band intensity, , “+” represents a less than a 50% reduction of upon addition of competitor (i.e. some competition occurred) , “++”represents >50% band intensity reduction with the competitor (i.e. strong competition occurred with the competitor).

Cell line	SW1353				TC28a2				U2OS				MDA-MB-231			
Probe	Unmethylated		Methylated		Unmethylated		Methylated		Unmethylated		Methylated		Unmethylated		Methylated	
Competitor 50x	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.	Unmethyl.	Methyl.
I	0	0	0	0	X	X	X	X	X	X	X	X	X	X	0	0
I+1	X	X	X	X	X	X	+	++	X	X	X	X	X	X	X	X
J	+	++	+	++	+	++	+	++	+	++	+	++	+	++	+	++
K	+	++	+	++	+	++	+	++	+	++	+	++	+	++	+	++
K+1	X	X	X	X	X	X	X	X	++	++	++	++	X	X	X	X
K+2	X	X	X	X	X	X	X	X	X	X	X	X	++	++	++	++
K+4	++	0	X	X	++	0	X	X	X	X	X	X	X	X	X	X
L	+	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++
L+1	X	X	X	X	X	X	X	X	+	++	+	++	++	++	++	++
L+2	X	X	X	X	++	++	++	++	X	X	X	X	X	X	X	X
L+4	X	X	X	X	X	X	X	X	X	X	0	++	X	X	0	++
L+5	X	X	X	X	X	X	0	++	X	X	X	X	X	X	X	X
L+6	X	X	X	X	+	++	+	++	X	X	X	X	0	++	0	++
L+7	X	X	X	X	X	X	X	X	+	++	+	++	0	++	0	++
M	+	+	+	+	+	+	+	+	X	X	X	X	+	+	+	+
N	++	++	++	++	X	X	X	X	X	X	X	X	++	++	++	++
O	++	++	++	++	++	++	++	++	+	+	+	+	++	++	++	++
O+1	X	X	X	X	++	++	++	++	X	X	X	X	X	X	X	X
P	++	0	++	0	X	X	X	X	X	X	X	X	X	X	X	X
P+1	X	X	X	X	X	X	X	X	++	++	++	++	X	X	X	X

5.3.3.5 Binding position determination of factors that bind the P2 region.

For any particular non-artefact band, the pattern will be that of disappearance of the signal due to competition until the reappearance where the binding site is located and disappearance again while the site is moved successively on the mutated competitor. As for probe 1, a series of mutated competitor sequences containing the original sequence replaced by AAAAAA was used to narrow down the region of the probe that each protein complex was binding to. For the complex bands J,K and L that bind both versions of the probe in all four nuclear lysates, competitors containing mutations located between 9-27bp of the probe had little effect on complex binding when added at 50X access assays (Figure 5. 7) and the binding sequence was determined to be 5'- GAGGCGACGCTCCCG - 3' for J and K for both version of the probe, containing cg06639320+6bp and cg06639320+9bp and cg06639320 (turquoise) whereas, band L has the binding sequence 5'- CGATAACT -3' containing cg06639320. All the bands found contained at least a CpG unit in their binding sequence (Table 5. 4) which can explain how the signal strength can be modulated by DNA methylation.

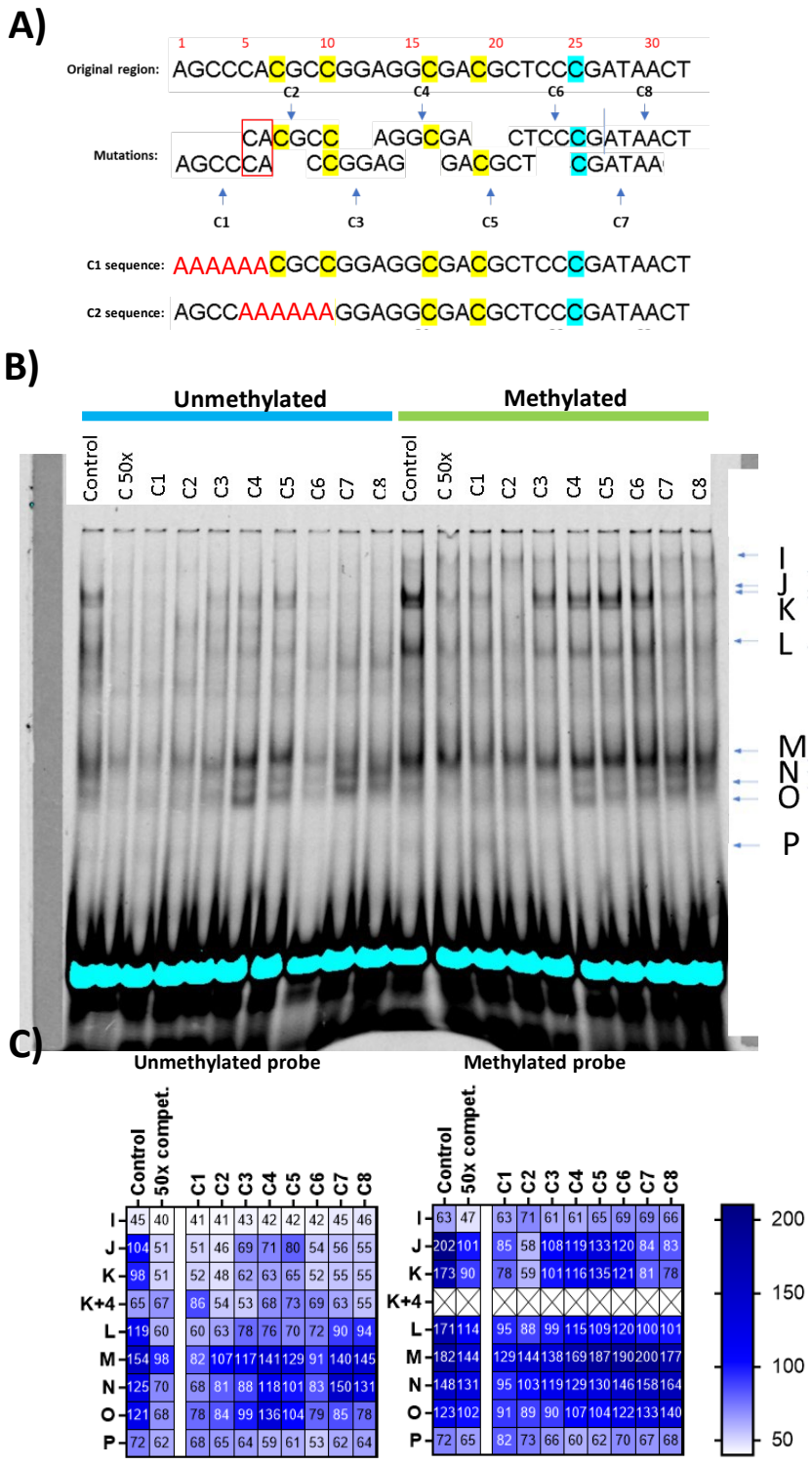


Figure 5. 7. P2 mutated competitor assay. P2 region determination of the binding location of the protein factor bands found in SW1353, a combination of eight competitors were used to determine the binding position. A) Schematic of the original probe sequence and the mutated combinations used for the identification of the binding position. Each mutated competitor had six consecutive bases replaced with adenine, with each combination of

competitor having an overlap of two bases. This gave rise to a total of eight combinations marked by C and a digit represented in the schematic B) Representative EMSA result using the different mutated competitors on the C) Densiometric quantification of the EMSA result to determine the preferred binding site for each identified band in SW1353 nuclear extract.

Table 5. 4. Summary of the findings from the mutated competitor assay on the P2 region. The position of binding site found is indicated for the complex bands studied as well as the determined binding sequence of the complex for each type of probe.

P2 Probe	Unmethylated probe		Methylated probe	
Band name	Position of binding	Determined binding sequence	Position of binding	Determined binding sequence
I	Unknown		x	
J	12-26	GAGGCGACGCTCCCG	12-26	GAGGCGACGCTCCCG
K	12-26	GAGGCGACGCTCCCG	12-26	GAGGCGACGCTCCCG
K+4	12-22	GAGGCGACGCT		
L	20-30	GCTCCCGATAA	20-30	GCTCCCGATAA
M	12-32	GAGGCGACGCTCCCGATAACT	12-32	GAGGCGACGCTCCCGATAACT
N	20-32	GCTCCCGATAACT	20-32	GCTCCCGATAACT
O	20-32	GCTCCCGATAACT	20-32	GCTCCCGATAACT
P	1-10	AGCCCACGCC	1-10	AGCCCACGCC

5.3.3.6 Identification of binding transcription factors and TF consensus competition assay

The prediction tools determined that the factors of E2F, NF-κB, Zfx, CREB, AP2, STAT family, Sp1+3, cMYC, ELK1 potentially bind the P2 region (Figure 5. 8 A) based on consensus TF binding sequence and ChIpseq. Transcription factors that were identified by at least two tools were considered for further analysis, giving a total of 23 factors, only 14 of which are actually expressed in SW1353 cells based on RNAseq data.

Using competitors containing their consensus binding sites (Figure 5. 8 B) in the position at which these proteins are expected to bind, the Sp1+Sp3 competitor decreased the signal of band J, band K and band L for both the methylated and unmethylated competitor with the NF-κB competitor also decreasing the signal of band J, K, L. For the E2F competitor there was a general reduction in the signal of all the bands for both versions of the probe. The AP2 competitor decreased the signal of band O. The competition with the STAT competitor resulted in the disappearance of the band N. All these findings seemed to impact both versions of the probe. No competition of any band was observed for the Zfx, CREB, cMyc and Elk1 competitors for both the unmethylated and methylated probe. A semi-quantitative summary of these findings is represented in Table 5. 5 .

A)

	Frequency of prediction	Probe position	TF Consensus sequence	Sequence 5'-3'
Original sequence				AGCCCACGCCGGAGGCGACGCTCCCGATAACT
Factor				
E2F	2	18-31	NNGCGCGAAANTK	AGCCCACGCCGGAGGCGACTCGCGCGAAATTG
NFKB	2	13-27	NGGGGAMTTTCCNN	AGCCCACGCCGGTGGGGACTTTCCGTATAACT
Zfx	2	18-26	AGGCCTGG	AGCCCACGCCGGAGGCGAGGCCTGGGATAACT
CREB	2	12-24	NNGNTGACGYNN	AGCCCACGCCGTCGCTGACGTACCCGATAACT
AP2	5	1-13	MKCCCSCNGGCG	AGCCCGCTGGCGAGGCGACGCTCCCGATAACT
STAT	2	22-30	TTCCCRKAA	AGCCCACGCCGGAGGCGACGCTTCCCGTAACT
Sp1+3	2	7-18	GGGGGCGGGGG	AGCCCAGGGGGCGGGGGACGCTCCCGATAACT
cMYC	2	2-14	NNNCACGTGNNN	AATGCACGTGCTGGGCGACGCTCCCGATAACT
ELK1	4	5-21	NNNACMGGAAGTNCNN	AGCCTTCACCGGAAGTTCGACTCCCGATAACT

B)

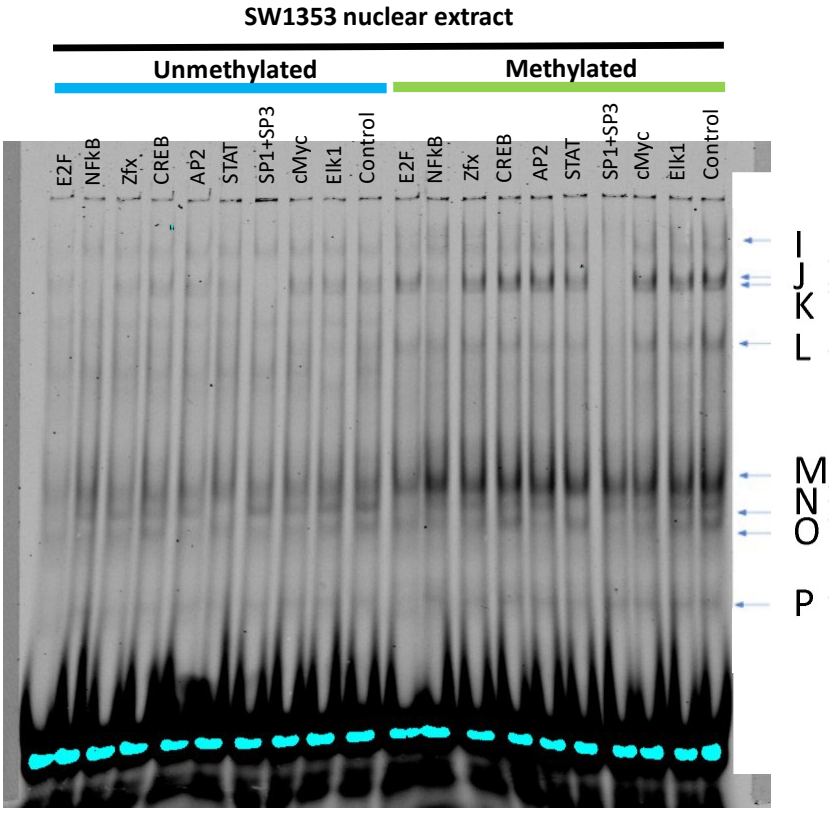


Figure 5. 8. P2 TF consensus competitor assay. Identification of and consensus competitor testing of transcription factors predicted to the P2 region. A) Table summarising the nine TF predicted by at least 2/7 tools the predicted TFs together with the frequency at which the

factor was identified, the binding position on the probe, consensus sequence, and binding sequence on the probe where the consensus sequence was added. Highlighting in turquoise is cg06639320 and in red the other CpG sites present. Red underlined sequence corresponds to the binding position of the transcription factor on the original probe sequence. B) The representative EMSA result for the incubation with the 50x TF consensus competitors.

Table 5. 5. Summary table of the EMSA competition assay with transcription factor consensus competitors on the P2 region. “No” represent that there was no competition to the respective band, “+” (light green)) represents mild competition that occurred with <50% of the original control signal being diminished, “++” (dark green) represents strong competition that occurred with >50% of the original control signal being diminished.

		Unmethylated probe								
Rf	**Letter assigned**	**E2F**	**NFkB**	**Zfx**	**CREB**	**AP2**	**STAT**	**Sp1+3**	**cMyc**	**Elk1**
0.09	I	+	No	No	No	No	No	No	No	No
0.15	J	+	+	No	No	No	No	++	No	No
0.17	K	+	+	No	No	No	No	++	No	No
0.23	K+4	+	No	No	No	No	No	No	No	No
0.27	L	+	+	No	No	No	No	++	No	No
0.53	M	+	No	No	No	No	No	No	No	No
0.57	N	+	No	No	No	No	++	No	No	No
0.6	O	+	No	No	No	+	No	No	No	No
0.74	P	+	No	No	No	No	No	No	No	No
		Methylated probe								
Rf	**Letter assigned**	**E2F**	**NFkB**	**Zfx**	**CREB**	**AP2**	**STAT**	**Sp1+3**	**cMyc**	**Elk1**
0.09	I	+	No	No	No	No	No	No	No	No
0.15	J	+	+	No	No	No	No	++	No	No
0.17	K	+	+	No	No	No	No	++	No	No
0.23	K+4	X	X	X	X	X	X	X	X	X
0.27	L	+	+	No	No	No	No	++	No	No
0.53	M	+	No	No	No	No	No	No	No	No
0.57	N	+	No	No	No	No	++	No	No	No
0.6	O	+	No	No	No	+	No	No	No	No
0.74	P	+	No	No	No	No	No	No	No	No

5.3.3.7 Confirming the identify of protein complexes binding P2 using antibody supershifts

The previously experiment suggested that several of the complexes that bind probe 2 contain the TFs SP1, SP3, E2F, NF-KB and AP2, or proteins with a similar consensus sequence

to these. To investigate this further, antibody supershift assays, whereby an antibody against the transcription factor is added to the EMSA reaction were performed with antibodies again STAT1, STAT2, STAT3, TFAP2 (AP2), SP1, SP3 and NF-KB using SW1353 nuclear lysate. Upon binding to the transcription factor, the band containing the factor is expected to 'shift' upwards in the gel or disappear. Supershifts were only observed Sp1 and Sp3 antibodies in the SW1335 nuclear extract, causing shifting g of band J with the SP1 antibody and bands K and L with the SP3 antibody (Figure 5. 9). This was observed for both the unmethylated and methylated version of the P2 probe. To confirm the identify of bands J, K and L the Sp1 and Sp3 supershifts were repeated using the nuclear extract from TC28a2, U2Os and MDA-MB-231 cells singularly and in combination. (Figure 5. 10). Through this result it confirmed that complex J had a supershift reaction with the SP1 antibody whereas the K and L complexes had a supershift with SP3, this indicating that these complexes contain SP1 and Sp3 respectively.

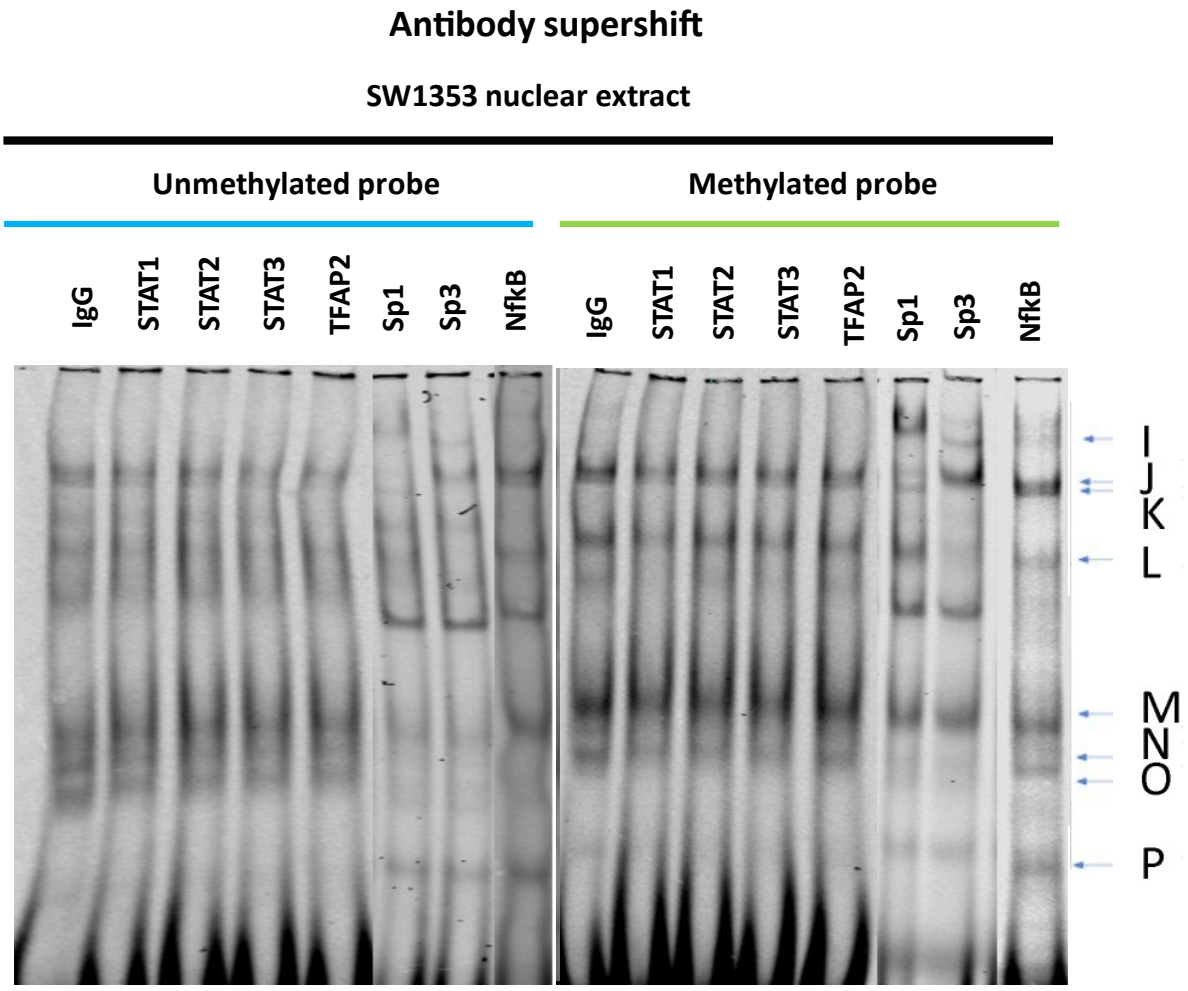


Figure 5. 9. P2 antibody supershift reaction. Supershift reaction was performed using control IgG, STAT1, STAT2, STAT3, Sp1, Sp3 and NF-κB (RelA) antibodies for both the unmethylated probe and methylated probe in SW1353 nuclear extracts.

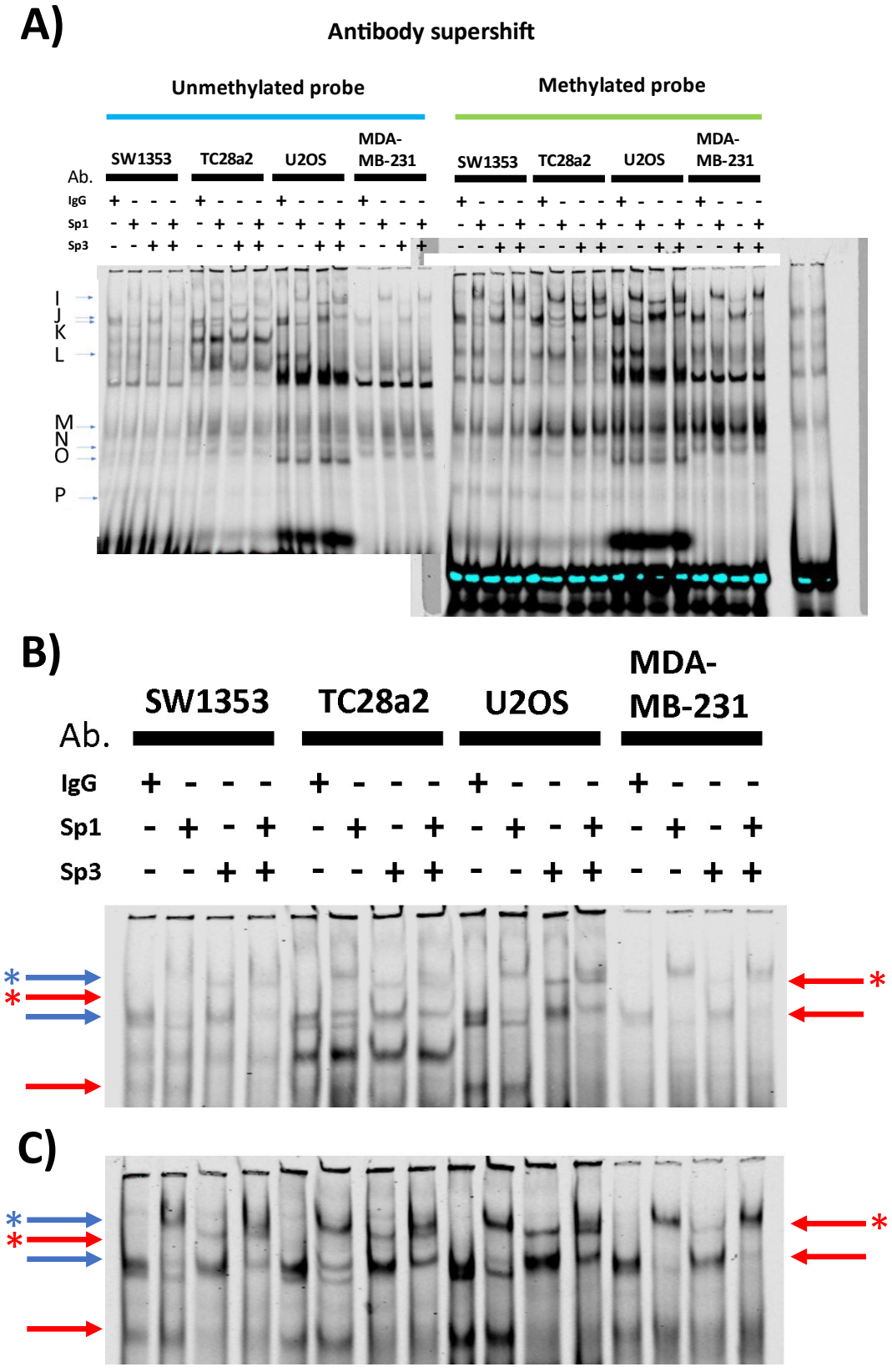

A)
Antibody supershift
Unmethylated probe
Methylated probe
SW1353
TC28a2
U2OS
MDA-MB-231
Ab.
IgG
Sp1
Sp3
I
J
K
L
M
N
O
P
B)
C)

Figure 5. 10. P2 region probe Sp1 and Sp3 antibody supershift using different cell nuclear extracts with the unmethylated and methylated probe. A) Representative image of the EMSA supershift with the Sp1 and Sp3 antibodies. Magnified imaged of the gel result from A) are B) Unmethylated probe supershift C) Methylated probe supershift. Blue arrow marks the Sp1 band and * marks the band that appears after the supershift reaction. Red arrows mark the Sp3 bands and * marks the band that appears after the supershift reaction.

5.3.3.8 Optical densitometry analysis of impact of DNA methylation

For the Sp1 and Sp3 containing complex bands J, K, L, optical densiometry was used to study compare their relative strength of protein binding between the methylated and unmethylated probe (Figure 5. 11). DNA methylation was seen to significantly increase the signal of Sp1 by 1.73 fold (p= 0.000053) (n=6) in SW1353, 1.75 fold (p= 0.009) in TC28a2 (n=3), 1.88 fold (p= 0.007) in U2OS (n=3) and 1.38 fold (p= 0.029) in MDA-MB-231 (n=3) paired ratio t-test analysis. For the signal corresponding to Sp3 (band K) the signal increased was observed to be by 1.80 fold (p= 0.0008) in SW1353 (n=6), 1.99 fold (p= 0.0078) in TC28a2 (n=3),1.89 fold (p= 0.0240) in U2OS (n=3) and 2.35 fold (p= 0.0308) in MDA-MB-231 (n=3) paired t-test analysis. Further to this, for the signal corresponding to Sp3 (band L) the signal increased was observed to be by 1.69 fold (p=0.0004) in SW1353 (n=6), 1.77 fold (p= 0.0146) in TC28a2 (n=3), 2.31 fold (p= 0.0394) in U2OS (n=3) and 1.92 (p= 0.0037) fold in MDA-MB-231 (n=3) paired ratio t-test analysis. These results indicating that the binding to methylated DNA is much stronger for the complexes studied across the cell lines studied.

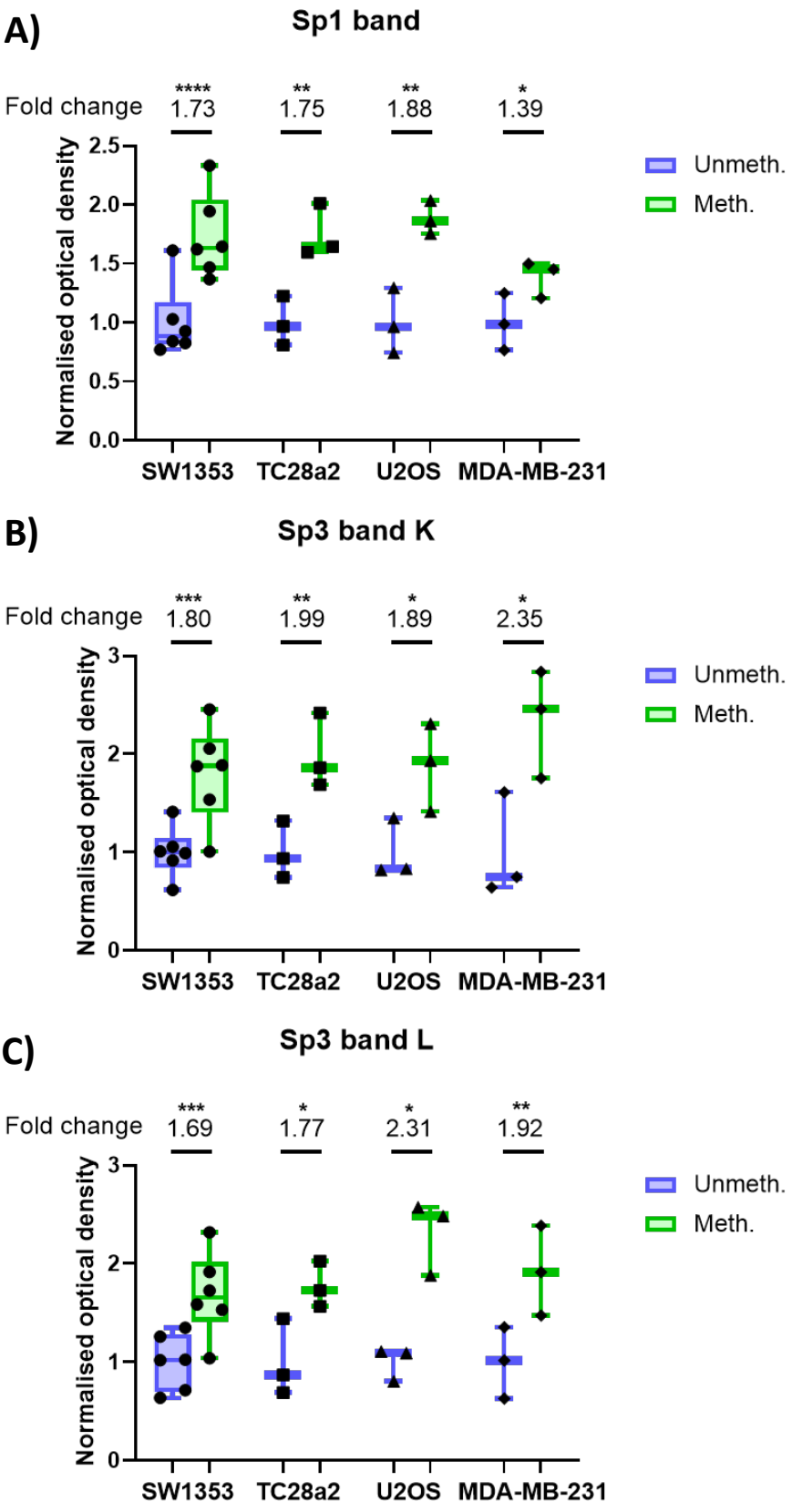


Figure 5. 11. Impact of the status of methylation of the probe on the complex signal strength. Signal strength was determined for the Sp1 complex (A), Sp3 complex band K (B)

and Sp3 band L (C) bands quantified by densiometric analysis in different cell nuclear extracts. Normalized by the mean of the unmethylated probe for each cell type nuclear extract. Individual points represents data from one replicate and the average fold change in presented for the group. Paired ratio t-test statistical analysis, $p<0.05$(*), $p<0.01$(**), $p<0.001$(**), $p<0.0001$(****).

5.3.3.9 Competition assay using un-methylated and methylated probes and binding kinetics

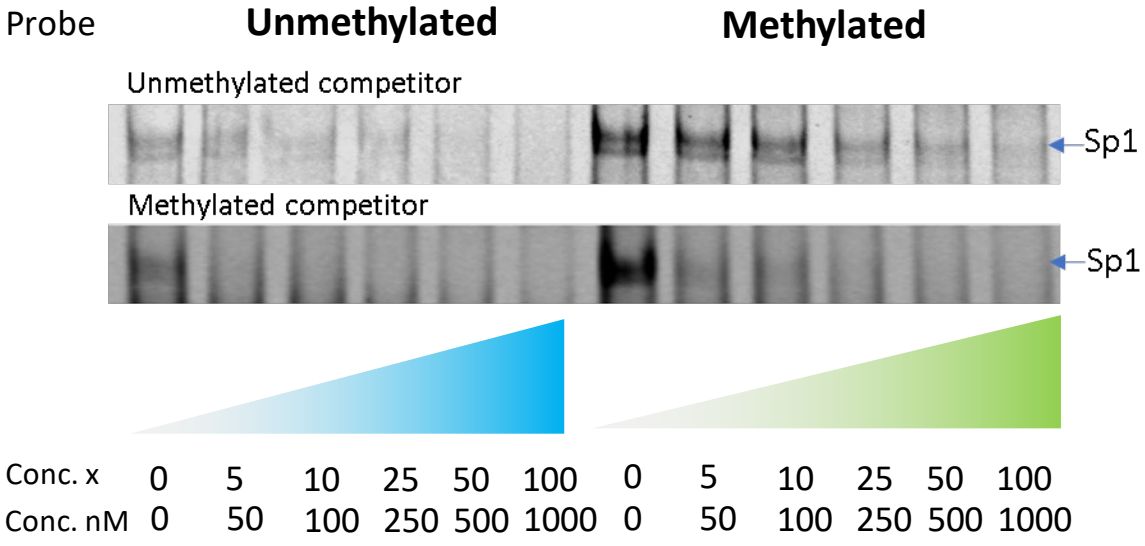
Based on the previous observation that the complexes for Sp1 and Sp3 display a higher binding strength for methylated DNA, a competition assay was performed using unmethylated and methylated non-tagged oligomers competitors to have further evidence for the difference in affinity between the probes (Table 2. 8). The incubation was done using an increasing gradient fold excess at 5x, 10x, 25x, 50x and 100x of competitor in order to investigate the response across a wide range of concentrations and to be able to infer binding kinetics for the two competitor versions. Densiometric quantification of the signal from the EMSA gel permitted the analysis of the decrease in signal associated with the different concentrations of competitor and the determination the binding dissociation constant (Kd), this corresponding to the concentration of competitor required to decrease the control signal by 50%. The analysis that is presented is for the bands of whose identity was revealed by the supershift experiment, these being Sp1 (band J) and Sp3 (band K and band L).

For the Sp1 band complex J (Figure 5. 12) properties starting with the unmethylated probe there is a much lower Kd value at 12 nM (4-23 nM C.I. 95%) when the methylated competitor is used compared to 103 nM (19-554 nM C.I. 95%) of the unmethylated competitor. A similar relationship is observed for the methylated probe where for the methylated competitor there is a Kd with a value of 21 nM (11-34 nM C.I. 95%) compared to a Kd of 281 nM (98-1665 nM C.I. 95%) for the unmethylated competitor. Overall, the impact of the methylated competitor is of a stronger affinity for competition compared to the unmethylated version, with the statistical analysis being presented in figure C). For the Sp3

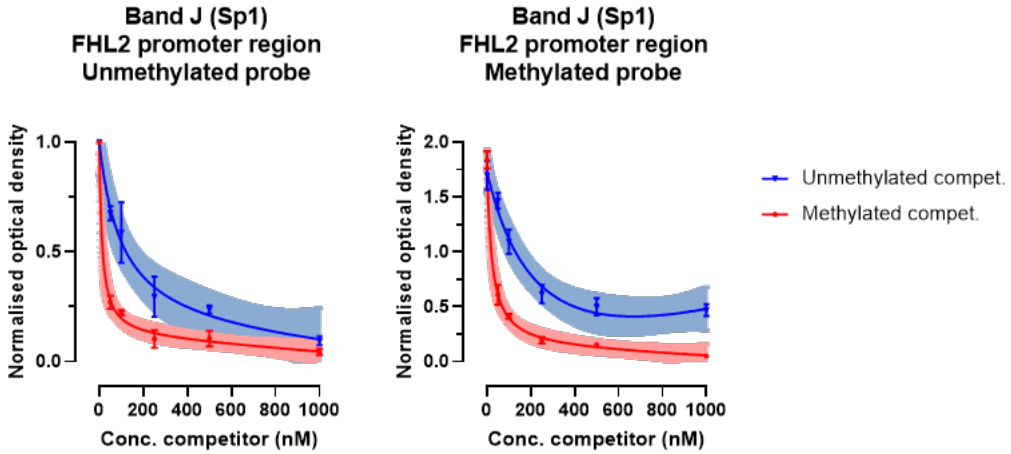
complex (band K) properties are presented in Figure 5. 13. Starting with the unmethylated probe there is a much lower Kd value at 21 nM (12-32 nM C.I. 95%) when the methylated competitor is used compared to 122 nM (40-427 nM C.I. 95%) of the unmethylated competitor. A similar relationship is observed for the methylated probe where for the methylated competitor there is a Kd with a value of 17 nM (8-29 nM C.I. 95%) compared to a Kd of 3305 nM (106-?? nM C.I. 95%) for the unmethylated competitor.Overall, the impact of the methylated competitor is of a stronger affinity for competition compared to the unmethylated version, with the statistical analysis being presented in figure C). For the Sp3 complex (band L) properties (Figure 5. 14), starting with the unmethylated probe there is a much lower Kd value at 59 nM (0-388 nM C.I. 95%) when the methylated competitor is used compared to 107 nM (0-?? nM C.I. 95%) of the unmethylated competitor. A similar relationship is observed for the methylated probe where for the methylated competitor there is a Kd with a value of 69 nM (23-181 nM C.I. 95%) compared to a Kd of 317 nM (73-?? nM C.I. 95%) for the unmethylated competitor. Overall, the impact of the methylated competitor is of a stronger affinity for competition compared to the unmethylated version, with the statistical analysis being presented in figure C).

Consistently across band J, K, L a statistically significant relationship was observed where the methylated version reduced the signal by a much higher magnitude than the unmethylated competitor at most of the concentrations of competitor used. The EMSA analysis for probe 2 has identified the transcription factor Sp1 as being part of complex J where mutation competition assays indicate that the binding occurs at bases 12-26 containing the age-CpGs of cg06639320, cg06639320+6bp, cg06639320+9bp and additionally DNA methylation increases the binding affinity of this complex. Another identified transcription factor was Sp3, and this was observed to be part of complex K and complex L. Mutation competition assays indicate that band K binds at the 12-26 position on the probe this containing the age-CpGs of cg06639320, cg06639320+6bp, cg06639320+9bp with competition assays indicating that DNA methylation increases the binding strength of this complex as observe for band J. For band L the position of binding was determine to be at position 20-30 on the probe where the only age-CpG present at this region is cg06639320 and in addition the complex was observed to have stronger binding when the CpGs were methylated rather than the unmethylated condition.

A)



B)



C)

Probe	Unmethylated		Methylated	
Competitor	Unmeth.	Meth.	Unmeth.	Meth.
Kd (nM)	103	12	281	21
95% C.I.	19 to 554	4 to 23	98 to 1665	11 to 34
R squared	0.90	0.98	0.92	0.99

D)

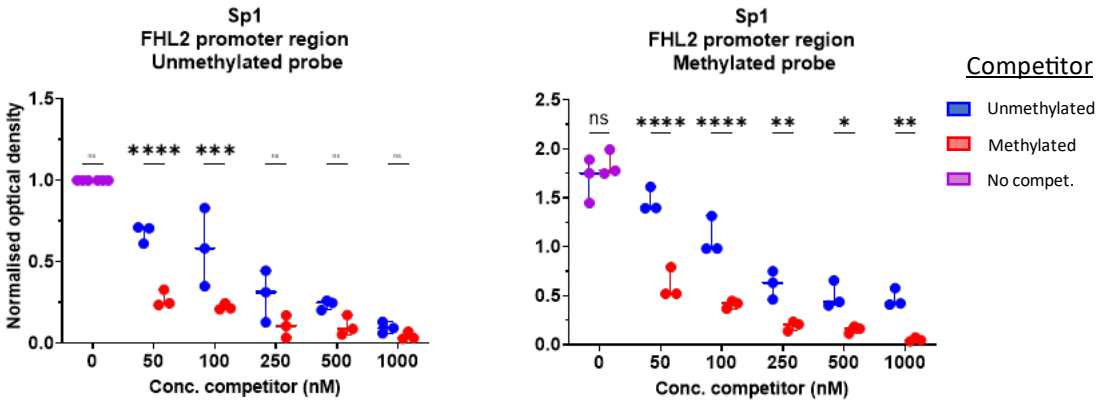
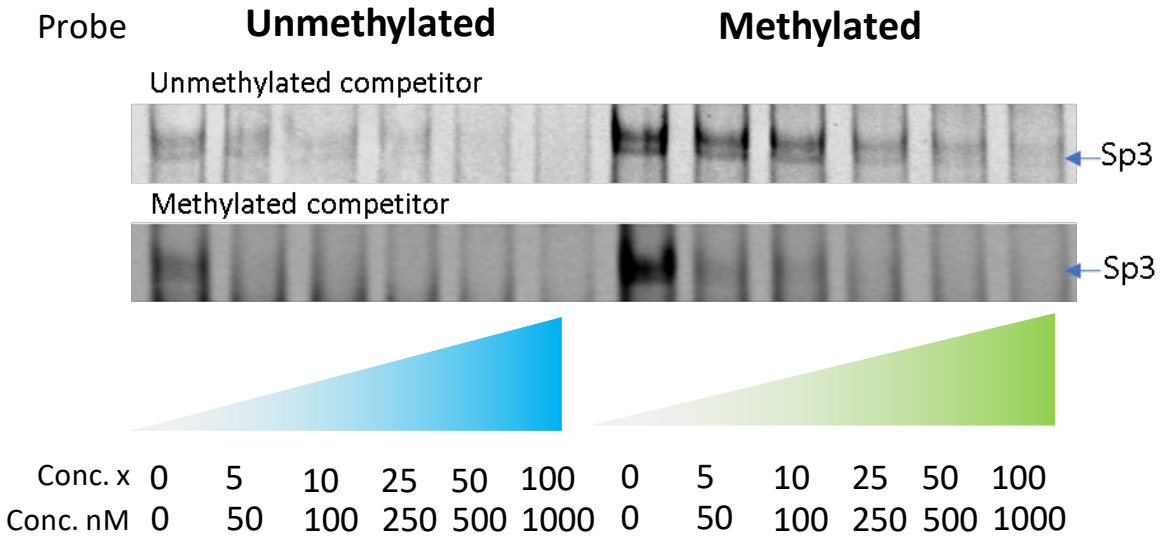
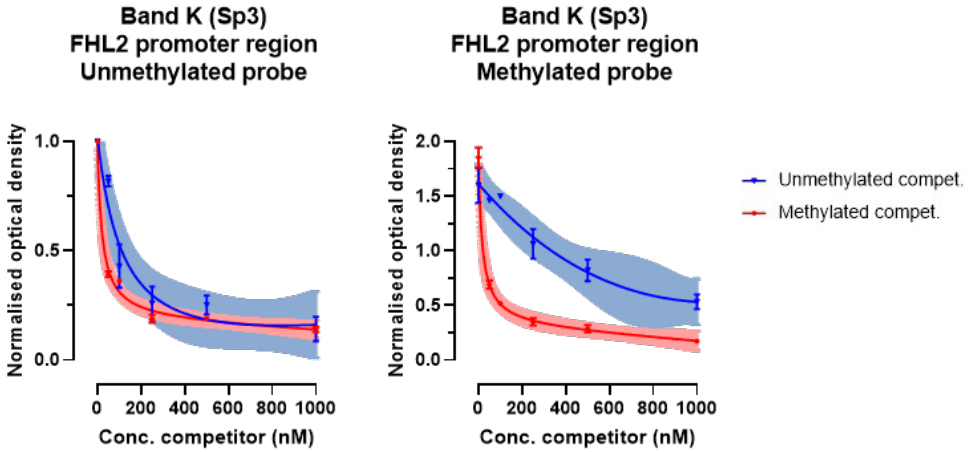


Figure 5. 12. Sp1 complex band J EMSA analysis. A) Representative image of the impact of competition with unmethylated and methylated competitors on the unmethylated and methylated probes on the Sp1 band (band J). B) Densiometric quantification of the signal from the competition assay and One site – Total binding kinetics analysis for the combination of the probes and competitors mentioned. C) Tabular results of the One site – Total kinetics analysis. D) Statistical analysis using Two-way ANOVA using the Sidak correction for multiple comparisons of the impact of the two different competitor types, $p<0.05$(*), $p<0.01$(**), $p<0.001$(**), $p<0.0001$(****).

A)



B)



C)

Probe	Unmethylated		Methylated	
Competitor	Unmeth.	Meth.	Unmeth.	Meth.
Kd	122	21	3305	17
95% C.I.	40 to 427	12 to 32	106 to ??	8 to 29
R squared	0.89	0.99	0.87	0.99

D)

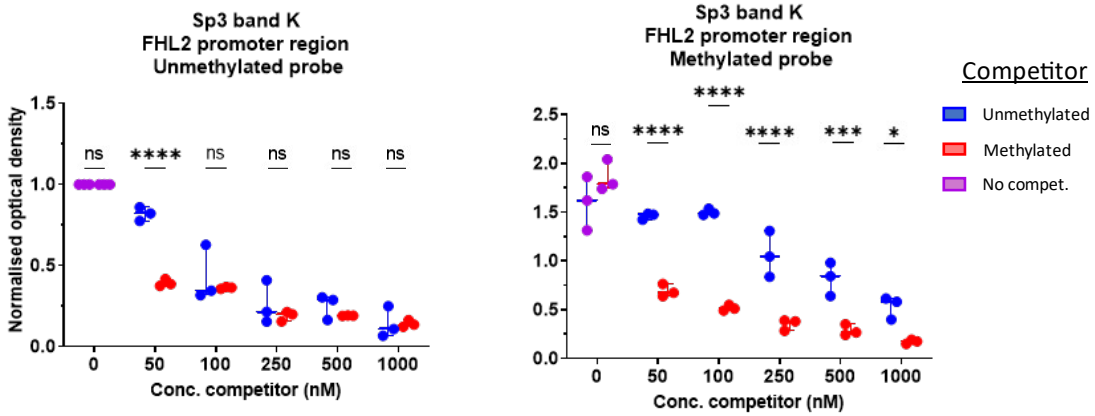
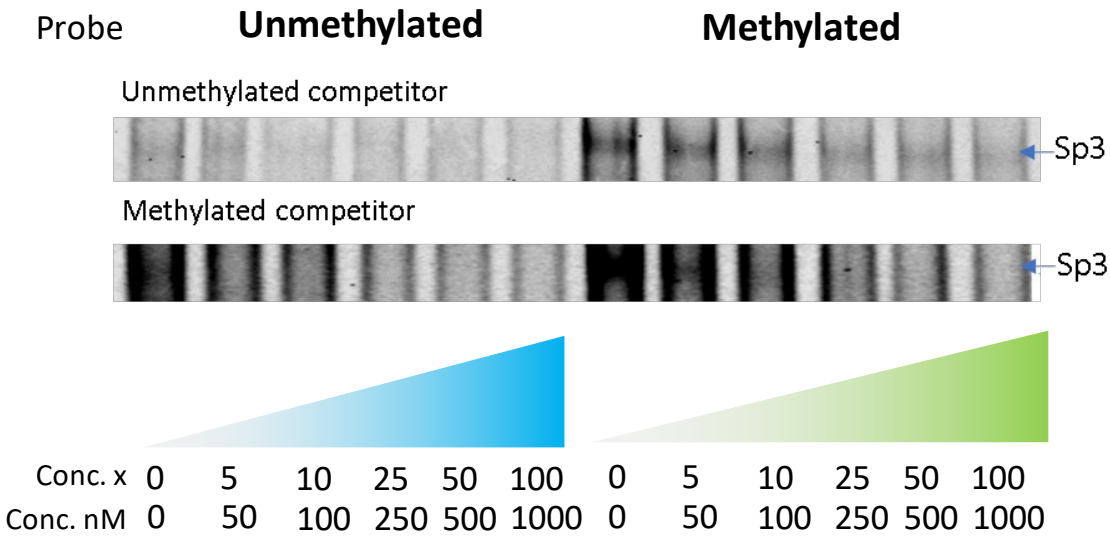
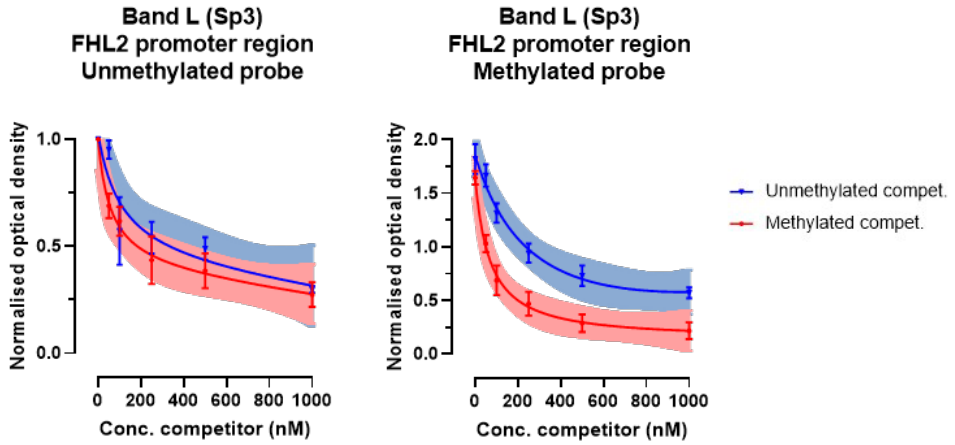


Figure 5. 13. Sp3 complex band K EMSA analysis. A) Representative image of the impact of competition with unmethylated and methylated competitors on the unmethylated and methylated probes on the Sp3 band (band K). B) Densiometric quantification of the signal from the competition assay and One site – Total binding kinetics analysis for the combination of the probes and competitors mentioned. C) Tabular results of the One site – Total kinetics analysis. D) Statistical analysis using Two-way ANOVA using the Sidak correction for multiple comparisons of the impact of the two different competitor types, $p<0.05$(*), $p<0.01$(**), $p<0.001$(**), $p<0.0001$(****).

A)



B)



C)

Probe	Unmethylated		Methylated	
Competitor	**Unmeth.**	**Meth.**	**Unmeth.**	**Meth.**
Kd	107	59	317	69
95% C.I.	0 to ??	0 to 388	73 to ??	23 to 181
R squared	0.75	0.85	0.91	0.93

D)

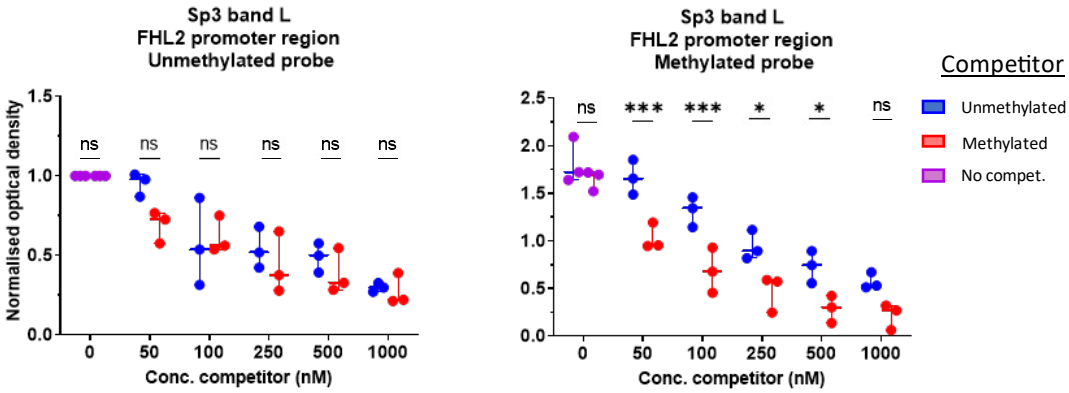


Figure 5. 14. Sp3 complex band L EMSA analysis. A) Representative image of the impact of competition with unmethylated and methylated competitors on the unmethylated and methylated probes on the Sp3 band (band L). B) Densiometric quantification of the signal from the competition assay and One site – Total binding kinetics analysis for the combination of the probes and competitors mentioned. C) Tabular results of the One site – Total kinetics analysis. D) Statistical analysis using Two-way ANOVA using the Sidak correction for multiple comparisons of the impact of the two different competitor types, $p<0.05$(*), $p<0.01$(**), $p<0.001$(**), $p<0.0001$(****).

5.4 Discussion

The impact of DNA methylation on the binding of transcription factors to the age-CpG region of the *FHL2* promoter was investigated through EMSA. The EMSA technique has several advantages over mass spectrometry and some of the disadvantages can be mitigated through the use of densiometric quantification and migration distance quantification through the retention factor scaling method. Due to technical aspects the *FHL2* promoter region with age-CpGs was divided in two probes that corresponded to the two DNA methylation correlation clusters identified in Chapter 3. The quantification in turn has several advantages such as increase in resolution as well as permitting the study of binding sequences of factors as well as reducing the overall complexity of the analysis and quantification. Unfortunately, most publication do not take full advantage of all these innovations in technology and the ability for quantification of the signal instead opting to just visual inspection. Interpretation restricted to visual inspection can not only introduce human associated error but is limited in signal strength assessment to only a semi-quantitative level. Another issue associated with the EMSA technique can be poor reproducibility in terms of migration pattern identification due to the inherent issues of gel electrophoresis such as gel heating and uneven cathode electrode wire arrangement. To overcome this, a method of migration distance standardization is introduced based on a Retention factor (Rf) system. This not only serves in increasing the reproducibility when paired with optical densiometric quantification but serves as a key method of protein complex identification based on the intrinsic property of protein-DNA complex migration.

For the P1 region this containing the following CpGs: cg22454769, cg24079702, cg24079702+2bp, cg24079702+9bp and cg24079702+13bp. The experiments focused on the use of nuclear protein extracts from multiple cell lines of different lineages where a total of 14 complexes were observed to bind to the p1 probe sequence where the probe with the methylated CpG sites having much stronger band signals for all the complexes as well as more complexes being observed for the methylated probe. Unique complexes were identified this being evidence that the region can be under different regulation based on the presence of different cell specific transcription factors and also complexes that were in common across the cell lines tested with these indicating that the region has also regulators

that are non-cell line specific. For some of the bands there is a maximum of signal that sometimes exceeds the initial control levels, a possible explanation for this kind of effect could be that under these conditions where only the site of binding is available while competition still occurs at other sites this could indicate that there are other protein factors that could inhibit binding in normal conditions. Without these proteins that would interfere normally the effect observed is of greater binding to the probe of some proteins. Such an observation indicates that there must be other factors present at the P1 region that inhibits the binding associated with the band F complex and that this can be competed in order to support the binding. Based on the web prediction tools available, possible transcription factors that can bind the P1 region these factors can be Zfx, NF-κB, TP53, Sp family, Myc, AP2, ELK1, E2F family and p300 however, none of the bands seem to change with competition from any of the transcription factor competitor. This could either be due to incorrect prediction from the web prediction tools or some other unknown factors impacting this competition. Unfortunately, this leaves the P1 region with no identified factors with future investigations to focus on the use of a different approach to determine the complexes such as through western blotting (Waby *et al.*, 2010) or mass-spectrometry (Stead *et al.*, 2006).

For the P2 region this contained the following CpGs: cg06639320, cg06639320+6bp, cg06639320+9bp, cg06639320+14bp, cg24079702-14 bp. The region presented a total of 20 complexes across all the cell line nuclear extracts used with more transcription factor complexes associated with the methylated version of the probe with these also showing higher signal strength. Methyl binding domain transcription factors are known to bind only when the DNA is methylated at certain CpG site from their consensus sequence (Héberlé and Bardet, 2019a) and these factors discovered to be present only with the methylated probe may well be in the family of methyl binding domain TFs that regulate the region of interest. These factors seemed to be cell line specific in some instances, informing that different regulation can be done on the methylated DNA. From the banding pattern observed non-cell line specific transcription factors are also present these being complex band J, band K, band L and band O. Since in previous chapters, DNA methylation changes were linked with the changes in *FHL2* expression then the findings from this chapter could

represent a potential intermediary mechanism through which this link can occur. Considering that these factors have a higher affinity for methylated DNA this informs that the age-associated changes could result in increased binding of these factors resulting in potential gene expression modulation.

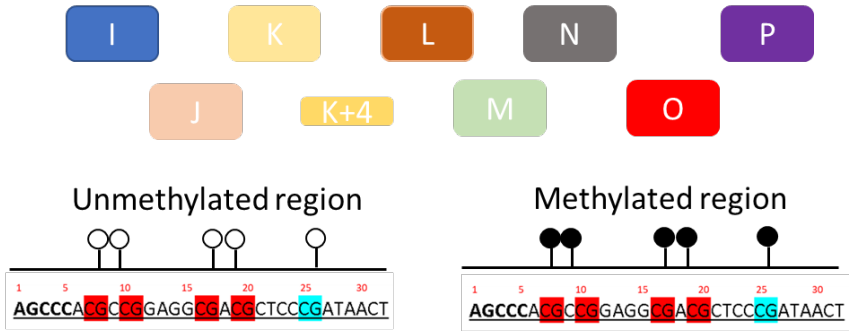
Based on the web prediction tools available, possible transcription factors that can bind the P2 region range from E2F family, NF-κB, Zfx, CREB, AP2, STAT family, Sp family, cMyc and Elk1. The result of the TF consensus competition assay for the P2 region indicates that only the competitors for the Sp family, STAT family, NF-κB and E2F family had an effect on the banding pattern compared to the control condition. The band J, K and L were impacted by the Sp family competitor as well as by the NF-κB competitor; Band O was impacted by the STAT family competitor and band P by the AP2 competitor. Interestingly, the E2F competitor had an effect on all the bands present (Figure 5. 8). The role of factors such as the Sp & Krüppel-like factor family, their role has been linked with the functionality mainly of regulating cell cycle associated genes, the formation transcription initiation complexes, histone modifying enzyme complexes and chromatin remodelling complexes (Li *et al.*, 2004; Li and Davie, 2010). DNA methylation seems to have varying effects on the binding of these factors based on the underlying binding sequence, these range from inhibition by DNA methylation (Reynard *et al.*, 2014; Tian *et al.*, 2015) to increased binding with DNA methylation (Boumber *et al.*, 2008). The STAT family of transcription factors is linked with the JAK family of proteins forming the JAK-STAT signalling pathway, this is associated with response to growth and inflammatory cytokines resulting in gene expression regulation and when ectopic a link with carcinogenesis (Clevenger, 2004; O'Shea *et al.*, 2015; Loh *et al.*, 2019). For AP2 (TFAP2A) its known role is in renal cellular differentiation (Chambers *et al.*, 2019) and carcinogenetic or suppressive activity in various tissues (Kołat *et al.*, 2019). NF-κB plays a role in modulating cell survival and inflammation responses, proliferation and carcinogenesis (Hayden and Ghosh, 2008) with DNA methylation also affecting the binding of this factors with high dependency on the underlying sequence (Wang *et al.*, 2017). E2F family of transcription factors are thought to be key regulators in cell cycle progression (Attwooll *et al.*, 2004) with some members of the family (e.g. E2F3) showing preferential binding to methylated DNA (Campanero *et al.*, 2000).

Taking forward the results from the P2 TF competitors, antibody supershift assays were then attempted in order to validate the results from the TF competitor assay. The antibodies tested were IgG (Control), STAT1, STAT2, STAT3, TFAP2, Sp1, Sp3, and NF-κB with only Sp1 and Sp3 antibodies producing the expected supershift. This confirms that the identity of at least one of the factors present in the complex for band J is Sp1, for band K is Sp3 as well as band L as Sp3. Band K and L being different complexes that show different binding sequences, yet both contain the Sp3 transcription factor. A possibly reason for why the NF-κB antibody did not seem to produce a supershift can be due to the nature of the consensus found for NF-κB competitor, this is a CG rich consensus being similar to the Sp family competitor and it could just act as a competitor for SP family factors without involving the NF-κB multi subunit family. A similar explanation why none of the STAT antibodies tested produced a supershift; the consensus sequence for the STAT family can be of a different unknown factor. Finally, another explanation is a technical issue where the antibodies not always targeting the intended target. For Sp1 and Sp3 a positive EMSA control was obtained from Dr. Reynard that was tested and validated in a previous publication (Reynard *et al.*, 2014). Unfortunately, no such positive control was found for the other antibodies to confirm that indeed they can perform the supershift assay. Testing the other cell nuclear extracts with the Sp1 and Sp3 antibodies revealed that as identified by their Rf values, bands identified as being possibly the same as J, K and L in the other nuclear extracts showed the same supershift effect with the respective antibodies. This confirms that these factors are present in all the nuclear extracts tested and that they can form similar transcription factors complexes. Following the identification of Sp1 in one complex and Sp3 in two complexes for the P2 region, quantification of the band strength is used to determine the overall affinity difference between the unmethylated probe and methylated probe in all the cell type nuclear extracts. DNA methylation was seen to increase the signal of the Sp1 and Sp3 complexes by more than 1.5 fold in all the nuclear extracts.

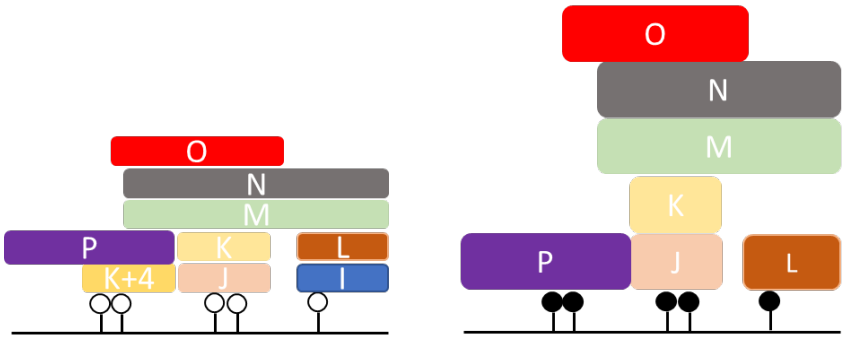
To further test this increase in affinity, competitors that themselves can be either unmethylated or methylated and with no fluorophore tag are used together with densiometric analysis of the signal. Another advantage of this type of quantification is the possibility of applying statistical analysis to determine the significance of the decrease of signal associated with competition which is not commonly observed in the analysis of EMSA

results. Using different concentrations of the competitor the kinetics can be modelled which in return can generate a dissociation constant value (Kd) for each combination indicated based on their binding properties. The Kd value informs of the overall competition ability for the competitors where lower Kd values indicate higher affinity (Salahudeen and Nishtala, 2017). For the Sp1 complex starting with the unmethylated probe there is a much lower Kd value when the methylated competitor is used compared to the unmethylated competitor. A similar relationship is observed for the methylated probe where for the methylated competitor there is a lower Kd compared to the unmethylated competitor. This validating the previous observation that the Sp1 complex has a higher affinity DNA methylation. For the Sp3 complex band K and L the unmethylated probe with the methylated competitor had smaller Kd value in comparison to the unmethylated competitor with a similar relationship being observed for the methylated probe with the methylated competitor compared to a higher Kd of the unmethylated competitor. This again validating the previous results that the Sp3 complexes have increased affinity for methylated DNA. A summary of the overall findings for the P2 sequence is shown Figure 5. 15.

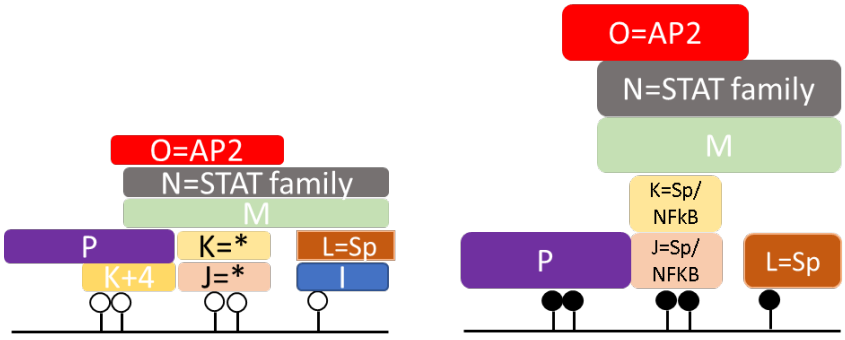
EMSA for P2 region in SW1353



Mutation competition assay



TF competition assay



Antibody supershift assay

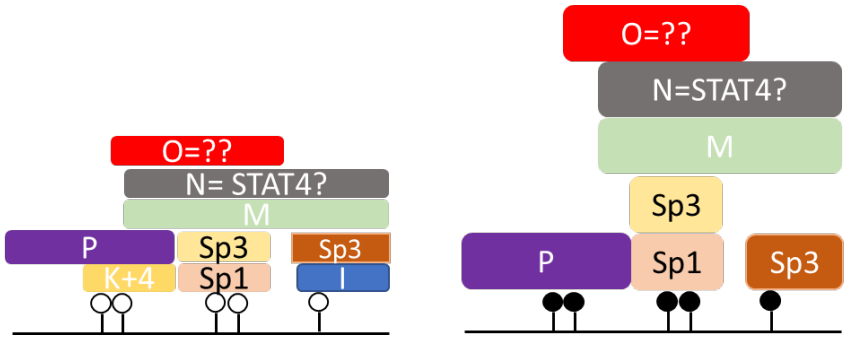


Figure 5. 15. Summary of the EMSA experimental findings for the P2 region.*= Sp1+3/ NF-κB. Thicknes size of the bands indicates the overal strength of binding of the complexes.

The Sp1 and Sp3 are mammalian transcription factors belonging to a family of transcription factors that can either have enhancing or repressive abilities and are generally implicated in controlling cell proliferation genes. Both of these factors are associated with processes such as transcription initiation complex recruitment and histone modifying enzyme recruitment and chromatin remodelling mechanisms (Li and Davie, 2010). The presence of these factors at a promoter region that changes with ageing can result in different cellular signalling associated with the role of FHL2. Interestingly Sp1 has been previously identified to potentially regulate *FHL2* expression however in a different position of the promoter that is located approx. 400 bp upstream of the region with age-CpGs (Guo *et al.*, 2010). This region was found to bind Sp1 and moreover it provided evidence that the absence of Sp1 using a siRNA knock down method resulted in downregulation of *FHL2* expression at both gene and protein level. Interestingly some links have been found between Sp1 activity and ageing, where Sp1 has been previously identified to be involved with ageing dependent nuclear-cytoplasmic protein trafficking impacting cellular senescence (SY *et al.*, 2012). Sp1 and Sp3 binding has been previously reported to be influenced by DNA methylation based on what CpG dinucleotides from the consensus sequence are methylated, these findings indicated that DNA methylation could either increase (Wang *et al.*, 2003), decrease(Aoyama *et al.*, 2004; Reamon-Buettner and Borlak, 2008; Reynard *et al.*, 2014; Tian *et al.*, 2015) or have no impact (Zhu *et al.*, 2003; Honda *et al.*, 2006) on binding. This mechanism could be also modulated on the availability and diversity of the transcription factors giving rise to the different sometimes opposing effects observed for DNA methylation.

The potential identification that NF-κB could be contained in the complexes of J,K and L ,based on the TF consensus competitor experiment was not confirmed by the supershift assay where this could be due to the antibody used where it could only target only certain isoforms of RelA (Spinelli *et al.*, 2021) that are not the ones participating in the complexes or that other subunits of NF-κB are present at these complexes (Oeckinghaus and Ghosh, 2009). Another explanation is that due to the similarity of the found consensus for NF-κB to the consensus of Sp1+Sp3, it could be that this acted as a competitor for the determined Sp1+Sp3 and not an indication that NF-κB is involved. However there is evidence of such

complexes being previously described in other works, for Sp1 and NF-κB having been previously associated to participate in complexing to induce gene expression changes (Perkins *et al.*, 1993; Perkins *et al.*, 1994) as well as for Sp3 and NF-κB (Carver *et al.*, 2013).This could potential indicate that FHL2 expression could be a target of these mechanisms that seem to be influence by the CpG methylation state.

Unfortunately, the identity of the cell line specific transcription factors could not be confirmed however the binding sequence of these is know which gives rise to a potential list of factors that can be tested in the future for further identification. The region of interest of the *FHL2* promoter could therefore be under more regulation once DNA methylation establishes at the region. Since in the bone tissue DNA methylation is associated with increase in *FHL2* expression while in tissues such as cartilage and synovium DNA methylation is associated with a decrease in *FHL2* expression, then DNA methylation changes could also activate cell type specific transcription factors in order to modulate this genes expression change and therefore give rise to different results in cell signalling. Therefore, the identity of these factors is important to the understanding of the different impacts on *FHL2* gene expression the methylated DNA can have. Another way to validate that these factors indeed bind is through the use of cell knock downs for a particular factor that is believed to bind and then to test the nuclear extract for the complex in question, however factors that are fundamental for normal cell functionality might be hard to produce due to the cell death that can result. To avoid this issue another method would be to do the crosslinking of DNA and protein complexes and then co-immunoprecipitation for the factors in question followed by qRTPCR or sequencing to see whether the region here investigated can be detected.

Understanding these findings in the context of age-associated DNA methylation changes at the promoter of *FHL2* informs of a serial and successive shift in the population of transcription factors that regulate potentially the activity of *FHL2* from a younger individual that has a hypomethylated region to an older with a hypermethylated region. This would indicate that the whole region of interest's functionality can be changed by the DNA methylation changes observed with ageing, this mechanism could therefore be an important mediator between the associated changes and correlation between the expression of *FHL2* and the DNA methylation changes with ageing. With Sp1 and Sp3 being

common factors that can be impacted slowly by ageing through increased binding and moreover through small changes to the cellular pathways.

Chapter 6: What are the functions of *FHL2* and what are the consequences of expression changes?

6.1 Introduction

From the previous chapter, *FHL2* has been revealed as a gene linked with the process of ageing: age-associated DNA methylation changes at its promoter correlate with gene expression in some tissues. DNA methylation was observed to influence the transcription factors as well as their binding affinity, potentially linked with the effect on transcription. To further understand the consequences of altered expression of this gene, an understanding of its function is necessary.

The function of *FHL2* is still being investigated as this appears dependent on its activation state based on phosphorylation as well as cellular location. Discovered functions include as a transcription co-activator (Morlon and Sassone-Corsi, 2003), stabilizer of transcription factor complexes (Tran *et al.*, 2016) and a complex stabilizer for focal adhesion (Park *et al.*, 2008). FHL2 was found to be a key member during the development of digits and a regulator of fibrogenic differentiation in the limb. Its activity is in mesodermal progenitors is in influencing the expression of SCX, TGIF1 and Tenomodulin, with itself being positively regulated by modulators such as TGFB, retinoic acid and Wnt signalling as well as being negatively regulated by the BMP pathway (Lorda-Diez *et al.*, 2018).

FHL2 is mainly cytoplasmic (Morlon and Sassone-Corsi, 2003; Verset *et al.*, 2013; Nakazawa *et al.*, 2016) with transitions in the nucleus being reported (Verset *et al.*, 2013). In the cytoplasm FHL2 has been found to be located at the periphery of focal adhesion points in a wide range of cells of different tissues (Hua *et al.*, 2016; Nakazawa *et al.*, 2016). FHL2 is seen to interact with actin filaments at these sites being involved in the assembly of the focal adhesion complex as well as in reinforcing these by accumulation during periods of mechanical stress on the cytoskeleton (Sun *et al.*, 2020). The cytoplasmic state of FHL2 is seen as an un-phosphorylated version of the protein however in order to act as a mechanical transducer as well as in cellular signalling the protein is activated by phosphorylation, followed by nuclear translocation. Factors influencing translocation are of chemical and physical. Physical factors include cellular substrate rigidity impacting the focal

adhesion mechanical receptors and mechanochemical signals that lead to changes in gene expression to accommodate cellular spatial position changes in the 3D architecture or substrate changes of the extra cellular matrix. These mechanisms are critical in normal cellular functions for pathways such as cellular differentiation and growth (Engler *et al.*, 2006; Cui *et al.*, 2015). The chemical factors include signalling from the Ras pathway (Labalette *et al.*, 2008a), drugs which destabilize the actin filaments and focal adhesion points e.g. Wiskostatin (Nakazawa *et al.*, 2016) and possibly IL1 mediated signalling (Joos *et al.*, 2008) and other pro-inflammatory cytokines and chemokines (Wixler, 2019).

Nuclear translocation of FHL2 is started by phosphorylation of tyrosine residues catalysed by the focal adhesion active kinases (FAK) (Gabriel *et al.*, 2004), however it is not currently known what mechanism is implicated in bringing FHL2 to the nucleus. A difficulty in determining this is due to the fact that FHL2 does not have a normal nuclear localisation signal. At least two distinct pathways have been found, passive diffusion through the nuclear membrane pores (Nakazawa *et al.*, 2016) and a CRM1- dependent (i.e. nuclear exporting protein) nuclear transport mechanism (Morlon and Sassone-Corsi, 2003). As phosphorylated FHL2 migrates from the cytoplasm through the nuclear membrane, once in the nucleus it was found to be involved in a series of functions involving the transcription machinery. FHL2 does not possess a DNA binding domain and therefore no direct binding to DNA is observed, however the role that it has in gene transcription is linked with the ability to participate in transcription factor complexes (Kleiber *et al.*, 2007). Several proteins are known to interact as with FHL2, these include: androgen receptor(Heemers *et al.*, 2007), Ap-1 (Morlon and Sassone-Corsi, 2003), CREB (Fimia *et al.*, 2000), WT1 (Du *et al.*, 2002), beta-catenin (Y *et al.*, 2003), CDC47 (Chan *et al.*, 2000), BRCA1 (Yan *et al.*, 2003) and a co-repressor for E4F1 (Paul *et al.*, 2006),ERK2 (Purcell *et al.*, 2004), FOXO1 (Yang *et al.*, 2005), TGFB1 (Dahan *et al.*, 2017). Following nuclear translocation, studies have seen that the activated FHL2 interacts indirectly with the promoter of genes such as *CDKN1A* (Nakazawa *et al.*, 2016) and *CCND1* (Labalette *et al.*, 2008b)where it positively and negatively regulates their expression, respectively thus FHL2 is involved in the cell cycle and therefore the control of proliferation.

Many of these pathways and functions are implicated in wound healing. The roles of FHL2 during wound healing have been summarized and reviewed in (Wixler, 2019) where its

expression is directly linked with the different stages of wound healing. Stimulants of *FHL2* include TGF-beta and an above normal expression of FHL2 was associated with "overhealed" wounds and increased tumorigenesis predisposition (Hinz, 2015). Inhibitors for its expression include the mentioned Il-1beta and TNF-alpha. Sustained down-regulation of FHL2 results in higher expression of enzyme remodelers such as MMPs with the transition into the chronic wound stage (Wixler *et al.*, 2015). Ectopic expression of *FHL2* was detected in cancer rhabdomyosarcoma tissue under the name (down-regulated in rhabdomyosarcoma LIM- protein, DRAL) (Kleiber *et al.*, 2007). FHL2 has been consistently found to be a key factor that is abnormally regulated in carcinogenesis with its role being found in regulation of cellular survival, adhesion, motility, proliferation, and signal transduction (Kleiber *et al.*, 2007; Hua *et al.*, 2016). FHL2 is also implicated in the cellular apoptosis pathways where, once activated, p53 stimulates its transcription (Scholl *et al.*, 2000). This again emphasises its importance in cancer biology and its potential as a therapeutic target in these conditions. Other diseases in which *FHL2* expression is abnormal are hypertrophic cardiomyopathies (Friedrich *et al.*, 2014) and myelopoiesis (Qian *et al.*, 2009).

Due to its complex roles based on cellular location and phosphorylation status, FHL2's role in ageing can be very difficult to estimate. Ageing is associated with a chronic inflammation state termed "inflamm-ageing" (Ferrucci and Fabbri, 2018). This state is characterised by the low grade but increasing presence of inflammatory markers such as IL-1, IL-18. Due to the presence of these factors certain functions can be impacted such as cellular repair, immune response, redox balance and metabolism which can lead to disease predisposition (e.g. cardiovascular disease and cancer) (Rea *et al.*, 2018). Unfortunately, there are limited reports into the function of FHL2 in articular joint tissues and less so of its link with ageing in these tissues. Therefore, an understanding of some possible key functions as well as to model the observed changes using relevant cell lines models is necessary to understand more about it.

6.2 Aims

1. Understand the impact of FHL2 knockdown and overexpression on the cellular transcriptome.

2. Study the localisation pattern of FHL2 in the mouse articular joint.

3. Investigate intra-cellular localisation of FHL2.

4. Determine the impact Interleukin-1 on FHL2 translocation.

6.3 Results

6.3.1 Impact of FHL2 knock down and overexpression in SW1353 using RNA-sequencing

Since the function of *FHL2* seems is not understood in articular joint tissues, a *FHL2* modulation experiment was performed in SW1353 cells. The levels of *FHL2* were knocked down using a CRISPR-Cas9 system and overexpressed using a tagged-plasmid vector (*FHL2-GFP*). Following the confirmation of the knockdown and overexpression of *FHL2* by western blotting, RNA sequencing analysis was performed on these cells to determine the transcriptional consequences of the modulation of *FHL2* expression.

For the knock-down of *FHL2*, two cell lines were made either using just one gRNA or two gRNAs both targeting the third common exon of the gene (i.e. common between all isoforms of *FHL2*) this representing a pool of cells for each respective gRNA combination. This producing a FHL2 knockdown line though the *FHL2* transcript seemingly increasing by 67% ($p= 0.0008$) and +184% ($p<0.0001$) (Figure 6. 1 B I) respectively even though at the protein level there was a reduction of 43% ($p<0.0001$) and 94% ($p<0.0001$), respectively (Figure 6. 1 C). In addition, the cell line where the *FHL2* age-CpG region (Guide pair 1; Chapter 4) was deleted was also included in the analysis and the expression of *FHL2* decreased by 26% ($p= 0.03$) and the FHL2 protein decreased by 18% ($p= 0.0002$) (Figure 6. 1 B I). The *FHL2* overexpressing line saw an increase in total *FHL2* expression by +348% ($p= 0.0007$) (Figure 6. 1 B II) and the contribution of the FHL2-GFP to the total level of cellular FHL2 saw an increase of 74% ($p= 0.002$) (Figure 6. 1 D).

RNA samples from the SW1353 cell lines that were FHL2 knock down (section 2.5.5) or overexpressing of FHL2 (section 2.3.3) and their controls were treated with DNase and submitted for RNA-sequencing analysis at the Genomics Core Facility (Newcastle University) with the reads per sample at $20x10^6$ reads and with read length of 75 bp per sample. After the receipt of the data, this was aligned to the Gencode human V38 index basic gene annotation dataset on the reference chromosomes using Salmon and analysed and normalised for differential expression using Deseq2 tool in Rstudio. A principal component analysis was performed (Figure 6. 2). Differential expression analysis indicates that for the

knockdown variant gRNA2 (Figure 6. 3 A) there were 99 genes that were downregulated at a threshold of fold change (FC) <0.66 (log_2FC< -0.599) and adjusted p value<0.05. Gene ontology analysis indicating that these genes are implicated in pathways related to "extracellular matrix organisation", "regulation of cellular proliferation", "actin filament based process" (i.e. pathway that can alter the cellular cytoskeleton) and "cellular adhesion" (Figure 6. 3 B). Seventy-four genes were upregulated >1.5 fold (log_2FC> 0.585) at an adjusted p value<0.05, with gene ontology analysis indicating that these are involved in metabolic pathways such as "sterol, steroid and cholesterol biosynthesis" (Figure 6. 3 B). For the knockdown guide pair 2 (Figure 6. 4) that also presented the highest level of *FHL2* knockdown, there were 133 downregulated genes FC<0.66 (log_2FC< -0.599) at an adjusted p value <0.05, where these genes were found through gene ontology analysis to be in pathways relating to "extracellular matrix and structure organisation", "regulation of cellular proliferation", "cell adhesion" and "collagen fibril organisation". For the upregulated genes a total of 146 genes were found to pass the significance threshold >1.5 fold (log_2FC> 0.585) at an adjusted p value<0.05, with these being involved in pathways such as "sterol and secondary alcohol metabolism" and "cell junction organisation". For the *FHL2-GFP* overexpression variant (Figure 6. 6), there were 66 genes that were found to be down regulated at FC<0.66 (log_2FC< -0.599) adjusted p value <0.05, with the gene ontology analysis revealing that these genes are implicated in the activity of an "L-amino acid transmembrane transporter"," extracellular matrix and structure organisation". For the upregulated genes there were 104 genes >1.5 fold (log_2FC> 0.585) at an adjusted p value<0.05 of its respective *GFP* control with these genes being implicated in pathways relating to "cellular and biological adhesion", "cellular motility and neural development".

Since there are two *FHL2* knockdown lines of different knockdown levels it would be important to understand what the intersection between the list of downregulated and upregulated genes is (Figure 6. 8). For the list of upregulated genes there were 40 genes in common between the two lines and for the downregulated genes 31 were found (Figure 6. 8 B). On the other hand, looking at genes that go in opposing directions between the conditions of knockdown and overexpression, for such analysis not many genes were found to exhibit this characteristic. Looking at the *FHL2* K.D. gRNA2 line vs. the *FHL2-GFP* overexpression line (Figure 6. 9) only 8 genes were found to go in an opposing direction and

for the *FHL2* K.D. guide pair 2 line vs. the *FHL2-GFP* overexpression line (Figure 6. 10) only 12 genes. From the gene ontology analysis, a pathway that stood up to be in common between conditions is the potential impact on cellular proliferation and metabolism.

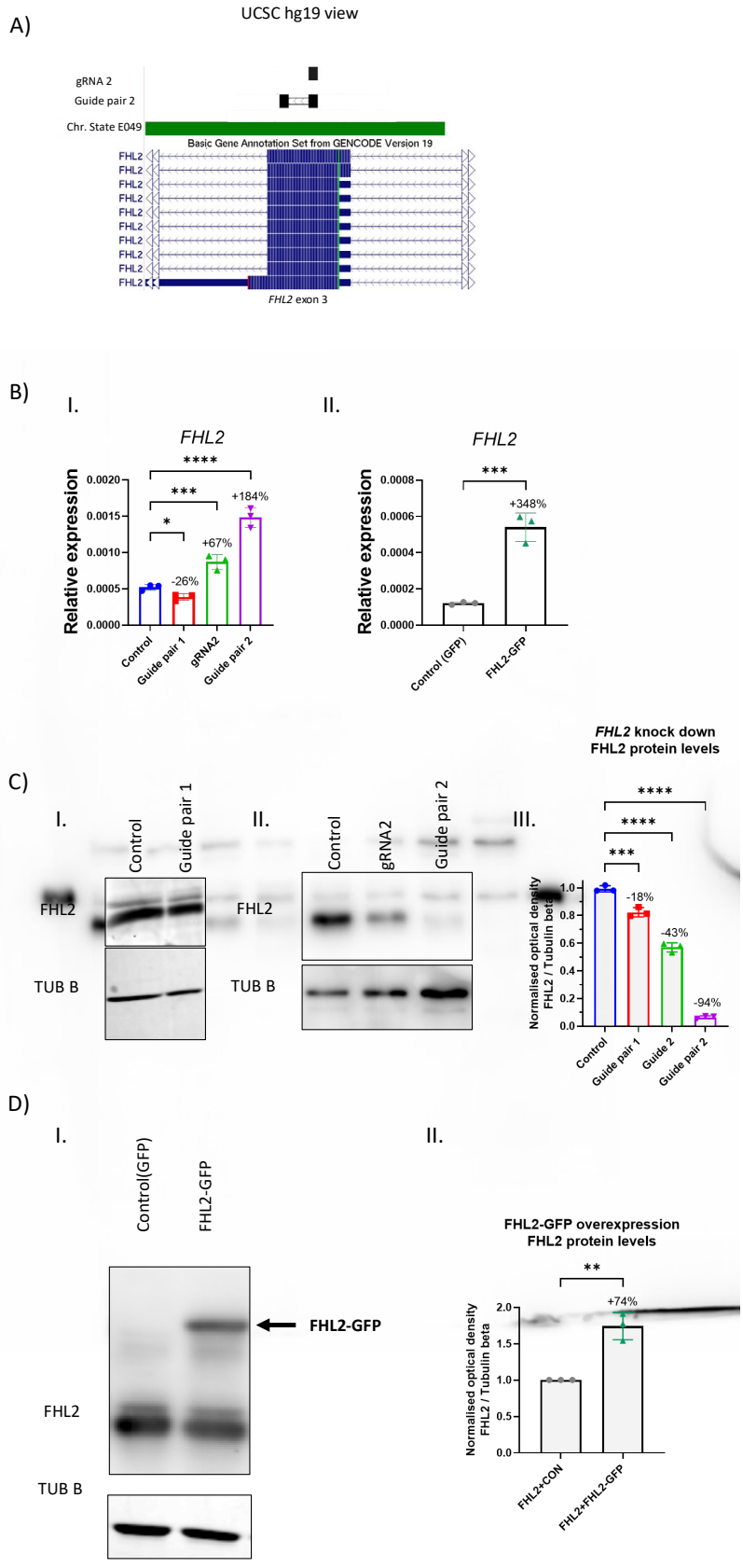


Figure 6. 1. FHL2 knockdown and overexpression cell line production. A) UCSC genome browser view in hg19 indicating the binding location of the gRNAs used for the knockdown

B) *FHL2* expression qRT-PCR analysis I. *FHL2* knockdown lines II. FHL2-GFP overexpression line. FHL2 protein size was determined to be ~32kDa with tubulin at ~55kDa C) Western blotting results of the FHL2 knockdown line I. for guide pair 1, II. gRNA2 & guide pair 2 combinations and III. densiometric analysis of western blotting result with one-way ANOVA n=3, ***p<0.001, ****p<0.0001. D) I) Western blotting results of the FHL2-GFP overexpression line, FHL2-GFP construct was determined to have a size of ~58kDa, II. densiometric quantification with unpaired t-test analysis n=3, **p<0.01.

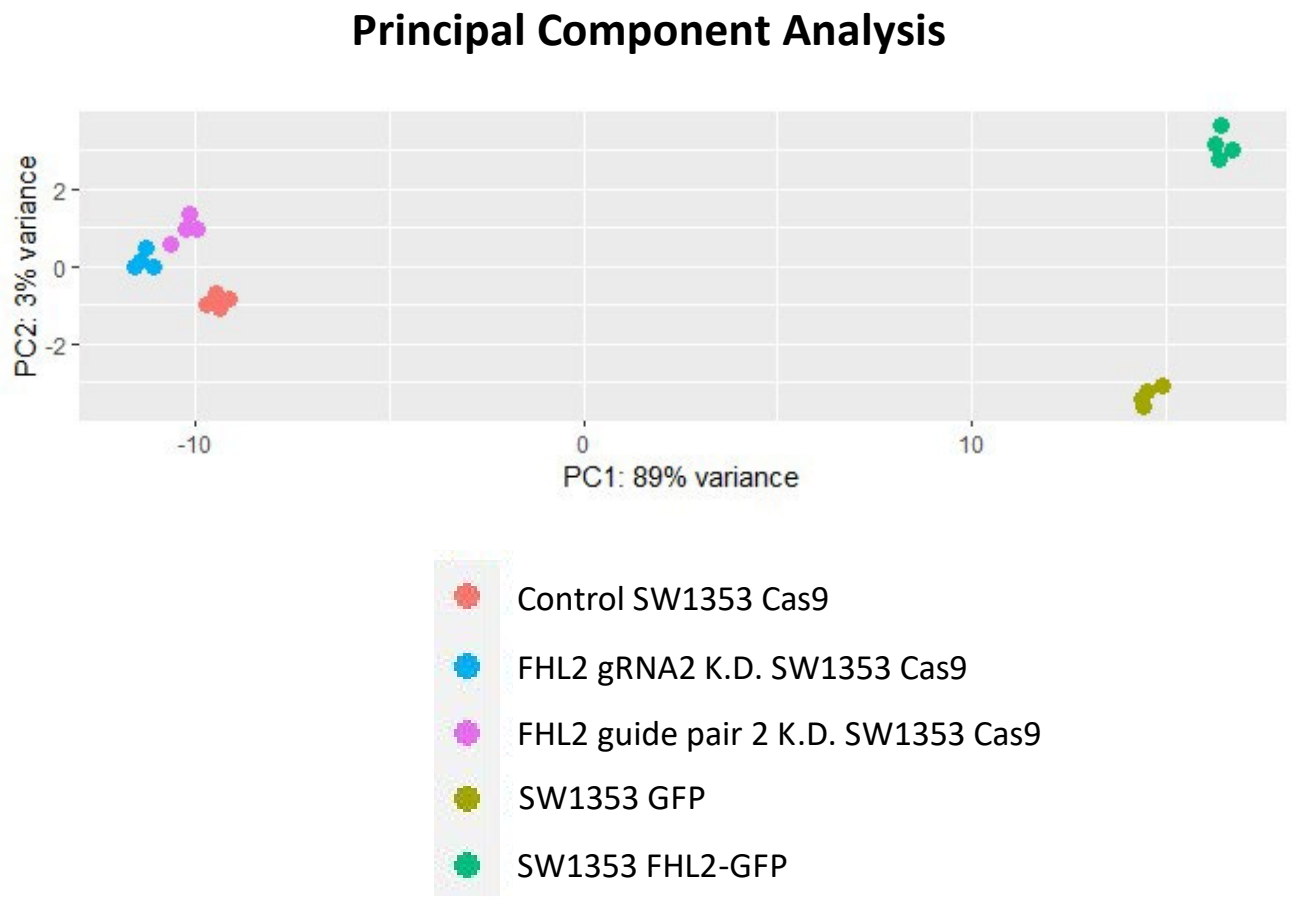
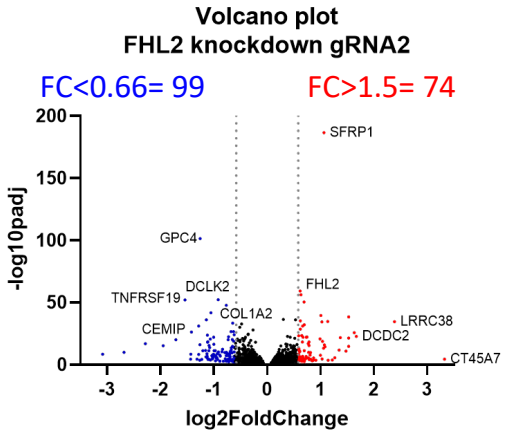


Figure 6. 2. Principal component analysis (PCA) for the *FHL2* knockdown lines gRNA2 (blue) and guide pair 2 (magenta) and FHL2-GFP overexpression (green) lines together with their respective controls as control Cas9 SW1353 cells (red) and SW1353 GFP line (olive) respectively. The PCA was performed on the Deseq2 normalised count data for each line in Rstudio.

A)



B)

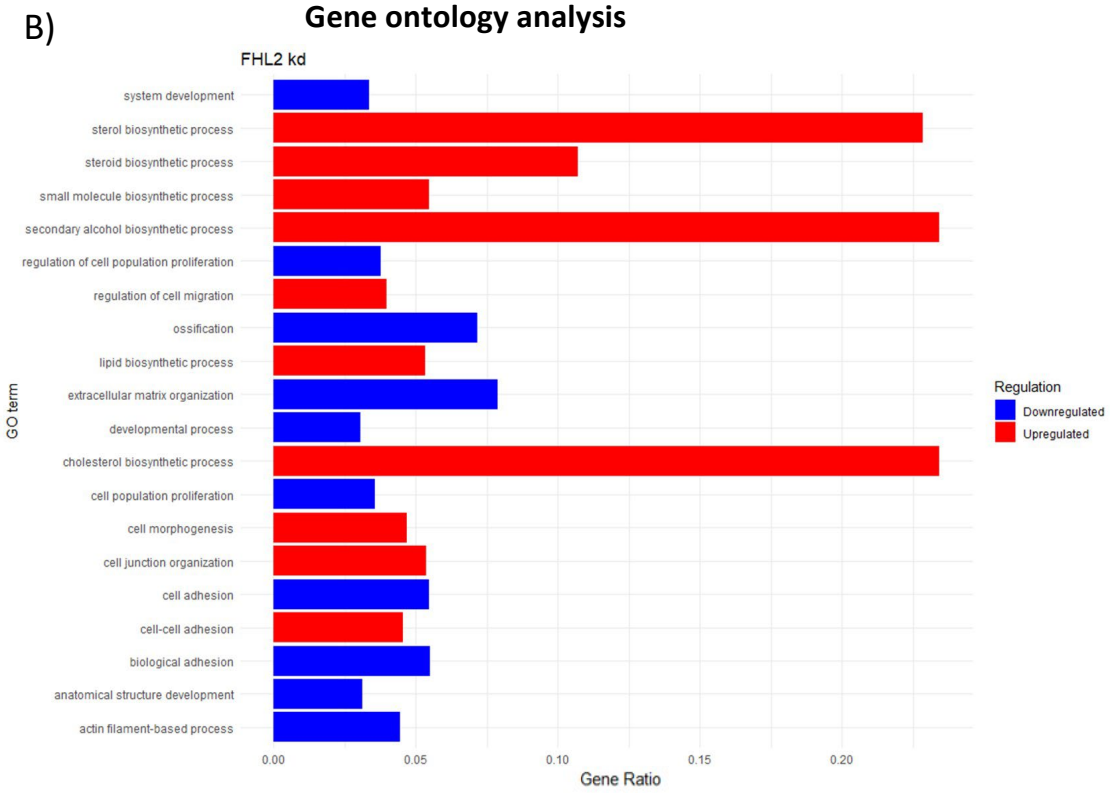
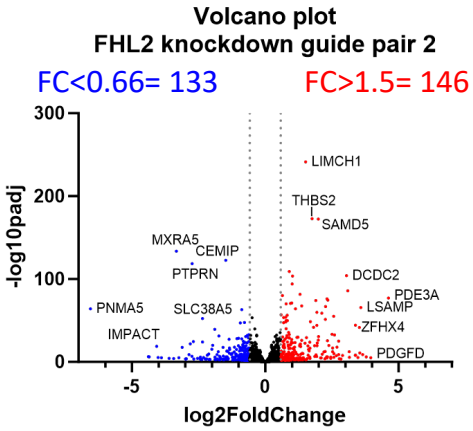


Figure 6. 3. *FHL2* knockdown gRNA2 line differential expression analysis. A) Volcano plot of differentially expressed genes for FHL2 knock down gRNA2 SW1353 Cas9 cells, down-regulated genes are marked in blue and up-regulated genes are red with indication of the number of statistically significant downregulated and upregulated genes B) Gene ontology analysis for the significantly downregulated (blue) and upregulated (red) genes for the *FHL2* knockdown experiment.

A)



B)

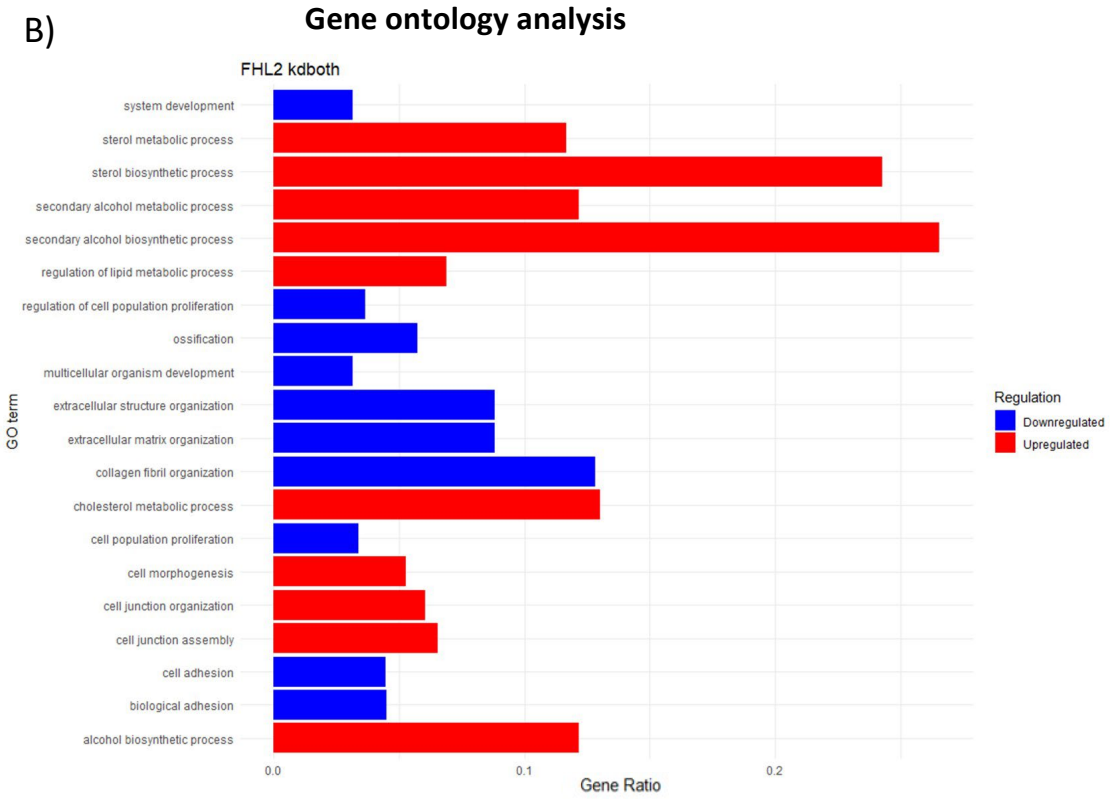


Figure 6. 4. *FHL2* knockdown guide pair 2 line differential expression analysis. A) Volcano plot of differentially expressed genes for FHL2 knock down guide pair 2 SW1353 Cas9 cells with down-regulated genes are marked in blue and up-regulated genes are red with indication of the number of statistically significant downregulated and upregulated genes B) Gene ontology analysis for the significantly downregulated (blue) and upregulated (red) genes for the *FHL2* knockdown experiment.

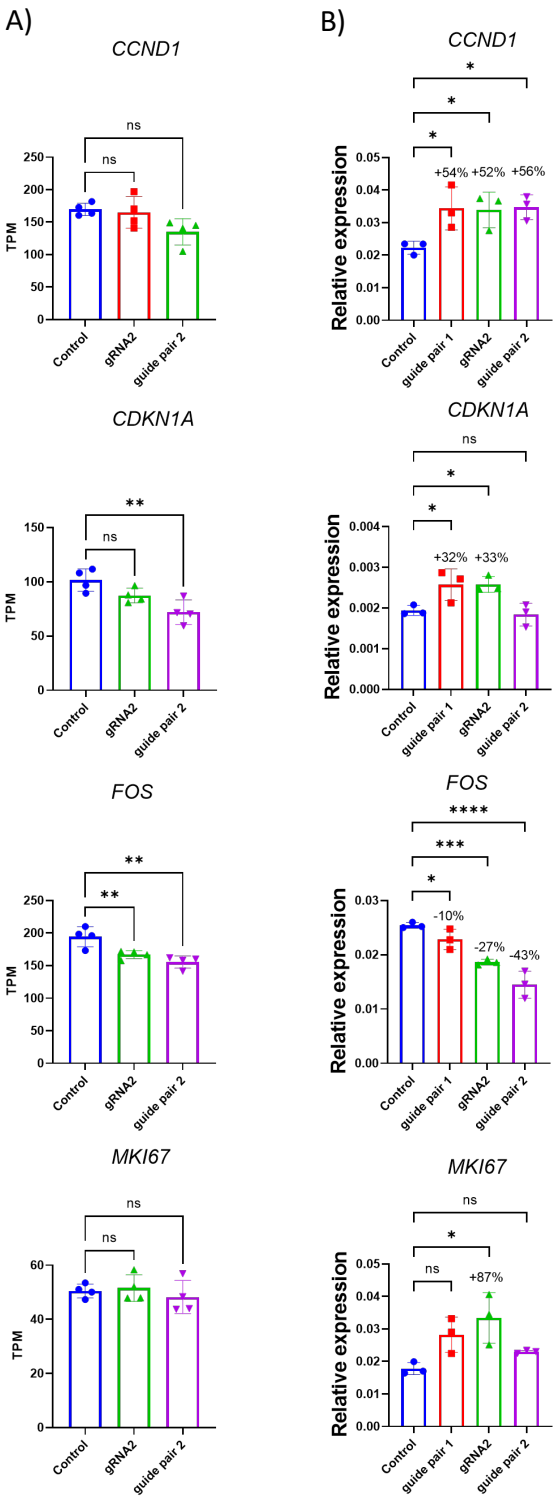
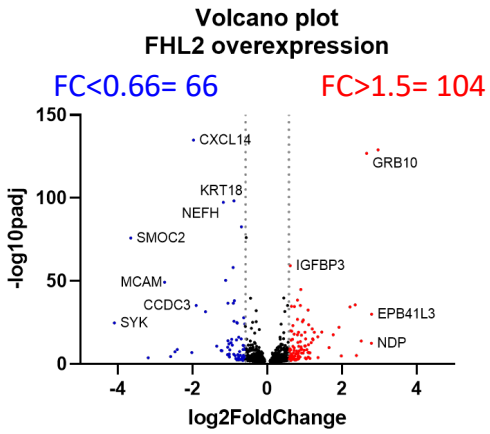


Figure 6. 5. Gene expression quantification for the SW1353 Cas9 *FHL2* knockdown lines. A) Transcripts per million as given by RNA-sequencing and B) validation by qRT-PCR. One-Way ANOVA Dunnett method for adjusting for multiple comparisons, n=3; *p<0.05, **p<0.01,***p<0.001, ****p<0.0001.

A)



B)

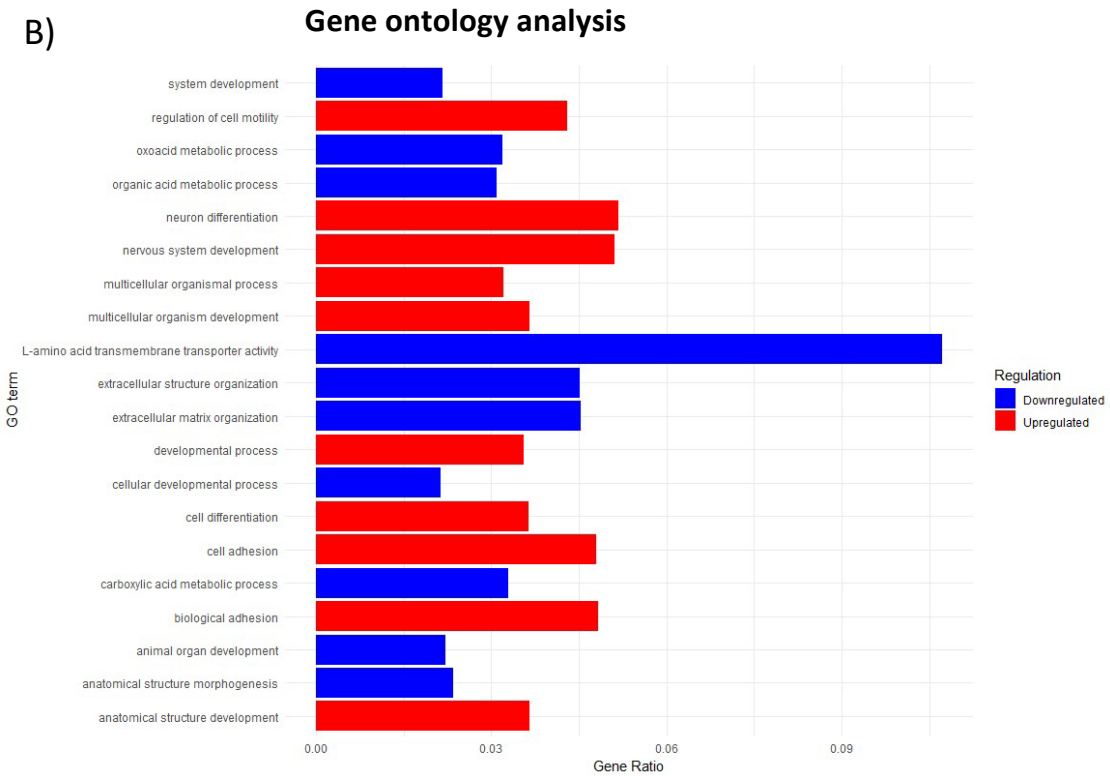


Figure 6. 6. FHL2-GFP overexpression line differential expression analysis. A) Volcano plot of differentially expressed genes for FHL2-GFP line, down-regulated genes are marked in blue and up-regulated genes are red with indication of the number of statistically significant downregulated and upregulated genes B) Gene ontology analysis for the significant downregulated (blue) and upregulated (red) genes for the FHL2 overexpression experiment.

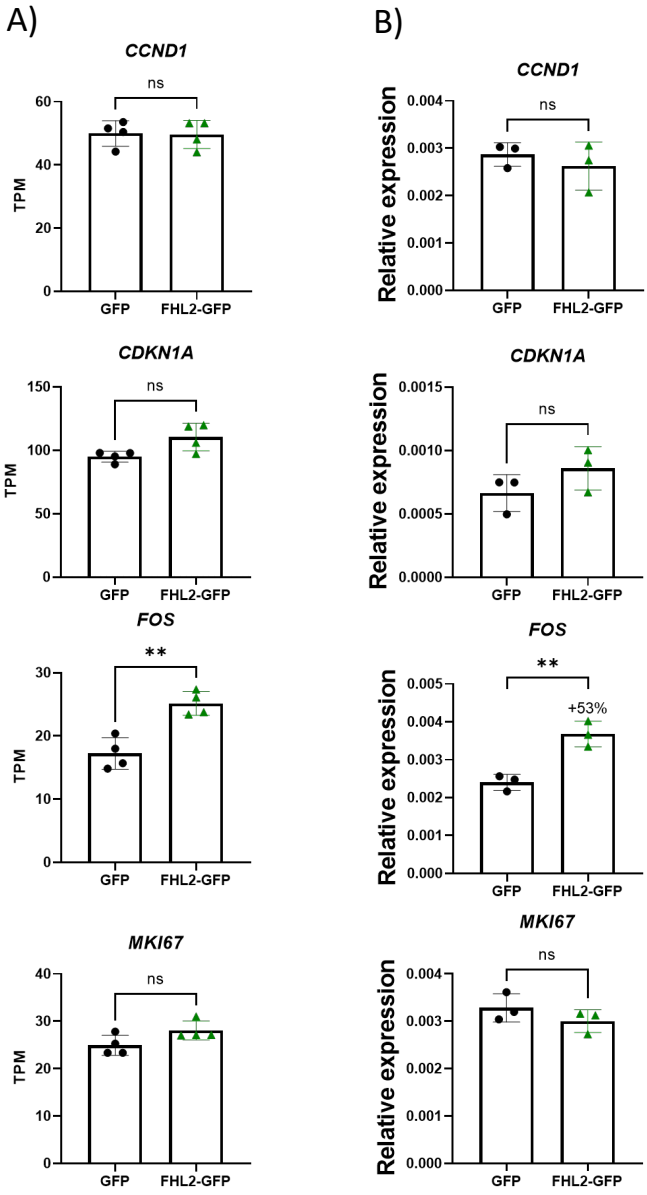


Figure 6. 7. Gene expression quantification for the SW1353 FHL2-GFP overexpression line A) Transcripts per million as given by RNA-sequencing and B) validation by qRT-PCR. Unpaired t-test analysis, n=3; **p<0.01

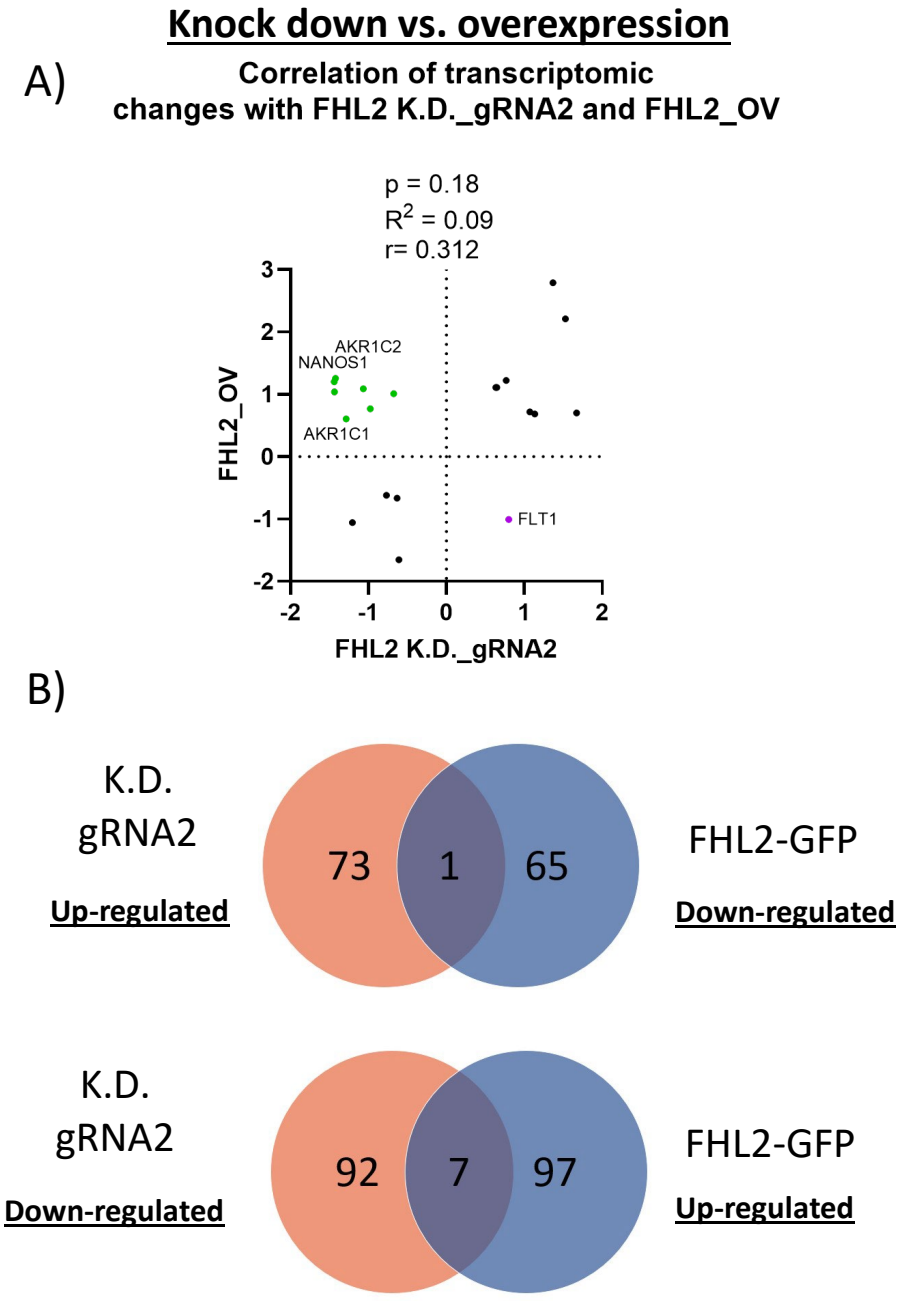


Figure 6. 8. Correlation analysis Venn diagram representation between FHL2-GFP and K.D. gRNA2 lines. A) Down-regulated genes for the FHL2 K.D. gRNA2 variant and that of the up-regulated genes for the FHL2-GFP overexpression line with the intersection between the two. B) Up-regulated genes for the FHL2 K.D. gRNA2 variant and that of the down-regulated genes for the FHL2-GFP overexpression line with the intersection between the two.

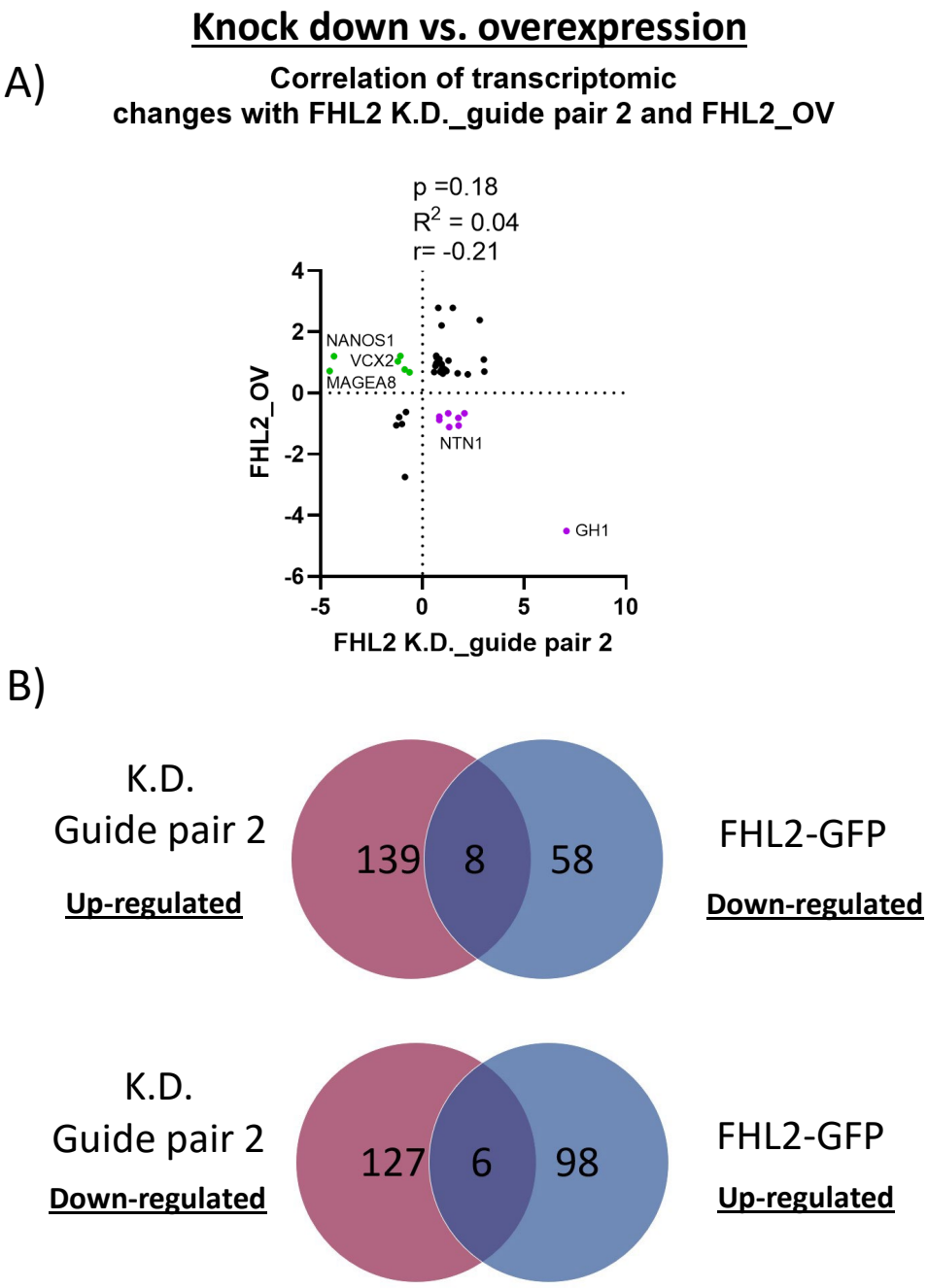


Figure 6. 9. Venn diagram representation between FHL2-GFP and K.D. guide pair 2 lines. A) Down-regulated genes for the FHL2 K.D. guide pair 2 variant and that of the up-regulated genes for the FHL2-GFP overexpression line with the intersection between the two. B) Up-regulated genes for the FHL2 K.D. guide pair 2 variant and that of the down-regulated genes for the FHL2-GFP overexpression line with the intersection between the two.

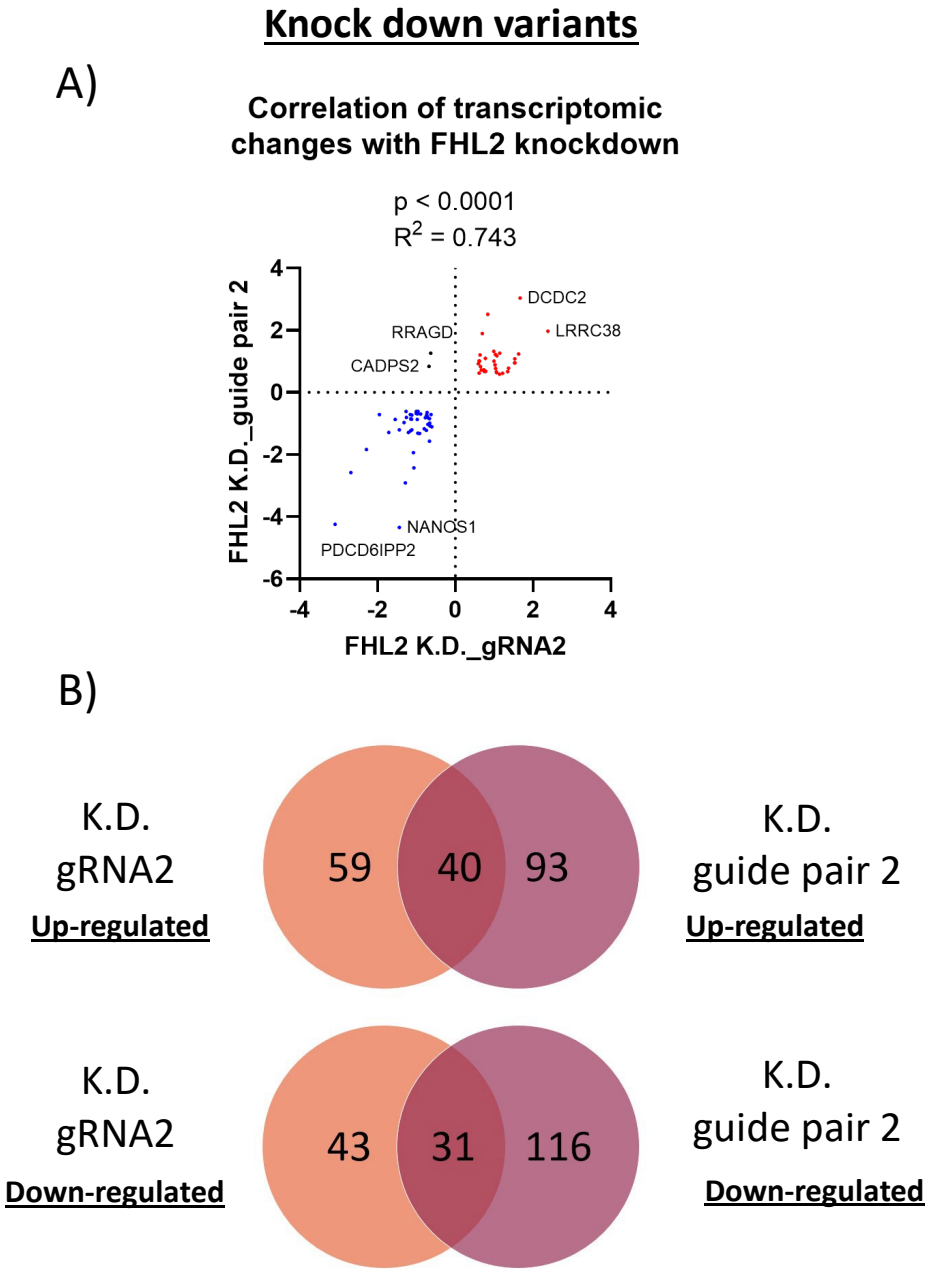
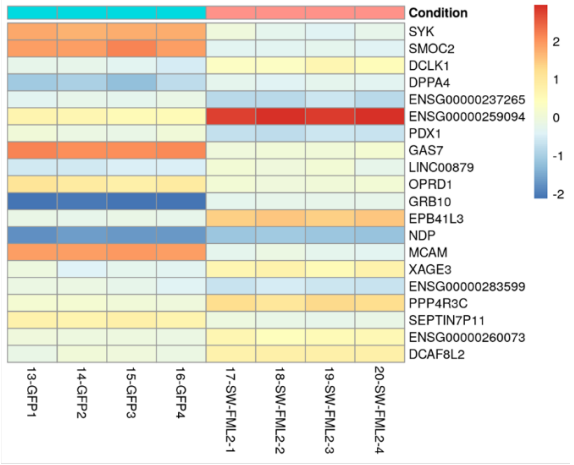
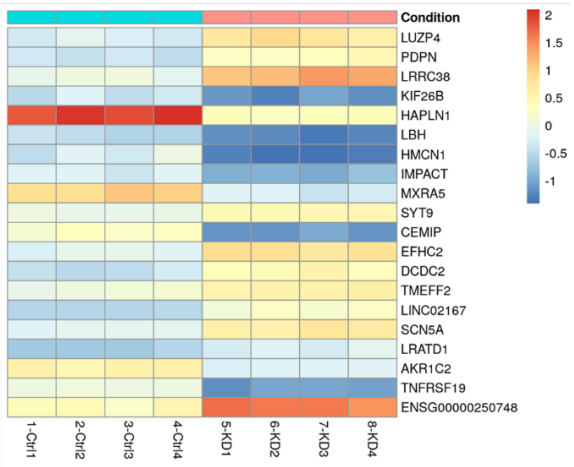


Figure 6. 10. Venn diagram representation between K.D. gRNA2 and K.D. gRNA2 lines. A) Up-regulated genes for the FHL2 K.D. gRNA2 variant and that of the down-regulated genes for the FHL2 K.D. guide pair 2 variant with the intersection between the two. B) Down-regulated genes for the FHL2 K.D. gRNA2 variant and that of the up-regulated genes for the FHL2 K.D. guide pair 2 variant with the intersection between the two.

A)


Condition
SYK
SMOC2
DCLK1
DPPA4
ENSG00000237265
ENSG00000259094
PDX1
GAS7
LINC00879
OPRD1
GRB10
EPB41L3
NDP
MCAM
XAGE3
ENSG00000283599
PPP4R3C
SEPTIN7P11
ENSG00000260073
DCAF8L2
2
1
0
-1
-2
13-GFP1
14-GFP2
15-GFP3
16-GFP4
17-SW-FML2-1
18-SW-FML2-2
19-SW-FML2-3
20-SW-FML2-4

B)


Condition
LUZP4
PDPN
LRRC38
KIF26B
HAPLN1
LBH
HMCN1
IMPACT
MXRA5
SYT9
CEMIP
EFHC2
DCDC2
TMEFF2
LINC02167
SCN5A
LRATD1
AKR1C2
TNFRSF19
ENSG00000250748
2
1.5
1
0.5
0
-0.5
-1
1-Ctrl1
2-Ctrl2
3-Ctrl3
4-Ctrl4
5-KD1
6-KD2
7-KD3
8-KD4

C)

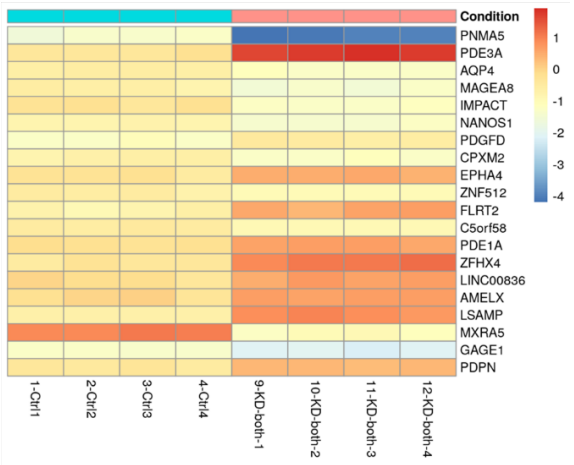

Condition
PNMA5
PDE3A
AQP4
MAGEA8
IMPACT
NANOS1
PDGFD
CPXM2
EPHA4
ZNF512
FLRT2
C5orf58
PDE1A
ZFHX4
LINC00836
AMELX
LSAMP
MXRA5
GAGE1
PDPN
1
0
-1
-2
-3
-4
1-Ctrl1
2-Ctrl2
3-Ctrl3
4-Ctrl4
9-KD-both-1
10-KD-both-2
11-KD-both-3
12-KD-both-4

Figure 6. 11. Heatmap representation of sample variability within each group for the FHL2 overexpression condition. A) gRNA2 knockdown of FHL2, B) guide pair 2 knockdown of FHL2 and C) Overexpression of FHL2 using FHL2-GFP plasmid vector.

6.3.2 Impact of FHL2 knock down and overexpression on cell proliferation

Based on the results from the RNA-sequencing these suggested that some of the pathways impacted are linked with cellular proliferation and therefore further testing was done in order to assess this.

To investigate cellular proliferation two methods were employed, the BrdU integration assay followed by cellular counting, and the colorimetric WST-1 assay. The WST-1 assay measures cellular proliferation through the formation of a coloured product through the mitochondrial dehydrogenase enzyme cleavage on the WST-1 compound. The BrdU assay measures cellular integration by the integration of BrdU into the replicating DNA and the subsequent detection of this by fluorescently labelled anti-BrdU antibodies.

For the *FHL2* knockdown lines (Figure 6. 12 A & B), both cell proliferation BrdU and WST-1 assays are in agreement of an increase in cellular proliferation associated with the knockdown and moreover this seems to be related with the level of *FHL2* knockdown. The cell line with the lowest *FHL2* knockdown had an increased proliferation rate of +14% for both assays ($p<0.05$) and with the cell line with the highest *FHL2* knockdown level showing an increase of 38% ($p<0.0001$) by the BrdU method and 49% ($p<0.0001$) by the WST-1 assay. A reason for the discrepancy between the results of the proliferation assays could be the experimental error and the resolution of the techniques.

For the *FHL2-GFP* overexpression line (Figure 6. 12 C & D) proliferation was seen to decrease by 15% ($p<0.05$) by the BrdU assay method compared to its control, however the WST-1 assay method indicated an increased proliferation rate of 8% ($p<0.05$).

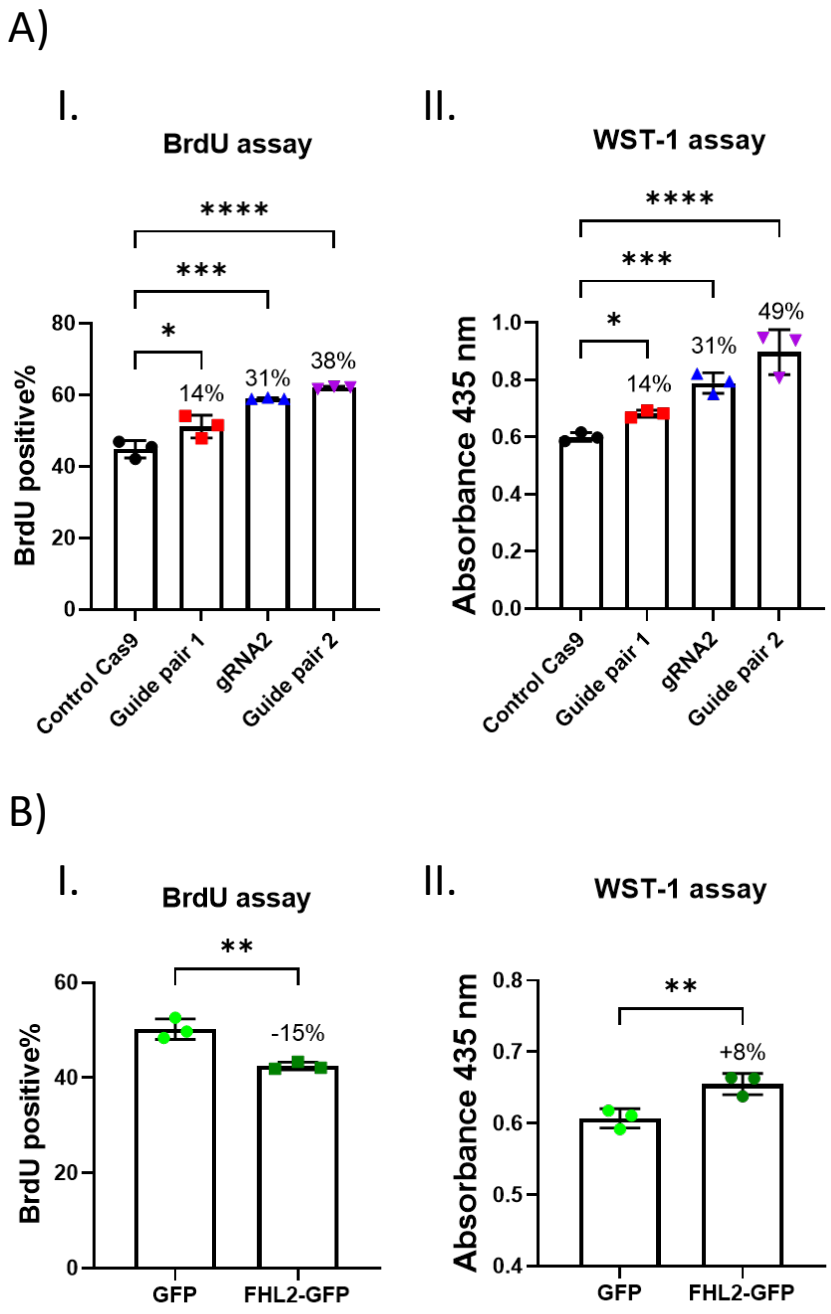


Figure 6. 12. Cell proliferation analysis of the FHL2 knockdown and FHL2-GFP overexpression lines. FHL2 knockdown lines proliferation results A) with BrdU assay (I.) and WST-1 assay (II.); One-way ANOVA with Dunnett correction method, *p<0.05, ***p<0.001, ****p<0.0001. FHL2-GFP overexpression line proliferation results B) by the BrdU assay (I.) and WST-1 assay (II.); unpaired t-test analysis, **p<0.01.

6.3.3 Impact of partial restoration of FHL2 in FHL2 knock down line on gene expression and cell proliferation

To understand whether the impact of the *FHL2* knockdown is reversible, *FHL2* was reintroduced into the cell line with the highest knockdown using a *FHL2-GFP* plasmid vector. The reintroduction was successful as verified by both qRTPCR to check the transcription of *FHL2* where an increase of 443% was observed compared to the FHL2 knockdown control line (Figure 6. 13 A), and western blotting to check for the presence of the FHL2-GFP construct (Figure 6. 13B).

Following the successful reintroduction of *FHL2* the genes identified as modulated by FHL2 levels were measured along with the cellular proliferation rate. For the gene expression changes, the re-introduction of FHL2 caused a statistically significant increase for *CCND1* (+68%)*, CDKN1A* (+147%) and *FOS* (+170%). No significant change was observed for the *MKI67* gene (p =0.07) although this shows +64% increased expression. On the proliferation rate, the re-introduction of FHL2 resulted in a reduction by 22% in proliferation measured by the BrdU assay (p<0.001) and for the WST-1 a reduction in absorbance by 6% (p <0.01) which correlates to reduced proliferation (Figure 6. 14).

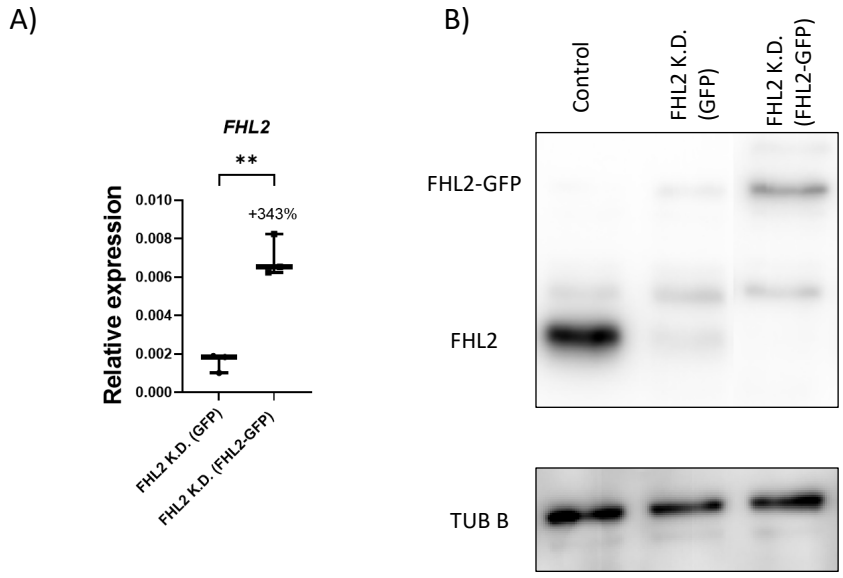


Figure 6. 13. Re-introduction of FHL2 in the Cas9 SW1353 K.D. guide pair 2 line. A) Gene expression quantification for FHL2 after introduction of the FHL2-GFP construct. Unpaired t-

test analysis. B) Western blot probing with anti-FHL2 antibody and loading control as beta tubulin for SW1353 Cas9 (Control) cells, SW1353 Cas9.

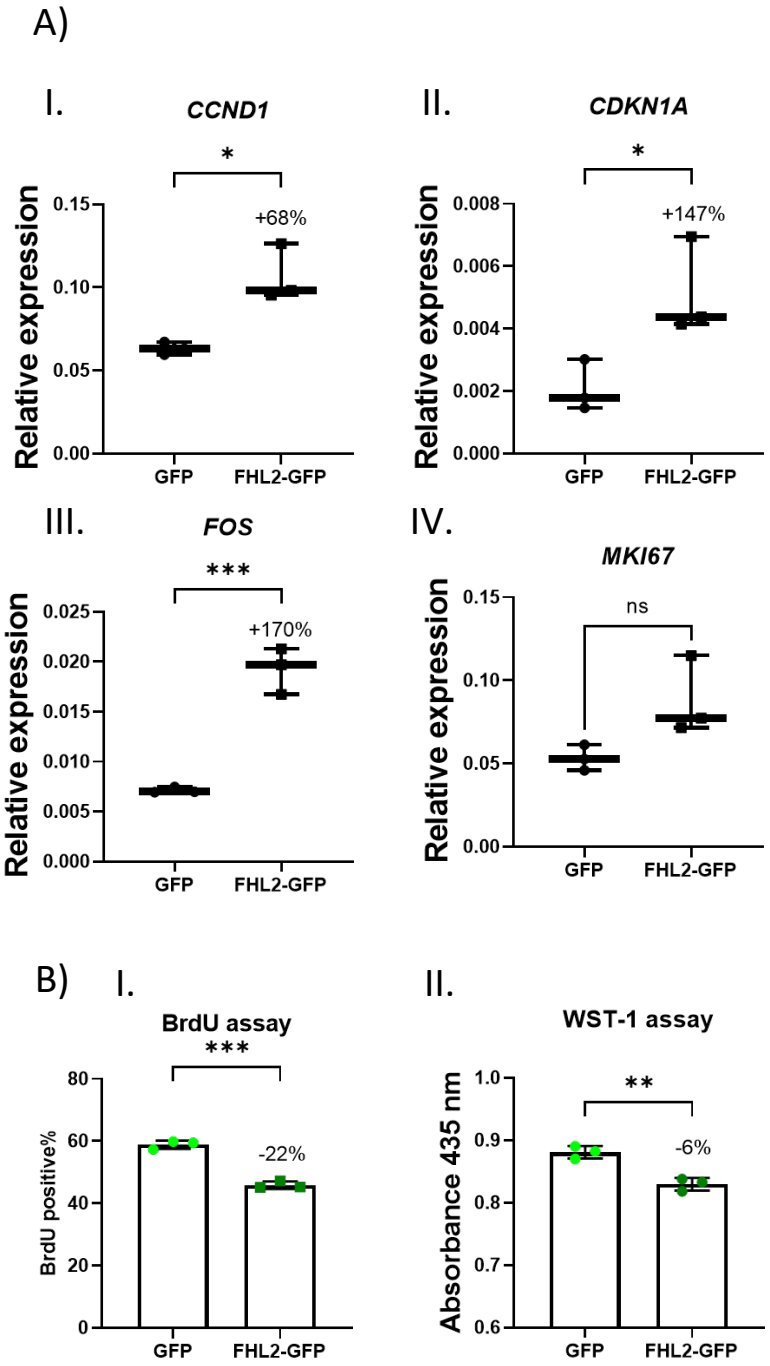


Figure 6. 14. Re-introduction of FHL2 in the Cas9 SW1353 K.D. guide pair 2 line impact on gene expression. A) qRT-PCR results for the impact of FHL2 restoration in the FHL2 guide

pair 2 knockdown line. Genes analysed are CCND1 (I.), CDKN1A (II.), FOS (III.) and MKI67 (IV.); upaired t-test analysis, *p<0.05, ***p<0.001. B) Proliferation analysis for the impact of FHL2 restoration in the FHL2 guide pair 2 knockdown line, BrdU assay (I.) and II. WST-1 assay (II.) results. Unpaired t-test analysis, **p<0.01, ***p<0.001.

6.3.4 Immunohistochemical localisation of FHL2 with age in mouse knee joint

FHL2 is known to be expressed throughout the human body tissues (Zheng and Zhao, 2007; Cao *et al.*, 2015). According to data available through the GTEX depository, that contains expression data of various tissues of different lineages for human adults, *FHL2* is expressed with a TPM>2 in 42/56 of the data sets of different tissues including: liver, heart, testis, ovary etc. TPM<2 is observed for many neuronal tissues, with exemption of brain cortex (Figure 6. 15) (GTExConsortium, 2013).

Little information is available about the role and localisation of FHL2 within the articular joint and therefore a preliminary analysis was performed on mouse knee joint of different ages to determine the tissue expression pattern through immunohistological techniques. Mouse tissue of ages 7-day, 4-months and 12-months were probed with an anti-FHL2 antibody. The four months tissue was taken from animals having experienced OA-induction by having destabilisation of median meniscus (DMM) surgical model at 12 weeks of age. Tissue samples were provided by Hua Lin from spare samples used within the Young group. Additional histological samples corresponding to the same sample joints were provided by Hua Lin with histological stains such as Safranin O & Fast Green and Toluidine Blue.

At day seven (Figure 6. 16), FHL2 localisation is predominantly observed within the skeletal muscle and bone marrow tissues, tissues known to express FHL2 and therefore acted a positive staining control (Paul *et al.*, 2006; Hamidouche *et al.*, 2008; Kurakula *et al.*, 2014). Toluidine blue staining marks glycosaminoglycan rich tissues such as the cartilage (Sridharan and Shankar, 2012). Inspection at the cartilage shows FHL2 positive cells. By visual division of the cartilage region, within the articular cartilage positive FHL2 cells are observed in the superficial zone, intermediate zone. At the growth plate region FHL2 is absent from the resting zone and proliferation zone with some positive cells in the hypertrophic zone –

though cells here are large, and the nucleus may not be present in all sections. In other tissues of the articular joint, FHL2 positive cells are observed in the fat pad, synovium and bone tissues (Figure 6. 17).

Looking at the four months old mouse (Figure 6. 18), within the cartilage tissue, stained red by the Safranin O + Fast Green staining, FHL2 is present mainly in the superficial zone of the articular cartilage with some presence observed for the growth plate cartilage. Within bone FHL2 positive osteocytes cells are visible but staining is not clear in for the other bone cell types. Again, there is a strong presence of FHL2 positive cells observed for the skeletal muscle and bone marrow. At this age, the fat pad and synovium tissues only show a few positive cells with weak staining for FHL2. This specimen was from an animal also subjected to DMM surgery eight weeks previously. This surgery generally causes OA-like disease initially on the medial side of the joint at this time point. The degradation of the articular cartilage caused by the procedure reduces the number of cells present in the superficial zone and therefore results in a decreased presence of FHL2 positive cells compared to the opposite non-eroded lateral side of the joint. In order to understand whether the DMM could impact *Fhl2* expression, RNA-sequencing data between DMM and control was provided by Prof. David Young, this indicated that the procedure had no significant impact on the expression of this gene. Additionally, published RNA-sequencing data (Ajekigbe *et al.*, 2019) (Figure 6. 18 D) for neck of femur fracture controls and OA hip samples again showed that the OA condition does not change the expression of *FHL2* at a statistically significant level ($p>0.05$).

At the 12 months age group, FHL2 is still strongly present in the skeletal muscle and bone marrow. For the articular cartilage, FHL2 positive cells are only seen for the superficial zone with positive cells only at the periphery. Closer inspection at the bone tissue again shows a strong presence in osteocytes, with some FHL2 positive cells in the fat pad and synovial membrane (Figure 6. *19*).

This analysis was only performed with an n=1 for each age group and are therefore only preliminary findings.

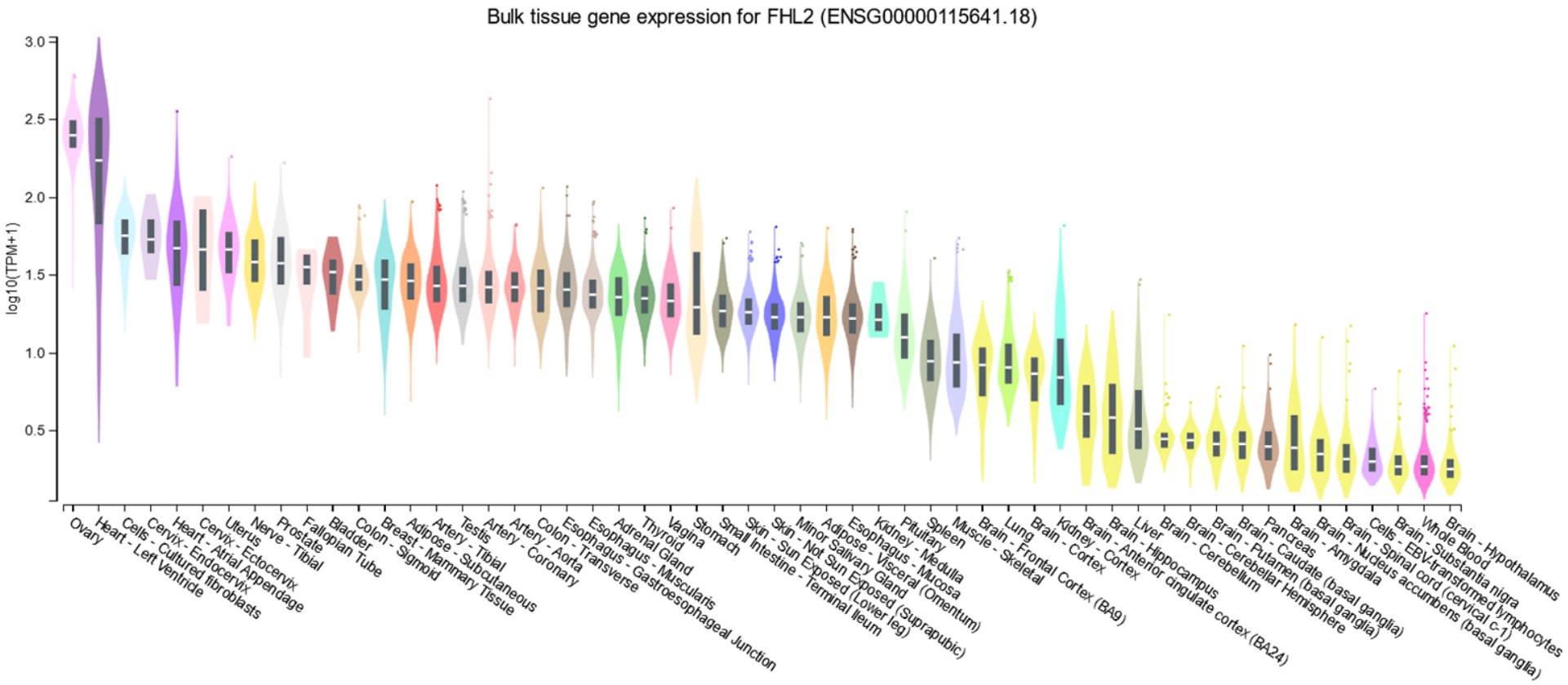


Figure 6. 15. Ranked tissue expression of FHL2 in the database available through the GTEx repository. TPM values are taken as the log10 of the values for visualisation of all of the tissue expression values. Violin box plot where the median and the 25th and 75th percentiles are displayed. Data accessed 25.09.2022 (Consortium, 2013)

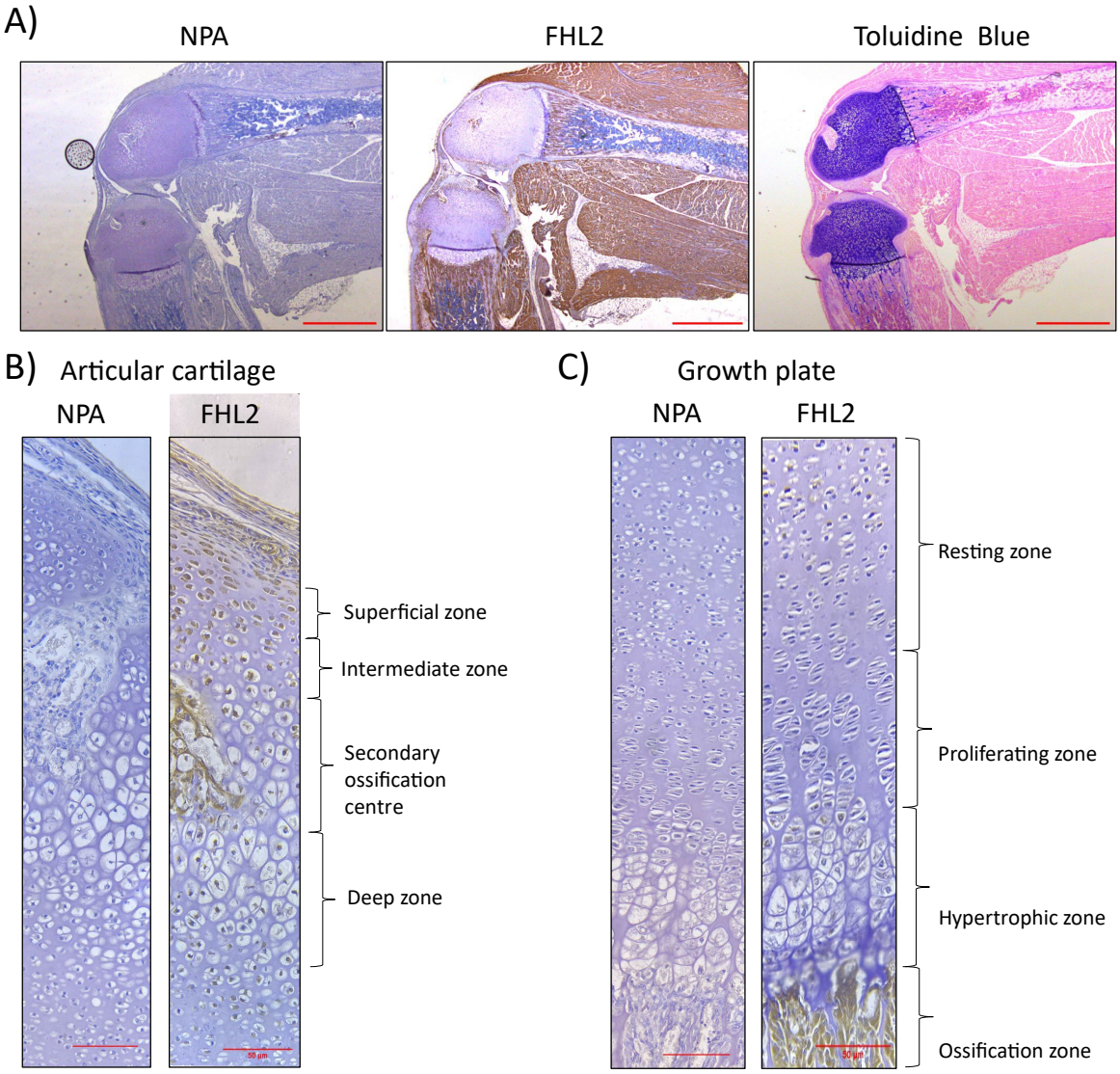


Figure 6. 16. Immunohistochemical analysis for the distribution FHL2 in the mouse knee joint at day seven. A) Whole joint view under no primary antibody (NPA), FHL2 and toluidine blue staining. Scale bar represents 500 µm. B) Higher magnificent examination of the articular cartilage distribution with the approximate zones; 50 µm scale bar C) growth plate cartilage distribution with the approximate zones; 50 µm scale bar.

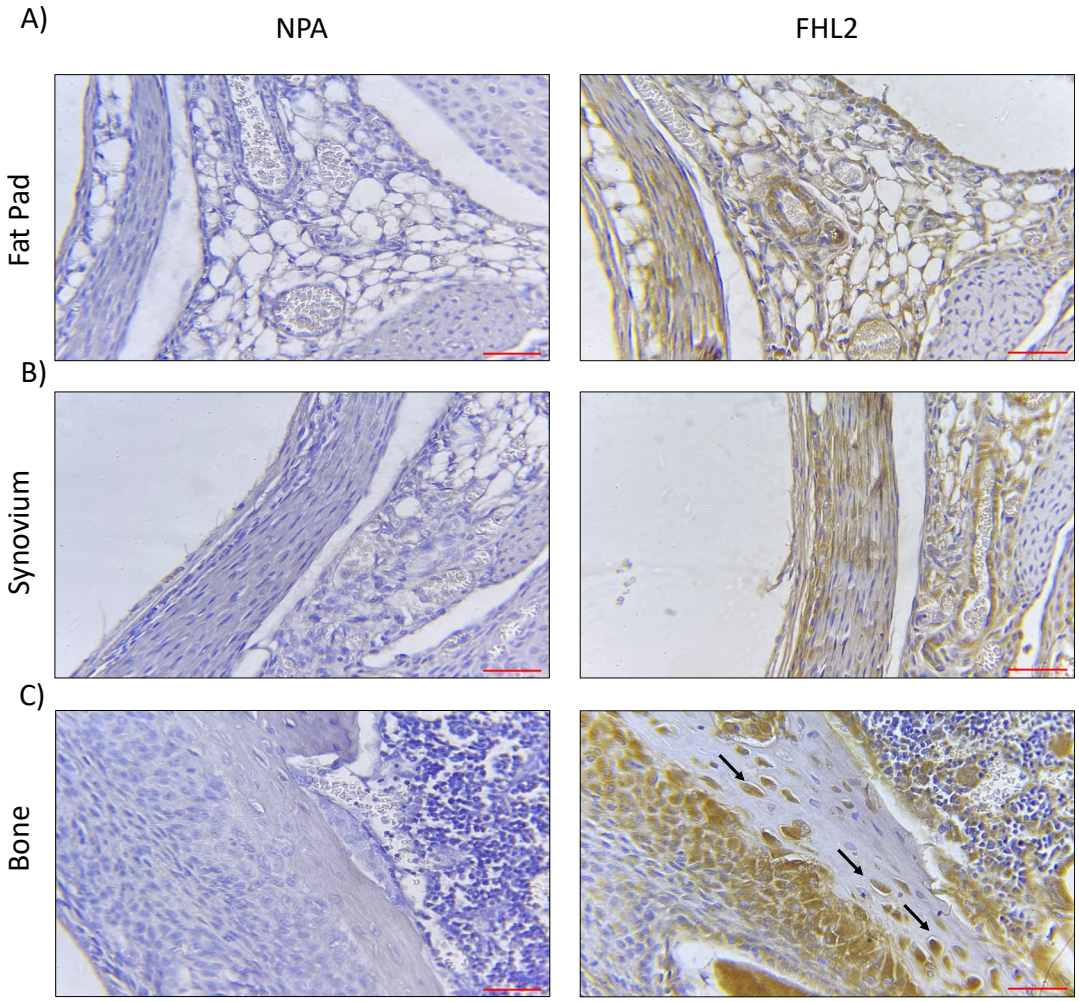


Figure 6. 17. Immunohistochemical investigation for the distribution of FHL2 in the tissues of day seven mouse A) fat pad, B) synovium and C) bone (arrows mark the position of osteocytes) with the respective NPA controls. Scale bar represents 50 µm.

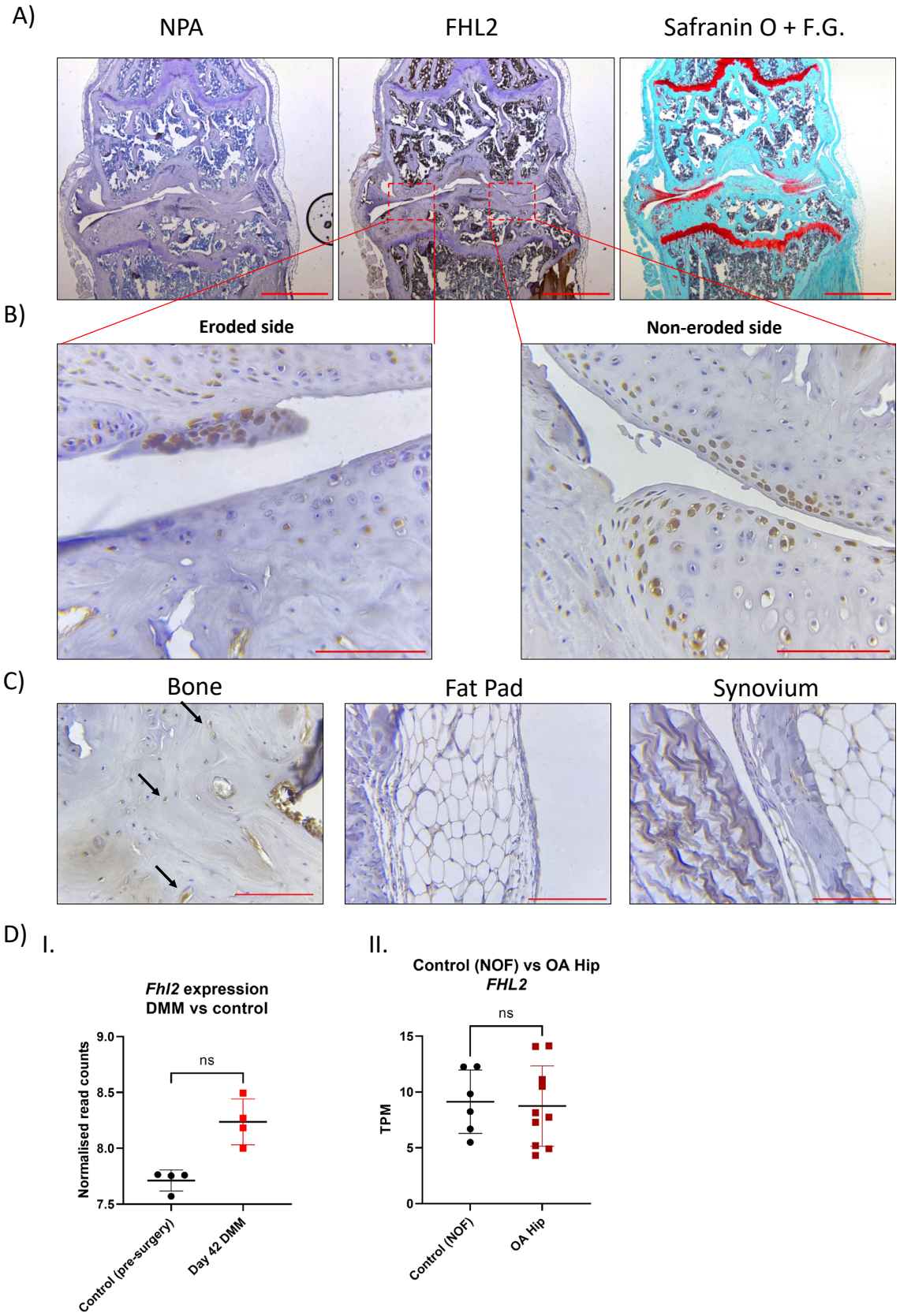


Figure 6. 18. Immunohistochemical investigation of FHL2 distribution at four months of age in the mouse which underwent destabilisation of medial meniscus surgery (DMM). A) Whole

joint view with NPA, FHL2 and Safranin O + Fast Green staining; 500 µm scale bar. B) Articular cartilage at the site of erosion due to the destabilisation of the joint and the other side where no-erosion can be seen. Scale bar represents 50 µm. C) Distribution of FHL2 in the tissues of bone (arrows mark the position of osteocytes), fat pad. Scale bar represents 50 µm. D) I. RNA-sequencing normalized counts for Fhl2 comparison between control (pre surgery) and DMM at day 42 mouse articular cartilage samples, differential expression analysis using Deseq2. II. RNA-sequencing TPM data comparison of expression of FHL2 in control neck of femur fracture (NOF) (non-OA) (n= 6) and end-stage OA (n= 10) hip human articular cartilage, differential expression analysis using Deseq2.

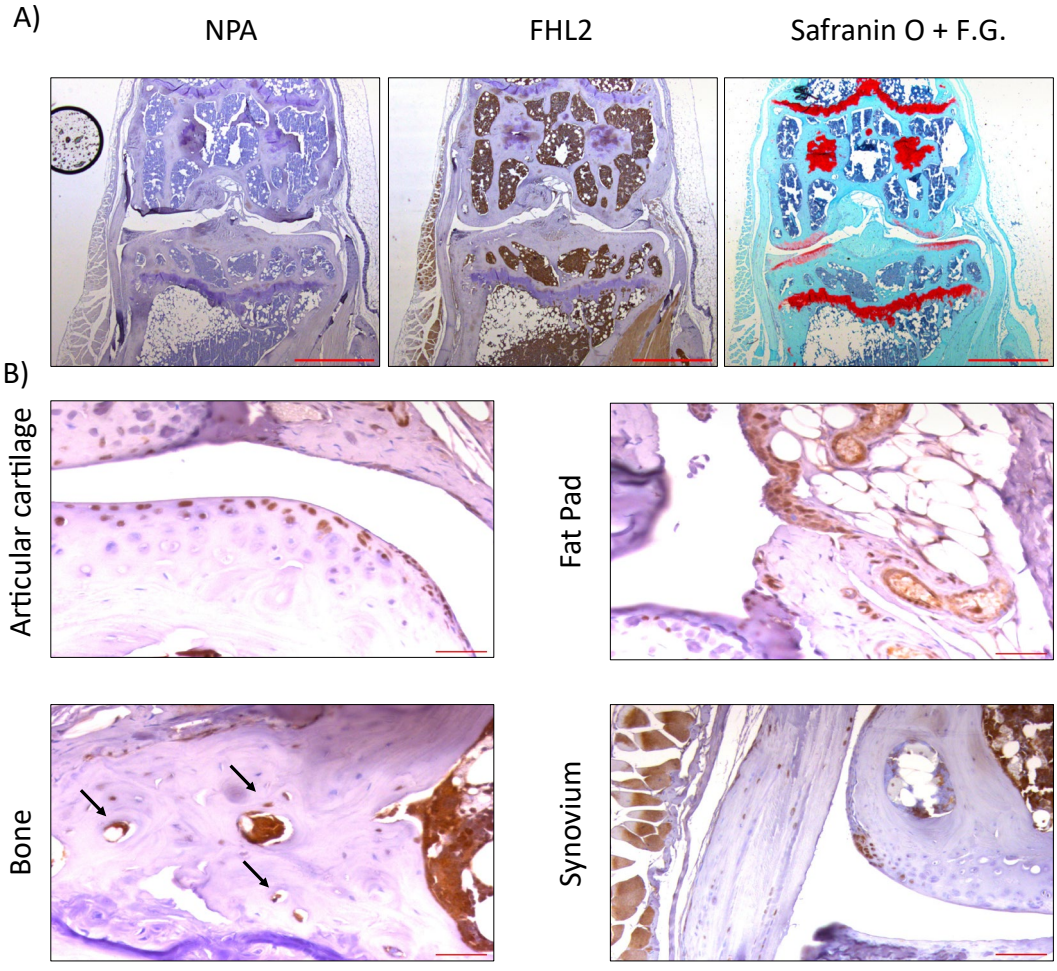


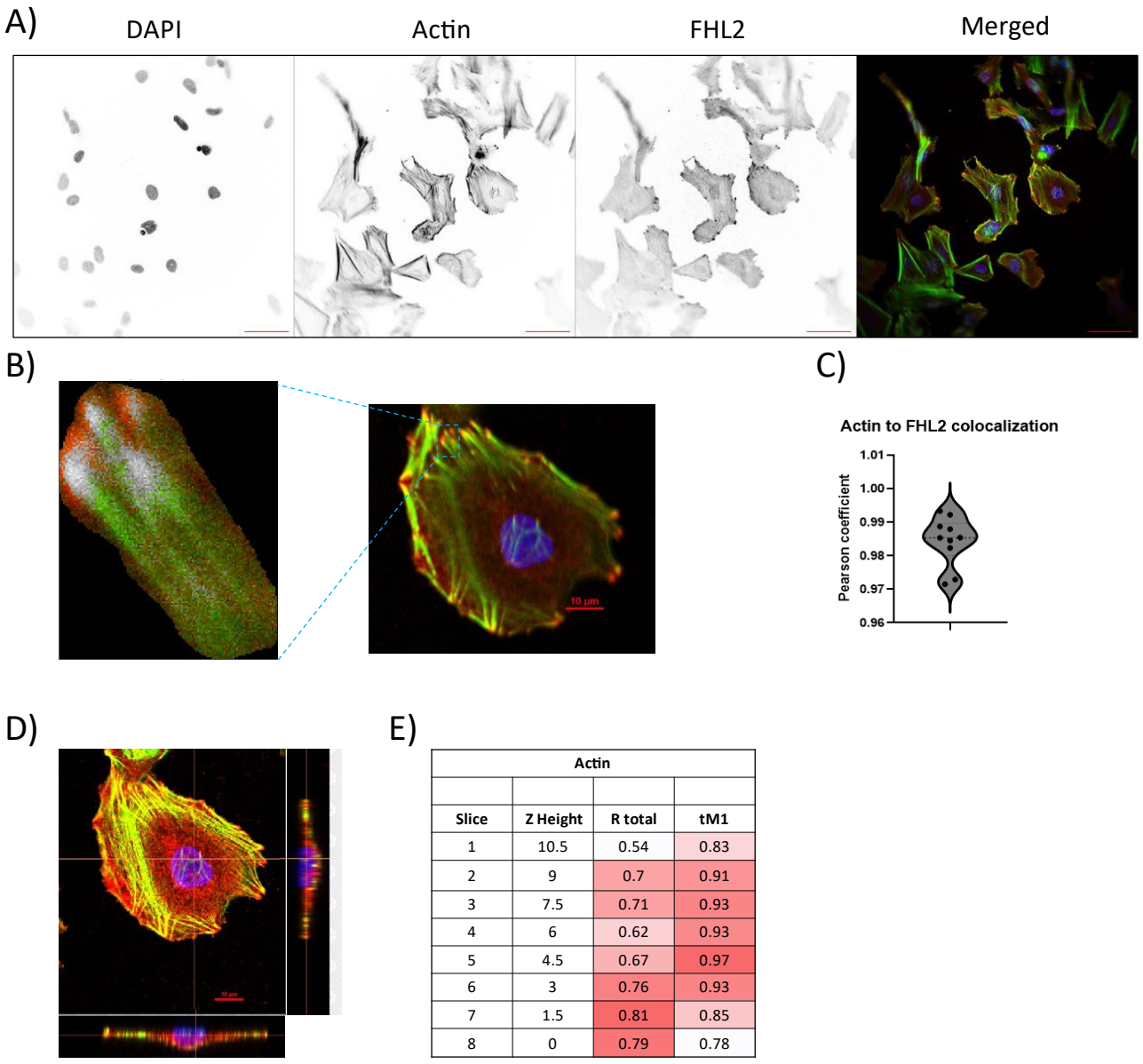
Figure 6. 19. Immunohistochemical investigation of FHL2 distribution at 12 months of age in the mouse A) Whole joint view with NPA, FHL2 and Safranin O + Fast Green staining; 500 µm scale bar B) Figure. Distribution of FHL2 in the tissues of A) fat pad, B) synovium and C) bone (arrows mark the position of osteocytes) with the respective NPA controls. Scale bar represents 50 µm.

6.3.5 Co-localisation of FHL2 with mitochondria and actin filaments

From the determination of the intra-articular pattern of FHL2, further investigation focused to determine the intracellular localisation of FHL2 in SW1353 chondrosarcoma cells together with the possible co-localisation to other cellular components. Previous reports on the localisation of intracellular FHL2 are conflicting, with several reports (Canault *et al.*, 2006; Tran *et al.*, 2016; Liu *et al.*, 2019) stating cytoplasmic, nuclear localisation. Based on the RNA-seq gene ontology analysis (section 6.3.1) and previous reports (Sun *et al.*, 2020; Basu *et al.*, 2021) I examined the co-localisation between FHL2 and either actin filament or mitochondria in SW1353 cells.

Co-localisation between actin and FHL2 was analysed using confocal microscopy and the Pearson correlation coefficient method to correlate the pixel based signal from the two proteins. FHL2 was mainly seen to co-localize with actin filaments near the periphery of the cell membrane scoring a Pearson correlation coefficient ~1 (Figure 6. 20). FHL2 is observed to have a higher co-localisation near the end of the actin filament at potentially focal adhesion points. Observing the correlation, analysis of the z stack images indicates that the main correlation is present near the bottom of the cell from 0-3 µm.

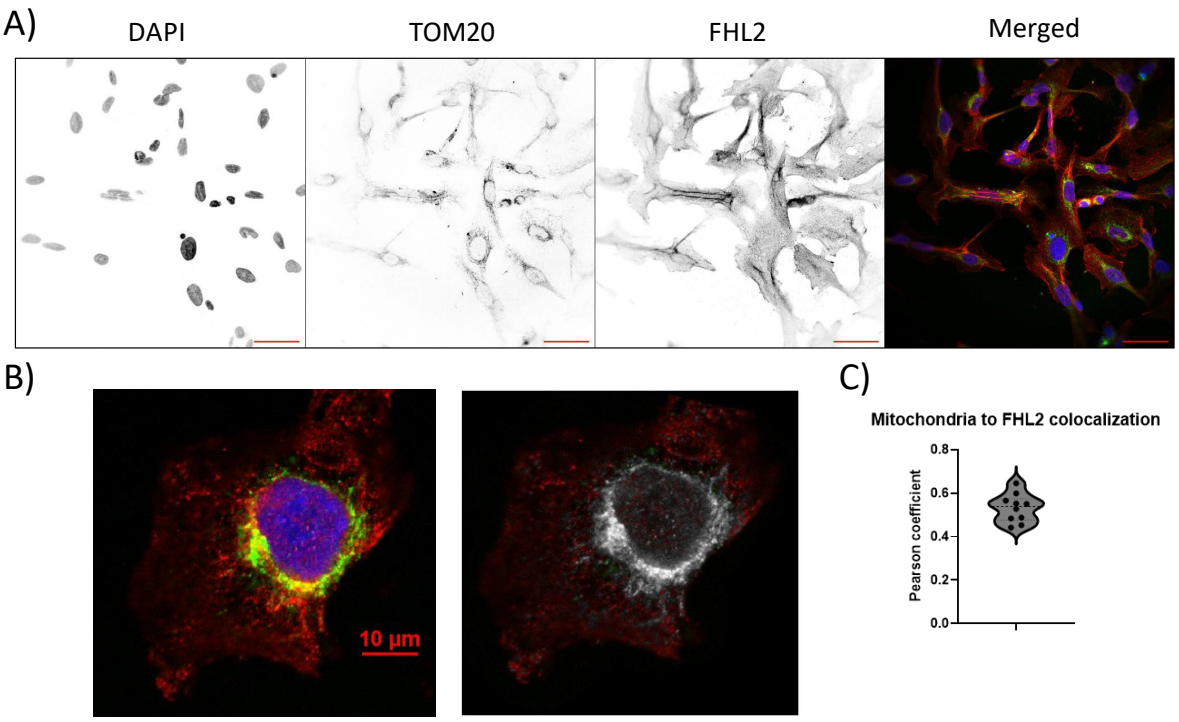
Co-localisation between mitochondria (TOM20 marker) and FHL2 (Figure 6. 21) was mainly seen to co-localize around the peri-nuclear region scoring a Pearson correlation coefficient ~0.5. Observing the correlation, at different heights of the cell reveals that the main correlation is present near the middle of the cell from 0-4.5 um. The overall correlation coefficient is lower for the mitochondria than the one observed for the actin filaments suggesting FHL2 presence in the peri-nuclear region could co-localise with other cellular components.



Actin			
Slice	Z Height	R total	tM1
1	10.5	0.54	0.83
2	9	0.7	0.91
3	7.5	0.71	0.93
4	6	0.62	0.93
5	4.5	0.67	0.97
6	3	0.76	0.93
7	1.5	0.81	0.85
8	0	0.79	0.78

Figure 6. 20. Co-localisation analysis between FHL2 and actin filaments within the cell. A) Fluorescent micrograph showing representative view split in the channels for DAPI, Actin, FHL2 and a merged view; 50 µm scale bar B) Co-localisation assay highlighted for the highest co-localized pixels represented in white with actin filament staining represented in green and FHL2 staining in red; 10 µm scale bar. C) Pearson correlation coefficient for the co-localisation analysis between Actin and FHL2 staining. D) Z height slice view for the X and Y axis of a representative cell showing the distribution of the FHL2 and Actin staining; 10 µm scale bar. E) Summary of the co-localisation analysis as a function of height (µm) with the

Pearson correlation coefficient (R total) and Mander's overlap coefficient (tM1) at each slice.



E)

Mitochondria			
Slice	Z Height	R total	tM1
1	10.5	0.06	0.2
2	9	0.11	0.16
3	7.5	0.19	0.24
4	6	0.37	0.37
5	4.5	0.5	0.56
6	3	0.54	0.51
7	1.5	0.49	0.28
8	0	0.47	0.12

Figure 6. 21. Co-localisation analysis between FHL2 and TOM20 (mitochondria) within the cell. A) Fluorescent micrograph showing representative view split in the channels for DAPI, TOM20, FHL2 and a merged view; 50 µm scale bar B) Co-localisation assay highlighted for the highest co-localized pixels represented in white with TOM20 staining represented in green and FHL2 distribution in red; 10 µm scale bar. C) Pearson correlation coefficient for the colocalisation analysis between TOM20 and FHL2 staining. D) Z height slice view for the X and Y axis of a representative cell showing the distribution of the FHL2 and TOM20 staining; 10 µm scale bar. E) Summary of the co-localisation analysis as a function of height with the Pearson correlation coefficient (R total) and Mander's overlap coefficient at each slice.

6.3.6 IL-1 impact on FHL2 expression and FHL2 nuclear translocation.

Another focus point is the understanding of the factors that can impact the expression of *FHL2* and whether these can activate FHL2 to translocate into the nucleus to act as a transcription co-factor.

Skeletalvis is a data portal that provides skeletal biology transcriptomic datasets and contains extensive information on the impact of stimuli or gene perturbation experiments on skeletal tissue related cell lines (Soul *et al.*, 2019). The database was consulted to understand whether factors such as Interleukin-1 (IL-1), Tumour necrosis factor (TNF) and Transforming Growth factor beta 1 (TGFB1) can impact the expression of *FHL2*. Several datasets were found where either synovial fibroblasts or chondrocytes were stimulated with IL-1, TNF or TGFB1 and the impact observed was that of reducing the expression of *FHL2* at a statistically significant level (Table 6. 1).

Following this investigation, SW1353 cells were stimulated with IL-1 (Figure 6. 22 A) to understand whether the expression of *FHL2* can be impacted by this treatment. Moreover, RNA from H.A.C. treated with IL-1 (Figure 6. 22 B) was provided by Dr. Louise Reynard and Dr. William Thompson to check for the expression of *FHL2.* In SW1353 cells, IL-1 decreased the expression of *FHL2* by 22% ($p<0.05$) whereas in the H.A.C. a trend for decreasing expression was observed however not for all the patient samples. Certain patients (6/17) responded quite robustly to the treatment resulting in decreased *FHL2* expression whereas the other 11/17 had only a small or no decrease in *FHL2* expression levels. In spite of this, the paired t-test analysis on the H.A.C. saw statistically significant relationship at $p<0.01$.

To investigate the impact of IL-1, SW1353 cells were stimulated with IL-1 with the translocation of FHL2 tracked by fluorescent microscopy across a time interval of one hour with data points at 20-minute intervals (Figure 6. 23). The movement of FHL2 to the nucleus

from the cytoplasmic compartment is found to be statistically significant after 40 minutes of IL-1 stimulation with 20% of the cells showing nuclear FHL2 ($p<0.001$), which increased after 60 minutes to 33% of the cells ($p<0.0001$). Investigating the region from which FHL2 translocated indicated or the actin filaments (Figure 6. 24) the co-localisation Pearson coefficient for actin control (mean= 0.98, S.D. = 0.007) vs IL-1 (mean= 0.98, S.D. = 0.009) stimulated did not show any change in significant change with $p = 0.22$, whereas for the correlation with the mitochondria (Figure 6. 25), the Pearson coefficient control (mean= 0.53, S.D. = 0.06) vs IL-1 stimulated (mean= 0.41, S.D. = 0.05) ($p <0.001$) indicating that the translocation happened from this region.

Table 6. 1. Skeletalvis investigation on stimuli that impact FHL2 expression. Significant p values are highlighted in red. The cell types investigated are synovial fibroblast (i.e. components of the synovial membrane) and articular chondrocytes.

Skeletalvis investigations			***FHL2* expression**	
Dataset	**Cell type**	**Condition**	**Log2 F.C.**	**Adj. p value**
GSE58203	syn fibroblast	Control vs IL1b	-1.2	3.00E-13
GSE15615	syn fibroblast	Control vs TNF stim +IL1	-1.2	2.00E-02
GSE13837	syn fibroblast	Control vs TNF	-0.9	4.00E-02
		Control vs TGFB1	-1.4	1.00E-03
GSE68428	chondrocytes	Control vs IL1	-1.1	7.00E-04
GSE75181	chondrocytes	Control vs IL1b	-1.9	7.00E-16

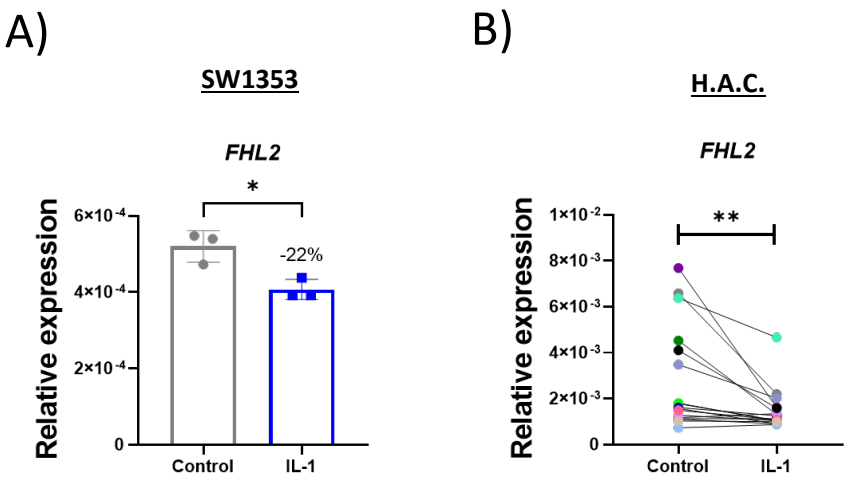


Figure 6. 22. IL-1 stimulation and impact on FHL2 gene expression analysis using RT-PCR in A) SW1353 cells IL-1 stimulation for 24 hours at a concentration of 0.5 ng/ml ;paired t-test *p<0.05; B) Human articular chondrocytes (H.A.C.) IL-1 stimulation for 22 days at a concentration of 0.1 ng/ml ;each patient sample is represented by an unique coloured dot, paired t-test **p<0.01. Relative expression is based on the expression of the housekeeper genes *GAPDH*, *HPRT1* and *18S.*

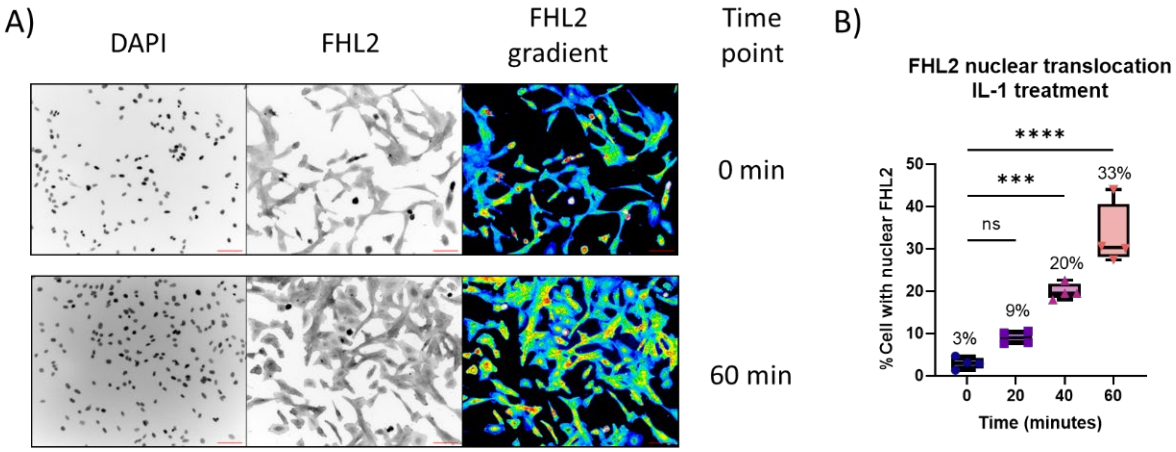


Figure 6. 23. Impact of IL-1 stimulation on FHL2 translocation in SW1353 cells. A) Fluorescent micrograph showing DAPI, FHL2 and a FHL2 signal heatmap (low signal blue, high signal-red) under control conditions and after 60 minutes of IL1 stimulation; 50 µm scale bar B) Quantification of FHL2 nuclear translocation effect at different time points after IL-1 stimulation. One-way ANOVA analysis with Dunnette correction method, ***p<0.001, ****p<0.0001.

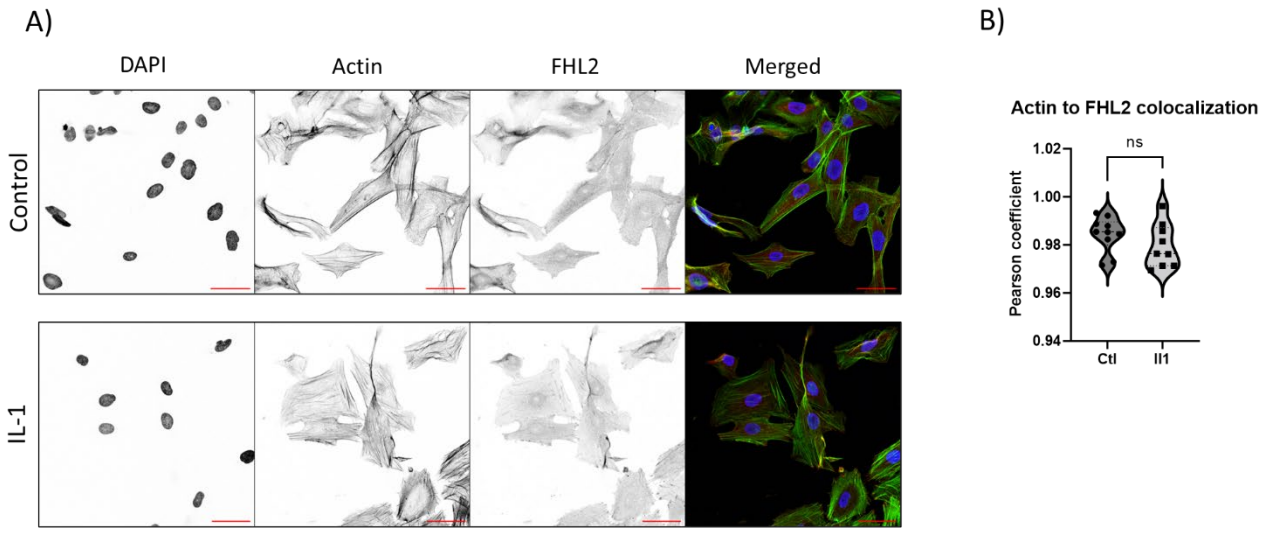


Figure 6. 24. Impact of IL-1 stimulation on FHL2 translocation and colocalization with actin filaments in SW1353 cells. A) Fluorescent micrograph showing DAPI, actin, FHL2 and a merged image for the co-localisation analysis under control and IL-1 stimulation conditions at a concentration of 0.5ng/ml for 60 minutes; 50 µm scale bar B) Pearson correlation coefficient analysis between the localisation of the FHL2 signal and that of the actin signal under control and IL-1 stimulation conditions, unpaired t-test analysis.

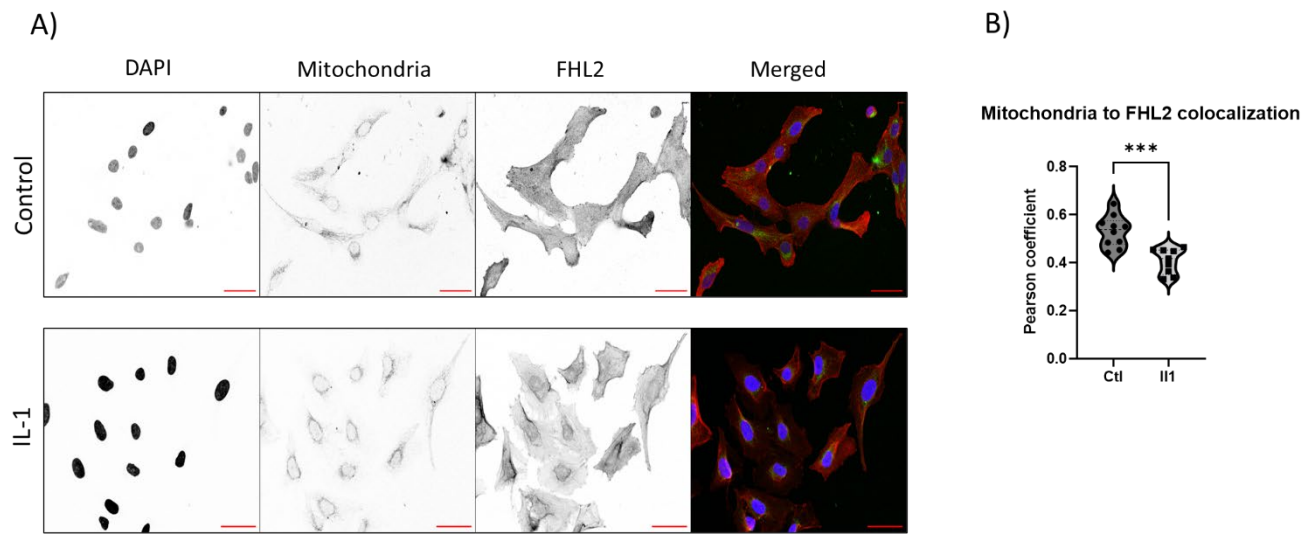


Figure 6. 25. Impact of IL-1 stimulation on FHL2 translocation and colocalization with mitochondria in SW1353 cells. A) Fluorescent micrograph showing DAPI, mitochondria, FHL2 and a merged image for the co-localisation analysis under control and IL-1 stimulation conditions at a concentration of 0.5ng/ml for 60 minutes; 50 µm scale bar B) Pearson correlation coefficient analysis between the localisation of the FHL2 signal and that of the mitochondrial signal under control and IL-1 stimulation conditions, unpaired t-test ***p<0.001.

6.4 Discussion

To understand more about the function of FHL2, knockdown and overexpression experiments were performed where two knock down lines and one overexpression line were produced (Figure 6. 1). The results of this modulation of *FHL2,* either by overexpression or knockdown, revealed several hundred genes that are changed in expression as a consequence. Key pathways for the knock down of *FHL2* that seem to be impacted according to gene ontology analysis (Figure 6. 3 & Figure 6. 4) are pathways relating to "cellular proliferation", "extracellular matrix synthesis" and "extracellular matrix organisation" as well as pathways relating to the cellular metabolism of cellular components such as sterols, cholesterol and lipids. There were some differences in the two knockdown lines referring to the genes that are differentially expressed with one cause for this being the technique and the fact that using two gRNAs to target the region also introduced different off targets this being still a big limitation of the technique. Additionally, they had different levels of *FHL2* knockdown where the number of impacted genes is greater with the higher knockdown line.

The overexpression of *FHL2* induced transcriptomic changes relating to pathways for the "metabolic processing of organic acids", "cellular motility", "cell adhesion and "extracellular matrix organisation" (Figure 6. 6). In the overexpression line FHL2 is fused to a GFP tag, this could interfere with some of its functions as suggested from other works (Weill *et al.*, 2019), although no issues were previously reported in the literature for this construct (Nakazawa *et al.*, 2016; Clemente-Olivo *et al.*, 2023). Such pathway changes give insight into the possible functionality of FHL2 and one functional assessment on cellular proliferation indicated that FHL2 could be involved in controlling cellular proliferation as it was also suggested previously (Hua *et al.*, 2016; Wu *et al.*, 2016; Jin and Liu, 2018). For the knock down experiments, the BrdU and WST-1 assays were in agreement that the knock down causes an increase in cellular proliferation which was dependent the level of FHL2 depletion (Figure 6. 12 A & B). Interestingly, for the overexpression line the two assays are not in agreement where the BrdU indicates a decrease in proliferation as measured by the integration of BrdU (Figure 6. 12 C) whereas the WST-1 assay reports an increase in proliferation (Figure 6. 12 D). One possible reason for the discrepancy could be in the

mechanism of the two assays where the BrdU relies on the integration of the BrdU nucleotide at cellular division and subsequent detection of this and the WST-1 assay is based on the cleavage of the dye molecule by mitochondrial dehydrogenases. For the WST-1 assay an assumption is made that the treatment or condition does not alter mitochondrial metabolism and therefore the assay readings would reflect cellular numbers. However, since the RNA-sequencing results indicate changes in metabolic pathways with some intersecting with mitochondrial metabolism changes, this could be a cause for interference where the measurements would not reflect only cellular proliferation changes but also the metabolic changes. Therefore, the BrdU results could be considered more indicative of the proliferation changes with the WST-1 assay being a reflection of the metabolic changes as well. Previously it was reported in the cell lines such as COV434 (human ovarian granulosa cells) (Hua *et al.*, 2016), TSCC (tongue squamous cell carcinoma) (Wang *et al.*, 2016), MEF (mouse embryonic fibroblasts) (Labalette *et al.*, 2008a), K562 (lymphoblast cell line) (Lu *et al.*, 2017) that the proliferation rate directly correlates with *FHL2* levels where increased levels of *FHL2* results in increased cellular proliferation rate which directly contradicts the findings from the current BrdU experiments in SW1353 cells. The reason for why SW1353 cells responded differently to previous reports for the impact on cell proliferation effect is unknown, however current experiments would indicate that re-introduction of *FHL2* in SW1353 cells would decrease this impact.

In the differential gene expression analysis revealed that not many genes seem to follow opposing direction of change between the knock-down and overexpression. Such assessment is limited as some genes might not be responsive to increased levels of FHL2 as well as some becoming responsive due to the much higher levels forcing an abnormal cellular response. Other pathways could have other compensatory proteins that take over some of the functions of FHL2 (e.g. other members of the FHL family) with this further increasing the complexity of understanding the functionality of this protein. An experiment to re-introduce *FHL2* was performed where on the cell line with the highest knockdown level of *FHL2* a FHL2-GFP vector was introduced into the cells (Figure 6. 13). Through the attempt of the re-introduction of *FHL2* the construct was successfully detected by qRT-PCR (Figure 6. 13 A), western blotting (Figure 6. 13 B) and cell proliferation was measured to ascertain the impact of this. Through this restoration, certain characteristic can be reversed

such as cellular proliferation and selected transcriptomic changes (Figure 6. 14), however no large-scale transcriptomic investigation was performed to examine the extent of the reversal. Such an investigation however remains an interesting point for future research as well as to study if through additional overexpression vectors the initial knockdown impact can be reversed and in what areas predominantly. With the *FHL2* overexpression, *FOS* expression was found to increase (Figure 6. 7), whereas *FOS* was observed to decrease in expression upon *FHL2* (Figure 6. 5) and moreover the effect of the decrease seemed proportional to the *FHL2* levels. Partial restoration of *FHL2* experiment also showed that levels can be partially restored after the knockdown associated decrease (Figure 6. 14 A). This would suggest that the expression of *FOS* is linked to the expression of *FHL2. FOS* is a member of the AP-1 transcription factor complex with critical functions in regulating signal transduction and cellular proliferation (Hess *et al.*, 2004; Fayos *et al.*, 2018). Due to the observed effects on cellular proliferation when *FHL2* is modulated it could be that *FHL2* is a key factor through which *FOS* and the AP-1 complex control the proliferation rate resulting in this phenotype. Many other pathways and genes involved in the control of cell proliferation seem to be linked with *FHL2* and have been reported including AKT1 (Hua *et al.*, 2016), Cyclin D1 (Labalette *et al.*, 2008a) and beta-catenin (Cai *et al.*, 2018). Other pathways were also observed as being changed and the direct impact of the modulation is difficult to ascertain from the indirect consequences of this change. In the literature the pathways described to be impacted by *FHL2* knockdown include cell cycle, lipid metabolism, cellular organisation and assembly and protein synthesis (Labalette *et al.*, 2008a; Kurakula *et al.*, 2015) most of these also being found in the current study.

FHL2 is potentially expressed ubiquitously across several tissues in the human body (Figure 6. 15) according to the GTEX dataset although no cell type specific data can be found for these and additionally there is not much data given for the articular joint tissues in GTEX. Looking at the distribution of *FHL2* expression in the mouse articular joint it seems to be present in most tissues at various levels in the mouse model, with the presence of the protein changing based on the age of the sample, albeit with only one sample representative per age group (Figure 6. 16-20). The characteristic across the age groups would be the presence of FHL2 positive cells at the periphery of the articular joint as well as strong presence in the tissues of skeletal muscle, bone marrow and bone osteocytes. The

work with the DMM four months old sample (Figure 6. 18) served to look whether the procedure could impact the expression pattern of *FHL2* where no such observation could be made. The RNA-sequencing data provided (Figure 6. 18 D & E) also would suggest that the procedure does not seem to influence any significant changes in its expression and with the OA RNA-sequencing data for control NOF samples and OA hip again changes in the expression of *FHL2* cannot be observed. This could indicate that either this gene is not involved in any pathways associated with OA or is under a compensatory regulatory mechanism. Since there is not much data available for the role of FHL2 in the articular joint this investigation serves as preliminary data for further analysis into the cell specific functionality. Previously there was some work into the understanding of tissue damage resolution and FHL2 activity in the articular joint where an increase in *FHL2* expression was observed during inflammatory arthritis, with the conclusion that FHL2 could have a more protective and pro-regenerative role in the articular joint (Wixler *et al.*, 2015). FHL2 activity has been linked with wound healing where its expression being directly linked with the different stages of wound healing. Stimulants of *FHL2* include TGF-beta and the higher expression of FHL2 to an abnormal level was associated with “overhealed” wounds and increased tumorigenesis predisposition (Hinz, 2015). Inhibitors for its expression include the mentioned Il-1beta and TNF-alpha and with sustained down-regulation results in higher expression of enzyme remodelers such as MMPs with the transition into the chronic wound stage (Wixler *et al.*, 2015).

Further investigation to determine the functionality of FHL2 in the context of a chondrocytes like cell were done through co-localisation experiments. As targets for the investigation of the co-localisation the RNA-seq data had flagged that there were modifications in pathways relating to focal adhesion points and actin filament processing and therefore the co-localisation with actin was firstly attempted. The result of this analysis suggests that indeed FHL2 seems to have a high degree of co-localisation with actin filaments (Figure 6. 20) and indeed the highest correlation seems to be found at focal adhesion points, with this being previously reported in the literature for Paxilin and Actin filaments at focal adhesion points (Nakazawa *et al.*, 2016). Other pathways such as the changes in cellular and possible mitochondrial metabolism pointed towards mitochondria as another target for the co-localisation analysis. For these experiments, the mitochondria

indeed seem to also co-localise with the FHL2 present at the peri-nuclear region in SW1353 (Figure 6. 21) also confirming previous observation in COS-7 cells (Basu *et al.*, 2021). The default location of FHL2 in SW1353 cells seems to be mainly cytoplasmic with a high degree of co-localisation with actin filaments and mitochondria with somewhat absent signal in the nuclear region. Z stack height analysis reveals additional details on the extent of the co-localisation with these components, where for the actin filaments it would seem mainly at the bottom of the cell to be occurring and for the mitochondria the extent is from the bottom to the middle point of the peri-nuclear region. The presence of FHL2 at these sites might suggest that it could be performing a function at this location especially at the focal adhesion points in addition to its reported function as a transcription co-factor. The literature reports high associations of FHL2 with many proteins such as beta-catenin (Cai *et al.*, 2018), ADAM-17 (Verset *et al.*, 2017), RNA Polymerase II and Paxillin (Nakazawa *et al.*, 2016), with such findings adding to the complexity of the diverse functions of FHL2. Moreover, there could be other compartments or complexes that could co-localise with FHL2, and this remain as ideas for future work.

Cytoplasmic un-phosphorylated FHL2 is known to react to certain stimuli such as IL-1 (Cao *et al.*, 2015), Blebbistatin (myosin inhibitor), Y-27632 (Rho signaling inhibitor), Cytochalasin D (inhibitor of actin polymerisation) and Wiskostatin (inhibitor of actin polymerisation) (Nakazawa *et al.*, 2016), where an enzyme catalysed reaction occurs at focal adhesion points driven by FAK (PTK2) (Tran *et al.*, 2016; Sun *et al.*, 2020). This then phosphorylating FHL2 causing nuclear through a CRM1 dependent transporter (Morlon and Sassone-Corsi, 2003) and once in the nucleus the phosphorylated FHL2 acts as a transcription co-factor for several genes such as *p21*. The interaction influences the expression of *p21* by the stabilisation of the transcription factor complex (Nakazawa *et al.*, 2016).

The expression of *FHL2* is repressed by the stimulation of IL-1, TNF and TGFB1 according to the data from Skeletalvis (Table 6. 1) and further confirmed by the treatment of SW1353 and primary human chondrocytes with IL-1 (Figure 6. 22). This reaction to the treatment further was investigated based on the report of IL-1 causing translocation changes to FHL2 (Joos *et al.*, 2008). Such movement was investigated at different time points, and it confirmed that indeed FHL2 seems to respond to IL-1 treatment by translocating from the cytoplasm to the nucleus (Figure 6. 23) and moreover one of the compartments from which

FHL2 is translocating seems to be the mitochondrial peri-nuclear region (Figure 6. 25) where a significant correlation was found and not from the actin filaments (Figure 6. 24). This could indicate that following IL-1 stimulation the previously described phosphorylation of FHL2 catalysed by FAK at the focal adhesion point might not be the current pathway followed as the FHL2 located on the actin filaments at the focal adhesion points had no significant change in the correlation coefficient. The movement of the FHL2 from the peri-nuclear region would indicate that there could be a protein population of FHL2 that is more readily activated by some stimuli and then the translocation to the nucleus occurs. Future work for this would investigate additional comportments and complexes that FHL2 could form in the cytoplasm as well as the ability of different stimuli to possible cause translocation of FHL2 from different comportments.

Chapter 7: Discussion

7.1 Summary and main findings

This research project was based on a previous study within the group that identified age-associated DNA methylation changes at 716 age-CpG sites in the cartilage tissues using an Illumina 450k DNA methylation array. The majority of these sites were found in regions with a promoter chromatin state (~70%) with 9% in regions with an enhancer signature based on E049 chromatin state data. The aim of this thesis was to validate selected cartilage age-CpGs, examine if these age-related DNA methylation changes occurred in other joint tissues and investigate the transcriptional consequences of these methylation changes.

Chapter 3 focussed on validating these previously identified age-CpGs. In order to validate these cartilage age-CpGs, originally *in silico* replication analysis was performed using two additional cartilage DNA methylation datasets. These 716 age-CpGs were further studied in a validation cohort where DNA methylation was measured using the newer Illumina EPIC array, where about 11% of the probes identified as age-CpGs in the discovery cohort were not present on this new array and from the remaining available probes only 473 CpGs (74%) were found to significantly changes with age. Another analysis was performed using a smaller cohort of cartilage samples on the Illumina 450k array and the comparison indicated not all the initial age-related CpG changes were observed to occur with only 51% of the initial age-CpG. Further investigation was also performed on publicly available datasets (Hannum *et al.*, 2013; Zykovich *et al.*, 2014; Vandiver *et al.*, 2015b; Lien *et al.*, 2018; Panjarian *et al.*, 2021a; Kurokawa *et al.*, 2022) that used the Illumina 450k array in other musculoskeletal tissues (e.g., bone, skeletal muscle and tendon) and non-musculoskeletal origin (e.g.,whole blood, liver, breast) where the results indicated that more than 50% of the 716 cartilage age-CpGs were also observed in the other tissues, these age-CpGs not being specific cartilage age-CpGs. The results of the Illumina 450k have been reported to potentially have poor correlation when compared to other platforms (Logue *et al.*, 2017) and to techniques such as bisulphite sequences (Cheung *et al.*, 2020), therefore follow-up

analysis focused on validating some of the cartilage age-CpGs observed using bisulphite pyrosequencing.

From the list of 716 cartilage age-CpGs, 15 age-CpGs were selected for bisulphite pyrosequencing based validation in cartilage, fat pad and synovium samples from OA knee patients. An additional 26 CpG sites that were neighbouring the target age-CpGs were also included in the pyrosequencing assay this resulting in 41 CpGs studied in total. The investigation by pyrosequencing looked in a new cohort of cartilage samples and found that DNA methylation levels at nine out of the 15 age-CpGs were observed to change with age with linear regression analysis and moreover 11 of the additional CpGs significantly correlated with age. Following this, the analysis was performed in other articular tissues such as synovium where 20 age-CpGs have been observed and in the fat pad tissue also 20 age-CpGs with some similarities to the cartilage age-CpGs. Across all the tissues investigated, 17 age-CpGs were observed to occur in all the articular tissues. For the regions that presented multiple neighbouring CpGs (e.g. *FHL2* and *ELOVL2* promoter age-CpGs) further analysis was performed to study the correlation of DNA methylation levels between neighbouring CpGs and if these are linked with the tissue of origin. The analysis indicated that there is a high degree of correlation between the CpGs contained in a region and a pattern of clusters were observed for this as well. The analysis also indicated that the cartilage tissue shows hypomethylation in comparison to the other tissues and that synovium tissue had no distinguishable methylation differences to the fat pad tissue at multiple sites. The DNA methylation changes would indicate that the tissues of the articular joint all show age-related changes with the consequences of these on tissue functionality being unknown. The studied age-CpG sites at the promoter region of *FHL2* and *ELOVL2* are one of the most reported age-CpG sites in the literature across various tissues such as saliva, buccal epithelia, dermal fibroblasts, whole blood, pancreas and brain (Garagnani *et al.*, 2012; Hannum *et al.*, 2013; Vandiver *et al.*, 2015a; Bacos *et al.*, 2016; Bacalini *et al.*, 2017; Gasparoni *et al.*, 2018; Jung *et al.*, 2019), yet few investigations have followed on the potential impact on gene expression. Due to the CpG DNA methylation changes associated with age an epigenetic clock tool was developed using multivariate regression analysis on the methylation values of the top 10 most significant age-CpGs together with the chronological age of the samples to generate a tool that can determine a type of biological

age. Additionally, another clock was generated only using data from the age-CpGs located in the promoter of *FHL2* as to make the clock only based on two pyrosequencing assays to increase the simplicity of using this. Both of these clocks were trained on samples such as cartilage, synovium samples + fat pad samples as well as a combined dataset with all tissues where a maximum age prediction error of ~10 years was observed. Similar clocks using pyrosequencing data were developed for mouse tissues(Han *et al.*, 2018; Han *et al.*, 2020) and for forensic applications (Paparazzo *et al.*, 2023)and the have proved powerful tools for the study of the epigenetic biological age and age unknown samples respectively.

In chapter 4, to study what impact DNA methylation can have on gene expression, 5-Aza 2'deoxycytidine a global demethylating agent was used to treat the cell lines SW1353 and TC28a2 as well as for the isolated primary human articular chondrocytes. These experiments revealed altered expression upon 5-Aza treatment of several predicted target genes of the age-CpGs, with an increase in expression following the demethylation treatment. Further study on the impact of DNA methylation as well as the promoter or enhancer strength of the regions. The regions containing the age-CpGs were cloned in CpG free luciferase vectors in either the position of promoters or enhancers respectively after which *in vitro* DNA methylation was performed. All the studied regions displayed reduced luciferase activity with the increase in DNA methylation suggesting DNA methylation is associated with the repression of the target gene expression. Further investigation for the impact of DNA methylation at the regions of interest was attempted by using a CRISPR-Cas9 tool based on the dCas9 fusion to epigenetic modifiers such as DNMT3a or TET1cd. Production of dCas9-DNMT3a and dCas9-TET1cd stable expressing cells was successful, however when tested using gRNAs that were previously used in the literature (Choudhury *et al.*, 2016; Pflueger *et al.*, 2018a) these failed to induce any DNA methylation change at the target regions. Several attempts have been made, however this still produced no DNA methylation changes at the expected sites. Following the unsuccessful attempt to modulate DNA methylation another attempt was made using a CRISPR-Cas9 deletion system. This system was used to delete the regions containing the age-CpGs to observe whether this might have an impact on target gene expression. CRISPR-Cas9 region deletions were successful in achieving over 50% deletion at the regions of interest, however only the regions of *FHL2* promoter and *ELOVL2* promoter produced changes in their target's gene expression. Deletion of the age-CpG

region of the *FHL2* promoter resulted in decreased expression of *FHL2* and in contrast the deletion of the age-CpG region of *ELOVL2* resulted in increased expression of *ELOVL2*. This would indicate that these regions are important for the functionality of their respective promoter region, where the *FHL2* age-CpG region having an activatory role for the expression of *FHL2* and the *ELOVL2* age-CpG region having an inhibitory one for the expression of *ELOVL2*. Next, gene expression was measured in the same samples for which DNA methylation data was available and the relationship between gene expression and DNA methylation was explored. This indicated mainly a repressive impact of DNA methylation at some CpG sites such as cg06639320 and *FHL2* gene however such link was not observed for all the tissues studied despite the age-related DNA methylation changes that were occurring in these. A significant correlation with DNA methylation was observed for the following genes of *ELOVL2* in synovium and fat pad, *FHL2* in cartilage and synovium, *PCDHB1/PCDHB6* in synovium and fat pad respectively, *ZNF577* in synovium and fat pad. This observation follows the predicted pattern from the previous 5-Aza stimulation experiment and the luciferase assay where DNA methylation was associated with target gene repression. Interestingly, since not all tissues show a significant link between gene expression and DNA methylation this would suggest that there are other factors that further contribute towards the establishment of such an effect as seen for the cell line experiments.

In chapter 5, the region containing the age-CpGs within the promoter of *FHL2* was further analysed to examine the transcription factors binding to this region and what impact DNA methylation has. Analysis was performed in the cell lines of SW1353, TC28a2, U2OS and MDA-MB-23 where a total of 34 complexes were observed. The data suggested that most factors bind at the age-CpG sites as well as that DNA methylation increases by several fold the binding affinity. Sp1 and Sp3 were identified as components of three complexes that bind the region containing three age-CpGs with these being present across the cell nuclear extracts tested, suggesting that this binding is not cell line specific. Additional complexes were present at this region with some of them being unique for some cell lines, however these remained unidentified. The Sp family of transcription factors are known to exhibit both repressive and enhancing effects at the target gene and are known to be implicated in cellular proliferation control (Li and Davie, 2010), where their presence at the promoter of *FHL2* suggest how the gene's expression can be modulated by these factors. Since the DNA

methylation changes were not always linked with *FHL2* expression changes in the studied tissues it would suggest that other factors are needed for the effect, and these could be tissue specific transcription factors that facilitate the link between DNA methylation and *FHL2* expression. In the literature one such study in the pancreas tissue correlated the increase in DNA methylation of the *FHL2* age-CpG sites with increased expression of *FHL2* (Bacos *et al.*, 2016), this being similar to what was observed for the bone tissues in the analysis done in the current study and in contrast to the results from cartilage and synovium where DNA methylation was observed to be repressive to *FHL2* expression.

The function of *FHL2* is not completely known especially in the context of musculoskeletal tissues and further understand FHL2's role in chondrocyte like cells, CRISPR-Cas9 technology was used in chapter 6 to knock down *FHL2* by targeting the common exon of the *FHL2* isoforms. This produced several cell lines with different levels of knockdown of *FHL2.* To complement the knockdown study an overexpression of *FHL2* was performed using a plasmid vector to further understand the role of this in SW1353 cells. Gene ontology analysis revealed that genes dysregulated upon FHL2 loss were part of pathways relating to cellular proliferation, extracellular matrix synthesis and cellular metabolism. The impact on proliferation being tested using two cell proliferation assays that confirmed that for the knockdown proliferation is increased. For the *FHL2-GFP* overexpression line, gene ontology analysis indicated that impacted pathways are involved in cell adhesion, cell motility and extra cellular matrix organisation control. Proliferation was also observed to be decreased for this line, this being opposite to the increase associated with the knockdown. Following the results from the RNA-sequencing, a rescue experiment was performed where *FHL2* was reintroduced as an *FHL2-GFP* vector in the cell line that presented the highest level of *FHL2* knockdown. The impact of the restoration was a reduction in cellular proliferation according to the BrdU assay, this indicating that some of the induced changes could be recovered. These results indicated a role for *FHL2* in a chondrocyte-like cell in the control of pathways for cellular proliferation, cytoskeleton, extracellular matrix synthesis and cellular metabolism, where the impact caused by the DNA methylation with age resulting in the potential alteration of these pathways.

The cellular distribution of FHL2 is unknown within the context of the knee articular joint and this was investigated in a mouse knee joint model at various ages using

immunohistochemistry. The analysis had only one sample per group, and this indicated the FHL2 is observed to be strongly present in tissues such as skeletal muscle and bone marrow across the age range, with tissues such as cartilage ageing resulted in the localisation shifting towards the peripheral cartilage region with age progression. Investigating the intra-cellular distribution of FHL2 in SW1353 cells, this was determined to be mainly cytoplasmic and associated with actin filaments and mitochondria. Additionally, Z stack analysis indicated that FHL2 the association with actin filaments occurs near the bottom of the cell and with the mitochondrial around the perinuclear region and from the bottom towards the middle of the cell. Such associations further add evidence to the RNA-sequencing results which suggested that *FHL2* could be linked with the cytoskeleton and mitochondrial metabolism pathways with these results indicating that FHL2 is present at sites such as the actin filaments that are part of the cellular cytoskeleton and mitochondria. Further to this, a recent study had similar findings where they found that FHL2 co-localises with both actin and mitochondria in neuronal cells and that FHL2 functions as an anchoring adaptor for mitochondria to the actin filaments in response to changes in glucose levels (Basu *et al.*, 2021).

The cytokine IL-1 has been used in the literature as an inducer of cartilage degradation and studied as a potential target for OA treatment (Vincent, 2019) and therefore examined whether *FHL2* could be modulated by this. Database searches on Skeletalvis (Soul *et al.*, 2019) indicated that *FHL2* expression could be impacted by this cytokine with further confirmation from the SW1353 and HACs experiments suggesting a repressive effect on gene expression. Further to this, FHL2 translocation to the nucleus was analysed in SW1353 cells that were stimulated with IL-1, where a significant translocation was observed after 40 minutes of stimulation. Additional analysis indicated that FHL2 is observed to translocate from the previously co-localised mitochondrial peri-nuclear region. These observations establish a link between *FHL2* functionality in the context of IL-1 presence in the articular environment.

With the observed correlation that *FHL2* could be impacted in tissues such as cartilage, synovium and bone due to age-related DNA methylation changes, based on the cellular pathways that are changed with the knockdown and overexpression would suggest that

there could be some links with OA-associated changes in the articular tissues. The impact of the altered extracellular matrix synthesis and collagen synthesis pathways directly related with the decrease in cartilage quality that is observed with age progression (Brittberg *et al.*, 2016). *FHL2* expression is also reduced upon IL-1 stimulation as well as causing nuclear translocation effect of FHL2, with IL-1 being a strong inducer of cartilage degradation and synovitis (Vincent, 2019). This would suggest that there could be potential for *FHL2* to be involved is some way with OA progression. However, *FHL2* was not significantly differentially expressed in both the DMM vs control dataset (mouse) and the control (NOF) vs OA hip dataset (human). Further work would be required to understand whether there could be a link with OA progression associated with these age-related changes, however since these DNA methylation changes have been reported for several non-musculoskeletal tissues this would be make FHL2 as a candidate for future study to understand the changes resulted from this.

7.2 Study limitations

7.2.1 Bisulphite pyrosequencing

With bisulphite pyrosequencing only a targeted approach to study DNA methylation at CpG sites can be made with the need for other techniques to study multiple CpGs or genome wide CpG changes. The approach only works once a small list of CpGs has been selected on various criteria (e.g. chromatin state location of CpG) and moreover calibration is necessary in order to get reliable data, this being done per individual CpG site (Moskalev *et al.*, 2011) as was done in this study. Additionally with the reported 5% in DNA methylation resolution limit of the technique (Jing-bin *et al.*, 2014) any subtle changes that are below this threshold will be difficult to detect reliably and study.

7.2.2 Age of patient samples for analysis

Another limitation of the investigation of the DNA methylation changes with age in the articular tissues is the reliance on samples from individuals age >45 years, where the study of the DNA

methylation changes in younger individuals being currently unknown. Therefore, the study of the extent of these age-related changes as well as when these occur is proposed for future work. Although this present a limitation based on the reliance on OA samples, to overcome this issue the acquisition of articular tissue samples from cadaver origin or from amputations could be obtained to provide younger samples for this assessment.

7.2.3 Epigenetic clock

The use of the current design of the epigenetic clock is limited to tissues such as for cartilage, synovium and fat pad samples for which this has been trained. Depending on the application for such a tool there is a relatively high error in age prediction for both clocks generated. Improvements can be made by increasing the sample size of the dataset with this also reducing the age error.

7.2.4 Luciferase limitations

For the investigation on the consequences of the DNA methylation increases the luciferase assay was used to study the strength of promoter and enhancers as well as to determine the impact DNA methylation has on their activity, however the there are several disadvantages to this technique. One limitation was the fact that the full DNA methylation status of the plasmids used had not been fully determined. The methods of evaluating the methylation status using methylation sensitive restriction enzymes relies on certain sites recognised by the enzymes where not all CpG sites within the regions of interest having the required sequence. This results introduces some uncertainty, although the *in vitro* CpG Methyltransferase (M.SssI) enzyme that was used is also used to methylate CpG sites at a genome wide scale with success (Mehrnaz *et al.*, 2005). To mitigate this bisulphite sequencing could have been used to confirm that all the region's CpG sites have been methylated.

Another disadvantage with this method is the fact the plasmid vectors do not have nucleosomes and therefore any interaction of the transcription machinery with histone marks at either the promoter or enhancers cannot be truly replicated making the luciferase activity reads somewhat artificial(Zhu *et al.*, 2016). Additionally, there are certain enhancers that require the presence of certain histone marks such as acetylation at the H3K27 position

for their activity to be observed (Charlet *et al.*, 2016), this could possibly mean that potential enhancers could remain undetected by this method.

7.2.5 CRISPR limitations

Deletion of the CpGs contained in promoter regions is not always a viable strategy to study the CpG region due to their proximity to transcription start sites which could greatly impact gene transcription or lead to gene silencing (Farzadfard *et al.*, 2013; Wu *et al.*, 2014). A further limitation of the CRISPR system is the requirement of the NGG PAM sequence to be available at the target region, where this could in some cases impact the direct targeting of sequences of interest (Xu and Li, 2020; Yang *et al.*, 2021a).

7.2.6 Correlation between gene expression and DNA methylation in tissues

The analysis that correlated gene expression and DNA methylation at the age-CpG sites for matched samples only reflects on a correlation that could occur with the DNA methylation changes in articular tissues. Further analysis is necessary to establish a link between DNA methylation and the impact it has on gene expression with experiments such as the modulation of DNA methylation at the target CpG with subsequent investigation of gene expression changes as well as to determine the necessary transcription factors involved.

7.2.7 EMSA densiometric assay quantification limitations

This was developed to quantify the signal to study binding kinetics based on densiometric analysis. Since the quantification method has been developed in the current work no information is available for the detection threshold limit as well as the possible resolution of such technique. Since the quantification procedure is similar to the one observed for western blotting quantification that uses IR-fluorophores it could be assumed that the signal acquisition will have similar limitations as observed for this other application. Additionally other limitations include the absence of controls or better normalisation procedures that are still to be developed for this, which will result in increased data reliability and reproducibility. The RF standardisation for the migration of the complexes provided reliable data even on EMSA results with un-uniform migration care must be taken not to overrun the free probe signal this being at almost equal migration capacity as the visible dye included in the reaction mixture. Another identified issues with the EMSA is the reliance on

the manufacturers validation of the antibody when it comes to addressing whether a transcription factor is binding to the region of interest. As stated in Chapter 5, a positive control was available for Sp1 and Sp3 for the antibody supershift, however no such control was available for the other antibodies used. Initially the Sp3 antibody received did not produce a supershift reaction and only after informing the manufacturer and receiving a replacement antibody was this able to produce the expected effect. Such reliance increases the difficulty of identification of the transcription factors present.

7.2.8 Immunohistochemistry work on mouse knee joint

The study of the intra-articular localization of FHL2 was done on mouse tissues due to firstly the difficulty of obtaining human tissue during the project and secondly, due to the overall size of the human knee joint in comparison to the mouse this. Further to this, the analysis only used one sample per age group and therefore no statistical analysis could be done, this serving more as a preliminary study. An additional disadvantage of such a study is the fact that it is currently unknown whether mice also have similar age-related changes to the corresponding *Fhl2,* and therefore if any observations made on this tissue could be reflect the human changes. The usefulness of this analysis is the identification of the expression pattern of *Fhl2* in the tissues of the knee joint and how age might impact this.

7.2.9 Proliferation assays

Through the use of the colorimetric WST-1 assay to determine cellular proliferation a potential limit was observed in conditions where mitochondrial metabolism is believed to be impacted by the knock down and overexpression of FHL2. The WST- assays relies on mitochondrial dehydrogenase cleavage of the tetrazolium salt WST-1 to formazan that shows a strong absorbance at 435nm. For this assay an assumption is made that the levels of the enzyme remain stable across the treatment and this would then indicate the activity of the cell and therefore a measure of cell number. However, with mitochondrial metabolism changes this is believed to impact the readings of the WST-1 assay, where it will no longer correspond only to cellular numbers and therefore such analysis would be prone to a false result on the impact of cell proliferation changes. Due to such concern with the WST-1 assay, the BrdU assays was also performed as this is not impacted by these changes

and the proliferation analysis of this would be more indicative of the proliferation changes with the knockdown or overexpression conditions.

7.3 Future research

7.3.1 Age-related DNA methylation changes mechanism

Currently it remains unknown what causes the increase in DNA methylation that is associated with ageing. DNMT3a and DNMT3b are *de novo DNA* methyl transferases however the factors that are bringing them to the flagged CpG sites are unknown. CpG methylation can be influenced by environmental factors such as smoking, diet and pollution, however some of these age-CpG sites have been reported to occur even in isolated African tribal hunter populations (Gopalan *et al.*, 2017). This population has a much different genetical and ecological background than the typical developed nation and the presence of these markers displays how well linked these age-CpGs are with human ageing despite any genetical and environmental diversity. As many age-CpGs are reported for many tissues it is therefore important to study what the consequences of these changes are on the transcriptome, proteome as well as cellular phenotype.

7.3.2 Epigenetic clock

The epigenetic clock developed using the pyrosequencing data can serve as a relative cheap and quick tool to study the epigenetic age of a sample from musculoskeletal tissues however the same methodology can easily be extended for other tissues in order to study the impact of treatments or diseases that can modify the results of the clock as the biological age. The epigenetic clocks as generated in the literature (Hannum *et al.*, 2013; Horvath, 2013) have been used to study the impact of growth hormone administration in post-pubertal individuals with a detectable decrease in the epigenetic age (Bartke *et al.*, 2021) as well as to study the impact of obesity on the liver epigenetic age (Horvath *et al.*, 2014).

7.3.3 Function of the other studied genes

Similar to the analysis for the age-CpGs located in the *FHL2* promoter from Chapter 5 and 6, this work can then be applied for the study of the other age-CpG sites and moreover

extended to include various other tissues from outside the musculoskeletal system. For the genes of*, ZNF577, ZNF551* their role remains largely unknown, yet DNA methylation could impact their expression, and these have age-related DNA methylation changes. Future work for this would be to generate similar knockdown and overexpression line followed by transcriptomic analysis of the changes to determine the involved pathways as well as functional studies for these changes. The target of the enhancers of *PCDHB1 and PCDHGA1* is still unknown, and these nearby genes were initially presumed to be possible targets however since enhancers do not always target the nearby target gene it could be that these are not targets. Additionally, since these are part of a gene family cluster it would be difficult to assess their potential target only through RT-PCR and for future work through the use of techniques such as Capture Hi-C would aid in the determination of the target genes for this region. Once this will be established a similar workflow could be performed where the impact of DNA methylation can be assessed as well as the study of the gene's functionality.

Identify the other transcription factors that bind the age-CpGs of *FHL2*

EMSA complex identification was only able to identify two transcription factors (i.e. Sp1 and Sp3 present in a total of three complexes) present at the age-CpG region of *FHL2* with 32 remaining complexes to be identified. Through the use of additional techniques coupled to EMSA such to western blotting (Waby *et al.*, 2010) or with mass- spectrometry (Stead *et al.*, 2006) will greatly aid in the identification of the remaining factors.

7.3.4 Investigate FHL2 protein levels change with age

Since the data presented here suggest that the expression of *FHL2* drops with the increase in the methylation of the age-CpGs within the gene's promoter, further work would need to focus to understand whether this also translates in less protein levels for this as well. Additionally, the data for the *FHL2* expression used a RT-PCR assay that targeted the common exons of all the isoforms of the gene however it is not currently known whether the age-related effects impact only certain isoforms for this gene and what importance this could present.

7.3.5 CRISPRa and CRISPRi

Additional functional analysis can be done to study the age-CpG regions using CRISPRa (e.g. Cas9-VP64) and CRISPRi (Larson *et al.*, 2013; Replogle *et al.*, 2022) (e.g. dCas9 and Cas9-KRAB) systems where localised activation/enhancement and inhibition of the target region can be achieved to study their role. Such systems have been previously used to temporary perturb gene expression (Kampmann, 2018) and provide an additional tool for the study of promoters (Heidersbach *et al.*, 2023) and enhancers (Matharu *et al.*, 2019).

7.3.6 Accessing more patient characteristics

Due to the limited data available through the current ethical approval only a small investigation of the DNA methylation changes was performed with only age and sex being investigated. Other variables to investigate the DNA methylation changes would be variables such as smoking status of the patient, Kellgren-Lawrence (K-L) score, BMI, past physical trauma incident for articular joints and weekly physical activity. This would permit the determination of their contribution towards these reported age-related effects and could further provide with points for the stratification of the data.

7.3.7 Function of work using conditional knock out experiments in mice for articular joint tissues and human articular chondrocytes

Further to the work presented here, future work could also focus on performing conditional knockouts of FHL2 in mice for the articular tissues such as the cartilage tissue as well as for isolated human chondrocytes. This could provide further evidence towards the function of FHL2 in the cartilage tissue and would further complement the data present here for the SW1353 cell line.

Bibliography

Acevedo, N., Scala, G., Merid, S., Frumento, P., Bruhn, S., Andersson, A., Ogris, C., Bottai, M., Pershagen, G., Koppelman, G. and Melén, E. (2021) 'DNA Methylation Levels in Mononuclear Leukocytes from the Mother and Her Child Are Associated with IgE Sensitization to Allergens in Early Life', *International journal of molecular sciences*, 22(2).

Acosta, J.C., Banito, A., Wuestefeld, T., Georgilis, A., Janich, P., Morton, J.P., Athineos, D., Kang, T.-W., Lasitschka, F., Andrulis, M., Pascual, G., Morris, K.J., Khan, S., Jin, H., Dharmalingam, G., Snijders, A.P., Carroll, T., Capper, D., Pritchard, C., Inman, G.J., Longerich, T., Sansom, O.J., Benitah, S.A., Zender, L. and Gil, J. (2013) 'A complex secretory program orchestrated by the inflammasome controls paracrine senescence', *Nature Cell Biology*, 15(8), pp. 978-990.

Adams, P. (2007) 'Remodeling of chromatin structure in senescent cells and its potential impact on tumor suppression and aging', *Gene*, 397(1-2).

Adli, M. (2018) 'The CRISPR tool kit for genome editing and beyond', *Nat Commun*, 9.

Aigner, T., Zien, A., Gehrsitz, A., Gebhard, P. and McKenna, L. (2001) 'Anabolic and catabolic gene expression pattern analysis in normal versus osteoarthritic cartilage using complementary DNA-array technology', *Arthritis and rheumatism*, 44(12).

Ajekigbe, B., Cheung, K., Xu, Y., Skelton, A., Panagiotopoulos, A., Soul, J., Hardingham, T., Deehan, D., Barter, M. and Young, D. (2019) 'Identification of long non-coding RNAs expressed in knee and hip osteoarthritic cartilage', *Osteoarthritis and cartilage*, 27(4).

Akkiraju, H. and Nohe, A. (2015) 'Role of Chondrocytes in Cartilage Formation, Progression of Osteoarthritis and Cartilage Regeneration', *J Dev Biol*, 3(4), pp. 177-92.

Alfaddagh, A., Martin, S., Leucker, T., Michos, E., Blaha, M., Lowenstein, C., Jones, S. and Toth, P. (2020) 'Inflammation and cardiovascular disease: From mechanisms to therapeutics', *American journal of preventive cardiology*, 4.

Alvarez-Garcia, O., Fisch, K., Wineinger, N., Akagi, R., Saito, M., Sasho, T., Su, A. and Lotz, M. (2016a) 'Increased DNA Methylation and Reduced Expression of Transcription Factors in Human Osteoarthritis Cartilage', *Arthritis & rheumatology (Hoboken, N.J.)*, 68(8).

Alvarez-Garcia, O., Fisch, K.M., Wineinger, N.E., Akagi, R., Saito, M., Sasho, T., Su, A.I. and Lotz, M.K. (2016b) 'Increased DNA Methylation and Reduced Expression of Transcription Factors in Human Osteoarthritis Cartilage', *Arthritis Rheumatol*, 68(8), pp. 1876-86.

Anastasiadi, D., Esteve-Codina, A. and Piferrer, F. (2018) 'Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species', *Epigenetics & Chromatin*, 11(1), pp. 1-17.

Andarawis-Puri, N., Flatow, E. and Soslowsky, L. (2015) 'Tendon basic science: Development, repair, regeneration, and healing', *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, 33(6).

Angeloni, A. and Bogdanovic, O. (2019) 'Enhancer DNA methylation: implications for gene regulation', *Essays in biochemistry*, 63(6).

Aoyama, T., Okamoto, T., Nagayama, S., Nishijo, K., Ishibe, T., Yasura, K., Nakayama, T., Nakamura, T. and Toguchida, J. (2004) 'Methylation in the core-promoter region of the chondromodulin-I gene determines the cell-specific expression by regulating the binding of transcriptional activator Sp3', *The Journal of biological chemistry*, 279(27).

Archer, C.W. and Francis-West, P. (2003) 'The chondrocyte', *Int J Biochem Cell Biol*, 35(4), pp. 401-4.

Aref-Eshghi, E., Zhang, Y., Liu, M., Harper, P.E., Martin, G., Furey, A., Green, R., Sun, G., Rahman, P. and Zhai, G. (2015) 'Genome-wide DNA methylation study of hip and knee cartilage reveals embryonic organ and skeletal system morphogenesis as major pathways involved in osteoarthritis', in *BMC Musculoskelet Disord*.

Ashraf, S. and Walsh, D.A. (2008) 'Angiogenesis in osteoarthritis', *Curr Opin Rheumatol*, 20(5), pp. 573-80.

Attwooll, C., Lazzerini, D., E and Helin, K. (2004) 'The E2F family: specific functions and overlapping interests', *The EMBO journal*, 23(24).

Azofeifa, J., Allen, M., Hendrix, J., Read, T., Rubin, J. and Dowell, R. (2018) 'Enhancer RNA profiling predicts transcription factor activity', *Genome research*, 28(3).

Bacalini, M., Boattini, A., Gentilini, D., Giampieri, E., Pirazzini, C., Giuliani, C., Fontanesi, E., Remondini, D., Capri, M., Del, R., A, Luiselli, D., Vitale, G., Mari, D., Castellani, G., Di, B., AM, Salvioli, S., Franceschi, C. and Garagnani, P. (2015) 'A meta-analysis on age-associated changes in blood DNA methylation: results from an original analysis pipeline for Infinium 450k data', *Aging*, 7(2).

Bacalini, M., Deelen, J., Pirazzini, C., De, C., M, Giuliani, C., Lanzarini, C., Ravaioli, F., Marasco, E., van, H., D, Suchiman, H., Slieker, R., Giampieri, E., Recchioni, R., Marcheselli, F., Salvioli, S., Vitale, G., Olivieri, F., Spijkerman, A., Dollé, M., Sedivy, J., Castellani, G., Franceschi, C., Slagboom, P. and Garagnani, P. (2017) 'Systemic Age-Associated DNA Hypermethylation of ELOVL2 Gene: In Vivo and In Vitro Evidences of a Cell Replication Process', *The journals of gerontology. Series A, Biological sciences and medical sciences*, 72(8).

Bacos, K., Gillberg, L., Volkov, P., Olsson, A.H., Hansen, T., Pedersen, O., Gjesing, A.P., Eiberg, H., Tuomi, T., Almgren, P., Groop, L., Eliasson, L., Vaag, A., Dayeh, T. and Ling, C. (2016) 'Blood-based biomarkers of age-associated epigenetic changes in human islets associate with insulin secretion and diabetes', *Nature Communications*, 7(1), pp. 1-13.

Bailey, T., Johnson, J., Grant, C. and Noble, W. (2015) 'The MEME Suite', *Nucleic acids research*, 43(W1).

Baker, D.J., Wijshake, T., Tchkonia, T., LeBrasseur, N.K., Childs, B.G., van de Sluis, B., Kirkland, J.L. and van Deursen, J.M. (2011) 'Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders', *Nature*, 479(7372), pp. 232-236.

Barter, M., Cheung, K., Falk, J., Panagiotopoulos, A., Cosimini, C., O'Brien, S., Teja-Putri, K., Neill, G., Deehan, D. and Young, D. (2021) 'Dynamic chromatin accessibility landscape changes following interleukin-1 stimulation', *Epigenetics*, 16(1).

Bartke, A., Hascup, E., Hascup, K. and Masternak, M. (2021) 'Growth Hormone and Aging: New Findings', *The world journal of men's health*, 39(3).

Bassil, C., Huang, Z. and Murphy, S. (2013) 'Bisulfite pyrosequencing', *Methods in molecular biology (Clifton, N.J.)*, 1049.

Basu, H., Pekkurnaz, G., Falk, J., Wei, W., Chin, M., Steen, J. and Schwarz, T. (2021) 'FHL2 anchors mitochondria to actin and adapts mitochondrial dynamics to glucose supply', *The Journal of cell biology*, 220(10).

Bedson, J., Chen, Y., Hayward, R.A., Ashworth, J., Walters, K., Dunn, K.M. and Jordan, K.P. (2016) 'Trends in long-term opioid prescribing in primary care patients with musculoskeletal conditions: an observational database study', *Pain*, 157(7), pp. 1525-31.

Bellido, M., Lugo, L., Roman-Blas, J.A., Castañeda, S., Caeiro, J.R., Dapia, S., Calvo, E., Largo, R. and Herrero-Beaumont, G. (2010) 'Subchondral bone microstructural damage by increased remodelling aggravates experimental osteoarthritis preceded by osteoporosis', *Arthritis Res Ther*, 12(4), p. R152.

Benayoun, B.A., Pollina, E.A. and Brunet, A. (2015) 'Epigenetic regulation of ageing: linking environmental inputs to genomic stability', *Nature Reviews Molecular Cell Biology*, 16(10), p. 593.

Benito, M.J., Veale, D.J., FitzGerald, O., van den Berg, W.B. and Bresnihan, B. (2005) 'Synovial tissue inflammation in early and late osteoarthritis', *Ann Rheum Dis*, 64(9), pp. 1263-7.

Berben, L., Floris, G., Wildiers, H. and Hatse, S. (2021) 'Cancer and Aging: Two Tightly Interconnected Biological Processes', *Cancers*, 13(6).

Bhutto, A. and Morley, J. (2008) 'The clinical significance of gastrointestinal changes with aging', *Current opinion in clinical nutrition and metabolic care*, 11(5).

Biran, A., Zada, L., Abou, K., P, Vadai, E., Roitman, L., Ovadya, Y., Porat, Z. and Krizhanovsky, V. (2017) 'Quantitative identification of senescent cells in aging and disease', *Aging cell*, 16(4).

Bird, A. (2002) 'DNA methylation patterns and epigenetic memory', *Genes & development*, 16(1).

Bird, A. (2007) 'Perceptions of epigenetics', *Nature*, 447(7143), p. 396.

Bitto, A., Ito, T.K., Pineda, V.V., LeTexier, N.J., Huang, H.Z., Sutlief, E., Tung, H., Vizzini, N., Chen, B., Smith, K., Meza, D., Yajima, M., Beyer, R.P., Kerr, K.F., Davis, D.J., Gillespie, C.H., Snyder, J.M., Treuting, P.M. and Kaeberlein, M. (2016) 'Transient rapamycin treatment can increase lifespan and healthspan in middle-aged mice', *eLife*, 5.

Bitton, R. (2009) 'The economic burden of osteoarthritis', *Am J Manag Care*, 15(8 Suppl), pp. S230-5.

Bjornsson, H.T., Fallin, M.D. and Feinberg, A.P. (2004) 'An integrated epigenetic and genetic approach to common human disease', *Trends Genet*, 20(8), pp. 350-8.

Blagosklonny, M.V. (2010) 'Why the disposable soma theory cannot explain why women live longer and why we age', *Aging (Albany NY)*, 2(12), pp. 884-7.

Blagosklonny, M.V. and Hall, M.N. (2009) 'Growth and aging: a common molecular mechanism', *Aging (Albany NY)*, 1(4), pp. 357-62.

Blalock, D., Miller, A., Tilley, M. and Wang, J. (2015) 'Joint Instability and Osteoarthritis', *Clin Med Insights Arthritis Musculoskelet Disord*, 8, pp. 15-23.

Bocklandt, S., Lin, W., Sehl, M.E., Sanchez, F.J., Sinsheimer, J.S., Horvath, S. and Vilain, E. (2011) 'Epigenetic predictor of age', *PLoS One*, 6(6), p. e14821.

Bordoni, B. and Varacallo, M. (2018) 'Anatomy, Tendons', *StatPearls*.

Botto, N., Rizza, A., Colombo, M., Mazzone, A., Manfredi, S., Masetti, S., Clerico, A., Biagini, A. and Andreassi, M. (2001) 'Evidence for DNA damage in patients with coronary artery disease', *Mutation research*, 493(1-2).

Boumber, Y., Kondo, Y., Chen, X., Shen, L., Guo, Y., Tellez, C., Estécio, M., Ahmed, S. and Issa, J. (2008) 'An Sp1/Sp3 binding polymorphism confers methylation protection', *PLoS genetics*, 4(8).

Bracken, A., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C., Theilgaard-Mönch, K., Minucci, S., Porse, B., Marine, J., Hansen, K. and Helin, K. (2007) 'The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells', *Genes & development*, 21(5).

Breton, C.V., Landon, R., Kahn, L.G., Enlow, M.B., Peterson, A.K., Bastain, T., Braun, J., Comstock, S.S., Duarte, C.S., Hipwell, A., Ji, H., LaSalle, J.M., Miller, R.L., Musci, R., Posner, J., Schmidt, R., Suglia, S.F., Tung, I., Weisenberger, D., Zhu, Y. and Fry, R. (2021) 'Exploring the evidence for epigenetic regulation of environmental influences on child health across generations', *Communications Biology*, 4(1), pp. 1-15.

Brittberg, M., Gomoll, A., Canseco, J., Far, J., Lind, M. and Hui, J. (2016) 'Cartilage repair in the degenerative ageing knee', *Acta orthopaedica*, 87(sup363).

Brown, T.D., Johnston, R.C., Saltzman, C.L., Marsh, J.L. and Buckwalter, J.A. (2006) 'Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease', *J Orthop Trauma*, 20(10), pp. 739-44.

Brunet, A. and Berger, S. (2014) 'Epigenetics of aging and aging-related disease', *The journals of gerontology. Series A, Biological sciences and medical sciences*, 69 Suppl 1(Suppl 1).

Buckwalter, J.A. (2003) 'Sports, joint injury, and posttraumatic osteoarthritis', *J Orthop Sports Phys Ther*, 33(10), pp. 578-88.

Buckwalter, J.A., Glimcher, M.J., Cooper, R.R. and Recker, R. (1996) 'Bone biology. I: Structure, blood supply, cells, matrix, and mineralization', *Instr Course Lect*, 45, pp. 371-86.

Bui, C., Barter, M., Scott, J., Xu, Y., Galler, M., Reynard, L., Rowan, A. and Young, D. (2012) 'cAMP response element-binding (CREB) recruitment following a specific CpG demethylation leads to the elevated expression of the matrix metalloproteinase 13 in human articular chondrocytes and osteoarthritis', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 26(7).

Bulger, M. and Groudine, M. (2011) 'Functional and mechanistic diversity of distal transcription enhancers', *Cell*, 144(3), pp. 327-39.

Burr, D.B. and Radin, E.L. (2003) 'Microfractures and microcracks in subchondral bone: are they relevant to osteoarthrosis?', *Rheum Dis Clin North Am*, 29(4), pp. 675-85.

Cai, T., Sun, D., Duan, Y., Qiu, Y., Dai, C., Yang, J. and He, W. (2018) 'FHL2 promotes tubular epithelial-to-mesenchymal transition through modulating β-catenin signalling', *Journal of cellular and molecular medicine*, 22(3).

Calo, E. and Wysocka, J. (2013) 'Modification of enhancer chromatin: what, how and why?', *Mol Cell*, 49(5).

Campanero, M., Armstrong, M. and Flemington, E. (2000) 'CpG methylation as a mechanism for the regulation of E2F activity', *Proceedings of the National Academy of Sciences of the United States of America*, 97(12).

Canault, M., Tellier, E., Bonardo, B., Mas, E., Aumailley, M., Juhan-Vague, I., Nalbone, G. and Peiretti, F. (2006) 'FHL2 interacts with both ADAM-17 and the cytoskeleton and regulates ADAM-17 localization and activity', *Journal of cellular physiology*, 208(2).

Cano-Rodriguez, D., Gjaltema, R.A.F., Jilderda, L.J., Jellema, P., Dokter-Fokkens, J., Ruiters, M.H.J. and Rots, M.G. (2016) 'Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner', *Nature Communications*, 7(1), pp. 1-11.

Cao, C., Mok, S., Cheng, V. and Tsui, S. (2015) 'The FHL2 regulation in the transcriptional circuitry of human cancers', *Gene*, 572(1).

Carver, B., Plosa, E., Stinnett, A., Blackwell, T. and Prince, L. (2013) 'Interactions between NF-κB and SP3 connect inflammatory signaling with reduced FGF-10 expression', *The Journal of biological chemistry*, 288(21).

Cesare, A. and Karlseder, J. (2012) 'A three-state model of telomere control over human proliferative boundaries', *Current opinion in cell biology*, 24(6).

Chambers, B., Gerlach, G., Clark, E., Chen, K., Levesque, A., Leshchiner, I., Goessling, W. and Wingert, R. (2019) 'Tfap2a is a novel gatekeeper of nephron differentiation during kidney development', *Development (Cambridge, England)*, 146(13).

Chan, K., Tsui, S., Ngai, S., Lee, S., Kotaka, M., Waye, M., Lee, C. and Fung, K. (2000) 'Protein-protein interaction of FHL2, a LIM domain protein preferentially expressed in human heart, with hCDC47', *Journal of cellular biochemistry*, 76(3).

Chao, D. and Skowronska-Krawczyk, D. (2020) 'ELOVL2: Not just a biomarker of aging', *Translational medicine of aging*, 4.

Charlet, J., Duymich, C.E., Lay, F.D., Mundbjerg, K., Dalsgaard Sorensen, K., Liang, G. and Jones, P.A. (2016) 'Bivalent Regions of Cytosine Methylation and H3K27 Acetylation Suggest an Active Role for DNA Methylation at Enhancers', *Mol Cell*, 62(3), pp. 422-431.

Chen, A., Gupte, C., Akhtar, K., Smith, P. and Cobb, J. (2012) 'The Global Economic Cost of Osteoarthritis: How the UK Compares', *Arthritis*, 2012, p. 698709.

Chen, D., Chao, D.L., Rocha, L., Kolar, M., Nguyen Huu, V.A., Krawczyk, M., Dasyani, M., Wang, T., Jafari, M., Jabari, M., Ross, K.D., Saghatelian, A., Hamilton, B.A., Zhang, K. and Skowronska-Krawczyk, D. (2020) 'The lipid elongation enzyme ELOVL2 is a molecular regulator of aging in the retina', *Aging Cell*, 19(2).

Chen, L. and Oshima, J. (2002) 'Werner Syndrome', *J Biomed Biotechnol*, 2(2), pp. 46-54.

Chen, M., Logan, T., Hochberg, M., Shelat, S., Yu, X., Wilding, G., Tan, W., Kujoth, G., Prolla, T., Selak, M., Kundu, M., Carroll, M. and Thompson, J. (2009) 'Erythroid dysplasia, megaloblastic anemia, and impaired lymphopoiesis arising from mitochondrial dysfunction', *Blood*, 114(19).

Cheng, J., Yoo, C., Weisenberger, D., Chuang, J., Wozniak, C., Liang, G., Marquez, V., Greer, S., Orntoft, T., Thykjaer, T. and Jones, P. (2004) 'Preferential response of cancer cells to zebularine', *Cancer cell*, 6(2).

Cheon, S., Kim, H., Rubinsztein, D. and Lee, J. (2019) 'Autophagy, Cellular Aging and Age-related Human Diseases', *Experimental neurobiology*, 28(6).

Cheung, K., Burgers, M., Young, D., Cockell, S. and Reynard, L. (2020) 'Correlation of Infinium HumanMethylation450K and MethylationEPIC BeadChip arrays in cartilage', *Epigenetics*, 15(6-7).

Cho, H., Walker, A., Williams, J. and Hasty, K.A. (2015) 'Study of osteoarthritis treatment with anti-inflammatory drugs: cyclooxygenase-2 inhibitor and steroids', *Biomed Res Int*, 2015, p. 595273.
Choudhury, S.R., Cui, Y., Lubecka, K., Stefanska, B. and Irudayaraj, J. (2016) 'CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter', *Oncotarget*, 7(29), pp. 46545-56.
Christman, J.K. (2002) '5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy', *Oncogene*, 21(35), pp. 5483-95.
Chuckpaiwong, B., Charles, H.C., Kraus, V.B., Guilak, F. and Nunley, J.A. (2010) 'Age-associated increases in the size of the infrapatellar fat pad in knee osteoarthritis as measured by 3T MRI', *J Orthop Res*, 28(9), pp. 1149-54.
Clemens, A., Wu, D., Moore, J., Christian, D., Zhao, G. and Gabel, H. (2020) 'MeCP2 Represses Enhancers through Chromosome Topology-Associated DNA Methylation', *Molecular cell*, 77(2).
Clemente-Olivo, M., Hernández-Quiles, M., Sparrius, R., van der Stoel, M., Janssen, V., Habibe, J., van den Burg, J., Jongejan, A., Alcaraz-Sobrevals, P., van Es, R., Vos, H., Kalkhoven, E. and de Vries, C. (2023) 'Early adipogenesis is repressed through the newly identified FHL2-NFAT5 signaling complex', *Cellular signalling*, 104.
Clevenger, C. (2004) 'Roles and regulation of stat family transcription factors in human breast cancer', *The American journal of pathology*, 165(5).
Collino, S., Montoliu, I., Martin, F., Scherer, M., Mari, D., Salvioli, S., Bucci, L., Ostan, R., Monti, D., Biagi, E., Brigidi, P., Franceschi, C. and Rezzi, S. (2013) 'Metabolic signatures of extreme longevity in northern Italian centenarians reveal a complex remodeling of lipids, amino acids, and gut microbiota metabolism', *PloS one*, 8(3).
Coppé, J., Desprez, P., Krtolica, A. and Campisi, J. (2010) 'The senescence-associated secretory phenotype: the dark side of tumor suppression', *Annual review of pathology*, 5.
Correia-Melo, C. and Passos, J. (2015) 'Mitochondria: Are they causal players in cellular senescence?', *Biochimica et biophysica acta*, 1847(11).
Corso-Díaz, X., Gentry, J., Rebernick, R., Jaeger, C., Brooks, M., van, A., F, Kooragayala, K., Gieser, L., Nellissery, J., Covian, R., Cogliati, T., Mondal, A., Jiang, K. and Swaroop, A. (2020) 'Genome-wide Profiling Identifies DNA Methylation Signatures of Aging in Rod Photoreceptors Associated with Alterations in Energy Metabolism', *Cell reports*, 31(3).
Costa, C., Manaa, W., Duplan, E. and Checler, F. (2020) 'The Endoplasmic Reticulum Stress/Unfolded Protein Response and Their Contributions to Parkinson's Disease Physiopathology', *Cells*, 9(11).
Cross, M., Smith, E., Hoy, D., Nolte, S., Ackerman, I., Fransen, M., Bridgett, L., Williams, S., Guillemin, F., Hill, C.L., Laslett, L.L., Jones, G., Cicuttini, F., Osborne, R., Vos, T., Buchbinder, R., Woolf, A. and

March, L. (2014) 'The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study', *Ann Rheum Dis*, 73(7), pp. 1323-30.
Cui, Y., Hameed, F., Yang, B., Lee, K., Pan, C., Park, S. and Sheetz, M. (2015) 'Cyclic stretching of soft substrates induces spreading and growth', *Nature communications*, 6.
Curradi, M., Izzo, A., Badaracco, G. and Landsberger, N. (2002) 'Molecular mechanisms of gene silencing mediated by DNA methylation', *Molecular and cellular biology*, 22(9).
d'Adda di Fagagna, F. (2008) 'Living on a break: cellular senescence as a DNA-damage response', *Nature reviews. Cancer*, 8(7).
d'Adda di Fagagna, F., Reaper, P., Clay-Farrace, L., Fiegler, H., Carr, P., Von Zglinicki, T., Saretzki, G., Carter, N. and Jackson, S. (2003) 'A DNA damage checkpoint response in telomere-initiated senescence', *Nature*, 426(6963).
Dahan, J., Levillayer, F., Xia, T., Nouët, Y., Werts, C., Fanton d'Andon, M., Adib-Conquy, M., Cassard-Doulcier, A., Khanna, V., Chen, J., Tordjmann, T., Buendia, M., Jouvion, G. and Wei, Y. (2017) 'LIM-Only Protein FHL2 Is a Negative Regulator of Transforming Growth Factor β1 Expression', *Molecular and cellular biology*, 37(10).
Das, S.K. and Farooqi, A. (2008) 'Osteoarthritis', *Best Pract Res Clin Rheumatol*, 22(4), pp. 657-75.
Day, K., Waite, L.L., Thalacker-Mercer, A., West, A., Bamman, M.M., Brooks, J.D., Myers, R.M. and Absher, D. (2013) 'Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape', *Genome Biol*, 14(9), p. R102.
de Graff, A., Hazoglou, M. and Dill, K. (2016) 'Highly Charged Proteins: The Achilles' Heel of Aging Proteomes', *Structure (London, England : 1993)*, 24(2).
De Majo, F., Martens, L., Hegenbarth, J., Rühle, F., Hamczyk, M., Nevado, R., Andrés, V., Hilbold, E., Bär, C., Thum, T., de Boer, M., Duncker, D., Schroen, B., Armand, A., Stoll, M. and De Windt, L. (2021) 'Genomic instability in the naturally and prematurely aged myocardium', *Proceedings of the National Academy of Sciences of the United States of America*, 118(36).
De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L. and Natoli, G. (2010) 'A large fraction of extragenic RNA pol II transcription sites overlap enhancers', *PLoS Biol*, 8(5), p. e1000384.
de Toro-Martín, J., Guénard, F., Tchernof, A., Hould, F.-S., Lebel, S., Julien, F., Marceau, S. and Vohl, M.-C. (2019) 'Body mass index is associated with epigenetic age acceleration in the visceral adipose tissue of subjects with severe obesity', *Clinical Epigenetics*, 11(1), pp. 1-11.
den Hollander, W., Ramos, Y., Bomer, N., Elzinga, S., van der, B., R, Lakenberg, N., de, D., WJ, Suchiman, H., Duijnisveld, B., Houwing-Duistermaat, J., Slagboom, P., Bos, S., Nelissen, R. and

Meulenbelt, I. (2015) 'Transcriptional associations of osteoarthritis-mediated loss of epigenetic control in articular cartilage', *Arthritis & rheumatology (Hoboken, N.J.)*, 67(8).
Denic, A., Glassock, R. and Rule, A. (2016) 'Structural and Functional Changes With the Aging Kidney', *Advances in chronic kidney disease*, 23(1).
Dhar, G., Saha, S., Mitra, P. and Nag, C.R. (2021) 'DNA methylation and regulation of gene expression: Guardian of our health', *The Nucleus : an international journal of cytology and allied topics*, 64(3).
Dilling, C., Roewer, N., Förster, C. and Burek, M. (2017) 'Multiple protocadherins are expressed in brain microvascular endothelial cells and might play a role in tight junction protein regulation', *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, 37(10).
Ding, L., Heying, E., Nicholson, N., Stroud, N.J., Homandberg, G.A., Buckwalter, J.A., Guo, D. and Martin, J.A. (2010) 'Mechanical impact induces cartilage degradation via mitogen activated protein kinases', *Osteoarthritis Cartilage*, 18(11), pp. 1509-17.
Distel, E., Cadoudal, T., Durant, S., Poignard, A., Chevalier, X. and Benelli, C. (2009) 'The infrapatellar fat pad in knee osteoarthritis: an important source of interleukin-6 and its soluble receptor', *Arthritis Rheum*, 60(11), pp. 3374-7.
Divo, M.J., Martinez, C.H. and Mannino, D.M. (2014) 'Ageing and the epidemiology of multimorbidity', *Eur Respir J*, 44(4), pp. 1055-68.
Djeghloul, D., Kuranda, K., Kuzniak, I., Barbieri, D., Naguibneva, I., Choisy, C., Bories, J., Dosquet, C., Pla, M., Vanneaux, V., Socié, G., Porteu, F., Garrick, D. and Goodhardt, M. (2016) 'Age-Associated Decrease of the Histone Methyltransferase SUV39H1 in HSC Perturbs Heterochromatin and B Lymphoid Differentiation', *Stem cell reports*, 6(6).
Domcke, S., Bardet, A.F., Adrian Ginno, P., Hartl, D., Burger, L. and Schübeler, D. (2015) 'Competition between DNA methylation and transcription factors determines binding of NRF1', *Nature*, 528(7583), pp. 575-579.
Du, P., Zhang, X., Huang, C., Jafari, N., Kibbe, W., Hou, L. and Lin, S. (2010) 'Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis', *BMC bioinformatics*, 11.
Du, Q., Luu, P., Stirzaker, C. and Clark, S. (2015) 'Methyl-CpG-binding domain proteins: readers of the epigenome', *Epigenomics*, 7(6).
Du, X., Hublitz, P., Günther, T., Wilhelm, D., Englert, C. and Schüle, R. (2002) 'The LIM-only coactivator FHL2 modulates WT1 transcriptional activity during gonadal differentiation', *Biochimica et biophysica acta*, 1577(1).

Dupont, C., Armant, D.R. and Brenner, C.A. (2009) 'Epigenetics: Definition, Mechanisms and Clinical Perspective', *Semin Reprod Med*, 27(5), pp. 351-7.
Dziedzic, K.S. and Allen, K.D. (2018) 'Challenges and controversies of complex interventions in osteoarthritis management: recognizing inappropriate and discordant care', *Rheumatology (Oxford)*, 57(Suppl 4), pp. iv88-98.
Edwards, P.K., Mears, S.C., Stambough, J.B., Foster, S.E. and Barnes, C.L. (2018) 'Choices, Compromises, and Controversies in Total Knee and Total Hip Arthroplasty Modifiable Risk Factors: What You Need to Know', *J Arthroplasty*, 33(10), pp. 3101-3106.
Egloff, C., Hugle, T. and Valderrabano, V. (2012) 'Biomechanics and pathomechanisms of osteoarthritis', *Swiss Med Wkly*, 142, p. w13583.
ElSayed, S.A. and Varacallo, M. (2018) 'Physiology, Bone', *StatPearls*.
Engler, A., Sen, S., Sweeney, H. and Discher, D. (2006) 'Matrix elasticity directs stem cell lineage specification', *Cell*, 126(4).
Fan, K., Moore, J., Zhang, X. and Weng, Z. (2021) 'Genetic and epigenetic features of promoters with ubiquitous chromatin accessibility support ubiquitous transcription of cell-essential genes', *Nucleic acids research*, 49(10).
Farzadfard, F., Perli, S. and Lu, T. (2013) 'Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas', *ACS synthetic biology*, 2(10).
Favero, M., El-Hadi, H., Belluzzi, E., Granzotto, M., Porzionato, A., Sarasin, G., Rambaldo, A., Iacobellis, C., Cigolotti, A., Fontanella, C.G., Natali, A., Ramonda, R., Ruggieri, P., De Caro, R., Vettor, R., Rossato, M. and Macchi, V. (2017) 'Infrapatellar fat pad features in osteoarthritis: a histopathological and molecular study', *Rheumatology (Oxford)*, 56(10), pp. 1784-1793.
Fayos, G.d.L., Alonso, I., Liang, H., Turner, S., Lagger, S., Merkel, O. and Kenner, L. (2018) 'The Role of Activator Protein-1 (AP-1) Family Members in CD30-Positive Lymphomas', *Cancers*, 10(4).
Fernandes, G.S., Parekh, S.M., Moses, J., Fuller, C., Scammell, B., Batt, M.E., Zhang, W. and Doherty, M. (2017) 'Prevalence of knee pain, radiographic osteoarthritis and arthroplasty in retired professional footballers compared with men in the general population: a cross-sectional study', *British Journal of Sports Medicine*.
Fernandez-Jimenez, N., Allard, C., Bouchard, L., Perron, P., Bustamante, M., Bilbao, J. and Hivert, M. (2019) 'Comparison of Illumina 450K and EPIC arrays in placental DNA methylation', *Epigenetics*, 14(12).
Ferrucci, L. and Fabbri, E. (2018) 'Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty', *Nature reviews. Cardiology*, 15(9).

Feser, J., Truong, D., Das, C., Carson, J., Kieft, J., Harkness, T. and Tyler, J. (2010) 'Elevated histone expression promotes life span extension', *Molecular cell*, 39(5).
Fimia, G., De Cesare, D. and Sassone-Corsi, P. (2000) 'A family of LIM-only transcriptional coactivators: tissue-specific expression and selective activation of CREB and CREM', *Molecular and cellular biology*, 20(22).
Fleming, B.C., Hulstyn, M.J., Oksendahl, H.L. and Fadale, P.D. (2005) 'Ligament Injury, Reconstruction and Osteoarthritis', *Curr Opin Orthop*, 16(5), pp. 354-62.
Flint, B. and Tadi, P. (2021) 'Physiology, Aging'.
Florath, I., Butterbach, K., Müller, H., Bewerunge-Hudler, M. and Brenner, H. (2014) 'Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites', *Human molecular genetics*, 23(5).
Florencio-Silva, R., Sasso, G.S., Sasso-Cerri, E., Simões, M.J. and Cerri, P.S. (2015) 'Biology of Bone Tissue: Structure, Function, and Factors That Influence Bone Cells', *Biomed Res Int*, 2015.
Fortier, L.A., Barker, J.U., Strauss, E.J., McCarrel, T.M. and Cole, B.J. (2011) 'The role of growth factors in cartilage repair', *Clin Orthop Relat Res*, 469(10), pp. 2706-15.
Fortin, J., Labbe, A., Lemire, M., Zanke, B., Hudson, T., Fertig, E., Greenwood, C. and Hansen, K. (2014) 'Functional normalization of 450k methylation array data improves replication in large cancer studies', *Genome biology*, 15(12).
Fraga, M.F., Agrelo, R. and Esteller, M. (2007) 'Cross-talk between aging and cancer: the epigenetic language', *Ann N Y Acad Sci*, 1100, pp. 60-74.
Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T.D., Wu, Y.Z., Plass, C. and Esteller, M. (2005) 'Epigenetic differences arise during the lifetime of monozygotic twins', *Proc Natl Acad Sci U S A*, 102(30), pp. 10604-9.
Franceschi, C. and Campisi, J. (2014) 'Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases', *The journals of gerontology. Series A, Biological sciences and medical sciences*, 69 Suppl 1.
Franceschi, C., Garagnani, P., Morsiani, C., Conte, M., Santoro, A., Grignolio, A., Monti, D., Capri, M. and Salvioli, S. (2018) 'The Continuum of Aging and Age-Related Diseases: Common Mechanisms but Different Rates', *Frontiers in medicine*, 5.
Friedrich, F., Reischmann, S., Schwalm, A., Unger, A., Ramanujam, D., Münch, J., Müller, O., Hengstenberg, C., Galve, E., Charron, P., Linke, W., Engelhardt, S., Patten, M., Richard, P., van der Velden, J., Eschenhagen, T., Isnard, R. and Carrier, L. (2014) 'FHL2 expression and variants in hypertrophic cardiomyopathy', *Basic research in cardiology*, 109(6).

Frietze, S. and Farnham, P. (2011) 'Transcription factor effector domains', *Sub-cellular biochemistry*, 52.

Gabriel, B., Mildenberger, S., Weisser, C., Metzger, E., Gitsch, G., Schüle, R. and Müller, J. (2004) 'Focal adhesion kinase interacts with the transcriptional coactivator FHL2 and both are overexpressed in epithelial ovarian cancer', *Anticancer research*, 24(2B).

Garagnani, P., Bacalini, M., Pirazzini, C., Gori, D., Giuliani, C., Mari, D., Di, B., AM, Gentilini, D., Vitale, G., Collino, S., Rezzi, S., Castellani, G., Capri, M., Salvioli, S. and Franceschi, C. (2012) 'Methylation of ELOVL2 gene as a new epigenetic marker of age', *Aging cell*, 11(6).

Garner, M. and Revzin, A. (1986) 'The use of gel electrophoresis to detect and study nucleic acid—protein interactions', *Trends in Biochemical Sciences*, 11(10), pp. 395-396.

Gasparoni, G., Bultmann, S., Lutsik, P., Kraus, T., Sordon, S., Vlcek, J., Dietinger, V., Steinmaurer, M., Haider, M., Mulholland, C., Arzberger, T., Roeber, S., Riemenschneider, M., Kretzschmar, H., Giese, A., Leonhardt, H. and Walter, J. (2018) 'DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex', *Epigenetics & chromatin*, 11(1).

Geiger, H. and Van Zant, G. (2002) 'The aging of lympho-hematopoietic stem cells', *Nature immunology*, 3(4).

Ghosh, D. (2000) 'Object-oriented transcription factors database (ooTFD)', *Nucleic acids research*, 28(1).

Ghoshal, K., Datta, J., Majumder, S., Bai, S., Kutay, H., Motiwala, T. and Jacob, S. (2005) '5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal', *Molecular and cellular biology*, 25(11).

Goldring, M.B., Birkhead, J.R., Suen, L.F., Yamin, R., Mizuno, S., Glowacki, J., Arbiser, J.L. and Apperley, J.F. (1994) 'Interleukin-1 beta-modulated gene expression in immortalized human chondrocytes', *J Clin Invest*, 94(6), pp. 2307-16.

Goldring, M.B. and Otero, M. (2011) 'Inflammation in osteoarthritis', *Curr Opin Rheumatol*, 23(5), pp. 471-8.

Goldring, M.B., Otero, M., Tsuchimochi, K., Ijiri, K. and Li, Y. (2008) 'Defining the roles of inflammatory and anabolic cytokines in cartilage metabolism', *Ann Rheum Dis*, 67 Suppl 3, pp. iii75-82.

Goldring, S.R. (2012) 'Alterations in periarticular bone and cross talk between subchondral bone and articular cartilage in osteoarthritis', *Ther Adv Musculoskelet Dis*, 4(4), pp. 249-58.

Goodwin, W., McCabe, D., Sauter, E., Reese, E., Walter, M., Buckwalter, J.A. and Martin, J.A. (2010) 'Rotenone prevents impact-induced chondrocyte death', *J Orthop Res*, 28(8), pp. 1057-63.
Gopalan, S., Carja, O., Fagny, M., Patin, E., Myrick, J.W., McEwen, L.M., Mah, S.M., Kobor, M.S., Froment, A., Feldman, M.W., Quintana-Murci, L. and Henn, B.M. (2017) 'Trends in DNA Methylation with Age Replicate Across Diverse Human Populations', *Genetics*, 206(3), pp. 1659-1674.
Gorbunova, V., Seluanov, A., Mita, P., McKerrow, W., Fenyö, D., Boeke, J., Linker, S., Gage, F., Kreiling, J., Petrashen, A., Woodham, T., Taylor, J., Helfand, S. and Sedivy, J. (2021) 'The role of retrotransposable elements in ageing and age-associated diseases', *Nature*, 596(7870).
Gorgoulis, V., Adams, P., Alimonti, A., Bennett, D., Bischof, O., Bishop, C., Campisi, J., Collado, M., Evangelou, K., Ferbeyre, G., Gil, J., Hara, E., Krizhanovsky, V., Jurk, D., Maier, A., Narita, M., Niedernhofer, L., Passos, J., Robbins, P., Schmitt, C., Sedivy, J., Vougas, K., von Zglinicki, T., Zhou, D., Serrano, M. and Demaria, M. (2019) 'Cellular Senescence: Defining a Path Forward', *Cell*, 179(4).
Gregg, S.Q., Gutierrez, V., Robinson, A.R., Woodell, T., Nakao, A., Ross, M.A., Michalopoulos, G.K., Rigatti, L., Rothermel, C.E., Kamileri, I., Garinis, G.A., Stolz, D.B. and Niedernhofer, L.J. (2012) 'A mouse model of accelerated liver aging caused by a defect in DNA repair', *Hepatology*, 55(2), pp. 609-21.
Grove, C. and Walhout, A. (2008) 'Transcription factor functionality and transcription regulatory networks', *Molecular bioSystems*, 4(4).
GTExConsortium (2013) 'The Genotype-Tissue Expression (GTEx) project', *Nature genetics*, 45(6).
Gu, L., Chen, J., Lin, Q. and Yang, Y. (2021) 'Roles of mitochondrial unfolded protein response in mammalian stem cells', *World journal of stem cells*, 13(7).
Guerrero, L., Marco-Ferreres, R., Serrano, A.L., Arredondo, J.J. and Cervera, M. (2010) 'Secondary enhancers synergise with primary enhancers to guarantee fine-tuned muscle gene expression', *Dev Biol*, 337(1), pp. 16-28.
Guo, Z., Zhang, W., Xia, G., Niu, L., Zhang, Y., Wang, X., Zhang, Y., Jiang, B. and Wang, J. (2010) 'Sp1 upregulates the four and half lim 2 (FHL2) expression in gastrointestinal cancers through transcription regulation', *Molecular carcinogenesis*, 49(9).
Haberle, V. and Stark, A. (2018) 'Eukaryotic core promoters and the functional basis of transcription initiation', *Nature reviews. Molecular cell biology*, 19(10).
Habiballa, L., Salmonowicz, H. and Passos, J. (2019) 'Mitochondria and cellular senescence: Implications for musculoskeletal ageing', *Free radical biology & medicine*, 132.
Habibe, J.J., Clemente-Olivo, M.P. and de Vrie, s.C.J. (2021) 'How (Epi)Genetic Regulation of the LIM-Domain Protein FHL2 Impacts Multifactorial Disease', *Cells*, 10(10).

Hamidouche, Z., Haÿ, E., Vaudin, P., Charbord, P., Schüle, R., Marie, P. and Fromigué, O. (2008) 'FHL2 mediates dexamethasone-induced mesenchymal cell differentiation into osteoblasts by activating Wnt/beta-catenin signaling-dependent Runx2 expression', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 22(11).

Hamilton, J. (2011) 'Epigenetics: principles and practice', *Digestive diseases (Basel, Switzerland)*, 29(2).

Hampsey, M. (1998) 'Molecular genetics of the RNA polymerase II general transcriptional machinery', *Microbiology and molecular biology reviews : MMBR*, 62(2).

Han, W., Cai, S., Liu, Z., Jin, X., Wang, X., Antony, B., Cao, Y., Aitken, D., Cicuttini, F., Jones, G. and Ding, C. (2014) 'Infrapatellar fat pad in the knee: is local fat good or bad for knee osteoarthritis?', *Arthritis Res Ther*, 16(4), p. R145.

Han, X., Xiao, C., Yi, S., Li, Y., Chen, M. and Huang, D. (2022) 'Accurate age estimation from blood samples of Han Chinese individuals using eight high-performance age-related CpG sites', *International journal of legal medicine*, 136(6).

Han, Y., Eipel, M., Franzen, J., Sakk, V., Dethmers-Ausema, B., Yndriago, L., Izeta, A., de Haan, G., Geiger, H. and Wagner, W. (2018) 'Epigenetic age-predictor for mice based on three CpG sites', *eLife*, 7.

Han, Y., Nikolić, M., Gobs, M., Franzen, J., de Haan, G., Geiger, H. and Wagner, W. (2020) 'Targeted methods for epigenetic age predictions in mice', *Scientific Reports*, 10(1), pp. 1-10.

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T. and Zhang, K. (2013) 'Genome-wide methylation profiles reveal quantitative views of human aging rates', *Mol Cell*, 49(2), pp. 359-367.

Hanzawa, N., Hashimoto, K., Yuan, X., Kawahori, K., Tsujimoto, K., Hamaguchi, M., Tanaka, T., Nagaoka, Y., Nishina, H., Morita, S., Hatada, I., Yamada, T. and Ogawa, Y. (2020) 'Targeted DNA demethylation of the Fgf21 promoter by CRISPR/dCas9-mediated epigenome editing', *Scientific Reports*, 10(1), pp. 1-14.

Hariprakash, J. and Ferrari, F. (2019) 'Computational Biology Solutions to Identify Enhancers-target Gene Pairs', *Computational and structural biotechnology journal*, 17.

Harley, C., Futcher, A. and Greider, C. (1990) 'Telomeres shorten during ageing of human fibroblasts', *Nature*, 345(6274).

Hayden, M. and Ghosh, S. (2008) 'Shared principles in NF-kappaB signaling', *Cell*, 132(3).

He, H.H., Meyer, C.A., Shin, H., Bailey, S.T., Wei, G., Wang, Q., Zhang, Y., Xu, K., Ni, M., Lupien, M., Mieczkowski, P., Lieb, J.D., Zhao, K., Brown, M. and Liu, X.S. (2010) 'Nucleosome dynamics define transcriptional enhancers', *Nat Genet*, 42(4), pp. 343-7.

Heemers, H., Regan, K., Dehm, S. and Tindall, D. (2007) 'Androgen induction of the androgen receptor coactivator four and a half LIM domain protein-2: evidence for a role for serum response factor in prostate cancer', *Cancer research*, 67(21).

Heidari, B. (2011) 'Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I', *Caspian J Intern Med*, 2(2), pp. 205-12.

Heidersbach, A.J., Dorighi, K.M., Gomez, J.A., Jacobi, A.M. and Haley, B. (2023) 'A versatile, high-efficiency platform for CRISPR-based gene activation', *Nature Communications*, 14(1), pp. 1-10.

Hellman, L. and Fried, M. (2007) 'Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions', *Nature protocols*, 2(8).

Helmick, C.G., Felson, D.T., Lawrence, R.C., Gabriel, S., Hirsch, R., Kwoh, C.K., Liang, M.H., Kremers, H.M., Mayes, M.D., Merkel, P.A., Pillemer, S.R., Reveille, J.D. and Stone, J.H. (2008) 'Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I', *Arthritis Rheum*, 58(1), pp. 15-25.

Henrotin, Y., Pesesse, L. and Sanchez, C. (2012) 'Subchondral bone and osteoarthritis: biological and cellular aspects', *Osteoporos Int*, 23 Suppl 8, pp. S847-51.

Hernandez-Gonzalez, F., Faner, R., Rojas, M., Agustí, A., Serrano, M. and Sellarés, J. (2021) 'Cellular Senescence in Lung Fibrosis', *International journal of molecular sciences*, 22(13).

Hess, J., Angel, P. and Schorpp-Kistner, M. (2004) 'AP-1 subunits: quarrel and harmony among siblings', *Journal of cell science*, 117(Pt 25).

Hetz, C., Zhang, K. and Kaufman, R. (2020) 'Mechanisms, regulation and functions of the unfolded protein response', *Nature reviews. Molecular cell biology*, 21(8).

Hewitt, G., Jurk, D., Marques, F.D.M., Correia-Melo, C., Hardy, T., Gackowska, A., Anderson, R., Taschuk, M., Mann, J. and Passos, J.F. (2012) 'Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence', *Nature Communications*, 3(1), pp. 1-9.

Heyn, H., Li, N., Ferreira, H.J., Moran, S., Pisano, D.G., Gomez, A., Diez, J., Sanchez-Mut, J.V., Setien, F., Carmona, F.J., Puca, A.A., Sayols, S., Pujana, M.A., Serra-Musach, J., Iglesias-Platas, I., Formiga, F., Fernandez, A.F., Fraga, M.F., Heath, S.C., Valencia, A., Gut, I.G., Wang, J. and Esteller, M. (2012) 'Distinct DNA methylomes of newborns and centenarians', *Proc Natl Acad Sci U S A*, 109(26), pp. 10522-7.

Hinz, B. (2015) 'The extracellular matrix and transforming growth factor-β1: Tale of a strained relationship', *Matrix biology : journal of the International Society for Matrix Biology*, 47.

Hoeijmakers, J.H. (2009) 'DNA damage, aging, and cancer', *N Engl J Med*, 361(15), pp. 1475-85.
Holden, N. and Tacon, C. (2011) 'Principles and problems of the electrophoretic mobility shift assay', *Journal of pharmacological and toxicological methods*, 63(1).
Honda, H., Pazin, M., Ji, H., Wernyj, R. and Morin, P. (2006) 'Crucial roles of Sp1 and epigenetic modifications in the regulation of the CLDN4 promoter in ovarian cancer cells', *The Journal of biological chemistry*, 281(30).
Hop, P., Zwamborn, R., Hannon, E., Dekker, A., van, E., KR, Walker, E., Iacoangeli, A., Jones, A., Shatunov, A., Khleifat, A., Opie-Martin, S., Shaw, C., Morrison, K., Shaw, P., McLaughlin, R., Hardiman, O., Al-Chalabi, A., Van, D., Berg, LH, Mill, J. and Veldink, J. (2020) 'Cross-reactive probes on Illumina DNA methylation arrays: a large study on ALS shows that a cautionary approach is warranted in interpreting epigenome-wide association studies', *NAR genomics and bioinformatics*, 2(4).
Horvath, Erhart, W., Brosch, M., Ammerpohl, O., Schönfels, W.v., Ahrens, M., Heits, N., Bell, J.T., Tsai, P.-C., Spector, T.D., Deloukas, P., Siebert, R., Sipos, B., Becker, T., Röcken, C., Schafmayer, C. and Hampe, J. (2014) 'Obesity accelerates epigenetic aging of human liver'.
Horvath, S. (2013) 'DNA methylation age of human tissues and cell types', *Genome Biol*, 14(10), p. R115.
Hou, Y., Dan, X., Babbar, M., Wei, Y., Hasselbalch, S., Croteau, D. and Bohr, V. (2019) 'Ageing as a risk factor for neurodegenerative disease', *Nature reviews. Neurology*, 15(10).
Hsieh, Y., Alqadah, A. and Chuang, C. (2016) 'An Optimized Protocol for Electrophoretic Mobility Shift Assay Using Infrared Fluorescent Dye-labeled Oligonucleotides', *Journal of visualized experiments : JoVE*, (117).
Hsu, E., Zemke, N. and Berk, A. (2021) 'Promoter-specific changes in initiation, elongation, and homeostasis of histone H3 acetylation during CBP/p300 inhibition', *eLife*, 10.
Hua, G., He, C., Lv, X., Fan, L., Wang, C., Remmenga, S., Rodabaugh, K., Yang, L., Lele, S., Yang, P., Karpf, A., Davis, J. and Wang, C. (2016) 'The four and a half LIM domains 2 (FHL2) regulates ovarian granulosa cell tumor progression via controlling AKT1 transcription', *Cell death & disease*, 7(7).
Huang, R.C., Lillycrop, K.A., Beilin, L.J., Godfrey, K.M., Anderson, D., Mori, T.A., Rauschert, S., Craig, J.M., Oddy, W.H., Ayonrinde, O.T., Pennell, C.E., Holbrook, J.D. and Melton, P.E. (2019) 'Epigenetic age acceleration in adolescence associates with BMI, inflammation and risk score for middle age cardiovascular disease', *J Clin Endocrinol Metab*.
Hämäläinen, R., Ahlqvist, K., Ellonen, P., Lepistö, M., Logan, A., Otonkoski, T., Murphy, M. and Suomalainen, A. (2015) 'mtDNA Mutagenesis Disrupts Pluripotent Stem Cell Function by Altering Redox Signaling', *Cell reports*, 11(10).

Héberlé, É. and Bardet, A. (2019a) 'Sensitivity of transcription factors to DNA methylation', *Essays in biochemistry*, 63(6).
Héberlé, É. and Bardet, A.F. (2019b) 'Sensitivity of transcription factors to DNA methylation', *Essays in biochemistry*, 63(6).
Ioan-Facsinay, A. and Kloppenburg, M. (2013) 'An emerging player in knee osteoarthritis: the infrapatellar fat pad', *Arthritis Res Ther*, 15(6), p. 225.
Ishibashi, J., Perry, R.L., Asakura, A. and Rudnicki, M.A. (2005) 'MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions', in *J Cell Biol*. pp. 471-82.
Ishibashi, R. and Taguchi, Y. (2021) 'Identification of Enhancers and Promoters in the Genome by Multidimensional Scaling', *Genes*, 12(11).
Islam, Z., Ali, A., Naik, A., Eldaw, M., Decock, J. and Kolatkar, P. (2021) 'Transcription Factors: The Fulcrum Between Cell Development and Carcinogenesis', *Frontiers in oncology*, 11.
Jang, H.S., Shin, W.J., Lee, J.E. and Do, J.T. (2017) 'CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function', *Genes (Basel)*, 8(6).
Jin, Z. and Liu, Y. (2018) 'DNA methylation in human diseases', *Genes Dis*, 5(1), pp. 1-8.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. and Charpentier, E. (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science (New York, N.Y.)*, 337(6096).
Jing-bin, Y., Rong, Z., Can, X., Cheng, H., Yao, L., Cong-rong, W., Wei-ping, J. and Fanyi, Z. (2014) 'Pyrosequencing is an accurate and reliable method for the analysis of heteroplasmy of the A3243G mutation in patients with mitochondrial diabetes', *The Journal of molecular diagnostics : JMD*, 16(4).
Jintaridth, P. and Mutirangura, A. (2010) 'Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences', *Physiol Genomics*, 41(2), pp. 194-200.
Johnson, A.A., Akman, K., Calimport, S.R., Wuttke, D., Stolzing, A. and de Magalhães, J.P. (2012) 'The Role of DNA Methylation in Aging, Rejuvenation, and Age-Related Disease', in *Rejuvenation Res*. pp. 483-94.
Joos, H., Albrecht, W., Laufer, S., Reichel, H. and Brenner, R. (2008) 'IL-1beta regulates FHL2 and other cytoskeleton-related genes in human chondrocytes', *Molecular medicine (Cambridge, Mass.)*, 14(3-4).
Jorgensen, B.G., Berent, R.M., Ha, S.E., Horiguchi, K., Sasse, K.C., Becker, L.S. and Ro, S. (2018) 'DNA methylation, through DNMT1, has an essential role in the development of gastrointestinal smooth muscle cells and disease', *Cell Death Dis*, 9(5).

Jung, S., Lim, S., Hong, S., Lee, E., Shin, K. and Lee, H. (2019) 'DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples', *Forensic science international. Genetics*, 38.

Jurk, D., Wilson, C., Passos, J.F., Oakley, F., Correia-Melo, C., Greaves, L., Saretzki, G., Fox, C., Lawless, C., Anderson, R., Hewitt, G., Pender, S.L., Fullard, N., Nelson, G., Mann, J., van de Sluis, B., Mann, D.A. and von Zglinicki, T. (2014) 'Chronic inflammation induces telomere dysfunction and accelerates ageing in mice', *Nature Communications*, 5(1), pp. 1-14.

Kadonaga, J. (2012) 'Perspectives on the RNA polymerase II core promoter', *Wiley interdisciplinary reviews. Developmental biology*, 1(1).

Kampmann, M. (2018) 'CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine', *ACS chemical biology*, 13(2).

Kanthawang, T., Bodden, J., Joseph, G., Lane, N., Nevitt, M., McCulloch, C. and Link, T. (2021) 'Obese and overweight individuals have greater knee synovial inflammation and associated structural and cartilage compositional degeneration: data from the osteoarthritis initiative', *Skeletal radiology*, 50(1).

Kapoor, M., Martel-Pelletier, J., Lajeunesse, D., Pelletier, J.P. and Fahmi, H. (2011) 'Role of proinflammatory cytokines in the pathophysiology of osteoarthritis', *Nat Rev Rheumatol*, 7(1), pp. 33-42.

Kazmi, N., Elliott, H., Burrows, K., Tillin, T., Hughes, A., Chaturvedi, N., Gaunt, T. and Relton, C. (2020) 'Associations between high blood pressure and DNA methylation', *PloS one*, 15(1).

Keller, K. and Engelhardt, M. (2014) 'Strength and muscle mass loss with aging process. Age and strength loss', *Muscles, ligaments and tendons journal*, 3(4).

Kim, G.H., Ryan, J.J., Marsboom, G. and Archer, S.L. (2011) 'Epigenetic mechanisms of pulmonary hypertension', *Pulm Circ*, 1(3), pp. 347-56.

Kirkwood, T.B. (2005) 'Understanding the odd science of aging', *Cell*, 120(4), pp. 437-47.

Kirkwood, T.B.L. and Austad, S.N. (2000) 'Why do we age?', *Nature*, 408(6809), pp. 233-238.

Kleiber, K., Strebhardt, K. and Martin, B. (2007) 'The biological relevance of FHL2 in tumour cells and its role as a putative cancer target', *Anticancer research*, 27(1A).

Klose, R.J. and Bird, A.P. (2006) 'Genomic DNA methylation: the mark and its mediators', *Trends Biochem Sci*, 31(2), pp. 89-97.

Knott, G. and Doudna, J. (2018) 'CRISPR-Cas guides the future of genetic engineering', *Science (New York, N.Y.)*, 361(6405).

Koch, F., Fenouil, R., Gut, M., Cauchy, P., Albert, T.K., Zacarias-Cabeza, J., Spicuglia, S., de la Chapelle, A.L., Heidemann, M., Hintermair, C., Eick, D., Gut, I., Ferrier, P. and Andrau, J.C. (2011) 'Transcription

initiation platforms and GTF recruitment at tissue-specific enhancers and promoters', *Nat Struct Mol Biol*, 18(8), pp. 956-63.
Kollman, C., Howe, C., Anasetti, C., Antin, J., Davies, S., Filipovich, A., Hegland, J., Kamani, N., Kernan, N., King, R., Ratanatharathorn, V., Weisdorf, D. and Confer, D. (2001) 'Donor characteristics as risk factors in recipients after transplantation of bone marrow from unrelated donors: the effect of donor age', *Blood*, 98(7).
Kolovos, P., Knoch, T.A., Grosveld, F.G., Cook, P.R. and Papantonis, A. (2012) 'Enhancers and silencers: an integrated and simple model for their function', *Epigenetics Chromatin*, 5, p. 1.
Kowald, A., Passos, J. and Kirkwood, T. (2020) 'On the evolution of cellular senescence', *Aging cell*, 19(12).
Kołat, D., Kałuzińska, Ż., Bednarek, A. and Płuciennik, E. (2019) 'The biological characteristics of transcription factors AP-2α and AP-2γ and their importance in various types of cancers', *Bioscience reports*, 39(3).
Kujoth, G., Hiona, A., Pugh, T., Someya, S., Panzer, K., Wohlgemuth, S., Hofer, T., Seo, A., Sullivan, R., Jobling, W., Morrow, J., Van, R., H, Sedivy, J., Yamasoba, T., Tanokura, M., Weindruch, R., Leeuwenburgh, C. and Prolla, T. (2005) 'Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging', *Science (New York, N.Y.)*, 309(5733).
Kumar, P.L. and James, P.F. (2015) 'Identification and characterization of methylation-dependent/independent DNA regulatory elements in the human SLC9B1 gene', *Gene*, 561(2), pp. 235-48.
Kurakula, K., Sommer, D., Sokolovic, M., Moerland, P., Scheij, S., van Loenen, P., Koenis, D., Zelcer, N., van Tiel, C. and de Vries, C. (2015) 'LIM-only protein FHL2 is a positive regulator of liver X receptors in smooth muscle cells involved in lipid homeostasis', *Molecular and cellular biology*, 35(1).
Kurakula, K., Vos, M., Otermin, R., I, Marinković, G., Buettner, R., Heukamp, L., Stap, J., de Waard, V., van Tiel, C. and CJ, d.V. (2014) 'The LIM-only protein FHL2 reduces vascular lesion formation involving inhibition of proliferation and migration of smooth muscle cells', *PloS one*, 9(4).
Kurokawa, S., Yoneda, M., Ogawa, Y., Honda, Y., Kessoku, T., Imajo, K., Saito, S., Nakajima, A. and Hotta, K. (2022) 'Two differentially methylated region networks in nonalcoholic fatty liver disease, viral hepatitis, and hepatocellular carcinoma', *BMC Gastroenterology*, 22(1), pp. 1-9.
Kyrchanova, O. and Georgiev, P. (2021) 'Mechanisms of Enhancer-Promoter Interactions in Higher Eukaryotes', *International journal of molecular sciences*, 22(2).
Labalette, C., Nouët, Y., Levillayer, F., Armengol, C., Renard, C., Soubigou, G., Xia, T., Buendia, M. and Wei, Y. (2008a) 'The LIM-only protein FHL2 mediates ras-induced transformation through cyclin D1 and p53 pathways', *PloS one*, 3(11).

Labalette, C., Nouët, Y., Sobczak-Thepot, J., Armengol, C., Levillayer, F., Gendron, M., Renard, C., Regnault, B., Chen, J., Buendia, M. and Wei, Y. (2008b) 'The LIM-only protein FHL2 regulates cyclin D1 expression and cell proliferation', *The Journal of biological chemistry*, 283(22).
Lambert, S., Jolma, A., Campitelli, L., Das, P., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. and Weirauch, M. (2018) 'The Human Transcription Factors', *Cell*, 172(4).
Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S. and Qi, L.S. (2013) 'CRISPR interference (CRISPRi) for sequence-specific control of gene expression', *Nature Protocols*, 8(11), pp. 2180-2196.
Lawrence, R.C., Felson, D.T., Helmick, C.G., Arnold, L.M., Choi, H., Deyo, R.A., Gabriel, S., Hirsch, R., Hochberg, M.C., Hunder, G.G., Jordan, J.M., Katz, J.N., Kremers, H.M. and Wolfe, F. (2008) 'Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II', *Arthritis Rheum*, 58(1), pp. 26-35.
Lee, C. and Huang, C. (2013) 'LASAGNA: a novel algorithm for transcription factor binding site alignment', *BMC bioinformatics*, 14.
Lee, J.-H., Kim, E.W., Croteau, D.L. and Bohr, V.A. (2020) 'Heterochromatin: an epigenetic point of view in aging', *Experimental & Molecular Medicine*, 52(9), pp. 1466-1474.
Lee, J.H., Ort, T., Ma, K., Picha, K., Carton, J., Marsters, P.A., Lohmander, L.S., Baribaud, F., Song, X.Y. and Blake, S. (2009) 'Resistin is elevated following traumatic joint injury and causes matrix degradation and release of inflammatory cytokines from articular cartilage in vitro', *Osteoarthritis Cartilage*, 17(5), pp. 613-20.
Lee, T. and Young, R. (2013) 'Transcriptional regulation and its misregulation in disease', *Cell*, 152(6).
Leontieva, O.V., Paszkiewicz, G.M. and Blagosklonny, M.V. (2012) 'Mechanistic or mammalian target of rapamycin (mTOR) may determine robustness in young male mice at the cost of accelerated aging', *Aging (Albany NY)*, 4(12), pp. 899-916.
Li, B., Ren, N., Yang, L., Liu, J. and Huang, Q. (2019) 'A qPCR method for genome editing efficiency determination and single-cell clone screening in human cells', *Scientific Reports*, 9(1), pp. 1-15.
Li, E., Bestor, T.H. and Jaenisch, R. (1992) 'Targeted mutation of the DNA methyltransferase gene results in embryonic lethality', *Cell*, 69(6), pp. 915-26.
Li, G., Yin, J., Gao, J., Cheng, T.S., Pavlos, N.J., Zhang, C. and Zheng, M.H. (2013a) 'Subchondral bone in osteoarthritis: insight into risk factors and microstructural changes', *Arthritis Research & Therapy*, 15(6), p. 223.
Li, L. and Davie, J. (2010) 'The role of Sp1 and Sp3 in normal and cancer cell biology', *Annals of anatomy = Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft*, 192(5).

Li, L., He, S., Sun, J. and Davie, J. (2004) 'Gene regulation by Sp1 and Sp3', *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 82(4).
Li, M., Hada, A., Sen, P., Olufemi, L., Hall, M., Smith, B., Forth, S., McKnight, J., Patel, A., Bowman, G., Bartholomew, B. and Wang, M. (2015) 'Dynamic regulation of transcription factors by nucleosome remodeling', *eLife*, 4.
Li, S., Shanghai Medical College, F.U., Shanghai, China, Zhang, J., Shanghai Medical College, F.U., Shanghai, China, Huang, S., Shanghai Medical College, F.U., Shanghai, China, He, X. and Shanghai Medical College, F.U., Shanghai, China (2017) 'Genome-wide analysis reveals that exon methylation facilitates its selective usage in the human transcriptome', *Briefings in Bioinformatics*, 19(5), pp. 754-764.
Li, X., Wang, J., Wang, L., Gao, Y., Feng, G., Li, G., Zou, J., Yu, M., Li, Y.F., Liu, C., Yuan, X.W., Zhao, L., Ouyang, H., Zhu, J.-K., Li, W., Zhou, Q. and Zhang, K. (2022) 'Lipid metabolism dysfunction induced by age-dependent DNA methylation accelerates aging', *Signal Transduction and Targeted Therapy*, 7(1), pp. 1-12.
Li, Y., Wei, X., Zhou, J. and Wei, L. (2013b) 'The Age-Related Changes in Cartilage and Osteoarthritis', *Biomed Res Int*, 2013.
Li, Z., Zhang, Z., Ren, Y., Wang, Y., Fang, J., Yue, H., Ma, S. and Guan, F. (2021) 'Aging and age-related diseases: from mechanisms to therapeutic strategies', *Biogerontology*, 22(2).
Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., Pop, R., Reyon, D., Tsai, S.Q., Mallard, W., Joung, J.K., Rinn, J.L., Gnirke, A. and Meissner, A. (2015) 'Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells', *Nat Genet*, 47(5), pp. 469-78.
Lien, T.G., Borgan, Ø., Reppe, S., Gautvik, K. and Glad, I.K. (2018) 'Integrated analysis of DNA-methylation and gene expression using high-dimensional penalized regression: a cohort study on bone mineral density in postmenopausal women', *BMC Medical Genomics*, 11(1), pp. 1-11.
Lima, T., Li, T.Y., Mottis, A. and Auwerx, J. (2022) 'Pleiotropic effects of mitochondria in aging', *Nature Aging*, 2(3), pp. 199-213.
Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., Yu, M., Tonti-Filippini, J., Heyn, H., Hu, S., Wu, J.C., Rao, A., Esteller, M., He, C., Haghighi, F.G., Sejnowski, T.J., Behrens, M.M. and Ecker, J.R. (2013) 'Global epigenomic reconfiguration during mammalian brain development', *Science*, 341(6146), p. 1237905.
Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A. and Jaenisch, R. (2016) 'Editing DNA Methylation in the Mammalian Genome', *Cell*, 167(1), pp. 233-247.e17.

Liu, Z., Han, S., Wang, Y., Cui, C., Zhu, Q., Jiang, X., Yang, C., Du, H., Yu, C., Li, Q., He, H., Shen, X., Chen, Y., Zhang, Y., Ye, L., Zhang, Z., Li, D., Zhao, X. and Yin, H. (2019) 'The LIM-Only Protein FHL2 is involved in Autophagy to Regulate the Development of Skeletal Muscle Cell', *International journal of biological sciences*, 15(4).

Loeser, R. (2017) 'The Role of Aging in the Development of Osteoarthritis', *Transactions of the American Clinical and Climatological Association*, 128.

Loeser, R.F., Goldring, S.R., Scanzello, C.R. and Goldring, M.B. (2012) 'Osteoarthritis: a disease of the joint as an organ', *Arthritis Rheum*, 64(6), pp. 1697-707.

Logue, M., Smith, A., Wolf, E., Maniates, H., Stone, A., Schichman, S., McGlinchey, R., Milberg, W. and Miller, M. (2017) 'The correlation of methylation levels measured using Illumina 450K and EPIC BeadChips in blood samples', *Epigenomics*, 9(11).

Loh, C., Arya, A., Naema, A., Wong, W., Sethi, G. and Looi, C. (2019) 'Signal Transducer and Activator of Transcription (STATs) Proteins in Cancer and Inflammation: Functions and Therapeutic Implication', *Frontiers in oncology*, 9.

Lombard, D.B., Chua, K.F., Mostoslavsky, R., Franco, S., Gostissa, M. and Alt, F.W. (2005) 'DNA repair, genome stability, and aging', *Cell*, 120(4), pp. 497-512.

Lopes, E., Valls, E., Figueroa, M., Mazur, A., Meng, F., Chiosis, G., Laird, P., Schreiber-Agus, N., Greally, J., Prokhortchouk, E. and Melnick, A. (2008) 'Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines', *Cancer research*, 68(18).

Lorda-Diez, C., Montero, J., Sanchez-Fernandez, C., Garcia-Porrero, J., Chimal-Monroy, J. and Hurle, J. (2018) 'Four and a half domain 2 (FHL2) scaffolding protein is a marker of connective tissues of developing digits and regulates fibrogenic differentiation of limb mesodermal progenitors', *Journal of tissue engineering and regenerative medicine*, 12(4).

Lorenz, J. and Grassel, S. (2014) 'Experimental osteoarthritis models in mice', *Methods Mol Biol*, 1194, pp. 401-19.

Lorenzo, P., Izquierdo, A., Diaz-Lagares, A., Carreira, M., Macias-Gonzalez, M., Sandoval, J., Cueva, J., Lopez-Lopez, R., Casanueva, F. and Crujeiras, A. (2020) 'ZNF577 Methylation Levels in Leukocytes From Women With Breast Cancer Is Modulated by Adiposity, Menopausal State, and the Mediterranean Diet', *Frontiers in endocrinology*, 11.

Lu, W., Yu, T., Liu, S., Li, S., Li, S., Liu, J., Xu, Y., Xing, H., Tian, Z., Tang, K., Rao, Q., Wang, J. and Wang, M. (2017) 'FHL2 interacts with iASPP and impacts the biological functions of leukemia cells', *Oncotarget*, 8(25).

Lyko, F. (2018) 'The DNA methyltransferase family: a versatile toolkit for epigenetic regulation', *Nat Rev Genet*, 19(2), pp. 81-92.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M. and Kroemer, G. (2013) 'The Hallmarks of Aging', *Cell*, 153(6), pp. 1194-217.

Ma, A., Wang, H., Guo, R., Wang, Y., Li, W., Cui, J., Wang, G., Hoffman, A. and Hu, J. (2014) 'Targeted gene suppression by inducing de novo DNA methylation in the gene promoter', *Epigenetics & chromatin*, 7.

Magnusson, M., Lu, E.X., Larsson, P., Ulfhammer, E., Bergh, N., Caren, H. and Jern, S. (2015) 'Dynamic Enhancer Methylation--A Previously Unrecognized Switch for Tissue-Type Plasminogen Activator Expression', *PLoS One*, 10(10), p. e0141805.

Magnusson, S.P., Langberg, H. and Kjaer, M. (2010) 'The pathogenesis of tendinopathy: balancing the response to loading', *Nat Rev Rheumatol*, 6(5), pp. 262-8.

Maierhofer, A., Flunkert, J., Oshima, J., Martin, G.M., Haaf, T. and Horvath, S. (2017) 'Accelerated epigenetic aging in Werner syndrome', *Aging (Albany NY)*, 9(4), pp. 1143-52.

Man, G. and Mologhianu, G. (2014) 'Osteoarthritis pathogenesis – a complex process that involves the entire joint', in *J Med Life*. pp. 37-41.

Marioni, R.E., Suderman, M., Chen, B.H., Horvath, S., Bandinelli, S., Morris, T., Beck, S., Ferrucci, L., Pedersen, N.L., Relton, C.L., Deary, I.J. and Hagg, S. (2019) 'Tracking the Epigenetic Clock Across the Human Life Course: A Meta-analysis of Longitudinal Cohort Data', *J Gerontol A Biol Sci Med Sci*, 74(1), pp. 57-61.

Martel-Pelletier, J. and Pelletier, J.P. (2010) 'Is osteoarthritis a disease involving only cartilage or other articular tissues?', *Eklem Hastalik Cerrahisi*, 21(1), pp. 2-14.

Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A., Hardin, A., Eckalbar, W., Vaisse, C. and Ahituv, N. (2019) 'CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency', *Science (New York, N.Y.)*, 363(6424).

Mathiessen, A. and Conaghan, P.G. (2017) 'Synovitis in osteoarthritis: current understanding with therapeutic implications', in *Arthritis Res Ther*.

Maynard, S., Fang, E., Scheibye-Knudsen, M., Croteau, D. and Bohr, V. (2015) 'DNA Damage, DNA Repair, Aging, and Neurodegeneration', *Cold Spring Harbor perspectives in medicine*, 5(10).

McCartney, D.L., Stevenson, A.J., Walker, R.M., Gibson, J., Morris, S.W., Campbell, A., Murray, A.D., Whalley, H.C., Porteous, D.J., McIntosh, A.M., Evans, K.L., Deary, I.J. and Marioni, R.E. (2018) 'Investigating the relationship between DNA methylation age acceleration and risk factors for Alzheimer's disease', *Alzheimers Dement (Amst)*, 10, pp. 429-437.

McDonald, J.I., Celik, H., Rois, L.E., Fishberger, G., Fowler, T., Rees, R., Kramer, A., Martens, A., Edwards, J.R. and Challen, G.A. (2016) 'Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation', *Biol Open*, 5(6), pp. 866-74.

McHugh, D. and Gil, J. (2018) 'Senescence and aging: Causes, consequences, and therapeutic avenues', *The Journal of cell biology*, 217(1).
McKinney, B.C., Lin, C.-W., Rahman, T., Oh, H., Lewis, D.A., Tseng, G. and Sibille, E. (2019) 'DNA methylation in the human frontal cortex reveals a putative mechanism for age-by-disease interactions', *Translational Psychiatry*, 9(1), p. 39.
McMahon, K., Karunasena, E. and Ahuja, N. (2017) 'The Roles of DNA Methylation in the Stages of Cancer', *Cancer journal (Sudbury, Mass.)*, 23(5).
Mehrnaz, F., M, P.M., Shinwu, J., N, G.-Y.E., Gerda, E., J, W.D. and A, J.P. (2005) 'Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level', *Nucleic acids research*, 33(20).
Mendelson, M.M. (2018) 'Epigenetic Age Acceleration: A Biological Doomsday Clock for Cardiovascular Disease?', *Circ Genom Precis Med*, 11(3), p. e002089.
Messeguer, X., Escudero, R., Farré, D., Núñez, O., Martínez, J. and Albà, M. (2002) 'PROMO: detection of known transcription regulatory elements using species-tailored searches', *Bioinformatics (Oxford, England)*, 18(2).
Messerschmidt, D., Knowles, B. and Solter, D. (2014) 'DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos', *Genes & development*, 28(8).
Metcalf, M., Higuchi-Sanabria, R., Garcia, G., Tsui, C. and Dillin, A. (2020) 'Beyond the cell factory: Homeostatic regulation of and by the UPRER', *Science advances*, 6(29).
Mirza, S., Sharma, G., Pandya, P. and Ralhan, R. (2010) 'Demethylating agent 5-aza-2-deoxycytidine enhances susceptibility of breast cancer cells to anticancer agents', *Mol Cell Biochem*, 342(1-2), pp. 101-9.
Miyake, K., Hirasawa, T., Soutome, M., Itoh, M., Goto, Y.-i., Endoh, K., Takahashi, K., Kudo, S., Nakagawa, T., Yokoi, S., Taira, T., Inazawa, J. and Kubota, T. (2011) 'The protocadherins, PCDHB1 and PCDH7 , are regulated by MeCP2 in neuronal cells and brain tissues: implication for pathogenesis of Rett syndrome', *BMC Neuroscience*, 12(1), p. 81.
Moarii, M., Boeva, V., Vert, J. and Reyal, F. (2015) 'Changes in correlation between promoter methylation and gene expression in cancer', *BMC genomics*, 16.
Mohrin, M., Shin, J., Liu, Y., Brown, K., Luo, H., Xi, Y., Haynes, C. and Chen, D. (2015) 'Stem cell aging. A mitochondrial UPR-mediated metabolic checkpoint regulates hematopoietic stem cell aging', *Science (New York, N.Y.)*, 347(6228).
Molnár, B., Galamb, O., Péterfia, B., Wichmann, B., Csabai, I., Bodor, A., Kalmár, A., Szigeti, K., Barták, B., Nagy, Z., Valcz, G., Patai, Á., Igaz, P. and Tulassay, Z. (2018) 'Gene promoter and exon DNA

methylation changes in colon cancer development - mRNA expression and tumor mutation alterations', *BMC cancer*, 18(1).
Molofsky, A., Slutsky, S., Joseph, N., He, S., Pardal, R., Krishnamurthy, J., Sharpless, N. and Morrison, S. (2006) 'Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing', *Nature*, 443(7110).
Momparler, R.L. (2005) 'Epigenetic therapy of cancer with 5-aza-2'-deoxycytidine (decitabine)', *Semin Oncol*, 32(5), pp. 443-51.
Moore, L.D., Le, T. and Fan, G. (2012) 'DNA Methylation and Its Basic Function', *Neuropsychopharmacology*, 38(1), pp. 23-38.
Moore, L.D., Le, T. and Fan, G. (2013) 'DNA Methylation and Its Basic Function', in *Neuropsychopharmacology*. pp. 23-38.
Moorthy, S. and Mitchell, J. (2016) 'Generating CRISPR/Cas9 Mediated Monoallelic Deletions to Study Enhancer Function in Mouse Embryonic Stem Cells', *Journal of visualized experiments : JoVE*, (110).
Morita, S., Horii, T. and Hatada, I. (2018) 'Editing of DNA Methylation Using dCas9-Peptide Repeat and scFv-TET1 Catalytic Domain Fusions', *Methods in molecular biology (Clifton, N.J.)*, 1767.
Morlon, A. and Sassone-Corsi, P. (2003) 'The LIM-only protein FHL2 is a serum-inducible transcriptional coactivator of AP-1', *Proceedings of the National Academy of Sciences of the United States of America*, 100(7).
Moskalev, E., Zavgorodnij, M., Majorova, S., Vorobjev, I., Jandaghi, P., Bure, I. and Hoheisel, J. (2011) 'Correction of PCR-bias in quantitative DNA methylation studies by means of cubic polynomial regression', *Nucleic acids research*, 39(11).
Muir, P., McCarthy, J., Radtke, C.L., Markel, M.D., Santschi, E.M., Scollay, M.C. and Kalscheur, V.L. (2006) 'Role of endochondral ossification of articular cartilage and functional adaptation of the subchondral plate in the development of fatigue microcracking of joints', *Bone*, 38(3), pp. 342-9.
Musumeci, G., Aiello, F.C., Szychlinska, M.A., Di Rosa, M., Castrogiovanni, P. and Mobasheri, A. (2015a) 'Osteoarthritis in the XXIst Century: Risk Factors and Behaviours that Influence Disease Onset and Progression', in *Int J Mol Sci*. pp. 6093-112.
Musumeci, G., Szychlinska, M.A. and Mobasheri, A. (2015b) 'Age-related degeneration of articular cartilage in the pathogenesis of osteoarthritis: molecular markers of senescent chondrocytes', *Histol Histopathol*, 30(1), pp. 1-12.
Myers, S.L., Brandt, K.D., Ehlich, J.W., Braunstein, E.M., Shelbourne, K.D., Heck, D.A. and Kalasinski, L.A. (1990) 'Synovial inflammation in patients with early osteoarthritis of the knee', *J Rheumatol*, 17(12), pp. 1662-9.

Myung, N., Zhu, X., Kruman, I., Castellani, R., Petersen, R., Siedlak, S., Perry, G., Smith, M. and Lee, H. (2008) 'Evidence of DNA damage in Alzheimer disease: phosphorylation of histone H2AX in astrocytes', *Age (Dordrecht, Netherlands)*, 30(4).

Nakazawa, N., Sathe, A., Shivashankar, G. and Sheetz, M. (2016) 'Matrix mechanics controls FHL2 movement to the nucleus to activate p21 expression', *Proceedings of the National Academy of Sciences of the United States of America*, 113(44).

Nam, A.-R., Lee, K.-H., Hwang, H.-J., Schabort, J.J., An, J.-H., Won, S.-H. and Cho, J.-Y. (2020) 'Alternative methylation of intron motifs is associated with cancer-related gene expression in both canine mammary tumor and human breast cancer', *Clinical Epigenetics*, 12(1), pp. 1-15.

Narita, M., Nũnez, S., Heard, E., Narita, M., Lin, A., Hearn, S., Spector, D., Hannon, G. and Lowe, S. (2003) 'Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence', *Cell*, 113(6).

Negrini, S., Gorgoulis, V. and Halazonetis, T. (2010) 'Genomic instability--an evolving hallmark of cancer', *Nature reviews. Molecular cell biology*, 11(3).

Nevalainen, T., Kananen, L., Marttila, S., Jylhävä, J., Mononen, N., Kähönen, M., Raitakari, O.T., Hervonen, A., Jylhä, M., Lehtimäki, T. and Hurme, M. (2017) 'Obesity accelerates epigenetic aging in middle-aged but not in elderly individuals', *Clin Epigenetics*, 9.

O'Connor, L., Gilmour, J. and Bonifer, C. (2016) 'The Role of the Ubiquitously Expressed Transcription Factor Sp1 in Tissue-specific Transcriptional Regulation and in Disease', *The Yale journal of biology and medicine*, 89(4).

O'Shea, J., Schwartz, D., Villarino, A., Gadina, M., McInnes, I. and Laurence, A. (2015) 'The JAK-STAT pathway: impact on human disease and therapeutic intervention', *Annual review of medicine*, 66.

Oeckinghaus, A. and Ghosh, S. (2009) 'The NF-kappaB family of transcription factors and its regulation', *Cold Spring Harbor perspectives in biology*, 1(4).

Oehler, S., Neureiter, D., Meyer-Scholten, C. and Aigner, T. (2002) 'Subtyping of osteoarthritic synoviopathy', *Clin Exp Rheumatol*, 20(5), pp. 633-40.

Ordoñez, R., Martínez-Calle, N., Agirre, X. and Prosper, F. (2019) 'DNA Methylation of Enhancer Elements in Myeloid Neoplasms: Think Outside the Promoters?', *Cancers*, 11(10).

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., Guigo, R. and Shiekhattar, R. (2010) 'Long noncoding RNAs with enhancer-like function in human cells', *Cell*, 143(1), pp. 46-58.

O'Geen, H., Ren, C., Nicolet, C.M., Perez, A.A., Halmai, J., Le, V.M., Mackay, J.P., Farnham, P.J. and Segal, D.J. (2017) 'dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression', *Nucleic Acids Res*, 45(17), pp. 9901-16.

Pagano, J., Clingman, C. and Ryder, S. (2011) 'Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes', *RNA (New York, N.Y.)*, 17(1).
Pal, S. and Tyler, J.K. (2016) 'Epigenetics and aging', *Sci Adv*, 2(7).
Palii, S., Van Emburgh, B., Sankpal, U., Brown, K. and Robertson, K. (2008) 'DNA methylation inhibitor 5-Aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B', *Molecular and cellular biology*, 28(2).
Pan, C., Yi, S., Xiao, C., Huang, Y., Chen, X. and Huang, D. (2020) 'The evaluation of seven age-related CpGs for forensic purpose in blood from Chinese Han population', *Forensic science international. Genetics*, 46.
Panjarian, S., Madzo, J., Keith, K., Slater, C., Sapienza, C., Jelinek, J. and Issa, J. (2021a) 'Accelerated aging in normal breast tissue of women with breast cancer', *Breast cancer research : BCR*, 23(1).
Panjarian, S., Madzo, J., Keith, K., Slater, C.M., Sapienza, C., Jelinek, J. and Issa, J.-P.J. (2021b) 'Accelerated aging in normal breast tissue of women with breast cancer', *Breast Cancer Research*, 23(1), pp. 1-14.
Paparazzo, E., Lagani, V., Geracitano, S., Citrigno, L., Aceto, M., Malvaso, A., Bruno, F., Passarino, G. and Montesanto, A. (2023) 'An ELOVL2-Based Epigenetic Clock for Forensic Age Prediction: A Systematic Review', *International journal of molecular sciences*, 24(3).
Park, J., Will, C., Martin, B., Gullotti, L., Friedrichs, N., Buettner, R., Schneider, H., Ludwig, S. and Wixler, V. (2008) 'Deficiency in the LIM-only protein FHL2 impairs assembly of extracellular matrix proteins', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 22(7).
Passos, J., Saretzki, G., Ahmed, S., Nelson, G., Richter, T., Peters, H., Wappler, I., Birket, M., Harold, G., Schaeuble, K., Birch-Machin, M., Kirkwood, T. and von Zglinicki, T. (2007) 'Mitochondrial dysfunction accounts for the stochastic heterogeneity in telomere-dependent senescence', *PLoS biology*, 5(5).
Patel, K., Dickson, J., Din, S., Macleod, K., Jodrell, D. and Ramsahoye, B. (2010) 'Targeting of 5-aza-2'-deoxycytidine residues by chromatin-associated DNMT1 induces proteasomal degradation of the free enzyme', *Nucleic Acids Res*, 38(13), pp. 4313-24.
Pathania, R., Ramachandran, S., Elangovan, S., Padia, R., Yang, P., Cinghu, S., Veeranan-Karmegam, R., Arjunan, P., Gnana-Prakasam, J.P., Fulzele, S., Pei, L., Chang, C.S., Choi, H., Shi, H., Manicassamy, S., Prasad, P.D., Sharma, S., Ganapathy, V., Jothi, R. and Thangaraju, M. (2015) 'DNMT1 is essential for mammary and cancer stem cell maintenance and tumorigenesis', *Nat Commun*, 6, p. 6910.
Paul, C., Lacroix, M., Iankova, I., Julien, E., Schäfer, B., Labalette, C., Wei, Y., Le Cam, A., Le Cam, L. and Sardet, C. (2006) 'The LIM-only protein FHL2 is a negative regulator of E4F1', *Oncogene*, 25(40).

Pearle, A.D., Warren, R.F. and Rodeo, S.A. (2005) 'Basic science of articular cartilage and osteoarthritis', *Clin Sports Med*, 24(1), pp. 1-12.

Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A. and Bejerano, G. (2013) 'Enhancers: five essential questions', *Nat Rev Genet*, 14(4), pp. 288-95.

Perkins, N.D., Agranoff, A.B., Pascal, E. and Nabel, G.J. (1994) 'An interaction between the DNA-binding domains of RelA(p65) and Sp1 mediates human immunodeficiency virus gene activation', *Mol Cell Biol*, 14(10), pp. 6570-83.

Perkins, N.D., Edwards, N.L., Duckett, C.S., Agranoff, A.B., Schmid, R.M. and Nabel, G.J. (1993) 'A cooperative interaction between NF-kappa B and Sp1 is required for HIV-1 enhancer activation', *EMBO J*, 12(9), pp. 3551-8.

Pflueger, C., Tan, D., Swain, T., Nguyen, T., Pflueger, J., Nefzger, C., Polo, J.M., Ford, E. and Lister, R. (2018a) 'A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs', *Genome Res*, 28(8), pp. 1193-206.

Pflueger, C., Tan, D., Swain, T., Nguyen, T.V., Pflueger, J., Nefzger, C., Polo, J.M., Ford, E. and Lister, R. (2018b) 'A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs', *Genome Research*.

Philp, A., Collier, R., Grover, L., Davis, E. and Jones, S. (2017) 'Resistin promotes the abnormal Type I collagen phenotype of subchondral bone in obese patients with end stage hip osteoarthritis', *Scientific reports*, 7(1).

Purcell, N., Darwis, D., Bueno, O., Müller, J., Schüle, R. and Molkentin, J. (2004) 'Extracellular signal-regulated kinase 2 interacts with and is negatively regulated by the LIM-only protein FHL2 in cardiomyocytes', *Molecular and cellular biology*, 24(3).

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) 'Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression', *Cell*, 152(5), pp. 1173-83.

Qian, Z., Mao, L., Fernald, A., Yu, H., Luo, R., Jiang, Y., Anastasi, J., Valk, P., Delwel, R. and Le Beau, M. (2009) 'Enhanced expression of FHL2 leads to abnormal myelopoiesis in vivo', *Leukemia*, 23(9).

Qin, T., Lee, C., Li, S., Cavalcante, R., Orchard, P., Yao, H., Zhang, H., Wang, S., Patil, S., Boyle, A. and Sartor, M. (2022) 'Comprehensive enhancer-target gene assignments improve gene set level interpretation of genome-wide regulatory data', *Genome biology*, 23(1).

Raddatz, G., Hagemann, S., Aran, D., Sohle, J., Kulkarni, P.P., Kaderali, L., Hellman, A., Winnefeld, M. and Lyko, F. (2013) 'Aging is associated with highly defined epigenetic changes in the human epidermis', *Epigenetics Chromatin*, 6(1), p. 36.

Ramos, M.P., Wijetunga, N.A., McLellan, A.S., Suzuki, M. and Greally, J.M. (2015) 'DNA demethylation by 5-aza-2′-deoxycytidine is imprinted, targeted to euchromatin, and has limited transcriptional consequences', *Epigenetics Chromatin*, 8.

Raney, B., Dreszer, T., Barber, G., Clawson, H., Fujita, P., Wang, T., Nguyen, N., Paten, B., Zweig, A., Karolchik, D. and Kent, W. (2014) 'Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser', *Bioinformatics (Oxford, England)*, 30(7).

Rattan, S.I. (2006) 'Theories of biological aging: genes, proteins, and free radicals', *Free Radic Res*, 40(12), pp. 1230-8.

Rauluseviciute, I., Drabløs, F. and Rye, M. (2020) 'DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation', *BMC medical genomics*, 13(1).

Rea, I., Gibson, D., McGilligan, V., McNerlan, S., Alexander, H. and Ross, O. (2018) 'Age and Age-Related Diseases: Role of Inflammation Triggers and Cytokines', *Frontiers in immunology*, 9.

Reamon-Buettner, S. and Borlak, J. (2008) 'Epigenetic silencing of cell adhesion molecule 1 in different cancer progenitor cells of transgenic c-Myc and c-Raf mouse lung tumors', *Cancer research*, 68(18).

Redman, M., King, A., Watson, C. and King, D. (2016) 'What is CRISPR/Cas9?', *Archives of disease in childhood. Education and practice edition*, 101(4).

Reiter, F., Wienerroither, S. and Stark, A. (2017) 'Combinatorial function of transcription factors and cofactors', *Current opinion in genetics & development*, 43.

Replogle, J., Bonnar, J., Pogson, A., Liem, C., Maier, N., Ding, Y., Russell, B., Wang, X., Leng, K., Guna, A., Norman, T., Pak, R., Ramos, D., Ward, M., Gilbert, L., Kampmann, M., Weissman, J. and Jost, M. (2022) 'Maximizing CRISPRi efficacy and accessibility with dual-sgRNA libraries and optimal effectors', *eLife*, 11.

Reynard, L.N. (2017) 'Analysis of genetics and DNA methylation in osteoarthritis: What have we learnt about the disease?', *Semin Cell Dev Biol*, 62, pp. 57-66.

Reynard, L.N., Bui, C., Syddall, C.M. and Loughlin, J. (2014) 'CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383', *Hum Genet*, 133(8), pp. 1059-73.

Reynard, L.N. and Loughlin, J. (2012) 'Genetics and epigenetics of osteoarthritis', *Maturitas*, 71(3), pp. 200-4.

Reynard, L.N., Ratnayake, M., Santibanez-Koref, M. and Loughlin, J. (2016) 'Functional Characterization of the Osteoarthritis Susceptibility Mapping to CHST11-A Bioinformatics and Molecular Study', *PLoS One*, 11(7), p. e0159024.

RH, H. (2019) 'Mitochondrial Dysfunction in Aging and Diseases of Aging', *Biology*, 8(2).
Ringrose, L. and Paro, R. (2004) 'Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins', *Annu Rev Genet*, 38, pp. 413-43.
Rippe, K., Schrader, A., Riede, P., Strohner, R., Lehmann, E. and Längst, G. (2007) 'DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes', *Proceedings of the National Academy of Sciences of the United States of America*, 104(40).
Robertson, K.D. (2005) 'DNA methylation and human disease', *Nat Rev Genet*, 6(8), pp. 597-610.
Rodier, F., Coppe, J.P., Patil, C.K., Hoeijmakers, W.A., Munoz, D.P., Raza, S.R., Freund, A., Campeau, E., Davalos, A.R. and Campisi, J. (2009) 'Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion', *Nat Cell Biol*, 11(8), pp. 973-9.
Roessler, J., Ammerpohl, O., Gutwein, J., Hasemeier, B., Anwar, S., Kreipe, H. and Lehmann, U. (2012) 'Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc', *BMC research notes*, 5.
Rosenbloom, K., Sloan, C., Malladi, V., Dreszer, T., Learned, K., Kirkup, V., Wong, M., Maddren, M., Fang, R., Heitner, S., Lee, B., Barber, G., Harte, R., Diekhans, M., Long, J., Wilder, S., Zweig, A., Karolchik, D., Kuhn, R., Haussler, D. and Kent, W. (2013) 'ENCODE data in the UCSC Genome Browser: year 5 update', *Nucleic acids research*, 41(Database issue).
Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J. and Letellier, T. (2003) 'Mitochondrial threshold effects', *The Biochemical journal*, 370(Pt 3).
Rushton, M.D., Reynard, L.N., Barter, M.J., Refaie, R., Rankin, K.S., Young, D.A. and Loughlin, J. (2014) 'Characterization of the Cartilage DNA Methylome in Knee and Hip Osteoarthritis', *Arthritis Rheumatol*, 66(9), pp. 2450-60.
Rushton, M.D., Reynard, L.N., Young, D.A., Shepherd, C., Aubourg, G., Gee, F., Darlay, R., Deehan, D., Cordell, H.J. and Loughlin, J. (2015) 'Methylation quantitative trait locus analysis of osteoarthritis links epigenetics with genetic risk', *Hum Mol Genet*, 24(25), pp. 7432-44.
Rönn, T., Volkov, P., Gillberg, L., Kokosar, M., Perfilyev, A., Jacobsen, A., Jørgensen, S., Brøns, C., Jansson, P., Eriksson, K., Pedersen, O., Hansen, T., Groop, L., Stener-Victorin, E., Vaag, A., Nilsson, E. and Ling, C. (2015) 'Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood', *Human molecular genetics*, 24(13).
Saddik, D., McNally, E.G. and Richardson, M. (2004) 'MRI of Hoffa's fat pad', *Skeletal Radiol*, 33(8), pp. 433-44.
Sakabe, N.J., Savic, D. and Nobrega, M.A. (2012) 'Transcriptional enhancers in development and disease', *Genome Biol*, 13(1), p. 238.

Salahudeen, M. and Nishtala, P. (2017) 'An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice', *Saudi pharmaceutical journal : SPJ : the official publication of the Saudi Pharmaceutical Society*, 25(2).
Salama, R., Sadaie, M., Hoare, M. and Narita, M. (2014) 'Cellular senescence and its effector programs', *Genes & development*, 28(2).
Sandelin, A., Alkema, W., Engström, P., Wasserman, W. and Lenhard, B. (2004) 'JASPAR: an open-access database for eukaryotic transcription factor binding profiles', *Nucleic acids research*, 32(Database issue).
Santos, K.F., Mazzola, T.N. and Carvalho, H.F. (2005) 'The prima donna of epigenetics: the regulation of gene expression by DNA methylation', *Braz J Med Biol Res*, 38(10), pp. 1531-41.
Santra, M., Dill, K. and de Graff, A. (2019) 'Proteostasis collapse is a driver of cell aging and death', *Proceedings of the National Academy of Sciences of the United States of America*, 116(44).
Saul, D. and Kosinsky, R. (2021) 'Epigenetics of Aging and Aging-Associated Diseases', *International journal of molecular sciences*, 22(1).
Scanzello, C.R. and Goldring, S.R. (2012) 'The role of synovitis in osteoarthritis pathogenesis', *Bone*, 51(2), pp. 249-57.
Schaffner, W. (2015) 'Enhancers, enhancers - from their discovery to today's universe of transcription enhancers', *Biol Chem*, 396(4), pp. 311-27.
Schmitz, M., Shaban, M., Albert, B., Gökçen, A. and Kracht, M. (2018) 'The Crosstalk of Endoplasmic Reticulum (ER) Stress Pathways with NF-κB: Complex Mechanisms Relevant for Cancer, Inflammation and Infection', *Biomedicines*, 6(2).
Scholl, F., McLoughlin, P., Ehler, E., de, G., C and Schäfer, B. (2000) 'DRAL is a p53-responsive gene whose four and a half LIM domain protein product induces apoptosis', *The Journal of cell biology*, 151(3).
Schumacher, B., Garinis, G.A. and Hoeijmakers, J.H. (2008) 'Age to survive: DNA damage and aging', *Trends Genet*, 24(2), pp. 77-85.
Scott, A.J., Ellison, M. and Sinclair, D.A. (2021) 'The economic value of targeting aging', *Nature Aging*, 1(7), pp. 616-623.
Sellam, J. and Berenbaum, F. (2010) 'The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis', *Nat Rev Rheumatol*, 6(11), pp. 625-35.
Seo, M., Lei, L. and Egli, M. (2019) 'Label-Free Electrophoretic Mobility Shift Assay (EMSA) for Measuring Dissociation Constants of Protein-RNA Complexes', *Current protocols in nucleic acid chemistry*, 76(1).

Shah, P., Donahue, G., Otte, G., Capell, B., Nelson, D., Cao, K., Aggarwala, V., Cruickshanks, H., Rai, T., McBryan, T., Gregory, B., Adams, P. and Berger, S. (2013) 'Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape', *Genes & development*, 27(16).

Shane Anderson, A. and Loeser, R.F. (2010) 'Why is osteoarthritis an age-related disease?', *Best Pract Res Clin Rheumatol*, 24(1), pp. 15-26.

Sharma, G. and Goodwin, J. (2006) 'Effect of aging on respiratory system physiology and immunology', *Clinical interventions in aging*, 1(3).

Shintaku, J., Peterson, J., Talbert, E., Gu, J., Ladner, K., Williams, D., Mousavi, K., Wang, R., Sartorelli, V. and Guttridge, D. (2016) 'MyoD Regulates Skeletal Muscle Oxidative Metabolism Cooperatively with Alternative NF-κB', *Cell reports*, 17(2).

Sierra, M.I., Fernández, A.F. and Fraga, M.F. (2015) 'Epigenetics of Aging', *Curr Genomics*, 16(6), pp. 435-40.

Slieker, R.C., Relton, C.L., Gaunt, T.R., Slagboom, P.E. and Heijmans, B.T. (2018) 'Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception', *Epigenetics Chromatin*, 11.

Slieker, R.C., van Iterson, M., Luijk, R., Beekman, M., Zhernakova, D.V., Moed, M.H., Mei, H., van Galen, M., Deelen, P., Bonder, M.J., Zhernakova, A., Uitterlinden, A.G., Tigchelaar, E.F., Stehouwer, C.D.A., Schalkwijk, C.G., van der Kallen, C.J.H., Hofman, A., van Heemst, D., de Geus, E.J., van Dongen, J., Deelen, J., van den Berg, L.H., van Meurs, J., Jansen, R., 't Hoen, P.A.C., Franke, L., Wijmenga, C., Veldink, J.H., Swertz, M.A., van Greevenbroek, M.M.J., van Duijn, C.M., Boomsma, D.I., Slagboom, P.E. and Heijmans, B.T. (2016) 'Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms', in *Genome Biol*.

Smith, R.G., Pishva, E., Shireby, G., Smith, A.R., Roubroeks, J.A.Y., Hannon, E., Wheildon, G., Mastroeni, D., Gasparoni, G., Riemenschneider, M., Giese, A., Sharp, A.J., Schalkwyk, L., Haroutunian, V., Viechtbauer, W., van den Hove, D.L.A., Weedon, M., Brokaw, D., Francis, P.T., Thomas, A.J., Love, S., Morgan, K., Walter, J., Coleman, P.D., Bennett, D.A., De Jager, P.L., Mill, J. and Lunnon, K. (2021) 'A meta-analysis of epigenome-wide association studies in Alzheimer's disease highlights novel differentially methylated loci across cortex', *Nature Communications*, 12(1), pp. 1-13.

Smith, Z.D. and Meissner, A. (2013) 'DNA methylation: roles in mammalian development', *Nat Rev Genet*, 14(3), pp. 204-20.

Soares, J.P., Cortinhas, A., Bento, T., Leitão, J.C., Collins, A.R., Gaivã, I. and Mota, M.P. (2014) 'Aging and DNA damage in humans: a meta-analysis study', *Aging (Albany NY)*, 6(6), pp. 432-9.

Sophia Fox, A.J., Bedi, A. and Rodeo, S.A. (2009) 'The Basic Science of Articular Cartilage: Structure, Composition, and Function', *Sports Health*, 1(6), pp. 461-8.
Soul, J., Hardingham, T., Boot-Handford, R. and Schwartz, J. (2019) 'SkeletalVis: an exploration and meta-analysis data portal of cross-species skeletal transcriptomics data', *Bioinformatics (Oxford, England)*, 35(13).
Spinelli, G., Biddeci, G., Artale, A., Valentino, F., Tarantino, G., Gallo, G., Gianguzza, F., Conaldi, P.G., Corrao, S., Gervasi, F., Aronica, T.S., Di Leonardo, A., Duro, G. and Di Blasi, F. (2021) 'A new p65 isoform that bind the glucocorticoid hormone and is expressed in inflammation liver diseases and COVID-19', *Sci Rep*, 11(1), p. 22913.
Spitz, F. and Furlong, E. (2012) 'Transcription factors: from enhancer binding to developmental control', *Nature reviews. Genetics*, 13(9).
Sridharan, G. and Shankar, A. (2012) 'Toluidine blue: A review of its chemistry and clinical utility', *Journal of oral and maxillofacial pathology : JOMFP*, 16(2).
Stabellini, G., De Mattei, M., Calastrini, C., Gagliano, N., Moscheni, C., Pasello, M., Pellati, A., Bellucci, C. and Gioia, M. (2003) 'Effects of interleukin-1beta on chondroblast viability and extracellular matrix changes in bovine articular cartilage explants', *Biomed Pharmacother*, 57(7), pp. 314-9.
Stead, J., Keen, J. and McDowall, K. (2006) 'The identification of nucleic acid-interacting proteins using a simple proteomics-based approach that directly incorporates the electrophoretic mobility shift assay', *Molecular & cellular proteomics : MCP*, 5(9).
Steegenga, WT, Boekschoten, M., Lute, C., Hooiveld, G., de Groot, P., Morris, T., Teschendorff, A., Butcher, L., Beck, S. and Müller, M. (2014) 'Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs', *Age (Dordrecht, Netherlands)*, 36(3).
Stepper, P., Kungulovski, G., Jurkowska, R.Z., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., Jeltsch, A. and Jurkowski, T.P. (2017) 'Efficient targeted DNA methylation with chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase', *Nucleic Acids Res*, 45(4), pp. 1703-13.
Stoczynska-Fidelus, E., Węgierska, M., Kierasińska, A., Ciunowicz, D. and Rieske, P. (2022) 'Role of Senescence in Tumorigenesis and Anticancer Therapy', *Journal of oncology*, 2022.
Stresemann, C. and Lyko, F. (2008) 'Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine', *Int J Cancer*, 123(1), pp. 8-13.
Sun, N., Youle, R. and Finkel, T. (2016) 'The Mitochondrial Basis of Aging', *Molecular cell*, 61(5).
Sun, X., Phua, D., Axiotakis, L., Smith, M., Blankman, E., Gong, R., Cail, R., Espinosa, d., Los, Reyes, S, Beckerle, M., Waterman, C. and Alushin, G. (2020) 'Mechanosensing through Direct Binding of Tensed F-Actin by LIM Domains', *Developmental cell*, 55(4).
Swindell, W.R. (2017) 'Rapamycin in mice', *Aging (Albany NY)*, 9(9), pp. 1941-2.

SY, K., HT, K., JA, H. and SC, P. (2012) 'The transcription factor Sp1 is responsible for aging-dependent altered nucleocytoplasmic trafficking', *Aging cell*, 11(6).

Szyf, M. (2008) 'The role of dna hypermethylation and demethylation in cancer and cancer therapy', *Curr Oncol*, 15(2), pp. 72-5.

T'Jonck, W., Guilliams, M. and Bonnardel, J. (2018) 'Niche signals and transcription factors involved in tissue-resident macrophage development', *Cellular immunology*, 330.

Teschendorff, A.E., Menon, U., Gentry-Maharaj, A., Ramus, S.J., Weisenberger, D.J., Shen, H., Campan, M., Noushmehr, H., Bell, C.G., Maxwell, A.P., Savage, D.A., Mueller-Holzner, E., Marth, C., Kocjan, G., Gayther, S.A., Jones, A., Beck, S., Wagner, W., Laird, P.W., Jacobs, I.J. and Widschwendter, M. (2010) 'Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer', *Genome Res*, 20(4), pp. 440-6.

Tian, H., Lun, S., Huang, H., He, R., Kong, P., Wang, Q., Li, X. and Feng, Y. (2015) 'DNA Methylation Affects the SP1-regulated Transcription of FOXF2 in Breast Cancer Cells', *The Journal of biological chemistry*, 290(31).

Tomás-Daza, L., Rovirosa, L., López-Martí, P., Nieto-Aliseda, A., Serra, F., Planas-Riverola, A., Molina, O., McDonald, R., Ghevaert, C., Cuatrecasas, E., Costa, D., Camós, M., Bueno, C., Menéndez, P., Valencia, A. and Javierre, B. (2023) 'Low input capture Hi-C (liCHi-C) identifies promoter-enhancer interactions at high-resolution', *Nature communications*, 14(1).

Tran, M., Kurakula, K., Koenis, D. and de Vries, C. (2016) 'Protein-protein interactions of the LIM-only protein FHL2 and functional implication of the interactions relevant in cardiovascular disease', *Biochimica et biophysica acta*, 1863(2).

Tsalamandris, S., Antonopoulos, A., Oikonomou, E., Papamikroulis, G., Vogiatzi, G., Papaioannou, S., Deftereos, S. and Tousoulis, D. (2019) 'The Role of Inflammation in Diabetes: Current Concepts and Future Perspectives', *European cardiology*, 14(1).

Tsunoda, T. and Takagi, T. (1999) 'Estimating transcription factor bindability on DNA', *Bioinformatics (Oxford, England)*, 15(7-8).

Tsurumi, A. and Li, W. (2012) 'Global heterochromatin loss: a unifying theory of aging?', *Epigenetics*, 7(7).

Uddin, M., Tewari, D., Sharma, G., Kabir, M., Barreto, G., Bin-Jumah, M., Perveen, A., Abdel-Daim, M. and Ashraf, G. (2020) 'Molecular Mechanisms of ER Stress and UPR in the Pathogenesis of Alzheimer's Disease', *Molecular neurobiology*, 57(7).

UK, A.R. (2018) *State of Musculoskeletal Health 2018 | Arthritis Research UK*. Available at: https://www.arthritisresearchuk.org/arthritis-information/data-and-statistics/state-of-musculoskeletal-health.aspx.

Urra, H., Dufey, E., Avril, T., Chevet, E. and Hetz, C. (2016) 'Endoplasmic Reticulum Stress and the Hallmarks of Cancer', *Trends in cancer*, 2(5).
Valdez, B.C., Li, Y., Murray, D., Ji, J., Liu, Y., Popat, U., Champlin, R.E. and Andersson, B.S. (2015) 'Comparison of the cytotoxicity of cladribine and clofarabine when combined with fludarabine and busulfan in AML cells: Enhancement of cytotoxicity with epigenetic modulators', *Exp Hematol*, 43(6), pp. 448-61.e2.
van den Beld, A., Kaufman, J., Zillikens, M., Lamberts, S., Egan, J. and van der Lely, A. (2018) 'The physiology of endocrine systems with ageing', *The lancet. Diabetes & endocrinology*, 6(8).
van Deursen, J. (2014) 'The role of senescent cells in ageing', *Nature*, 509(7501).
van Otterdijk, S.D., Mathers, J.C. and Strathdee, G. (2013) 'Do age-related changes in DNA methylation play a role in the development of age-related diseases?', *Biochem Soc Trans*, 41(3), pp. 803-7.
van Saase, J.L., van Romunde, L.K., Cats, A., Vandenbroucke, J.P. and Valkenburg, H.A. (1989) 'Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations', *Ann Rheum Dis*, 48(4), pp. 271-80.
Vandiver, A.R., Irizarry, R.A., Hansen, K.D., Garza, L.A., Runarsson, A., Li, X., Chien, A.L., Wang, T.S., Leung, S.G., Kang, S. and Feinberg, A.P. (2015a) 'Age and sun exposure-related widespread genomic blocks of hypomethylation in nonmalignant skin', *Genome Biol*, 16, p. 80.
Vandiver, A.R., Irizarry, R.A., Hansen, K.D., Garza, L.A., Runarsson, A., Li, X., Chien, A.L., Wang, T.S., Leung, S.G., Kang, S. and Feinberg, A.P. (2015b) 'Age and sun exposure-related widespread genomic blocks of hypomethylation in nonmalignant skin', *Genome Biology*, 16(1), pp. 1-15.
Vanyushin, B.F., Nemirovsky, L.E., Klimenko, V.V., Vasiliev, V.K. and Belozersky, A.N. (1973) 'The 5-methylcytosine in DNA of rats. Tissue and age specificity and the changes induced by hydrocortisone and other agents', *Gerontologia*, 19(3), pp. 138-52.
Verset, L., Tommelein, J., Decaestecker, C., De Vlieghere, E., Bracke, M., Salmon, I., De Wever, O. and Demetter, P. (2017) 'ADAM-17/FHL2 colocalisation suggests interaction and role of these proteins in colorectal cancer', *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*, 39(3).
Verset, L., Tommelein, J., Moles, L., X, Decaestecker, C., Mareel, M., Bracke, M., Salmon, I., De, W., O and Demetter, P. (2013) 'Epithelial expression of FHL2 is negatively associated with metastasis-free and overall survival in colorectal cancer', *British journal of cancer*, 109(1).
Vincent, T. (2019) 'IL-1 in osteoarthritis: time for a critical review of the literature', *F1000Research*, 8.
Vinson, C., Chatterjee, R. and Fitzgerald, P. (2011) 'Transcription factor binding sites and other features in human and Drosophila proximal promoters', *Sub-cellular biochemistry*, 52.

Waby, J.S., Chirakkal, H., Yu, C., Griffiths, G.J., Benson, R.S., Bingle, C.D. and Corfe, B.M. (2010) 'Sp1 acetylation is associated with loss of DNA binding at promoters associated with cell cycle arrest and cell death in a colon cell line', *Molecular Cancer*, 9(1), pp. 1-16.

Wang, G., Wei, L. and Loh, H. (2003) 'Transcriptional regulation of mouse delta-opioid receptor gene by CpG methylation: involvement of Sp3 and a methyl-CpG-binding protein, MBD2, in transcriptional repression of mouse delta-opioid receptor gene in Neuro2A cells', *The Journal of biological chemistry*, 278(42).

Wang, Q., Wang, X., Tian, X., Tang, R. and Xu, X. (2016) 'Four and a half LIM domains 2 contributes to the development of human tongue squamous cell carcinoma', *Journal of molecular histology*, 47(2).

Wang, T., Li, J., Ding, K., Zhang, L., Che, Q., Sun, X., Dai, Y., Sun, W., Bao, M., Wang, X., Yang, L. and Li, Z. (2017) 'The CpG Dinucleotide Adjacent to a κB Site Affects NF-κB Function through Its Methylation', *International journal of molecular sciences*, 18(3).

Weill, U., Krieger, G., Avihou, Z., Milo, R., Schuldiner, M. and Davidi, D. (2019) 'Assessment of GFP Tag Position on Protein Localization and Growth Fitness in Yeast', *Journal of molecular biology*, 431(3).

Wenham, C.Y.J. and Conaghan, P.G. (2010) 'The Role of Synovitis in Osteoarthritis', in *Ther Adv Musculoskelet Dis*. pp. 349-59.

Williams, G.C. (1957) 'Pleiotropy, Natural Selection, and the Evolution of Senescence'. Available at: http://dx.doi.org/.

Wixler, V. (2019) 'The role of FHL2 in wound healing and inflammation', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 33(7).

Wixler, V., Cromme, C., Retser, E., Meyer, L., Smyth, N., Mühlenberg, K., Korb-Pap, A., Koers-Wunrau, C., Sotsios, Y., Bassel-Duby, R., Baeten, D., Tak, P., Niederreiter, B., Redlich, K., Bertrand, J., Skryabin, B., Ludwig, S. and Pap, T. (2015) 'FHL2 regulates the resolution of tissue damage in chronic inflammatory arthritis', *Annals of the rheumatic diseases*, 74(12).

Wodrich, A., Scott, A., Shukla, A., Harris, B. and Giniger, E. (2022) 'The Unfolded Protein Responses in Health, Aging, and Neurodegeneration: Recent Advances and Future Considerations', *Frontiers in molecular neuroscience*, 15.

Wong, K., Morgan, G., Boyle, E., Cheng, A., Yip, K. and Chim, C. (2021) 'A proof-of-concept study for the pathogenetic role of enhancer hypomethylation of MYBPHL in multiple myeloma', *Scientific reports*, 11(1).

Wu, C. and Morris, J.R. (2001) 'Genes, genetics, and epigenetics: a correspondence', in *Science*. United States, pp. 1103-5.

Wu, H. and Zhang, Y. (2014) 'Reversing DNA methylation: mechanisms, genomics, and biological functions', *Cell*, 156(1-2), pp. 45-68.
Wu, M., Wang, J., Tang, W., Zhan, X., Li, Y., Peng, Y., Huang, X., Bai, Y., Zhao, J., Li, A., Chen, C., Chen, Y., Peng, H., Ren, Y., Li, G., Liu, S. and Wang, J. (2016) 'FOXK1 interaction with FHL2 promotes proliferation, invasion and metastasis in colorectal cancer', *Oncogenesis*, 5(11).
Wu, S., Tan, M., Hu, Y., Wang, J., Scheuner, D. and Kaufman, R. (2004) 'Ultraviolet light activates NFkappaB through translational inhibition of IkappaBalpha synthesis', *The Journal of biological chemistry*, 279(33).
Wu, X., Kriz, A. and Sharp, P. (2014) 'Target specificity of the CRISPR-Cas9 system', *Quantitative biology (Beijing, China)*, 2(2).
Wyss-Coray, T. (2016) 'Ageing, neurodegeneration and brain rejuvenation', *Nature*, 539(7628).
Xu, M., Palmer, A., Ding, H., Weivoda, M., Pirtskhalava, T., White, T., Sepe, A., Johnson, K., Stout, M., Giorgadze, N., Jensen, M., LeBrasseur, N., Tchkonia, T. and Kirkland, J. (2015) 'Targeting senescent cells enhances adipogenesis and metabolic function in old age', *eLife*, 4.
Xu, X., Gao, J., Dai, W., Wang, D., Wu, J. and Wang, J. (2019) 'Gene activation by a CRISPR-assisted trans enhancer', *eLife*, 8.
Xu, Y. and Li, Z. (2020) 'CRISPR-Cas systems: Overview, innovations and applications in human disease research and gene therapy', *Computational and structural biotechnology journal*, 18.
Y, W., CA, R., C, L., Y, W., L, L., C, N., X, P., M, F., S, P. and MA, B. (2003) 'Identification of the LIM protein FHL2 as a coactivator of beta-catenin', *The Journal of biological chemistry*, 278(7).
Yan, J., Zhu, J., Zhong, H., Lu, Q., Huang, C. and Ye, Q. (2003) 'BRCA1 interacts with FHL2 and enhances FHL2 transactivation function', *FEBS letters*, 553(1-2).
Yan, J.-b., Zhang, R., Xiong, C., Hu, C., Lv, Y., Wang, C.-r., Jia, W.-p. and Zeng, F. (2014) 'Pyrosequencing is an accurate and reliable method for the analysis of heteroplasmy of the A3243G mutation in patients with mitochondrial diabetes', *The Journal of molecular diagnostics : JMD*, 16(4).
Yang, J., Tian, X., Cui, J., Jiang, S., Shi, R., Liu, Y., Liu, X., Xu, W., Xie, W., Jia, X., Bade, R., Zhang, T., Zhang, M., Gong, K., Yan, S., Yang, Z. and Shao, G. (2017) '5-Aza-2'-deoxycytidine, a DNA methylation inhibitor, induces cytotoxicity, cell cycle dynamics and alters expression of DNA methyltransferase 1 and 3A in mouse hippocampus-derived neuronal HT22 cells', *J Toxicol Environ Health A*, 80(22), pp. 1222-1229.
Yang, X., Han, H., De Carvalho, D.D., Lay, F.D., Jones, P.A. and Liang, G. (2014) 'Gene Body Methylation can alter Gene Expression and is a Therapeutic Target in Cancer', *Cancer Cell*, 26(4), pp. 577-90.

Yang, X., Noushmehr, H., Han, H., Andreu-Vieyra, C., Liang, G. and Jones, P.A. (2012) 'Gene Reactivation by 5-Aza-2'-Deoxycytidine–Induced Demethylation Requires SRCAP–Mediated H2A.Z Insertion to Establish Nucleosome Depleted Regions', *PLoS Genet*, 8(3).

Yang, Y., Hou, H., Haller, E., Nicosia, S. and Bai, W. (2005) 'Suppression of FOXO1 activity by FHL2 through SIRT1-mediated deacetylation', *The EMBO journal*, 24(5).

Yang, Y., Xu, J., Ge, S. and Lai, L. (2021a) 'CRISPR/Cas: Advances, Limitations, and Applications for Precision Cancer Research', *Frontiers in medicine*, 8.

Yang, Y., Zheng, R. and Zhao, L. (2021b) 'Population Aging, Health Investment and Economic Growth: Based on a Cross-Country Panel Data Analysis', *International journal of environmental research and public health*, 18(4).

Yi, S. and Kim, K. (2020) 'New Insights into the Role of Histone Changes in Aging', *International journal of molecular sciences*, 21(21).

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P., Kivioja, T., Dave, K., Zhong, F., Nitta, K., Taipale, M., Popov, A., Ginno, P., Domcke, S., Yan, J., Schübeler, D., Vinson, C. and Taipale, J. (2017) 'Impact of cytosine methylation on DNA binding specificities of human transcription factors', *Science (New York, N.Y.)*, 356(6337).

Yuan, G.H., Tanaka, M., Masuko-Hongo, K., Shibakawa, A., Kato, T., Nishioka, K. and Nakamura, H. (2004) 'Characterization of cells from pannus-like tissue over articular cartilage of advanced osteoarthritis', *Osteoarthritis Cartilage*, 12(1), pp. 38-45.

Yucesoy, B., Charles, L.E., Baker, B. and Burchfiel, C.M. (2015) 'Occupational and genetic risk factors for osteoarthritis: A review', *Work*, 50(2), pp. 261-73.

Yun, C. and DasGupta, R. (2014) 'Luciferase reporter assay in Drosophila and mammalian tissue culture cells', *Current protocols in chemical biology*, 6(1).

Zampieri, M., Ciccarone, F., Calabrese, R., Franceschi, C., Burkle, A. and Caiafa, P. (2015) 'Reconfiguration of DNA methylation in aging', *Mech Ageing Dev*, 151, pp. 60-70.

Zentner, G.E., Tesar, P.J. and Scacheri, P.C. (2011) 'Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions', *Genome Res*, 21(8), pp. 1273-83.

Zhang, W., Li, J., Suzuki, K., Qu, J., Wang, P., Zhou, J., Liu, X., Ren, R., Xu, X., Ocampo, A., Yuan, T., Yang, J., Li, Y., Shi, L., Guan, D., Pan, H., Duan, S., Ding, Z., Li, M., Yi, F., Bai, R., Wang, Y., Chen, C., Yang, F., Li, X., Wang, Z., Aizawa, E., Goebl, A., Soligalla, R., Reddy, P., Esteban, C., Tang, F., Liu, G. and Belmonte, J. (2015) 'Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging', *Science (New York, N.Y.)*, 348(6239).

Zhang, Y. and Jordan, J.M. (2010) 'Epidemiology of Osteoarthritis', *Clin Geriatr Med*, 26(3), pp. 355-69.

Zhang, Y., Najmi, S. and Schneider, D. (2017) 'Transcription factors that influence RNA polymerases I and II: To what extent is mechanism of action conserved?', *Biochimica et biophysica acta. Gene regulatory mechanisms*, 1860(2).

Zhang, Z.M., Lu, R., Wang, P., Yu, Y., Chen, D.L., Gao, L., Liu, S., Ji, D., Rothbart, S.B., Wang, Y., Wang, G.G. and Song, J. (2018) 'Structural basis for DNMT3A-mediated de novo DNA methylation', *Nature*, 554(7692), pp. 387-91.

Zhao, L. and Sumberaz, P. (2020) 'Mitochondrial DNA Damage: Prevalence, Biological Consequence, and Emerging Pathways', *Chemical research in toxicology*, 33(10).

Zhao, Q., Wang, J., Levichkin, I., Stasinopoulos, S., Ryan, M. and Hoogenraad, N. (2002) 'A mitochondrial specific stress response in mammalian cells', *The EMBO journal*, 21(17).

Zheng, Q. and Zhao, Y. (2007) 'The diverse biofunctions of LIM domain proteins: determined by subcellular localization and protein-protein interaction', *Biology of the cell*, 99(9).

Zheng, Y., Lunetta, K., Liu, C., Katrinli, S., Smith, A., Miller, M. and Logue, M. (2022) 'An evaluation of the genome-wide false positive rates of common methods for identifying differentially methylated regions using illumina methylation arrays', *Epigenetics*, 17(13).

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

Zhu, W., Srinivasan, K., Dai, Z., Duan, W., Druhan, L., Ding, H., Yee, L., Villalona-Calero, M., Plass, C. and Otterson, G. (2003) 'Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter', *Molecular and cellular biology*, 23(12).

Zs-Nagy, I., Cutler, R.G. and Semsei, I. (1988) 'Dysdifferentiation hypothesis of aging and cancer: a comparison with the membrane hypothesis of aging', *Ann N Y Acad Sci*, 521, pp. 215-25.

Zykovich, A., Hubbard, A., Flynn, J.M., Tarnopolsky, M., Fraga, M.F., Kerksick, C., Ogborn, D., MacNeil, L., Mooney, S.D. and Melov, S. (2014) 'Genome-wide DNA methylation changes with age in disease-free human skeletal muscle', *Aging Cell*, 13(2), pp. 360-6.