

Epigenetic Mechanisms of Breast Cancer
Prevention by the Type II Diabetes Mellitus
Agent Metformin

Caitriona Tyndall

The Epigenetics Unit, Division of Cancer, Department of Surgery and Cancer, Imperial College London

PhD A3Z1DA Clinical Medicine Research (Surgery and Cancer) FT

WORD COUNT: 32,872

DECLARATION OF ORIGINALITY

This is to declare that the work included in this thesis has been conducted by myself, Caitriona Tyndall.

Work conducted by collaborators, including MRes students, have been acknowledged appropriately.

Additional work has been referenced.

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

JUSTIFICATION

Development of methods for cancer prevention was a growing concern in cancer research. Repurposing drugs as cancer preventative agents provided an attractive alternative to developing expensive methods. Epidemiological data suggested metformin, an anti-type II diabetes mellitus drug, may reduce breast cancer risk and therefore could be a potential breast cancer preventative agent. However, the mechanisms by which metformin interacts with normal breast epithelial cells was little understood. This study aimed to elucidate potential pathways metformin uses to interact with *in vitro* models of normal breast epithelial cells.

ABSTRACT

Breast cancer is the most commonly diagnosed cancer in the UK. Long-term use of metformin, a drug used to treat Type II Diabetes Mellitus, has been linked to a reduced risk of breast cancer in diabetic patients. This study aims to elucidate the mechanisms by which metformin acts as a potential breast cancer preventative agent in non-cancerous breast epithelial cells.

Two models of normal breast epithelial cells were used; MCF12A and MCF10A non-cancerous breast epithelial cells lines, treated with 0mM, 2.5mM and 5mM metformin for 3-days in a range of metabolite levels (glucose, acetate, insulin and L-glutamine). Results were validated by primary normal matched breast epithelial and myoepithelial cells obtained from Breast Cancer Now Tissue Bank, treated with 1.5mM and 1mM metformin, respectively, for 7-days. RNA-sequencing and RT-qPCR, Illumina MethylationEPIC array and pyrosequencing, crystal violet, western blot and HILIC-MS were used to assess the mechanisms of metformin response in normal breast epithelial and myoepithelial cells.

This study demonstrates sensitivity of normal breast epithelial cells to metformin (IC₅₀ 3mM-4mM), and low glucose (5mM) significantly modifies sensitivity to metformin (IC₅₀ 1mM-2mM). Gene expression is significantly altered in response to metformin, where 367 genes in MCF12A and 3 genes in MCF10A showed FDR<0.05. Overall, 7/8 selected genes were validated by RT-qPCR. Metformin subtly altered DNA methylation levels in a dose dependent manner. In addition, metformin appears to alter the expression of key methionine pathway genes. Metformin increases p-AMPK levels and decreases H3K27ac histone modification (0.5-fold, p=0.035). This is modified by glucose or acetate levels. However, effects of metformin appear to return to baseline following 24-hours.

It is estimated that 23% of breast cancer cases could be prevented. This study highlights potential pathways metformin interacts with in normal breast epithelial cells.

DEDICATION AND ACKNOWLEDGEMENTS

I would like to thank my supervisors, James and Marc. Your guidance helped me learn and develop as a scientist. Thank you to my lab at the IRDB, particularly Annelie, Emma, John and Francesca, for all their help and support for the last four years. Thank you for teaching me R and letting me pester you about basic functions (like subsetting data). It was great to have our little PhD support group. Thank you specifically to Emma for supporting me through my depression. You are a star.

I want to acknowledge all the work my three master students, Charlene, Fotini and Daisy, have done. Without their hard work I would not have had the time to finish a massive chunk of my thesis. They are all brilliant scientists in their own right, and I hope they go on to do great things.

Thank you to the CRUK research engagement team and everyone I met from CRUK. You inspired me so much and gave me the opportunity to explore the area I truly love. I will never forget the experiences I had over the last four years talking to patients and the public. Tamsin, Catriona and Sophie you are absolute legends.

Thank you to my friends and family for their support. Thank you for listening me complain about things you didn't understand, for visiting me constantly so I wasn't alone and for propping me up during the harder times. It's been a whirlwind and you've kept me grounded.

Finally, I would like to dedicate this thesis to Liam McConville. I struggled with mental health issues throughout my PhD experience and he was there for me every single step. We had to live apart for over three years and yet I always had his support. Without him, I would not be here. Without him, there would be no thesis. Thank you, Liam.

TABLE OF CONTENTS

Title

Abstract

Acknowledgements

Chapter 1: Introduction

1.1 Type II Diabetes Mellitus and metformin

1.2 Mechanism of metformin in Type II Diabetes Mellitus patients

1.3 Mechanisms of metformin in cancer

1.4 Epidemiological studies linking metformin or Type II Diabetes Mellitus with breast cancer

1.5 Breast cancer and current prevention practices

1.6 Clinical trials studying metformin effects in breast cancer patients

1.7 Epigenetic modifications associated with metformin and Type II Diabetes Mellitus

Chapter 2: Methods

2.1 Samples

2.1.1 Cell lines

2.1.1.1 Cell line media information

2.1.2 Primary Cells

2.1.2.1 Primary cell details

2.1.2.2 Primary cell media information

2.2 Culturing of cell lines

2.2.1 Preparation of MSC safety cabinet

2.2.2 Preparation of cell culture reagents

2.2.3 Resuscitating of cell lines

2.2.4 Subculturing of cell lines

2.2.5 Counting of cell lines

2.2.6 Mycoplasma testing of cell lines

2.2.7 Freezing of cell lines

2.3 Culturing of primary cells

2.3.1 Preparation of MSC safety cabinet

2.3.2 Preparation of cell culture reagents

2.3.3 Resuscitating of primary cells

2.3.4 Collagen coating of plastic ware for primary cell growth

2.4 Cell viability assays

2.4.1 Chemicals

2.4.2 Crystal Violet assay for cell lines

2.4.2.1 Cell line seeding for crystal violet assay

2.4.2.2 Crystal violet

2.4.3 Crystal Violet assay for primary cells

2.4.3.1 Primary cell seeding for crystal violet assay

2.4.3.2 Crystal violet

2.4.4 MTS assay for cell lines

2.5. Protein Study

2.5.1 Cell line seeding for protein extraction

2.5.2 Primary cell seeding for protein extraction

2.5.3 Cytosolic protein extraction

2.5.3.1 Cytosolic protein extraction solution

2.5.3.2 Cell line extraction

2.5.4 Acid extraction

2.5.4.1 Acid extraction solution

2.5.4.2 Cell line acid extraction

2.5.4.3 Primary cell acid extraction

2.5.5 Concentration analysis of protein extracts

2.6 Western Blot

2.6.1 Preparation of gels for SDS-PAGE Gel Electrophoresis

2.6.2 Preparation of protein for SDS-PAGE gel Electrophoresis

2.6.3 SDS-PAGE Gel Electrophoresis

2.6.4 Transfer to PVDF membrane

2.6.4.1 iBlot

2.6.4.2 Wet transfer

2.6.5 Blocking of unspecific binding

2.6.6 Antibody incubation

2.6.7 Chemiluminescence by enhanced chemiluminescence (ECL)

2.6.8 Stripping and re-staining

2.7 DNA Methylation study

2.7.1 Cell line seeding for DNA extraction

2.7.2 DNA Extraction

2.8 Bisulfite conversion

2.8.1 Preparation of whole genome amplified DNA – 100% methylated DNA control.

2.8.2 Purification of PCR product was conducted using the QIAquick PCR purification kit (D5003, Qiagen, 28106).

2.8.2 Bisulfite conversion (EZ DNA Methylation kit, D5002, Zymo).

2.8.3 Analysis of DNA quantity by Qubit

2.8.4 Quality Control of bisulfite converted DNA

2.8.5 PCR of bisulfite converted DNA

2.8.5.1 Preparation of primers

2.8.5.2 Preparation for PCR amplification

2.8.5.3 2% Gel Electrophoresis

2.9 Illumina MethylationEPIC Array

2.10 Pyrosequencing

2.10.1 Preparation of Sequencing Primer(s)

2.10.2 Preparation of PCR plate(s)

2.10.3 Preparation of sequencing plate(s)

2.10.4 Preparation for Pyrosequencing

2.10.5 Preparation of dispensing tips

2.11 RNA extraction

2.11.1 Cell line seeding for RNA extraction

2.11.2 Primary cell seeding for RNA extraction

2.11.3 RNA Extraction

2.11.4 Analysis of RNA quantity by NanoDrop 1000 Spectrophotometer (Thermo Scientific)

2.12 qPCR

2.12.1 RNA extraction

2.12.2 cDNA synthesis (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems, 4369914)

2.12.2 qPCR – primer validation

2.12.2.1 Preparation of primers

2.12.2.2 Preparation of SYBR green/primer master mix

2.12.2.3 Preparation of cDNA Standard Curve

2.12.2.4 Preparation of PCR plate

2.12.2.5 qPCR run

2.12.3 qPCR – target genes

2.12.3.1 Preparation of SYBR green/primer master mix

2.12.3.2 Preparation of PCR plate

2.12.3.3 qPCR run

2.13 RNA Sequencing

2.13.1 Cell line seeding for RNA extraction

2.13.2 RNA extraction

2.13.3 Assessing RNA quality

2.13.3.1 Qubit analysis of RNA concentration

2.13.3.2 Bioanalyzer analysis of RNA quality

2.13.3.3 RNA sequencing RNA quality conditions

2.13.4 Preparation of RNA Sequencing Library

2.13.4.1 rRNA depletion (E6310)

2.13.4.2 RNA-seq library preparation (NEBNext® Ultra™ II Directional RNA library prep kit for Illumina®, E7760)

2.13.5 RNA Sequencing

2.14 siRNA for Acss2 - MISSION® esiRNA Transfection Protocol

2.15 H3K27ac tracking

2.16 Liquid Chromatography Mass Spectrometry measurement of S-adenosyl-methionine

2.16.1 Cell line seeding for metabolite extraction

2.16.2 Preparation of reagents for methanol quenching extraction

2.16.3 Methanol quenching extraction

2.16.4 Liquid chromatography mass spectrometry

2.17 Statistical Analysis

2.17.1 Statistical analysis of cell viability assays

2.17.2 Statistical analysis of western blot

2.17.3 Statistical analysis of RNA-Sequencing data

2.17.4 Statistical analysis of qPCR gene expression data

2.17.5 Statistical analysis of Illumina MethylationEPIC Array

2.17.6 Statistical analysis of pyrosequencing data

2.17.7 Statistical analysis of liquid chromatography mass spectrometry data

2.17.8 Graphing tools

Chapter 3: Metformin alters gene expression in normal breast epithelial cells

3.1 Introduction

3.2 Results

3.2.1 Non-cancerous breast epithelial cells are sensitive to metformin

3.2.2 Metformin alters gene expression in non-cancerous breast epithelial cells

3.2.3 Validation of metformin's effects on gene expression

3.2.4 Differential gene expression changes *in vitro* and *in vivo* in response to metformin

3.2.5 Metformin induces differentially expressed exons

3.2.6 Metformin potentially modifies known breast cancer risk genes

3.3 Discussion

Chapter 4: Metformin modifies one-carbon metabolism and subtly alters DNA methylation in normal breast epithelial cell lines

4.1 Introduction

4.2 Results

4.2.1 Metformin alters gene expression of key methionine pathway enzymes

4.2.2 Metformin alters SAM/SAH axis

4.2.3 Metformin induces subtle changes to DNA methylation

4.2.4 DNA methylation changes are associated with gene expression changes in MCF10A non-cancerous breast epithelial cell line

4.3 Discussion

Chapter 5: Metformin alters histone acetylation in normal breast epithelial cells

5.1 Introduction

5.2 Results

5.2.1 Metformin reduces total H3K27ac in normal breast epithelial cells

5.2.2 H3K27ac levels are restored 24 hours following metformin treatment

5.2.3 Acetate supplementation can rescue H3K27ac levels in metformin-treated normal breast epithelial cells

5.2.4 esiRNA knockdown of ACSS2 replicates metformin's reduction of H3K27ac in the presence of acetate

5.2.5 Alterations in genes associated with acetyl co-A generation pathways are altered by metformin

5.3 Discussion

Chapter 6: Metabolite availability alters metformin efficacy

6.1 Introduction

6.2 Results

6.2.1 Glucose restriction enhances metformin's efficacy

6.2.2 Glucose restriction increases activation of p-AMPK in non-cancerous breast epithelial cell lines

6.2.3 GLUT1 gene expression is modified by metformin in normal breast epithelial cells

6.2.4 Glucose restriction on DNA methylation levels in glucose linked genes

6.2.5 Glucose restriction leads to decreased H3K27ac in non-cancerous cell lines

6.2.6 L-Glutamine does not alter non-cancerous breast epithelial cell lines' response to metformin

6.2.7 Metformin alters key enzymes in the glutamine metabolism

6.2.8 Insulin does not alter non-cancerous breast epithelial cell lines' response to metformin

6.2.9 Metformin alters key enzymes in the insulin pathway

6.3 Discussion

Chapter 7: Discussion and Conclusions

Bibliography

Appendix

ABBREVIATIONS

T2DM	Type II Diabetes Mellitus
TNBC	Triple Negative Breast Cancer
BMI	Body mass index
T1DM	Type I Diabetes Mellitus
IRS	Insulin signalling substrates
PTMs	Post-translational modifications
OCT	Organic cation transporters
AMP	Adenosine monophosphate
CSC	Cancer stem cells
TAM	Tumour associated macrophages
FASN	Fatty acid synthase
TPM	Transcripts per million
RPKM	Reads per kilo base million
TSD	Tumour sample data
DE	Differential expression
POF	Premature ovarian failure
SAM	S-adenosylmethionine
DMNT	DNA methyltransferase
HATs	Histone acetylation transferases
HDACs	Histone deacetylases
TSS	Transcription start site
BCAAs	Branched chain amino acids
MSUD	Maple Syrup Urine Disease
KIC	Ketoisocaproic acid
GPDH	Glycerophosphate dehydrogenase
HILIC-MS	Hydrophilic Interaction Liquid Chromatography Mass Spectrometry

CHAPTER 1

1. Introduction

The objective of this investigation was to determine pathways metformin, an anti-type II diabetes mellitus drug, uses to interact with non-cancerous breast epithelial cells. The examination of these pathways could potentially contribute to our understanding of the effects of metformin in normal breast epithelial cells.

1.1 Type II Diabetes Mellitus and Metformin

Metformin is the most commonly prescribed T2DM drug and has been in use since the 1950s. Diabetes accounts for a large proportion of health burden in the UK. It is estimated that over 3 million people in the UK suffer with diabetes and this number increases each year, with one in sixteen diabetic patients going undiagnosed (Diabetes UK, 2015, 2016). T2DM is known as the insulin resistant disease. This is due to cells becoming less sensitive to insulin signalling over time and is therefore acquired during the patient's lifetime. T2DM accounts for 90% of diabetic patients and is responsible for a larger proportion of health burden compared to T1DM (Kahn, Cooper and Del Prato, 2014; Diabetes UK, 2015). Most patients are asymptomatic when diagnosed with T2DM. The common diagnostic tests for T2DM show fasting plasma glucose $\geq 7\text{mmol/l}$, leading to hyperglycaemia (World Health Organization, 2019). The ability for cells to respond to insulin signalling is also impaired, leading to hyperinsulinemia (American Diabetes, 2010). There are many treatments for T2DM. Initial treatment focuses on lifestyle modifications, such as changing diet and increasing exercise. If patients do not respond to these changes the next stage of treatment is drug intervention. The most commonly prescribed drug for T2DM is the biguanide metformin. Other drugs prescribed include sulfonylureas, α -glucosidase inhibitors and Thiazolidinediones (shown in table 1.1).

Class	Example Drug Name	Mechanism of Action	Side Effects
Biguanide	Metformin	↑ Gluconeogenesis	Gastrointestinal effects Lactic acidosis
Sulfonylureas	Tolazamide Glibenclamide	↑ Insulin secretion	Hypoglycaemia Cardiovascular Events Weight Gain
α-glucosidase inhibitors	Acarbose, Voglibose	↓ Carbohydrate absorption in intestine	Gastrointestinal effects
Thiazolidinediones	Pioglitazone, Rosiglitazone, Troglitazone	Insulin Sensitizer	Hypoglycaemia Cardiovascular events
GLP-1 agonists	Exenatide Liraglutide	↑ Insulin secretion	Gastrointestinal effects Hypoglycaemia
DPP-4 inhibitors	Sitagliptin Saxagliptin	↓ Blood glucose levels	Gastrointestinal effects Reduced blood pressure Cardiovascular events Hypoglycaemia

Table 1.1: Anti-Type II Diabetes Mellitus Drugs (Hauner, 2002; Proks *et al.*, 2002; van de Laar, 2008; Dicker, 2011; Garber, 2011; Pernicova and Korbonits, 2014)

Metformin is a biguanide, consisting of a basic $C_2H_7N_5$ formula with two additional methyl groups on one of the guanide side chains. Metformin is taken up into cells by organic cation transporters (OCT), OCT 1 – 3 and is not metabolised but excreted through the urine. The half-life of this drug is 1.5 – 6.5 hours in adults however food delays absorption. The standard drug dosage is approximately 1000mg – 1500mg once per day. Metformin is a relatively non-cytotoxic drug however side effects can include nausea and diarrhoea. Adverse effects include lactic acidosis, which occurs in 0.03 cases per 1000 patient years, and approximately 30% patients have mild gastrointestinal disturbance. In T2DM patients Metformin is used as an “insulin-sensitizing” drug. The main function of Metformin is to inhibit gluconeogenesis in the liver through inhibition of the mitochondrial complex I. This reduces ATP levels and induces metabolic stress (Segal *et al.*, 2011; Pernicova and Korbonits, 2014; Wang and Weinshilboum, 2014).

1.2 Mechanisms of metformin in Type II Diabetes Mellitus patients

The mechanisms of metformin action have been extensively studied in diabetic patients. It is well established in diabetic liver metformin induces metabolic stress. As previously stated, Metformin is taken into cells by OCT 1-3. OCT1, coded by the SLC22A1 gene, is exclusively expressed in the liver and is responsible for hepatic uptake of Metformin. OCT2 is expressed in the kidney while OCT3 is expressed in a wide range of tissue such as skeletal muscle, brain and heart. It has been shown that OCT1 is highly expressed in cancer allowing Metformin to be readily taken up into cells (Hayer-Zillgen, Brüss and Bönisch, 2002). Metformin causes metabolic stress in a number of ways. The primary mechanism of action is inhibition of mitochondrial complex I, shown in figure 1.1. Mitochondria provide energy to cells through a process called oxidative phosphorylation. A major hallmark of cancer is dysregulated energy function (Hanahan and Weinberg, 2011). Tumour cells use high levels of energy in the form of ATP to support many hallmarks of cancer such as unlimited growth, sustaining proliferative signalling and angiogenesis (Vander Heiden, Cantley and Thompson, 2009). Mitochondria

also play roles in different cellular pathways, for example mitochondria inhibit intrinsic apoptosis through cytochrome c inhibition (Cai, Yang and Jones, 1998). It is unknown how Metformin inhibits this complex. It is hypothesised that metformin inhibits mitochondrial glycerol-3-phosphate dehydrogenase (mGPD), reducing the effectiveness of cytosolic GPD (cGPD) (Madiraju *et al.*, 2014; Petersen, Vatner and Shulman, 2017). This leads to increased NADH levels as NADH cannot be converted to NAD⁺. This could potentially result in a number of downstream effects such as reduced pyruvate uptake into the citric acid cycle (TCA). The known result of mitochondrial inhibition by metformin is reduced ATP production and increased AMP production (Pernicova and Korbonits, 2014). Decrease in ATP levels reduces energy needed for vital cancer pathways, leading to a slowing down of cellular division, growth and proliferation.

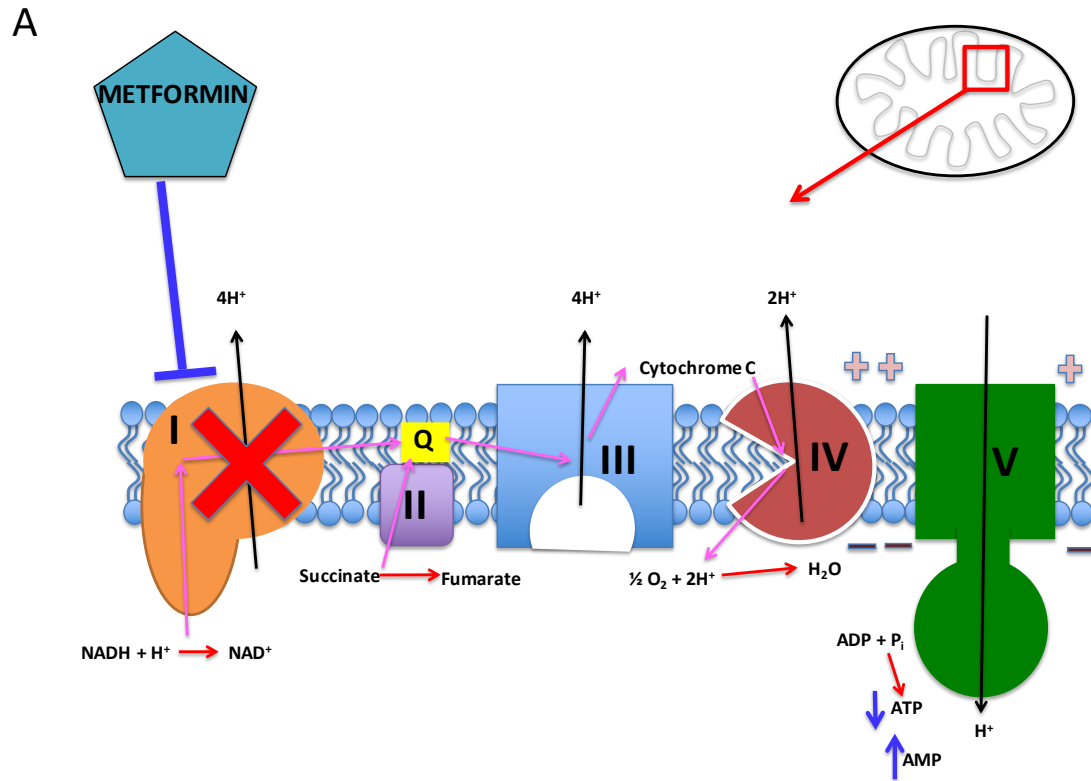


Figure 1.1: Metformin causes metabolic stress. (A) Metformin Action on Mitochondrial Respiratory Chain Complex I. [Adapted from (Jiang and Zhang, 2003; Miller *et al.*, 2013; Madiraju *et al.*, 2014; Pernicova and Korbonits, 2014)].

1.3 Mechanisms of metformin in cancer

Metformin has been investigated as an anti-cancer drug in multiple cancer types for decades. There are a number of mechanisms of action metformin is thought to be involved with to inhibit cancer cell growth and increase cancer cell death.

(A) Inducing metabolic stress

In cancer models, metformin has been shown to induce a similar effect on complex I of the electron transport chain in the mitochondria as seen in diabetic liver, increasing AMP levels while reducing ATP levels. Metformin can also inhibit transcription of gluconeogenic genes, shown in figure 1.2, by inducing phosphorylation of CRTC2. This brings CRTC2 into the cytoplasm, preventing transcription of genes such as PEP, PCK2 and G6Pase. Reducing expression of these genes leads to decreased L-alanine and lactate uptake. This is important as hypoxic tumour cells use high volumes of L-alanine and lactate as energy sources. Metformin can also inhibit glycerophosphate dehydrogenase (GPDH) activity and reduce cAMP levels, leading to further metabolic stress (Jiang and Zhang, 2003; Miller *et al.*, 2013; Madiraju *et al.*, 2014).

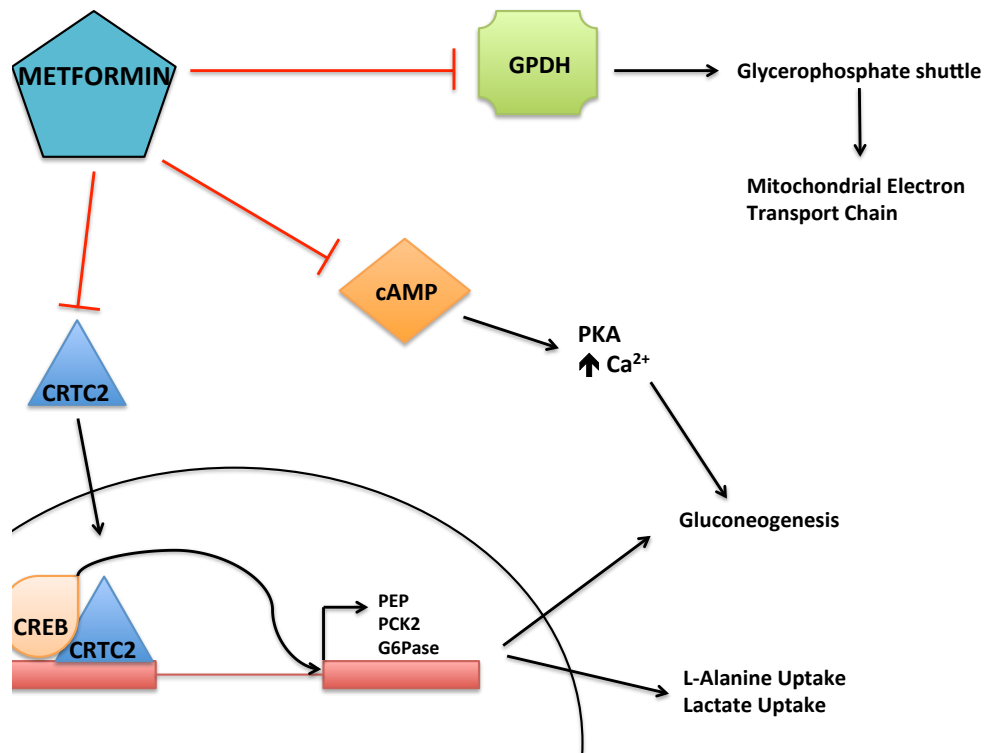


Figure 1.2: Metformin causes metabolic stress. Metformin causes phosphorylation of CRTC2, preventing translocation into the nucleus and initiation of transcription of gluconeogenic genes. Metformin further inhibits GPDH to impair the mitochondrial electron transport chain. Inhibition of cAMP prevents to activation of PKA and increased calcium, thus inhibiting gluconeogenesis [Adapted from (Jiang and Zhang, 2003; Miller *et al.*, 2013; Madiraju *et al.*, 2014; Pernicova and Korbonits, 2014)].

(B) Inhibiting Growth Pathways

Growth pathways are over-active in tumour cells. Constant activation of these pathways coupled with reduction of growth pathway inhibitors leads to dysregulated tumour growth. AMPK is a protein kinase that regulates key growth and metabolism pathways. AMPK activation leads to carbohydrate metabolism and translocation of GLUT4 glucose transporter to plasma membranes. AMPK is also involved in lipid and protein metabolism and growth regulation by inhibition of pathways such as mTOR. Activation of AMPK is through high levels of adenosine monophosphate (AMP), leading to phosphorylation of AMPK via LKB1. AMPK can then inhibit RPTOR, a component of the mTORC1 complex, or induce TSC1 and TSC2 inhibitory molecules to prevent mTOR activation. In cancer, AMPK is downregulated leading to over activation of downstream targets of AMPK, such as mTOR. AMPK showed reduced expression in breast cancer cells. This allows mTOR pathways to be highly activated contributing to breast cancer growth and inhibition of cell death pathways such as apoptosis. Metformin, through inhibition of mitochondrial complex 1, induces increased levels of AMP, shown in figure 1.3 (A). This raises the AMP:ATP ratio, activating AMPK. AMPK activation leads to mTOR inhibition. This reduces tumour cell growth and proliferation and drives tumour cell death (Zakikhani *et al.*, 2006; Kuhajda, 2008; Wu *et al.*, 2011; Korsse, Peppelenbosch and van Veelen, 2013; Faubert *et al.*, 2014). Metformin has also been shown to indirectly activate the protein REDD1, which inhibits mTOR component of the mTORC1 complex, shown in figure 1.3 (A) (Ben Sahra *et al.*, 2011). Metformin can also augment the insulin-signalling pathway, inhibiting Akt pathway activation, shown in figure 1.3 (B). Metformin causes phosphorylation at serine residues of the IRS-1 signalling molecule, a component of the insulin-signalling cascade. This inhibits IRS-1 from inducing downstream Akt activity but also prevents insulin-binding leading to Akt activation (Dong *et al.*, 2006; Karnevi *et al.*, 2013).

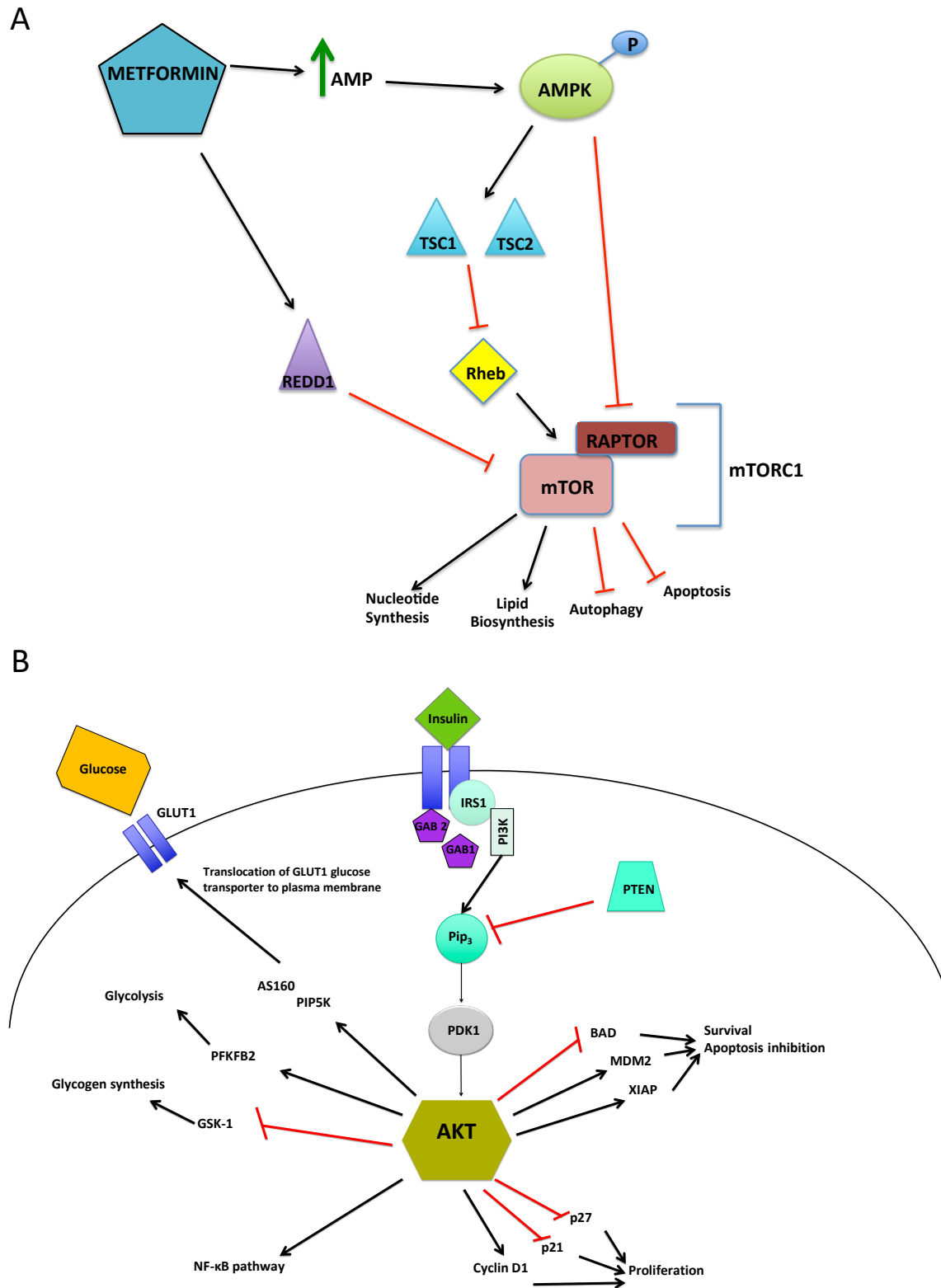


Figure 1.3: Metformin inhibits key growth pathways. (A) Metformin increases AMPK phosphorylation, which inhibits mTOR signaling. mTOR can also be inhibited by REDD1 activation by metformin. (B) Metformin causes phosphorylation of IRS1, suppressing downstream Akt pathway activation [Adapted

from (Dong *et al.*, 2006; Zakikhani *et al.*, 2006; Kuhajda, 2008; Ben Sahra *et al.*, 2011; Wu *et al.*, 2011; Karnevi *et al.*, 2013; Korsse, Peppelenbosch and van Veelen, 2013)].

(C) Inducing Apoptosis

A major hallmark of cancer is inability of cells to induce cell death pathways. One of these pathways is apoptosis. Apoptosis can be activated by internal cues or by external signalling. Apoptosis is a response to a number of stresses including DNA damage, metabolic stress and environmental stress (Taylor, Cullen and Martin, 2008). Cancer cells prevent apoptosis by downregulating and/or inhibiting key apoptosis inducers (Hanahan and Weinberg, 2011). Metformin can induce apoptosis in a number of indirect mechanisms, shown in figure 1.4. Metformin has been shown in cancer to interrupt Cyclin D1 binding. Cyclin D1 is a key regulator of cyclin-dependent kinases (CDKs) in cell cycle progression. When Cyclin D1 forms a complex with a CDK, for example CDK4, this drives cells through G1/S phase transition in the cell cycle (Bièche *et al.*, 2002; Bartek and Lukas, 2011). Inhibition of Cyclin D1 leads to cell cycle arrest and subsequent apoptosis (Ben Sahra *et al.*, 2008; Zhuang and Miskimins, 2008). Metformin has also been shown to indirectly activate the intrinsic apoptotic pathway. Metformin treatment can lead to cleavage of PARP1, which induces cytochrome c release from the mitochondria leading ultimately to cleavage of executioner caspases such as caspase 7. Alternatively, Metformin has been demonstrated to lead to cleavage of caspase 7. Both of these mechanisms lead to apoptosis and cell death (Zhuang and Miskimins, 2011; Shang *et al.*, 2016).

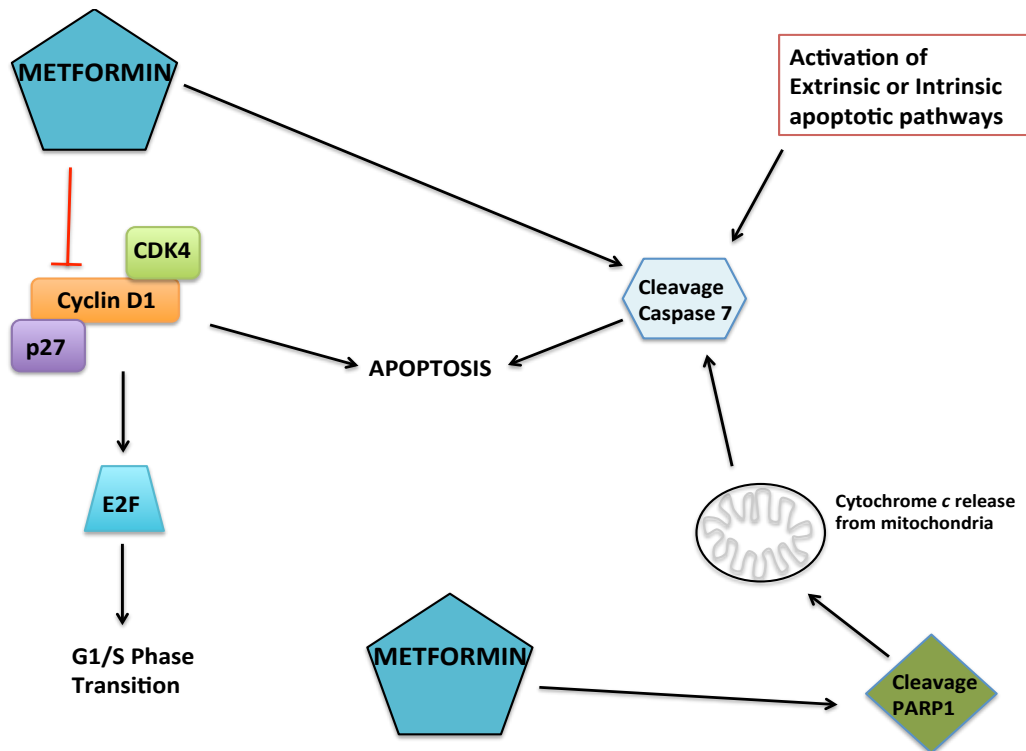


Figure 1.4: Metformin induces apoptosis by inhibiting Cyclin D1 leading to cell cycle arrest and increasing caspase activity through cleavage of PARP1 or direct cleavage of executioner caspases [Adapted from (Ben Sahra *et al.*, 2008; Taylor, Cullen and Martin, 2008; Zhuang and Miskimins, 2011; Shang *et al.*, 2016)].

(D) Inducing Stem Cell Death

The tumour microenvironment is heterogeneous with multiple different cell types making up the tumour mass. For example tumours contain fibroblasts, immune cells and endothelial cells which all aid in tumour development (Whiteside, 2008). In many tumours there is also a small subset of “cancer stem cells” (CSCs). CSCs are cancer cells that are undifferentiated with the potential to become different cell types. This allows the tumour to constantly adapt to the microenvironment. These CSCs maintain the gene mutations and cellular dysfunctions associated with the main tumour cells (Beck and Blanpain, 2013). CSCs are important not only in tumour development but also treatment. It has been shown that CSCs are more resistant to radiotherapy and chemotherapy and contribute to more aggressive tumours (Baumann, Krause and Hill, 2008; Vidal *et al.*, 2014). Metformin has been shown to not only kill CSCs but also specifically target CSCs, shown in figure 1.4. For example in breast cancer, CD44^{High}/CD24^{Low} cancer stem cells are selectively targeted by Metformin, leading to CSC death (Cufi *et al.*, 2012).

There are multiple mechanisms by which it is thought Metformin contributes to CSC death. The first mechanism is by causing metabolic stress leading to apoptosis, as previously described. Another mechanism is by inhibiting CSC maintenance. EZH2 is a histone methyltransferase, which leads to repression of gene expression. It is thought the genes, which EZH2 regulates, are associated with differentiation. Metformin has been shown to inhibit EZH2 activity and has also been shown to induce a microRNA, mir26a, which represses EZH2, subsequently inhibiting EZH2 action (Bao *et al.*, 2012; Chang and Hung, 2012; Jansen *et al.*, 2012) Metformin not only can inhibit the repression of differentiation-related genes but can also inhibit the transcription of “stem-like” genes. Metformin treatment has been shown to lead to inhibition of NFκB and STAT3 transcription factors from initiating transcription (Hirsch, Iliopoulos and Struhl, 2013; Zhu *et al.*, 2014). Inhibition of NFκB is also important in regulating the immune response. The immune system plays a vital role in the maintenance of CSCs. CSCs hijack immune cells such as M2-like macrophages called tumour associated macrophages (TAMs). TAMs dampen the immune system by releasing cytokines, which suppress the adaptive

immune response. NFκB is an important activator of TAMs (Hagemann *et al.*, 2008). NFκB is also a key regulator of metabolism and inhibition of NFκB/RelA has been shown to induce cell death. While it is unknown if Metformin induces this cell death pathway through inhibition of NFκB, it would be an interesting avenue to investigate (Mauro *et al.*, 2011). In breast cancer, Metformin has been shown to inhibit Oct4 expression. Oct4 is an important transcription factor that maintains stem cell pluripotency. It is a key gene in maintaining CSCs (Jung *et al.*, 2011; Kim and Nam, 2011).

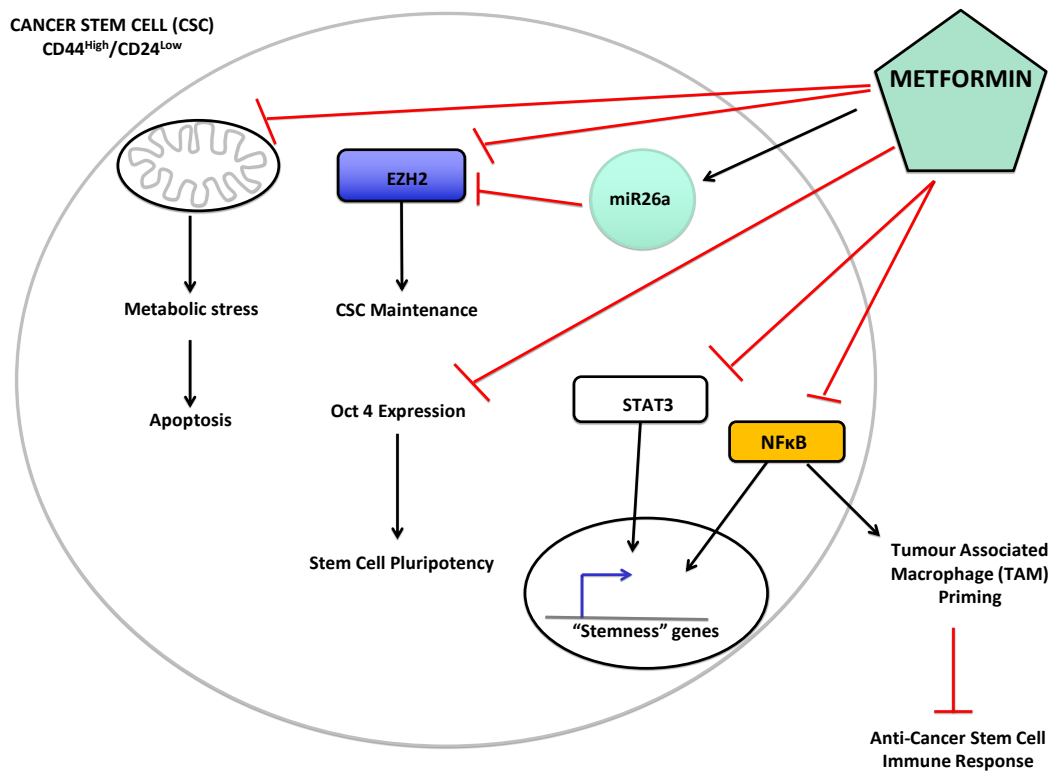


Figure 1.4: Metformin induces cancer stem cell death. Metformin treatment leads to inhibition of important regulators of cancer stem cells such as EZH2 and Oct4. Metformin can also inhibit pro-stem cell genes via inhibition of NFκB and STAT3. Inhibition of NFκB can also inhibit the pro-stem cell immune response [Adapted from (Isoda *et al.*, 2006; Hagemann *et al.*, 2008; Jung *et al.*, 2011; Bao *et al.*, 2012; Hirsch, Iliopoulos and Struhl, 2013; Pernicova and Korbonits, 2014)].

1.4 Epidemiological studies linking metformin or Type II Diabetes Mellitus with breast cancer

Metformin has been shown to potentially reduce breast cancer risk in diabetic patients (Bodmer *et al.*, 2010; Bosco *et al.*, 2010). While a number of epidemiological studies have shown that breast cancer risk is reduced by Metformin therapy, a number of analyses show little to no effect on breast cancer risk reduction by Metformin, shown in table 1.2. Bodmer *et al.* performed a nested case-control study analysing Metformin users vs. non-users and incident cases of breast cancer and found reduced risk, most predominantly reduction was seen in long-term users of Metformin, suggesting there Metformin has a stable, longitudinal effect [OR = 0.44 (95% CI, 0.24 – 0.82)] (Bodmer *et al.*, 2010). Bosco *et al.* also saw a modest reduction in breast cancer risk following a nested case-control analysis [OR = 0.82 (95% CI, 0.64–1.08)] (Bosco *et al.*, 2010). It must be noted this includes former users, recent users and users >5years. When long-term metformin use was only considered the results were similar [OR = 0.82 (95% CI, 0.56–1.21)]. In contrast, a meta-analysis of a number of randomized control trials and observational studies showed no reduction in breast cancer risk [OR = 0.97 (95% CI, 0.88 – 1.08)] (Franciosi *et al.*, 2013). No association between Metformin and breast cancer risk has also been seen in retrospective cohort studies when comparing to other T2DM drugs [HR = 1.04 (95% CI, 0.83, 1.31)] (Qiu *et al.*, 2013). A major consideration that may affect these studies is the duration of Metformin use. As previously stated, long term Metformin use has been shown to reduce risk in contrast to short-term use. Many studies use total person-years, which include short-term users, possibly biasing the results.

T2DM is a multifactorial disease that is associated with many risk factors including age, BMI, genetic factors, smoking, altered circadian rhythms, sedentary lifestyle and diet (Helmrich *et al.*, 1991; Hu *et al.*, 2001; Koppes *et al.*, 2005; Sladek *et al.*, 2007; Willi *et al.*, 2007). Many of these risk factors overlap with common risk factors for many types of cancer. In fact, T2DM has been associated with a number of cancers, including colorectal, liver and breast cancer (Adami *et al.*, 1996; Yang, Hennessy and Lewis,

2004; Vona-Davis, Howard-McNatt and Rose, 2007). Larsson *et al.* analysed 5 case-control studies and 15 cohort studies and established a 20% increased risk of incident cancer when diabetics were compared to non-diabetics. This study also noted the strongest connection was seen in postmenopausal women. Studies from Italy and Japan show similar results, with a 24% and 51% increased risk of breast cancer incidence respectively (Jee *et al.*, 2005; Larsson, Mantzoros and Wolk, 2007; Gini *et al.*, 2016). Epidemiological studies demonstrate that there is a link between T2DM and breast cancer, but the ultimate question focuses on the mechanisms by which T2DM can increase breast cancer risk. T2DM has a number of biological effects, which can promote tumourgenesis. The two major characteristics of T2DM - hyperinsulinemia and hyperglycaemia, are thought to contribute to cancer, discussed further in the review (Rapp *et al.*, 2006; Gunter *et al.*, 2009). However, T2DM is also associated with other pro-cancer effects for example chronic inflammation. In the adipose tissue of obese and diabetic patients there is a major increase in macrophage populations. Pro-inflammatory macrophages, M1 subtype, are activators of the adaptive immune system such as T cells. While increasing the pro-inflammatory immune system, signalling cytokines are released and dampen the anti-inflammatory immune system such as M2 macrophages and regulatory T cells. This promotes chronic inflammation leading to increased release of inflammatory cytokines such as IL-6 and TNF α . These cytokines can induce NF κ B transcription factor signalling and MAPK signalling, eventually leading to decreased insulin signalling substrates (IRS) function. IRS proteins are important components of the insulin receptor complex. IRS proteins transmit ligand binding signalling to intracellular pathways. Inhibition of IRS proteins by phosphorylation effectively renders the receptor inactive, preventing insulin activating intracellular signalling when bound to its receptor (Liang, Jiang and Frank, 2000; Dhawan and Richmond, 2002; Pickup, 2004; Gupta *et al.*, 2005). T2DM can also lead to endoplasmic reticulum stress and mitochondrial dysfunction, contributing to insulin resistance (Lowell and Shulman, 2005; Back and Kaufman, 2012).

	Study Type	Result	Interpretation
T2DM	Retrospective population-based cohort study	SIR = 1.24 (95% CI, 1.00–1.52) [Observed Vs. Expected]	24% greater risk of developing breast cancer in diabetic patients compared to non-diabetic patients
T2DM	Meta-analysis of case–control and cohort studies	RR = 1.20 (95% CI, 1.12–1.28) [Q4 Vs. Q1]	20% greater risk of developing breast cancer in diabetic patients compared to non-diabetic patients
T2DM	10-year prospective cohort study	HR = 1.51 (95% CI, 1.26–1.80) [Diabetic Vs. Non Diabetic]	51% greater risk of developing breast cancer in diabetic patients compared to non-diabetic patients
Metformin	Nested Case-Control Analysis	OR = 0.44 (95% CI, 0.24 – 0.82)	2-fold reduction in the risk of developing Breast Cancer.
Metformin	Nested Case-Control Analysis	OR = 0.82 (95% CI, 0.64–1.08)	22% reduction in Breast Cancer incidence in T2DM patients
Metformin	Meta Analysis of randomized control trails (n = 12) and observational studies (n = 41)	OR = 0.97 (95% CI, 0.88 – 1.08)	No evident reduction in Breast Cancer Risk
Metformin	Retrospective cohort study	HR = 1.04 (95% CI, 0.83, 1.31)	No evident reduction in Breast Cancer Risk
Metformin	Retrospective cohort study Cox Regression Analysis	HR = 0.88 (95% CI 0.56, 1.39)	No evident reduction in Breast Cancer Risk
Metformin	Retrospective cohort study	HR = 1.00 (0.98 to 1.02) [3 months – 6 years treatment time]	No evident reduction in Breast Cancer Risk
Metformin	Retrospective cohort study in elderly women 65 – 85 years	HR = 1.2 (95% CI 0.94, 1.6)	No evident reduction in Breast Cancer Risk
Metformin	Systematic review and meta-analysis	OR = 0.93 (95% CI 0.85, 1.03)	No evident reduction in Breast Cancer Risk
Metformin	Retrospective population based cohort study	IRR = 0.69 (95% CI 0.32 , 1.48) [Metformin only vs diet-only]	No evident reduction in Breast Cancer Risk
Metformin	Retrospective population study	HR = 0.94 (95% CI 0.62, 1.43)	No evident reduction in Breast Cancer Risk
Metformin	Nested case-control analysis	HR = 1.02 (95% CI 0.93–1.11)	No evident reduction in Breast Cancer Risk

Table 1.2: Summary of epidemiological studies linking metformin or Type II Diabetes Mellitus with incident cases of Breast Cancer. Studies linking hyperglycaemia and hyperinsulinemia are included in appendix table 1.1 (Jee *et al.*, 2005; Larsson, Mantzoros and Wolk, 2007; Bodmer *et al.*, 2010; Bosco *et al.*, 2010; Qiu *et al.*, 2013; Franciosi *et al.*, 2013; Gini *et al.*, 2016; Bowker *et al.*, 2017; Hong *et al.*,

2017; Tang *et al.*, 2018; Vicentini *et al.*, 2018; Dankner *et al.*, 2019; Farmer *et al.*, 2019; Hosio *et al.*, 2019).

1.5 Breast cancer and current prevention practices

Breast cancer is the most commonly diagnosed cancer in the UK. Approximately 50,000 cases are diagnosed each year (CRUK, 2015a). Breast cancer accounts for 7% of total cancer related deaths with a 5-year survival rate of 86% (CRUK, 2015a, 2015b). While the survival rates of breast cancer are improving, there is a difference in survival between subtypes. The majority of breast cancers are classed as carcinomas, which originate from epithelial cells. 1% of breast tumours originate from stromal cells and are classed as sarcomas. The most common form of breast carcinoma; oestrogen receptor positive breast cancer (luminal A and luminal B) has a relatively good prognosis whereas prognosis for the most aggressive form of breast cancer, triple negative basal-like breast cancer (TNBC) is very poor (Onitilo *et al.*, 2009). Approximately 10% of breast cancer risk can be attributed to inherited factors such as common and rare genetic variants while the remaining 90% is likely due to lifestyle factors and their interaction with genetic factors. There are many risk factors that are related to breast cancer susceptibility including age, age at menarche, age at first live birth, alcohol consumption, smoking, family history (e.g. *BRCA1/2* mutations), body mass index (BMI), hormone replacement therapy, obesity and type II diabetes mellitus (McPherson, Steel and Dixon, 2000).

There are a number of breast cancer prevention practices currently used. Breast checks by mammogram has been available for post-menopausal women for decades. Mammography is an imaging technique based on x-ray which can identify calcifications and tumours (National Health Service (NHS), 2019). While the success rate of mammography has meant more breast cancer cases are diagnosed each year, it may not be the most effective imaging technique for prevention (Bleyer and Welch, 2012). MRI has been recommended as a better technique for breast cancer identification

in post-menopausal women (Kuhl *et al.*, 2005). The age at which mammograms are available excludes a cohort of breast cancer patients who have either not reached the minimum age for mammogram checks or are pre-menopausal. There are preventative practices available for patients who have a familial history of breast cancer, accounting for 10% of all breast cancer cases. The most common driver of familial breast cancer is *BRCA1/2* mutations. These patients can be followed by early monitoring combined with mammograms (Slawski, Daumer and Boulesteix, 2008). Prophylactic mastectomy has been shown to reduce breast cancer risk for these patients. However, this practice requires major surgery, which has associated risks, as well as an impact on the quality of life for the patient(s) (Scheuer *et al.*, 2002; Geiger *et al.*, 2007).

Prevention practices are not only implemented by national health services. Breast cancer awareness has created a number of prevention campaigns run by charities that focuses on self-checking and self-care. For example, the charity Coppafeel!, founded by a pre-menopausal breast cancer patient, actively encourages self-checking and implements a social media-based campaign to explain and demonstrate proper breast checking practices (Coppafeel!, 2009). Cancer Research UK (CRUK) run a campaign of understanding and modifying breast cancer risk factors. For example, CRUK runs campaigns on the effects of alcohol, exercise and healthy diet on breast cancer risk. These campaigns provide information on how to reduce breast cancer risk while placing the onus to change lifestyle factors on the patient(s) (Cancer Research UK, 2019). Drug interventions can also be used as preventative techniques. For example, high-risk patients can be given Tamoxifen prophylactically in order to reduce breast cancer risk (Powles *et al.*, 2007). Aromatase inhibitors, such as anastrozole have also been studied and have been shown to reduce breast cancer more effectively than tamoxifen (Cuzick *et al.*, 2014).

1.6 Clinical trials studying metformin effects in breast cancer risk

Few clinical trials are focusing on metformin as a cancer preventative measure in breast cancer. Two trials, which commenced in 2013 and 2014, have been recently completed and results are not available (Anna Maria Storniolo, 2013; University of Arizona, 2014). These studies (details of which are shown in table 1.3) aim to treat pre-menopausal women with metformin hydrochloride or placebo to determine the physical and biological changes to breast tissue. The third trial included in table 1.3 is a longitudinal study started in 2003 to assess the development of diabetes in a large population when treated with metformin ((NIDDK), 2002). This study focuses on metformin as a preventative agent for T2DM however cancer incidence is included as a secondary outcome.

Metformin's effects on breast tumours have been investigated *in vivo*. For example, a neoadjuvant "window" of opportunity study (NCT00897884), treated newly diagnosed breast cancer patients with metformin for 1-3 weeks before treatment with conventional therapies. While there was the expected drop in weight and BMI post treatment, there was also a significant decrease in glucose levels ($p < 0.045$) and increase in cell death ($p < 0.004$). Further analyses of the samples led the investigators to believe the effect of Metformin was due to altered insulin signalling, leading to reduced PI3K activation (Niraula *et al.*, 2012; Dowling *et al.*, 2015).

These studies pre- and post- breast cancer give an insight into the molecular effects of metformin in breast tissue.

ClinicalTrials.gov Identifier	Study Type	Inclusion Criteria	Trial Design	Primary Outcomes	Secondary Outcomes	Study Start Date	Estimated Study End Date
NCT00038727	Randomized Interventional Clinical Trial	Females and Males 25 years +	2779 participants Three Arms: 1 – Original Lifestyle 2 – Metformin treatment (850mg twice per day) 3 – Placebo treatment	To determine if participants develop Diabetes To assess the prevalence of aggregate microvascular complication To identify total cancer incidence (excludes non-melanoma skin cancer) To assess the prevalence of major adverse cardiovascular events (MACE)	To assess prevalence of subclinical atherosclerosis To assess cognitive function To assess short physical performance battery To assess frailty in participants	2002	2025
NCT01793948	Randomized double-blind interventional trial	Pre-menopausal women BMI >25 Waist circumference >35 inches At least one additional indicator of metabolic syndrome	151 Patients Two Arms 1 – Metformin 2 - placebo Metformin: First 4 weeks one 850 mg tablet taken by mouth. Following this one 850 mg tablet taken twice daily.	To determine changes in breast density	To assess changes in serum insulin levels To assess changes in serum IGF-1 to IGFBP-3 ratio To assess changes in serum testosterone levels To assess changes in serum leptin to adiponectin ratio To assess changes in body weight To assess changes in waist circumference To assess changes in serum IGF-2 levels	2013	2018
NCT02028221	Phase II Randomized double-blind interventional trial	Pre-menopausal women with elevated risk of breast cancer BMI >25	24 Patients Two experimental arms: 1 – Metformin Hydrochloride 2 – Placebo First 30 days: metformin/placebo by mouth once daily Second 30 days: metformin/placebo by mouth twice daily 12 courses of treatment	To determine changes in signalling pathways on breast tissue Using reverse phase proteomics	To determine changes in ordinal level of breast density To determine the effect of metformin on the insulin axis in serum To assess the proportion of patients experiencing adverse events	2014	2020

Table 1.3: Clinical trials studying metformin effects in breast cancer risk ((NIDDK), 2002; Anna Maria Storniolo, 2013; University of Arizona, 2014)

1.7 Epigenetic modifications associated with metformin and Type II Diabetes Mellitus

Metformin's known mechanisms result in metabolic dysregulation. Metabolism is closely linked to epigenetics as a number of metabolites, such as acetyl co-A and S-adenosyl methionine, are required to modify the epigenome.

There are a number of epigenetic modifications including DNA methylation, histone post-translational modifications (PTMs) and microRNAs. DNA methylation allows for gene regulation by adding methyl groups to cytosines 5' of guanine and is regulated by DNA methyltransferase enzymes (DNMT1 and DNMT3a/3b). This addition adjusts how DNA is accessed either causing inactivation of genes in promoters or differential splicing of genes through gene body methylation (Robertson, 2005). Histone PTMs have been linked to modification of chromatin conformation, for example acetylation of histone marks can change chromatin into an open state, allowing gene expression whereas lack of acetylation closes chromatin, leading to gene repression. Histone PTMs are also important in interaction with DNA methylation to regulate gene transcription (Jenuwein and Allis, 2001). microRNAs (miRNAs) are small non-coding RNAs which regulate gene transcription by regulating RNA levels. miRNAs can degrade RNA, leading to reduced protein formation and thus inhibiting the final outcome of gene transcription (Ha and Kim, 2014).

Metformin has been associated with inducing epigenetic modifications, shown in table 1.4. Let-7 is a key tumour-suppressor miRNA important in reducing expression of invasion genes such as c-Myc. In fact, the interaction of H19 and Let-7 is key in regulating this activity. Increased H19 expression is associated with decreased Let-7 and poorer prognosis. Metformin has been shown to alter DNA methylation of H19 gene promoters, reducing H19 expression in combination with increasing Let-7 levels. This leads to more differentiated cancer cells. Cancer stem cells are more difficult to eliminate with conventional therapies such as radiotherapy. More differentiated phenotypes could aid in improving treatment effectiveness (Yan *et al.*, 2015).

Histone PTMs are important regulators of gene expression. Histone H2B mono-ubiquitination is an important histone PTM in regulating expression of cell cycle transition proteins such as p21 and cyclin D1. Du *et al.* demonstrated Metformin treatment in breast cancer cell lines could induce histone H2B mono-ubiquitination. Metformin inhibition of cyclin D1 can drive cancer cell cycle arrest and apoptosis (Du *et al.*, 2014).

As previously mentioned, Metformin can inhibit EZH2 activity. EZH2 is important for maintaining cancer stem cell phenotypes. Inhibition of EZH2 is through a number of miRNAs including Let-7 and miR-26-a (Bao *et al.*, 2012). Fatty acid synthesis is a key biological pathway, which is highly upregulated in tumours. Fatty acid synthase (FASN) is an important enzyme in this process. Metformin treatment has been shown to induce increased levels of miR-193-b. miR-193-b has been shown to regulate FASN. Reduction or inhibition of FASN in tumours has been associated specifically with cancer cell death and provides a key pathway for Metformin-induced apoptosis (Wahdan-Alaswad *et al.*, 2014). Many of the miRNAs altered by Metformin treatment are involved in maintaining or blocking “stem cell” phenotypes. miR-181-b was one of these “stem cell” associated miRNAs reduced by Metformin treatment, which has been shown to promote EMT and invasion in breast cancer cells (Oliveras-Ferraro *et al.*, 2011).

Further investigation into the epigenetic altering effects of Metformin is needed to fully elucidate Metformin’s anti-cancer potential.

Epigenetic Modification	Gene	
	MOA	Consequence
DNA Methylation	H19 Hypermethylation	Increased Let-7
Histone PTMs	Histone H2B Mono-Ubiquitination	Inhibition of Cyclin D1 transcription
miRNA	↑ miR-26-a	Inhibition of EZH2
	↑ miR-33-a	Regulation of fatty acid metabolism
	↑ miR-193-b	Reduction in FASN
	↑ Let-7-a	Promote differentiation
	↓ miR-181-b	Inhibit “stem cell” phenotype
	↓ miR-183	TGFβ-1 induced apoptosis

Table 1.4: Epigenetic modifications associated with Metformin (Oliveras-Ferraro *et al.*, 2011; Bao *et al.*, 2012; Blandino *et al.*, 2012; Du *et al.*, 2014; Wahdan-Alaswad *et al.*, 2014; Yan *et al.*, 2015).

T2DM has been linked to altered epigenetic patterns, shown in table 1.5. A major risk factor of T2DM is aging. It has been demonstrated that DNA methylation plays a role in repressing the *COX7A1* gene, which is an important part of the mitochondrial respiratory chain complex 4. *COX7A1* expression has been associated with glucose uptake and loss of this gene could contribute to the hyperglycaemic state of diabetic patients (Rönn *et al.*, 2008). It had also been demonstrated that there was a link between hypermethylation of glucokinase (Gck) and L-type pyruvate kinase (LPK) promoters, which contributes to obesity and insulin resistance in the liver in high-fat fed mice compared to normal

controls (Jiang *et al.*, 2011). In the pancreas of humans, the gene *PPARGC1A* is hypermethylated leading to down regulation of PGC-1 α expression. PGC-1 α is a key regulator of genes involved in cellular metabolism such as oxidative phosphorylation. Loss of this gene is seen in the skeletal muscle of diabetics compared to non-diabetics (Ling *et al.*, 2008). There is on-going research into the changes in DNA methylation associated with T2DM seen in different countries/regions, suggesting a key environmental factor in T2DM development (Al Muftah *et al.*, 2016). Conversely, hypomethylation has also been seen to increase T2DM risk as hypomethylation of the *FTO* gene is strongly related to T2DM and obesity (Dayeh *et al.*, 2014).

High levels of glucose have shown to induce H3K4 monomethylation and H3K9 monomethylation loss causing over activation of NF κ B-p65, a powerful pro-cancer transcription factor and enhancing cardiac complications associated with T2DM. In a rat model of intrauterine growth restriction (IUGR rats) chromatin remodelling was associated with decreased *Glut4* expression in muscle cells, a key glucose transporter. This could contribute to hyperglycaemic conditions seen in T2DM. This association has also been seen in exercise studies, demonstrating changes in *Glut4* histone modifications when at rest compared to during exercise (Raychaudhuri *et al.*, 2008; Brasacchio *et al.*, 2009).

A key component of insulin signalling is IRS-1. IRS-1 is a cytoplasmic signalling molecule, which when phosphorylated leads to downstream activation of the Akt pathway. Upregulation of miR-144 has been correlated with downregulation of IRS-1 activation and has been shown to inhibit IRS-1 phosphorylation, preventing insulin signalling from activating key growth pathways. This can contribute to insulin resistance seen in T2DM patients (Karolina *et al.*, 2011). miR-126 is an important regulator of angiogenesis in endothelial cells by increasing VEGF and FGF activity. In a recent study, reduced miR-126 levels in high glucose conditions were seen in apoptotic bodies of endothelial cells and has been associated with diabetic patients. Loss of miR-126 could contribute to adverse cardiac events associated with T2DM. NF κ B is an important transcription factor and cytokine for promoting

inflammation (Zampetaki *et al.*, 2010). miR-146a is a potent suppressor of NFκB, leading to reduced expression of target genes of NFκB. Loss of miR-146a reduces control on NFκB, contributing to the chronic inflammation state of diabetics (Balasubramanyam *et al.*, 2011).

Epigenetic Modification	Gene	
	MOA	Consequence
Methylation	COX7A1 Hypermethylation	Hyperglycaemia
	Gck Hypermethylation	Insulin resistance
	LPK Hypermethylation	Insulin resistance
	PPARGC1A Hypermethylation	Downregulation of PGC-1 α
	FTO Hypomethylation	Increased T2DM risk
	INS Hypermethylation	Hyperglycaemia
	PDX1 Hypermethylation	Hyperglycaemia
Histone PTMs	H3K4 Mono-Methylation	Hyperglycaemia, cardiac complications
	Glut4 H3K9 Di-Methylation	Hyperglycaemia
miRNAs	↓ miR-146a	Increased NF κ B activity
	↑ miR-144	Inhibition of IRS-1
	↓ miR-126	Reduced VEGF activation and endothelial cell protection
	↑ miR-124-a	Association with breast cancer
	↑ miR-30-d	Association with breast cancer

Table 1.5: Epigenetic modifications associated with Type II Diabetes, Hyperinsulinemia and Hyperglycaemia (Ling *et al.*, 2008; Raychaudhuri *et al.*, 2008; Rönn *et al.*, 2008; Brasacchio *et al.*, 2009; Zampetaki *et al.*, 2010; Balasubramanyam *et al.*, 2011; Jiang *et al.*, 2011; Karolina *et al.*, 2011; Yang *et al.*, 2011, 2012; Dayeh *et al.*, 2014; Zhang *et al.*, 2016)

This study hypothesised that metformin could potentially alter epigenetic patterns, potentially through metabolic dysregulation. This could potentially alter downstream gene expression and influence breast cancer susceptibility.

- **Aim 1:** Identify whether metformin caused gene expression changes in non-cancerous breast epithelial cells.
- **Aim 2:** Identify DNA methylation changes associated with metformin response in non-cancerous breast epithelial cells.
- **Aim 3:** Identify Histone post-translation modification (PTM) changes associated with metformin response in non-cancerous breast epithelial cells.
- **Aim 4:** Identify whether key metabolites such as glucose and acetate affected sensitivity of non-cancerous breast epithelial cells to metformin.

CHAPTER 2

2. Methods

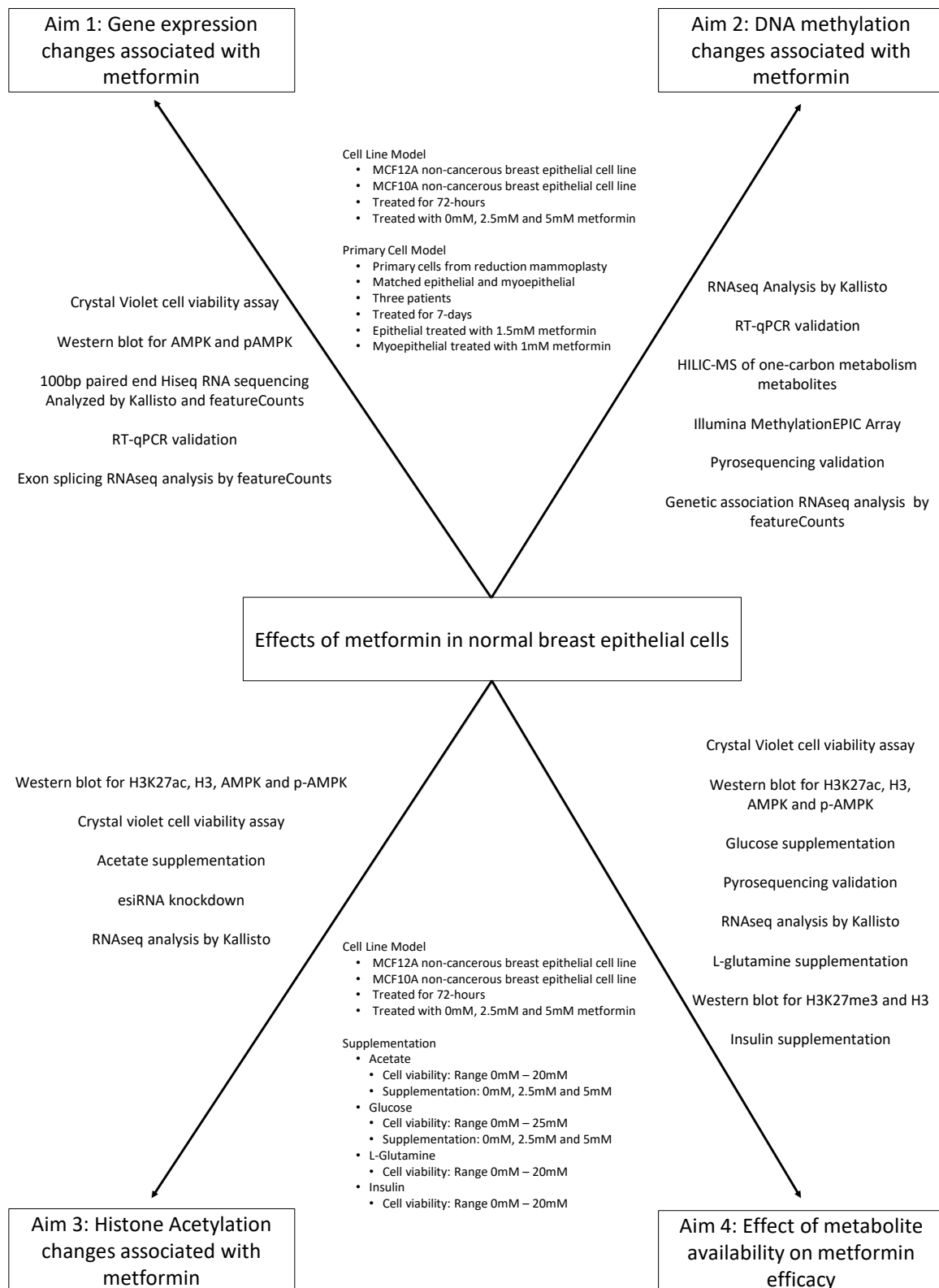


Figure 2.1: Methods overview

2.1 Samples

2.1.1 Cell lines

Cell Line	Cell type	Site of Origin	Doubling Time	Mutations	STR Profiling	Obtained from
MCF10A	Epithelial Fibrocystic disease from 36 years old adult female	Human mammary gland, breast	26 hours	ER-/PR- [Basal-like subtype]	Amelogenin: X CSF1PO: 12 D13S317: 8,9 D16S539: 11,12 D5S818: 10,13 D7S820: 10,11 THO1: 8,9,3 TPOX: 9,11 vWA: 15,17	Dr. Rachael Natrajan, Institute of Cancer Research (ICR), 123 Old Brompton Rd, Kensington, London SW7 3RP.
MCF12A	Epithelial Spontaneous immortalization from fibrocystic disease from 60 years old female adult	Human mammary gland, breast	33 hours	ER-/PR- [Basal-like subtype]	Amelogenin: TBA CSF1PO: TBA D13S317: TBA D16S539: TBA D5S818: TBA D7S820: TBA THO1: TBA TPOX: TBA vWA: TBA	Dr. Rachael Natrajan, Institute of Cancer Research (ICR), 123 Old Brompton Rd, Kensington, London SW7 3RP.
MCF7	Epithelial Adenocarcinoma from 69 years old female adult	Human mammary gland breast, obtained from metastatic site (pleural effusion).	30 hours	ER+/PR+ [Luminal A subtype]	Amelogenin: XX CSF1PO: 10,10 D13S317: 11,11 D16S539: 11,12 D5S818: 11,12 D7S820: 8,9 THO1: 6,6 TPOX: 9,12 vWA: 14,15	In lab stocks

Table 2.1: Details of cell lines

2.1.1.1 Cell line media information

Component	Stock Concentration	Final Concentration (500ml)	Obtained from
DMEM/F12	500ml	500ml	Life Technologies (11320074)
Foetal Calf Serum (FCS)	500ml	10%	First Link
Penicillin/Streptomycin	5000U/ml Penicillin 5000µg/ml Streptomycin	1%	Gibco (15070-063)
hEGF	100µg/ml	20ng/ml	Gibco (PHG0311L)
Hydrocortisone	1mg/ml	0.5µg/ml	Sigma-Aldrich (H0888)
Cholera Toxin	1mg/ml	100ng/ml	Sigma-Aldrich (C8052)
Insulin	4mg/ml	10µg/ml	Gibco (12585014)

Table 2.2: Details of media components for the subculturing of MCF10A and MCF12A non-cancerous breast epithelial cell lines – normal media conditions.

Component	Stock Concentration	Final Concentration (500ml)	Obtained from
DMEM (no glucose)	500ml	500ml	Gibco (11966-205)
Foetal Calf Serum (FCS)	500ml	10%	First Link
Penicillin/Streptomycin	5000U/ml Penicillin 5000µg/ml Streptomycin	1%	Gibco (15070-063)
hEGF	100µg/ml	20ng/ml	Gibco (PHG0311L)
Hydrocortisone	1mg/ml	0.5µg/ml	Sigma-Aldrich (H0888)
Cholera Toxin	1mg/ml	100ng/ml	Sigma-Aldrich (C8052)
Insulin	4mg/ml	10µg/ml	Gibco (12585014)

Table 2.3: Details of media components for the subculturing of MCF10A and MCF12A non-cancerous breast epithelial cell lines – glucose modified media conditions.

Component	Stock Concentration	Final Concentration (500ml)	Obtained from
DMEM/F12	500ml	500ml	Life Technologies (11320074)
Foetal Calf Serum (FCS)	500ml	10%	First Link
Penicillin/Streptomycin	5000U/ml Penicillin 5000µg/ml Streptomycin	1%	Gibco (15070-063)
hEGF	100µg/ml	20ng/ml	Gibco (PHG0311L)
Hydrocortisone	1mg/ml	0.5µg/ml	Sigma-Aldrich (H0888)
Cholera Toxin	1mg/ml	100ng/ml	Sigma-Aldrich (C8052)

Table 2.4: Details of media components for the subculturing of MCF10A and MCF12A non-cancerous breast epithelial cell lines – insulin modified media conditions.

Component	Stock Concentration	Final Concentration (500ml)	Obtained from
DMEM/F12 (no glutamine)	500ml	500ml	Life Technologies (21331020)
Foetal Calf Serum (FCS)	500ml	10%	First Link
Penicillin/Streptomycin	5000U/ml Penicillin 5000µg/ml Streptomycin	1%	Gibco (15070-063)
hEGF	100µg/ml	20ng/ml	Gibco (PHG0311L)
Hydrocortisone	1mg/ml	0.5µg/ml	Sigma-Aldrich (H0888)
Cholera Toxin	1mg/ml	100ng/ml	Sigma-Aldrich (C8052)
Insulin	4mg/ml	10µg/ml	Gibco (12585014)

Table 2.5: Details of media components for the subculturing of MCF10A and MCF12A non-cancerous breast epithelial cell lines – glutamine modified media conditions.

Component	Stock Concentration	Final Concentration (500ml)	Obtained from
DMEM	500ml	500ml	Sigma-Aldrich (D6429)
Foetal Calf Serum (FCS)	500ml	10%	First Link
Penicillin/Streptomycin	5000U/ml Penicillin 5000µg/ml Streptomycin	1%	Gibco (15070-063)

Table 2.6: Details of media components for the subculturing of MCF7 breast cancer cell line.

2.1.2 Primary Cells

2.1.2.1 Primary cell details

Patient ID	Tissue Type	Cell Type	Age	BMI	Diabetes
1989	Normal reduction	Epithelial	33	23.46	No
1989	Normal reduction	Myoepithelial	33	23.46	No
2357	Normal reduction	Epithelial	28	26.7	No
2357	Normal reduction	Myoepithelial	28	26.7	No
3002	Normal reduction	Epithelial	31	28.3	No
3002	Normal reduction	Myoepithelial	31	28.3	No
2585	Normal reduction	Epithelial	25	No information	No information
2585	Normal reduction	Myoepithelial	25	No information	No information

Table 2.7: Details of primary cell

2.1.2.2 Primary cell media information

Component	Stock Concentration	Final Concentration (500ml)	Obtained from
DMEM/F12 w/ 15mM HEPES	500ml	500ml	Sigma-Aldrich (D8437)
FCS	500ml	10%	First Link
Penicillin/Streptomycin	5000U/ml Penicillin 5000µg/ml Streptomycin	1%	Gibco (15070-063)
Amphotericin B (Fungizone)	250µg/ml	1% - 2.5µg/ml	Gibco/Life Technologies (15290026)
Hydrocortisone	0.5mg/ml	0.5µg/ml	Sigma-Aldrich (H0888)
Apo-transferrin	100mg	10µg/ml	Sigma-Aldrich (T1147)
hEGF	100µg/ml	10ng/ml	Gibco (PHG0311L)
Insulin	4mg/ml	5µg/ml	Gibco (12585014)

Table 2.8: Details of media components for the subculturing of primary non-cancerous breast epithelial cells

Component	Stock Concentration	Final Concentration (50ml)	Obtained from
HuMEC Basal Serum Free Media	500ml	50ml	Life Technologies (12753018)
Bovine Pituitary Extract (BPE)	500ml	0.4% - 52µg/ml	Gibco (13028014)
Amphotericin B (Fungizone)	250µg/ml	0.2% - 0.5µg/ml	Gibco/Life Technologies (15290026)
Gentamicin	10ml	10µg/ml	Sigma-Aldrich (G1397)
Hydrocortisone	0.5mg/ml	0.5µg/ml	Sigma-Aldrich (H0888)
hEGF	100µg/ml	10ng/ml	Gibco (PHG0311L)
Insulin	4mg/ml	5µg/ml	Gibco (12585014)

Table 2.9: Details of media components for the subculturing of primary non-cancerous breast myoepithelial cells

2.2 Culturing of cell lines

2.2.1 Preparation of MSC safety cabinet

A MSC safety cabinet hood was prepared for cell culture work by wiping with 70% ethanol and allowing sterile airflow to establish for 15 minutes.

2.2.2 Preparation of cell culture reagents

Appropriate media, 10% trypsin/EDTA (T4549 Sigma-Aldrich, 15070-63 Gibco) and phosphate buffer saline (PBS) were warmed in a water bath at 37°C.

2.2.3 Resuscitating of cell lines

Cryovials were removed from liquid nitrogen storage and placed in -80°C for 30 minutes to acclimatise cells to the increase temperature and pressure. Samples were allowed to defrost by gently swirling in a water bath at 37°C. 1ml of sample was added to a 10ml falcon and 9ml pre-warmed media was added in a drop wise fashion. The sample was centrifuged for 3 minutes at 900g Heraeus Multifuge X1 centrifuge (Thermo Scientific). The supernatant was removed, and the cell pellet was resuspended in 6ml fresh pre-warmed media. Sample was added to a pre-labelled T25 flask. The flask was gently rocked to allow the cells to cover the bottom of the flask. Cells were incubated at 37°C at 5% CO₂.

2.2.4 Subculturing of cell lines

Morphology and confluency of cells were checked before subculturing. Cell lines were subcultured when 70%-80% confluent. Media was removed and flasks were washed twice with pre-warmed PBS. PBS was removed and 1ml-2ml of 10% trypsin/EDTA was added to the bottom flask. Non-cancerous breast epithelial cell lines MCF10A and MCF12A were incubated at 37°C at 5% CO₂ for 15 minutes.

Cancerous breast epithelial cell line MCF7 was incubated at 37°C at 5% CO₂ for 3-5 minutes. Once cells had dissociated from the flask bottom, 6-8ml of pre-warmed media was added. Cells were centrifuged for 3 minutes at 900g using a Heraeus Multifuge X1 centrifuge (Thermo Scientific). The media was removed, and cell pellets were resuspended in 4-6ml of pre-warmed media. Cells were added to labelled T75 flasks at appropriate dilution. Cells were incubated at 37°C at 5% CO₂.

2.2.5 Counting of cell lines

1ml of cell suspension was removed and added to a fresh, labelled 1.5ml eppendorf centrifuge tube. 20µl cell suspension was added to 20µl trypan blue stain 0.4% (T10282, Life Technologies) in a fresh, labelled 1.5ml centrifuge eppendorf tube. 20µl of cell suspension/trypan blue stain 0.4% solution was added to a Countess™ cell counting chamber slide (C10283, Invitrogen). Cells were counted using the Countess™ automated cell counter (Invitrogen).

2.2.6 Mycoplasma testing of cell lines

1ml of cell free media from flask with cells growing for minimum 24 hours was taken from mycoplasma testing. Mycoplasma testing was carried out in-house using MycoAlert™ mycoplasma detection kit (Lonza, #LT07-318).

2.2.7 Freezing of cell lines

A freezing solution was made up for the number of desired cryovials. For 1ml of freezing solution 0.7ml appropriate pre-warmed media was added to 0.2ml foetal calf serum (FCS) and 0.1ml dimethyl sulfoxide (DMSO). Cells were trypsinised as detailed in 2.2.4. Following centrifugation of cell suspension, media was removed. Cell pellets were resuspended in 1ml freezing solution and added to

a labelled cryovial. Cells were stored at -80°C. Following 24 hours incubation at -80°C, cell samples were placed in liquid nitrogen storage.

2.3 Culturing of primary cells

2.3.1 Preparation of MSC safety cabinet

A MSC safety cabinet hood was prepared for cell culture work by wiping with 70% ethanol and allowing sterile airflow to establish for 15 minutes.

2.3.2 Preparation of cell culture reagents

Appropriate media, 10% trypsin/EDTA (T4549 Sigma-Aldrich, 15070-63 Gibco) and phosphate buffer saline (PBS) were warmed in a water bath at 37°C.

2.3.3 Resuscitating of primary cells

Cryovials were removed from liquid nitrogen storage and placed on ice. 5ml of appropriate medium for epithelial or myoepithelial primary cells was added to a 15ml tube. Cryovials were allowed to defrost by warming in hands. 1ml of appropriate medium was added to the cryovial. The contents of the cryovial were poured into the appropriate 15ml tube. Cryovials were washed twice with 1ml appropriate medium⁹. Appropriate medium was added to 15ml tubes to make total volume of 14ml. The sample was centrifuged for 3 minutes at 900g (Centrifuge 5810 R, Eppendorf). The supernatant was removed, and the cell pellet was gently resuspended using a p200 pipette set to 180µl. 1 ml of appropriate media was added to the cell suspension. 12µl cell suspension was added to 12µl trypan blue stain 0.4% (T10282, Life Technologies) in a fresh, labelled 1.5ml centrifuge eppendorf tube. 20µl of cell suspension/trypan blue stain 0.4% solution was added to a Countess™ cell counting chamber

slide (C10283, Invitrogen). Cells were counted using the Countess™ automated cell counter (Invitrogen).

2.3.4 Collagen coating of plastic ware for primary cell growth

0.02M acetic acid solution was made by adding glacial acetic acid with sterile PBS and filtered sterilized through a 0.22µm filter (Sartorius™ Minisart™ NML Syringe Filters, 16534). 0.02M acetic acid solution was added to 3.6mg/ml collagen I (Rat Tail 354236, BD Biosciences) for the following plastic ware:

Plate Type	Collagen per well (µg/ml)	0.02M acetic acid per well (ml)	Collagen I per well (ml)	Volume per well (ml)
96 Well	16 µg in 1ml	0.99545	0.00455	0.1
6 Well	47.5 µg in 2ml	1.98658	0.01342	2
6cm Dish	142.85 µg in 5ml	4.9544	0.04056	5

2.4 Cell viability assays

2.4.1 Chemicals

Metformin (PHR1084), D-Glucose (G7021) and Sodium acetate (S5636) were obtained from Sigma-Aldrich Inc., St. Louis, MA, USA and reconstituted to a concentration of 1M in sterile water. L-Glutamine 200nM (25030-024) was obtained from Gibco Life Technologies.

2.4.2 Crystal Violet assay for cell lines

2.4.2.1 Cell line seeding for crystal violet assay

Cells were seeded at 5×10^3 for 72 hours or 1×10^3 for 7 days in triplicate in a 96-well plate and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with increasing doses of Metformin (0mM – 25mM and 0mM – 15mM respectively). For 7-day culture, cells were re-treated every 3 days.

For glucose in brief, cells were plated as previously described in DMEM/F12 media (2.2) supplemented with increasing concentrations of glucose (0mM – 20mM). Following 24-hour incubation cells were treated with 2.5mM and 5mM metformin for 3 days.

For acetate in brief, cells were plated as previously described in DMEM/F12 media (2.2) supplemented with increasing concentrations of sodium acetate (0mM – 20mM). Following 24-hour incubation cells were treated with 2.5mM and 5mM metformin for 3 days.

For insulin in brief, cells were plated as previously described in DMEM/F12 media (2.2) supplemented with increasing concentrations of insulin (10µg/ml - 20µg/ml). Following 24-hour incubation cells were treated with 2.5mM and 5mM metformin for 3 days.

For glutamine in brief, cells were plated as previously described in DMEM/F12 media (2.2) supplemented with increasing concentrations of glutamine (0mM – 20mM). Following 24-hour incubation cells were treated with 2.5mM and 5mM metformin for 3 days.

2.4.2.2 Crystal violet

Media was removed and plate(s) were washed twice with PBS. Crystal violet reagent (0.5% crystal violet in 25% methanol) was added to plate(s) and incubated on a rocker for 20 minutes at room temperature. Plate(s) were washed with water until water runs clear. Plate(s) were allowed to air-dry overnight at room temperature. 10% acetic acid was added to plate(s) and incubated on a rocker at

room temperature for 20 minutes. Absorbance was read at 595nm using an OPTImax microplate reader (Molecular Devices).

2.4.3 Crystal Violet assay for primary cells

2.4.3.1 Primary cell seeding for crystal violet assay

Cell proliferation was measured using Crystal Violet assay. In brief, cells were seeded at 5×10^3 for 7 days in triplicate in a 96-well plate and incubated at 37°C and 5% CO₂. Following 6-hour incubation, cells were treated with increasing doses of Metformin (0mM – 25mM). Cells were re-treated once following 4-day treatment.

2.4.3.2 Crystal violet

In a class II MSC safety cabinet, media was removed, and plate(s) were washed twice with PBS. Cells were fixed by adding 100% methanol for 30 seconds. Crystal violet reagent (0.5% crystal violet (C6158, Sigma Aldrich) in 25% methanol) was added to plate(s) and incubated on a rocker for 20 minutes at room temperature. Plate(s) were washed with water until water runs clear. Plate(s) were allowed to air-dry overnight at room temperature. 10% acetic acid was added to plate(s) and incubated on a rocker at room temperature for 20 minutes. Absorbance was read at 595nm using an OPTImax microplate reader (Molecular Devices).

2.4.4 MTS assay for cell lines

Experiment was seeded as in 2.4.2 for 3 day and 7-day treatments.

Media was removed and plate(s) were washed once with PBS. Fresh media was added to plate(s). 20μl MTS reagent (G3580, CellTiter 96® Aqueous One Solution, Promega) was added to plate(s). Plate(s)

were incubated at 37°C at 5% CO₂ for 2.5 hours. Absorbance was read at 490nm using an OPTImax microplate reader (Molecular Devices).

2.5. Protein Study

2.5.1 Cell line seeding for protein extraction

Cells were seeded at 2.5×10^5 for 3 days in a 6cm dish and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 2.5mM and 5mM metformin. Cells were incubated at 37°C and 5% CO₂ for 3 days.

For glucose studies, cells were seeded in normal DMEM/F12 media. Following 24-hour incubation prior to metformin treatment media was changed to DMEM/F12 media with no glucose (Gibco Life Technologies, 11966-205) or DMEM/F12 media with no glucose supplemented with 5mM or 15mM glucose. Cells were treated with 2.5mM metformin and incubated at 37°C and 5% CO₂ for 3 days.

2.5.2 Primary cell seeding for protein extraction

Cells were seeded at 2.5×10^5 for 7 days in a 6cm dish and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 1.5mM and 1mM Metformin, for epithelial and myoepithelial primary cells respectively. Primary cells were re-treated following 4 days incubation.

2.5.3 Cytosolic protein extraction

2.5.3.1 Cytosolic protein extraction solution

Radioimmunoprecipitation (RIPA) buffer for protein extraction was made using the following:

Component	Stock Concentration	Final Concentration	Obtained from
Tris pH 8	-	50mM	Made in house
Sodium Chloride (NaCl)	-	150mM	
Triton-X 100	100X	1%	Fisher Scientific (BP151-500)
SDS	0.1%	10%	Sigma Aldrich (L3771)
Sodium deoxycholate	0.5%	0.5%	Sigma Aldrich (D6750)
Glycerol	100%	10%	
Sodium fluoride (NaF)	0.5M	50mM	Sigma Aldrich (57920)
Sodium orthovanadate (Na_3VO_4)	40mM	0.1mM	Sigma Aldrich (450243/S6508)
Dithiothreitol (DTT)	1M	2mM	BioRad (161-0610)
Phenylmethylsulfonyl fluoride (PMSF)	100mM	1mM	Sigma Aldrich (P7626)
Aprotinin	1000,000 U/ml	100U/ml	Calbiochem (#616370)
Leupeptin	10mg/ml	10µg/ml	Calbiochem (#108975)

Table 2.9: RIPA buffer components

2.5.3.2 Cell line extraction

Media was removed from the cells. Cells were washed twice with ice-cold PBS. Plates were drained for 1 minute at room temperature and the remaining PBS was aspirated. RIPA buffer, made using table, was added to cells. Cells were scraped into the lysis buffer and transferred to labelled, chilled 1.5ml eppendorf tubes. Sample(s) were incubated on ice for 20 minutes with occasional vortexing. Sample(s) were centrifuged at 16,200g at 4°C. Following centrifugation, the lysate was transferred to a chilled, labelled 1.5ml eppendorf. Sample(s) were stored at -20°C

2.5.4 Acid extraction

2.5.4.1 Acid extraction solution

Acid extraction buffer was made using the following:

Component	Stock Concentration	Final Concentration	Obtained from
Phosphate buffer saline (PBS)	-	-	Made in house
Triton-X 100	100X	0.5%	Fisher Scientific (BP151-500)
Sodium Azide (NaN ₃)	500X	0.02%	Sigma Alrich (S2002)
cOmplete Protease inhibitor cocktail	25X	25X	Roche Diagnostics (11697498001)

Table 2.10: TEB buffer components

2.5.4.2 Cell line acid extraction

Media was removed from the cells. Cells were washed twice with ice-cold PBS supplemented with 5mM sodium butyrate. Plates were drained for 1 minute at room temperature and the remaining PBS was aspirated. TEB buffer, made using table, was added to cells at a cell density of 1×10^7 cells per ml. Cells were scraped into the lysis buffer and transferred to labelled, chilled 1.5ml eppendorf tubes. Sample(s) were incubated on ice for 10 minutes with occasional vortexing. Sample(s) were centrifuged at 6,900g at 4°C for 10 minutes. Following centrifugation, the lysate was transferred to a chilled, labelled 1.5ml eppendorf. The pellet was retained and washed with half the volume of TEB buffer. Sample(s) were centrifuged at 6,900g at 4°C for 10 minutes. Following centrifugation, the lysate was removed. The pellet was retained and resuspended 0.2N hydrochloric acid (HCl) at a cell density 4×10^7 cells per ml. Samples(s) were incubated overnight at 4°C on a rotator. Following incubation sample(s) were centrifuged at 6,900g at 4°C for 10 minutes. Following centrifugation, the lysate was transferred to a chilled, labelled 1.5ml eppendorf. Sample(s) were stored at -20°C.

2.5.4.3 Primary cell acid extraction

Performed as in 2.5.4.2

2.5.5 Concentration analysis of protein extracts

A standard curve was created using albumin protein at known concentration (2mg/ml) with the following dilutions:

Dilution	Volume known protein (μl)
0	0
0.4	2 – diluted 1:10 in DNase/RNase free water
1	5 – diluted 1:10 in DNase/RNase free water
2	1 – undiluted
4	2 – undiluted

500μl Quick Start Bradford 1X Dye Reagent (500-0205, BioRad) was added to 1μl protein sample. Sample(s) were vortexed for 5 seconds at medium-high speed. Sample(s) were incubated for 15 minutes at room temperature. Absorbance was read in triplicate at 595nm on an OPTImax microplate reader (Molecular Devices). A standard curve was generated, and the following equation was used to calculate unknown protein concentration.

2.6 Western Blot

2.6.1 Preparation of gels for SDS-PAGE Gel Electrophoresis

1mm glass cassettes, consisting of two glass plates, were washed with detergent and wiped with ethanol. Cassettes were prepared in a cassette clamp and gel-casting stand. Cassettes were checked for leaks before use. For cytosolic proteins a 10% Resolving Gel was prepared. For nuclear proteins a 12% resolving gel was made. The cassette(s) filled with resolving gel, leaving approximately 2cm gap. Resolving gel(s) were covered in approximately 1ml isopropanol. Gel(s) were left to set for 10-15 minutes at room temperature. The isopropanol was removed, and gel(s) were gently rinsed with water. A 5% Stacking Gel was prepared and added to the top of the resolving gel. A gel comb was placed into the top of the cassette. The stacking gel was allowed to set for 10-15 minutes at room temperature.

Component	Stock Concentration	Final Concentration	Obtained from
30% Acrylamide	37.5:1, 30% solution	37.5:1, 30% solution	VWR International Ltd., J61505.AP
Tris/SDS (pH 8.8)	1.5M	1.5M	Made in house
Tris/SDS (pH 6.8)	1.5M	1.5M	Made in house
Ammonium persulfate (APS)	10%	10%	Sigma Aldrich, A3678
Tetramethylethylenediamine (TEMED)	100%	0.1%	Sigma Aldrich, T9281

Table 2.11: Western blot gel components

Component	7% Gel (ml)	10% Gel (ml)	12.5% Gel (ml)
Water	7.75	4.75	2.25
30% Acrylamide	7	10	12.5
Tris/SDS (1.5M, pH8.8)	15	15	15
10% APS	0.225	0.225	0.225
TEMED	0.025	0.025	0.025

Table 2.12: Resolving gel components

Component	3% Gel (ml)	4% Gel (ml)	5% Gel (ml)
Water	5.9	5.4	4.9
30% Acrylamide	1.5	2	2.5
Tris/SDS (1.5M, pH6.8)	7.5	7.5	7.5
10% APS	0.09	0.09	0.09
TEMED	0.01	0.01	0.01

Table 2.13: Stacking gel components

2.6.2 Preparation of protein for SDS-PAGE gel Electrophoresis

10µg of protein was added into appropriate labelled 1.5 ml Eppendorf kept on ice. The volume of sample(s) was equalized in water. 5µl Add 3X Laemli loading buffer was added to each sample for a final volume of 20µl. Samples were boiled at 95°C for 5 minutes on a heating block. Sample(s) were placed on ice for 1 minute before centrifuging at 16,200g for 30 seconds. Sample(s) were kept on ice until used.

2.6.3 SDS-PAGE Gel Electrophoresis

Gel(s) were placed in a negative electrode chamber and clamped in with the front facing inwards. The chamber was placed in a tank filled ¼ way up with 1X SDS running buffer (made in house). The centre of the chamber was filled with 1X SDS running buffer. The comb was gently removed. Using a gel loading pipette tip round (Costar, #4853), sample(s) were added to appropriate wells. A ladder (Page Ruler Plus Prestained protein ladder. 26619, Thermo Scientific) was added. The tank was filled with 1X SDS running buffer and the gel(s) were run at 80V for 20 minutes and 120V for 1.5 hours.

2.6.4 Transfer to PVDF membrane

2.6.4.1 iBlot

Protein was transferred to a PVDF was conducted using the iBlot dry transfer system (Thermo Fisher Scientific Invitrogen). In brief, following SDS-PAGE gel electrophoresis gel(s) were removed carefully from cassette(s). Gels were placed on the bottom cathode (comprised of filter paper and PVDF membrane). Filter paper was wetted and placed over the gel. Air bubbles were removed using a roller. The top cathode was placed on top of gel(s), gel side down. Air bubbles were removed. The completed stack was placed in the iBlot machine.

The following programmes were used for transfer:

Protein Extraction Type	iBlot Programme	Voltage	Time
RIPA	P3	30V	5 minutes
TEB/Acid	P4	20V	5 minutes

2.6.4.2 Wet transfer

1X transfer buffer (10X, made in house) was made before transfer and placed at 4°C for 1 hour. The PVDF membrane was prepared by soaking the PVDF membrane (Amersham™ Hybond™ 0.2µm blotting membrane, GE Healthcare Life Sciences, 10600021) in 100% methanol (VWR Chemicals) for 15 seconds. PVDF membrane was washed in water for 30 seconds and covered in cooled 1X transfer buffer. Transfer cassette(s) were prepared by placing the cassette white side down. The PVDF membrane, filter paper and sponges were soaked in cooled 1X transfer buffer. The following were placed into the cassette(s); sponge, filter paper, PVDF membrane, gel(s), sponge. The cassette(s) were closed, and excess buffer was removed. The cassette(s) were placed in transfer stand(s). An ice pack was placed in the transfer stand. Transfer stand(s) were placed in a tank, which was filled with cooled 1X transfer buffer. Transfer(s) were run at 30V for 70 minutes. Following transfer PVDF membrane was placed in 100% methanol for 15 seconds. PVDF membrane(s) were washed in water for 2 minutes before blocking.

2.6.5 Blocking of unspecific binding

Once transfer was completed, the membrane was removed from the stack and placed in 5% Bovine Serum Albumin (BSA, ChemCruz sc2323) in 1X Tris-buffered Saline with Tween 20 (TBST, made in house) supplemented with sodium azide (NaN_3 , 500X, Sigma Alrich S2002). Membranes were incubated in 5% BSA/TBST/ NaN_3 solution for 1 hour at room temperature on a rocker. For H3 staining, membranes were blocked in 5% Milk/TBST/ NaN_3 solution for 1 hour at room temperature on a rocker.

2.6.6 Antibody incubation

Membrane(s) were incubated in primary antibody solutions at the following:

Antibody	Organism	Dilution	Solvent	Time point	Temperature	Rocking	Obtained From
AMPK α	Rabbit	1:1,000	5% BSA/TBST/NaN ₃	Overnight	4°C	Yes	Cell Signalling Technology, #2532
pAMPK α (Thr172)(40H9)	Rabbit	1:1,000	5% BSA/TBST/NaN ₃	Overnight	4°C	Yes	Cell Signalling Technology, #2535
Cyclin D1	Rabbit	1:1,000	5% BSA/TBST/NaN ₃	Overnight	4°C	Yes	Cell Signalling Technology, #2978
H3K27ac	Rabbit	1:1,000	5% BSA/TBST/NaN ₃	Overnight	4°C	Yes	Millipore #07-441
H3K27me3	Rabbit	1:1,000	5% BSA/TBST/NaN ₃	Overnight	4°C	Yes	
H3 (D1H2)	Rabbit	1:2,000	5% Milk/TBST/NaN ₃	Overnight	4°C	Yes	Cell Signalling Technology, #4499
β -actin (loading control)	Mouse	1:25,000	5% BSA/TBST/NaN ₃	1 hour	Room Temperature	Yes	Novus Biologicals, NB600-501

Table 2.14: Primary antibodies

Following incubation membrane(s) were washed three times in 1X TBST for 10-20 minutes.

Membrane(s) were incubated in secondary antibody solutions at the following:

Antibody	Organism	Dilution	Solvent	Time point	Temperature	Rocking	Obtained From
AMPK	Goat Anti-Rabbit IgG HRP Linked	1:5,000	5% Milk/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7074
pAMPK (Thr172)	Goat Anti-Rabbit IgG HRP Linked	1:1,000	5% BSA/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7074
Cyclin D1	Goat Anti-Rabbit IgG HRP Linked	1:5,000	5% Milk/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7074
H3K27ac	Goat Anti-Rabbit IgG HRP Linked	1:5,000	5% Milk/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7074
H3K27me3	Goat Anti-Rabbit IgG HRP Linked	1:5,000	5% Milk/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7074
H3	Goat Anti-Rabbit IgG HRP Linked	1:5,000	5% Milk/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7074
β -actin (loading control)	Horse Anti-Mouse IgG HRP Linked	1:50,000	5% Milk/TBST	1 hour	Room Temperature	Yes	Cell Signalling Technology, #7076

Table 2.15: Secondary antibodies

Following incubation membrane(s) were washed six times in 1X TBST for 5-10 minutes.

2.6.7 Chemiluminescence by enhanced chemiluminescence (ECL)

Membrane(s) were visualized using the Pierce ECL Western Blotting Substrate Kit (32106, Thermo Scientific). In brief, following TBST washing membrane(s) were added to a 1:1 solution of Detection Reagent 1 (peroxide solution) and Detection Reagent 2 (Luminol Enhancer). Membrane(s) were incubated for 1 minute at room temperature. Membrane(s) were placed between two layer of polyethylene plastic and bubbles were removed. Membrane(s) were placed in either the GE LAS4000 Chemiluminescence imager or the Fusion Solo ChemilMager for chemiluminescence analysis. Membrane(s) were exposed in 1-minute increments for 10-15 minutes. For the loading control, membrane(s) were exposed for 10 second increments for 2 minutes.

2.6.8 Stripping and re-staining

Membrane(s) were placed in Restore Plus Western Blot Stripping Buffer (Thermo Scientific, 46430). Membrane(s) were incubated for 15 minutes on a rocker at room temperature. The stripping buffer was removed, and the membrane(s) were washed with sterile water for 2-3 minutes. Membrane(s) were washed 3 times with TBST for 10 minutes. Membrane(s) were blocked with 5% BSA/TBST/ NaN_3 for 30 minutes on a rocker at room temperature. Membrane(s) were re-stained for proteins following steps 2.6.6 and 2.6.7.

2.7 DNA Methylation study

2.7.1 Cell line seeding for DNA extraction

Cells were seeded at 1.25×10^5 for 3 days in a 6cm dish and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 2.5mM and 5mM metformin. Cells were incubated at 37°C and 5% CO₂ for 3 days.

2.7.2 DNA Extraction

DNA was extracted using Puregene Core Kit B (1042608, Qiagen). In brief, cells were pelleted by trypsinisation (see 2.2.4). Pellets were washed with PBS, leaving 20µl. Pellets were vortexed vigorously. 300µl Cell Lysis Solution was added to sample(s) to resuspend cells. Cells were lysed at 37°C until solution is homogenous. 1.5µl RNase A Solution was added to sample(s) and mixed by inversion 25 times. Sample(s) were incubated at 37°C for 5 minutes. Sample(s) were incubated on ice for 1 minute. 100µl Protein precipitation solution was added and sample(s) were vortexed vigorously. Sample(s) were incubated on ice for 5 minutes. Sample(s) were centrifuged at 13,000-16,000g for 1 minute at room temperature. Supernatant was added by pouring to 300µl isopropanol in a fresh 1.5ml eppendorf tube. Sample(s) were mixed by inversion gently 50 times. Sample(s) were centrifuged at 13,000-16,000 for 1 minute at room temperature. Supernatant was discarded and sample(s) were washed twice with 70% ethanol. Sample(s) were centrifuged at 13,000-16,000g for 1 minute at room temperature. Supernatant was removed and pellet(s) were allowed to air dry for 10 minutes at room temperature. Pellet(s) were resuspended in 100µl DNA Hydration Solution and vortexed for 5 seconds at medium speed. Sample(s) were incubated at 65°C for 1 hour. Samples(s) were incubated shaking at 700rpm overnight at room temperature. Samples were stored at -20°C.

2.8 Bisulfite conversion

2.8.1 Preparation of whole genome amplified DNA – 100% methylated DNA control.

All steps were conducted in a pre-PCR hood. 10 ng/μl concentration of DNA samples were prepared with nuclease free water in a 0.2 ml tube. 9 μl Sample Buffer was added to 1 μl of 10ng template DNA. Template DNA was denatured by heating samples to 95°C for 3 minutes, then cool to 4°C on ice. 9 μl of Reaction Buffer was added with 1 μl of Enzyme mix on ice. 10 μl of the master mix was added to cooled samples on ice. Sample(s) were incubated at 30°C for 1 hour and 30 min. Samples were heated to 65°C for 10 min and then cooled to 4°C.

2.8.2 Purification of PCR product was conducted using the QIAquick PCR purification kit (D5003, Qiagen, 28106).

Reagents were prepared by adding 220 ml of Ethanol to the PE Buffer prior to use. 5 volumes Buffer PB was added to 1 volume of the PCR sample and mix. A QIAquick spin column provided in a 2ml collection tube. The sample mix from 2.8.1 was added to column(s) and centrifuged at 17,900g at room temperature for 30-60 seconds. The flow-through was discarded. 750 μl Buffer PE to the column was added to column(s) and centrifuged at 17,900g at room temperature for 30-60 seconds. The flow-through was discarded. Sample(s) were centrifuged at 17,900 x g at room temperature for 1 min. Column(s) were placed in a clean 1.5 ml micro centrifuge tube and label accordingly. 30 μl Buffer EB was added to the centre of the column(s). Sample(s) were incubated for 1 minute and centrifuged at 17,900g at room temperature for 1 minute.

2.8.2 Bisulfite conversion (EZ DNA Methylation kit, D5002, Zymo).

Reagents were prepared as follows - 750μl water and 210μl of M-Dilution Buffer were added to one tube of CT Conversion Reagent. CT Conversion Reagent was mixed at room temperature for 15

minutes. 24ml of 100% Ethanol was added to the M-Wash Buffer concentrate before use. In a 96 well PCR clear plate the following was added to each well and mixed by pipetting up and down:

Component	Volume per reaction (µl)
M-Dilution Buffer	5 µl
DNA	Variable (maximum 45 µl (500-2ug))
RNase-free water	To final volume 50 µl
Total Volume	50 µl

Sample(s) were incubated at 37 °C for 15 min. 100µl of prepared CT Conversion Reagent was added to each sample and mixed. Sample(s) were incubated in the following steps:

Temperature (°C)	Time (minutes)	Cycles
95 °C	30 seconds	16
50 °C	60 minutes	
4 °C	10 minutes	1

Bisulfite converted samples were cleaned in brief, 400 µl of M-Binding Buffer was added to each spin column. Converted sample(s) were added onto the columns and mixed by pipetting up and down. Sample(s) were centrifuged at $\geq 10,000g$ for 30 seconds and the flow-through was discarded. 100 µl of M-Wash Buffer was added to each sample and centrifuged at $\geq 10,000g$ for 30 seconds. 200µl of M-Desulphonation Buffer was added to sample(s) and incubated at room temperature for 15-20 min. Sample(s) were centrifuged at $\geq 10,000g$ for 30 seconds and the flow through was discarded. 200 µl of

M-Wash Buffer was added to sample(s) and centrifuged at $\geq 10,000g$ for 30 seconds. Flow through was discarded and an additional 200 μl was added to sample(s). Sample(s) were centrifuged at $\geq 10,000g$ for 30 seconds. Sample column(s) were placed on fresh, labelled 1.5ml eppendorf(s). 15 μl of M-Elution Buffer was added directly to the binding matrix in each column. Sample(s) were incubated for 1-2 minutes at room temperature and centrifuged at $\geq 10,000g$ for 3 minutes.

2.8.3 Analysis of DNA quantity by Qubit

A working solution was prepared by adding 1:200 reagent:buffer per sample. DNA standards were prepared by adding 190 μl working solution to 10 μl standard in a 0.5ml PCR tube (Axygen, 321-05-051). Samples were vortexed for 5 seconds and incubated at room temperature for 2 minutes. Fluorescence of standards were read using a Qubit 2.0 Fluorometer (Thermo Fisher) set for High Sensitivity dsDNA. For samples, 199 μl working solution was added to 1 μl sample. Sample(s) were vortexed for 5 seconds and incubated at room temperature for 2 minutes. Fluorescence of sample(s) were read using a Qubit 2.0 Fluorometer.

Concentration of sample(s) were noted in ng/ml. Concentration(s) in ng/ μl were obtained using the following calculation:

$$\text{Calculate Concentration} = [(QF \text{ Value}) \times (200/x)]/1000$$

Where QF = value given by machine

x = volume of sample for detection

2.8.4 Quality Control of bisulfite converted DNA

Confirmation of bisulfite conversion of DNA sample(s) was conducted using PCR for the calcium binding protein Calponin. PCR primers were specifically designed to only amplify bisulfite converted DNA (appendix).

2.8.5 PCR of bisulfite converted DNA

2.8.5.1 Preparation of primers

Oligo(s) were reconstituted with nuclease free water to 100 μ M stock solution. Primer(s) were diluted to 20 μ M working solution in PCR grade nuclease free water.

2.8.5.2 Preparation for PCR amplification

PCR reagents were defrosted on ice. A PCR master-mix was made by combining the following components:

Component	Volume X1 (μ l)	Final Concentration
10X PCR buffer	2.5	1X
MgCl ₂ (25mM)	3	3mM
2.5mM dNTP mixture	0.5	0.05mM
Forward primer	1	0.8 μ M
Reverse primer	1	0.8 μ M
Taq DNA polymerase (5U/ μ l)	0.2	1U
PCR grade water	15.8	-

24 μ l master mix was added to 1 μ l bisulfite converted DNA or PCR grade water to a final volume of 25 μ l in a 96 well PCR plate (Star Lab, I1402-9800). The plate was sealed using adhesive sealing sheets (Thermo Scientific, AB-0558) and briefly centrifuged. Sample(s) were placed in a thermal cycler using the following parameters:

Step	Temperature (°C)	Time (minutes)	Cycles
Initial Denaturation	95	6 minutes	1
Denaturation	95	30 seconds	40
Annealing	63	30 seconds	
Extension	72	30 seconds	
Final Extension	72	5 minutes	1
Hold	4	Forever	Forever

Sample(s) were kept at -20°C for long-term storage.

2.8.5.3 2% Gel Electrophoresis

Sample(s) generated in 2.8.5.2 were prepared by adding 3µl sample to 1µl Orange G-based loading buffer. 2% agarose gel electrophoresis was used to analyse quality of sample(s) generated in 2.8.5.2. 2g of agarose (Fisher Scientific, BP1356) was dissolved in 100ml 1X TBE (Tris-borate-EDTA, made in house) by microwaving for 1-2 minutes in 30 second bursts. Gloves were used to handle glassware. Dissolved agarose solution was allowed to cool slightly before (i) ethidium bromide (10mg/ml) in fume cupboard or (ii) Gel Red dye (1:10,000 dilution). A gel caster tray was assembled with combs positioned in the caster tray. The warm agarose gel solution was poured into the gel caster tray to approximately ¼ to ½ inch thick and allowed to cool within a fume cupboard. Once set, the gel caster tray was placed in a horizontal electrophoresis unit, combs at the cathode end, filled with 1X TBE

buffer. The comb was removed, and sample(s) were loaded into wells. A molecular weight marker was added to each lane of wells. The electrophoresis unit was attached to a power pack set to 75V. Gel(s) were resolved for 30 minutes – 1 hour. The gel was removed from the electrophoresis unit and placed in a transilluminating machine. Gel(s) were visualised by UV light. The molecular weight marker was used to estimate the size of DNA fragments.

2.9 Illumina MethylationEPIC Array

Illumina MethylationEPIC Array was conducted by the SPH Genotyping facility (Burlington Danes Building, Imperial College London, Hammersmith Campus, Du Cane Road, London, W12 0NN, UK). 500ng/ μ l bisulfite converted DNA was provided for Illumina MethylationEPIC Array.

2.10 Pyrosequencing

2.10.1 Preparation of Sequencing Primer(s)

Oligo(s) were reconstituted with nuclease free water to 100 μ M stock solution. Primer(s) were diluted to 10 μ M working solution in PCR grade nuclease free water.

2.10.2 Preparation of PCR plate(s)

A master mix with the following components was prepared:

Component	Volume X1
Pyromark Binding buffer (Qiagen, 979006)	39µl
Streptavidin Sepharose High Performance beads (GE Healthcare, 17-5113-01)	1µl
Water	35µl

In a clear 96 well plate 75µl master mix was added per well. 5µl PCR product generated in 2.8.5.2 was added per well. Three controls were added per plate; a non-template control, 100% methylated DNA control and 0% methylated DNA control. Plate(s) were sealed with adhesive sealing sheets (Thermo Scientific, AB-0558).

2.10.3 Preparation of sequencing plate(s)

Sequencing primer(s) were diluted to 0.3µM in annealing buffer (Qiagen, 1056481). 12µl diluted sequencing primer(s) was added to each well in a 96 well white sequencing plate (PyroMark Q96 HS).

2.10.4 Preparation for Pyrosequencing

Before use, the PyroMark Q96 MD pyrosequencing machine was turned on for 1 hour and a heat block was warmed to 80°C. The PCR plate was agitated for 10 minutes at 1400rpm at room temperature. The vacuum handset tool was washed with water before use. Wash stations were set up before use;

70% ethanol, denaturation buffer (0.2mM NaOH), Pyromark wash buffer (Qiagen, 979008) and water. A vacuum (450mmHg) was applied to the vacuum hand tool. The hand tool was carefully lowered into the PCR plate to capture beads with sample(s) attached. The beads were washed in 70% ethanol for 5 seconds, denaturation solution for 5 seconds and wash buffer for 10 seconds. The vacuum hand tool was raised and allowed to drain. The vacuum was turned off. The vacuum hand tool was carefully placed into the sequencing plate and gently shaken. The vacuum hand tool was removed and placed in water to wash. The sequencing plate was incubated at 80°C for 2 minutes and incubated at room temperature for 5-10 minutes.

2.10.5 Preparation of dispensing tips

Lyophilized enzyme and substrate mixtures were dissolved in 620µl high purity water and mixed by swirling. Reconstituted enzyme and substrate mixes were stored at -20°C. The dispensing tip holder was prepared with the following; enzyme, substrate, dATP, dCTP, dGTP, dTTP. Relevant volumes, calculated by Pyro Q-CpG program, was added to appropriate tips using unfiltered pipette tips. Dispensing tips were placed into the dispensing tip holder and placed in the PyroMark Q96 MD pyrosequencing machine. Dispensation was tested three times before use. Sequencing plate(s) were placed in the PyroMark Q96 MD pyrosequencing machine and the sequencing plate was run.

2.11 RNA extraction

2.11.1 Cell line seeding for RNA extraction

Cells were seeded at 1.25×10^5 for 3 days in a 6-well plate and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 2.5mM and 5mM metformin. Cell were incubated at 37°C and 5% CO₂ for 3 days.

2.11.2 Primary cell seeding for RNA extraction

Cells were seeded at 4×10^4 for 7 days in a 6-well plate and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 1.5mM and 1mM Metformin, for epithelial and myoepithelial primary cells respectively. Primary cells were re-treated following 4 days incubation.

2.11.3 RNA Extraction

RNA extraction was conducted using RNeasy Mini Kit (74104, Qiagen). Before extraction buffer RPE was prepared by adding 4 volumes of 96% ethanol to one volume buffer RPE. 1×10^7 cells were harvested per extraction. In brief, media was removed from plate(s) and plate(s) were washed twice with ice-cold PBS. PBs was discarded. For 6 well plates 350μL buffer RTL was added. For 6cm dishes 600μL buffer RTL was added. A cell scraper was used to lyse cells. Lysate(s) were added to QIAshredder (info) spin columns and centrifuged at 8000g for 2 minutes. Lysates were added to gDNA spin columns and centrifuged at 8000g for 1 minute. 1 volume of 70% ethanol was added to lysate(s). Lysate(s) were added to RNeasy mini spin column(s) on 2ml collection tube(s). Sample(s) were centrifuged at 8000g for 30 seconds. Supernatant(s) were discarded. 700μL buffer RW1 was added to RNeasy mini spin column(s). Sample(s) were centrifuged at 8000g for 30 seconds. Supernatant(s) were discarded. 500μL buffer RPE was added to RNeasy mini spin column(s). Sample(s) were centrifuged at 8000g for 30 seconds. Supernatant(s) were discarded. This step was repeated. Sample(s) were centrifuged at 16,200g for 1 minute. The RNeasy mini spin column(s) were placed on fresh labeled 1.5ml eppendorf tube(s). 30μL RNA elution buffer was added directly to the membrane to the spin column. Sample(s) were centrifuged at 8000g for 2 minutes.

2.11.4 Analysis of RNA quantity by NanoDrop 1000 Spectrophotometer (Thermo Scientific)

RNA sample(s) were analyzed by NanoDrop. The NanoDrop was blanked using RNA elution buffer. 1 μ L sample was added to the NanoDrop platform and RNA sample(s) were analyzed. Results were recorded.

2.12 qPCR

2.12.1 RNA extraction

RNA was extracted as in 2.11

2.12.2 cDNA synthesis (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems, 4369914)

RNA sample(s) were prepared to 100ng in RNase free water to a final volume of 10 μ L. cDNA synthesis master mix was prepared with the following components:

Component	Volume (μ l) per sample
10X RT Buffer	2
25x dNTP mix	0.8
10x RT Random primers	2
Multiscribe Reverse Transcriptase	1
Nuclease Free Water	4.2

To 0.2ml PCR tubes 10 μ l Master Mix was added to 10 μ l RNA mix. Samples were placed in a thermocycler using the following protocol:

Temperature (° C)	Time (minutes)
25° C	10 minutes
37° C	120 minutes
85° C	5 minutes
4° C	Hold

Sample(s) were diluted to 20ng/μl by adding 30μl nuclease free water.

2.12.2 qPCR – primer validation

2.12.2.1 Preparation of primers

Oligo(s) were reconstituted with nuclease free water to 100μM stock solution. Primer(s) were diluted to 5μM working solution in PCR grade nuclease free water.

2.12.2.2 Preparation of SYBR green/primer master mix

Master mix was prepared by adding 5μl SYBR green to 0.6μl 5μM primer working solution per sample.

2.12.2.3 Preparation of cDNA Standard Curve

For standard curve, combine 1μl each cDNA sample in a pool. Dilute cDNA 1, ¼, 1/16, 1/64, 1/256 in nuclease free water, where standard curve 1 equals 20ng/μl. cDNA standard curves were diluted to 10ng in nuclease free water.

2.12.2.4 Preparation of PCR plate

To a MicroAmp® Fast 96-Well Reaction Plate (Applied Biosystems, 4346907) plate 5.6µl master mix was added to 4.4µl cDNA standard pool in triplicate. Plate(s) were sealed using MicroAmp™ Optical Adhesive Film (Applied Biosystems, 4311971). Plate(s) were mixed by gentle vortexing and spun briefly.

2.12.2.5 qPCR run

qPCR plates were run on a 7900HT Real Time PCR machine with a 96 well plate stand. Plates were run with the following protocol:

Stage	Temperature (° C)	Time (minutes)	Cycles
1	95	0:20	1
2	95	0:01	40
	60	0:20	
	95	0:15	1
3	60	0:15	
4	95	0:15	1

2.12.3 qPCR – target genes

2.12.3.1 Preparation of SYBR green/primer master mix

Master mix was prepared by adding 5µl SYBR green to 0.6µl 5µM primer working solution per sample.

2.12.3.2 Preparation of PCR plate

To a white 96-well Multiplate® PCR Plate™ (Biorad, MLL9651) plate 5.6µl master mix was added to 4.4µl cDNA sample(s) in triplicate. Plate(s) were sealed using MicroAmp™ Optical Adhesive Film (Applied Biosystems, 4311971). Plate(s) were mixed by gentle vortexing and spun briefly.

2.12.3.3 qPCR run

qPCR plates were run on a CFX96™ Real Time System (Biorad) with a 96 well plate stand. Plates were run with the following protocol:

Stage	Temperature (° C)	Time (minutes)	Cycles
1	95	0:20	1
2	95	0:01	40
	60	0:20	
	95	0:15	1
3	60	0:15	
4	95	0:15	1

RQ value(s), used to estimate relative gene expression, was calculated using the following calculation:

$$\Delta Ct = Ct_{\text{target gene}} - Ct_{\text{endogenous control}}$$

$$\Delta\Delta Ct = \Delta Ct_{\text{sample}} - \Delta Ct_{\text{calibrator sample}}$$

$$RQ = 2^{-(\Delta\Delta Ct)}$$

2.13 RNA Sequencing

2.13.1 Cell line seeding for RNA extraction

Cell lines were seeded as in 2.11.1.

2.13.2 RNA extraction

RNA was extracted as in 2.11.3.

2.13.3 Assessing RNA quality

2.13.3.1 Qubit analysis of RNA concentration

A working solution was prepared by adding 1:200 reagent:buffer per sample. DNA standards were prepared by adding 190µl working solution to 10µl standard in a 0.5ml PCR tube (Axygen, 321-05-051). Samples were vortexed for 5 seconds and incubated at room temperature for 2 minutes. Fluorescence of standards were read using a Qubit 2.0 Fluorometer set for ssRNA. For samples, 199µl working solution was added to 1µl sample. Sample(s) were vortexed for 5 seconds and incubated at room temperature for 2 minutes. Fluorescence of sample(s) were read using a Qubit 2.0 Fluorometer. Concentration of sample(s) were noted in ng/ml. Concentration(s) in ng/µl were obtained using the following calculation:

$$\text{Calculate Concentration} = [(QF \text{ Value}) \times (200/x)]/1000$$

Where QF = value given by machine

x = volume of sample for detection

2.13.3.2 Bioanalyzer analysis of RNA quality

Before sample(s) were prepared, High Sensitivity RNA ScreenTape Sample buffer (5067-5580, Aligent Technologies), High Sensitivity RNA ScreenTape ladder (5067-5581, Aligent Technologies) and High Sensitivity RNA ScreenTape (5067-5579, Aligent Technologies) were allowed to reach room temperature. Sample(s) were prepared in optical tubes (Aligent technologies, #401428) by adding 2µl RNA to 1µl High Sensitivity RNA ScreenTape Sample buffer. Ladder was prepared in optical tubes (Aligent technologies, #401428) by adding 2µl High Sensitivity RNA ScreenTape ladder to 1µl High Sensitivity RNA ScreenTape Sample buffer. Sample(s) were centrifuged briefly and vortexed at 2000rpm for 60 seconds. Sample(s) were centrifuged briefly. Sample(s) were denatured at 72°C for 3 minutes and placed on ice for 2 minutes. Sample(s) were centrifuged briefly. Sample(s) were placed in an Aligent 2200 Tape Station system. Loading tips (Aligent Technologies, #5067-5153) were placed into the tape station and High Sensitivity RNA ScreenTape was placed into the tape station and run.

2.13.3.3 RNA sequencing RNA quality conditions

RNA sample(s) were considered for RNA sequencing library preparation if RIN value(s) were >7 and minimum RNA concentration of 500ng.

2.13.4 Preparation of RNA Sequencing Library

2.13.4.1 rRNA depletion (E6310)

rRNA depletion was conducted using the NEBNext rRNA Depletion Kit (E6310, New England Biolabs).

In brief, 500ng RNA sample(s) were diluted to a final volume of 12µl in nuclease free water. Probe hybridization master mix was made by combining the following:

RNA/Probe Master Mix	Volume (μ l)
NEBNext rRNA depletion solution	1
Probe hybridization buffer	2

3 μ l probe hybridization master mix was added to RNA sample(s) and mixed thoroughly by gently pipetting up and down. Sample(s) were briefly centrifuged before placing in a thermal cycler with the following programme with heated lid set at 105°C:

Temperature ($^{\circ}$ C)	Time
95	2 minutes
Ramp down to 22	0.1 $^{\circ}$ C/sec
22	5 minutes

Sample(s) were briefly centrifuged. RNase H digestion master mix was made by combining the following:

RNase H Master Mix	Volume (μ l)
NEBNext RNase H	2
NEBNext RNase H reaction buffer	2
Nuclease free water	1

5µl RNase H master mix was added to RNA sample(s) and mixed thoroughly by gently pipetting up and down. Sample(s) were briefly centrifuged before placing in a thermal cycler at 37°C for 30 minutes with heated lid set at 40°C.

Sample(s) were briefly centrifuged. DNase I digestion master mix was made by combining the following:

DNase I Master Mix	Volume (µl)
DNase I reaction buffer	5
DNase I (RNase-free)	2.5
Nuclease free water	22.5

30µl DNase I digestion master mix was added to RNA sample(s) and mixed thoroughly by gently pipetting up and down. Sample(s) were briefly centrifuged before placing in a thermal cycler at 37°C for 30 minutes with heated lid set at 40°C.

Sample(s) were briefly centrifuged. Sample(s) were cleaned with homemade paramagnetic beads (made using the protocol by Rohland *et al.* (Rohland and Reich, 2012)). In brief, 110 µl (2.2X) warmed homemade paramagnetic beads were added to RNA sample(s). Sample(s) were incubated for 15 minutes on ice. Sample(s) were placed on a magnetic rack and the solution was allowed to clear for 5 minutes. The supernatant was discarded. 200µl 80% ethanol was washed over the beads and allowed to incubate at room temperature on the magnetic rack for 30 seconds. Supernatant was discarded. The previous step was repeated once more. Beads were air-dried for 2-5 minutes. The sample(s) were removed from the magnetic rack. Sample bead(s) were eluted in 7µl nuclease free water and incubated at room temperature for 2 minutes. Sample(s) were placed on the magnetic rack and

incubated for 2 minutes at room temperature. 5µl of supernatant was removed and used for the next steps.

2.13.4.2 RNA-seq library preparation (NEBNext® Ultra™ II Directional RNA library prep kit for Illumina®, E7760)

RNA-seq library preparation was conducted using NEBNext® Ultra™ II Directional RNA library prep kit for Illumina®. In brief, RNA fragmentation and priming reaction master mix was made by combining the following:

RNA fragmentation and priming Master Mix	Volume (µl)
NEBNext First Strand Synthesis Reaction Buffer (5X)	4
Random primers	1

5µl RNA fragmentation and priming reaction master mix was added to RNA sample(s) obtained from 2.13.4.1 and mixed thoroughly by pipetting up and down. Samples were incubated at 94°C for 15 minutes.

First strand synthesis reaction master mix was made by combining the following:

First strand synthesis Master Mix	Volume (µl)
NEBNext strand specific specificity reagent	8
NEBNext first strand synthesis enzyme mix	2

10µl first strand synthesis reaction master mix was added to sample(s) and mixed thoroughly by pipetting up and down. Samples were incubated with the following programme with heated lid $\geq 80^{\circ}\text{C}$:

Temperature ($^{\circ}\text{C}$)	Time
25	10 minutes
42	15 minutes
70	15 minutes
4	Hold

Second strand synthesis reaction master mix was made by combining the following:

First strand synthesis Master Mix	Volume (μl)
NEBNext second strand synthesis reaction buffer with dUTP	8
NEBNext second strand synthesis enzyme mix	4
Nuclease free water	48

60µl second strand synthesis reaction master mix was added to sample(s) and mixed thoroughly by pipetting up and down. Samples were incubated at 16°C for 60 minutes with heated lid $\leq 40^{\circ}\text{C}$.

Sample(s) were briefly centrifuged. Sample(s) were cleaned with homemade paramagnetic beads. In brief, 144 μl (1.8X) warmed homemade paramagnetic beads were added to RNA sample(s). Sample(s) were incubated for 5 minutes on ice. Sample(s) were placed on a magnetic rack and the solution was allowed to clear for 3 minutes. The supernatant was discarded. 200µl 80% ethanol was washed over the beads and allowed to incubate at room temperature on the magnetic rack for 30 seconds. Supernatant was discarded. The previous step was repeated once more. Beads were air-dried for 2-5

minutes. The sample(s) were removed from the magnetic rack. Sample bead(s) were eluted in 53µl 0.1X TE and incubated at room temperature for 2 minutes. Sample(s) were placed on the magnetic rack and incubated for 2 minutes at room temperature. 50µl of supernatant was removed and used for the next steps.

End prep reaction master mix was made by combining the following:

End prep reaction Master Mix	Volume (µl)
NEBNext Ultra II End Prep Reaction Buffer	7
NEBNext ultra II End Prep Enzyme Mix	3

10µl end prep reaction master mix was added to sample(s) and mixed thoroughly by pipetting up and down. Samples were incubated with the following programme with heated lid $\geq 75^{\circ}\text{C}$:

Temperature ($^{\circ}\text{C}$)	Time
20	30 minutes
65	30 minutes
4	Hold

NEBNext adaptors were diluted 5-fold in adaptor dilution buffer. Ligation reaction master mix was made by combining the following:

Ligation reaction Master Mix	Volume (μ l)
Diluted adaptor	2.5
NEBNext Ligation enhancer	1
NEBNext ultra II ligation master mix	30

33.5 μ l ligation reaction master mix was added to sample(s) and mixed thoroughly by pipetting up and down. Samples were incubated 20°C for 15 minutes. 3 μ l USER enzyme was added to RNA sample(s), mixed thoroughly by pipetting and incubated at 37°C for 15 minutes.

Sample(s) were briefly centrifuged. Sample(s) were cleaned with homemade paramagnetic beads. In brief, 87 μ l (0.9 X) warmed homemade beads were added to RNA sample(s). Sample(s) were incubated for 10 minutes on ice. Sample(s) were placed on a magnetic rack and the solution was allowed to clear for 3 minutes. The supernatant was discarded. 200 μ l 80% ethanol was washed over the beads and allowed to incubate at room temperature on the magnetic rack for 30 seconds. Supernatant was discarded. The previous step was repeated once more. Beads were air-dried for 2-5 minutes. The sample(s) were removed from the magnetic rack. Sample bead(s) were eluted in 17 μ l 0.1X TE and incubated at room temperature for 2 minutes. Sample(s) were placed on the magnetic rack and incubated for 2 minutes at room temperature. 15 μ l of supernatant was removed and used for the next steps.

Before PCR enrichment, optimal PCR cycle numbers for each sample was determined by qPCR. PCR enrichment master mix was made by combining the following:

PCR enrichment Master Mix	Volume (μl)
NEBNext Ultra II Q5 master mix	25
Universal PCR primer	5
Index (X) primer (NEBNext® Multiplex Oligos for Illumina®, Index 1 E7335 and Index 2 E7500)	5

35μl PCR reaction master mix was to adaptor ligated DNA and mixed thoroughly by pipetting up and down. Sample(s) were incubated with the following programme:

Temperature (° C)	Time	Cycles
98	30 seconds	1
98	10 seconds	5
65	75 seconds	
4	Hold	Hold

qPCR to quantify optimal PCR enhancer cycles was conducted using NEBNext Library Quant kit. Master and primer mix were prepared by adding 100 μl NEBNext Library Quant Primer Mix to NEBNext Library Quant Master Mix. 20 μl of ROX was added to the NEBNext Library Quant Master Mix. NEBNext Library Quant Dilution Buffer (10X) was diluted 1:10 with nuclease-free water. Three library dilutions were prepared by diluting library sample(s) in 1X NEBNext Library Quant Dilution Buffer; 1:1000, 1:10,000 and 1:100,000. In a white 96-well Multiplate® PCR Plate™ (Biorad, MLL9651) plate 16 μl NEBNext Library Quant Master Mix was added to 4 μl diluted library sample or non-template control (NTC). qPCR Plate in a Real-time Thermal Cycler was run on a CFX96™ Real Time System (Biorad) with the following programme:

Temperature (° C)	Time	Cycles
95	1 minute	1
95	15 seconds	20
63	45 seconds	

Optimal cycles for PCR enhancing stage was determined. The PCR enhancing step was continued with the following programme with the heated lid set to 105°C:

Temperature (° C)	Time	Cycles
98	10 seconds	Determined by qPCR
65	75 seconds	
65	5 minutes	1
4	Hold	Hold

Sample(s) were briefly centrifuged. Sample(s) were cleaned with homemade paramagnetic beads. In brief, 45µl (0.9 X) warmed homemade beads were added to RNA sample(s). Sample(s) were incubated for 5 minutes on ice. Sample(s) were placed on a magnetic rack and the solution was allowed to clear for 3 minutes. The supernatant was discarded. 200µl 80% ethanol was washed over the beads and allowed to incubate at room temperature on the magnetic rack for 30 seconds. Supernatant was discarded. The previous step was repeated once more. Beads were air-dried for 2-5 minutes. The sample(s) were removed from the magnetic rack. Sample bead(s) were eluted in 23µl 0.1X TE and incubated at room temperature for 2 minutes. Sample(s) were placed on the magnetic rack and incubated for 2 minutes at room temperature. 20µl of supernatant was removed and store at -20°C.

Quantification of final library by PCR was conducted using NEBNext Library Quant kit. Master and primer mix were prepared by adding 100 µl NEBNext Library Quant Primer Mix to NEBNext Library Quant Master Mix. 20 µl of ROX was added to the NEBNext Library Quant Master Mix. NEBNext Library Quant Dilution Buffer (10X) was diluted 1:10 with nuclease-free water. Three library dilutions were prepared by diluting library sample(s) in 1X NEBNext Library Quant Dilution Buffer; 1:1000, 1:10,000 and 1:100,000. A range of standards was included. The range of standard concentrations in the NEBNext Library Quant Kit is 10–0.01pM. In a white 96-well Multiplate® PCR Plate™ (Biorad, MLL9651) plate 16 µl NEBNext Library Quant Master Mix was added to 4 µl diluted library sample, standard or non-template control (NTC). qPCR Plate in a Real-time Thermal Cycler was run on a CFX96™ Real Time System (Biorad) with the following programme:

Temperature (° C)	Time	Cycles
95	1 minute	1
95	15 seconds	35
63	45 seconds	

Quantification of final library by PCR was conducted using Qubit (as detailed in 2.8.3). The libraries were pooled to a final concentration of 2.5nM. Pooled library concentration was quantified by Qubit (as detailed in 2.8.3). Quality of library pool was assessed using Aligent 2200 Tape Station system. In brief, High Sensitivity D1000 ScreenTape Sample buffer (5190-6504, Aligent Technologies), High Sensitivity D1000 ScreenTape ladder (5067-5587, Aligent Technologies) and High Sensitivity D1000 ScreenTape (5067-5584, Aligent Technologies) were allowed to reach room temperature. Sample(s) were prepared in optical tubes (Aligent technologies, #401428) by adding 2µl DNA to 2µl High Sensitivity D1000 ScreenTape Sample buffer. Sample(s) were centrifuged briefly and vortexed at 2000rpm for 60 seconds. Sample(s) were centrifuged briefly. Sample(s) were placed in an Aligent 2200

Tape Station system. Loading tips (Aligent Technologies, #5067-5153) were placed into the tape station and High Sensitivity D1000 ScreenTape was placed into the tape station and run.

2.13.5 RNA Sequencing

HiSeq 100bp paired end RNA sequencing was carried out by the Medical Research Council (MRC) Genomic Facility, ICTEM, Imperial College London, Hammersmith Hospital, London.

2.14 siRNA for Acss2 - MISSION® esiRNA Transfection Protocol

Cell lines were seeded at 1×10^5 in normal DMEM/F12 media into 6-well plates for 3 days and incubated at 37°C and 5% CO₂. Following 4-hour incubation, media was replaced with DMEM/F12 media or DMEM/F12 media supplemented with 5mM acetate. esiRNA was added to cells. In brief, 12pmol esiRNA for Acss2 (esiRNA1, Sigma-Aldrich, EHU135341) was added to 10µl transfection reagent (Sigma-Aldrich, S1452) and 200µl serum free media to a final concentration of 10nM per well. The siRNA transfection mix was allowed to incubate for 10 minutes before adding to appropriate wells by drop-wise pipetting. The following controls were included; esiRNA for EGFP (esiRNA1, negative control, Sigma-Aldrich EHUEGFP) and no siRNA (transfection reagent only). Following 24-hour incubation cells were treated with 5mM metformin and incubated for 3 days at 37°C and 5% CO₂. Protein from esiRNA sample(s) were acid extracted as in 2.5.4, quantified as in 2.5.5 and analysed by western blot as in 2.6.

2.15 H3K27ac tracking

MCF12A non-cancerous breast epithelial cell line was seeded in normal DMEM/F12 media in 6cm dishes for the following time points:

Time Point (hours)	Seeding
24	1×10^6
48	5×10^5
72	2.5×10^5
96	1.25×10^5
120	6.25×10^4
144	3.125×10^4

Cells were seeded as above and incubated at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 2.5mM and 5mM metformin. Cells were incubated at 37°C and 5% CO₂ for 24-hours to 144-hours. At each time point protein sample(s) were acid extracted as in 2.5.4, quantified as in 2.5.5 and analysed by western blot as in 2.6.

2.16 Hydrophilic Interaction Liquid Chromatography Mass Spectrometry (HILIC-MS) measurement of S-adenosyl-methionine

2.16.1 Cell line seeding for metabolite extraction

Cells were seeded at 2×10^5 in 6 well plates in technical triplicates for cell counting and technical triplicates for extraction. Media only blanks in triplicate were included. Cells were incubated for 24 hours at 37°C and 5% CO₂. Following 24-hour incubation, cells were treated with 0mM, 2.5mM and 5mM metformin. Cells were incubated at 37°C and 5% CO₂ for 3 days.

2.16.2 Preparation of reagents for methanol quenching extraction

Internal samples were prepared as follows and stored at 4°C:

	µl from Stock (2mg/mL)	Concentration (µg/mL)
SAH d4	10	20
SAM 13C5	10	20
H2O	980	

Ringer's buffer (Merck Millipore, 1155250001) was prepared by adding 1 tablet of Ringer's buffer to 500ml distilled sterile water. Buffer was filtered sterilized through a 0.22µm filter (Sartorius™ Minisart™ NML Syringe Filters, 16534) and stored at 4°C. 80% liquid chromatography mass spectrometry (LC-MS) standard methanol was prepared in LC-MS grade water.

2.16.3 Methanol quenching extraction

To determine internal standard volume, cells were counted in triplicate and averaged. Cell counting wells were trypsinised as in 2.2.4. 20µl cell suspension was added to 20µl trypan blue stain 0.4% (T10282, Life Technologies) in a fresh, labelled 1.5ml centrifuge eppendorf tube. 20µl of cell suspension/trypan blue stain 0.4% solution was added to a Countess™ cell counting chamber slide (C10283, Invitrogen). Cells were counted using the Countess™ automated cell counter (Invitrogen). For extraction wells, plates were placed on ice and media was removed. Wells were washed twice with 1ml Ringer's buffer. 1mL 80% LC-MS grade methanol was added to each well and internal standard was added, 10µl of internal standard mix per 1,000,000 cells. Plates were incubated on dry ice for 20 minutes. Following incubation plates were placed on ice or a cooled surface. Cells were scraped and supernatant was transferred to 1.5ml eppendorf centrifuge tube. 0.5mL 100% LC-MS grade methanol was added to wells and scraped a second time. Supernatant(s) were pooled with 1ml supernatant. Pooled sample(s) were vortexed vigorously for 10 seconds. Sample(s) were centrifuged

at 16,200g at 4°C. Supernatant(s) were collected in LC-MS grade vial and stored at 4°C. Sample(s) were dried via nitrogen flow. Sample(s) were stored at -20°C to -40°C.

2.16.4 Hydrophilic interaction Liquid chromatography mass spectrometry

HILIC-MS was conducted by Eirini Kouloura, Hector Keun lab, Imperial College London, UK.

2.17 Statistical Analysis

2.17.1 Statistical analysis of cell viability assays

IC₅₀ was determined using Prism GraphPad software, version 7 for Mac OS X, GraphPad Software, San Diego, California USA, www.graphpad.com. In brief, data was transformed using Log(X) and analysed by non-linear regression (curve fit) sigmoidal, 4PL, where X is log(concentration).

2.17.2 Statistical analysis of western blot

Western blot analysis was conducted in three biological triplicates. Analysis of western blot data was conducted by measuring relative absorbance of bands using ImageJ (Image Processing and Analysis in Java) programme. Band(s) were measured, and corresponding background measurement(s) were taken. Measurements(s), both sample and background, were inverted by taking the ImageJ measurement(s) from 225. Inverted background measurement(s) were taken from inverted sample measurement(s) to give net absorbance value(s) of sample(s). Following measurement(s) of target protein and loading control, a ratio of target protein:loading control was obtained and plotted as relative abundance (A/U). One sample paired student's *t*-tests were conducted to determine differences between treatment conditions, where $p < 0.05$ was accepted as statistically significant.

Target Protein	Loading Control
AMPK α	β -actin
pAMPK α (Thr172)(40H9)	β -actin
Cyclin D1	β -actin
H3K27ac	H3 (D1H2)
H3K27me3	H3 (D1H2)

2.17.3 Statistical analysis of RNA-Sequencing data

Raw data was pre-processed using bbdutk and bbmerge to remove adaptor sequences. RNA-sequencing data was aligned using two methods (i) Kallisto pseudoalignment (Bray *et al.*, 2016) and (ii) TopHat2 (Kim *et al.*, 2013). Data aligned by Kallisto was analysed by Sleuth analysis software (Pimentel *et al.*, 2017) to obtain differential expression (b) values. Kallisto quantified abundances of transcripts using a combination hashing of k-mers and transcriptome de Bruijn graph. Sleuth used bootstrapping and response error linear modelling to conduct differential analysis of Kallisto estimates. This allowed biological variance to be separated from intentional variance.

Analysis of RNA sequencing data by TopHat 2 alignment and featureCounts was conducted by Daisy Stoner, MRes. TopHat 2 aligned each transcript to the target genome, Genome Reference Consortium Human Build 37 (GRCh37). This allowed TopHat2 to detect splice sites for introns and allowed genomic features, such as pseudogenes and chromosomal features such as indels, inversion, deletions and translocations, to be identified. This increased accuracy of the annotations and the sensitivity of the annotations. featureCounts used chromosome hashing and feature blocking to count reads associated with genomic features.

2.17.4 Statistical analysis of qPCR gene expression data

Ct value(s), representing the level of expression for the chosen gene probe based on the fluorescence and cycle number, was obtained following qPCR of target genes. Ct value(s) were also obtained for endogenous controls. The RQ value, representing log-fold change, was calculated using the following calculation:

$$\Delta Ct = Ct_{\text{target gene}} - Ct_{\text{endogenous control}}$$

$$\Delta\Delta Ct = \Delta Ct_{\text{sample}} - \Delta Ct_{\text{calibrator sample}}$$

$$RQ = 2^{-(\Delta\Delta Ct)}$$

One sample paired student's *t*-tests were conducted to determine differences between treatment conditions, where $p < 0.05$ was accepted as statistically significant.

2.17.5 Statistical analysis of Illumina MethylationEPIC Array

Raw beta values generated from Illumina MethylationEPIC array were normalised by functional normalization. Probes were filtered by detection p -value < 0.05 .

Model 1 – a mixed effects model on combined MCF12A and MCF10A non-cancerous breast epithelial cell line M-values generated from beta values ($\log_2(\text{betas} / (1 - \text{betas}))$) adjusting for dose, batch, position, cell type and replicate. Significant CpG sites were determined by $p < 0.05$ and $\text{FDR} < 0.05$.

Model 2 – a general linear model on individual MCF12A and MCF10A non-cancerous breast epithelial cell line beta values adjusting for dose and batch. Interesting CpG sites were chosen based on $> 10\%$ difference in methylation in either cell line.

Model 3 – average difference in beta values between treated and untreated within batches was determined. Interesting probes were chosen based on $> 10\%$ difference in methylation in the same

direction across batches. Conducted in MCF12A and MCF10A non-cancerous breast epithelial cell lines separately.

2.17.6 Statistical analysis of pyrosequencing data

Methylation percentages obtained from pyrosequencing for target CpG sites were averaged over 4 replicates. One sample paired student's *t*-tests were conducted to determine differences between treatment conditions, where $p < 0.05$ was accepted as statistically significant.

2.17.7 Statistical analysis of liquid chromatography mass spectrometry data

Analysis of LC-MS data was conducted by Eirini Kouloura, Hector Keun lab, Imperial College London, UK.

2.17.8 Graphing tools

Graphs were made using Prism GraphPad software (GraphPad, 2019), R-studio (RStudio, 2019) and ImageJ (Schneider, Rasband and Eliceiri, 2012).

CHAPTER 3

3. Metformin alters gene expression in normal breast epithelial cell lines

3.1 Introduction

Gene expression changes could be associated with modified risk of developing certain cancers. For example, the most common gene expression changes associated with breast cancer risk are *BRCA1/2* mutations. *BRCA1/2* genes encode for enzymes involved in double strand break DNA repair by homologous recombination (Powell and Kachnic, 2003). Global mutations induced by genetic instability contributes to cancer progression. In fact, patients with inherited *BRCA1/2* mutations are 71-84% more likely to develop cancer compared to patients with no hereditary predisposition (Levy-Lahad and Friedman, 2007).

Non-inherited mutations associated with breast cancer risk have also been studied. Wu *et al.* conducted a transcriptome-wide association study (TWAS) analysis from data obtained from the Breast Cancer Association Consortium (BCAC) (Wu *et al.*, 2018). This study identified 38 gene changes significantly associated with breast cancer risk. Smaller studies have additionally shown breast cancer risk genes (Ghoussaini *et al.*, 2014; Lawrenson *et al.*, 2016) (appendix table 3.1)

Genetic wide association studies (GWAS) have been conducted on diabetic patients taking metformin (Pawlyk *et al.*, 2014; Niu *et al.*, 2016). These studies have found a number of gene variants associated with metformin and metformin response. For example, gene variants in the cellular transporter molecules of metformin have been shown to modify patient responses to metformin. Polymorphisms in *SLC22A1* (OCT1) could lead to reduced uptake of metformin (Mato *et al.*, 2018). Additionally, single nucleotide polymorphisms (SNPs) in *SLC47A1* (MATE1) have been associated with patient response (Gong *et al.*, 2012; Florez, 2017). GWAS analyses have shown *ATM* mutations were associated with metformin treatment. *ATM* is involved in DNA repair and could play a role in patient response to metformin (GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group *et al.*, 2011). *SLC2A2*

(*GLUT2*) has also been shown to be modified in metformin treated samples, which may have an impact on metformin's action on inhibiting gluconeogenesis and reinstatement of glucose-associated pathways (Zhou *et al.*, 2016).

We hypothesize that metformin may be affecting gene expression in normal breast epithelial cells and metformin treatment could potentially be linked to known breast cancer risk genes.

3.2 Results

3.2.1 Non-cancerous breast epithelial cells are sensitive to metformin

The majority of breast cancers originate from epithelial cells within the breast tissue (Polyak, 2007). To identify potential mechanisms metformin uses to reduce breast cancer risk we chose to investigate models of normal breast epithelial cells. We chose two non-cancerous breast epithelial cell line models to study, MCF12A and MCF10A. Findings were validated using matched normal primary breast epithelial and myoepithelial cells derived from patients undergoing reduction mammoplasty and obtained from Breast Cancer Now.

To test whether normal breast epithelial cells were sensitive to metformin, MCF12A and MCF10A non-cancerous breast epithelial cell lines were treated in a range of metformin concentrations (0mM – 25mM) for 3 and 7-days. Cell viability was analysed by crystal violet assay and MTS assay. By crystal violet assay MCF12A and MCF10A cell lines were both sensitive to metformin, however MCF12A cell line appeared to be more sensitive compared to MCF10A (for example 3-day IC₅₀ 3.8mM and 2.8mM, respectively). Analysis by MTS assay showed MCF12A retained cell viability at 3-days which was lost at 7 days. MCF10A showed sensitivity at both timepoints (figure 3.1). Comparison of 3-day treated crystal violet data to a similar experiment conducted in the MCF-7 breast cancer cell line showed IC₅₀ 1.5mM, suggesting that non-cancerous breast epithelial cell lines were not as sensitive to metformin as cancer cell lines (appendix figure 3.1).

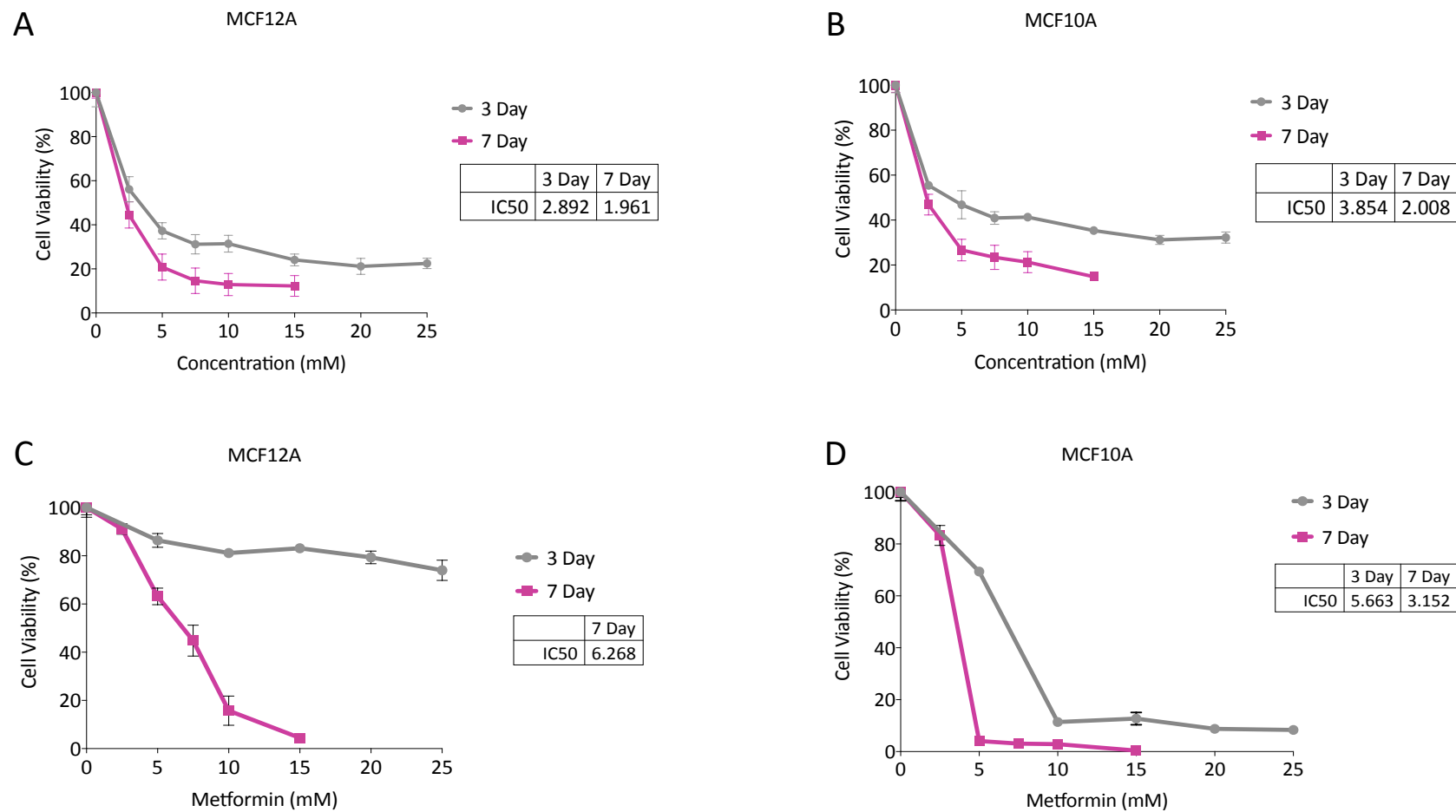


Figure 3.1 Cell viability assay of MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 3- and 7-days with a range of metformin concentrations (0mM – 25mM). (A) MCF12A crystal violet assay. IC50 2.8mM and 1.9mM for 3-day and 7-day treatment, respectively. (B) MCF10A crystal

violet assay. IC50 3.8mM and 2mM for 3- and 7-day treatment, respectively. (C) MCF12A MTS assay. IC50 undetermined and 6mM for 3- and 7-day treatment, respectively. (D) MCF10A MTS assay. IC50 5.6mM and 3.1mM for 3- and 7-day treatment, respectively.

To validate these findings, we conducted crystal violet assay on matched normal primary breast epithelial and myoepithelial cells treated at a range of metformin (0mM – 20mM) for 7-days. Epithelial cells showed significant sensitivity to metformin (passage 2) however there appeared to be variability in response in each patient, for example patient 3002 showed the highest sensitivity whereas patient 1957 showed an IC₅₀ almost twice as high as 3002. Crystal violet assay of passage 2 (P2) myoepithelial cells showed similar sensitivity as epithelial cells and again inter-patient variability was evident. We were able to obtain an additional passage of myoepithelial cells obtained from each patient. We conducted crystal violet analysis of passage 3 (P3) myoepithelial cells to assess whether passage greatly affected our previously shown results. P3 myoepithelial cells showed significant resistance to metformin, particularly when compared to P2 myoepithelial cells. Initial analysis of untreated P2 and P3 myoepithelial cellsshowed P3 myoepithelial cells potentially grew slower than P2 myoepithelial cells. Further analysis of this difference could highlight a limiting factor of metformin treatment (figure 3.2).

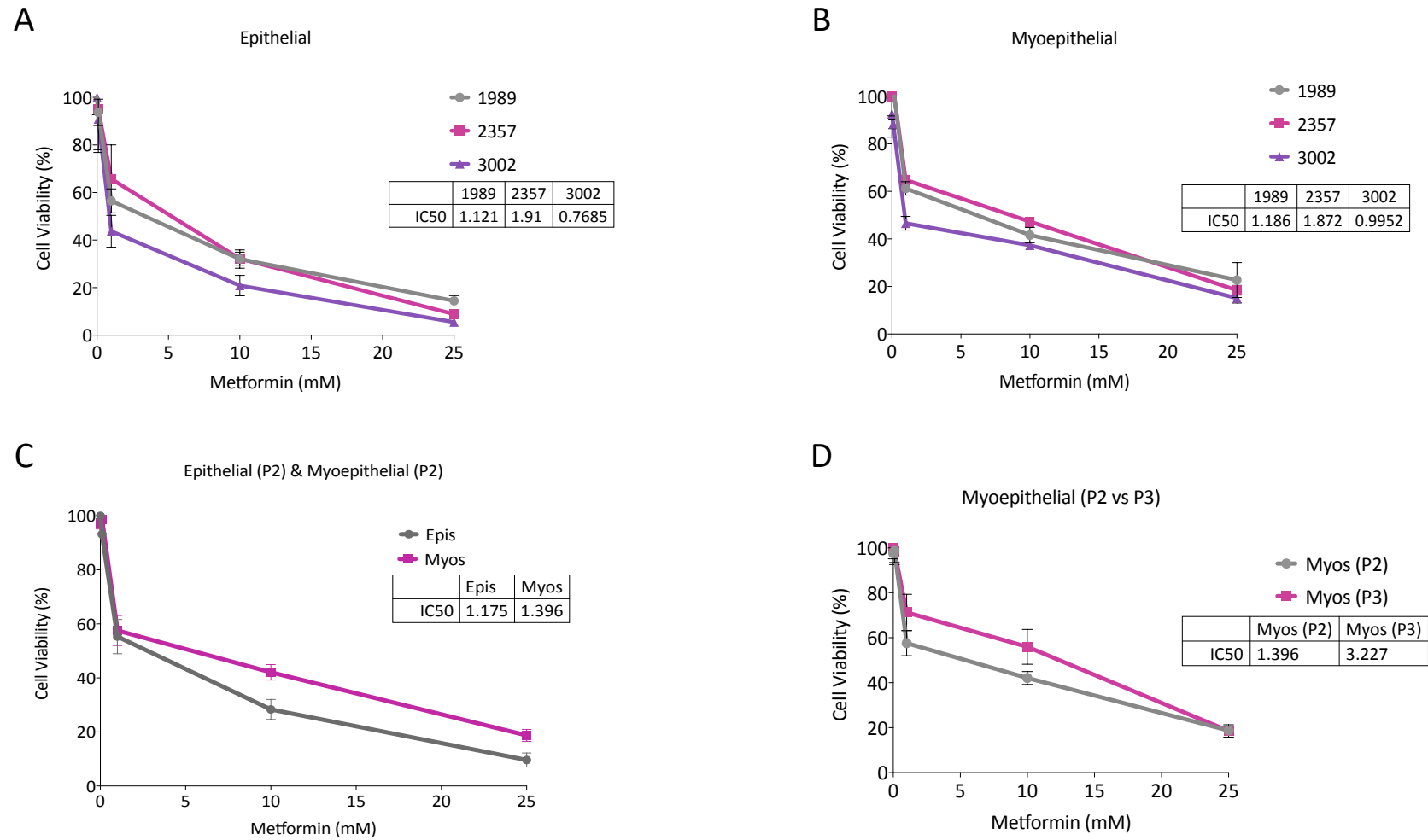


Figure 3.2 Crystal violet cell viability assay of three patient normal matched breast epithelial and myoepithelial primary cells treated for 7 days with a range of metformin concentrations (0mM – 25mM). (A) Epithelial primary cells (P2) from three patients. IC50 1.1mM, 1.9mM and 0.7mM for patients 1989, 2357

and 3002, respectively (n=3). (B) Myoepithelial primary cells (P2) from three patients. IC50 1.1mM, 1.8mM and 0.9mM for patients 1989, 2357 and 3002, respectively (n=3). (C) Total epithelial primary cells (P2) and total myoepithelial primary cells (P2). IC50 1.1mM and 1.3mM respectively (n = 9). (D) Comparison of P2 and P3 myoepithelial primary cells. IC50 1.2mM and 3.2mM respectively (n=9).

Metformin has been shown to activate specific pathways in both diabetic and breast cancer models. The activation of AMPK by phosphorylation is a downstream effect of metformin's inhibitory effect on complex I of the mitochondrial electron transport chain. Activation of this pathway by metformin has been shown in cell line models of diabetes and breast cancer and was the most commonly used pathway to indicate metformin's interaction with these models. Therefore, we chose to study the activation of AMPK in our non-cancerous breast epithelial models.

We conducted western blot analysis of p-AMPK and total AMPK with and without metformin in our normal breast epithelial cell line models. MCF12A and MCF10A cell lines were treated for 3-days with 0mM, 2.5mM and 5mM metformin. The treatment concentrations were chosen to encompass the respective IC50s for MCF12A (2.9 mM) and MCF10A (3.8 mM) at 3-days treatment. Western blot analysis showed a trend towards increased p-AMPK activation with a decrease in total AMPK levels (figure 3.3), however this was not significant.

These results showed that models of non-cancerous breast epithelial cells appeared to be sensitive to metformin however the known pathways metformin was interacting with may not be activated to the same degree as diabetic liver or breast cancer models.

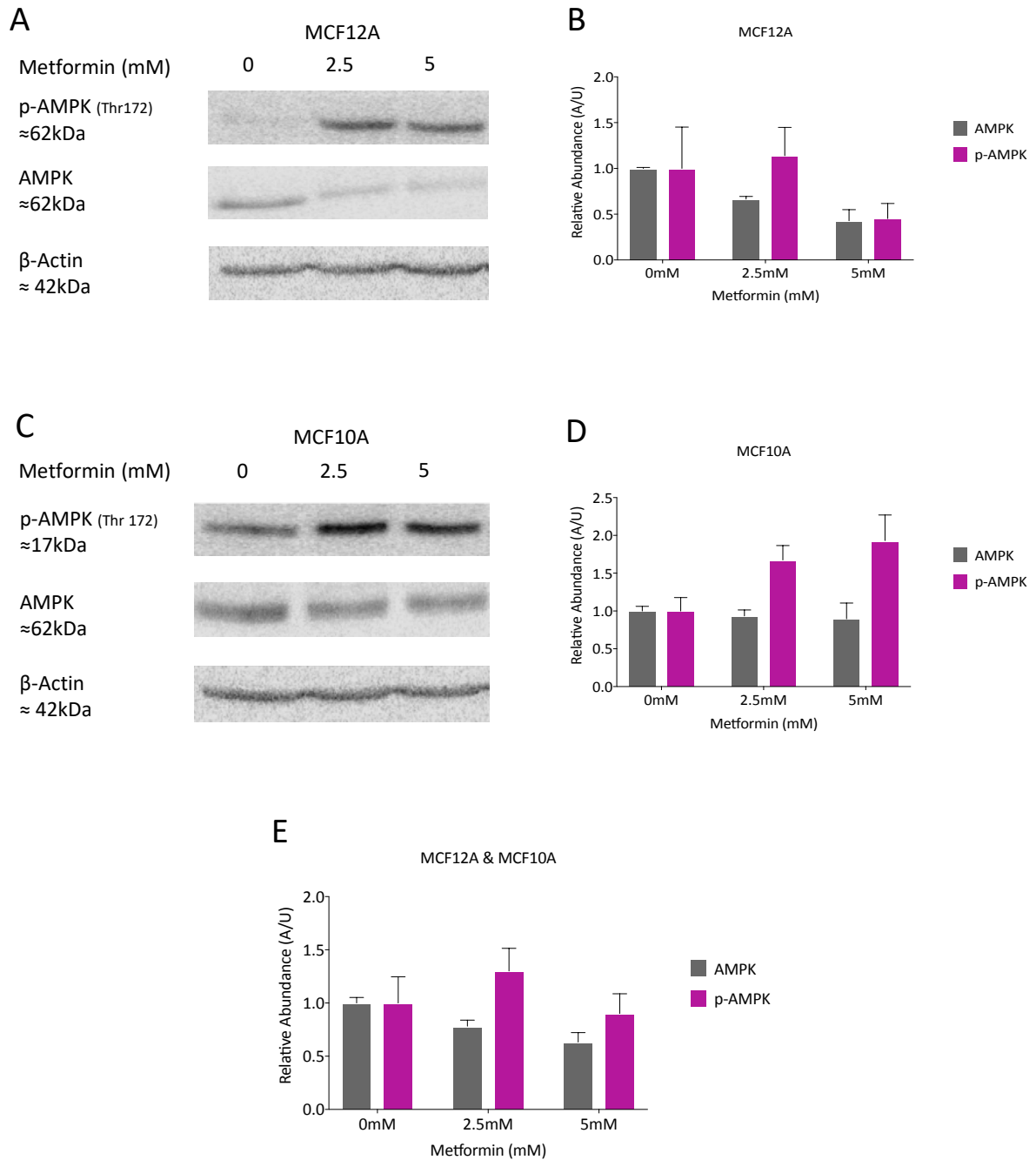


Figure 3.3 Western blot analysis of p-AMPK and AMPK in non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. Relative abundance measured as ratio to beta actin. (A-B) MCF12A (n=3) (C-D) MCF10A (n=3) (F) Relative abundance measurement of p-AMPK and AMPK in MCF12A and MCF10A averaged over six biological replicates (n = 6).

3.2.2 Metformin alters gene expression in non-cancerous breast epithelial cells

Our results have shown that metformin may not be activating the same pathways it would in breast cancer. To identify new pathways which metformin may be interacting with in normal breast epithelial cells, we chose to study gene expression changes in our non-cancerous breast epithelial models whose expression may be affected by metformin.

RNA sequencing data was obtained from MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin. Multiple statistical methods could be used to analyse this data. In order to quantify gene expression changes and more in depth changes such as exon splicing, the data was aligned using two techniques; Kallisto (Bray *et al.*, 2016), a “pseudoalignment” technique which quantified abundances of transcripts, and TopHat2 (Kim *et al.*, 2013), which aligned all sequences to the chosen genome, GRCh37.

Principle Component Analysis (PCA) of Kallisto aligned data on transcripts per million (TPM) showed that majority of variation in both MCF12A and MCF10A data sets came from cell type (figure 3.4). Statistical analysis of Kallisto data was conducted using Sleuth (Pimentel *et al.*, 2017). In MCF12A 361 genes showed significant changes in response to metformin in both doses, 2.5mM and 5mM (FDR<0.05). However only 3 genes showed significance (FDR<0.05) in MCF10A in response to metformin at both doses. Data shown in figure 3.5 showed statistically significant genes (FDR<0.05) which changed in expression under metformin treatment conditions in MCF12A cell line. Significant genes in MCF10A data set were highlighted in pink. In MCF12A data set while more significant genes were found in 5mM treated samples (768 genes) vs 2.5mM samples (550 genes) there did not appear to be a dose dependent effect.

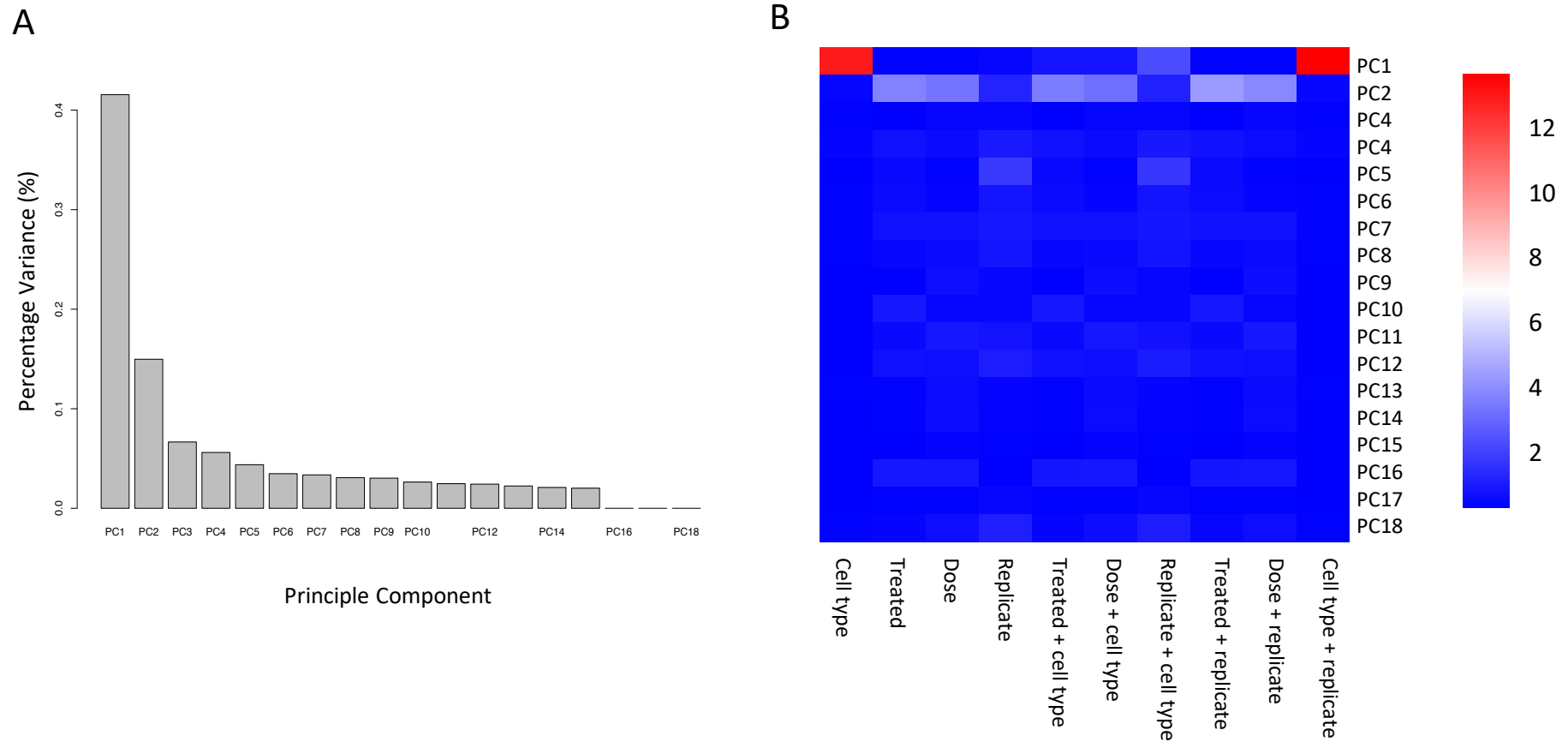


Figure 3.4 Principle Component Analysis (PCA) of MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin. RNA sequencing data analyzed by Kallisto. (A) Percentage of variance contributed by each principle component. (B) Heatmap of $-\log_{10} p$ values of univariate and multivariate linear models.

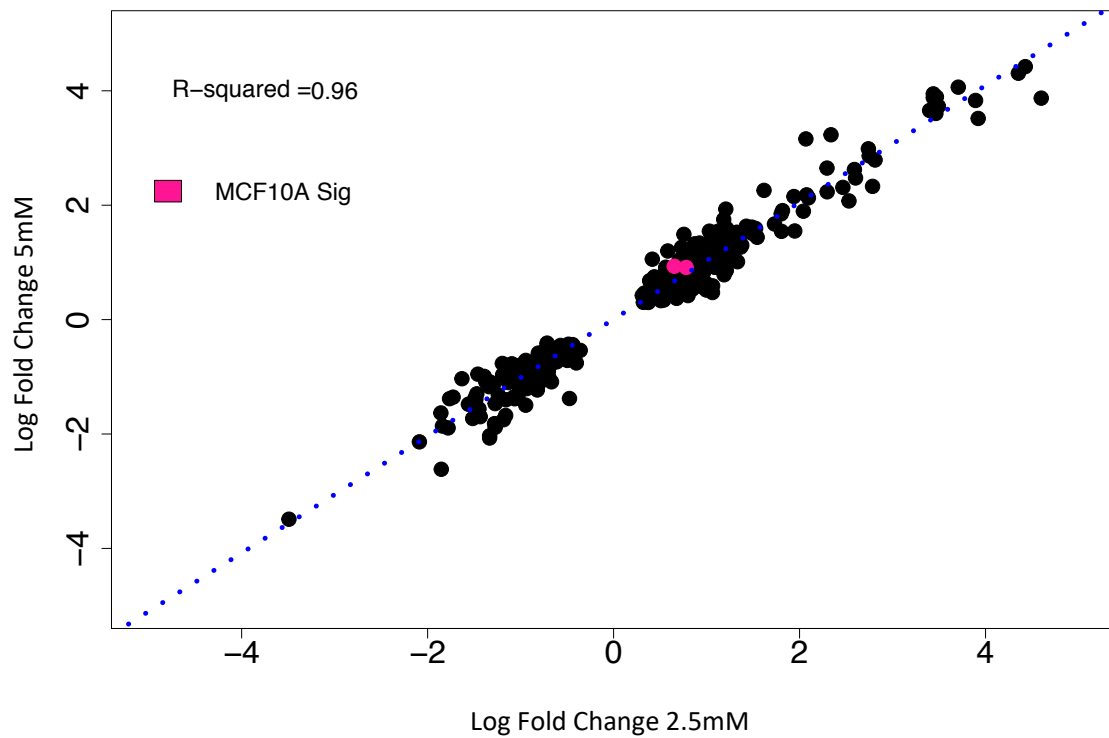


Figure 3.5 Correlation plot of genes showing significant Log Fold change of MCF12A cell lines treated with 2.5mM vs 5mM (FDR<0.05, Pearson correlation $r=0.96$). Pink points indicate genes showing significant Log Fold change of MCF10A cell lines treated with 2.5mM and 5mm (FDR<0.05) which overlap with MCF12A significant genes.

PCA analysis of TopHat aligned data on reads per kilo base million (RPKM) found the majority of variability in this data was from cell type. However, in contrast to Kallisto aligned data dose and treatment condition in both MCF12A and MCF10A cell lines contributed to a significant proportion of the variability and was seen after adjusting for cell type (figure 3.6). This suggested that gene expression changes seen within this data set could be attributed to metformin treatment. Statistical analysis of TopHat2 data was conducted using featureCounts function in the Rsubread package by Daisy Stoner, MRes. FeatureCounts assigned mapped sequencing reads to genomic features and therefore provided a more comprehensive list of significant genes accounting for features such as exon splice sites. Figure 3.7 showed a heatmap of 1368 and 2156 significant genes found to be altered by 2.5mM and 5mM metformin, respectively, in MCF12A data set compared to 130 and 65 significant genes found to be altered by 2.5mM and 5mM metformin, respectively, in MCF10A data set.

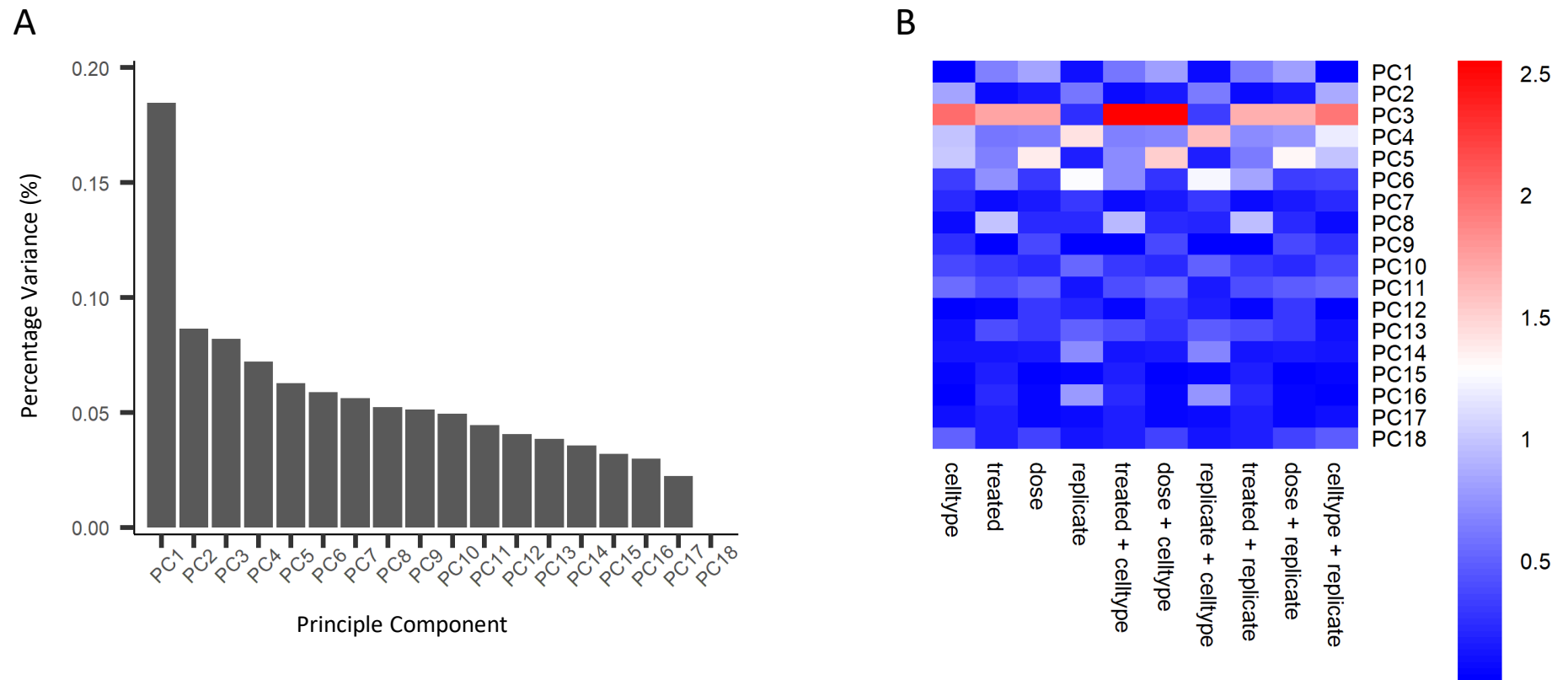


Figure 3.6 Principle Component Analysis (PCA) of MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin. RNA sequencing data analyzed by featureCounts. (A) Percentage of variance contributed by each principle component. (B) Heatmap of $-\log_{10}$ p values of univariate and multivariate linear models. Figure produced by Daisy Stoner, MRes.

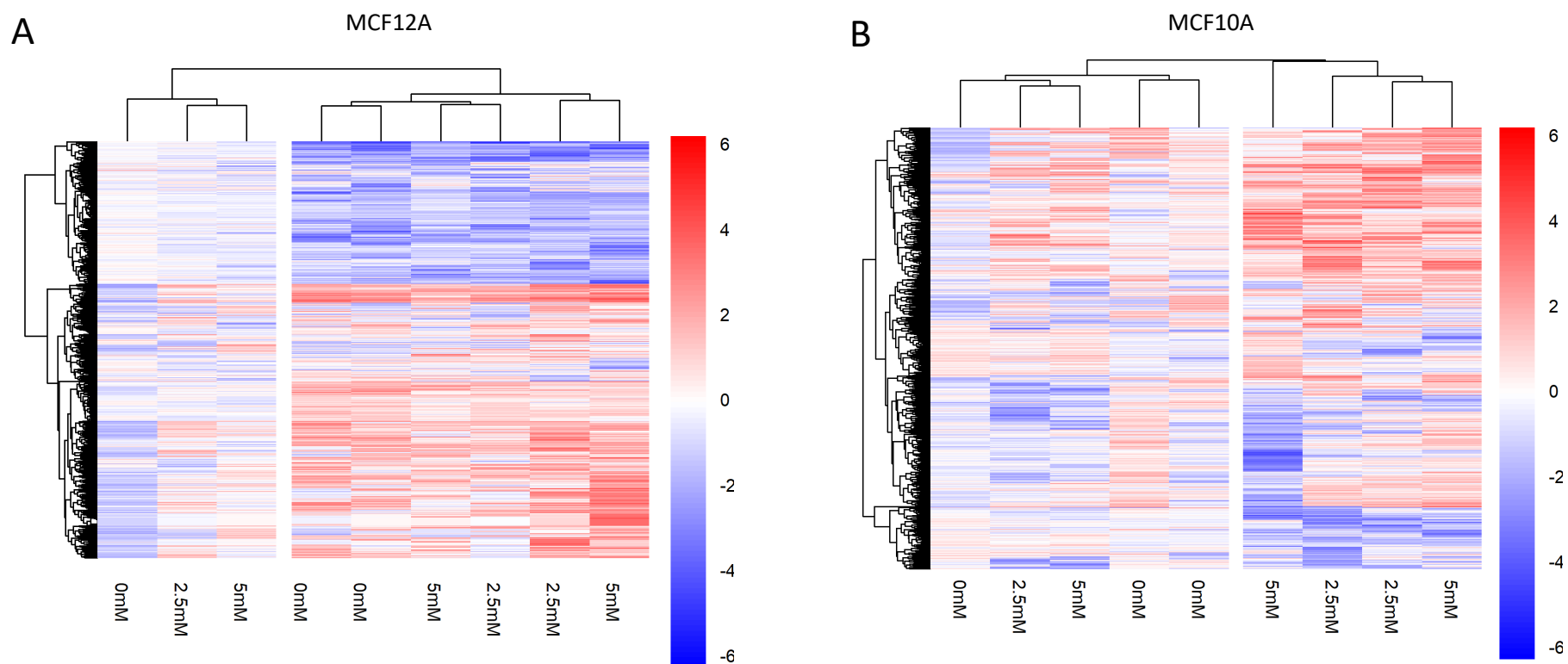


Figure 3.7 Heatmaps showing logFold change in gene expression of significant genes identified by featureCounts analysis (Pearson correlation clustering). (A) MCF12A non-cancerous breast epithelial cell line. (B) MCF10A non-cancerous breast epithelial cell line. Figure produced by Daisy Stoner, MRes.

Two techniques were chosen to analyse RNA sequencing data to ensure the validity of the data. Comparison of the two data sets analysed showed good correlation between the two data sets in all treatment conditions (figure 3.8). Analysis of the distribution of the data sets showed a more positive distribution in featureCounts analysed data compared to Kallisto. Additionally, comparison of the two techniques showed overlap of significant genes in MCF12A data set (figure 3.9). However, there was no overlap between the techniques in MCF10A data set. The lack of overlap may have been partially due to the methods used to align transcripts to the target genome (k-mer hashing vs chromosome hashing). Variability in results could additionally be due to the presence of exon splice sites in TopHat2 aligned data but not in Kallisto aligned data. This may potentially have changed the significant genes identified.

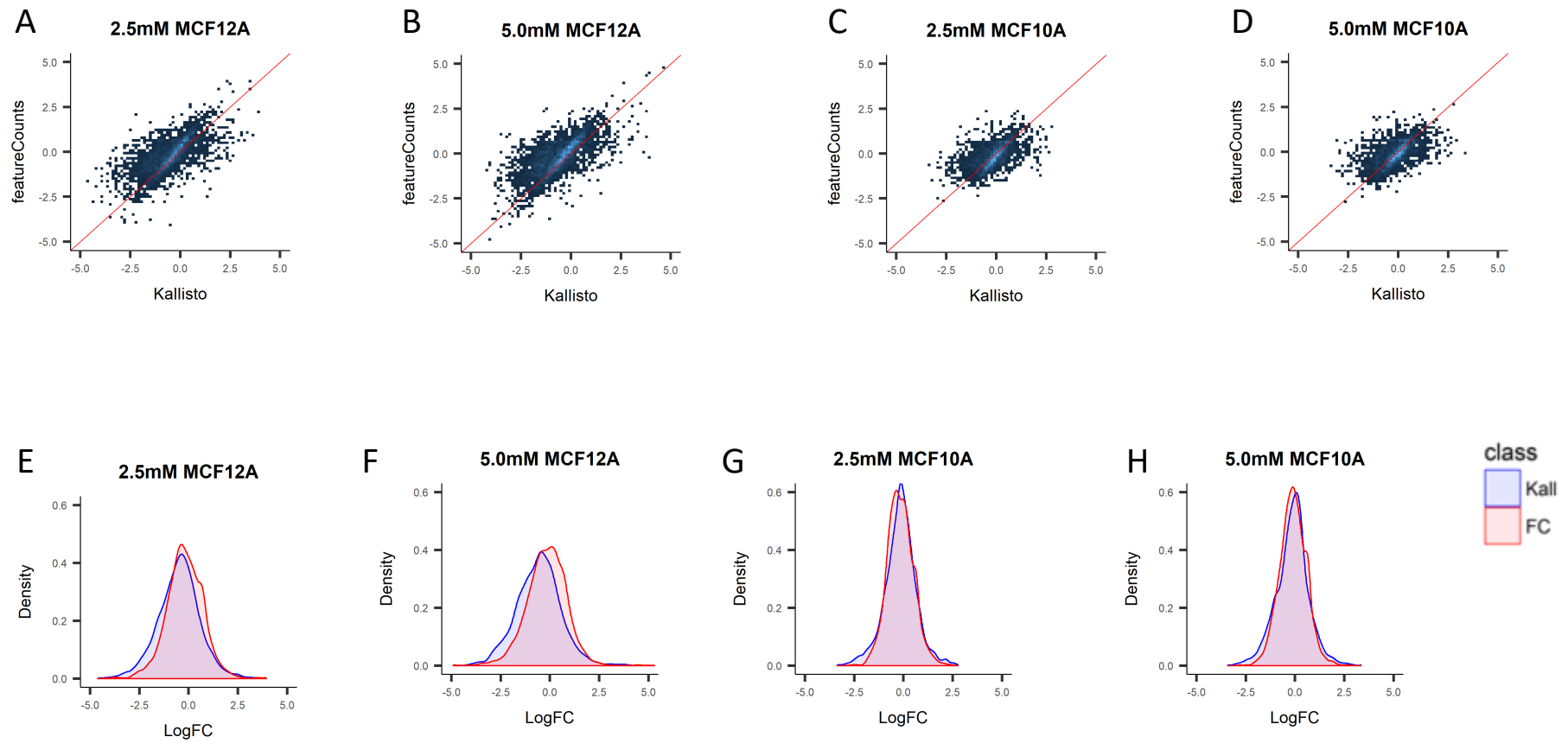


Figure 3.8 Scatter plots showing logFold change gene expression in transcripts overlapping Kallisto and featureCounts analyzed data in untreated vs treated metformin conditions. (A-B) MCF12A non-cancerous breast epithelial cell line. (C-D) MCF10A non-cancerous breast epithelial cell line. Distribution plots of

logFold change gene expression in transcripts overlapping Kallisto and featureCounts analyzed data in untreated vs treated metformin conditions. (E-F)

MCF12A non-cancerous breast epithelial cell line. (G-H) MCF10A non-cancerous breast epithelial cell line. Figure produced by Daisy Stoner, MRes.

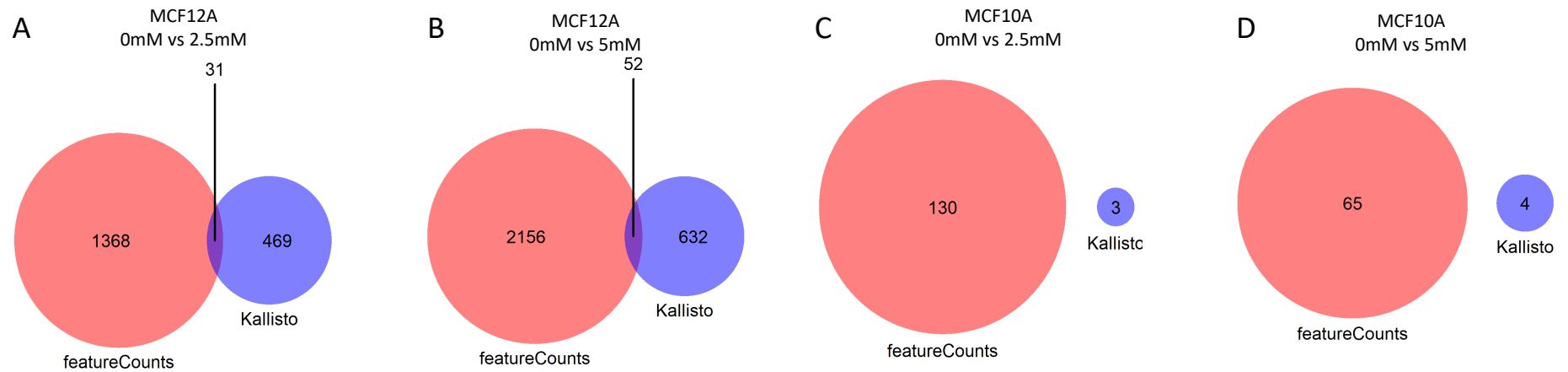


Figure 3.9 Venn diagrams of significant genes (logFC) identified by Kallisto analysis and featureCounts analysis. (A) MCF12A non-cancerous breast epithelial cell lines 0mM vs 2.5mM metformin. (B) MCF12A non-cancerous breast epithelial cell lines 0mM vs 5mM metformin. (C) MCF10A non-cancerous breast epithelial cell lines 0mM vs 2.5mM metformin. (D) MCF10A non-cancerous breast epithelial cell lines 0mM vs 5mM metformin. Figure produced by Daisy Stoner, MRes.

3.2.3 Validation of metformin's effects on gene expression

To validate significant data found in Sleuth analysed MCF12A and MCF10A data set, RT-qPCR was conducted in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72hours with 0mM, 2.5mM and 5mM metformin. Genes for validation were chosen in two ways; (1) three genes showing statistical significance ($FDR < 0.05$) in MCF10A data set and (2) five genes were chosen from MCF12A data set showing statistical significance ($FDR < 0.05$) and logFold change of 0.5 – 4, either gain or loss (appendix table 3.2). Figure 3.10 showed a heatmap of 8 validation genes analysed by Sleuth logFold change, featureCounts logFold change and RQ value obtained from qPCR clustered by dose (2.5mM and 5mM). 7 out of 8 genes validated between Sleuth and qPCR data. Further to this, there was significant correlation of the 7 validated genes between qPCR and Sleuth analysed data ($r=0.8$, $p < 0.05$). Similarly, 7 out of 8 genes validated between featureCounts and qPCR data. There was a strong correlation of 7 validated genes between qPCR and featureCounts analysed data ($r=0.9$, $p < 0.05$).

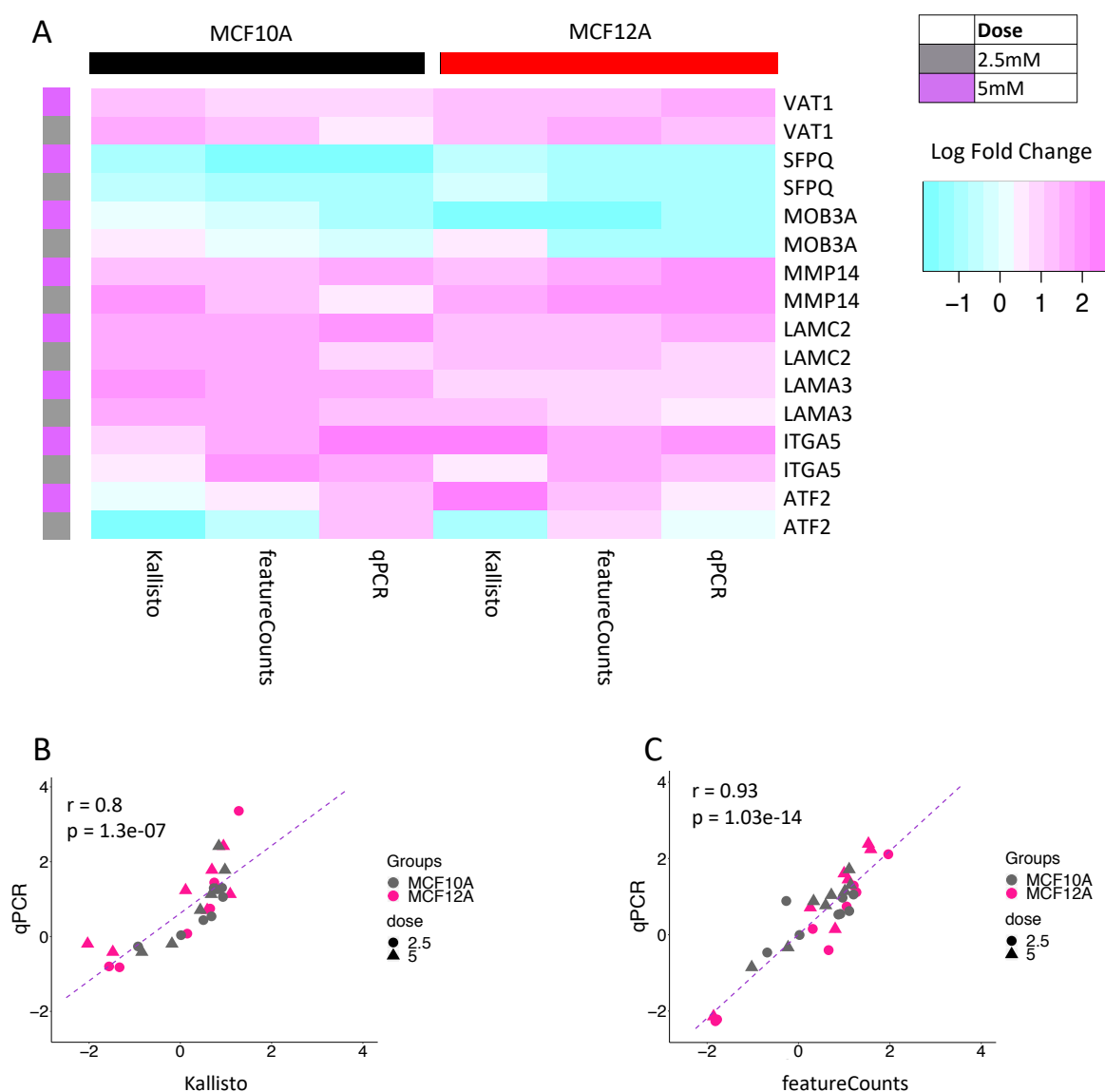


Figure 3.10 Validation of RNA-sequencing data in eight genes chosen showing 0.5 – 4-fold increased/decreased gene expression. (A) Heatmap of log fold change from RNA sequencing, analyzed by Kallisto and featureCounts, and qPCR validation for eight genes in two cell lines, MCF10A and MCF12A, treated with 2.5mM and 5mM metformin. Correlation plot of values from seven genes chosen from RNA-sequencing (B) Kallisto data in comparison to gene expression log fold change derived from qPCR (Pearson correlation $r=0.8$, $p<0.05$). (C) featureCounts data in comparison to gene expression log fold change derived from qPCR (Pearson correlation $r=0.9$, $p<0.05$).

To further validate these results, qPCR of 8 validation genes was conducted in three primary normal matched breast epithelial and myoepithelial patients (1989, 2357 and 3002) treated for 7-days in 1.5mM and 1mM metformin, respectively. Heatmap in figure 3.11 showed that 8 validation genes chosen from cell line data show similar logFold change in primary cells. However, there was heterogeneity between patients and within patients, suggesting there was intrinsic biological differences between cell types and how patients responded to metformin.

This small cohort suggested that metformin may be altering gene expression, which could be validated in primary cell samples. However not all genes in the validation cohort showed consistent gene expression changes suggesting the methods of analysis or validation may need to be considered when inferring conclusions from this data.

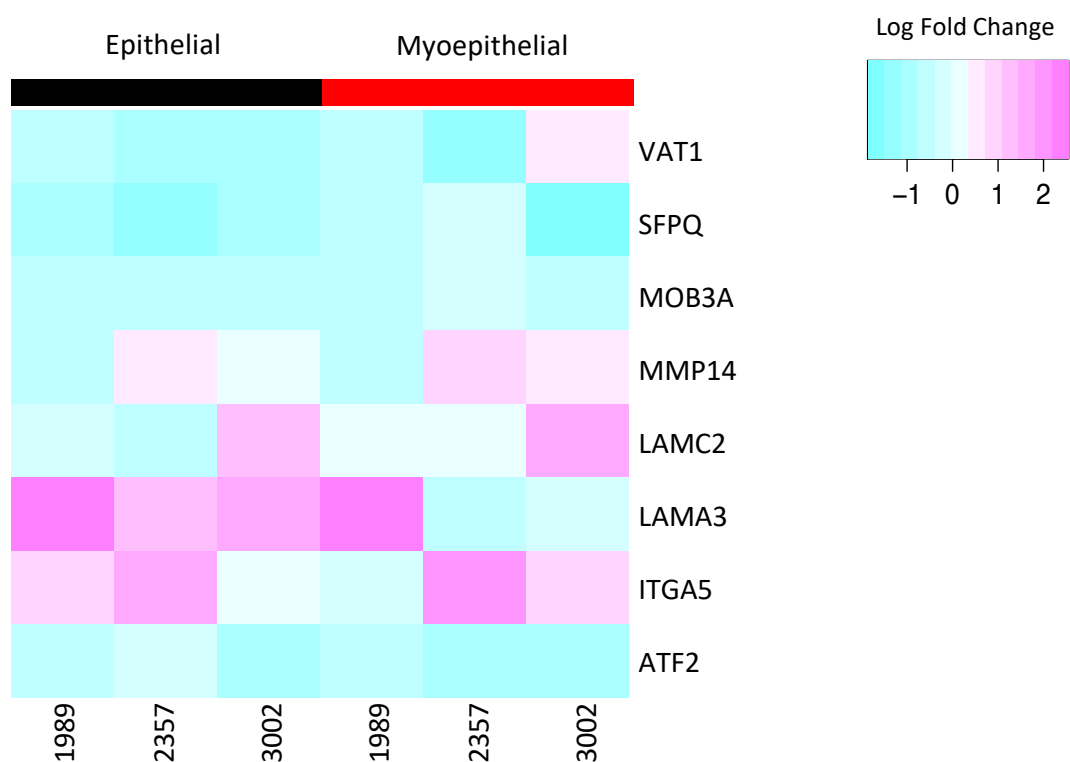


Figure 3.11 Validation of RNA-sequencing data in eight genes chosen showing 0.5 – 4-fold increased/decreased gene expression. Heatmap of log fold change from qPCR validation for eight genes in matched epithelial and myoepithelial primary normal breast cells from four patients, 1989, 2357 and 3002, treated with 1.5mM and 1mM metformin, respectively.

3.2.4 Differential gene expression changes *in vitro* and *in vivo* in response to metformin

Our data has shown that metformin may be altering gene expression in non-cancerous breast epithelial cell lines. However, our data has also shown metformin may activate similar pathways in non-cancerous breast epithelial cells as it did in breast cancer models. To assess whether metformin induced gene expression changes which could potentially be linked to breast cancer risk, we obtained data generated by a single arm window trial by Lord *et al.* In this trial, metformin was administered to women recently diagnosed with breast cancer for 2-3 weeks prior to conventional treatment (Lord *et al.*, 2015, 2018). This study found significantly differential gene expression changes induced by metformin treatment, including genes involved in key metabolic pathways. Data was obtained pre and post treatment and logFold gene expression changes were determined. Analysis was conducted by Daisy Stoner, MRes. Wilcox rank sum of 102 genes found to be significant in the tumour sample data (TSD) were significantly different between TSD to Kallisto (Wilcox ranked test, $p = 2.688e-07$) and TSD to featureCounts (Wilcox ranked test, $p = 1.274e-06$). 48 genes were shown to be statistically significant in one or more of the three data sets. Of the 48 genes only two genes – *MCM9* and *LTV1* – were found to be statistically significant in all three data sets (figure 3.12). Data suggested that in the 48 significant genes the majority of genes showed loss of gene expression in MCF12A and MCF10A non-cancerous cell lines. In contrast, the TSD showed more gain in genes shown to be reduced in the cell lines and more loss in genes shown to be gained in cell lines. While there were some similarities in direction of expression between all three data sets, both *MCM9* and *LTV1*, showed loss of expression in cell lines whereas in the TSD there was a gain of expression. This data suggested that metformin may induce significant gene expression changes in tumours which contrasted with those seen in non-cancerous breast epithelial models.

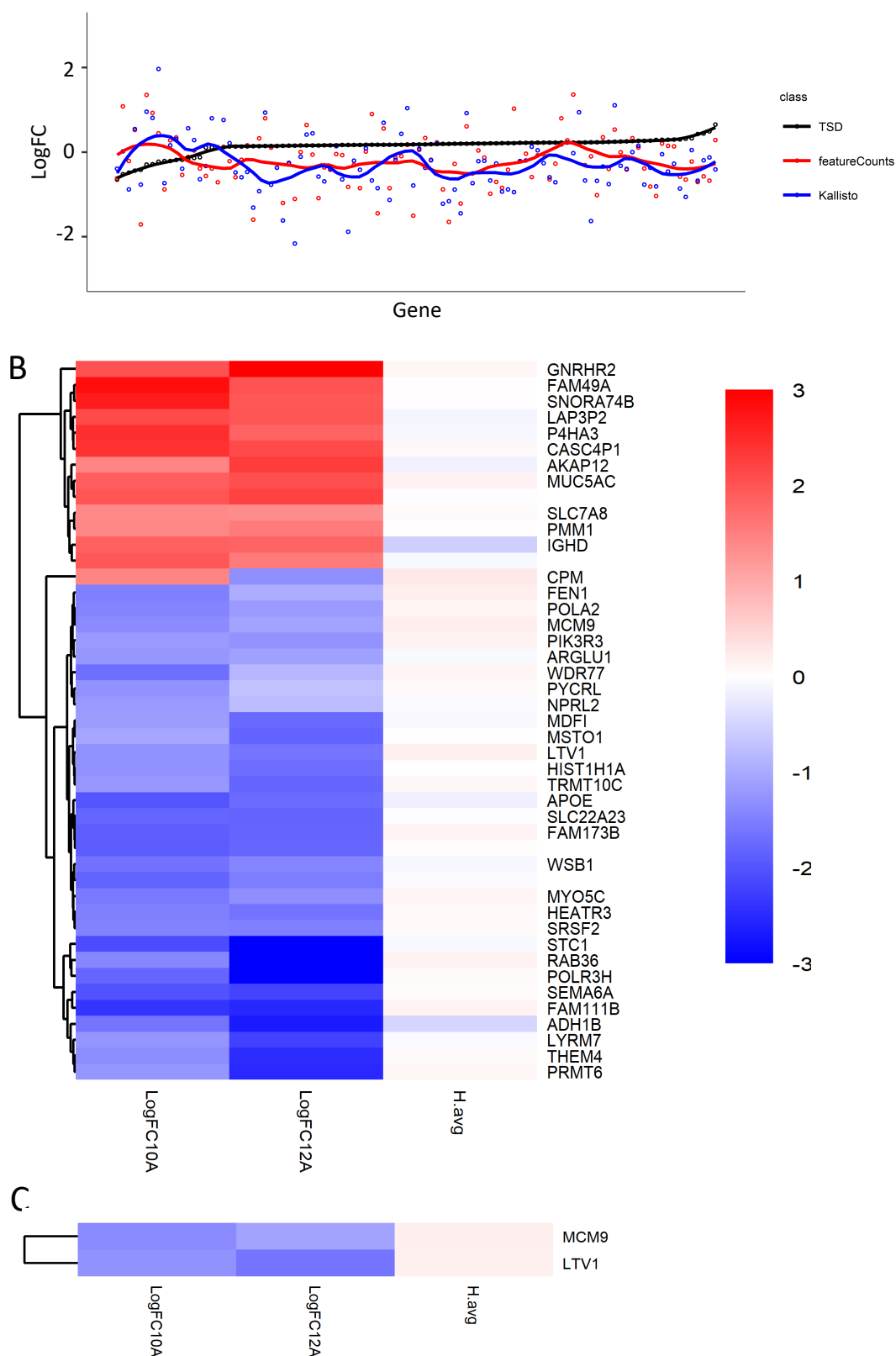


Figure 3.12 (A) Comparison of Tumour Sample Dataset (TSD) to Kallisto and featureCounts analysed data. LogFold change significantly different from zero (t-test). Wilcox rank test: featureCounts vs Kallisto $p = 0.363$; TSD vs Kallisto $p = 2.688\text{e-}07$; TSD vs featureCounts $p = 1.274\text{e-}06$. (B) Heatmap

showing genes found to be statistically significant in one or more of the three data sets (logFold change). (C) Heatmap showing genes found to be statistically significant in all three data sets (logFold change). Figure produced by Daisy Stoner, MRes.

3.2.5 Metformin induces differentially expressed exons

It was unclear whether metformin was driving the gene expression changes seen in this data or if the effects were downstream of metformin's main mechanism(s) of action. To study this further, differential exon splicing events were analysed in featureCounts dataset in MCF12A and MCF10A cell lines by Daisy Stoner, MRes. Figure 3.13 showed that there were exons showing significant differential expression (DE) in both MCF12A (n = 2684) and MCF10A (n = 6721). While MCF10A showed fewer statistically significantly altered genes, there were more differential exon splicing events in MCF10A compared to MCF12A. This could potentially be a reason for the variation seen in MCF10A cell line data. Only 42 DE exons overlapped between both cell lines (figure 3.13). The majority of the DE exon regions (91%) also showed loss of exons. This suggested alternative transcripts could contribute potentially to gene expression changes. Further to this, analysis of exons showing DE exons to genes found to have differentially expressed RNA levels found 5,107 DE exons and 3,081 DE RNA however only 425 of these DE regions overlapped (figure 3.13).

This data suggested that while gene expression changes may be contributing to the effects of metformin, it did not appear to be the only mechanism driving metformin's effects.

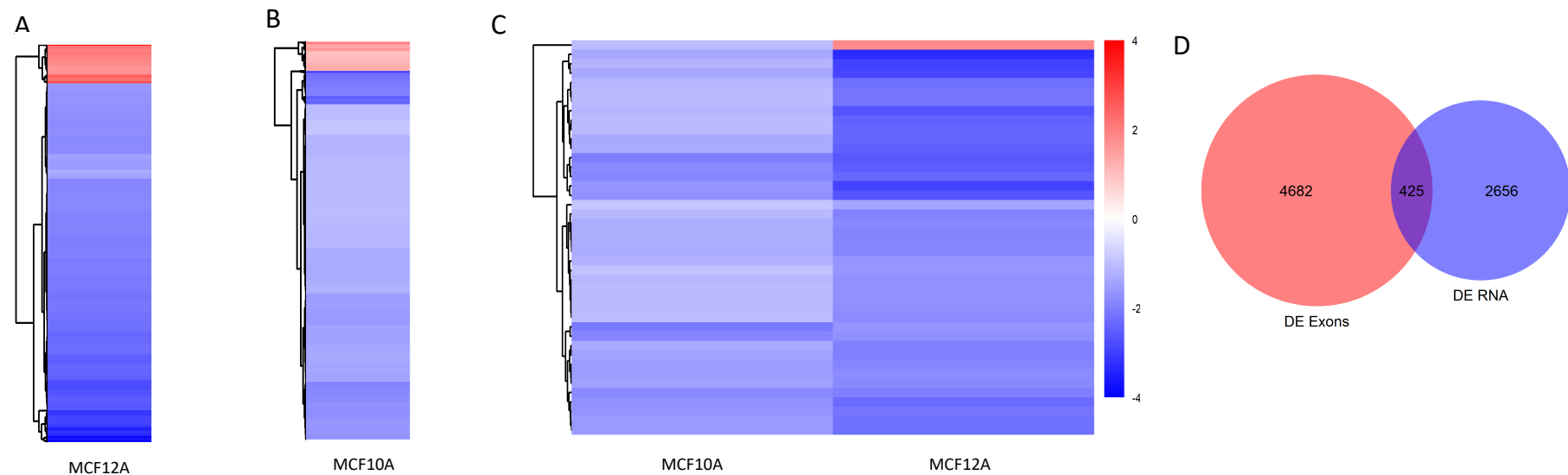


Figure 3.13 Analysis of differential exon splicing in featureCounts analyzed data. (A) Exons showing significant differential expression in MCF12A non-cancerous breast epithelial cell line ($n = 2694$, logFold change). (B) Exons showing significant differential expression in MCF10A non-cancerous breast epithelial cell line ($n = 6721$, logFold change). (C) Exons showing significant differential expression overlapping in MCF12A and MCF10A non-cancerous breast epithelial cell lines ($n = 42$, logFold change). (D) Venn diagram of genes found to have differentially expressed (DE) levels of RNA compared to genes showing differentially expressed exons. Calculated as untreated vs treated. Figure produced by Daisy Stoner, MRes.

3.2.6 Metformin potentially modifies known breast cancer risk genes

Metformin has been implicated as a preventative drug by reducing breast cancer risk. To assess whether metformin may be affecting known breast cancer risk genes a panel of genes was obtained from publications (shown in appendix table 3.1). The expression of these genes was analysed in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin. Analysis showed that out of 27 risk genes, 4 showed significant gene expression changes in MCF12A in response to metformin (table 3.1). For example, in breast cancer, *UBLCP1* showed significant increase in expression associated with increased breast cancer risk (z-score 5.93). In MCF12A, *UBLCP1* showed significant downregulation in metformin treated cells (logFold change -1.6, FDR = 6.56×10^{-5}). In MCF10A, 3 genes showed significant gene expression changes in response to metformin, however, did not show significance after multiple correction testing (appendix table 3.4).

This data suggested metformin may be interacting with genes known to modify breast cancer risk.

Gene	Breast Cancer Risk Alleles Derived from Breast Cancer Models	Metformin Treated Non- Cancerous Breast Epithelial Cell Line MCF12A	Metformin LogFold Change	Metformin FDR<0.05
ANKLE1	increase expression	reduce expression	-0.9730416	0.04357602
GALNT16	reduce expression	increase expression	1.72447626	0.01409484
ANKRD34A	reduce expression	increase expression	0.13858215	0.04601132
UBLCP1	increase expression	reduce expression	-1.5768576	6.56E-05

Table 3.1: Comparison of expression of known breast cancer risk genes in breast cancer models vs non-cancerous breast epithelial cell line MCF12A treated with metformin

3.3 Discussion

Little is understood about the preventative effects of metformin in normal breast epithelial cells. This study chose to focus on cell line models of non-cancerous breast epithelial cells (MCF12A cell line) and a pre-cancerous breast epithelial cell line (MCF10A cell line). This study showed both cell lines were affected by metformin treatment leading to reduced cell viability at increasing metformin doses, showing sensitivity to metformin. MCF12A cell line showed more sensitivity compared to the pre-cancerous model MCF10A (IC₅₀ 2.8 vs IC₅₀ 3.8, respectively). Further to this, normal primary matched breast epithelial and myoepithelial cells showed sensitivity to metformin (IC₅₀ 1.1 and IC₅₀ 1.3 following 7-days, respectively). Comparison of passage number found that in myoepithelial cells P2 cells were more sensitive to metformin compared to P3 cells. From observation of these cells we hypothesised that, as P3 myoepithelial cells appear to grow slower than P2 myoepithelial cells, cell growth speed could be a potential limiting factor for metformin treatment. To test this hypothesis analysis of cell cycle by flow cytometry could provide information about doubling time of each passage type. Using this information, a crystal violet cell viability assay could be conducted to analyse the effect of metformin at similar doses correlated to doubling time. Analysis of IC₅₀s could determine metformin sensitivity. Preliminary analysis of untreated P3 and P2 myoepithelial cells by crystal violet assay following 7-day incubation showed average 35% reduction in cell percentage in P3 myoepithelial cells compared to P2 (appendix figure 3.2). P3 myoepithelial cells grow slower than P2 myoepithelial cells as well as being more sensitive to metformin (IC₅₀ 3.2 vs IC₅₀ 1.3, respectively) suggested that slower growing cells were potentially more resistant to metformin. This could be important in using metformin as a preventative or anti-cancer agent and should be considered whether treatment in slow growing tumours would be beneficial.

Two major limitations of this study were the timepoints used for treatments and concentrations of metformin used. As few studies studied non-cancerous breast epithelial cells, review of metformin

treatment breast cancer cell line models was used to determine a range of metformin treatments and timepoints.

The average dose of metformin used was 5mM for 24-48hours. As we wanted to analyse DNA methylation and histone post-translational modifications (PTMs) we chose 3-day minimum treatment in cell line models based on previous published studies showing loss of DNA methyltransferases induced following 3-day treatment with 5-azacytidine (Hollenbach *et al.*, 2010). However, this treatment time may not fully elucidate the effects of long-term treatment of metformin on normal breast epithelial cells. A longitudinal study treating non-cancerous breast epithelial cells with low dose metformin over a number of months would be ideal to study the effects of metformin in a long-term setting. Choosing a lower dose of metformin could bring non-cancerous breast epithelial cells closer to physiological conditions as this study used supraphysiological doses of metformin. A recent published study has shown that long-term metformin treatment can lead to metformin resistance in breast cancer cell lines (Scherbakov *et al.*, 2016). This would also need to be considered when conducting such a study.

Doses of metformin used in this study did not reflect physiological levels. The average concentration of metformin in the liver *in vivo* was 220 μ M whereas the average concentration of metformin in serum *in vivo* was 10-40 μ M (He and Wondisford, 2015). This was very low compared to the doses chosen for this study (2.5mM and 5mM). There are no pharmacodynamic studies of metformin in normal breast epithelial cells *in vivo*, so it is unknown the concentration of metformin that reaches the breast tissue. Analysis of metformin concentrations in normal breast tissue could identify ideal concentrations of metformin to use.

One of the key hallmarks of tumourgenesis is distinct gene expression changes that lead to tumour formation and progression. Tumour types, such as breast cancer, can have gene mutations associated with specific subtypes allowing these tumours to be identified. It is hypothesised that manipulation of

pre-cancerous gene expression changes could potentially revert pre-cancerous cells to “normal” or kill pre-cancerous cells before they form a tumour. This hypothesis raises an interesting question about preventative techniques. Is the intervention targeting normal cells and preventing them becoming cancerous or are they targeting pre-cancerous cells?

Analysis of gene expression in metformin treated non-cancerous breast epithelial cells showed significant alterations in gene expression. 7 out of 8 genes chosen for validation were consistent across sequencing and qPCR results and correlated significantly for example, *SFPQ* and *VAT1*. *SFPQ* has recently been shown to be involved in double strand break repairs and low levels of *SFPQ* were associated with delays in DNA repair (Salton *et al.*, 2010). In this study *SFPQ* showed significant downregulation of expression. In contrast *VAT1* showed significant increase in gene expression when treated with metformin. *VAT1* plays a role in synaptic vesicles (Linial, Miller and Scheller, 1989). Metformin treatment induced significant differential expression of exons, suggesting while gene expression changes may not be driving metformin’s effects, a portion of genes could potentially be associated one or more of metformin’s mechanisms. Further analysis of genes associated with differential expression of exons may identify pathways for further study.

A number of genes associated with breast cancer risk have been identified in recent years. This study showed metformin may be interacting with some but not all of these genes to potentially reverse the expression of these genes in non-cancerous breast epithelial cells compared to breast cancer models. Further analysis of these genes is required, potentially by identifying which pathway(s) these gene(s) are associated with and modifying that pathway.

Comparison with clinical trial data obtained from breast tumours did not show consistent overlap with cell lines data. In fact, the two genes that showed significance across all three data sets - *MCM9* and *LTV1* – were increased in clinical trial samples and decreased in cell line samples. Loss of *LTV1*, which

encodes for a ribosomal assembly factor, has been implicated in impaired apoptosis in breast cancer (Collins *et al.*, 2018). *MCM9* mutations, which encodes for a chromosome maintenance factor for homologous recombination, have been associated with early-onset colorectal cancer. *MCM9* mutations have also been associated with premature ovarian failure (POF) (Wood-Trageser *et al.*, 2014; Goldberg *et al.*, 2015). These results could give indications into whether metformin is targeting pre-cancerous cells. Further study into the progression of breast cancer and metformin's role would be important to elucidate this further. A model of breast cancer progression could be developed from a number of breast cancer cell lines, from pre-cancer to metastatic cell line models using a common tissue bank. Tracking of gene expression of key genes in tumourgenesis with and without metformin could give insight into where/if metformin is interfering with this pathway. Conducting similar experiments in primary normal and cancerous breast epithelial cells would provide additional validation. As shown in this study, analysis of gene expression of validation genes in primary normal breast epithelial and myoepithelial cells showed similar expression patterns.

A limitation in analysis of sequencing was the choice of alignment and downstream analysis as statistical analyses techniques have been shown to produce differing results. To attempt to address this potential limitation in the analysis of our RNA sequencing data, we chose to use two alignment techniques with two additional statistical analysis techniques. While few overall significant genes overlapped in MCF12A and no overall significant genes overlapped in MCF10A, the data sets correlated well with each other. Importantly, when Kallisto/Sleuth and featureCounts results were compared between the validation gene cohort there was significant correlation ($r = 0.7$, $p < 0.05$, appendix figure 3.3). This suggested that while there were differences in significant results found by the two analysis techniques the inferences made from either technique could be reliable.

Results from this study suggested that metformin may be inducing significant gene expression changes in cell line models of non-cancerous breast epithelial cells. These changes could help identify pathways

metformin may be interacting with, further elucidating metformin's mechanisms of action in normal breast epithelial cells.

CHAPTER 4

4. Metformin modifies one-carbon metabolism and subtly alters DNA methylation in normal breast epithelial cell lines

4.1 Introduction

DNA methylation is one of the key regulators of gene expression, chromatin confirmation and protein-DNA interactions (Robertson, 2005; Bernstein, Meissner and Lander, 2007). The addition of a methyl group on a CpG site has been associated with an oppressive effect on gene expression. However, there was evidence to suggest that while DNA methylation in gene promoter regions were associated with reduced expression, DNA methylation in gene bodies or non-coding areas may not be linked to gene repression. The main donor for DNA methylation S-adenosylmethionine, SAM, is generated in the methionine cycle which is a part of one-carbon metabolism. Methionine is converted to SAM by methionine adenosyltransferase 2A (MAT2A). This is recycled to S-adenosylhomocysteine (SAH), homocysteine and finally back to methionine. Methionine is also generated as serine is metabolised in the folate cycle to generate nucleic acids, for example thymidine. These two cycles link to the form one-carbon metabolism pathway (figure 4.1, (Mentch and Locasale, 2016; Ducker and Rabinowitz, 2017; Newman and Maddocks, 2017)).

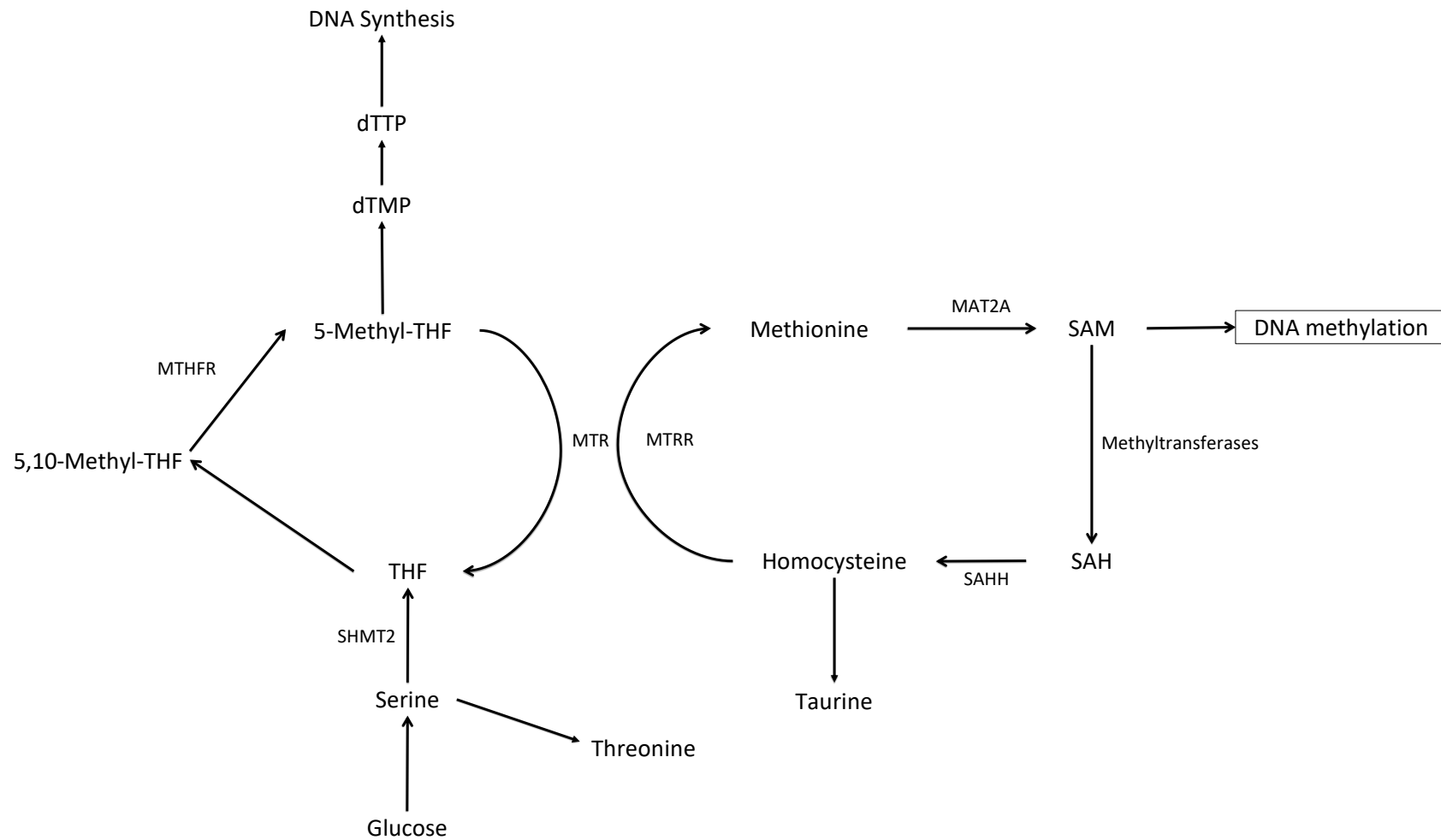


Figure 4.1 One carbon metabolism pathway. THF; Tetrahydrofolate, 5-Methyl-THF; 5-Methyl-tetrahydrofolate, dTMP; Thymidine monophosphate, dTTP; thymidine triphosphate, SAM; S-adenosylmethionine, SAH; S-adenosylhomocysteine, MAT2A; methionine adenosyltransferase 2A, MTHFR;

methylenetetrahydrofolate reductase, MTRR; 5-methyltetrahydrofolate-homocysteine methyltransferase reductase, MTR; 5-methyltetrahydrofolate-homocysteine methyltransferase, SAHH; S-adenosylhomocysteinase, SHMT2; serine hydroxymethyltransferase 2 [Adapted from (Mentch and Locasale, 2016; Ducker and Rabinowitz, 2017; Newman and Maddocks, 2017)].

DNA methylation modifications have been associated with breast cancer risk factors, such as T2DM, hyperinsulinemia and hyperglycaemia. For example hypermethylation of the insulin gene (*INS*) promoter has been associated with hyperglycaemia (Yang *et al.*, 2011). A major risk factor of T2DM is aging. It has been demonstrated that DNA methylation plays a role in repressing the *COX7A1* gene, which is an important part of the mitochondrial respiratory chain complex 4. *COX7A1* expression was associated with glucose uptake and loss of this gene could contribute to the hyperglycaemic state of diabetic patients (Rönn *et al.*, 2008). It had also been demonstrated that there was a link between hypermethylation of glucokinase (*Gck*) and L-type pyruvate kinase (*LPK*) promoters, which contributed to obesity and insulin resistance in the liver in high-fat fed mice compared to normal controls (Jiang *et al.*, 2011). In the pancreas of humans, the gene *PPARGC1A* was hypermethylated leading to down regulation of PGC-1 α expression. PGC-1 α is a key regulator of genes involved in cellular metabolism such as oxidative phosphorylation. Loss of this gene was seen in the skeletal muscle of diabetics compared to non-diabetics (Ling *et al.*, 2008). There is on-going research into the changes in DNA methylation associated with T2DM seen in different countries/regions, suggesting a key environmental factor in T2DM development (Al Muftah *et al.*, 2016). Conversely, hypomethylation was also seen to increase T2DM risk as hypomethylation of the *FTO* gene was strongly related to T2DM and obesity (Dayeh *et al.*, 2014). These DNA methylation changes and others were associated with disease progression and could be used as biomarkers for T2DM severity.

Metformin has been linked to DNA methylation changes. A number of studies by Yan *et al.* and Zhong *et al.* focus on the effect of metformin on H19/SAHH axis (Yan *et al.*, 2015; Zhong *et al.*, 2017). H19 is regulated by let-7 miRNA. These studies showed metformin treatment decreasing H19 expression, increasing let-7 levels and increasing S-adenosylhomocysteinase (SAHH) enzyme levels. SAHH hydrolyses SAH which can in turn inhibit DNMT3B. This study also showed decreased SAH levels associated with metformin treatment in non-cancerous breast epithelial cells. This links with previous studies which show decreased DNA methylation of H19 has been associated with T2DM progression.

Genome wide methylation changes were also identified in metformin treated cells. SAHH is a key enzyme in methionine metabolism. This pathway has recently been shown to be altered by metformin treatment, leading to increased SAM/SAH ratio and increased DNA methylation in MCF10A non-cancerous breast epithelial cell lines (Cuyàs *et al.*, 2018). In lymphoblastoid cell lines, DNA methylation changes were associated with metformin IC50, leading to changes in over 200 probes (Niu *et al.*, 2016). In diabetic patients, metformin was shown to reduce DNA methylation levels of three metformin-related genes. *SLC22A1* (OCT1) and *SLC22A3* (OCT3) transport metformin into the cell while *SLC47A1* (MATE1) transports metformin out of the cell (García-Calzón *et al.*, 2017). Metformin alone was shown to reduce DNA methylation levels of these sites in diabetic patients to almost the same level of DNA methylation around the same genes in non-diabetic controls. Insulin alone did not recapitulate this effect. Indirect evidence additionally suggested metformin's effect on DNA methylation. As previously discussed, metformin can activate AMPK through modified ATP:AMP ratio. AMPK activation has been shown to inhibit DNA methyltransferases such as DNMT1 (Marin *et al.*, 2017).

In this chapter we explore the hypothesis that metformin may be altering DNA methylation in non-cancerous breast epithelial cell lines.

4.2 Results

4.2.1 Metformin alters gene expression of key methionine pathway enzymes

Previous studies, for example by Anderson *et al.* have shown a link between one-carbon metabolism and DNA methylation, where modification of this metabolic pathway leads to significant changes in DNA methylation levels (Anderson, Sant and Dolinoy, 2012). Metformin has been linked to changes in metabolic pathways as well as DNA methylation. To investigate this relationship, we chose to explore whether metformin may be affecting SAM generation pathways. Analysis of RNA sequencing Kallisto-generated data was conducted on folate and methionine pathway enzymes. Figure 4.2 showed gene expression LogFold change untreated vs treated of a number of enzymes in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 2.5mM and 5mM. In the methionine pathway there was a decrease of expression of *MAT2A*, *MTR* and *MTRR* in both MCF12A and MCF10A. These enzymes are involved in the generation of SAM from homocysteine. However, the gene *ACLY* (which codes for the enzyme SAHH) showed increased gene expression in MCF12A and was lost at higher doses in MCF10A. In the other part of one-carbon metabolism, the folate pathway, there was increased gene expression of *SHMT2* and *MTHFR* in MCF10A cell line. However, in the MCF12A cell line while there was increased gene expression in *SHMT2* there was decreased expression in *MTHFR*. This data was further validated by qPCR (figure 4.3), where similarly to sequencing data *MAT2A*, *MTR* and *MTRR* showed a significant decrease in gene expression when treated with metformin ($p < 0.05$). Only *MTHFR* showed significant changes in MCF10A and again was consistent with sequencing data which showed a significant increase in response to metformin ($p < 0.05$). Sequencing data and qPCR showed significant correlation ($r = 0.9$, $p < 0.05$). This data suggested that metformin may be affecting the one-carbon pathway by altering gene expression of key enzymes in this pathway.

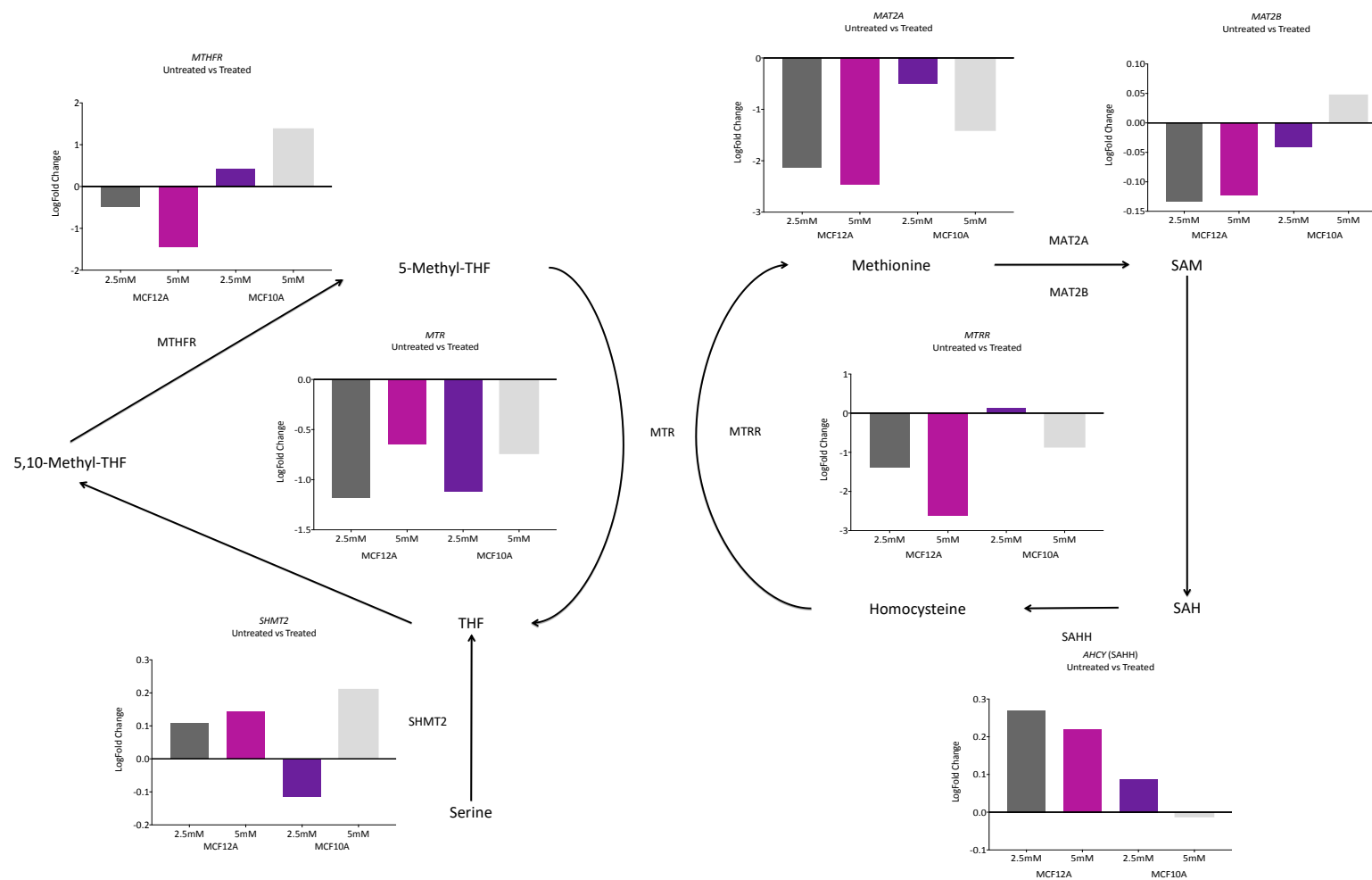


Figure 4.2 LogFold change gene expression untreated vs treated of one carbon metabolism pathway related genes – *SHMT2*, *MTHFR*, *MTR*, *MTRR*, *MAT2A*, *MAT2B* and *AHCY* – in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. Averaged over six biological replicates (n = 6).

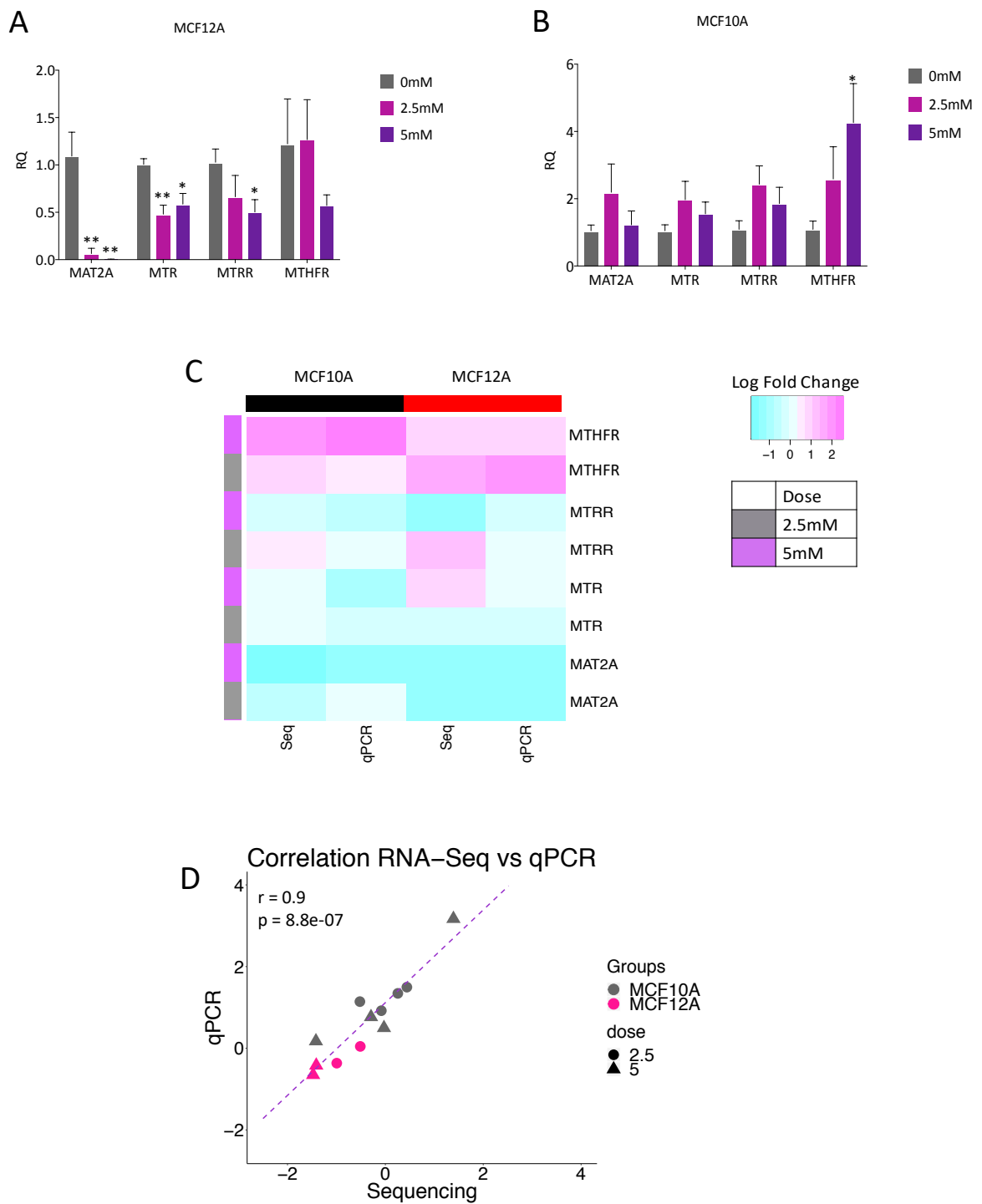


Figure 4.3 Analysis of methionine pathway enzymes – *MAT2A*, *MTR*, *MTRR* and *MTHFR* – in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin. Student's t-test * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 . (C) Heatmap of Sleuth

logFold change sequencing data and qPCR results for four methionine pathway enzymes. (D)

Correlation of sequencing data vs qPCR data ($r = 0.9$, $p = 8.8e-06$).

4.2.2 Metformin alters SAM/SAH axis

As metformin has been shown to alter the gene expression of key enzymes in the methionine pathway, we hypothesised that metformin may be subsequently altering key metabolites in this pathway. To assess whether metformin may be affecting DNA methylation through alterations of the methionine pathway hydrophilic interaction liquid chromatography mass spectrometry (HILIC-MS) was conducted on MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in biological and technical triplicates. MCF12A cell line showed significant increase in SAM levels in the highest dose of metformin and showed significant dose dependency ($p < 0.05$) however MCF10A cell line showed no difference in SAM levels. In contrast there was no change in SAH levels in MCF12A and MCF10A cell lines. This result showed significant increase in SAM:SAH ratio in both cell lines ($p < 0.05$, figure 4.4).

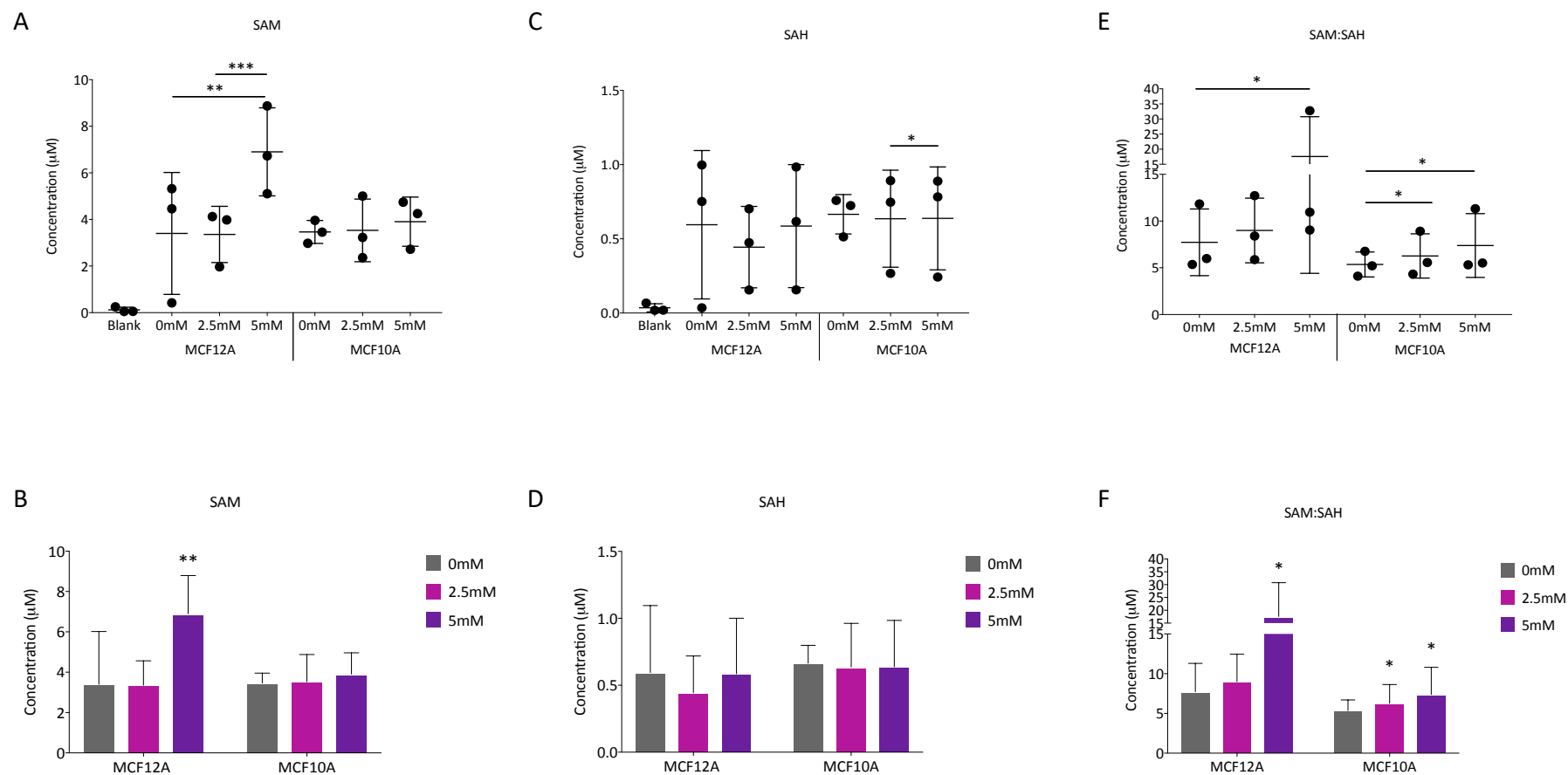


Figure 4.4 HILIC-MS analysis of SAM/SAH axis in MCF12A and MCF10A non-cancerous breast epithelial cell line treated for 72 hours with 0mM, 2.5mM and 5mM metformin. (A-B) S-adenosylmethionine iodine, (C-D) S-adenosylhomocysteine and (E-F) SAM:SAH ratio. Student's t-test * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 .

Further analysis of key metabolites generated in one carbon metabolism showed significant changes in response to metformin (figure 4.5). Metabolites from the methionine pathway were analysed initially. In addition to significant increase in SAM levels in MCF12A there was significant increase in methionine levels. Homocysteine levels significantly decrease in MCF12A following metformin treatment. In contrast, while methionine levels significantly increase in MCF10A ($p < 0.05$), there was no change in homocysteine levels in MCF10A. Other metabolites related to this pathway were analysed including serine, threonine and taurine. Serine, which can be generated from glucose, showed significant increase in both cell lines and showed dose dependency in MCF12A ($p < 0.05$). Taurine metabolite levels significantly increase in both cell lines ($p < 0.05$). Taurine can be generated from homocysteine in an additional pathway. Threonine metabolite levels, which can be generated from serine, showed no change in MCF12A but significant increase in MCF10A at the higher dose ($p < 0.05$). This data suggested that metformin may be modifying metabolites in the methionine pathway and related pathways.

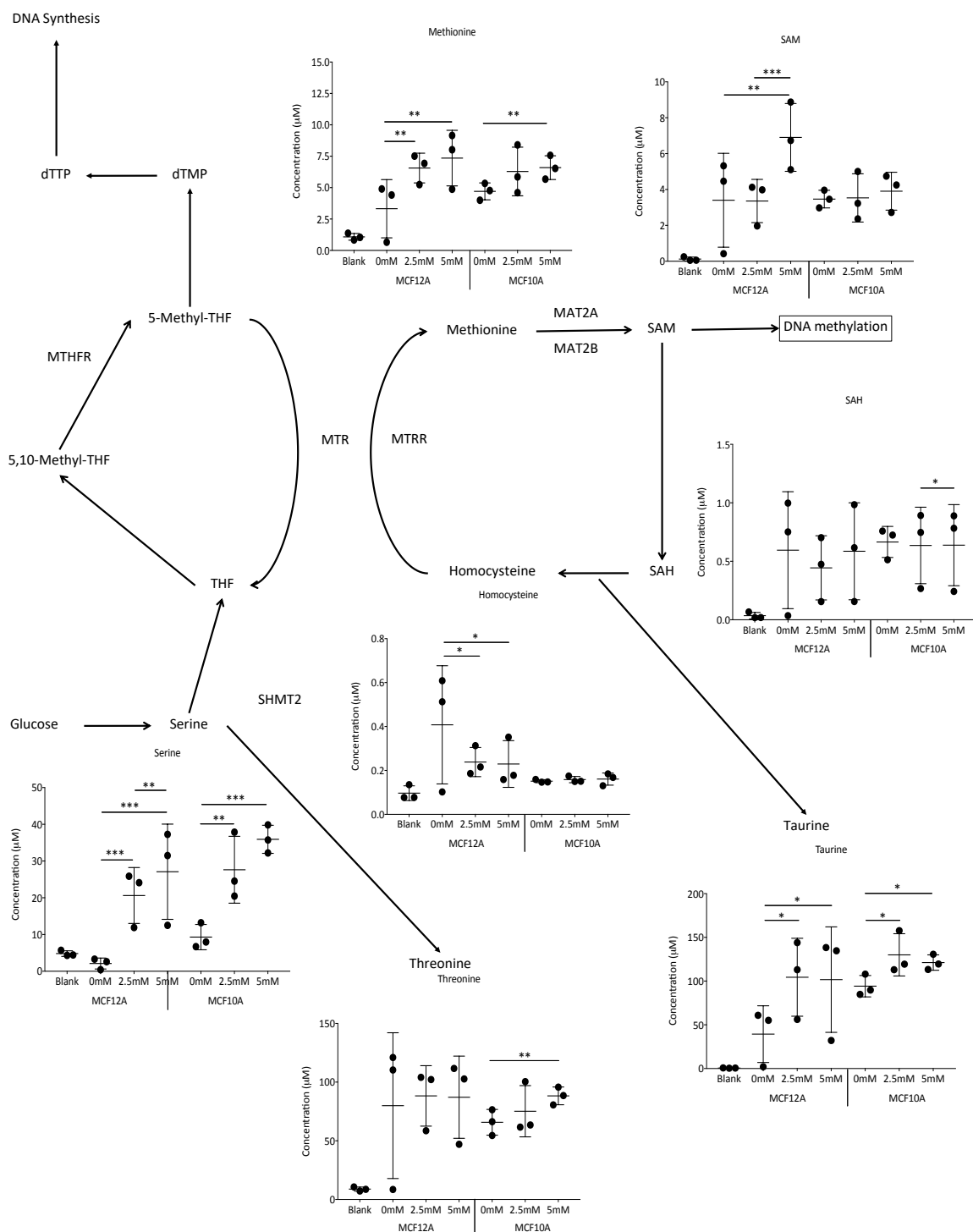


Figure 4.5 HILIC-MS analysis of one carbon metabolism intermediates – Serine, Methionine, S-adenosylmethionine, S-adenosylhomocysteine, Homocysteine, Taurine and Threonine - in MCF12A and MCF10A non-cancerous breast epithelial cell line treated for 72 hours with 0mM, 2.5mM and 5mM metformin.

4.2.3 Metformin induces subtle changes to DNA methylation

SAM was a key methyl donor to DNA methylation. As SAM loses a methyl group it is converted to SAH. The ratio of SAM:SAH could be used as a surrogate marker of DNA methylation. Our study has shown in MCF12A SAM:SAH ratio increases where in MCF10A SAM:SAH ratio remains unaltered. We hypothesised that DNA methylation levels may be changing in response to this. To test this hypothesis, Illumina MethylationEPIC array was conducted in MCF12A and MCF10A cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in five and four biological replicates, respectively. Initial analysis of results (figure 4.6) showed the majority of the variation in the combined MCF12A and MCF10A data was within principle component 1 (PC1) and as seen in the corresponding heatmap, this variation appeared to be driven by the two groups (MCF12A and MCF10A). Analysis of global mean methylation (β -values) across both cell lines and cell line separately showed no significant change in DNA methylation levels.

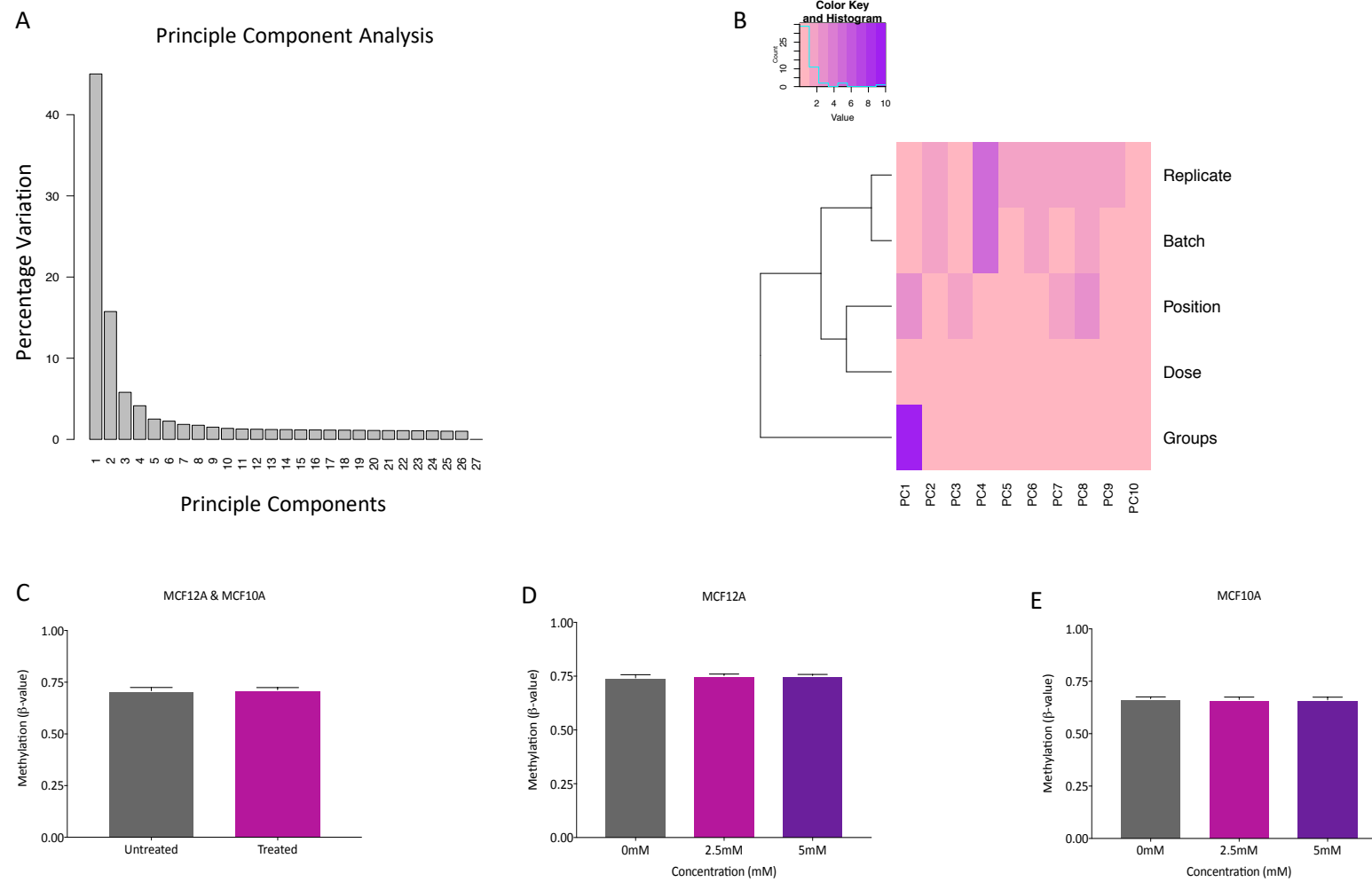


Figure 4.6 Illumina MethylationEPIC array analysis of MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin. (A) Principle component analysis (PCA) histogram of beta values over both cell types and all doses. (B) Heatmap of variables in relation

to principle components. (C) Global methylation MCF12A and MCF10A untreated vs treated. (D) Global methylation of MCF12A cell line. (E) Global methylation of MCF10A cell line.

Data generated from Illumina MethylationEPIC array was analysed by three different models.

The first analysis was a mixed effects model on M-values. M-values were generated from β -values obtained from the Illumina MethylationEPIC array. This analysis combined the data obtained from both MCF12A and MCF10A data sets. The results showed 41 CpG sites were significantly altered by metformin (FDR<0.05). The Manhattan plot in figure 4.7 showed that a number of significantly altered genes were located in the X chromosome and of these 41 CpG sites, the majority showed loss of methylation (figure 4.7). A selection of the top 20 significant probes were chosen for validation (FDR<0.05, appendix table 4.1).

A second linear model of untreated vs treated was conducted on β -values. While M-values were more statistically powerful they don't represent biological effects, whereas β -values represent biological changes in methylation. This analysis was conducted in MCF12A and MCF10A data set separately (figure 4.8). In MCF12A, 944 CpG sites showed >10% difference in methylation in response to metformin and the majority of these CpG sites gain methylation. In contrast, in MCF10A 1755 CpG sites showed >10% difference in methylation however the majority of these CpG sites lose methylation. Comparison of both data sets showed 6 CpG sites demonstrating methylation changes in either the same or opposite direction and magnitude. These CpG sites were chosen for further validation (appendix table 4.3).

The final analysis was conducted in MCF12A and MCF10A data sets separately. Due to batch effects discovered in model 1 and model 2 after adjusting for batch, this analysis was conducted in two parts; methylation differences between doses was determined within four batches. Following this, average methylation differences were identified in combined batches (figure 4.9). CpG sites showing >10% difference in methylation in the same direction across all batches were identified. Two of these CpG sites were chosen for further validation (appendix table 4.5). Figure 4.9 showed there was subtle dose dependent changes in both MCF12A (slope = 0.57) and MCF10A (slope = 0.71).

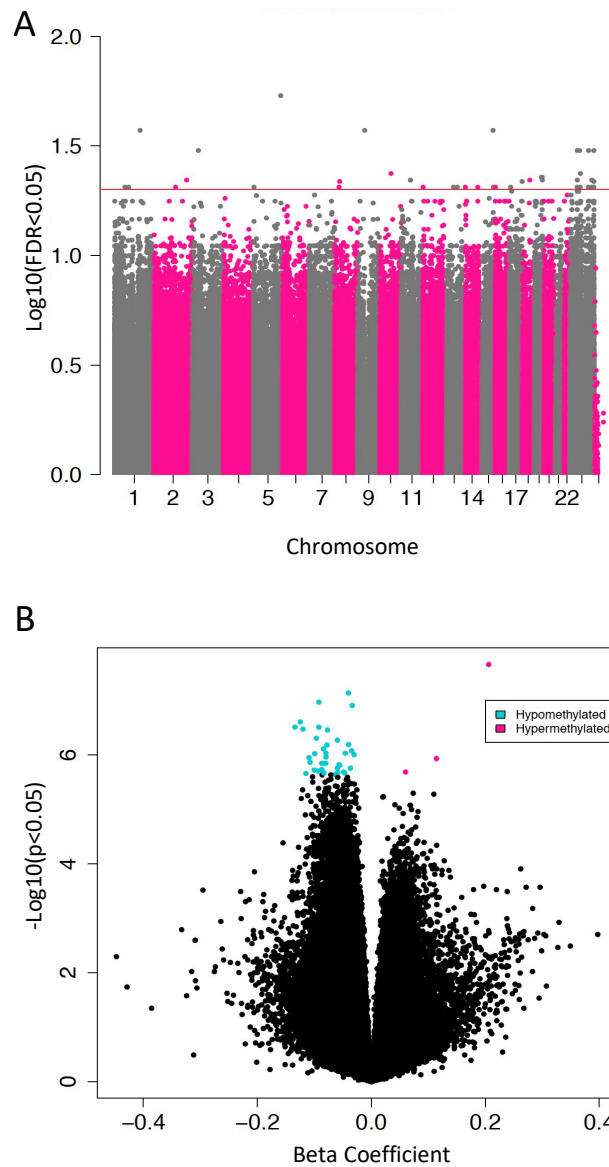


Figure 4.7 Model 1 analysis of Illumina MethylationEPIC array. Mixed effects model analyzing M-values of combined MCF12A and MCF10A non-cancerous breast epithelial data. Metformin induces genome-wide hypermethylation in two non-cancerous breast epithelial cell lines. (A) Manhattan plot of significant probes (FDR<0.05) where 13 out of the 41 probes were located within the X chromosome and show decreased DNA methylation. (B) Volcano plot of significant probes, the majority of which show significant DNA hypomethylation (38 out of 41, FDR <0.05).

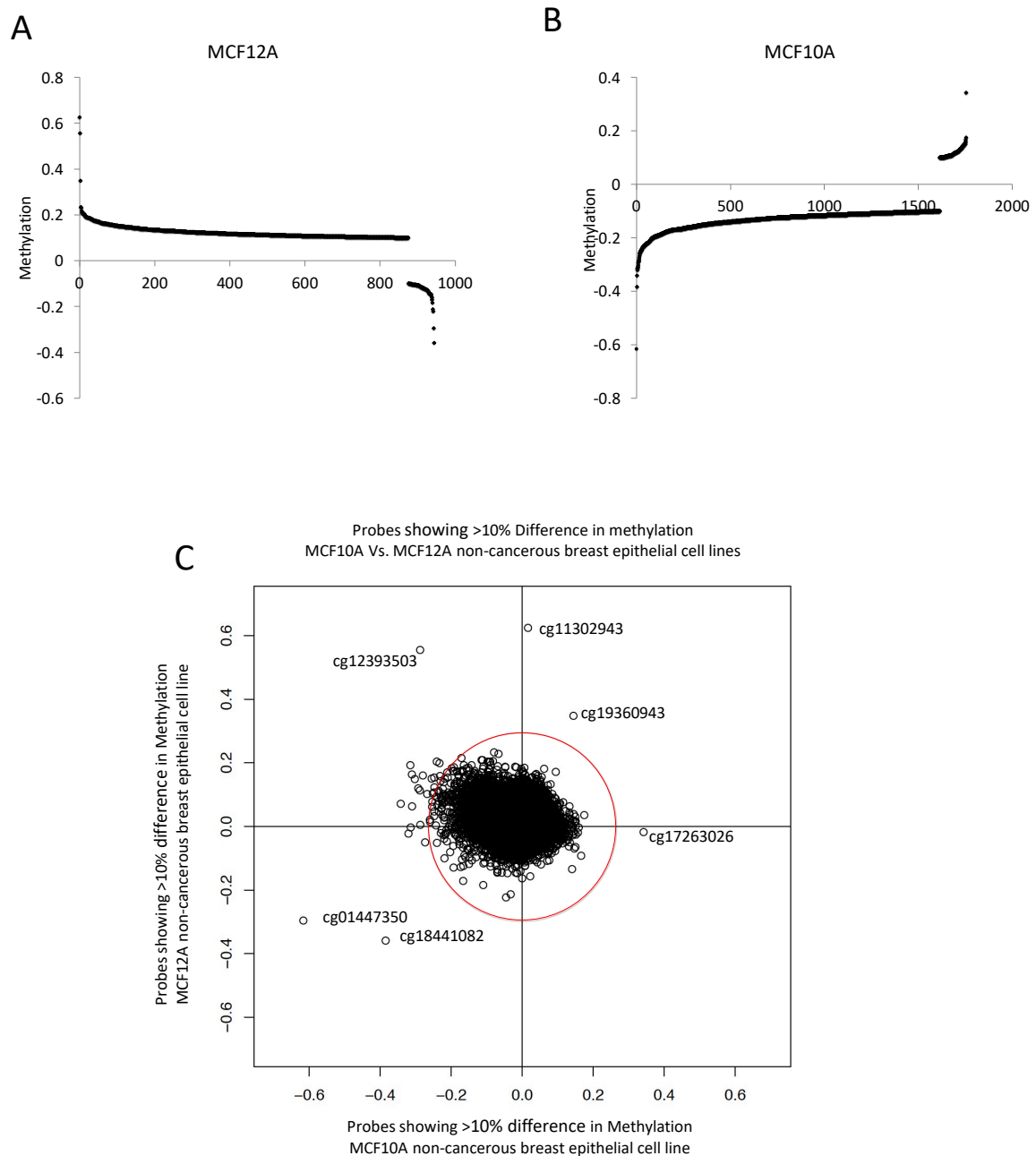


Figure 4.8 Model 2 analysis of Illumina MethylationEPIC array. Linear model of beta-values analyzing magnitude of effect in MCF12A and MCF10A non-cancerous breast epithelial cell lines separately. Metformin treatment alters DNA methylation in two non-cancerous breast epithelial cell lines, MCF10A and MCF12A. (A) Following metformin treatment, 944 probes were found in MCF12A to show >10% difference in DNA methylation and 875 of these probes show hypermethylation. (B) In contrast to MCF12A, 1755 probes were found to show >10% difference in DNA methylation and of these 1614 show hypomethylation in MCF10A. (C) Six probes show differential DNA methylation

changes. A number of probes show DNA methylation changes in the same direction however additional probes show DNA methylation changes in opposite directions.

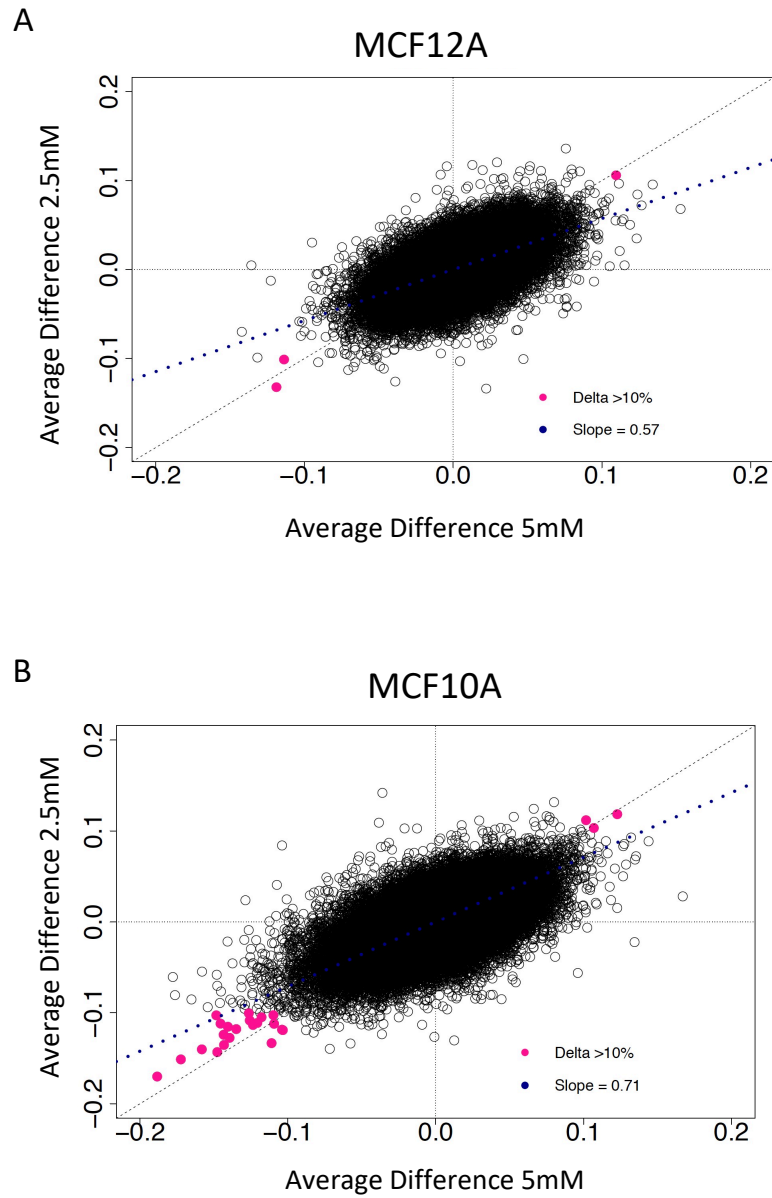


Figure 4.9 Model 3 analysis of Illumina MethylationEPIC array. Scatter plot of all CpG sites showing average difference between 0mM and 2.5mM vs. 0mM and 5mM (beta-values) in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines. Blue dotted line = slope of the line, pink dots = CpG sites showing >10% change in beta value.

Validation of three models was conducted by pyrosequencing of chosen CpG sites (figure 4.10). While model 1 showed a high correlation in values generated from the MethylationEPIC array and pyrosequencing, the difference between untreated and treated samples was 0.5-1% (as shown in figure 4.10 B). This was an example of the statistical power of M-values versus the biological impact of the methylation changes. In model 2 there was little correlation between MethylationEPIC sequencing data and pyrosequencing validation. This model was based on β -values, which as biologically relevant but were not statistically relevant. Similarly, in model 3 there was no correlation between sequencing data and validation data. As this data set was analysed to manually adjust for batch effects seen in model 1 and model 2, we concluded that while there appeared to be changes in DNA methylation following metformin treatment the difference between batches was too great to make concrete statements about the data. However, this did show that metformin may not affect each cell in the same way.

This data suggested while there may be subtle changes to DNA methylation induced by metformin treatment, there appeared to be no significant changes to DNA methylation at individual CpG sites in response to metformin.

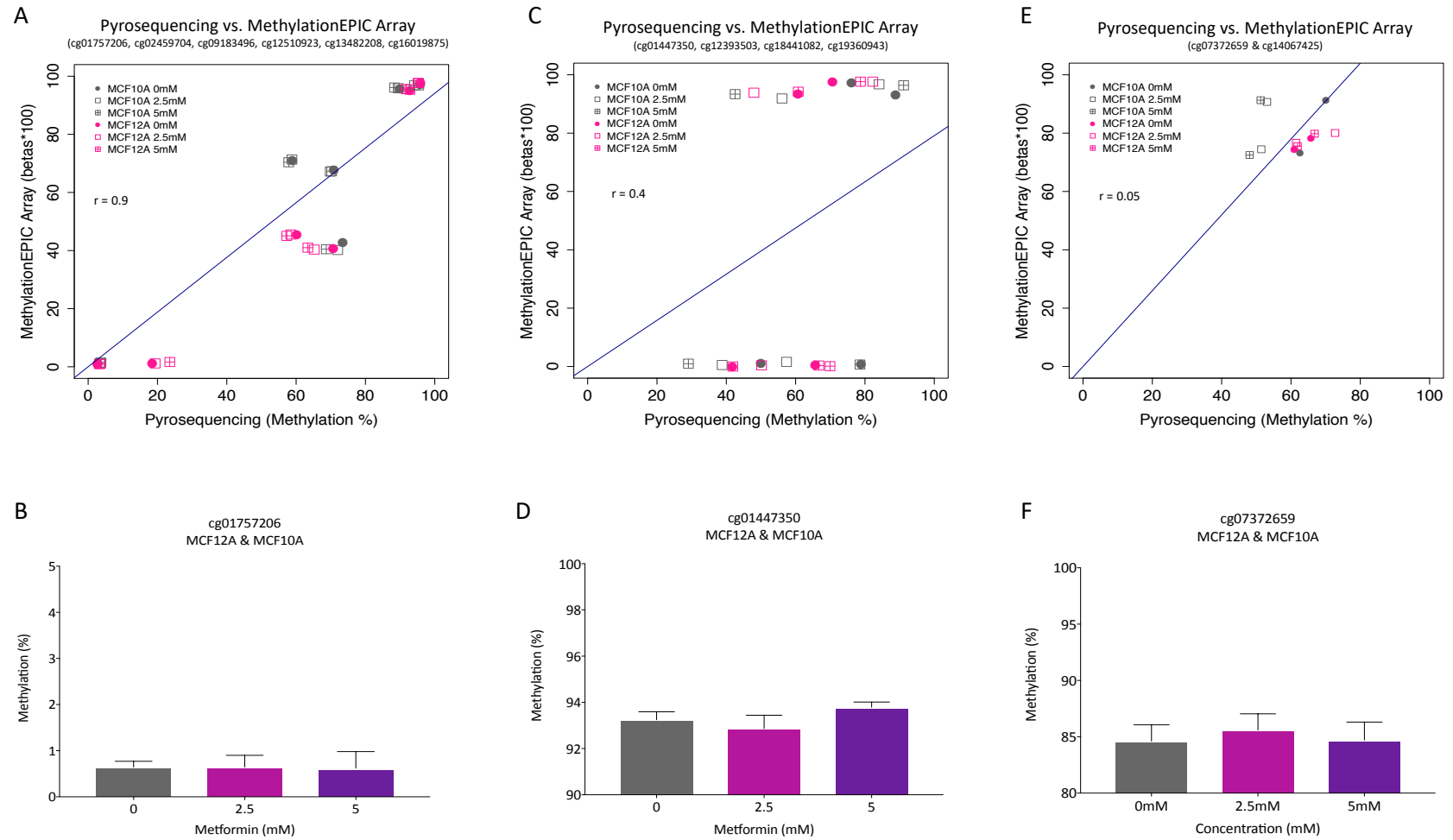


Figure 4.10 Validation of Illumina MethylationEPIC array analyses. (A) Correlation of Illumina MethylationEPIC array vs pyrosequencing model 1 ($r = 0.9$). (B) Example validation CpG site cg01757206 pyrosequencing model 1. (C) Correlation of Illumina MethylationEPIC array vs pyrosequencing model 2 ($r = 0.4$). (D)

Example validation CpG site cg01447350 pyrosequencing model 2. (E) Correlation of Illumina MethylationEPIC array vs pyrosequencing model 3 ($r = 0.05$). (F)

Example validation CpG site cg07372659 pyrosequencing model 3.

4.2.4 DNA methylation changes are associated with gene expression changes in non-cancerous breast epithelial cell line

DNA methylation played a key role in gene expression modification. For example, increased methylation at promoter areas could lead to decreased gene expression. Our data shows a shift in SAM:SAH ratios however very subtle changes to global DNA methylation. In contrast metformin appeared to induce significant gene expression changes in MCF12A and MCF10A cell lines. We hypothesised that metformin may be altering gene expression by subtly modifying DNA methylation. Comparison of RNA sequencing data analysed by featureCounts and MethylationEPIC data by Daisy Stoner, MRes in MCF12A and MCF10A. Results showed no significant association between gene expression changes and DNA methylation changes in 48 genes (identified in 3.2.4) in MCF12A (figure 4.11). In contrast, in MCF10A cell line three genes showed significant association between gene expression and DNA methylation (Wilcoxon rank test, $p < 0.05$). Further analysis of these genes – *POLA2*, *MCM9* and *TRMT10C* – showed a number of CpG sites within these genes were altered in response to metformin. However, there was variation in the gain and loss of methylation across these CpG sites and magnitude of change was very small ($< 0.5\%$ difference, figure 4.12). Analysis of alternative exon splicing (figure 4.13) in these genes – *POLA2*, *MCM9* and *TRMT10C* – showed no significant changes in exon differential splicing in MCF10A however there was significant differential exon splicing in MCF12A in these genes. This suggested that differential exon splicing may play a role in where DNA methylation occurs.

This data suggested that while metformin induces DNA methylation changes associated with gene expression of three genes, the subtle alteration in DNA methylation levels may not be sufficient to fully regulate the gene expression changes.

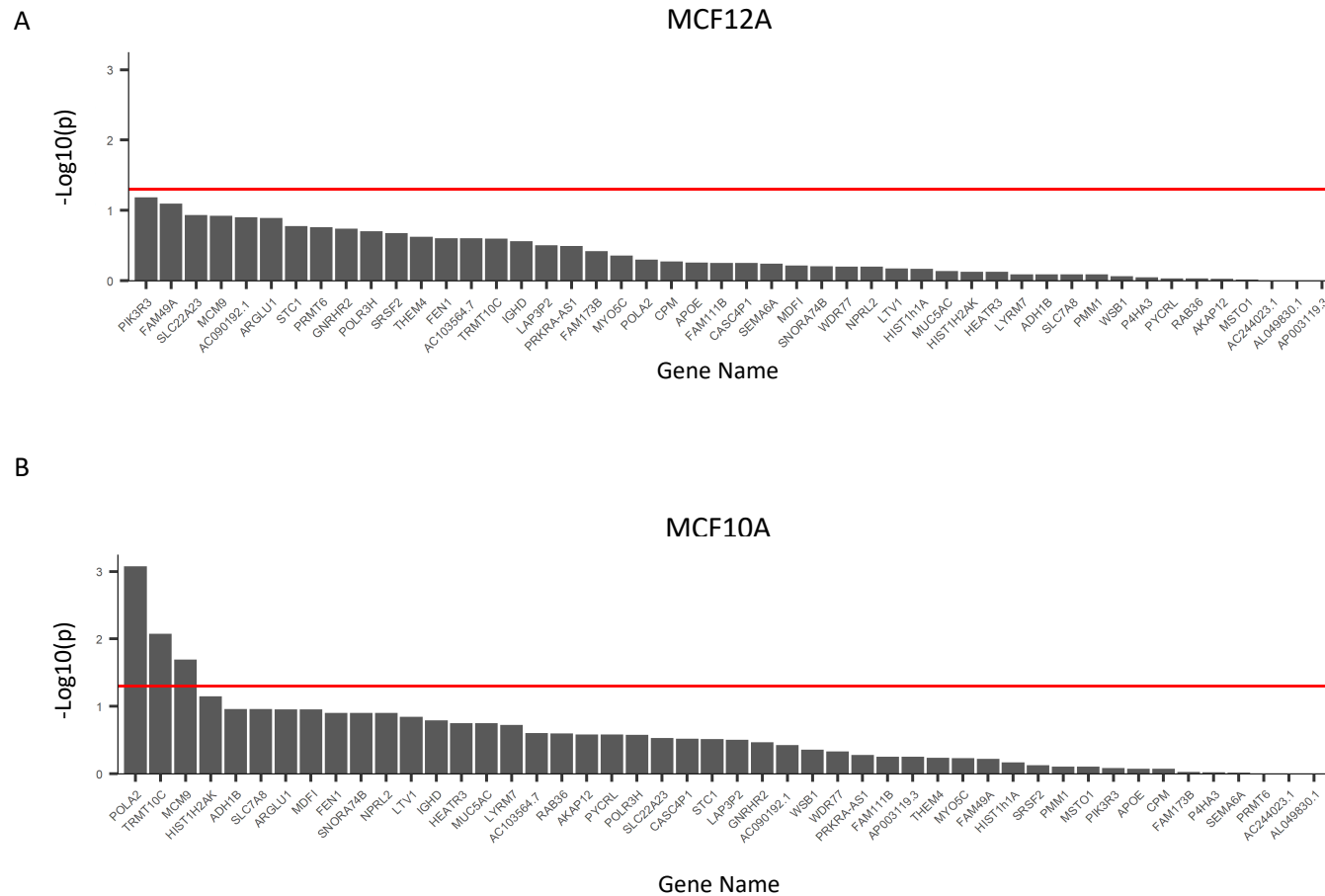


Figure 4.11 Analysis of 48 genes (FeatureCounts analysis) shown to associated with DNA methylation changes (Illumina MethylationEPIC array). (A) MCF12A non-cancerous breast epithelial cell line (Wilcoxon rank test) and (B) MCF10A non-cancerous breast epithelial cell line (Wilcoxon rank test). Red line; $p=0.05$. Data produced by Daisy Stoner, MRes.

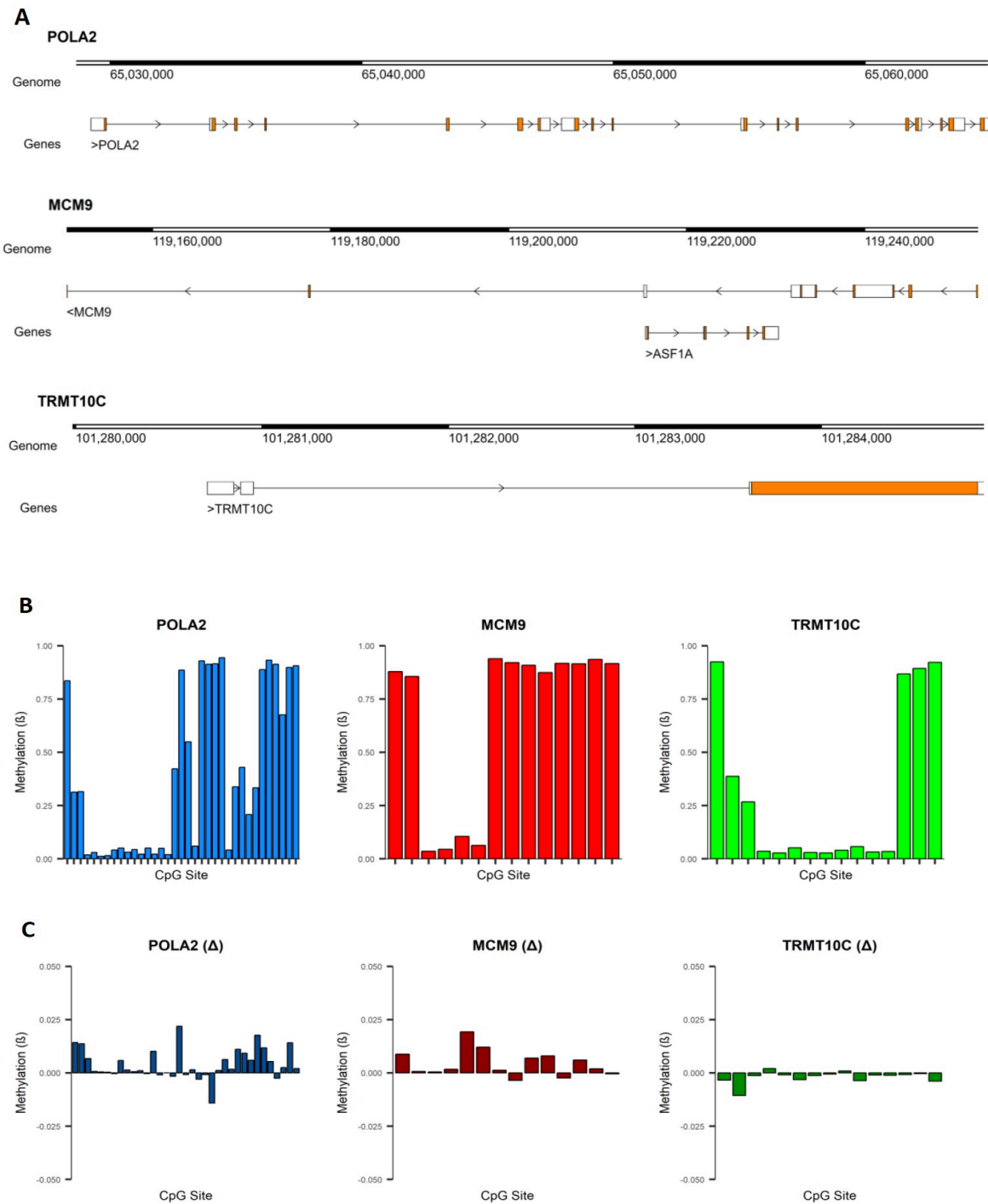


Figure 4.12 (A) Genes where gene expression changes associate with methylation changes – POLA2, MCM9 and TRMT10C (Wilcoxon test) in MCF10A non-cancerous breast epithelial cell line. White box - untranslated regions (UTRs), orange box - protein coding regions and black lines - introns. (B) Average methylation per CpG site in untreated vs treated. (C) Magnitude of DNA methylation changes per CpG site in untreated vs treated. Data produced by Daisy Stoner, MRes.

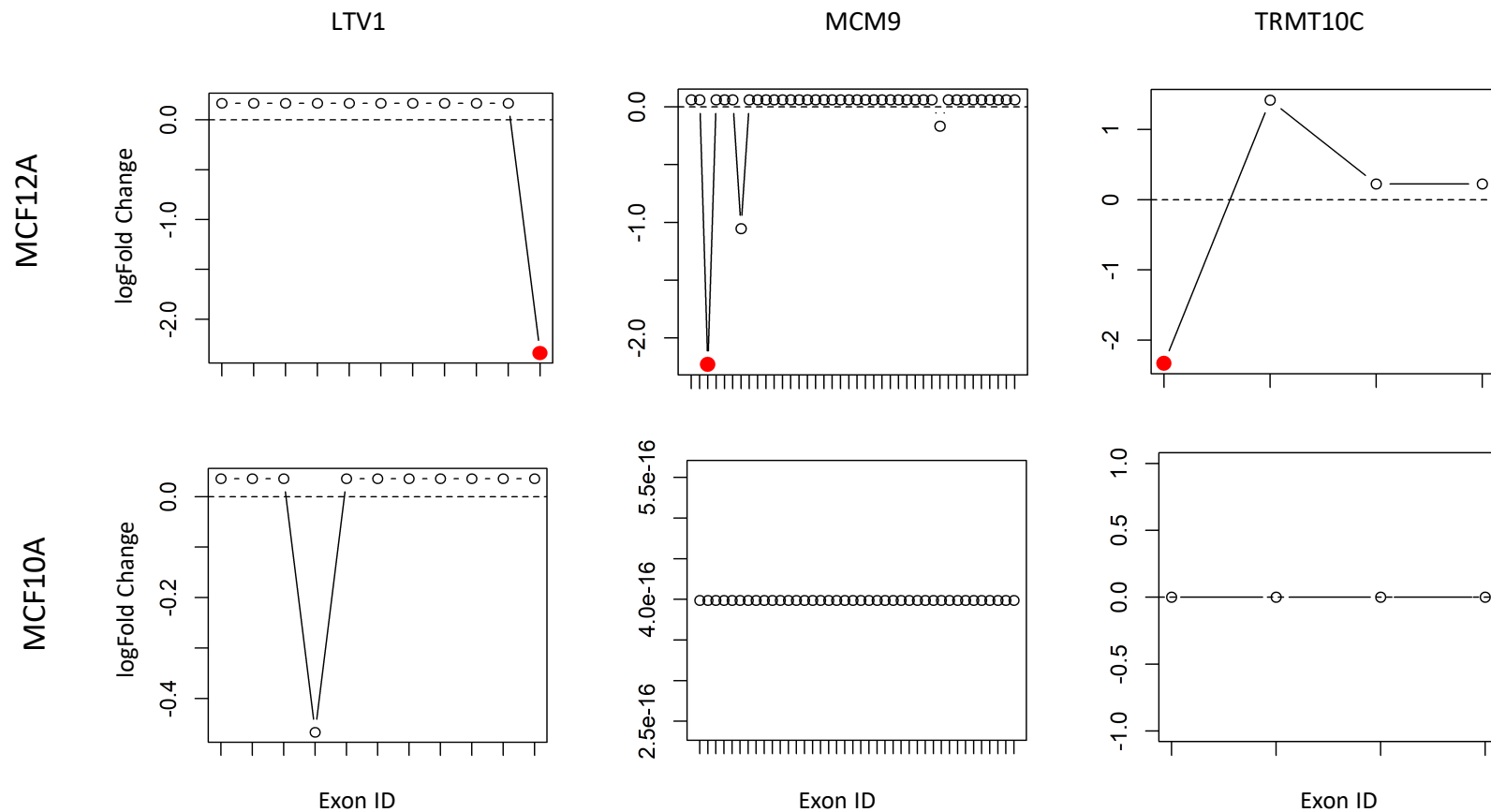


Figure 4.13 Genes showing significant changes in RNA levels and DNA methylation levels – LTV1, MCM9, TRMT10C in both MCF12A and MCF10A non-cancerous breast epithelial cell lines. These three genes were found to have significant changes in exon expression in MCF12A non-cancerous breast epithelial cell line only. Significant exons - red dot. Data produced by Daisy Stoner, MRes.

4.3 Discussion

DNA methylation patterns have previously been shown to indicate breast cancer risk. For example Brennan *et al.* have recently shown that *ATM* hypermethylation, detectable in white blood cells, can indicate increased risk of breast cancer and can be used as an important biomarker (Brennan *et al.*, 2012). Similarly, *BRCA1* promoter methylation levels could also indicate breast cancer risk (Iwamoto *et al.*, 2011). In recent years metformin has been associated with inducing DNA methylation changes. These changes have been linked to both changes in cellular functions and gene expression. However, the mechanisms by which metformin was altering DNA methylation levels was not well understood. We hypothesised that DNA methylation changes were induced by metformin and these changes could potentially drive the downstream metformin mechanism(s), for example gene expression changes.

Recent studies have shown that metformin can affect the DNA methylation of gene promoters leading to reduced gene expression, for example methylation at *H19* promoter region lowers *H19* levels in liver cells of diabetic patients leading to modification of pro-cancerous pathways (Zhong *et al.*, 2017). Our study suggested that while metformin may be affecting DNA methylation subtly, DNA methylation did not appear to be the driving force behind the significant gene expression changes seen following metformin treatment in non-cancerous breast epithelial cells. This result implies that metformin may have an indirect effect on DNA methylation in normal breast epithelial cells, leading to subtle changes. A study conducted by Cuyàs *et al.* on one-carbon metabolites and DNA methylation in metformin treated MCF10A non-cancerous breast epithelial cell line demonstrated not only an increase in global DNA methylation but also a shift in the SAM/SAH axis (Cuyàs *et al.*, 2018). The SAM/SAH axis is important in regulating the availability of methyl groups for DNA and protein methylation. This study found metformin increased SAM levels, increasing the SAM:SAH ratio and increasing DNA methylation. Results shown here demonstrate a similar effect in MCF12A and MCF10A non-cancerous breast epithelial cell lines. There was significant increase in SAM:SAH ratio in the higher doses of metformin. Further to the study by Cuyàs *et al.*, our results not only showed the shift in SAM:SAH ratio

but also significant changes in a number of metabolites in the methionine and folate pathways. This suggested metformin may be having a wider effect on cellular metabolism than our hypothesis initially proposed.

Metformin has been shown to increase GLUT1 translocation to the cellular membrane in diabetic patients (Fischer *et al.*, 1995). This increases the intake of external glucose, leading to changes in downstream pathways, including serine generation. As shown here, serine metabolite levels significantly increase in response to metformin. This may be potentially due to the increase in glucose availability. Further analysis of glucose concentrations intercellularly as well as glucose-carbon tracking would give insight into how glucose is used in the one-carbon metabolism pathway. In addition to increased serine metabolite levels there was also increased *SHMT2* gene expression in both cell lines. Knockout of *SHMT2* in Hap1 cells by Cuyàs *et al* abrogated metformin's effects (Cuyàs *et al.*, 2018). In breast cancer, *SHMT2* levels have been associated with overall survival (OS) where high protein levels of *SHMT2*, independent of conventional subtypes of breast cancer, can predict poor OS [HR=1.53, 95% CI = 1.10 – 2.12] (Bernhardt *et al.*, 2017). Modification of this enzyme, such as esiRNA knockdown, could aid in determining if modifications in folate pathway intermediates could affect methionine pathway intermediates in non-cancerous breast epithelial cells. One key consideration was the increase of threonine following metformin treatment. Serine could potentially be shuttled into pathways not related to SAM generation.

Analysis of methionine pathway also showed significant modifications. There was significant increase in methionine levels however homocysteine levels significantly decrease. Further to this the gene expression levels of enzymes required for the conversion of homocysteine to methionine, *MTR* and *MTRR*, significantly decrease. As homocysteine is required for methionine recycling within the methionine pathway, this suggested there may be alternative sources of methionine, for example serine metabolism. CRISPR gene knock-in could be used to reinstate *MTR* and *MTRR* gene expression, which may modify the effects seen by metformin.

Similar to serine, there was a significant decrease in homocysteine metabolite levels with a subsequent significant increase in taurine metabolite levels. Taurine is an important amino acid used for a number of biological processes including regulating mitochondrial function and has been implicated in cardiac and skeletal muscle health (Schaffer *et al.*, 2010). SAHH (ACLY) enzyme gene expression was significantly increased. This suggested that homocysteine was generated from SAH by SAHH but may be shuttled into taurine generation rather than being recycled back into the methionine pathway. A study by Bishnu *et al.* in chemoresistant ovarian cancer cell lines showed metformin sensitizes resistant cells to conventional treatments, potentially through to increased taurine levels in stem cell-like cells (Bishnu *et al.*, 2019). Further analysis of this pathway, potentially through inhibition of SAHH, could identify how homocysteine is being used.

While there were consistencies between gene expression data and metabolite levels, one section of the methionine pathway poses interesting questions. While SAM metabolite levels in MCF12A were significantly increased in response to metformin, the gene expression of one of the key enzymes involved in SAM generation from methionine, namely *MAT2A*, showed significant decrease in expression in MCF12A. One hypothesis could be while *MAT2A* loses expression, other enzymes in this pathway may still be functional. Analysis of *MAT2B* gene expression in MCF12A showed decreased gene expression in both doses of metformin. In contrast *MAT1A* gene expression was unchanged by metformin treatment. Further modification of this section of the methionine pathway is required to properly elucidate what is occurring. Western blot analysis of protein levels of *MAT1A*, *MAT2A* and *MAT2B* would be able to link gene expression loss with protein loss. esiRNA knockdown of these enzymes in metformin treated and untreated conditions in combination with SAM measurements may also identify if metformin induces alternative SAM generation and whether these enzymes were involved.

A limitation in this study was the variability seen in the analysis of Illumina MethylationEPIC array. This variability allowed identification of a small number of significant CpG sites however the methylation

changes associated with these sites were very small (0.5% - 1%) and may not be biologically relevant. We identified large batch effects (appendix figure 4.1) within our data after extensive filtering and pre-processing. This suggested there were technical issues or metformin may induce DNA methylation changes in a random way, leading to high biological variation. One clue to elucidate this was seen in the HILIC-MS data. The samples were generated in combined technical triplicates and individual biological triplicates in highly controlled conditions to minimise variability. Even so, in some cases there was large variability between triplicates regardless of metformin treatment. Modification of the experiment, taking into account the intrinsic biological variability in metabolite levels, could potentially reduce the effect of variability on Illumina MethylationEPIC generated data.

While it was still unclear the specific mechanisms metformin was using to modify metabolites in the one-carbon metabolism pathway, particularly in the generation of SAM leading to DNA methylation changes, this data highlighted the potential metabolic shift of this pathway metformin incurred in non-cancerous breast epithelial cells.

CHAPTER 5

5. Metformin alters histone acetylation in normal breast epithelial cells

5.1 Introduction

Histone acetylation is a key regulator of gene expression. The main donor of acetyl groups for histone acetylation is acetyl coenzyme A (acetyl co-A) (Pietrocola *et al.*, 2015). Acetyl co-A is generated in both the mitochondria and cytosol (figure 5.1). In the mitochondria acetyl co-A is generated through glycolysis where pyruvate is converted to acetyl co-A through decarboxylation by pyruvate dehydrogenase complex (PDC). In normal conditions, mitochondrial acetyl co-A continues through the TCA cycle to produce NADH. In fasting conditions, mitochondrial acetyl co-A can be converted to ketone bodies. In the cytosol, acetyl co-A is generated through glutamine reductive carboxylation to produce citrate. Citrate can be converted to acetyl co-A by ATP citrate lyase (ACLY) (Wellen *et al.*, 2009). Additionally, cytosolic acetyl co-A can be generated by conversion of acetate by acyl-CoA synthetase short chain family member 2 (ACSS2). Cytosolic acetyl co-A can be used to generate fatty acids, steroids, glutamate and amino acids (for example proline and arginine). As histone acetylation occurs in the nucleus, acetyl co-A can freely diffuse through the nuclear membrane. It can also be generated in the nucleus where nuclear PDC can convert pyruvate to acetyl co-A (Pietrocola *et al.*, 2015).

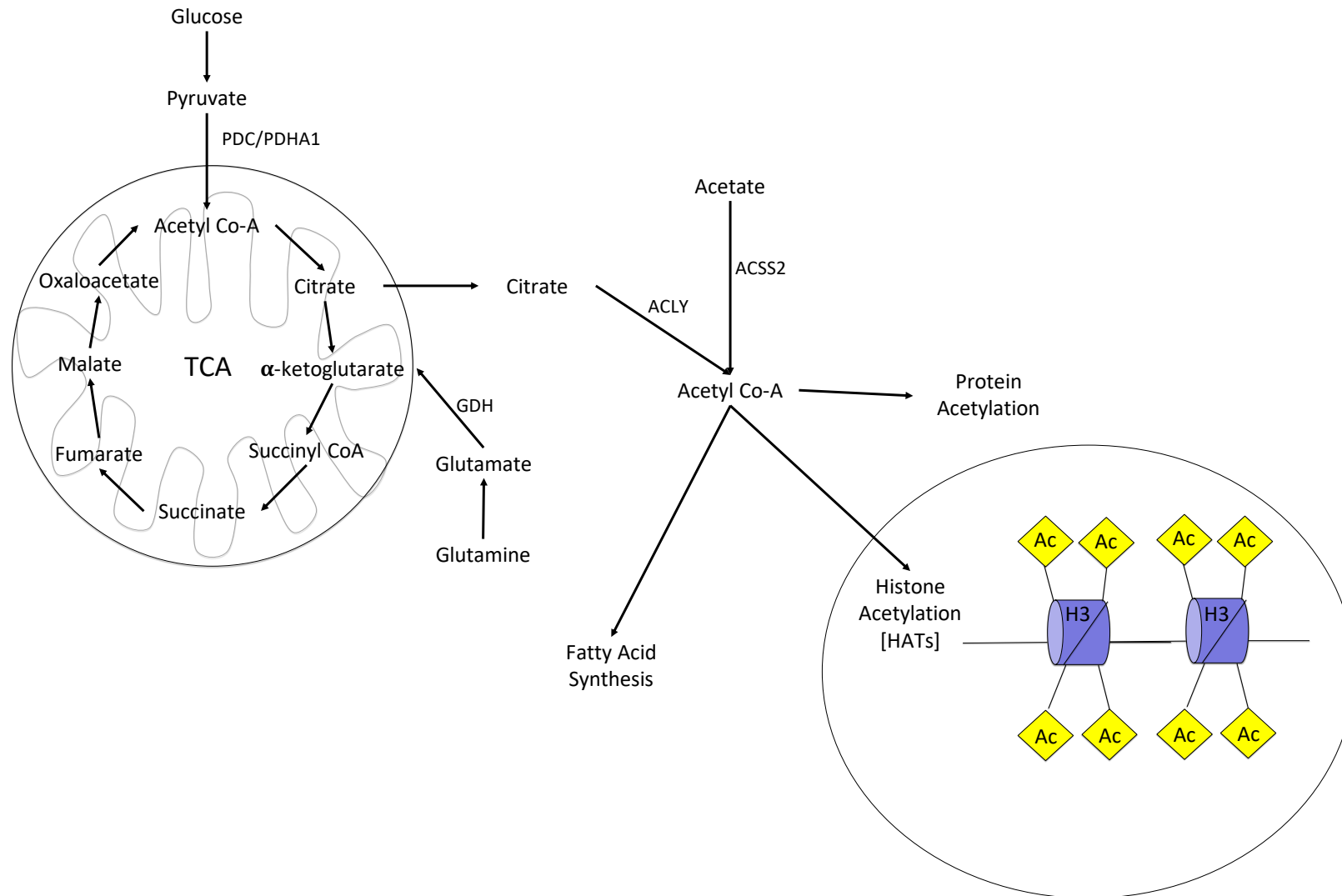


Figure 5.1 Mechanisms of acetyl co-A generation in normal breast cells [Adapted from (Wellen *et al.*, 2009; Pietrocola *et al.*, 2015)].

Histone acetylation is controlled by histone acetylation transferases (HATs) and histone deacetylases (HDACs) (Yang and Seto, 2007). One of the most widely studied histone modification is the acetylation of lysine 27 on histone 3 (H3K27ac). H3K27ac is responsible for regulating gene promoter activity as well as enhancer activity (Karlić *et al.*, 2010). The presence of H3K27ac can indicate gene expression and is used as a surrogate marker for gene expression (Creyghton *et al.*, 2010). Modifications in H3K27ac have been associated with cancer progression, for example increased H3K27ac around the TSS of oncogenes was seen in a number of cancers. In breast cancer, histone modifications have been shown to be associated with breast cancer progression. Studies have shown a gain in H3K4ac, a mark associated with active transcription at gene promoters, in cancerous breast cell lines compared to the non-cancerous breast epithelial cell line MCF10A (Messier *et al.*, 2016). Histone acetylation could identify breast tumours compared to normal breast tissue (Elsheikh *et al.*, 2009). For example, H4K16ac showed low levels in breast tumours compared to normal breast tissue. In contrast H3K18ac showed high levels. Additionally, low levels of H4K16ac and H3K9ac were shown to correlate with tumour size and vascular invasion whereas high levels of H3K9ac were associated with better tumour outcomes.

Metformin has also been linked to histone acetylation. Firstly, it has been shown to increase HAT1 activity (Marin *et al.*, 2017). Metformin could also activate SIRT1, a class III HDAC (Caton *et al.*, 2010). Increased SIRT1 has been associated with decreased serum glucose and improved insulin sensitivity in mouse models (Milne *et al.*, 2007). In contrast, metformin induced AMPK activation was associated with reduced class II HDAC activation, for example HDAC5 deactivation was associated with increased GLUT4 expression (McGee *et al.*, 2008; Galdieri *et al.*, 2016). GLUT4 is an important glucose transporter and is also regulated by metformin. Translocation of GLUT4 to cellular membrane induced by metformin is an important mechanism in inhibiting gluconeogenesis and reinstating insulin signalling in T2DM (Giannarelli *et al.*, 2003; McGee *et al.*, 2008; Lee *et al.*, 2012). Therefore, if

metformin was modifying histone acetylation this could provide a mechanism by which cancer prevention can occur.

In this chapter we explore the hypothesis that metformin may be affecting histone acetylation in non-cancerous breast epithelial cells.

5.2 Results

5.2.1 Metformin reduces total H3K27ac in normal breast epithelial cells

H3K27ac was one of the most commonly studied histone PTM in relation to gene expression changes.

Our study has shown that metformin may be significantly altering gene expression in non-cancerous breast epithelial cells and that DNA methylation modifications may be potentially contributing to some but not all of these changes. Our study has also shown that metformin may affect metabolic pathways associated with epigenetic modifications. We hypothesised total H3K27ac levels may be modified with metformin treatment, potentially through modification of the acetyl co-A pathway.

To assess whether metformin potentially modifies this histone PTM mark, we conducted western blot analysis for H3K27ac and total H3 in MCF12A and MCF10A non-cancerous breast epithelial cells lines (figure 5.2). Cell lines were treated with 0mM, 2.5mM and 5mM metformin for 72 hours. In MCF12A there was a significant decrease in total H3K27ac at both doses with no significant reduction in total H3, quantified over biological triplicates (figure 5.2, student's t-test, $p < 0.05$). In contrast MCF10A, while showing a trend for reduced H3K27ac, showed no significant reduction in H3K27ac. When combined together (figure 5.2 E), metformin significantly decreases H3K27ac in both doses, where 5mM induces 0.5-fold decrease in H3K27ac ($p = 0.035$).

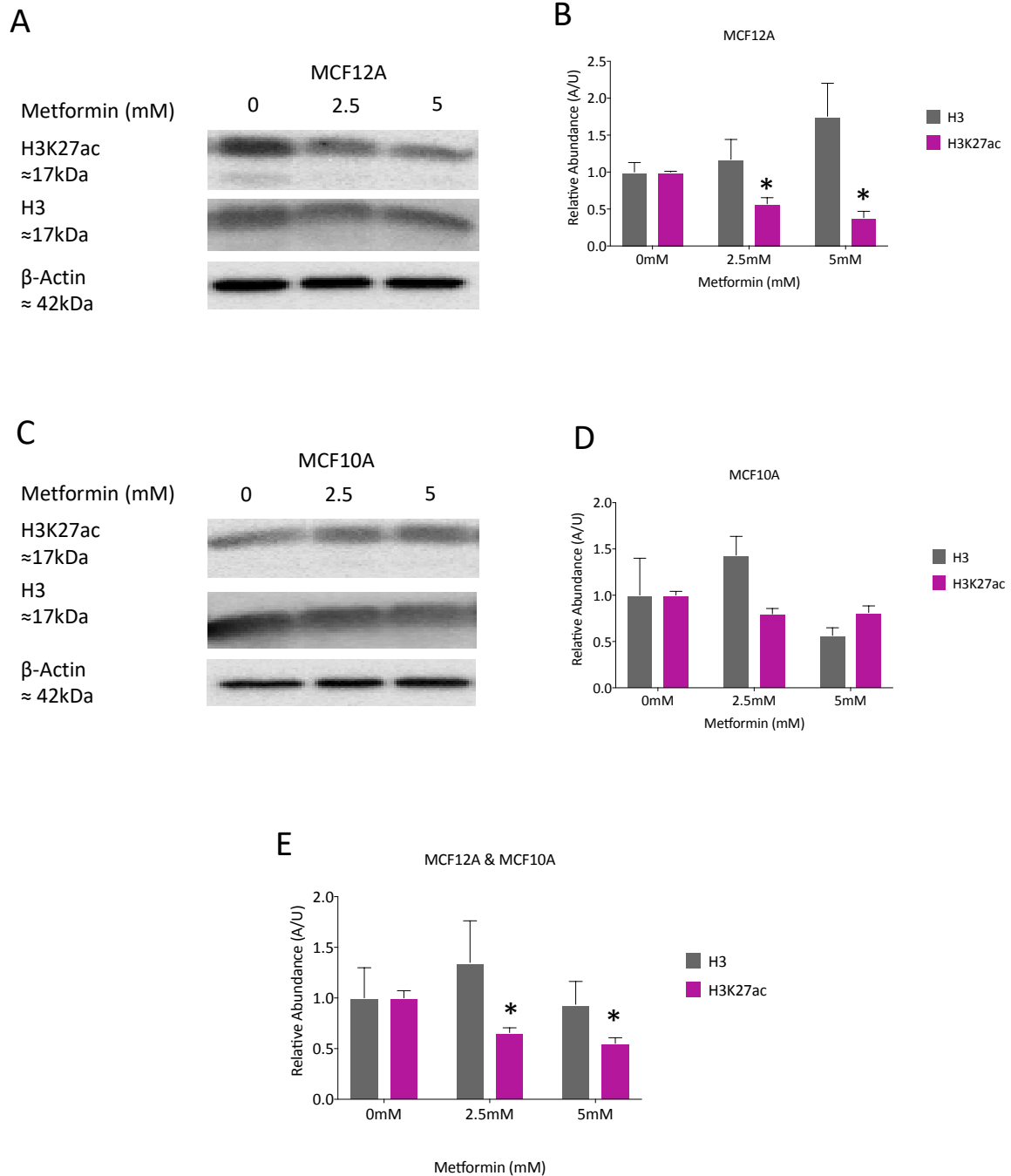


Figure 5.2 (A-B) Western blot of H3K27ac and H3 in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. Relative abundance measured as ratio to beta actin (n=3) (C-D) Western blot of H3K27ac and H3 in MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. Relative abundance measured as ratio to beta actin (n=3) (E) Relative abundance measurement of H3K27ac and H3 in MCF12A and MCF10A

averaged over six biological replicates (n = 6). Relative abundance calculated as ratio to beta actin.

Student's t-test * ≤ 0.05 .

To validate this, we conducted western blot for H3K27ac and total H3 in three matched primary breast epithelial and myoepithelial patient samples (beta actin not obtained). While there was a trend in individual patients (n=1), H3K27ac levels did not significantly decrease in either epithelial or myoepithelial cells, potentially due to the large variability between patients (figure 5.3). For example, patient 3002's epithelial cells showed significant decrease in H3K27ac, however, the myoepithelial cells from the same patient showed less loss of H3K27ac. This may potentially be due to the levels of H3K27ac in untreated samples in each cell type.

This data suggested that metformin may be affecting histone post-translational modifications, however, the mechanisms by which metformin was influencing histone acetylation was not fully known.

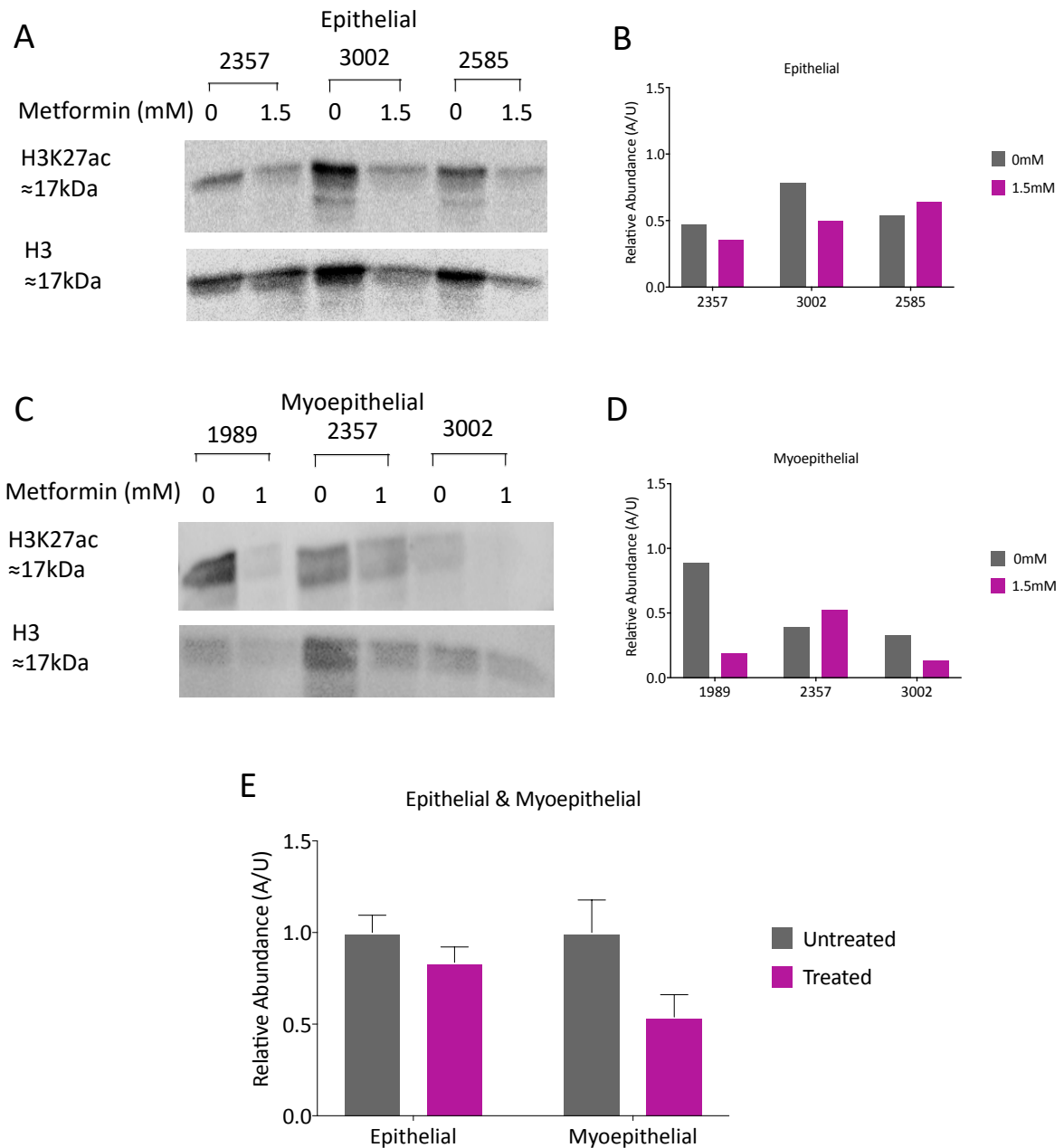


Figure 5.3 Western blot of H3K27ac and H3 in three primary normal matched breast epithelial and myoepithelial cells treated 7-days metformin. (A-B) Three primary normal epithelial cells treated with 0mM and 1.5mM for 7 days (n=1). (C-D) Three primary normal myoepithelial cells treated with 0mM and 1mM for 7 days (n=1). (E) Relative abundance measurement of epithelial vs myoepithelial cells averaged over three patients (n = 3). Relative abundance calculated as ratio to H3.

5.2.2 H3K27ac levels are restored 24 hours following metformin treatment

Studies have shown that histone PTM marks can be flexible (Clayton, Hazzalin and Mahadevan, 2006; Greer and Shi, 2012). With multiple readers, writers and erasers for histone PTMs there could potentially be a high turnover of epigenetic marks. To study how quickly metformin modified H3K27ac and whether constant metformin treatment was required to maintain H3K27ac loss we conducted H3K27ac tracking in MCF12A. Cells were treated for 72 hours with 0mM, 2.5mM and 5mM metformin and a further 72 hours (144 hours in total) without metformin. Histone protein was extracted every 24 hours and western blot for H3K27ac and H3 were conducted. Figure 5.4 showed total H3K27ac levels appeared to decrease following 48 hours treatment with metformin at both doses, however, H3K27ac levels returned to baseline levels following 24 hours of no treatment with metformin. This suggested that non-cancerous breast epithelial cells require constant treatment with metformin to maintain the decrease in total H3K27ac levels.

Total H3 appeared to follow a similar pattern when treated with metformin however this was not seen in every western blot conducted in non-cancerous breast epithelial cell lines or primary cells. This may potentially be due to global changes to H3 associated with metformin rather than just loss of H3K27ac. Further analysis would be required to identify the cause of total H3 decrease.

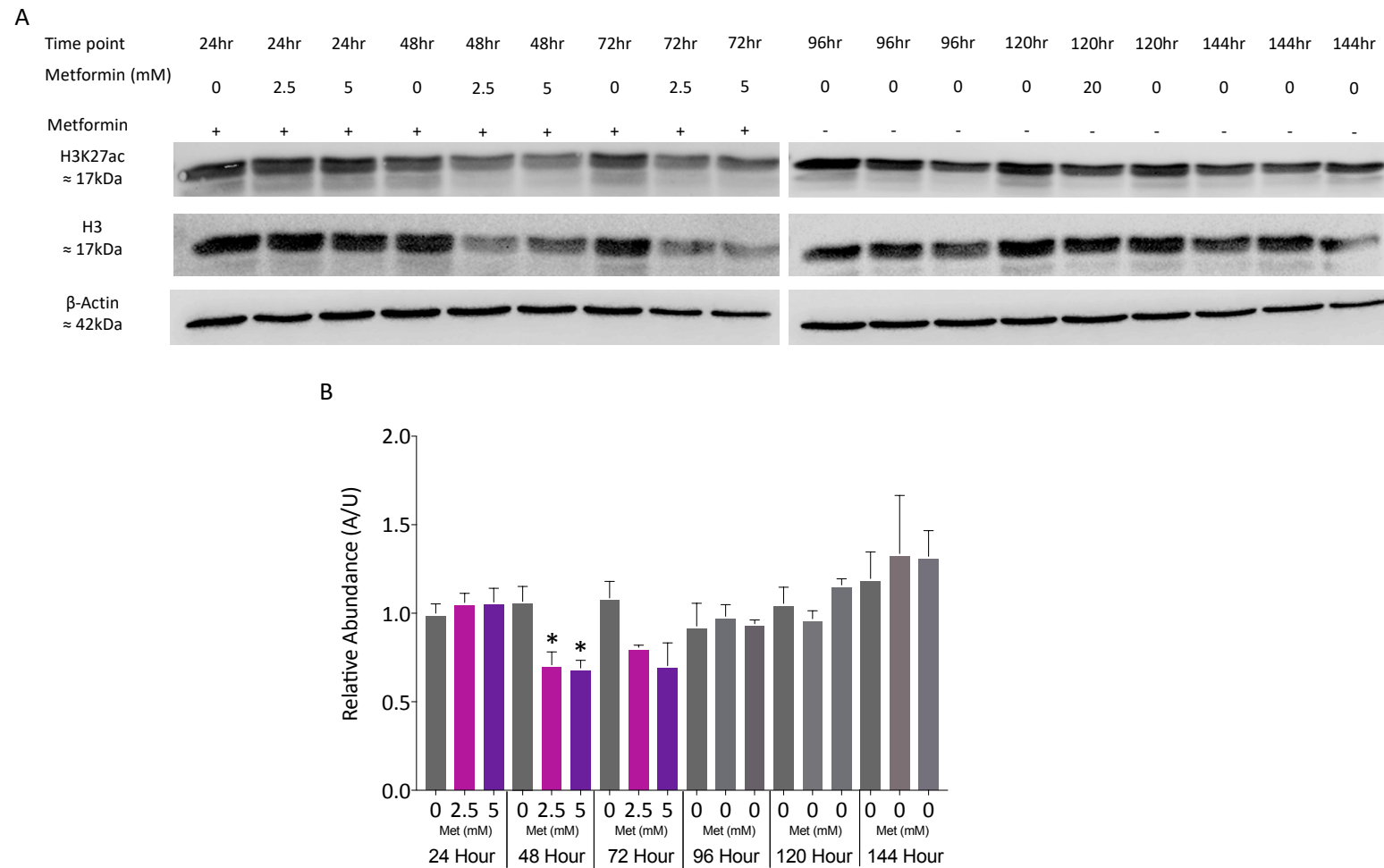


Figure 5.4 (A) Western blot of H3K27ac and H3 in MCF12A non-cancerous breast epithelial cell lines treated 144 hours with 0mM, 2.5mM and 5mM metformin.

(B) Relative abundance measurement. Averaged over three biological replicates (n = 3). Relative abundance calculated as ratio to H3. Student's t-test * ≤ 0.05 .

5.2.3 Acetate supplementation could potentially rescue H3K27ac levels in metformin-treated normal breast epithelial cells

Under hypoxic conditions the normal generation of acetyl co-A through the TCA cycle and citrate metabolism by ACLY was switched to acetate metabolism by ACSS2 (Schug *et al.*, 2015). As acetyl co-A was the acetyl donor for H3K27ac we hypothesised that acetate supplementation could potentially alter metformin's effects. To assess this in non-cancerous breast epithelial cell lines, cell culture media was supplemented with acetate. Traditional cell culture media for MCF12A and MCF10A cell lines did not contain acetate. Initially we assessed whether acetate supplementation had an impact on cell viability and conducted crystal violet cell viability assays in MCF12A and MCF10A. Cell lines were treated with 0mM, 2.5mM and 5mM metformin for 72 hours in a range of acetate concentrations (0mM – 20mM). Figure 5.5 showed that while there was significant reduction in cell viability when treated with metformin, certain concentration(s) of acetate did not significantly modify this effect. As it was believed the impact on cell viability may be due to p-AMPK activation, further investigation into this pathway was conducted by western blot analysis of p-AMPK and AMPK. MCF12A and MCF10A cell lines were treated for 72 hours in 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Figure 5.6 showed that while p-AMPK was activated upon metformin treatment, acetate supplementation did not significantly alter this activation.

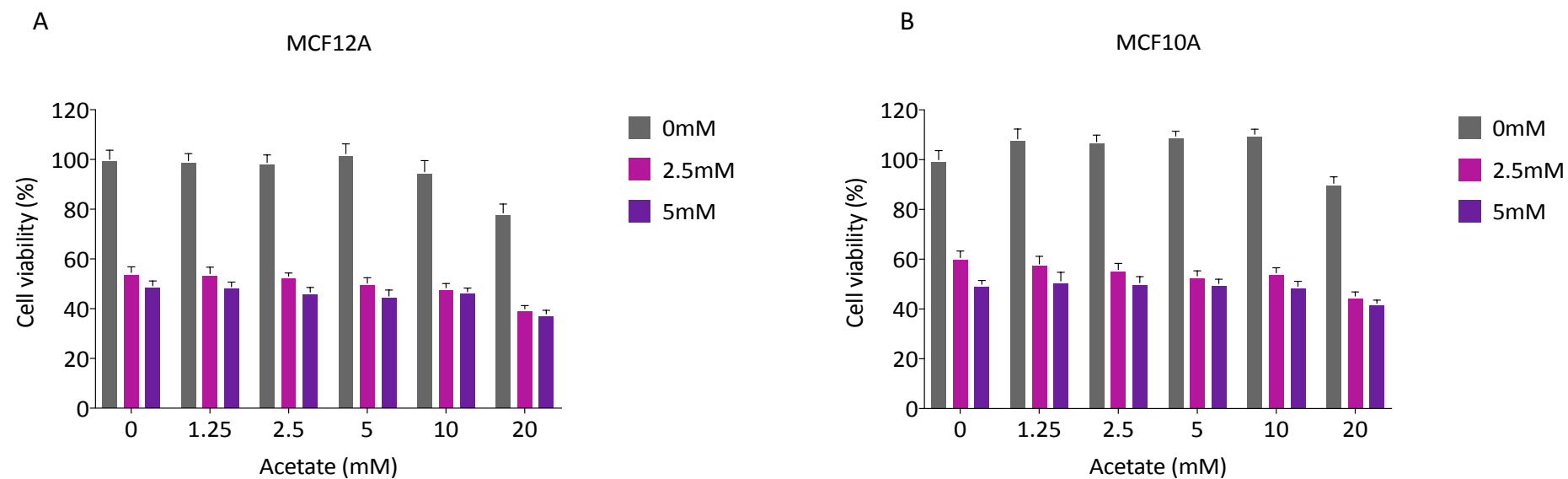


Figure 5.5 Crystal violet cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 1.25mM, 2.5mM, 5mM, 10mM, 20mM acetate.

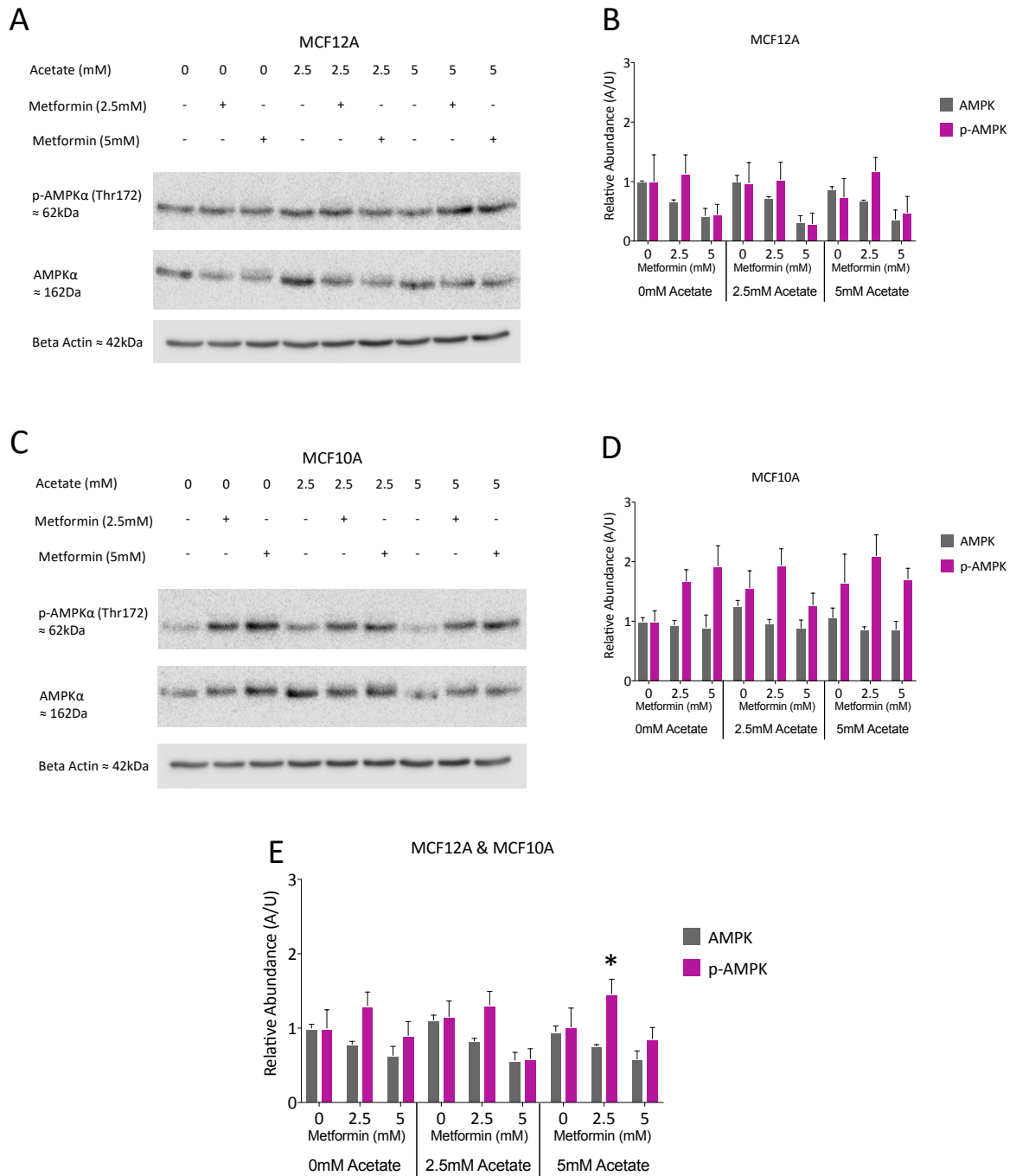


Figure 5.6 (A-B) Western blot of p-AMPK and AMPK in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. (C-D) Western blot of p-AMPK and AMPK in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. (F) Relative abundance measurement of MCF12A and MCF10A averaged over six biological replicates (n = 6).

Relative abundance calculated as ratio to loading control beta actin. Student's t-test * ≤ 0.05 Data produced by Fotini Poyia, MRes.

As acetate can be converted to acetyl co-A, independent of the TCA cycle, we conducted western blot analysis of H3K27ac and H3 in MCF12A and MCF10A cell lines treated for 72 hours in 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Figure 5.7 showed in the absence of acetate metformin reduces H3K27ac levels in both doses, however when acetate was added H3K27ac did not decrease and could in some cases increase. This suggested that acetate could potentially replenish the acetyl co-A pool required for histone acetylation. Overall this data suggested that metformin may be reducing acetyl co-A pools potentially through modification of TCA cycle metabolites that contribute to acetyl co-A generation. Supplementation of this pathway allowed partial H3K27ac rescue in metformin treated normal breast epithelial cells.

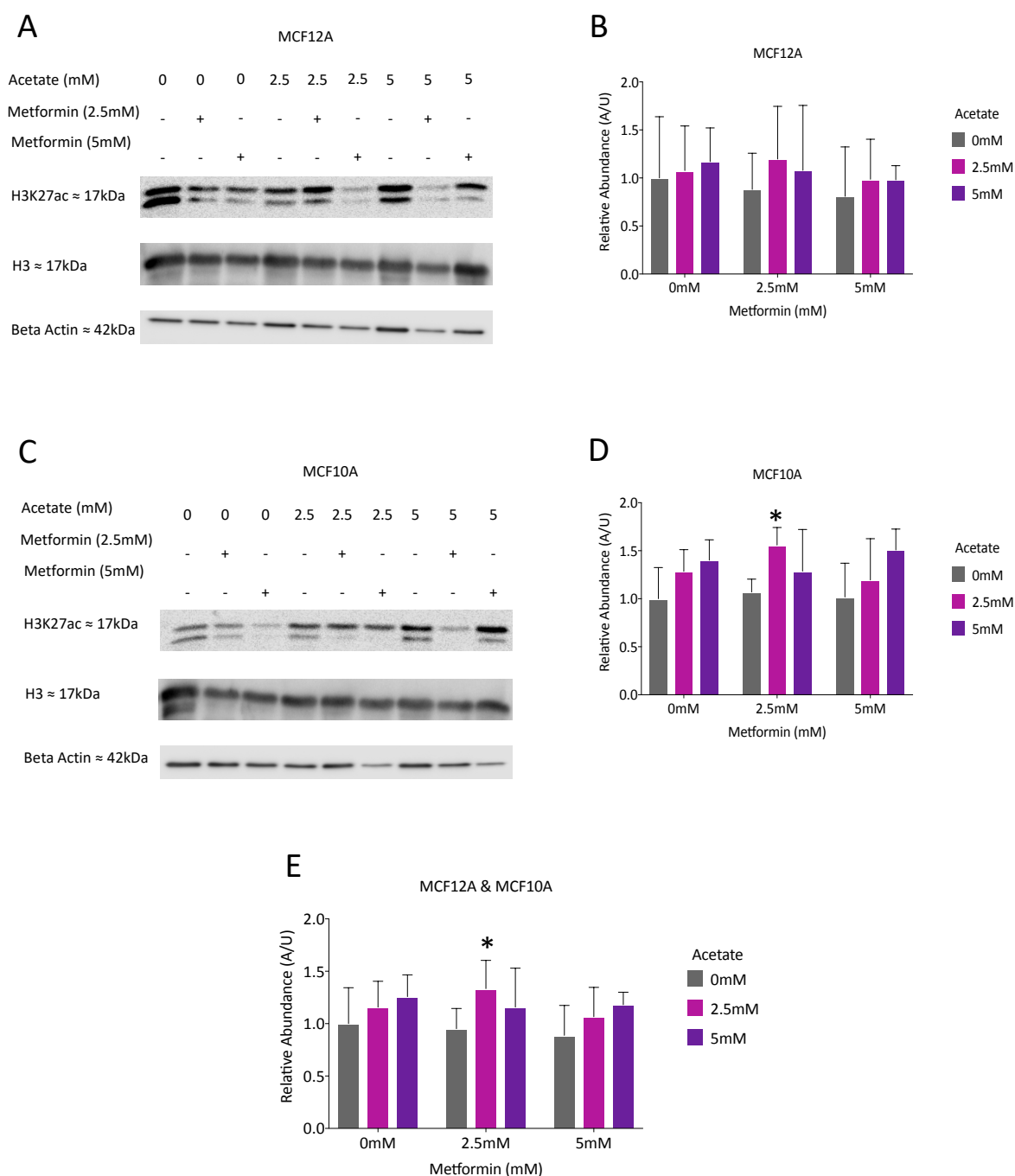


Figure 5.7 (A-B) Western blot of H3K27ac in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Relative abundance measurement calculated as ratio to H3 (n=3). (C-D) Western blot of H3K27ac in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Relative abundance measurement calculated as ratio to H3 (n=3). (F) Relative abundance measurement of MCF12A and MCF10A averaged over six biological

replicates (n = 6). Relative abundance calculated as ratio to H3. Student's t-test * ≤ 0.05 . Data produced by Fotini Poyia, MRes.

5.2.4 esiRNA knockdown of ACSS2 enzyme can replicate metformin's reduction of H3K27ac in the presence of acetate

To further test the hypothesis that acetate could replenish acetyl co-A pools, we chose to knockdown *AceCS1* (ACSS2). As previously mentioned, ACSS2 was the key enzyme that converts acetate to acetyl co-A in the cytosol. Knockdown was conducted using esiRNA, a pool of siRNAs against *AceCS1* (ACSS2), obtained from Sigma Aldrich. MCF12A cells were treated for 72hours in 0mM or 5mM acetate and 5mM metformin with or without 10nM esiRNA for *AceCS1* (ACSS2) followed by western blot analysis of H3K27ac and H3. Figure 5.8 showed reduced ACSS2 protein levels when treated with esiRNA for *AceCS1*. Figure 5.9 showed in no esiRNA conditions metformin reduces H3K27ac levels in acetate negative conditions, however, acetate partially rescued this upon incubation of cells in acetate rich media. In the esiRNA treated cells metformin reduced H3K27ac in acetate negative media, however, in acetate rich media H3K27ac levels remained low, suggesting the acetyl co-A pool was not replenished.

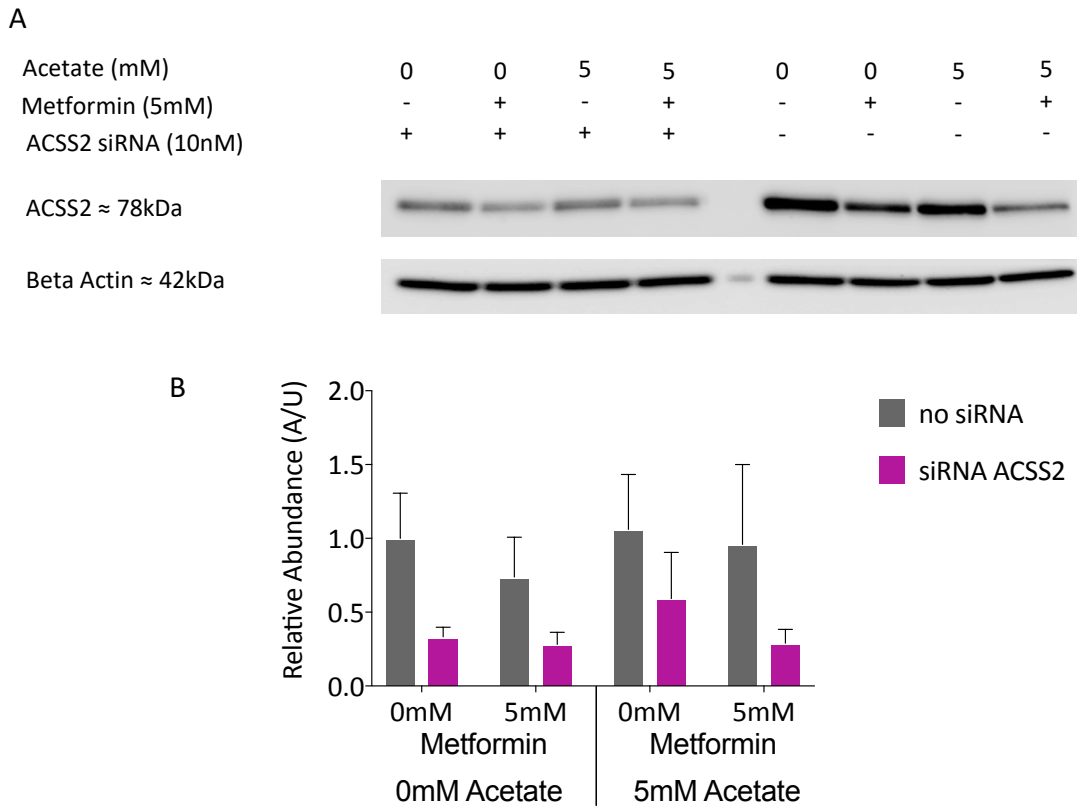


Figure 5.8 Western blot of ACSS2 in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Treated with 10nm siRNA for AceS1 (ACSS2). (B) Relative abundance measurement averaged over three biological replicates ($n = 3$). Relative abundance calculated as ratio to loading control beta actin. Data produced by Fotini Poyia, MRes.

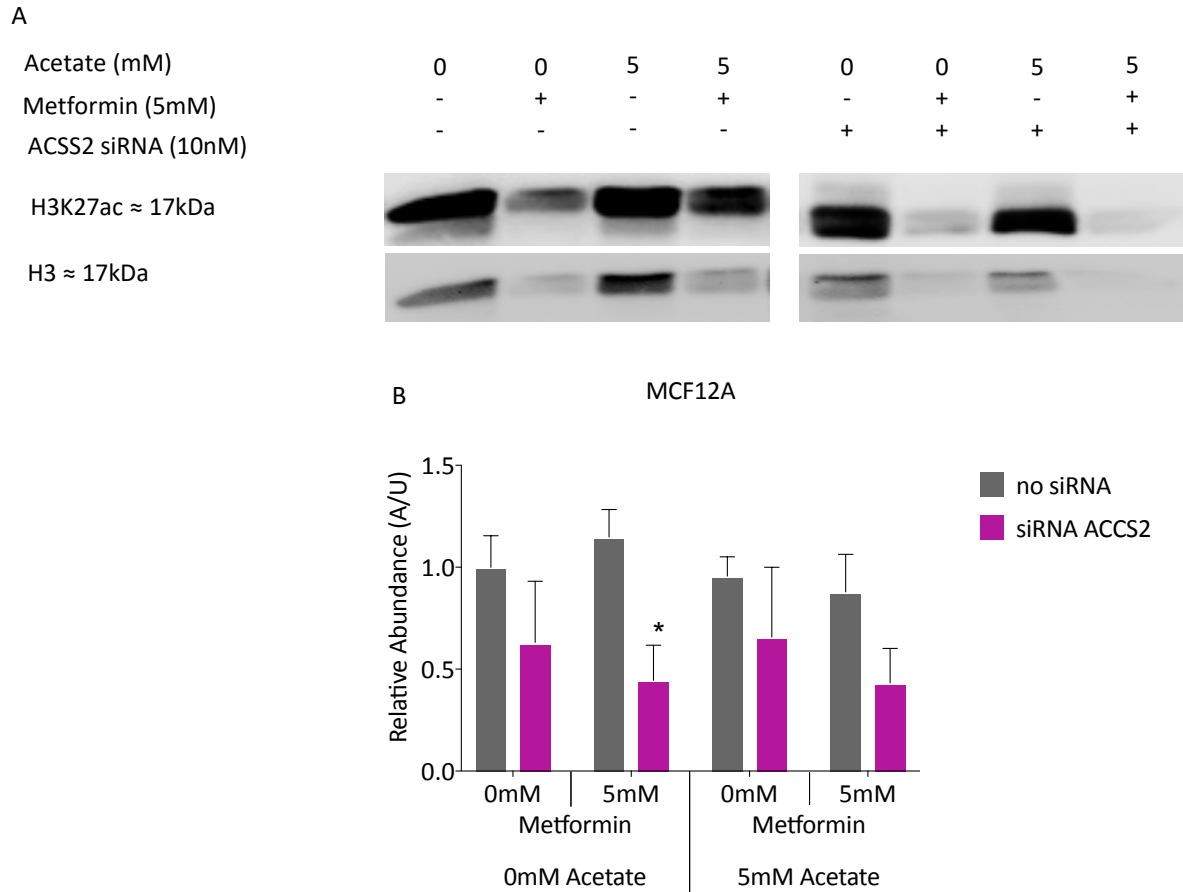


Figure 5.9 (A) Western blot of H3K27ac in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Treated with 10nm siRNA for AceS1 (ACSS2). (B) Relative abundance measurement averaged over three biological replicates (n = 3). Relative abundance calculated as ratio to H3. Student's t-test * $\leq$ 0.05.

5.2.5 Genes expression of genes associated with acetyl co-A generation pathways are altered by metformin

As shown previously, metformin can alter gene expression of a number of pathways associated with metabolism. To further assess the impact of metformin treatment on acetyl co-A metabolism, RNA sequencing data was analysed for key enzymes in this pathway. In figure 5.10, there was a significant increase in expression of ACLY ($p < 0.05$), the enzyme that converts citrate from the TCA cycle into acetyl co-A. Similarly, there was a significant increase in ACSS2 enzyme gene expression ($p < 0.05$). This has previously been shown in this study to be important in maintaining acetyl co-A pools in metformin-treated non-cancerous breast epithelial cells. There was no statistically significant change in the gene expression of the enzyme PDHA1, an enzyme that converts pyruvate to acetyl co-A.

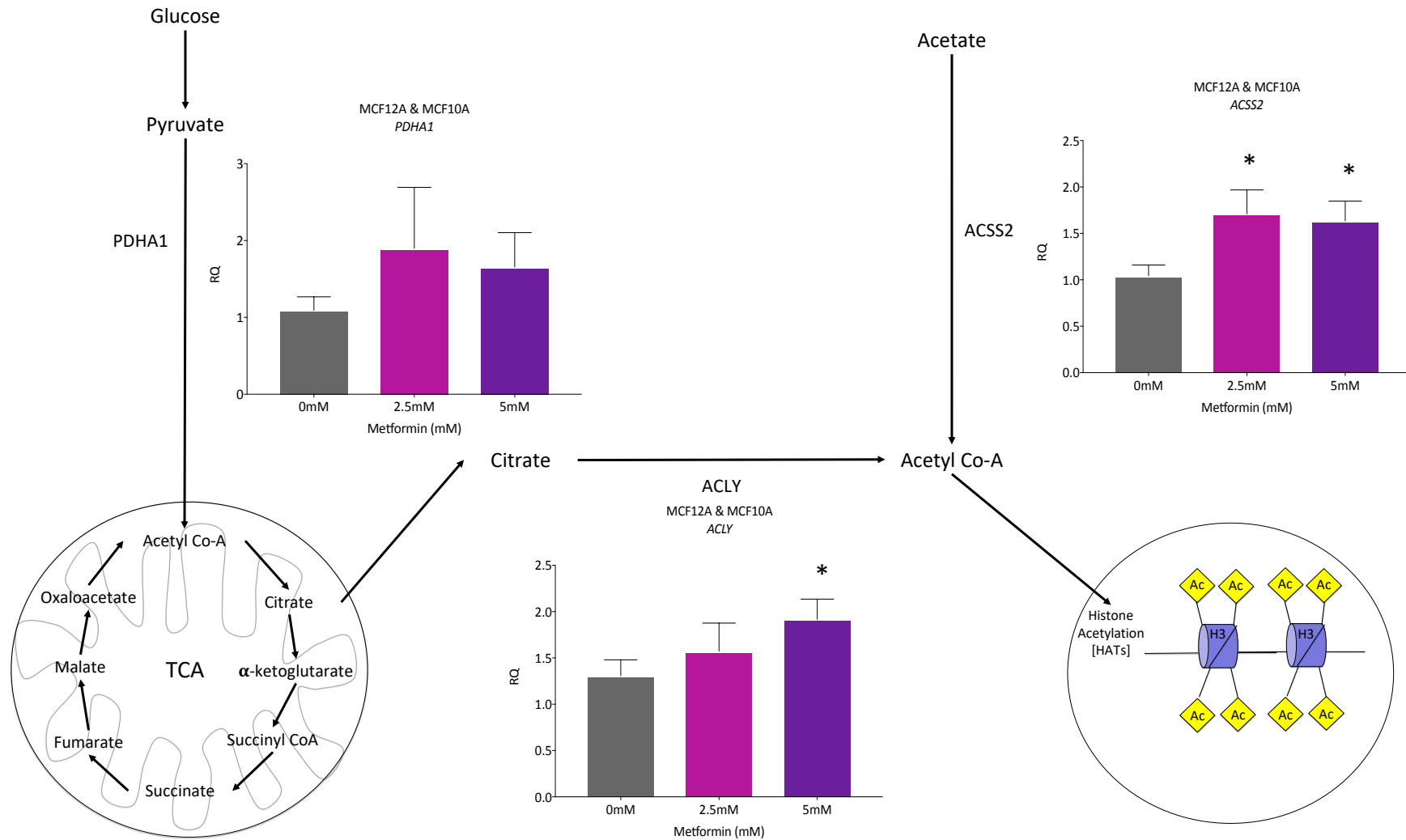


Figure 5.10 qPCR RQ values of acetyl co-A pathway related genes – *ACSS2*, *ACLY* and *PDHA1* – in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. Averaged over six biological replicates (n = 6).

5.3 Discussion

Metformin has been shown to affect a number of histone modifications in breast cancer cells, including acetylation. In breast cancer models, acetyl co-A was generated during normoxic conditions through the TCA cycle and citrate metabolism. During cancer progression, acetyl co-A biosynthesis shifts dramatically. Under hypoxic conditions glucose utilization decreases however fatty acid production via acetyl co-A was maintained (Schug *et al.*, 2015). This suggested an alternative method of acetyl co-A generation that did not depend on glucose. Acetate, an alternative source for acetyl co-A production, has been shown to be consumed by cancer cells under hypoxia (Kamphorst *et al.*, 2014). Modifications in acetyl co-A generating pathways has been shown to alter histone acetylation. Wellen *et al.* have shown modifications of ACLY can decrease total histone acetylation of core histones. As previously discussed, in normal conditions ACLY generates acetyl co-A from citrate where in metabolite restricted conditions acetyl co-A can be generated from acetate. Similar to our study, Wellen *et al.* have shown acetate supplementation in ACLY deficient cells can rescue histone acetylation levels. Additionally, this study showed inhibition of histone deacetylase enzymes (HDACs) could potentially rescue histone acetylation in ACLY deficient cells (Wellen *et al.*, 2009; Zhao *et al.*, 2016). Results from our study indicate modification of acetate availability could overcome metformin induced H3K27ac loss and perturbation of this pathway through knockdown of ACSS2 could mimic metformin's effects. ACSS2 has been shown to be activated by AMPK in glucose deprived glioma cells (Li *et al.*, 2017). It would be interesting to assess whether glucose deprivation could potentially modify metformin's effects. Additionally, investigation of MCF12A and MCF10A cell line response to metformin under hypoxic conditions may give more insight into this interplay.

The potential loss of H3K27ac in non-cancerous breast epithelial cells in response to metformin could provide another pathway in which metformin could modify gene expression. This may be contrary to what was seen in diabetic liver samples, where metformin induces increase H3K27ac around the transcription start site (TSS) of a number of metformin-associated gene promoters (Luizon *et al.*,

2016). This study by Luizon *et al.*, however, only focused on samples taken from one patient. As we have shown in previous chapters, metformin response varies between patients. This was a potential limiting factor for this study. It would be interesting to conduct ChIP-seq analysis of H3K27ac in MCF12A and MCF10A cells treated for 72hours with 0mM, 2.5mM and 5mM metformin to see where these H3K27ac changes occur. This could be validated with ChIP-PCR in primary matched normal breast epithelial and myoepithelial cells used previously. This may provide insight into where metformin is affecting histone acetylation.

In this study we demonstrated that H3K27ac decreased in relation to total H3, however, in some cases total H3 itself also decreases. This may be due to technical and/or biological variation or potentially through a more global effect of metformin on histones. It would be important to study the impact of metformin on total histone levels of H2A, H2B and H4 as well as the total acetylation levels of these histones. Western blot data could be combined with ChIP-seq data to determine if metformin is affecting histone acetylation and if this was global. H3K27ac decrease appeared to be dependent on continuous metformin treatment. This provides an interesting novel theory as to breast cancer prevention in diabetic women. A number of epidemiological studies into metformin's role in breast cancer risk have shown little to no association (Franciosi *et al.*, 2013; Qiu *et al.*, 2013). These studies designate metformin treatment as a yes or no condition. Studies that have separated patients by treatment time have shown decreased breast cancer risk associated with longer metformin treatment.

A limitation of this study was the 72-hour timepoint. This study could be taken further by treating non-cancerous breast epithelial cell lines with metformin for an extended period, for example 3-6 months, with metformin and then a similar time period without metformin. H3K27ac tracking could be used as a surrogate marker for metformin response as well as assessing whether metformin's effects on histone acetylation are time dependent.

Enzymes involved in the generation of acetyl co-A showed significant upregulation in metformin treated cells. It was unknown how this affected total acetyl co-A pools and whether these gene expression changes were caused by metformin or were in response to low acetyl co-A levels. We have shown that perturbation of one pathway leading to acetyl co-A generation, namely acetate metabolism, could potentially replicate results of metformin alone. Acetyl co-A levels could be measured in metformin treated cells to provide a baseline and when these enzymes are modified. This data combined with acetyl co-A measurements could potentially provide more insight into this pathway. Carbon tracking through these pathways could additionally highlight how acetyl co-A is being used and whether these cells are sacrificing certain pathways for other pro-growth pathways. As shown previously, to maintain H3K27ac loss induced by metformin, cells need to be continuously treated. Removal of metformin allowed H3K27ac levels to return to untreated levels. This could potentially be in part due to upregulation of acetyl co-A generating enzymes. Monitoring of enzyme levels over a prolonged period of metformin treatment may answer whether enzyme upregulation is in response to metformin or not. Overlap of RNA-seq and ChIP-seq data would further help elucidate whether gene expression changes were associated with histone modifications and whether metformin is driving these changes. This has been shown in other studies to identify AMPK-dependent and AMPK-intendent gene expression changes, where there is an increase of H3K27ac around the TSS of AMPK-dependent genes (Luizon *et al.*, 2016).

A number of other pathways associated with this study could also be explored. Acetyl co-A is not only used for histone acetylation but can be used for fatty acid synthesis. Studies in heart muscle have shown that p-AMPK activation can inhibit acetyl co-A carboxylase (ACC) in fatty acid synthesis preventing malonyl co-A generation from acetyl co-A (Hopkins, Dyck and Lopaschuk, 2003). We hypothesise that metformin could potentially modify fatty acid synthesis, which could be acetyl co-A dependent. One way to test this hypothesis to perturb the fatty acid synthesis pathway through

esiRNA of ACC without metformin treatment to assess whether this recapitulates what was seen in metformin treated conditions. Further modification of acetyl co-A availability could measure the dependency of fatty acid synthesis on acetyl co-A pools and how metformin plays a role in this pathway. Branched chain amino acids (BCAAs) – valine, isoleucine and leucine – are essential amino acids used for a number of pathways including glucose metabolism, a pathway known to be modified by metformin (Adeva-Andany *et al.*, 2017). Metformin has been shown to inhibit BCAA metabolism in mouse models of Maple Syrup Urine Disease (MSUD) through decrease of ketoisocaproic acid (KIC). We hypothesise that metformin may alter BCAA function and one way to test this would be to modify BCAA levels in MCF12A and MCF10A cell lines to identify whether metformin's effects are perturbed by these amino acid levels (Sonnet *et al.*, 2016).

This data suggests that metformin may be affecting epigenetic marks, namely histone acetylation, and this could be in part due to acetyl co-A availability.

CHAPTER 6

6. Metabolite availability alters metformin efficacy

6.1 Introduction

As previously discussed, Type II Diabetes Mellitus is a risk factor for breast cancer. Two main features of T2DM are hyperglycaemia and hyperinsulinemia. As skeletal muscle, liver or fat cells become more resistant to insulin, cells stop taking up extracellular glucose, preferring to make glucose internally through gluconeogenesis (American Diabetes, 2010). Lack of insulin activation of the AKT pathway leads to decreased GLUT1 translocation to plasma membranes thus reducing extracellular glucose transport into cells (Chang, Chiang and Saltiel). Glucose is the main contributor to energy metabolism in cells. Glucose can be used in energy producing pathways such as oxidative phosphorylation or glycolysis leading to the production of ATP. In comparison to normal cells, cancer cells are more active consumers of glucose. As a tumour is actively growing and dividing at rapid rates the amount of glucose required for cellular functions is far greater compared to normal tissues. Glucose uptake contributes to key growth pathways required for cancer cell survival such as nucleotide biosynthesis and lipogenesis. During cancer progression, there is a major switch from oxidative phosphorylation to glycolysis, known as the Warburg effect (figure 6.1 A, (Vander Heiden, Cantley and Thompson, 2009)). Hyperglycaemia is a major risk factor for breast cancer. A number of prospective studies have demonstrated the relationship between increased glucose and breast cancer incidence. Rapp *et al.* saw a 38% increase in breast cancer risk [HR = 1.38 (95% CI, 1.02 – 1.86)] independent of BMI (Rapp *et al.*, 2006). Therefore, it is likely that if metformin was reducing breast cancer incidence it might be doing so via reducing this risk associated with hyperglycaemia.

Another key characteristic of T2DM is hyperinsulinemia. Hyperinsulinemia is caused when insulin-interacting cells such as liver, fat and/or skeletal muscle cells become insulin resistant leading to large

circulating levels of insulin (DeFronzo and Tripathy, 2009). A number of epidemiological studies demonstrate an increased risk of incident breast cancer in women with hyperinsulinemia. Gunter *et al.* demonstrated a 2-fold increase in risk of developing breast cancer in women with high insulin levels [HR = 2.40 (95% CI, 1.30 – 4.41)], independent of BMI and oestradiol (Gunter *et al.*, 2009). While the epidemiological evidence supports a relationship between insulin and breast cancer, the mechanisms were not completely understood. One hypothesis is that the greater risk of breast cancer observed in hyperinsulinemic women is due to the increase in circulating oestrogen in postmenopausal women. Insulin is an inhibitor of sex-hormone binding globulin (SHBG) expression, which transports oestradiol through the circulatory system. Increased circulating insulin levels have been associated with decreased SHBG which could lead to increased levels of free oestradiol (Wallace *et al.*, 2013). Insulin is also a powerful signalling hormone. Insulin binding to insulin receptors leads to IRS protein phosphorylation and activation of the AKT pathway. This is a key growth regulatory pathway in cells and was over activated in breast cancer with a simultaneous decline in inhibitory proteins towards this pathway. Activation of AKT leads to inhibition of apoptosis, activation of NF κ B pathways, translocation of GLUT1 glucose transporter to plasma membranes as well as a host of other pro-cancer functions (figure 6.1 B, (Chang, Chiang and Saltiel; Fresno Vara *et al.*, 2004)). Therefore, additional breast cancer prevention mechanisms could be via regulation of insulin and circulating hormone levels.

While glucose and insulin have been associated with T2DM and cancer risk, other metabolites have been linked with cancer progression. For example, glutamine could potentially alter pro-cancer pathways. Glutamine metabolism plays a key role in many normal cellular pathways including nucleotide biosynthesis, amino acid synthesis and reactive oxygen species (ROS) regulation. Glutamine-derived metabolites can progress through the TCA cycle and lead to acetyl co-A production (Newsholme *et al.*, 2003). Glutamine is a key metabolite that is utilised by cancer cells in glucose deficient conditions. Glutamine can enter the TCA cycle as glutamate and converted to α -

ketoglutarate. This is cycled through the TCA to produce ATP (Yang *et al.*, 2014). Glutamate can also be converted into 2-hydroxyglutarate (2-HG) by IDH1 and IDH2 in the cytoplasm and mitochondria, respectively. In the mitochondria, IDH2 converts α -ketoglutarate to 2-HG. IDH1 and IDH2 mutations were common in many cancer types, leading to accumulation of 2-HG both intra- and extracellularly. 2-HG, known as an “oncometabolite”, has been shown to significantly effect epigenetic marks such as histone PTMs and DNA methylation (Reitman and Yan, 2010; Xu *et al.*, 2011). Increased expression of glutamine-related genes has been associated with breast cancer progression; this was particularly associated with triple negative breast cancer (TNBC) cells (Lampa *et al.*, 2017; El Ansari *et al.*, 2018). In breast cancer cells 2-HG can cause inhibition of KDM2A (Terunuma *et al.*, 2014) leading to increased 2-HG, which has been associated with DNA hypermethylation through inhibition of TET enzymes (figure 6.1 C, (Xu *et al.*, 2011)). Therefore, regulation of cancer-linked metabolites such as glutamine could provide insight into breast cancer prevention mechanisms.

As shown previously in Chapter 3, acetate levels can greatly alter the efficacy of metformin in reducing H3K27ac. In this chapter we explore the hypothesis that availability of key metabolites such as glucose, insulin or glutamine could potentially affect the sensitivity of non-cancerous breast epithelial cells to metformin.

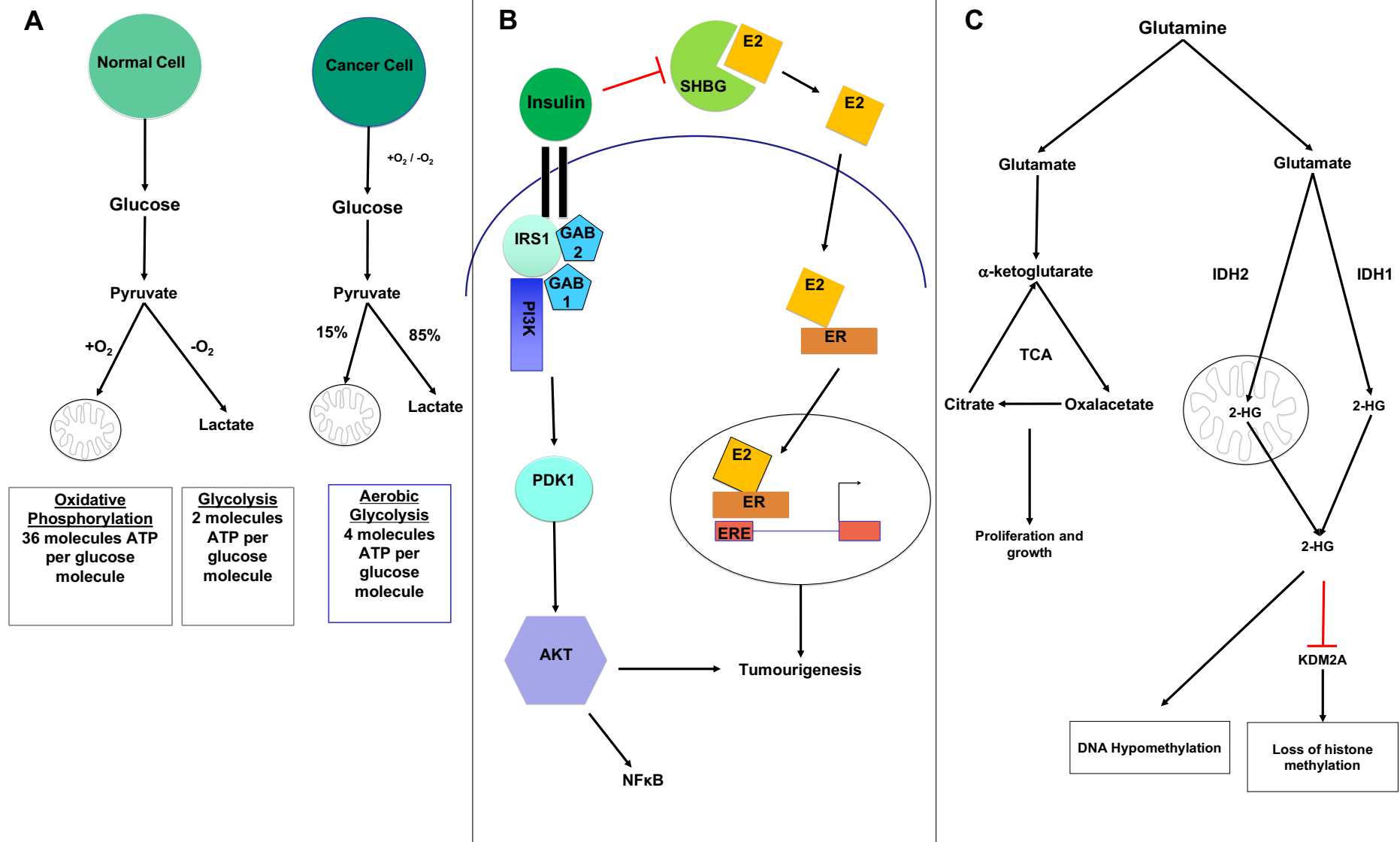


Figure 6.1 (A) The Warburg Effect (B) Insulin-related mechanisms to oestrogen pathways and cell growth pathways and (C) Glutamine metabolism effects on epigenetic marks and cell growth [Adapted from (Chang, Chiang and Saltiel; Fresno Vara *et al.*, 2004; Vander Heiden, Cantley and Thompson, 2009; Xu *et al.*, 2011)].

6.2 Results

6.2.1 Glucose restriction enhances metformin's efficacy

Previous studies in non-cancerous breast epithelial cell line MCF10A have shown no change in effect of metformin efficacy when treated at a range of glucose doses following 24-hour treatment (Zhuang *et al.*, 2014). To investigate whether this was seen in MCF12A non-cancerous breast epithelial cell line we chose to repeat this experiment with MCF12A and MCF10A cell lines. As MCF12A and MCF10A cells were typically cultured in media containing 17.5mM glucose the effect of glucose modification was assessed by crystal violet assay following treatment for 24hrs and 72hr in a range of glucose concentrations supplemented into glucose free media (0mM – 25mM). Modification of glucose concentrations in these cell lines did not significantly decrease cell viability (figure 6.2).

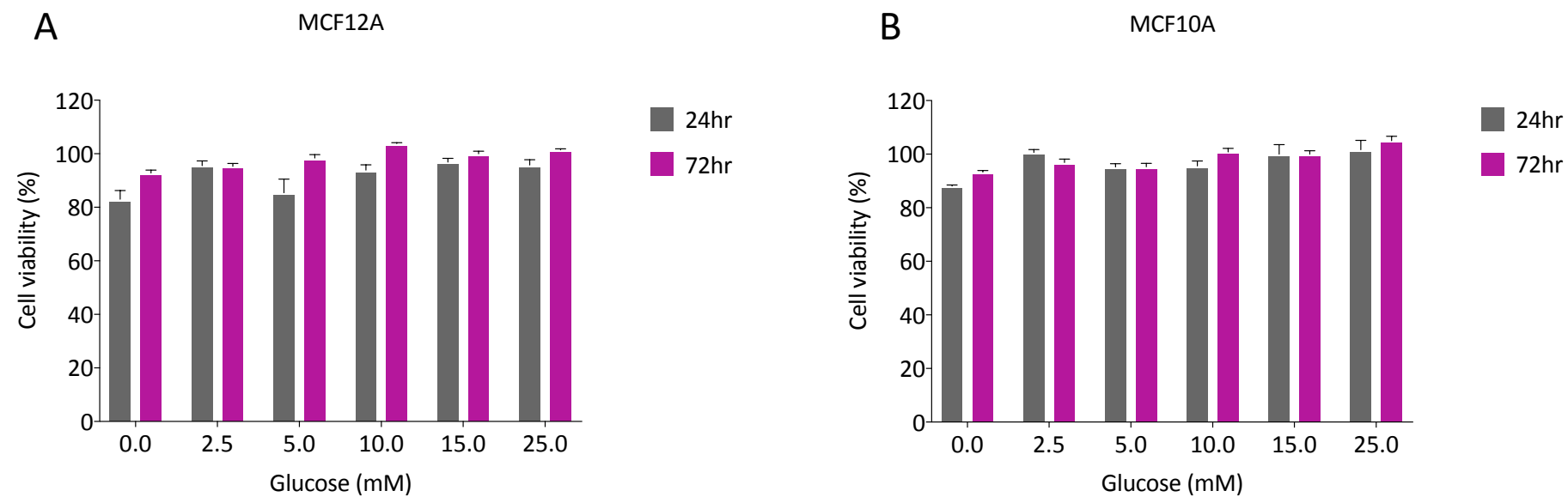


Figure 6.2 Crystal violet cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 24 and 72 hours in 0mM 2.5mM, 5mM, 10mM, 15mM, 20mM glucose. Data produced by Charlene Lam, MRes.

Following this cell lines were treated for 24- or 72-hours with 0mM, 2.5mM and 5mM metformin in a range of glucose concentrations (0mM – 25mM). Following treatment cell viability was assessed by crystal violet and MTS assays. Results of the crystal violet assay suggested there was a significant reduction in cell viability following 24hr in the samples treated in lower glucose concentrations (figure 6.3 A-B). This was seen in both MCF12A and MCF10A cell lines. In line with previous MTS assays, there was more reduction in cell viability compared with untreated and crystal violet assays (figure 6.4 A-B). As this was contrary to results shown by Zhuang *et al.*, we chose to investigate further the effect of glucose restriction on non-cancerous breast epithelial cells treated with metformin for 72 hours. This led to further reduced cell viability in samples treated with metformin. Samples treated in low glucose conditions 0mM and 5mM showed almost complete loss of cell viability when treated with metformin over this period (figure 6.3 C-D). It was important to note that in crystal violet assays the untreated cells were able to maintain cell viability, even in low glucose conditions. However, MTS assays showed significant reduction in cell viability in untreated cells in low glucose conditions (figure 6.4 C-D). As the MTS reagent was metabolised in the mitochondria, this may suggest that while the cells were able to adhere to the plates, and remain dormant, there may have been reduced mitochondrial function.

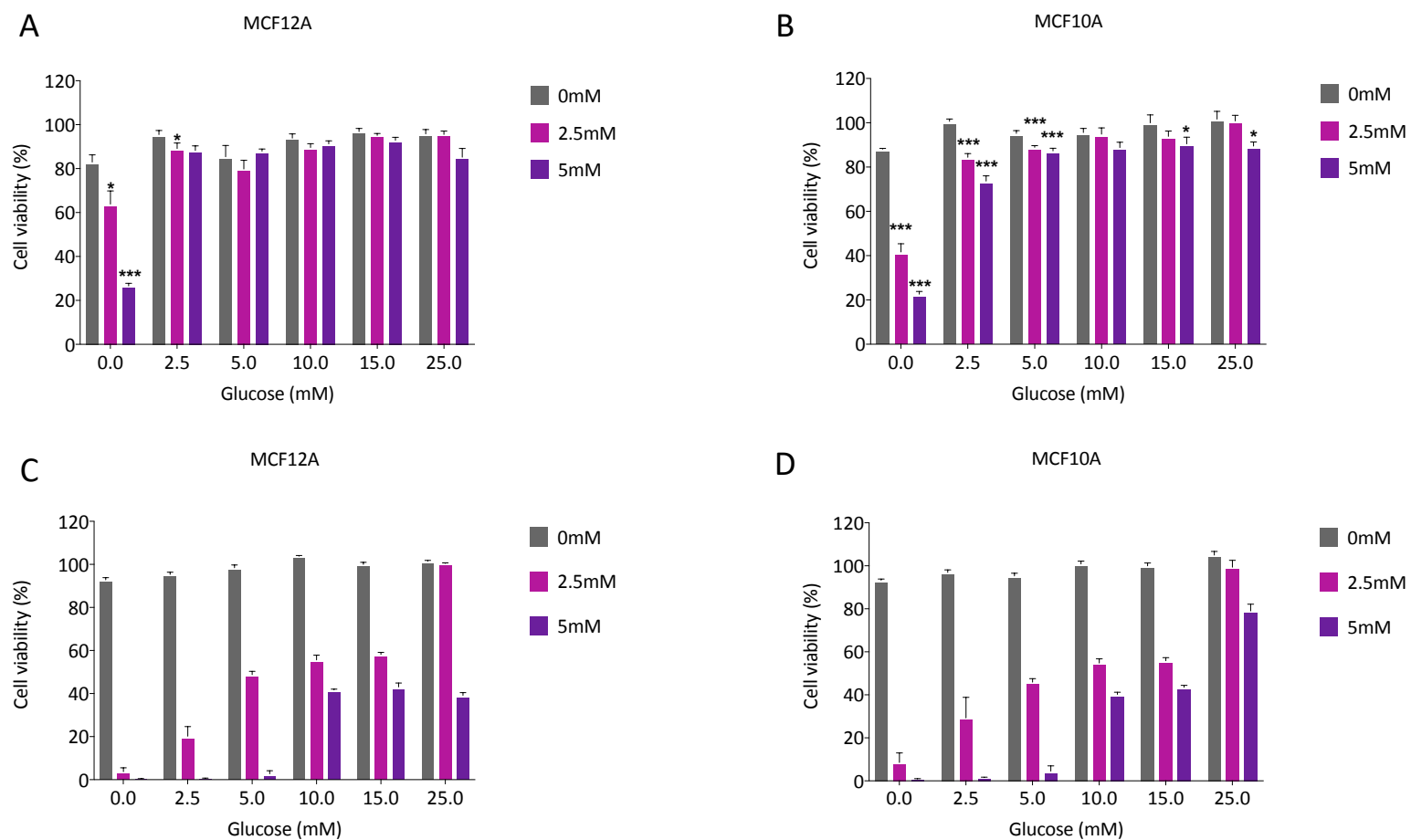


Figure 6.3 Crystal violet cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 24 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 2.5mM, 5mM, 10mM, 15mM, 20mM glucose. (C) MCF12A and (D) MCF10A non-cancerous breast epithelial cell lines treated

for 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 2.5mM, 5mM, 10mM, 15mM, 20mM glucose. Paired student's t-test * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 . Data produced by Charlene Lam, MRes.

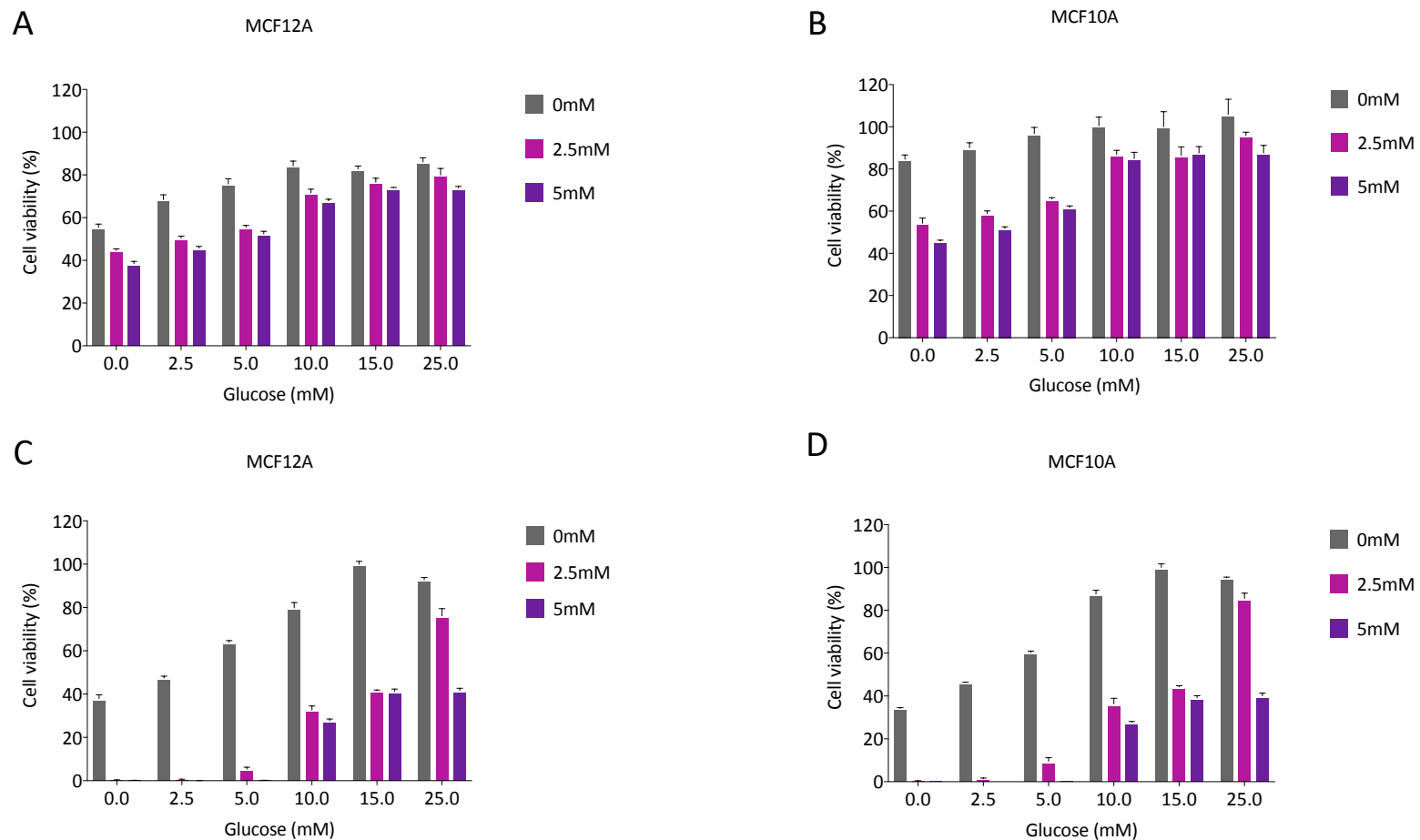


Figure 6.4 MTS cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 24 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 2.5mM, 5mM, 10mM, 15mM, 20mM glucose. (C) MCF12A and (D) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 2.5mM, 5mM, 10mM, 15mM, 20mM glucose. Data produced by Charlene Lam, MRes.

We conducted crystal violet cell viability assays in low glucose conditions (5mM) where MCF12A and MCF10A were treated in a lower concentration range of metformin (0mM – 2.5mM) for 72 hours. Results showed a significant loss of cell viability at 62.5 μ M (figure 6.5). MTS analysis of MCF12A and MCF10A cell lines treated with metformin in low glucose conditions showed similar reduction in cell viability in MCF12A cell lines. However, there appeared to be little effect on cell viability with this technique in MCF10A cell lines. This suggested that non-cancerous breast epithelial cells in normal physiological levels of glucose (5mM serum glucose) were more sensitive to metformin and lower doses of metformin could potentially remain effective.

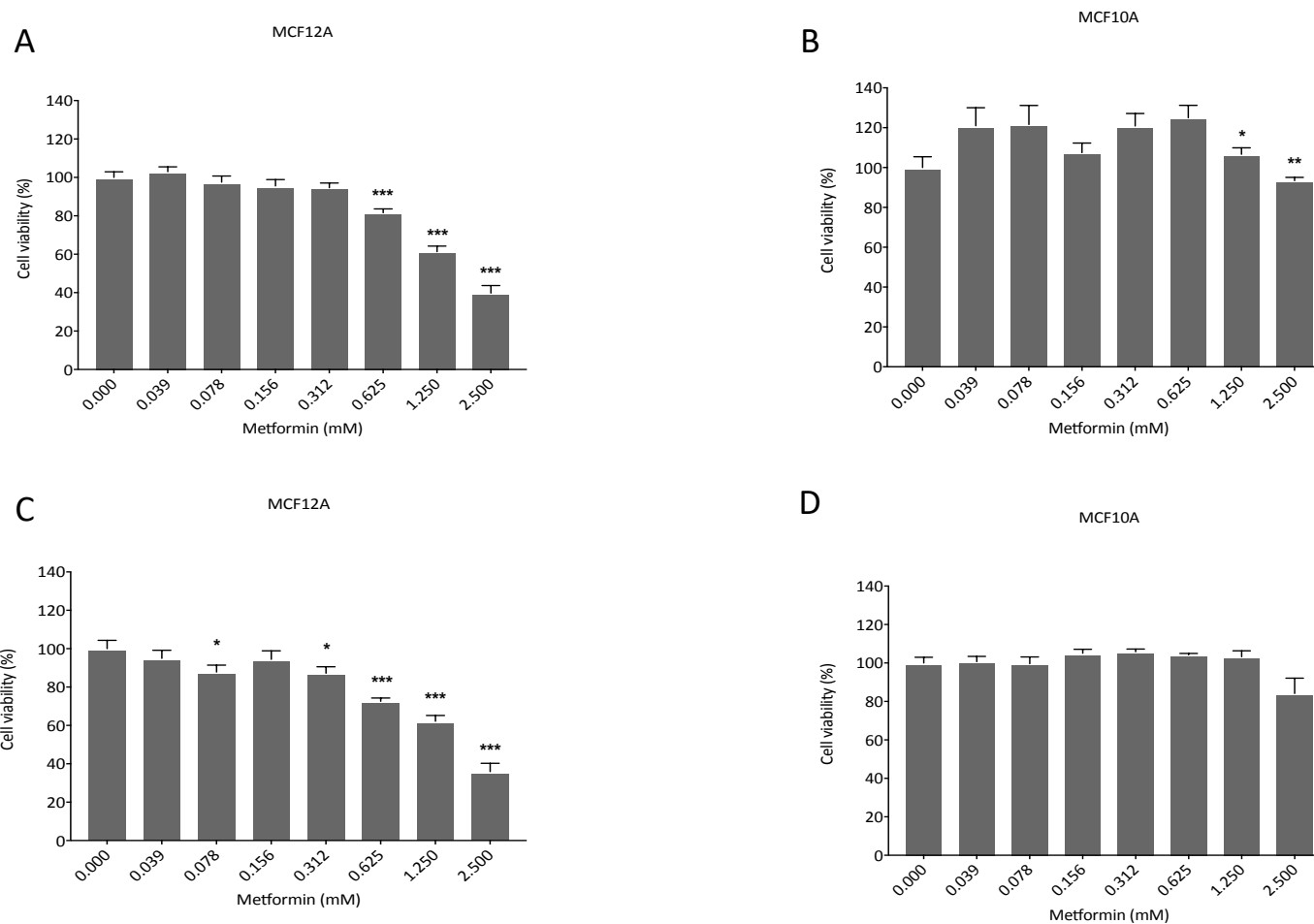


Figure 6.5 Crystal violet cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM – 2.5mM metformin in 5mM glucose. MTS cell viability assay in (C) MCF12A and (D) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM – 2.5mM metformin in 5mM glucose. Student's t-test * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 . Data produced by Charlene Lam, MRes.

6.2.2 Glucose restriction increases activation of p-AMPK in non-cancerous breast epithelial cell lines

As previously shown in chapter 3, metformin was able to activate p-AMPK while also decreasing total AMPK levels. To assess whether glucose restriction could alter p-AMPK activation, a western blot for p-AMPK and AMPK in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated with 2.5mM metformin at 5mM and 15mM glucose in biological triplicates was conducted. Figure 6.6 showed that in low glucose conditions (5mM) there was a significant increase in pAMPK protein activation when treated with 2.5mM metformin compared to untreated samples ($p < 0.05$). A drop in total AMPK was seen however, while this was not significant between treated and untreated samples it suggested that p-AMPK activation could reduce total AMPK levels. This data suggested that glucose restriction enhances the activation of p-AMPK and potentially the inhibition of cell growth.

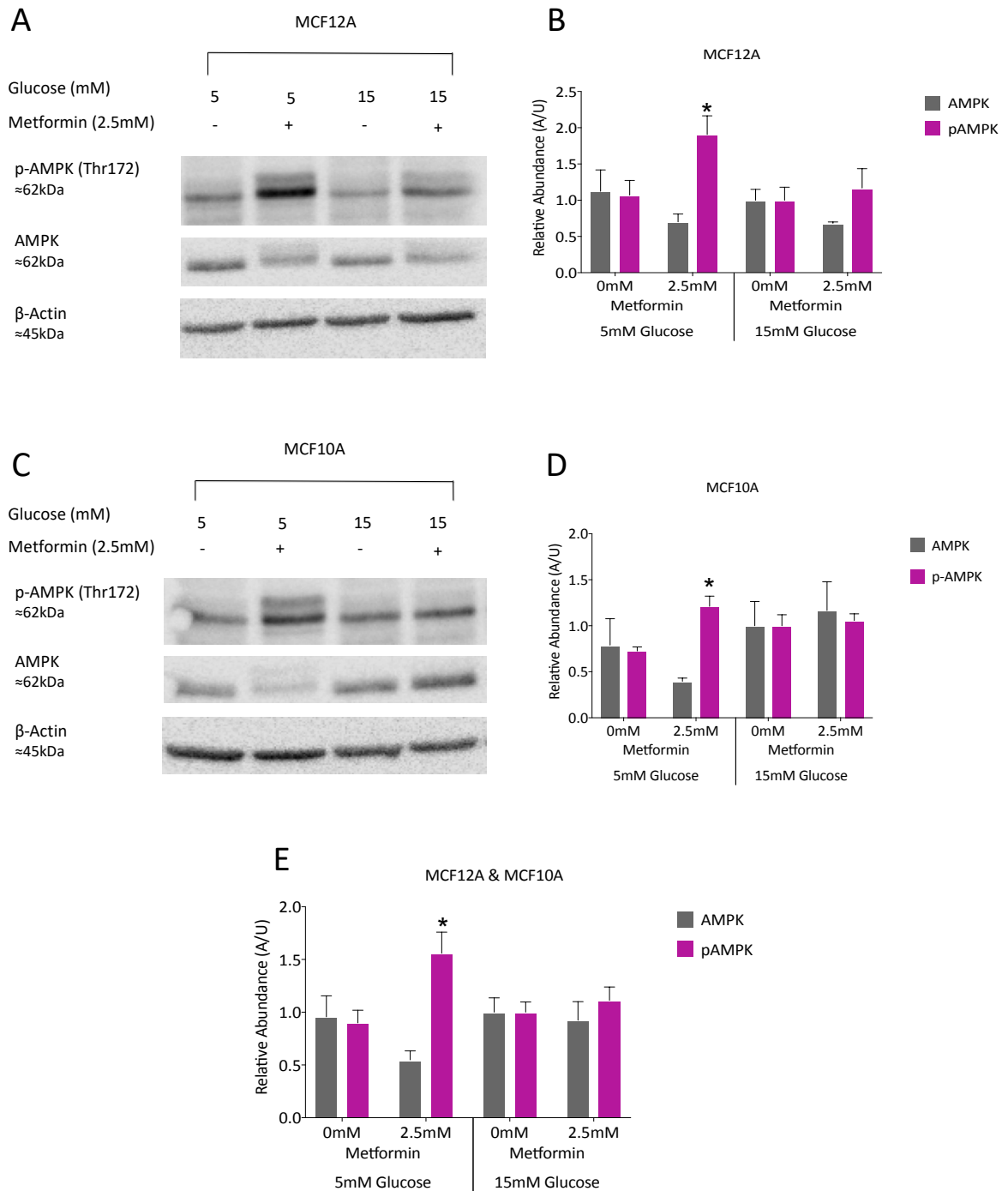


Figure 6.6 (A) Representative western blot for p-AMPK and AMPK for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Loading control beta actin. Relative abundance measurement of p-AMPK and AMPK of (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Averaged over three

biological replicates (n = 3). Relative abundance calculated as ratio to loading control beta actin.

Student's t-test * ≤ 0.05 .

6.2.3 GLUT1 gene expression is modified by metformin in normal breast epithelial cells

In diabetic conditions, metformin induces GLUT1 translocation to the cellular membrane as a result of mitochondrial complex I inhibition. This leads to external glucose uptake in parallel with metformin inhibiting gluconeogenesis (Fischer *et al.*, 1995). To assess whether metformin alters gene expression of GLUT1 (*SLC2A1*), RNA sequencing data analysed by Kallisto was probed - LogFold change of untreated vs treated (2.5mM or 5mM metformin). As shown in figure 6.7, Metformin reduced expression of GLUT1 in MCF10A. In contrast metformin induces increased GLUT1 gene expression in MCF12A. As previously shown, MCF12A cell line was more sensitive to metformin treatment compared to MCF10A cell line. This difference in GLUT1 gene expression may play a role in this sensitivity.

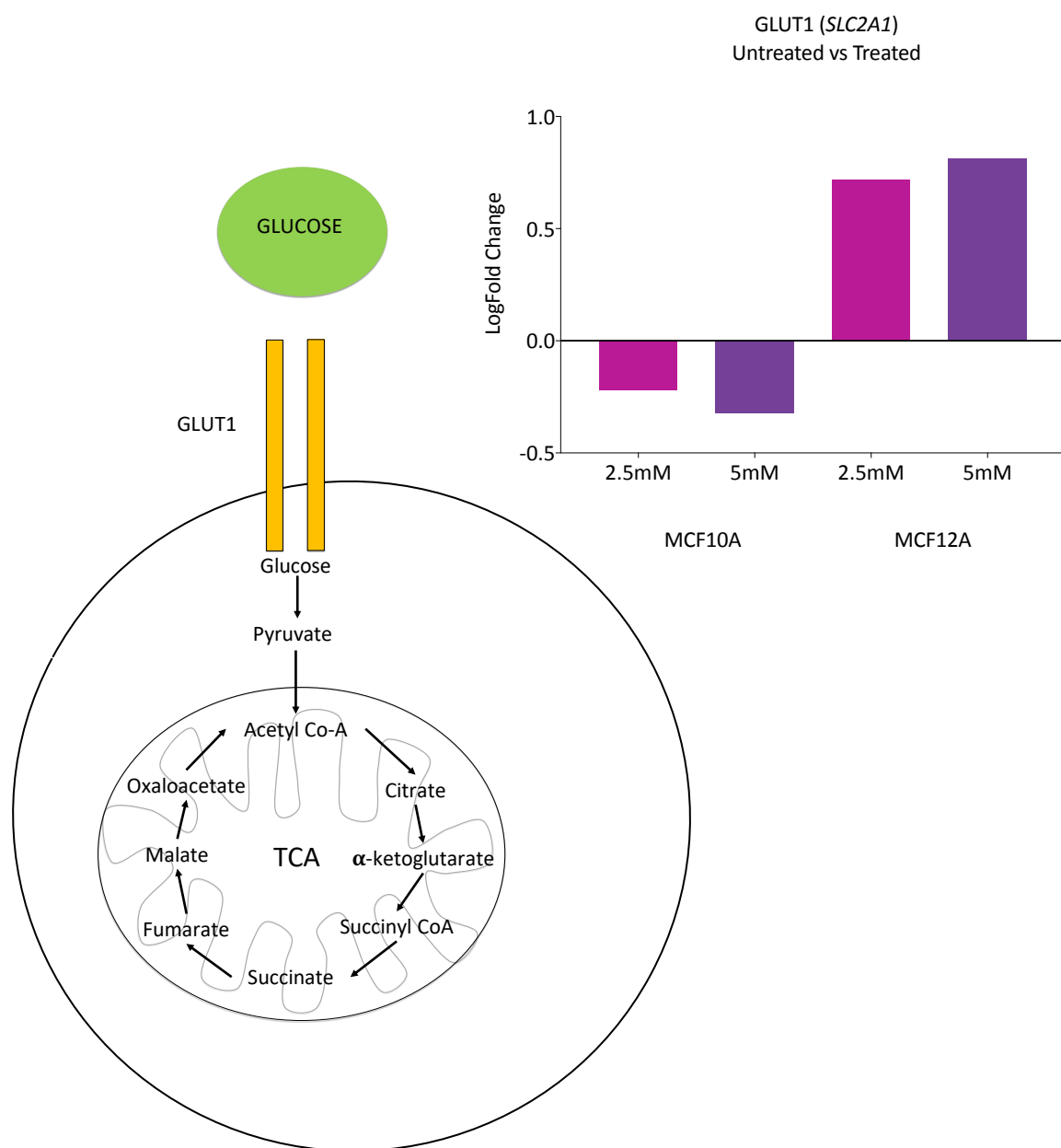


Figure 6.7 LogFold change gene expression untreated vs treated of *SLC2A1* (GLUT1) glucose transport gene in MCF12A and MCF12A non-cancerous breast epithelial cell lines treated for 72 hours with 2.5mM and 5mM metformin.

6.2.4 Glucose restriction on DNA methylation levels in glucose linked genes

As shown previously, metformin subtly alters DNA methylation levels and was associated with gene expression changes. To assess whether glucose restriction could potentially increase the effects of metformin on DNA methylation in non-cancerous breast epithelial cell lines, pyrosequencing analysis was conducted on MCF12A and MCF10A cell lines in low glucose conditions (5mM) treated for 72 hours with 0mM, 2.5mM, 5mM and 17.5mM metformin. LINE-1 methylation was initially assessed. There was no change in DNA methylation levels between treated and untreated samples in lower glucose conditions (figure 6.8 A-B). However, there was significant loss of LINE-1 methylation in the highest glucose conditions (17.5mM) compared to the lower glucose conditions ($p < 0.05$). MCF12A showed an average difference of $< 10\%$ and MCF10A showed an average difference of $< 9.5\%$. This suggested that glucose restriction may affect global LINE-1 methylation. Pyrosequencing analysis was also conducted on two CpG sites linked with glucose-related genes identified in Chapter 4 - cg01757206 and cg09183496 (table 6.1). There was no significant change in DNA methylation levels between treated and untreated samples in these CpG sites in across glucose conditions (figure 6.8 C-F). This suggested that glucose restriction may not alter metformin's effects on DNA methylation in these genes.

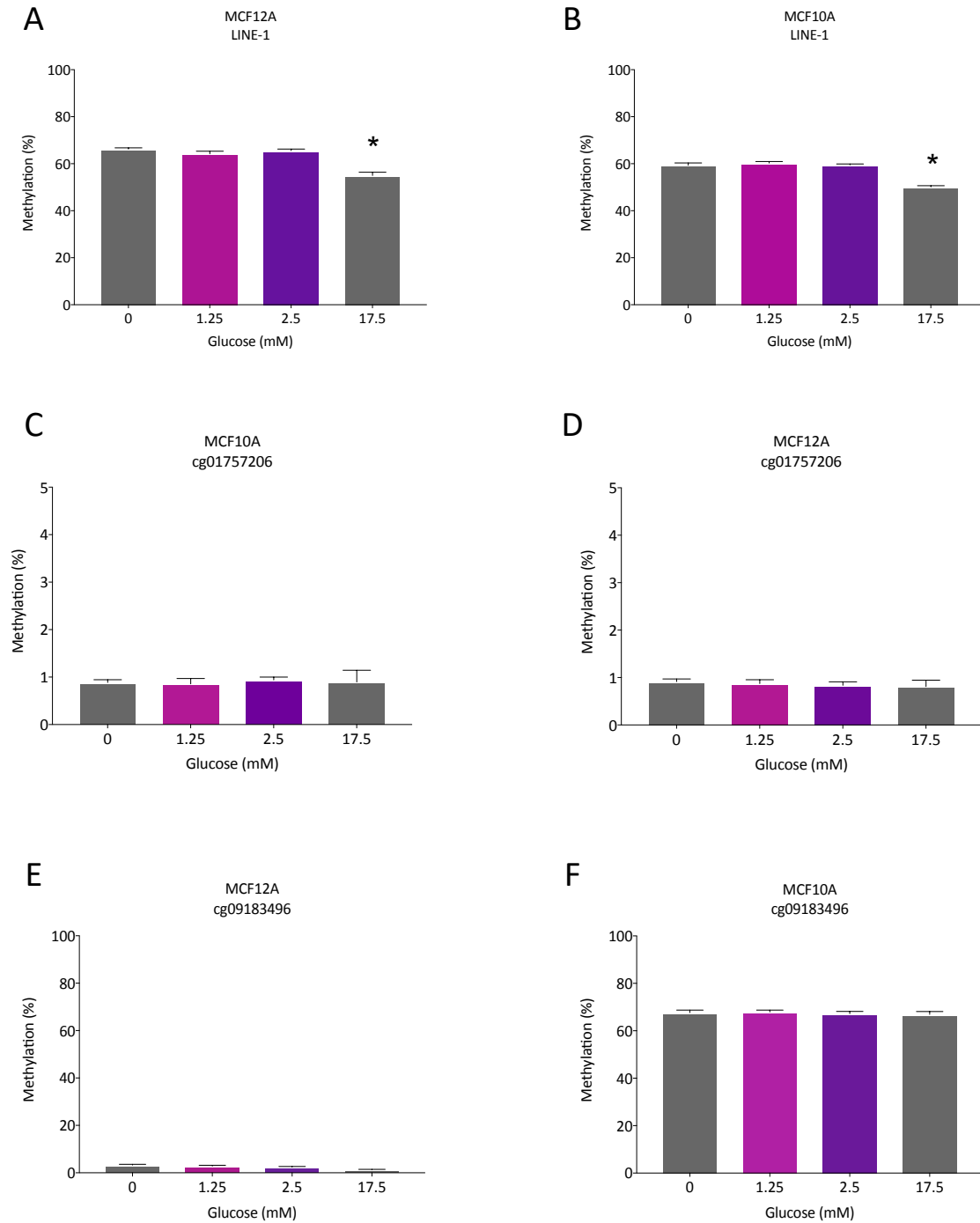


Figure 6.8 Pyrosequencing measurement of percentage methylation (%) in non-cancerous breast epithelial cell lines treated 72 hours with 5mM metformin in 0mM, 1.5mM and 2.5mM and 17.5mM glucose. Averaged over four biological replicates ($n = 4$). (A) LINE-1 in MCF12A, (B) LINE-1 in MCF10A, (C) cg01757206 in MCF12A, (D) cg01757206 in MCF10A, (E) cg09183496 in MCF12A and (F) cg09183496 in MCF10A. Student's t-test $* \leq 0.05$. Data produced by Charlene Lam, MRes.

cg Site	Closest Gene	Literature	Reference
cg09183496	DOCK2;FAM196B	DOCK2 Involved in chemokine signaling and lymphocyte migration DOCK2 Expressed in AML	(Nishikimi et al., 2013; Wu et al., 2017)
cg01757206	SLC2A4	Glucose transporter, known as GLUT4. Translocation to cell membrane induced by insulin signaling Metformin modulates GLUT4 translocation to cellular membrane	(Huang and Czech, 2007; Lee et al., 2012)

Table 6.1 Information of CpG sites cg09183496 and cg01757206

6.2.5 Glucose restriction leads to decreased H3K27ac in non-cancerous cell lines

It was demonstrated in Chapter 3 that metformin can decrease H3K27ac levels in normal breast epithelial cells. Further to this, acetate supplementation could restore normal H3K27ac levels. This suggested that metformin may affect acetyl co-A pools in normal breast epithelial cells. Glucose is the precursor for the TCA cycle where a downstream metabolite citrate can be converted in acetyl co-A. To further investigate this pathway, western blots in biological triplicates were conducted on MCF12A and MCF10A cells treated for 72 hours with 2.5mM metformin in low (5mM) and high (15mM) glucose conditions (figure 6.9). In low glucose conditions metformin significantly reduces H3K27ac protein levels compared to high glucose levels ($p < 0.05$). There was less H3K27ac in the untreated samples in low glucose compared to high glucose which could account for the high significance of loss. However, this data suggested that glucose restriction enhances the loss of H3K27ac seen with metformin treatment.

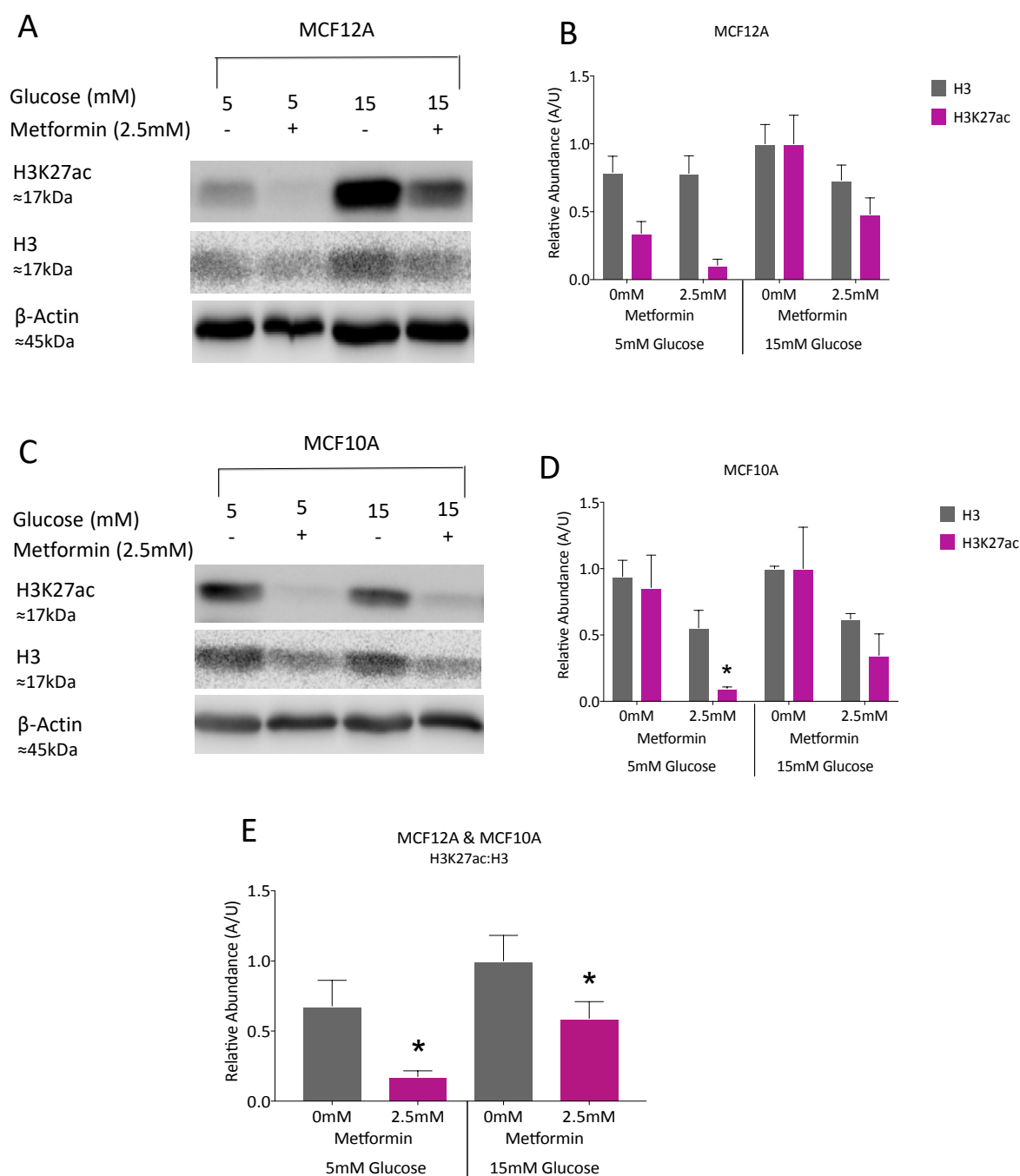


Figure 6.9 (A) Representative western blot for H3K27ac and H3 for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Loading control beta actin. Relative abundance measurement H3K27ac of (B) MCF12A and (C) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Averaged over three biological replicates (n = 3). (D) Relative abundance measurement of H3K27ac in MCF12A and MCF10A

combined treated for 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Averaged over six biological replicates (n = 6) Relative abundance calculated as ratio to H3. Student's t-test * ≤ 0.05 .

6.2.6 L-Glutamine does not alter non-cancerous breast epithelial cell lines' response to metformin

As previously discussed, glutamine metabolism has been linked with breast cancer progression. Recent studies have linked metformin treatment with inhibition of glutamine-related enzyme glutaminase (GLS) in breast cancer and cervical cancer cell lines (Saladini *et al.*, 2019). Crystal violet assays were used to assess whether L-glutamine could potentially alter sensitivity of non-cancerous breast epithelial cell lines to metformin. Treatment of MCF12A and MCF10A cell lines was conducted for 72 hours with 0mM, 2.5mM and 5mM metformin in a range of L-glutamine concentrations (0mM – 20mM). Similar to what was seen in chapter 3, there was a significant decrease in cell viability in metformin treated samples, however, there was no dose or doses of L-glutamine that showed more or less change (figure 6.10).

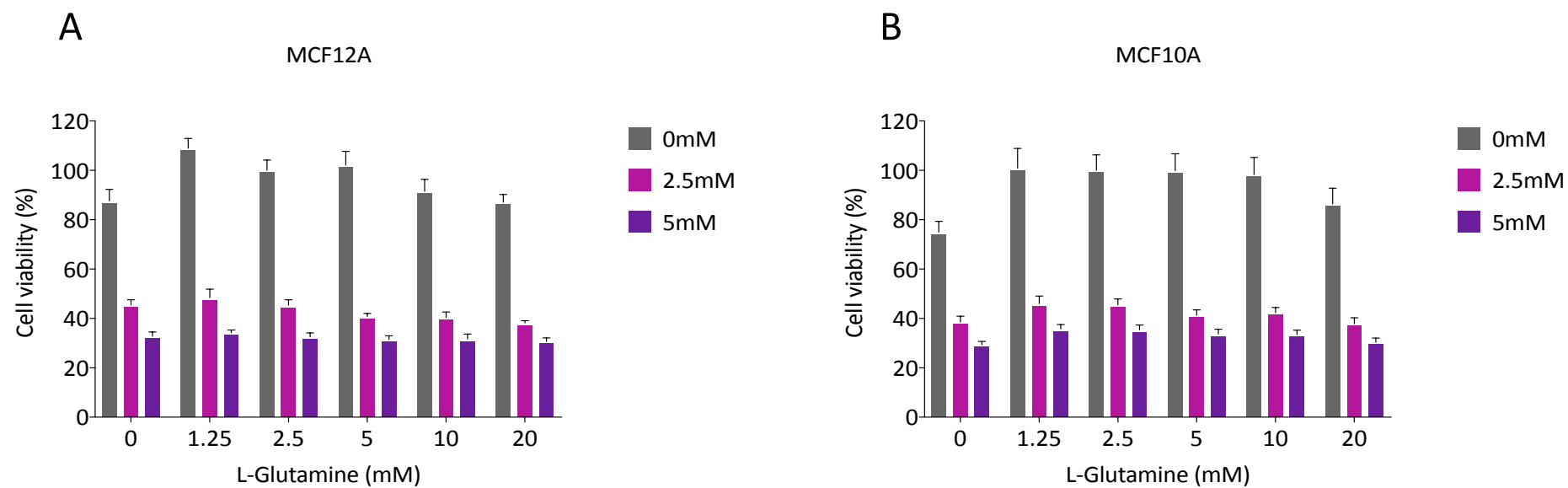


Figure 6.10 Crystal violet cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 1.25mM, 2.5mM, 5mM, 10mM, 20mM L-Glutamine (n=9).

6.2.7 Metformin alters key enzymes in the glutamine metabolism

Glutamine metabolism has been linked to alterations in DNA methylation and histone methylation. Increase in 2-HG production from glutamine could lead to DNA hypermethylation and increases histone methylation. To assess whether metformin impacts key enzymes in this pathway, RNA sequencing data analysed by Kallisto was assessed – LogFold change in untreated vs treated (2.5mM or 5mM metformin). Figure 6.11 showed gene expression of IDH1, a key enzyme that converts glutamate to 2-HG in the cytosol, increases with metformin treatment in both cell lines. This was consistent with clinical trial data which showed increased IDH1 expression post-metformin treatment in breast cancer tissue. In contrast, IDH2 gene expression, a key enzyme that converts glutamate to 2-HG in the mitochondria, decreased in MCF12A and MCF10A cell lines with metformin treatment. This may be potentially due to metformin's mitochondrial inhibiting mechanisms, leading to reduced functionality of this pathway. This result was in contrast to clinical trial data showing significantly increased IDH2 expression post-metformin treatment however this was only seen in OXPHOS transcriptional response (OTR) group. There appears to be little change in expression in FDG response (FR) group (Lord *et al.*, 2018).

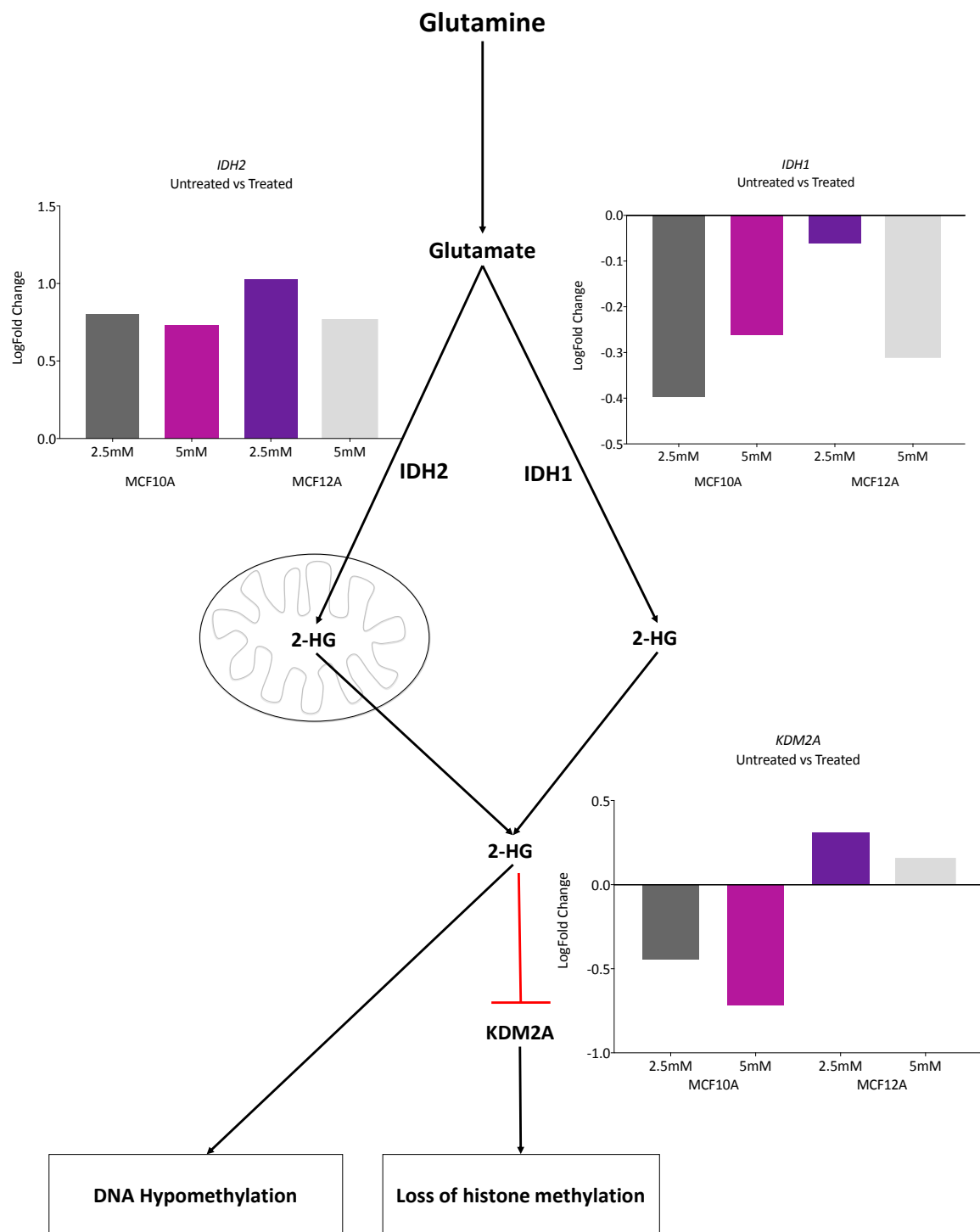


Figure 6.11 LogFold change gene expression untreated vs treated of glutamine-related genes – IDH1, IDH2 and KDM2A – in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 2.5mM and 5mM metformin.

Analysis of H3K27me3 by western blot in MCF12A and MCF10A cell lines treated for 72 hours with metformin (0mM, 2.5mM and 5mM) in 2.5mM L-glutamine showed H3K27me3 decreasing with metformin treatment (figure 6.12). Further studies into the effects of L-glutamine restriction/supplementation could further elucidate the effects of metformin on this pathway. Initial studies on L-glutamine concentration modification in metformin treated MCF12A and MCF10A of H3K27ac showed no significant difference in H3K27ac loss in L-glutamine normal or enriched conditions (figure 6.13). In L-glutamine deprived conditions there was no change in H3K27ac in MCF10A cell line (n=1).

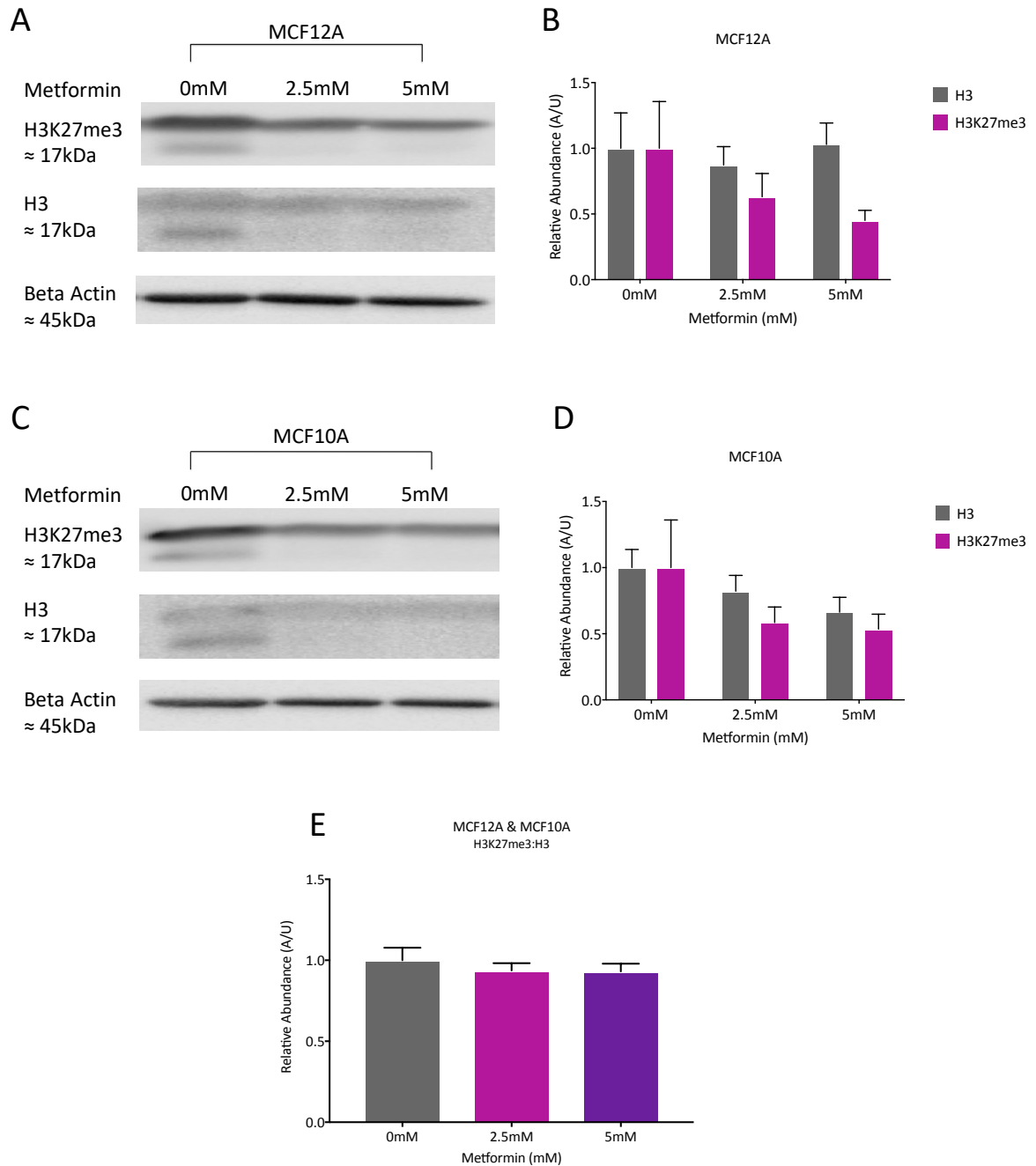


Figure 6.12 (A) Representative western blot for H3K27me3 and H3 for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Loading control beta actin. Relative abundance measurement H3K27ac of (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Averaged over three biological replicates (n = 3). Relative abundance calculated as ratio to H3.

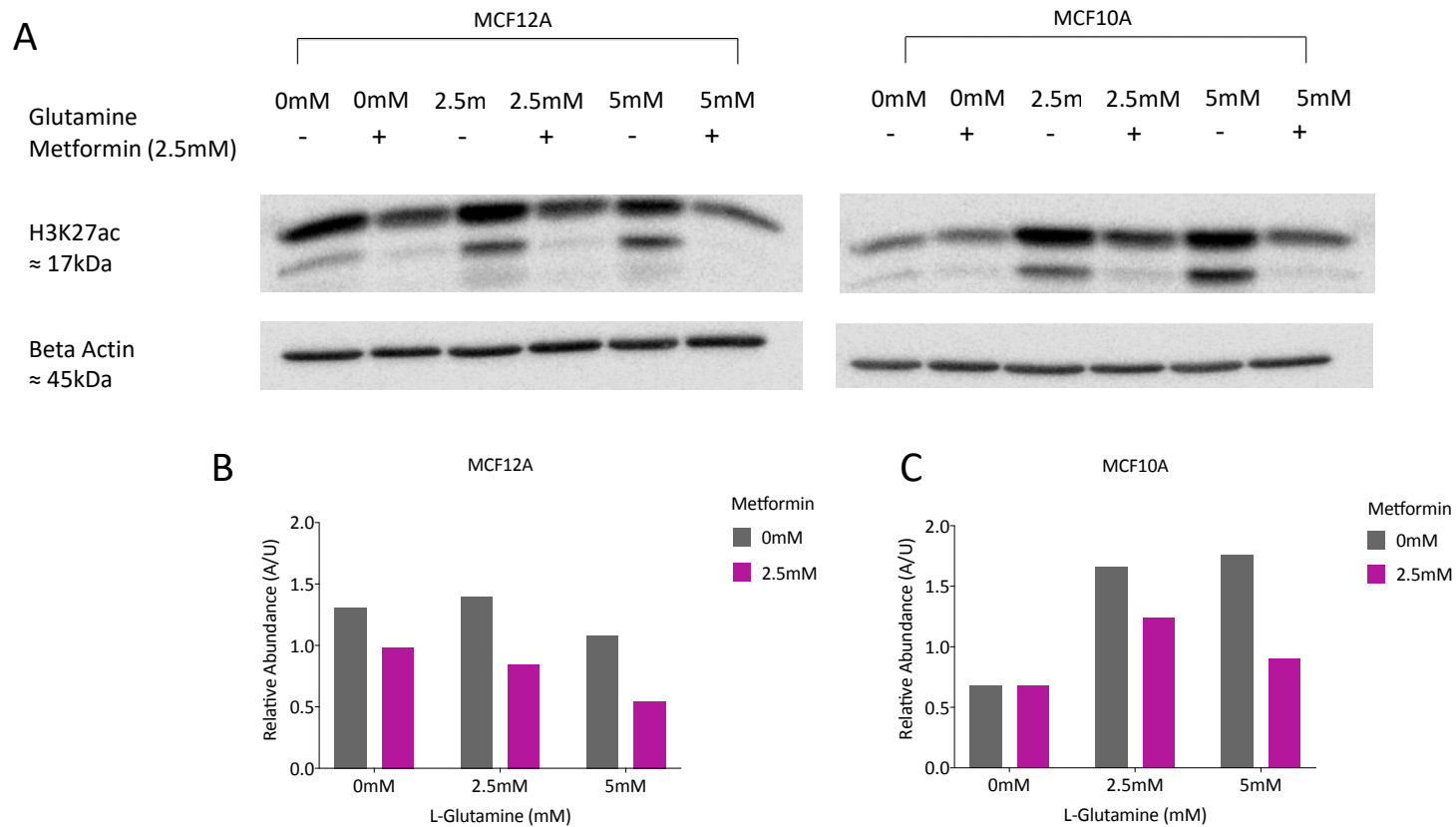


Figure 6.13 (A) Representative western blot for H3K27ac for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 2.5mM and 5mM L-glutamine (n=1). Loading control beta actin. (B) MCF12A and (C) MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 2.5mM and 5mM L-glutamine. Relative abundance calculated as ratio to loading control beta actin.

6.2.8 Insulin does not alter non-cancerous breast epithelial cell lines' response to metformin

To investigate whether insulin alterations have an effect on metformin treatment in non-cancerous breast epithelial cell lines, crystal violet assays were conducted in MCF12A and MCF10A cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in a range of insulin concentrations (0µg/ml - 20µg/ml). While metformin treatment induced significant loss in cell viability, consistent with studies shown in Chapter 3, there was no dose or doses that showed more or less change in cell viability in response to altered insulin levels (Figure 6.14 C-D). Additionally, investigation of crystal violet cell viability assay over 24 hours showed similar results (figure 6.14 A-B). This suggested that insulin concentrations did not affect the cell viability changes seen in metformin treated non-cancerous breast epithelial cell lines.

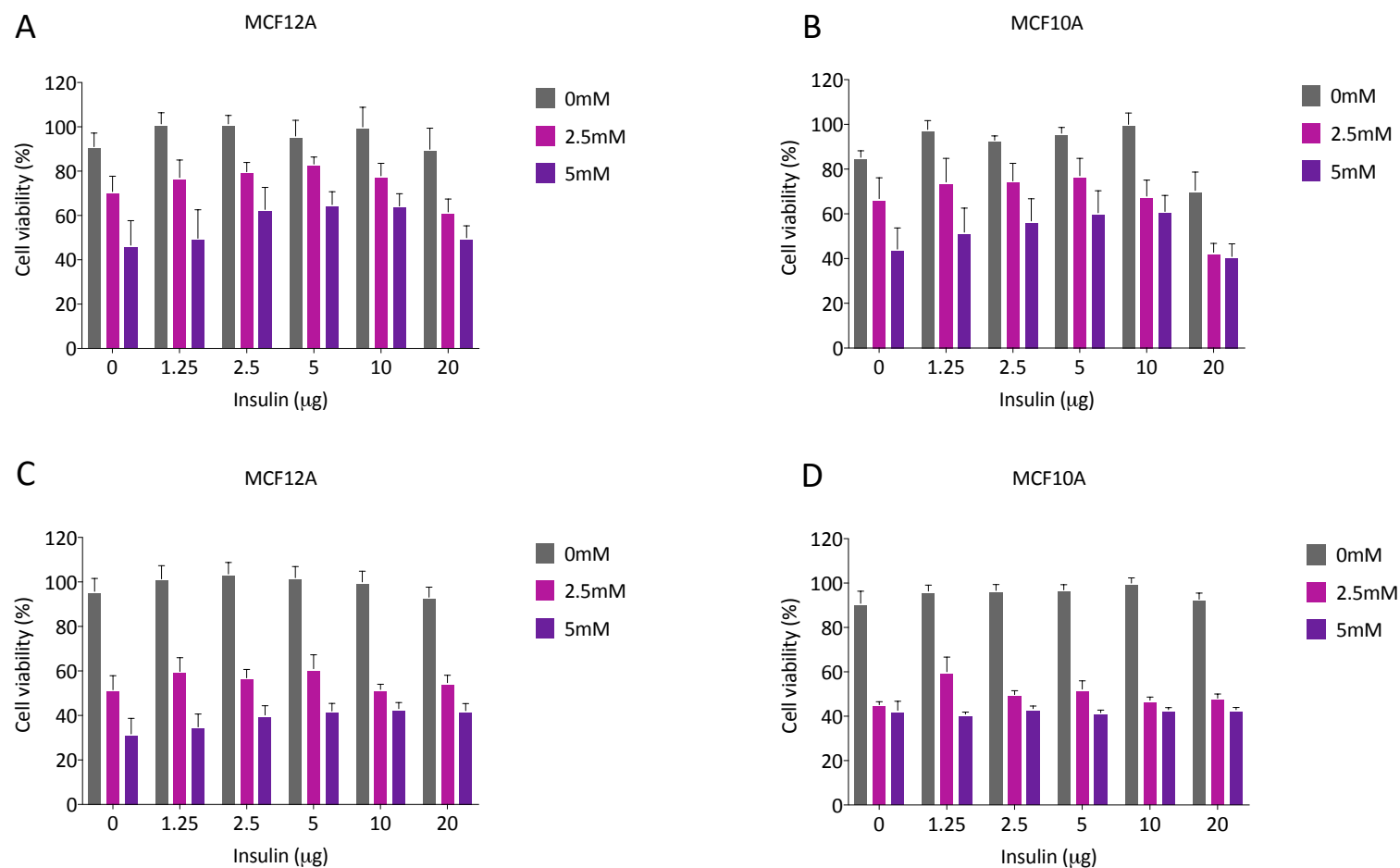


Figure 6.14 Crystal violet cell viability assay in (A) MCF12A and (B) MCF10A non-cancerous breast epithelial cell lines treated for 24 hours with 0mM, 2.5mM and 5mM metformin in 0µg/ml, 1.25µg/ml, 2.5µg/ml, 5µg/ml, 10µg/ml, 20µg/ml insulin (n=9). (C) MCF12A and (D) MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 0mM, 2.5mM and 5mM metformin in 0µg/ml, 1.25µg/ml, 2.5µg/ml, 5µg/ml, 10µg/ml, 20µg/ml insulin (n=9).

6.2.9 Metformin alters key enzymes in the insulin pathway

In diabetics, metformin re-establishes insulin signalling in order to normalise serum insulin levels. To assess the effects of metformin on insulin signalling in normal breast epithelial cells RNA sequencing data analysed by Kallisto was assessed – LogFold change in untreated vs treated (2.5mM or 5mM metformin). Figure 6.15 showed the insulin signalling pathway. *IRS1*, *GAB1* and *GAB2* are key signalling proteins which allow activation of PDK1 and downstream activation of insulin-related pathways. In MCF12A cell line there was a decrease in gene expression of *IRS1* in both treatment conditions. *GAB2* expression increased in 2.5mM conditions but this was lost in higher metformin concentrations. There appeared to be little effect on *GAB1*. In MCF10A cell line there was a loss of expression of *IRS1* between 2.5mM and 5mM treated conditions. *GAB2* showed increased expression, again dose dependent, while in contrast *GAB1* showed decreased expression. This result suggested that metformin may be affecting insulin signalling differently in the two cell lines.

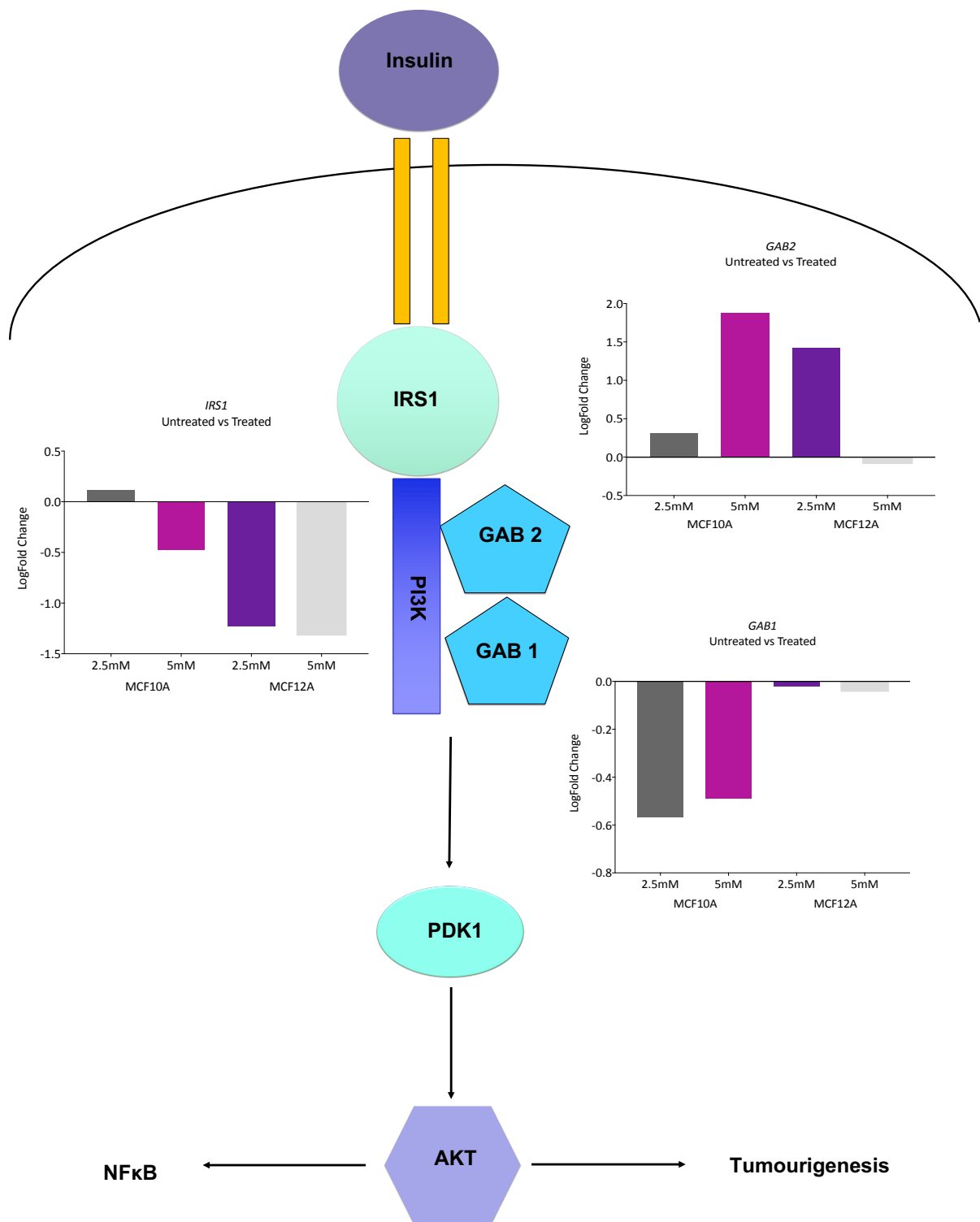


Figure 6.15 LogFold change gene expression untreated vs treated of insulin-related genes – IRS1, GAB1 and GAB2 – in MCF12A and MCF10A non-cancerous breast epithelial cell lines treated for 72 hours with 2.5mM and 5mM metformin.

6.3 Discussion

Metabolic Syndrome, characterised by metabolic dysfunction, is a disorder that increases the risk of patients developing a number of diseases including cardiovascular disease and T2DM (Kassi *et al.*, 2011). Metabolic dysfunction has become an emerging Hallmark of Cancer (Hanahan and Weinberg, 2011). This suggested that changes to metabolic pathways could have an impact in the development of disease. Data shown in chapter 4 and chapter 5 have indicated that metformin could potentially modify histone acetylation and DNA methylation. As metformin was known to dysregulate metabolism this may be potentially through modifying acetyl co-A pools and SAM levels. Glucose is the main precursor for many metabolites in the cell. Glucose conversion to serine could feed into the methionine pathway while glucose conversion to citrate could feed into acetyl co-A pools (Bouché *et al.*, 2004; Ducker and Rabinowitz, 2017). As these downstream metabolites are the main donors for epigenetic marks such as acetylation and methylation, it stands to reason that glucose concentrations may impact these pathways in metformin treated normal breast epithelial cells.

This study suggested that metformin efficacy could be modified by metabolite availability. MCF12A and MCF10A non-cancerous breast epithelial cell lines were grown in hyperglycaemic (17.5mM glucose) and hyperinsulinemic (10µg/ml insulin) conditions. Glucose restriction in these cell lines increased metformin efficacy in a number of mechanisms. In 5mM glucose, which is considered serum glucose levels in normal non-diabetic people, metformin activates p-AMPK significantly ($p < 0.05$) and showed significant increase in p-AMPK activation compared to hyperglycaemic conditions (15mM glucose). Recent studies in hepatocellular carcinoma (HCC) have shown glucose starvation as an important component in inducing metformin effects. A study by Ferretti *et al.* found that glucose starvation for 24 hours in 1mM metformin in C3A cell line induces significant reduction in the invasion and migratory ability of these cells. This study additionally showed glucose starvation can greatly increase p-AMPK activation compared to normal culture conditions in cells treated for 24 hours with 5mM metformin (Ferretti *et al.*, 2019). p-AMPK activation in 5mM glucose and 5mM metformin was

also enhanced in our study. Our study suggested that glucose restriction enhances metformin's effect on cell viability. This was contrary to a previous study by Zhuang *et al.* which showed glucose restriction (0mM – 25mM) over a 24-hour period in 0mM – 16mM metformin showed no alteration in cell viability (Zhuang *et al.*, 2014). Our study has shown that at both 24-hour and 72-hour time periods glucose restriction has a significant effect on metformin's efficacy and can, in lower glucose conditions, almost completely reduce cell viability when treated with 5mM metformin. For this reason, while the study by Ferretti *et al.* uses glucose starvation with and without metformin to study p-AMPK activation, our study focused on low glucose (5mM) rather than starvation. This data presented here showed glucose restriction could potentially enhance sensitivity of non-cancerous breast epithelial cells to metformin. In fact, doses of metformin closer to biological values can induce significant decreases in cell viability following 72 hours treatment. This was important, as previously discussed, the concentrations of metformin used in this study were not physiologically equivalent. It suggested that lower metformin concentrations could be used in normal glycaemic conditions and these future studies could potentially correspond more to *in vivo* responses. Future work into long-term exposure to metformin in glucose-restricted conditions would be interesting to understand potential *in vivo* responses. Metformin has been implicated as a cancer preventative drug, particular as a prophylactic therapy for non-diabetic women with a high risk of breast cancer. Our study suggested that lower metformin doses compared to diabetic patients should be considered in clinical trials. Glucose restriction was also shown to significantly reduce H3K27ac. As previously discussed, glucose is the main precursor for a number of metabolites including pyruvate and citrate. It would be interesting to modify this pathway through esiRNA knockdown of key enzymes to replicate metformin's effects.

Similar experiments conducted with insulin and L-glutamine restricted conditions have shown there was no additive or subtractive effect on metformin's efficacy. This may suggest there was no link between these metabolite levels and metformin efficacy in normal breast epithelial cells or, more

likely, that there were alternative pathways in which these metabolites interact with metformin that were not fully elucidated in this study. Glutamine metabolism in cancer has been implicated in changes in histone methylation and DNA hypomethylation. As we have shown previously metformin has a subtle effect on DNA methylation. Additional preliminary work has shown metformin may reduce H3K27me3. These epigenetic marks could be perturbed by glutamine restriction/addition and further assessed through esiRNA knockdown of enzymes in this pathway. It would be interesting to study these epigenetic marks following glutamine modifications. It is important to consider the cross-talk between these metabolites, specifically glucose and insulin. Future work into modifying both glucose and insulin in combination is required to fully explore this relationship. Metformin has been linked to insulin pathways in diabetic liver by re-introducing insulin signalling pathways. This was achieved by reducing gluconeogenesis however metformin has also been shown to increase insulin receptor function leading to a greater insulin sensitivity. As hyperinsulinemia is a risk factor for breast cancer this mechanism may not fit with the hypothesis that metformin reduces breast cancer risk. However, a key concept that is emerging is that normalising the energy balance is important in maintaining normal functioning cells and that perturbations to this balance could enhance cancer progression. Studies in the last decade have added to the idea that cancer cells cause deregulation of “cancer energetics” and this is quickly becoming a new hallmark of cancer. There may be an argument that metformin modifying the metabolism of MCF12A and MCF10A cell lines, typically seen as non-cancerous or precursor cells, could be returning these cells to a more “normal” phenotype. This may be a limitation as the cells studied were not fully “normal”. Previous data produced in our study has used normal matched primary epithelial and myoepithelial cells to validate what was seen in cell lines to reduce this limitation. Another technique that could be used to assess metformin’s effects is 3D mammosphere formation. This replicates the 3D *in vivo* environment of normal breast cells. Preliminary analysis in this study suggested MCF12A and MCF10A cell lines were capable of forming spheroids, however, further investigations into the structure and phenotypes of these spheroids is required (appendix figure 6.4).

This study raised key questions in understanding the effect of external metabolites on metformin efficacy. This could potentially impact results in models of non-cancerous breast epithelial cells treated with metformin and highlighted key areas that may need to be considered in future experiments.

CHAPTER 7

7. Discussion and Conclusions

As the most commonly diagnosed cancer in the UK, the drive to prevent breast cancer has been growing in recent years. Research into breast cancer prevention is required to bring feasible and sensitive measures to the general public. T2DM is a risk factor for breast cancer. Modifying risk factors such as this provides a feasible method of breast cancer prevention. Metformin is the first line therapy for type II diabetics, meaning almost every type II diabetic in the UK will have taken metformin at some point during their treatment. Metformin is also a highly tolerable drug, with the majority of complaints related to gastrointestinal discomfort. Finally, metformin is generically available and is relatively inexpensive. Metformin has been shown in epidemiological studies to reduce breast cancer risk by 2-fold. Treatment of high-risk cancer patients, for example those with familial history, with metformin could provide a possible reduction in breast cancer risk. Generation of polygenic risk scores (PRS) by Mavaddat *et al.* using common breast cancer associated SNPs have been able to estimate lifetime risk for breast cancer, for example the PRS of the highest risk patients with familial history is 24.4%. Treatment with metformin could hypothetically reduce this lifetime risk to 12.4%, which is less than the highest risk patients with no familial history (16.6%) (Mavaddat *et al.*, 2015). Taken together this makes metformin a feasible drug to potentially use as a preventative measure for breast cancer.

Epidemiological studies have found that metformin reduces breast cancer risk in diabetic women. The data, however, were not consistent, as many studies only take into account usage and not necessarily time of usage. For example, meta-analyses group metformin users as ever or never. These studies showed no reduced risk associated with metformin treatment (Franciosi *et al.*, 2013). However, the studies showing reduced risk separated metformin users by timepoint. The study by Bodmer *et al.* for example separated users by years and prescriptions issued and showed a 2-fold decrease in breast

cancer risk in the “5+ year or >40 prescription” group [OR = 0.44 (95% CI, 0.24 – 0.82)]. This is an important factor in metformin treatment which our study has started to address. Metformin induced histone modifications following 24-hour treatment however 24-hours after metformin was removed, the effect of metformin started to recover and was fully recovered after 48-hours. This short-term study provides a possible insight into the length of time needed to see effects by metformin and the possibility that consistent dosage with metformin is required. This could have implications for not only T2DM treatment but also breast cancer prevention. The initial aim of a preventative measure such as this would be to give this drug to high risk patients, for example those with familial breast cancer, prophylactically. High risk breast cancer patients may need to consider metformin at an earlier age than other conventional preventative measures in order to fully gain the benefit. Further study into the dosage and time effects of metformin would need to be fully explored before this observation could potentially impact treatment policy.

This study raised a number of questions in relation to dosage. The typical dose of metformin for a diabetic is 1000 - 1500mg/day. However, as metformin is used to reduce serum glucose levels this dose may not be appropriate for patients with normal glucose and/or insulin levels. This study showed that in glucose conditions closer to physiological levels (5mM glucose), metformin was more effective than in hyperglycaemic conditions, for example decreased H3K27ac and increased p-AMPK activation. Women with breast cancer and normal BMI have been treated with metformin before standard treatment in a number of clinical trials. Breast cancer patients were treated with metformin doses similar to doses given to diabetic patients and have shown reduced tumour size and metabolite modifications (Goodwin *et al.*, 2002; Lord *et al.*, 2018). Patients reported adverse effects including gastrointestinal complications. Our data suggested that a lower dose of metformin in “normal” glucose condition may be as effective as metformin in hyperglycaemic conditions.

Further to this, this study has shown that different levels of metabolites could greatly affect metformin’s efficacy. Metformin reduced H3K27ac levels in normal breast epithelial cells. We

hypothesised that this may be linked to altered acetyl co-A generation through the TCA cycle as we see a greater effect in low glucose conditions (5mM) compared to high (17.5mM). However, acetate supplementation abrogated this effect. In 2mM L-glutamine conditions treated with metformin there was reduced H3K27me3. Published studies have shown that glutamate availability, which is derived from glutamine, could greatly affect histone methylation and DNA methylation. We hypothesise from this data that modifying glutamine levels could potentially alter metformin efficacy.

Further study into the interaction of metabolite availability and metformin dose may potentially give insight into dosage for high risk, non-diabetic patients.

This study has shown genetic modifications associated with metformin treatment in normal breast epithelial cells, particularly in the expression of metabolic enzymes. This provided a potential connection between the epigenetic modifications seen in metformin treated breast epithelial cells and feasible mechanisms of action. Further study into these modifications could provide more insight into these pathways. This study showed that there is distinct heterogeneity between primary normal breast epithelial and myoepithelial samples from four different patients. Published studies have shown genetic variants of the key transporters of metformin, namely *SLC22A1* (OCT1) and *SLC47A1* (MATE1) phenotypes in diabetic patients greatly alters their responsiveness, where low expression of *SLC22A1* (OCT1) reduces the accumulation of metformin in cells leading to reduced response (Gong *et al.*, 2012; Florez, 2017; Mato *et al.*, 2018). Further study into the transporter status of these primary cells could potentially explain these heterogeneous effects.

Biomarkers are useful in stratifying patients and monitoring patient response. This study has shown that metformin may be interacting with a number of potential biomarkers. Gene expression of key enzymes in metformin associated pathways could be used to stratify patients, for example patients with high expression of *ACSS2* or *ACLY* may not respond to metformin compared to low expressors of these enzymes. Metabolite levels could be used to determine optimum metformin dose as well as

monitor patient response to metformin. For example, analysis of SAM or acetyl co-A levels could determine whether patients are responding to metformin. H3K27ac tracking may provide a specific biomarker for patient response, as this could be compared to baseline levels of H3K27ac across a number of months and/or years. Future work into these pathways could not only provide more information in metformin's mechanisms of action but also could provide sensitive and cost-effective biomarkers for breast cancer prevention.

This study had combined genomics, epigenomics and metabolomics in order to investigate potential mechanisms of metformin in normal breast epithelial cells. The data presented here highlighted potential avenues for further study with the hope that future studies could identify metformin's mechanisms and potential efficacy as a preventive agent for breast cancer.

BIBLIOGRAPHY

(NIDDK), N. I. of D. and D. and K. D. (2002) *Diabetes Prevention Program Outcomes Study (DPPOS) (NCT00038727)*, *ClinicalTrials.gov*. Available at:

<https://clinicaltrials.gov/ct2/show/NCT00038727?term=metformin+risk+cancer&draw=8&rank=24>
(Accessed: 19 September 2019).

Adami, H.-O. *et al.* (1996) 'Excess Risk of Primary Liver Cancer in Patients With Diabetes Mellitus', *JNCI Journal of the National Cancer Institute*, 88(20), pp. 1472–1477. doi: 10.1093/jnci/88.20.1472.

Adeva-Andany, M. M. *et al.* (2017) 'Enzymes involved in branched-chain amino acid metabolism in humans', *Amino Acids*, 49(6), pp. 1005–1028. doi: 10.1007/s00726-017-2412-7.

American Diabetes, A. (2010) 'Diagnosis and classification of diabetes mellitus.', *Diabetes care*, 33 Suppl 1, pp. S62-69. doi: 10.2337/dc10-S062.

Anderson, O. S., Sant, K. E. and Dolinoy, D. C. (2012) 'Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation', *The Journal of Nutritional Biochemistry*. Elsevier, 23(8), pp. 853–859. doi: 10.1016/J.JNUTBIO.2012.03.003.

Anna Maria Storniolo (2013) *Metformin Hydrochloride vs. Placebo in Overweight or Obese Patients at Elevated Risk for Breast Cancer (NCT01793948)*, *ClinicalTrials.gov*. Available at:
<https://clinicaltrials.gov/ct2/show/study/NCT01793948?term=metformin+risk+cancer&draw=2&rank=3> (Accessed: 19 September 2019).

El Ansari, R. *et al.* (2018) 'Altered glutamine metabolism in breast cancer; subtype dependencies and alternative adaptations', *Histopathology*, 72(2), pp. 183–190. doi: 10.1111/his.13334.

Back, S. H. and Kaufman, R. J. (2012) 'Endoplasmic reticulum stress and type 2 diabetes.', *Annual review of biochemistry*, 81, pp. 767–93. doi: 10.1146/annurev-biochem-072909-095555.

Balasubramanyam, M. *et al.* (2011) 'Impaired miR-146a expression links subclinical inflammation and insulin resistance in Type 2 diabetes', *Molecular and Cellular Biochemistry*, 351(1–2), pp. 197–205. doi: 10.1007/s11010-011-0727-3.

Bao, B. *et al.* (2012) 'Metformin inhibits cell proliferation, migration and invasion by attenuating CSC

function mediated by deregulating miRNAs in pancreatic cancer cells.’, *Cancer prevention research (Philadelphia, Pa.)*, 5(3), pp. 355–64. doi: 10.1158/1940-6207.CAPR-11-0299.

Bartek, J. and Lukas, J. (2011) ‘DNA repair: Cyclin D1 multitasks.’, *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 474(7350), pp. 171–2. doi: 10.1038/474171a.

Baumann, M., Krause, M. and Hill, R. (2008) ‘Exploring the role of cancer stem cells in radioresistance.’, *Nature reviews. Cancer*. Nature Publishing Group, 8(7), pp. 545–54. doi: 10.1038/nrc2419.

Beck, B. and Blanpain, C. (2013) ‘Unravelling cancer stem cell potential.’, *Nature reviews. Cancer*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 13(10), pp. 727–38. doi: 10.1038/nrc3597.

Bernhardt, S. *et al.* (2017) ‘Proteomic profiling of breast cancer metabolism identifies SHMT2 and ASCT2 as prognostic factors’, *Breast Cancer Research*, 19(1), p. 112. doi: 10.1186/s13058-017-0905-7.

Bernstein, B. E., Meissner, A. and Lander, E. S. (2007) ‘The mammalian epigenome.’, *Cell*, 128(4), pp. 669–81. doi: 10.1016/j.cell.2007.01.033.

Bièche, I. *et al.* (2002) ‘Prognostic value of CCND1 gene status in sporadic breast tumours, as determined by real-time quantitative PCR assays.’, *British journal of cancer*, 86(4), pp. 580–6. doi: 10.1038/sj.bjc.6600109.

Bishnu, A. *et al.* (2019) ‘Long term treatment of metformin impedes development of chemoresistance by regulating cancer stem cell differentiation through taurine generation in ovarian cancer cells’, *The International Journal of Biochemistry & Cell Biology*, 107, pp. 116–127. doi: 10.1016/j.biocel.2018.12.016.

Blandino, G. *et al.* (2012) ‘Metformin elicits anticancer effects through the sequential modulation of DICER and c-MYC.’, *Nature communications*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 3, p. 865. doi: 10.1038/ncomms1859.

Bleyer, A. and Welch, H. G. (2012) 'Effect of Three Decades of Screening Mammography on Breast-Cancer Incidence', *New England Journal of Medicine*. Massachusetts Medical Society, 367(21), pp. 1998–2005. doi: 10.1056/NEJMoa1206809.

Bodmer, M. *et al.* (2010) 'Long-term metformin use is associated with decreased risk of breast cancer.', *Diabetes care*, 33(6), pp. 1304–8. doi: 10.2337/dc09-1791.

Bosco, J. L. F. *et al.* (2010) 'Metformin and Incident Breast Cancer among Diabetic Women: A Population-Based Case-Control Study in Denmark', *Cancer Epidemiology Biomarkers & Prevention*, 20(1), pp. 101–111. doi: 10.1158/1055-9965.EPI-10-0817.

Bouché, C. *et al.* (2004) 'The Cellular Fate of Glucose and Its Relevance in Type 2 Diabetes', *Endocrine Reviews*. Narnia, 25(5), pp. 807–830. doi: 10.1210/er.2003-0026.

Bowker, S. L. *et al.* (2017) 'Time-Varying Risk for Breast Cancer Following Initiation of Glucose-Lowering Therapy in Women with Type 2 Diabetes: Exploring Detection Bias', *Canadian Journal of Diabetes*, 41(2), pp. 204–210. doi: 10.1016/j.jcjd.2016.08.227.

Brasacchio, D. *et al.* (2009) 'Hyperglycemia induces a dynamic cooperativity of histone methylase and demethylase enzymes associated with gene-activating epigenetic marks that coexist on the lysine tail.', *Diabetes*, 58(5), pp. 1229–36. doi: 10.2337/db08-1666.

Bray, N. L. *et al.* (2016) 'Near-optimal probabilistic RNA-seq quantification', *Nature Biotechnology*. Nature Publishing Group, 34(5), pp. 525–527. doi: 10.1038/nbt.3519.

Brennan, K. *et al.* (2012) 'Intragenic ATM methylation in peripheral blood DNA as a biomarker of breast cancer risk.', *Cancer research*, 72(9), pp. 2304–13. doi: 10.1158/0008-5472.CAN-11-3157.

Cai, J., Yang, J. and Jones, D. (1998) 'Mitochondrial control of apoptosis: the role of cytochrome c', *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1366(1–2), pp. 139–149. doi: 10.1016/S0005-2728(98)00109-1.

Cancer Research UK (2019) *Breast Cancer Risk*. Available at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/risk-factors> (Accessed: 3 April 2019).

Caton, P. W. *et al.* (2010) 'Metformin suppresses hepatic gluconeogenesis through induction of SIRT1 and GCN5.', *The Journal of endocrinology*, 205(1), pp. 97–106. doi: 10.1677/JOE-09-0345.

Chang, C.-J. and Hung, M.-C. (2012) 'The role of EZH2 in tumour progression.', *British journal of cancer*. Cancer Research UK, 106(2), pp. 243–7. doi: 10.1038/bjc.2011.551.

Chang, L., Chiang, S.-H. and Saltiel, A. R. 'Insulin signaling and the regulation of glucose transport.', *Molecular medicine (Cambridge, Mass.)*, 10(7–12), pp. 65–71. doi: 10.2119/2005-00029.Saltiel.

Clayton, A. L., Hazzalin, C. A. and Mahadevan, L. C. (2006) 'Enhanced Histone Acetylation and Transcription: A Dynamic Perspective', *Molecular Cell*. Cell Press, 23(3), pp. 289–296. doi: 10.1016/J.MOLCEL.2006.06.017.

Collins, J. C. *et al.* (2018) 'Ribosome biogenesis factor Ltv1 chaperones the assembly of the small subunit head', *The Journal of Cell Biology*, 217(12), pp. 4141–4154. doi: 10.1083/jcb.201804163.

Coppafeel! (2009) *Coppafeel*. Available at: <https://coppafeel.org/> (Accessed: 30 April 2019).

Creyghton, M. P. *et al.* (2010) 'Histone H3K27ac separates active from poised enhancers and predicts developmental state.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(50), pp. 21931–6. doi: 10.1073/pnas.1016071107.

CRUK (2015a) *Cancer Research UK: Breast Cancer Incidence UK*. Available at: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-Zero> (Accessed: 19 October 2015).

CRUK (2015b) *Cancer Research UK: Breast Cancer Survival UK*. Available at: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/survival#heading-Zero> (Accessed: 19 October 2015).

Cufi, S. *et al.* (2012) 'Metformin-induced preferential killing of breast cancer initiating CD44+CD24-/low cells is sufficient to overcome primary resistance to trastuzumab in HER2+ human breast cancer xenografts.', *Oncotarget*, 3(4), pp. 395–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3380574&tool=pmcentrez&rendertype=abstract> (Accessed: 6 October 2015).

- Cuyàs, E. *et al.* (2018) 'Metformin regulates global DNA methylation via mitochondrial one-carbon metabolism', *Oncogene*. Nature Publishing Group, 37(7), pp. 963–970. doi: 10.1038/onc.2017.367.
- Cuzick, J. *et al.* (2014) 'Anastrozole for prevention of breast cancer in high-risk postmenopausal women (IBIS-II): an international, double-blind, randomised placebo-controlled trial.', *Lancet (London, England)*. Elsevier, 383(9922), pp. 1041–8. doi: 10.1016/S0140-6736(13)62292-8.
- Dankner, R. *et al.* (2019) 'Metformin Treatment and Cancer Risk: Cox Regression Analysis, With Time-Dependent Covariates, of 320,000 Persons With Incident Diabetes Mellitus', *American Journal of Epidemiology*. doi: 10.1093/aje/kwz157.
- Dayeh, T. *et al.* (2014) 'Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion.', *PLoS genetics*, 10(3), p. e1004160. doi: 10.1371/journal.pgen.1004160.
- DeFronzo, R. A. and Tripathy, D. (2009) 'Skeletal muscle insulin resistance is the primary defect in type 2 diabetes.', *Diabetes care*, 32 Suppl 2, pp. S157-63. doi: 10.2337/dc09-S302.
- Dhawan, P. and Richmond, A. (2002) 'A novel NF-kappa B-inducing kinase-MAPK signaling pathway up-regulates NF-kappa B activity in melanoma cells.', *The Journal of biological chemistry*, 277(10), pp. 7920–8. doi: 10.1074/jbc.M112210200.
- Diabetes UK (2015) *Diabetes: Facts and Stats*, *diabetes.org*. Available at: [https://www.diabetes.org.uk/Documents/Position statements/Facts and stats June 2015.pdf](https://www.diabetes.org.uk/Documents/Position%20statements/Facts%20and%20stats%20June%202015.pdf) (Accessed: 4 May 2016).
- Diabetes UK (2016) *Diabetes Statistics UK*, *diabetes.org*. Available at: https://www.diabetes.org.uk/About_us/What-we-say/Statistics/2015-as-published-2016/ (Accessed: 4 May 2016).
- Dicker, D. (2011) 'DPP-4 inhibitors: impact on glycemic control and cardiovascular risk factors.', *Diabetes care*, 34 Suppl 2(Supplement_2), pp. S276-8. doi: 10.2337/dc11-s229.
- Dong, X. *et al.* (2006) 'Irs1 and Irs2 signaling is essential for hepatic glucose homeostasis and systemic growth.', *The Journal of clinical investigation*, 116(1), pp. 101–14. doi: 10.1172/JCI25735.

- Dowling, R. J. O. *et al.* (2015) 'Changes in insulin receptor signaling underlie neoadjuvant metformin administration in breast cancer: a prospective window of opportunity neoadjuvant study.', *Breast cancer research : BCR*. BioMed Central Ltd, 17(1), p. 32. doi: 10.1186/s13058-015-0540-0.
- Du, Y. *et al.* (2014) 'Metformin inhibits histone H2B monoubiquitination and downstream gene transcription in human breast cancer cells', *Oncology Letters*, 8(2), pp. 809–812. doi: 10.3892/ol.2014.2158.
- Ducker, G. S. and Rabinowitz, J. D. (2017) 'One-Carbon Metabolism in Health and Disease', *Cell Metabolism*, 25(1), pp. 27–42. doi: 10.1016/j.cmet.2016.08.009.
- Elsheikh, S. E. *et al.* (2009) 'Global Histone Modifications in Breast Cancer Correlate with Tumor Phenotypes, Prognostic Factors, and Patient Outcome', *Cancer Research*, 69(9), pp. 3802–3809. doi: 10.1158/0008-5472.CAN-08-3907.
- Farmer, R. E. *et al.* (2019) 'Metformin use and risk of cancer in patients with type 2 diabetes: a cohort study of primary care records using inverse probability weighting of marginal structural models.', *International journal of epidemiology*. Oxford University Press, 48(2), pp. 527–537. doi: 10.1093/ije/dyz005.
- Faubert, B. *et al.* (2014) 'Loss of the tumor suppressor LKB1 promotes metabolic reprogramming of cancer cells via HIF-1 α .', *Proceedings of the National Academy of Sciences of the United States of America*, 111(7), pp. 2554–9. doi: 10.1073/pnas.1312570111.
- Ferretti, A. C. *et al.* (2019) 'Metformin and glucose starvation decrease the migratory ability of hepatocellular carcinoma cells: targeting AMPK activation to control migration.', *Scientific reports*. Nature Publishing Group, 9(1), p. 2815. doi: 10.1038/s41598-019-39556-w.
- Fischer, Y. *et al.* (1995) 'Action of metformin on glucose transport and glucose transporter GLUT1 and GLUT4 in heart muscle cells from healthy and diabetic rats.', *Endocrinology*, 136(2), pp. 412–420. doi: 10.1210/endo.136.2.7835271.
- Florez, J. C. (2017) 'The pharmacogenetics of metformin', *Diabetologia*. Springer Berlin Heidelberg, 60(9), pp. 1648–1655. doi: 10.1007/s00125-017-4335-y.

- Franciosi, M. *et al.* (2013) 'Metformin therapy and risk of cancer in patients with type 2 diabetes: systematic review.', *PLoS one*. Public Library of Science, 8(8), p. e71583. doi: 10.1371/journal.pone.0071583.
- Fresno Vara, J. A. *et al.* (2004) 'PI3K/Akt signalling pathway and cancer.', *Cancer treatment reviews*, 30(2), pp. 193–204. doi: 10.1016/j.ctrv.2003.07.007.
- Galdieri, L. *et al.* (2016) 'Activation of AMP-activated Protein Kinase by Metformin Induces Protein Acetylation in Prostate and Ovarian Cancer Cells.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 291(48), pp. 25154–25166. doi: 10.1074/jbc.M116.742247.
- Garber, A. J. (2011) 'Long-acting glucagon-like peptide 1 receptor agonists: a review of their efficacy and tolerability.', *Diabetes care*, 34 Suppl 2(Supplement_2), pp. S279–84. doi: 10.2337/dc11-s231.
- García-Calzón, S. *et al.* (2017) 'Diabetes medication associates with DNA methylation of metformin transporter genes in the human liver', *Clinical Epigenetics*. BioMed Central, 9(1), p. 102. doi: 10.1186/s13148-017-0400-0.
- Geiger, A. M. *et al.* (2007) 'Quality of Life After Bilateral Prophylactic Mastectomy', *Annals of Surgical Oncology*. Springer-Verlag, 14(2), pp. 686–694. doi: 10.1245/s10434-006-9206-6.
- Ghoussaini, M. *et al.* (2014) 'Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation', *Nature Communications*, 5(1), p. 4999. doi: 10.1038/ncomms5999.
- Giannarelli, R. *et al.* (2003) 'Reducing insulin resistance with metformin: the evidence today.', *Diabetes & metabolism*, 29(4 Pt 2), pp. 6S28–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14502098> (Accessed: 21 June 2017).
- Gini, A. *et al.* (2016) 'Cancer among patients with type 2 diabetes mellitus: A population-based cohort study in northeastern Italy.', *Cancer epidemiology*, 41, pp. 80–87. doi: 10.1016/j.canep.2016.01.011.
- GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group *et al.* (2011) 'Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes', *Nature Genetics*, 43(2),

pp. 117–120. doi: 10.1038/ng.735.

Goldberg, Y. *et al.* (2015) 'Mutated MCM9 is associated with predisposition to hereditary mixed polyposis and colorectal cancer in addition to primary ovarian failure', *Cancer Genetics*, 208(12), pp. 621–624. doi: 10.1016/j.cancergen.2015.10.001.

Gong, L. *et al.* (2012) 'Metformin pathways: pharmacokinetics and pharmacodynamics.', *Pharmacogenetics and genomics*. NIH Public Access, 22(11), pp. 820–7. doi: 10.1097/FPC.0b013e3283559b22.

Goodwin, P. J. *et al.* (2002) 'Fasting Insulin and Outcome in Early-Stage Breast Cancer: Results of a Prospective Cohort Study', *Journal of Clinical Oncology*. American Society of Clinical Oncology, 20(1), pp. 42–51. doi: 10.1200/JCO.20.1.42.

GraphPad, P. (2019) 'Prism GraphPad software'. San Diego, California USA. Available at: www.graphpad.com.

Greer, E. L. and Shi, Y. (2012) 'Histone methylation: a dynamic mark in health, disease and inheritance', *Nature Reviews Genetics*. Nature Publishing Group, 13(5), pp. 343–357. doi: 10.1038/nrg3173.

Gunter, M. J. *et al.* (2009) 'Insulin, insulin-like growth factor-I, and risk of breast cancer in postmenopausal women.', *Journal of the National Cancer Institute*, 101(1), pp. 48–60. doi: 10.1093/jnci/djn415.

Gupta, S. *et al.* (2005) 'Role of NF-kappaB signaling pathway in increased tumor necrosis factor-alpha-induced apoptosis of lymphocytes in aged humans.', *Cell death and differentiation*, 12(2), pp. 177–83. doi: 10.1038/sj.cdd.4401557.

Ha, M. and Kim, V. N. (2014) 'Regulation of microRNA biogenesis.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 15(8), pp. 509–524. doi: 10.1038/nrm3838.

Hagemann, T. *et al.* (2008) "'Re-educating" tumor-associated macrophages by targeting NF-kappaB.', *The Journal of experimental medicine*, 205(6), pp. 1261–8. doi: 10.1084/jem.20080108.

- Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: the next generation.', *Cell*. Elsevier, 144(5), pp. 646–74. doi: 10.1016/j.cell.2011.02.013.
- Hauner, H. (2002) 'The mode of action of thiazolidinediones', *Diabetes/Metabolism Research and Reviews*. John Wiley & Sons, Ltd, 18(S2), pp. S10–S15. doi: 10.1002/dmrr.249.
- Hayer-Zillgen, M., Brüss, M. and Bönisch, H. (2002) 'Expression and pharmacological profile of the human organic cation transporters hOCT1, hOCT2 and hOCT3.', *British journal of pharmacology*, 136(6), pp. 829–36. doi: 10.1038/sj.bjp.0704785.
- He, L. and Wondisford, F. E. (2015) 'Metformin action: concentrations matter.', *Cell metabolism*. Elsevier, 21(2), pp. 159–162. doi: 10.1016/j.cmet.2015.01.003.
- Vander Heiden, M. G., Cantley, L. C. and Thompson, C. B. (2009) 'Understanding the Warburg effect: the metabolic requirements of cell proliferation.', *Science (New York, N.Y.)*, 324(5930), pp. 1029–33. doi: 10.1126/science.1160809.
- Helmrich, S. P. *et al.* (1991) 'Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus.', *The New England journal of medicine*, 325(3), pp. 147–52. doi: 10.1056/NEJM199107183250302.
- Hirsch, H. A., Iliopoulos, D. and Struhl, K. (2013) 'Metformin inhibits the inflammatory response associated with cellular transformation and cancer stem cell growth.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(3), pp. 972–7. doi: 10.1073/pnas.1221055110.
- Hollenbach, P. W. *et al.* (2010) 'A comparison of azacitidine and decitabine activities in acute myeloid leukemia cell lines.', *PloS one*. Public Library of Science, 5(2), p. e9001. doi: 10.1371/journal.pone.0009001.
- Hong, J.-L. *et al.* (2017) 'Comparative Effect of Initiating Metformin Versus Sulfonylureas on Breast Cancer Risk in Older Women.', *Epidemiology (Cambridge, Mass.)*. NIH Public Access, 28(3), pp. 446–454. doi: 10.1097/EDE.0000000000000635.
- Hopkins, T. A., Dyck, J. R. B. and Lopaschuk, G. D. (2003) 'AMP-activated protein kinase regulation of

fatty acid oxidation in the ischaemic heart.', *Biochemical Society transactions*, 31(Pt 1), pp. 207–12. doi: 10.1042/.

Hosio, M. *et al.* (2019) 'Association of antidiabetic medication and statins with breast cancer incidence in women with type 2 diabetes', *Breast Cancer Research and Treatment*. Springer, 175(3), pp. 741–748. doi: 10.1007/s10549-019-05185-0.

Hu, F. B. *et al.* (2001) 'Diet, lifestyle, and the risk of type 2 diabetes mellitus in women.', *The New England journal of medicine*, 345(11), pp. 790–7. doi: 10.1056/NEJMoa010492.

Isoda, K. *et al.* (2006) 'Metformin inhibits proinflammatory responses and nuclear factor-kappaB in human vascular wall cells.', *Arteriosclerosis, thrombosis, and vascular biology*, 26(3), pp. 611–7. doi: 10.1161/01.ATV.0000201938.78044.75.

Iwamoto, T. *et al.* (2011) 'BRCA1 promoter methylation in peripheral blood cells is associated with increased risk of breast cancer with BRCA1 promoter methylation.', *Breast cancer research and treatment*, 129(1), pp. 69–77. doi: 10.1007/s10549-010-1188-1.

Jansen, M. P. H. M. *et al.* (2012) 'High miR-26a and low CDC2 levels associate with decreased EZH2 expression and with favorable outcome on tamoxifen in metastatic breast cancer.', *Breast cancer research and treatment*, 133(3), pp. 937–47. doi: 10.1007/s10549-011-1877-4.

Jee, S. H. *et al.* (2005) 'Fasting serum glucose level and cancer risk in Korean men and women.', *JAMA*. American Medical Association, 293(2), pp. 194–202. doi: 10.1001/jama.293.2.194.

Jenuwein, T. and Allis, C. D. (2001) 'Translating the histone code.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 293(5532), pp. 1074–80. doi: 10.1126/science.1063127.

Jiang, G. and Zhang, B. B. (2003) 'Glucagon and regulation of glucose metabolism', *American Journal of Physiology - Endocrinology And Metabolism*, 284(4), pp. E671–E678. doi: 10.1152/ajpendo.00492.2002.

Jiang, M. *et al.* (2011) 'Hypermethylation of hepatic glucokinase and L-type pyruvate kinase promoters in high-fat diet-induced obese rats.', *Endocrinology*, 152(4), pp. 1284–9. doi:

10.1210/en.2010-1162.

Jung, J.-W. *et al.* (2011) 'Metformin represses self-renewal of the human breast carcinoma stem cells via inhibition of estrogen receptor-mediated OCT4 expression.', *PloS one*, 6(11), p. e28068. doi: 10.1371/journal.pone.0028068.

Kahn, S. E., Cooper, M. E. and Del Prato, S. (2014) 'Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future.', *Lancet (London, England)*. Elsevier, 383(9922), pp. 1068–83. doi: 10.1016/S0140-6736(13)62154-6.

Kamphorst, J. J. *et al.* (2014) 'Quantitative analysis of acetyl-CoA production in hypoxic cancer cells reveals substantial contribution from acetate', *Cancer & Metabolism*, 2(1), p. 23. doi: 10.1186/2049-3002-2-23.

Karlić, R. *et al.* (2010) 'Histone modification levels are predictive for gene expression.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(7), pp. 2926–31. doi: 10.1073/pnas.0909344107.

Karnevi, E. *et al.* (2013) 'Metformin-mediated growth inhibition involves suppression of the IGF-I receptor signalling pathway in human pancreatic cancer cells.', *BMC cancer*. BioMed Central, 13(1), p. 235. doi: 10.1186/1471-2407-13-235.

Karolina, D. S. *et al.* (2011) 'MicroRNA 144 impairs insulin signaling by inhibiting the expression of insulin receptor substrate 1 in type 2 diabetes mellitus.', *PloS one*. Public Library of Science, 6(8), p. e22839. doi: 10.1371/journal.pone.0022839.

Kassi, E. *et al.* (2011) 'Metabolic syndrome: definitions and controversies', *BMC Medicine*. BioMed Central, 9, p. 48. doi: 10.1186/1741-7015-9-48.

Kim, D. *et al.* (2013) 'TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions', *Genome Biology*. BioMed Central, 14(4), p. R36. doi: 10.1186/GB-2013-14-4-R36.

Kim, R.-J. and Nam, J.-S. (2011) 'OCT4 Expression Enhances Features of Cancer Stem Cells in a Mouse Model of Breast Cancer.', *Laboratory animal research*, 27(2), pp. 147–52. doi:

10.5625/lar.2011.27.2.147.

Koppes, L. L. J. *et al.* (2005) 'Moderate Alcohol Consumption Lowers the Risk of Type 2 Diabetes: A meta-analysis of prospective observational studies', *Diabetes Care*, 28(3), pp. 719–725. doi: 10.2337/diacare.28.3.719.

Korsse, S. E., Peppelenbosch, M. P. and van Veelen, W. (2013) 'Targeting LKB1 signaling in cancer.', *Biochimica et biophysica acta*, 1835(2), pp. 194–210. doi: 10.1016/j.bbcan.2012.12.006.

Kuhajda, F. P. (2008) 'AMP-activated protein kinase and human cancer: cancer metabolism revisited.', *International journal of obesity (2005)*. Macmillan Publishers Limited, 32 Suppl 4(S4), pp. S36–41. doi: 10.1038/ijo.2008.121.

Kuhl, C. K. *et al.* (2005) 'Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer', *Journal of Clinical Oncology*, 23(33), pp. 8469–8476. doi: 10.1200/JCO.2004.00.4960.

van de Laar, F. A. (2008) 'Alpha-glucosidase inhibitors in the early treatment of type 2 diabetes.', *Vascular health and risk management*, 4(6), pp. 1189–95. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2663450&tool=pmcentrez&rendertype=abstract> (Accessed: 10 March 2016).

Lampa, M. *et al.* (2017) 'Glutaminase is essential for the growth of triple-negative breast cancer cells with a deregulated glutamine metabolism pathway and its suppression synergizes with mTOR inhibition', *PLOS ONE*. Edited by M. Tan, 12(9), p. e0185092. doi: 10.1371/journal.pone.0185092.

Larsson, S. C., Mantzoros, C. S. and Wolk, A. (2007) 'Diabetes mellitus and risk of breast cancer: A meta-analysis', *International Journal of Cancer*, 121(4), pp. 856–862. doi: 10.1002/ijc.22717.

Lawrenson, K. *et al.* (2016) 'Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast–ovarian cancer susceptibility locus', *Nature Communications*, 7(1), p. 12675. doi: 10.1038/ncomms12675.

Lee, J. O. *et al.* (2012) 'Metformin regulates glucose transporter 4 (GLUT4) translocation through AMP-activated protein kinase (AMPK)-mediated Cbl/CAP signaling in 3T3-L1 preadipocyte cells.', *The*

Journal of biological chemistry. American Society for Biochemistry and Molecular Biology, 287(53), pp. 44121–9. doi: 10.1074/jbc.M112.361386.

Levy-Lahad, E. and Friedman, E. (2007) 'Cancer risks among BRCA1 and BRCA2 mutation carriers.', *British journal of cancer*. Nature Publishing Group, 96(1), pp. 11–5. doi: 10.1038/sj.bjc.6603535.

Li, X. *et al.* (2017) 'Nucleus-Translocated ACS2 Promotes Gene Transcription for Lysosomal Biogenesis and Autophagy', *Molecular Cell*, 66(5), pp. 684-697.e9. doi: 10.1016/j.molcel.2017.04.026.

Liang, L., Jiang, J. and Frank, S. J. (2000) 'Insulin receptor substrate-1-mediated enhancement of growth hormone-induced mitogen-activated protein kinase activation.', *Endocrinology*, 141(9), pp. 3328–36. doi: 10.1210/endo.141.9.7673.

Ling, C. *et al.* (2008) 'Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion.', *Diabetologia*, 51(4), pp. 615–22. doi: 10.1007/s00125-007-0916-5.

Linial, M., Miller, K. and Scheller, R. H. (1989) 'VAT-1: an abundant membrane protein from Torpedo cholinergic synaptic vesicles.', *Neuron*. Elsevier, 2(3), pp. 1265–73. doi: 10.1016/0896-6273(89)90311-5.

Lord, S. R. *et al.* (2015) 'Neoadjuvant Window Studies of Metformin and Biomarker Development for Drugs Targeting Cancer Metabolism.', *Journal of the National Cancer Institute. Monographs*, 2015(51), pp. 81–6. doi: 10.1093/jncimonographs/lgv011.

Lord, S. R. *et al.* (2018) 'Integrated Pharmacodynamic Analysis Identifies Two Metabolic Adaption Pathways to Metformin in Breast Cancer', *Cell Metabolism*. Cell Press. doi: 10.1016/J.CMET.2018.08.021.

Lowell, B. B. and Shulman, G. I. (2005) 'Mitochondrial dysfunction and type 2 diabetes.', *Science (New York, N.Y.)*, 307(5708), pp. 384–7. doi: 10.1126/science.1104343.

Luizon, M. R. *et al.* (2016) 'Genomic Characterization of Metformin Hepatic Response', *PLOS Genetics*. Edited by E. M. Mendenhall. Public Library of Science, 12(11), p. e1006449. doi: 10.1371/journal.pgen.1006449.

- Madiraju, A. K. *et al.* (2014) 'Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase.', *Nature*. NIH Public Access, 510(7506), pp. 542–6. doi: 10.1038/nature13270.
- Marin, T. L. *et al.* (2017) 'AMPK promotes mitochondrial biogenesis and function by phosphorylating the epigenetic factors DNMT1, RBBP7, and HAT1.', *Science signaling*. NIH Public Access, 10(464). doi: 10.1126/scisignal.aaf7478.
- Mato, E. P. M. *et al.* (2018) 'Genetic polymorphisms of organic cation transporters 1 (OCT1) and responses to metformin therapy in individuals with type 2 diabetes mellitus: a systematic review protocol', *Systematic Reviews*, 7(1), p. 105. doi: 10.1186/s13643-018-0773-y.
- Mauro, C. *et al.* (2011) 'NF- κ B controls energy homeostasis and metabolic adaptation by upregulating mitochondrial respiration', *Nature Cell Biology*. Europe PMC Funders, 13(10), pp. 1272–1279. doi: 10.1038/ncb2324.
- Mavaddat, N. *et al.* (2015) 'Prediction of Breast Cancer Risk Based on Profiling With Common Genetic Variants', *JNCI: Journal of the National Cancer Institute*, 107(5). doi: 10.1093/jnci/djv036.
- McGee, S. L. *et al.* (2008) 'AMP-Activated Protein Kinase Regulates GLUT4 Transcription by Phosphorylating Histone Deacetylase 5', *Diabetes*, 57(4), pp. 860–867. doi: 10.2337/db07-0843.
- McPherson, K., Steel, C. M. and Dixon, J. M. (2000) 'ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics.', *BMJ (Clinical research ed.)*, 321(7261), pp. 624–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1118507&tool=pmcentrez&rendertype=abstract> (Accessed: 20 January 2015).
- Mentch, S. J. and Locasale, J. W. (2016) 'One-carbon metabolism and epigenetics: understanding the specificity.', *Annals of the New York Academy of Sciences*. NIH Public Access, 1363(1), pp. 91–8. doi: 10.1111/nyas.12956.
- Messier, T. L. *et al.* (2016) 'Histone H3 lysine 4 acetylation and methylation dynamics define breast cancer subtypes.', *Oncotarget*. Impact Journals, LLC, 7(5), pp. 5094–109. doi:

10.18632/oncotarget.6922.

Miller, R. A. *et al.* (2013) 'Biguanides suppress hepatic glucagon signalling by decreasing production of cyclic AMP.', *Nature*, 494(7436), pp. 256–60. doi: 10.1038/nature11808.

Milne, J. C. *et al.* (2007) 'Small molecule activators of SIRT1 as therapeutics for the treatment of type 2 diabetes.', *Nature*. NATURE PUBLISHING GROUP, MACMILLAN BUILDING, 4 CRINAN ST, LONDON N1 9XW, ENGLAND, 450(7170), pp. 712–6. doi: 10.1038/nature06261.

Al Muftah, W. A. *et al.* (2016) 'Epigenetic associations of type 2 diabetes and BMI in an Arab population.', *Clinical epigenetics*, 8, p. 13. doi: 10.1186/s13148-016-0177-6.

National Health Service (NHS) (2019) *Breast cancer screening*. Available at:

<https://www.nhs.uk/conditions/breast-cancer-screening/> (Accessed: 30 April 2019).

Newman, A. C. and Maddocks, O. D. K. (2017) 'One-carbon metabolism in cancer'. doi: 10.1038/bjc.2017.118.

Newsholme, P. *et al.* (2003) 'Glutamine and glutamate?their central role in cell metabolism and function', *Cell Biochemistry and Function*. John Wiley & Sons, Ltd, 21(1), pp. 1–9. doi: 10.1002/cbf.1003.

Niraula, S. *et al.* (2012) 'Metformin in early breast cancer: a prospective window of opportunity neoadjuvant study.', *Breast cancer research and treatment*, 135(3), pp. 821–30. doi: 10.1007/s10549-012-2223-1.

Niu, N. *et al.* (2016) 'Metformin pharmacogenomics: a genome-wide association study to identify genetic and epigenetic biomarkers involved in metformin anticancer response using human lymphoblastoid cell lines', *Human Molecular Genetics*, p. ddw301. doi: 10.1093/hmg/ddw301.

Oliveras-Ferraro, C. *et al.* (2011) 'Micro(mi)RNA expression profile of breast cancer epithelial cells treated with the anti-diabetic drug metformin: induction of the tumor suppressor miRNA let-7a and suppression of the TGF β -induced oncomiR miRNA-181a.', *Cell cycle (Georgetown, Tex.)*, 10(7), pp. 1144–51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21368581> (Accessed: 16 March 2016).

Onitilo, A. A. *et al.* (2009) 'Breast cancer subtypes based on ER/PR and Her2 expression: comparison

of clinicopathologic features and survival.', *Clinical medicine & research*, 7(1–2), pp. 4–13. doi: 10.3121/cmr.2009.825.

Pawlyk, A. C. *et al.* (2014) 'Metformin pharmacogenomics: current status and future directions.', *Diabetes*. American Diabetes Association, 63(8), pp. 2590–9. doi: 10.2337/db13-1367.

Pernicova, I. and Korbonits, M. (2014) 'Metformin--mode of action and clinical implications for diabetes and cancer.', *Nature reviews. Endocrinology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 10(3), pp. 143–56. doi: 10.1038/nrendo.2013.256.

Petersen, M. C., Vatner, D. F. and Shulman, G. I. (2017) 'Regulation of hepatic glucose metabolism in health and disease.', *Nature reviews. Endocrinology*. NIH Public Access, 13(10), pp. 572–587. doi: 10.1038/nrendo.2017.80.

Pickup, J. C. (2004) 'Inflammation and Activated Innate Immunity in the Pathogenesis of Type 2 Diabetes', *Diabetes Care*, 27(3), pp. 813–823. doi: 10.2337/diacare.27.3.813.

Pietrocola, F. *et al.* (2015) 'Acetyl Coenzyme A: A Central Metabolite and Second Messenger', *Cell Metabolism*, 21(6), pp. 805–821. doi: 10.1016/j.cmet.2015.05.014.

Pimentel, H. *et al.* (2017) 'Differential analysis of RNA-seq incorporating quantification uncertainty', *Nature Methods*. Nature Publishing Group, 14(7), pp. 687–690. doi: 10.1038/nmeth.4324.

Polyak, K. (2007) 'Breast cancer: origins and evolution.', *The Journal of clinical investigation*. American Society for Clinical Investigation, 117(11), pp. 3155–63. doi: 10.1172/JCI33295.

Powell, S. N. and Kachnic, L. A. (2003) 'Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation', *Oncogene*. Nature Publishing Group, 22(37), pp. 5784–5791. doi: 10.1038/sj.onc.1206678.

Powles, T. J. *et al.* (2007) 'Twenty-Year Follow-up of the Royal Marsden Randomized, Double-Blinded Tamoxifen Breast Cancer Prevention Trial', *JNCI Journal of the National Cancer Institute*. Narnia, 99(4), pp. 283–290. doi: 10.1093/jnci/djk050.

Proks, P. *et al.* (2002) 'Sulfonylurea stimulation of insulin secretion.', *Diabetes*, 51 Suppl 3, pp. S368-

76. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12475777> (Accessed: 20 April 2016).
- Qiu, H. *et al.* (2013) 'Initial metformin or sulphonylurea exposure and cancer occurrence among patients with type 2 diabetes mellitus.', *Diabetes, obesity & metabolism*, 15(4), pp. 349–57. doi: 10.1111/dom.12036.
- Rapp, K. *et al.* (2006) 'Fasting blood glucose and cancer risk in a cohort of more than 140,000 adults in Austria.', *Diabetologia*, 49(5), pp. 945–52. doi: 10.1007/s00125-006-0207-6.
- Raychaudhuri, N. *et al.* (2008) 'Histone code modifications repress glucose transporter 4 expression in the intrauterine growth-restricted offspring.', *The Journal of biological chemistry*, 283(20), pp. 13611–26. doi: 10.1074/jbc.M800128200.
- Reitman, Z. J. and Yan, H. (2010) 'Isocitrate dehydrogenase 1 and 2 mutations in cancer: alterations at a crossroads of cellular metabolism.', *Journal of the National Cancer Institute*. Oxford University Press, 102(13), pp. 932–41. doi: 10.1093/jnci/djq187.
- Robertson, K. D. (2005) 'DNA methylation and human disease.', *Nature reviews. Genetics*, 6(8), pp. 597–610. doi: 10.1038/nrg1655.
- Rohland, N. and Reich, D. (2012) 'Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture.', *Genome research*. Cold Spring Harbor Laboratory Press, 22(5), pp. 939–46. doi: 10.1101/gr.128124.111.
- Rönn, T. *et al.* (2008) 'Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle.', *Diabetologia*, 51(7), pp. 1159–68. doi: 10.1007/s00125-008-1018-8.
- RStudio, T. (2019) 'RStudio: Integrated Development for R. RStudio', 2016. Inc., Boston, MA: RStudio, Inc. Available at: <http://www.rstudio.com/>.
- Ben Sahra, I. *et al.* (2008) 'The antidiabetic drug metformin exerts an antitumoral effect in vitro and in vivo through a decrease of cyclin D1 level.', *Oncogene*, 27(25), pp. 3576–86. doi: 10.1038/sj.onc.1211024.
- Ben Sahra, I. *et al.* (2011) 'Metformin, independent of AMPK, induces mTOR inhibition and cell-cycle arrest through REDD1.', *Cancer research*, 71(13), pp. 4366–72. doi: 10.1158/0008-5472.CAN-10-

1769.

Saladini, S. *et al.* (2019) 'Metformin Impairs Glutamine Metabolism and Autophagy in Tumour Cells', *Cells*, 8(1), p. 49. doi: 10.3390/cells8010049.

Salton, M. *et al.* (2010) 'Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response', *Cell Cycle*, 9(8), pp. 1568–1576. doi: 10.4161/cc.9.8.11298.

Schaffer, S. W. *et al.* (2010) 'Physiological roles of taurine in heart and muscle', *Journal of Biomedical Science*. BioMed Central, 17(Suppl 1), p. S2. doi: 10.1186/1423-0127-17-S1-S2.

Scherbakov, A. M. *et al.* (2016) 'The phenomenon of acquired resistance to metformin in breast cancer cells: The interaction of growth pathways and estrogen receptor signaling.', *IUBMB life*, 68(4), pp. 281–92. doi: 10.1002/iub.1481.

Scheuer, L. *et al.* (2002) 'Outcome of Preventive Surgery and Screening for Breast and Ovarian Cancer in *BRCA* Mutation Carriers', *Journal of Clinical Oncology*, 20(5), pp. 1260–1268. doi: 10.1200/JCO.2002.20.5.1260.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) 'NIH Image to ImageJ: 25 years of image analysis', *Nature Methods*, 9(7), pp. 671–675. doi: 10.1038/nmeth.2089.

Schug, Z. T. *et al.* (2015) 'Acetyl-CoA Synthetase 2 Promotes Acetate Utilization and Maintains Cancer Cell Growth under Metabolic Stress', *Cancer Cell*. Cell Press, 27(1), pp. 57–71. doi: 10.1016/J.CCELL.2014.12.002.

Segal, E. D. *et al.* (2011) 'Relevance of the OCT1 transporter to the antineoplastic effect of biguanides.', *Biochemical and biophysical research communications*, 414(4), pp. 694–9. doi: 10.1016/j.bbrc.2011.09.134.

Shang, F. *et al.* (2016) 'Cardiovascular Protective Effect of Metformin and Telmisartan: Reduction of PARP1 Activity via the AMPK-PARP1 Cascade.', *PloS one*, 11(3), p. e0151845. doi: 10.1371/journal.pone.0151845.

Sladek, R. *et al.* (2007) 'A genome-wide association study identifies novel risk loci for type 2 diabetes.', *Nature*, 445(7130), pp. 881–5. doi: 10.1038/nature05616.

Slawski, M., Daumer, M. and Boulesteix, A. L. (2008) 'CMA - A comprehensive Bioconductor package for supervised classification with high dimensional data', *BMC Bioinformatics*. NICE, 9, pp. 1–48. doi: 10.1186/1471-2105-9-439.

Sonnet, D. S. *et al.* (2016) 'Metformin inhibits Branched Chain Amino Acid (BCAA) derived ketoacidosis and promotes metabolic homeostasis in MSUD', *Scientific Reports*. Nature Publishing Group, 6(1), p. 28775. doi: 10.1038/srep28775.

Tang, G. H. *et al.* (2018) 'Association of metformin with breast cancer incidence and mortality in patients with type II diabetes: A GRADE-Assessed systematic review and meta-analysis', *Cancer Epidemiology Biomarkers and Prevention*. doi: 10.1158/1055-9965.EPI-17-0936.

Taylor, R. C., Cullen, S. P. and Martin, S. J. (2008) 'Apoptosis: controlled demolition at the cellular level.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 9(3), pp. 231–41. doi: 10.1038/nrm2312.

Terunuma, A. *et al.* (2014) 'MYC-driven accumulation of 2-hydroxyglutarate is associated with breast cancer prognosis', *Journal of Clinical Investigation*, 124(1), pp. 398–412. doi: 10.1172/JCI71180.

University of Arizona (2014) *Phase II Study of Metformin for Reduction of Obesity-Associated Breast Cancer Risk (NCT02028221)*, *ClinicalTrials.gov*. Available at:

<https://clinicaltrials.gov/ct2/show/NCT02028221?term=metformin+risk+cancer&draw=2&rank=2>
(Accessed: 19 September 2019).

Vicentini, M. *et al.* (2018) 'Effect of different glucose-lowering therapies on cancer incidence in type 2 diabetes: An observational population-based study', *Diabetes Research and Clinical Practice*, 143, pp. 398–408. doi: 10.1016/j.diabres.2018.04.036.

Vidal, S. J. *et al.* (2014) 'Targeting cancer stem cells to suppress acquired chemotherapy resistance.', *Oncogene*. Macmillan Publishers Limited, 33(36), pp. 4451–63. doi: 10.1038/onc.2013.411.

Vona-Davis, L., Howard-McNatt, M. and Rose, D. P. (2007) 'Adiposity, type 2 diabetes and the metabolic syndrome in breast cancer.', *Obesity reviews : an official journal of the International Association for the Study of Obesity*, 8(5), pp. 395–408. doi: 10.1111/j.1467-789X.2007.00396.x.

- Wahdan-Alaswad, R. S. *et al.* (2014) 'Metformin-induced killing of triple-negative breast cancer cells is mediated by reduction in fatty acid synthase via miRNA-193b.', *Hormones & cancer*, 5(6), pp. 374–89. doi: 10.1007/s12672-014-0188-8.
- Wallace, I. R. *et al.* (2013) 'Sex hormone binding globulin and insulin resistance.', *Clinical endocrinology*, 78(3), pp. 321–9. doi: 10.1111/cen.12086.
- Wang, L. and Weinshilboum, R. (2014) 'Metformin pharmacogenomics: biomarkers to mechanisms.', *Diabetes*, 63(8), pp. 2609–10. doi: 10.2337/db14-0609.
- Wellen, K. E. *et al.* (2009) 'ATP-Citrate Lyase Links Cellular Metabolism to Histone Acetylation', *Science*, 324(5930), pp. 1076–1080. doi: 10.1126/science.1164097.
- Whiteside, T. L. (2008) 'The tumor microenvironment and its role in promoting tumor growth.', *Oncogene*. Macmillan Publishers Limited, 27(45), pp. 5904–12. doi: 10.1038/onc.2008.271.
- Willi, C. *et al.* (2007) 'Active Smoking and the Risk of Type 2 Diabetes', *JAMA*, 298(22), p. 2654. doi: 10.1001/jama.298.22.2654.
- Wood-Trageser, M. A. *et al.* (2014) 'MCM9 mutations are associated with ovarian failure, short stature, and chromosomal instability.', *American journal of human genetics*. Elsevier, 95(6), pp. 754–62. doi: 10.1016/j.ajhg.2014.11.002.
- World Health Organization (2019) *Classification of Diabetes Mellitus, WHO*. Available at: <https://www.who.int/publications-detail/classification-of-diabetes-mellitus> (Accessed: 8 October 2019).
- Wu, L. *et al.* (2011) '5'-AMP-activated protein kinase (AMPK) regulates progesterone receptor transcriptional activity in breast cancer cells.', *Biochemical and biophysical research communications*, 416(1–2), pp. 172–7. doi: 10.1016/j.bbrc.2011.11.018.
- Wu, L. *et al.* (2018) 'A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer', *Nature Genetics*, 50(7), pp. 968–978. doi: 10.1038/s41588-018-0132-x.
- Xu, W. *et al.* (2011) 'Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α -

- Ketoglutarate-Dependent Dioxygenases', *Cancer Cell*, 19(1), pp. 17–30. doi: 10.1016/j.ccr.2010.12.014.
- Yan, L. *et al.* (2015) 'Regulation of tumor cell migration and invasion by the H19/let-7 axis is antagonized by metformin-induced DNA methylation.', *Oncogene*. Macmillan Publishers Limited, 34(23), pp. 3076–84. doi: 10.1038/onc.2014.236.
- Yang, B. T. *et al.* (2011) 'Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA(1c) levels in human pancreatic islets.', *Diabetologia*, 54(2), pp. 360–7. doi: 10.1007/s00125-010-1967-6.
- Yang, B. T. *et al.* (2012) 'Increased DNA methylation and decreased expression of PDX-1 in pancreatic islets from patients with type 2 diabetes.', *Molecular endocrinology (Baltimore, Md.)*, 26(7), pp. 1203–12. doi: 10.1210/me.2012-1004.
- Yang, C. *et al.* (2014) 'Glutamine oxidation maintains the TCA cycle and cell survival during impaired mitochondrial pyruvate transport.', *Molecular cell*. NIH Public Access, 56(3), pp. 414–24. doi: 10.1016/j.molcel.2014.09.025.
- Yang, X.-J. and Seto, E. (2007) 'HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention', *Oncogene*, 26(37), pp. 5310–5318. doi: 10.1038/sj.onc.1210599.
- Yang, Y.-X., Hennessy, S. and Lewis, J. D. (2004) 'Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients.', *Gastroenterology*, 127(4), pp. 1044–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15480982> (Accessed: 20 October 2015).
- Zakikhani, M. *et al.* (2006) 'Metformin is an AMP kinase-dependent growth inhibitor for breast cancer cells.', *Cancer research*, 66(21), pp. 10269–73. doi: 10.1158/0008-5472.CAN-06-1500.
- Zampetaki, A. *et al.* (2010) 'Plasma MicroRNA Profiling Reveals Loss of Endothelial MiR-126 and Other MicroRNAs in Type 2 Diabetes', *Circulation Research*, 107(6), pp. 810–817. doi: 10.1161/CIRCRESAHA.110.226357.
- Zhang, S. *et al.* (2016) 'Roles of microRNA-124a and microRNA-30d in breast cancer patients with

type 2 diabetes mellitus.', *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. doi: 10.1007/s13277-016-4981-6.

Zhao, S. *et al.* (2016) 'ATP-Citrate Lyase Controls a Glucose-to-Acetate Metabolic Switch.', *Cell reports*. NIH Public Access, 17(4), pp. 1037–1052. doi: 10.1016/j.celrep.2016.09.069.

Zhong, T. *et al.* (2017) 'Metformin alters DNA methylation genome-wide via the H19/SAHH axis', *Oncogene*. Nature Publishing Group, 36(17), pp. 2345–2354. doi: 10.1038/onc.2016.391.

Zhou, K. *et al.* (2016) 'Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin', *Nature Genetics*. Nature Publishing Group, 48(9), pp. 1055–1059. doi: 10.1038/ng.3632.

Zhu, P. *et al.* (2014) 'Metformin selectively targets tumor-initiating cells in ErbB2-overexpressing breast cancer models.', *Cancer prevention research (Philadelphia, Pa.)*, 7(2), pp. 199–210. doi: 10.1158/1940-6207.CAPR-13-0181.

Zhuang, Y. *et al.* (2014) 'Mechanisms by which low glucose enhances the cytotoxicity of metformin to cancer cells both in vitro and in vivo.', *PloS one*. Public Library of Science, 9(9), p. e108444. doi: 10.1371/journal.pone.0108444.

Zhuang, Y. and Miskimins, W. K. (2008) 'Cell cycle arrest in Metformin treated breast cancer cells involves activation of AMPK, downregulation of cyclin D1, and requires p27^{Kip1} or p21^{Cip1}', *Journal of Molecular Signaling*, 3, p. 18. doi: 10.1186/1750-2187-3-18.

Zhuang, Y. and Miskimins, W. K. (2011) 'Metformin induces both caspase-dependent and poly(ADP-ribose) polymerase-dependent cell death in breast cancer cells.', *Molecular cancer research : MCR*, 9(5), pp. 603–15. doi: 10.1158/1541-7786.MCR-10-0343.

APPENDIX

	Author	Study Type	Result	Interpretation
Hyperinsulinemia	Gunter <i>et al.</i>	Prospective	HR = 2.40 (95% CI, 1.30 – 4.41) [Q4 vs. Q1]	2-fold increased risk of women with high insulin developing breast cancer
Hyperinsulinemia	Kabat <i>et al.</i>	Prospective	HR = 2.22 (95% CI, 1.39 – 3.53) [T3 vs. T1]	2-fold increased risk of women with high insulin developing breast cancer
Hyperinsulinemia	Gunter <i>et al.</i>	Prospective	HR = 2.06 (95% CI, 1.01–4.22) [Q4 Vs. Q1]	2-fold increased risk of metabolically unhealthy women with high insulin developing breast cancer
Hyperinsulinemia	Del Giudice <i>et al.</i>	Case Control Study	OR = 2.83 (95% CI, 1.22– 6.58) [Q4 Vs. Q1]	2-fold increased risk of women with high insulin developing breast cancer
Hyperglycaemia	Stattin <i>et al.</i>	Prospective	RR = 1.06 (95% CI, 0.82 – 1.37) [Q4 vs. Q1] RR = 2.12 (95% CI, 1.2 – 4.1) [Q4 vs. Q1]	6% increased risk of breast cancer incidence 2-fold increased risk of breast cancer incidence
Hyperglycaemia	Muti <i>et al.</i>	Prospective	RR = 1.63 (95% CI, 0.59 – 4.46) [Q4 vs. Q1] RR = 2.76 (95% CI, 1.18, 6.46) [Q4 Vs. Q1]	63% increased risk of breast cancer incidence 2-fold increased risk of breast cancer incidence
Hyperglycaemia	Rapp <i>et al.</i>	Prospective	HR = 1.38 (95% CI, 1.02 – 1.86) [Q4 vs. Q1]	38% increased risk of breast cancer incidence

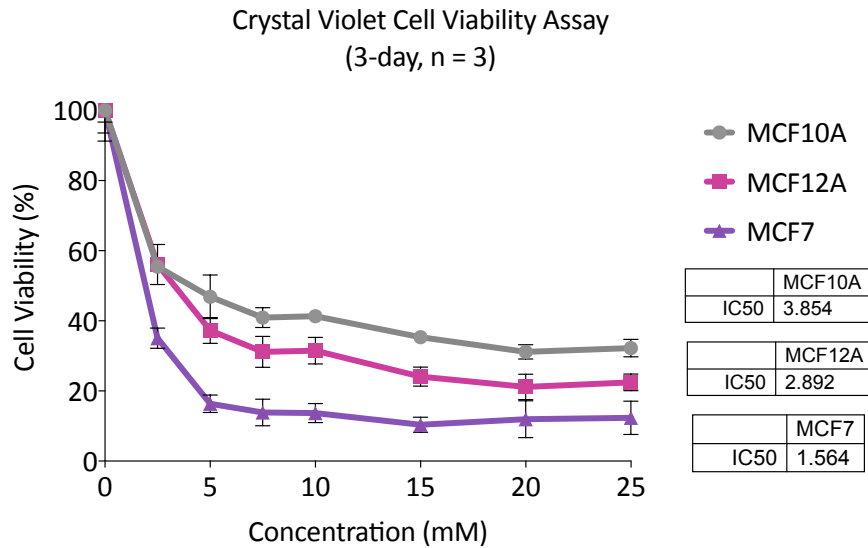
Appendix table 1.1 Summary of epidemiological studies linking hyperinsulinemia and hyperglycaemia with incident cases of Breast Cancer.

Calponin PCR primers:	
Forward Primer	5'-GGAAGGTAGTTGAGGTTGTG-3'
Reverse Primer	5'-CCCAAACCTCAAACCTAACCTAAC-3'

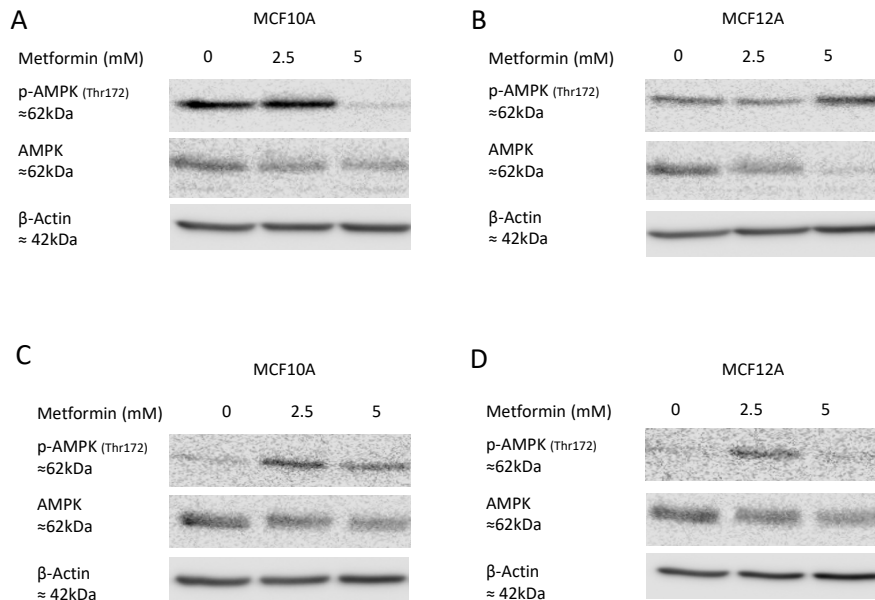
Appendix Table 2.1: Calponin PCR primer sequences

LINE-1 primers:	
Forward PCR Primer	5'-GGATTTTTTGAGTTAGGTGTGGG-3'
Reverse PCR Primer	5'-BIOTIN-CAAAAAATCAAAAAATCCCTTTCC-3'
Sequencing Primer	5'-AGGTGTGGGATATAGT-3'

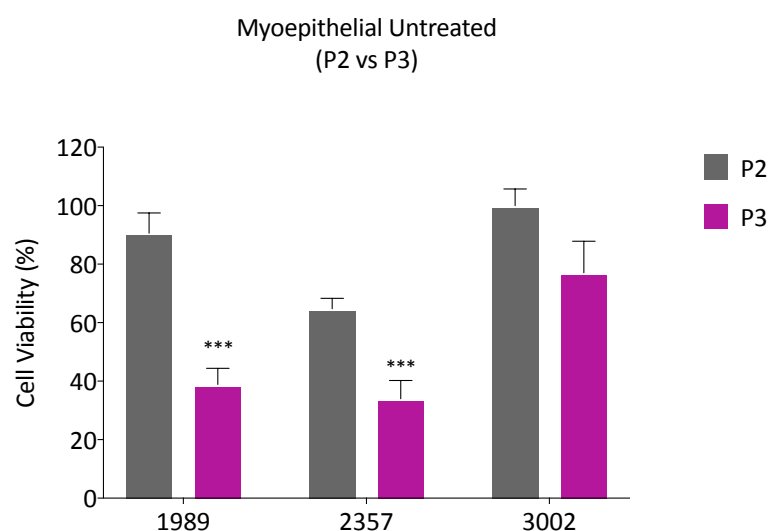
Appendix Table 2.2: LINE-1 PCR and pyrosequencing primer sequences



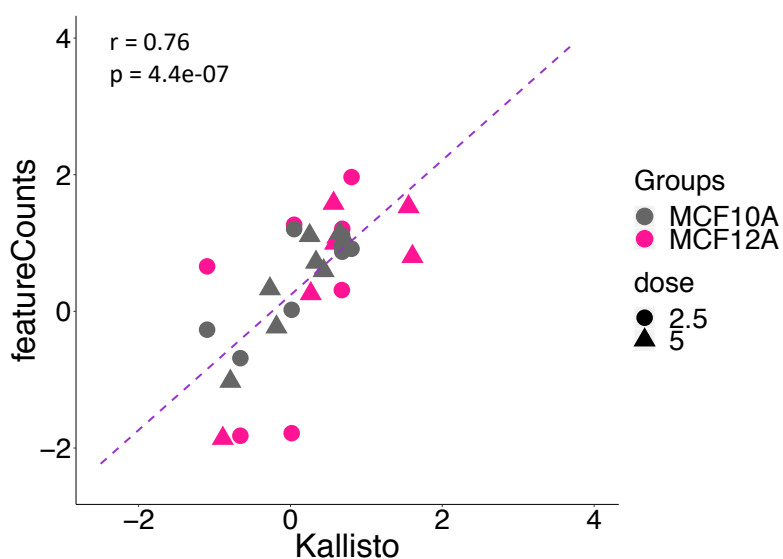
Appendix Figure 3.1 Cell viability assay of MCF12A and MCF10A non-cancerous breast epithelial cell lines and MCF7 breast epithelial cell line treated for 3-days with a range of metformin concentrations (0mM – 25mM).



Appendix Figure 3.2 Western blot analysis of p-AMPK and AMPK in non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. Relative abundance measured as ratio to beta actin. (A) MCF10A n=2, (B) MCF12A n=2, (C) MCF10A n=3, (D) MCF12A n=3.



Appendix Figure 3.3 Crystal violet cell viability assay of three patient normal matched breast epithelial and myoepithelial primary cells grown for 7 days. Comparison of P2 and P3 myoepithelial primary cells. Student's t-test * ≤ 0.05 , ** ≤ 0.01 , *** ≤ 0.001 .



Appendix Figure 3.4 Correlation plot of values from seven genes chosen from RNA-sequencing Kallisto data in comparison to featureCounts data (Pearson correlation $r=0.7$, $p<0.05$).

Locus	Gene	Type	Risk alleles	Zscores	Ref
9p21.3	MIR31HG	lncRNA	reduce expression	-5.02	(Wu et al., 2018)
12p13.33	RP11-218M22.1	lncRNA	increase expression	5.02	
15q24.2	CTD-2323K18.1f	lncRNA	reduce expression	-4.65	
1p11.2	RP11-439A17.7	lncRNA	reduce expression	-5.34	
3q12.2	RP11-114I8.4	lncRNA	reduce expression	-5.84	
5p12	RP11-530I9.1	lncRNA	increase expression	10.38	
5q33.3	RP11-32D16.1	lncRNA	reduce expression	-5.41	
6q23.1	RP11-73O6.3f	lncRNA	reduce expression	-6.61	
11p15.5	AP006621.6g	lncRNA	increase expression	5.61	
14q32.33	CTD-3051D23.1	lncRNA	reduce expression	-5.06	
16q12.2	RP11-467J12.4	lncRNA	increase expression	8.04	
16q12.2	CTD-3032H12.1	lncRNA	increase expression	4.92	
17q21.31	KANSL1-AS1g	lncRNA	reduce expression	-5.58	
17q21.31	LINC00671	lncRNA	reduce expression	-5.85	
19q13.31	RP11-15A1.7h	lncRNA	increase expression	5.45	
2q35	IGFBP5	Protein	reduce expression	-	(Ghoussaini et al., 2014)
19p13.11	ANKLE1	Protein	increase expression	-	(Lawrenson et al., 2016)
19p13.11	ADHD8	Protein	increase expression	-	
6q25.1	ESR1	Protein	reduce expression	-	(Dunning et al., 2016)
6q25.1	RMND1	Protein	reduce expression	-	
6q25.1	CCDC170	Protein	reduce expression	-	
1p34.1	ZSWIM5	Protein	increase expression	5.26	(Wu et al., 2018)
3p24.1	LRRC3B	Protein	reduce expression	-9.57	
4q12	SPATA18	Protein	reduce expression	-4.62	
6p22.1	UBD	Protein	reduce expression	-4.87	
7q32.2	KLHDC10	Protein	increase expression	5.21	
11p15.5	RIC8A	Protein	reduce expression	-5.27	
11q13.2	B3GNT1	Protein	reduce expression	-5.85	
14q24.1	GALNT16	Protein	reduce expression	-8.27	
14q24.1	PLEKHD1	Protein	increase expression	7.5	
15q24.2	MAN2C1f	Protein	reduce expression	-5.32	
1q21.1	NUDT17	Protein	reduce expression	-6.27	
1q21.1	ANKRD34A	Protein	reduce expression	-5.05	
2p23.1-2p23.2	ALK	Protein	increase expression	4.67	
3p21.31	PRSS46	Protein	reduce expression	-5.83	
5q33.3	UBLCP1	Protein	increase expression	5.93	
6p22.2	BTN3A2	Protein	increase expression	4.61	
11p15.5	RPLP2g	Protein	increase expression	4.64	
17q21.31	LRRC37Ag	Protein	reduce expression	-5.89	
17q21.31	CRHR1g	Protein	reduce expression	-5.29	
17q21.31	LRRC37A2	Protein	reduce expression	-5.77	
19p13.11	HAPLN4	Protein	reduce expression	-7.13	
11q13.2	RP11-867G23.10	Transcript	increase expression	4.71	

Appendix Table 3.1: Breast cancer risk alleles

Cell Line	Dose	Target ID	Ensemble ID	Gene Name	p-value	q_value	b
MCF12A	2.5mM	ENST00000345739	ENSG00000115966	ATF2	9.58E-07	0.00012655	4.42770683
MCF12A	2.5mM	ENST00000293379	ENSG00000161638	ITGA5	5.04E-14	2.68E-11	0.78021762
MCF12A	2.5mM	ENST00000269217	ENSG00000053747	LAMA3	0.12869219	0.65272851	0.15132809
MCF12A	2.5mM	ENST00000264144	ENSG00000058085	LAMC2	1.49E-10	4.20E-08	0.65223561
MCF12A	2.5mM	ENST00000311852	ENSG00000157227	MMP14	4.97E-20	5.76E-17	1.28046468
MCF12A	2.5mM	ENST00000357066	ENSG00000172081	MOB3A	3.44E-06	0.00038891	-1.3331694
MCF12A	2.5mM	ENST00000357214	ENSG00000116560	SFPQ	8.38E-34	2.06E-30	-1.5612592
MCF12A	2.5mM	ENST00000355653	ENSG00000108828	VAT1	6.83E-12	2.49E-09	0.7432944
MCF12A	5mM	ENST00000345739	ENSG00000115966	ATF2	9.84E-07	0.000102	4.42284018
MCF12A	5mM	ENST00000293379	ENSG00000161638	ITGA5	7.55E-20	6.46E-17	0.94479301
MCF12A	5mM	ENST00000269217	ENSG00000053747	LAMA3	0.2491141	0.7276822	0.11479534
MCF12A	5mM	ENST00000264144	ENSG00000058085	LAMC2	1.14E-11	3.40E-09	0.69107101
MCF12A	5mM	ENST00000311852	ENSG00000157227	MMP14	5.30E-15	3.07E-12	1.09254106
MCF12A	5mM	ENST00000357066	ENSG00000172081	MOB3A	1.49E-12	5.23E-10	-2.0317343
MCF12A	5mM	ENST00000357214	ENSG00000116560	SFPQ	1.54E-30	3.79E-27	-1.4797894
MCF12A	5mM	ENST00000355653	ENSG00000108828	VAT1	5.62E-08	8.71E-06	0.588312
MCF10A	2.5mM	ENST00000345739	ENSG00000115966	ATF2	NA	NA	NA
MCF10A	2.5mM	ENST00000293379	ENSG00000161638	ITGA5	8.54E-09	0.00013958	0.91259888
MCF10A	2.5mM	ENST00000269217	ENSG00000053747	LAMA3	8.15E-06	0.04435311	0.7327205
MCF10A	2.5mM	ENST00000264144	ENSG00000058085	LAMC2	9.41E-08	0.00076843	0.93292181
MCF10A	2.5mM	ENST00000311852	ENSG00000157227	MMP14	0.04449917	0.99991568	0.67877691
MCF10A	2.5mM	ENST00000357066	ENSG00000172081	MOB3A	0.94035038	0.99991568	0.01582324
MCF10A	2.5mM	ENST00000357214	ENSG00000116560	SFPQ	0.14970978	0.99991568	-0.3921899
MCF10A	2.5mM	ENST00000355653	ENSG00000108828	VAT1	0.03444035	0.99991568	0.68211643
MCF10A	5mM	ENST00000345739	ENSG00000115966	ATF2	NA	NA	NA
MCF10A	5mM	ENST00000293379	ENSG00000161638	ITGA5	1.05E-07	0.00085988	0.84285706
MCF10A	5mM	ENST00000269217	ENSG00000053747	LAMA3	1.60E-06	0.00871143	0.78807155
MCF10A	5mM	ENST00000264144	ENSG00000058085	LAMC2	2.30E-08	0.00037555	0.97662453
MCF10A	5mM	ENST00000311852	ENSG00000157227	MMP14	0.04507728	0.99973223	0.67694413
MCF10A	5mM	ENST00000357066	ENSG00000172081	MOB3A	0.38410021	0.99973223	-0.1840453
MCF10A	5mM	ENST00000357214	ENSG00000116560	SFPQ	0.00661685	0.99973223	-0.7393014
MCF10A	5mM	ENST00000355653	ENSG00000108828	VAT1	0.17529376	0.99973223	0.43715661

Appendix Table 3.2: Genes chosen for validation by RT-qPCR from Kallisto analysed data.

Gene	Primer	Primer Sequence
ATF2	Forward	GCATGGTTGTGCCTTCTTTT
ATF2	Reverse	TCTGCCTTGGAGGTTGAACT
ITGA5	Forward	ATCCTAGCCATCCTGTTTGG
ITGA5	Reverse	TGAAATTGGGAGGACTCAGG
LAMA3	Forward	ACACCGTCCGGTATACAAGC
LAMA3	Reverse	CTGGGCTACAGTTCACAGCA
LAMC2	Forward	CAGCACCTTCCAGAATCACA
LAMC2	Reverse	CGCAGCTCTGCAGAATACAG
MMP14	Forward	TCACGGATGTAGGCATAGGG
MMP14	Reverse	GCAGAAGTTTTACGGCTTGC
MOB3A	Forward	TGCACGTCTACATCCACCAC
MOB3A	Reverse	GGGCTCTCAGTGGCACAT
SFPQ	Forward	CGTCAACGTGAGATGGAAGA
SFPQ	Reverse	CTTCCCATCATGGAACCACT
VAT1	Forward	ATCAAGCCCCACATTGACTC
VAT1	Reverse	TCTCCCTTCGCTGGTCTCTA
RNF20	Forward	CCTGTCGTCGTTCCAGTTTT
RNF20	Reverse	GGAAGTTCTGCCTTGCTCC

Appendix Table 3.3: Primer information genes chosen for validation by RT-qPCR from Kallisto analysed data

Gene	Breast Cancer Risk Alleles Derived from Breast Cancer Models	Metformin Treated Non-Cancerous Breast Epithelial Cell Line MCF10A	Metformin LogFold Change	Metformin p<0.05	Metformin FDR<0.05
IGFBP5	reduce expression	increase expression	1.01147509	0.01109795	0.41547329
ZSWIM5	increase expression	reduce expression	-0.2735336	0.03669922	0.41547329
ANKRD34A	reduce expression	increase expression	0.08447733	0.03645666	0.41547329

Appendix Table 3.4: Comparison of expression of known breast cancer risk genes in breast cancer models vs non-cancerous breast epithelial cell line MCF12A treated with metformin

cg Site	Meth Diff Non-Cancerous	Value (Beta)	Value (M)	P-Value	FDR	Chromosome	Location	USCS Gene Name
cg12510923	-0.0551658	-0.0181494	-0.1340219	0.0000003	0.0332602	X	Island	HDAC6
cg02459704	-0.1209100	-0.0188855	-0.1199178	0.0000003	0.0332602	X	Island	PDZD4
cg13482208	0.0143677	-0.0051477	-0.1090956	0.0000011	0.0459102	X	Island	PDZD4
cg22016202	0.0001089	0.00077053	0.1141642	0.0000012	0.0459102	8	Island	NRG1
cg09183496	0.0003967	0.01357286	0.2061454	0.0000000	0.0186153	5	OpenSea	DOCK2;FAM196B
cg01757206	0.0065484	0.00129141	0.0599993	0.0000021	0.0488485	17	N_Shore	SLC2A4
cg24775300	-0.0086581	-0.0045338	-0.0593874	0.0000017	0.0488485	12	OpenSea	TEAD4
cg16019875	-0.0070608	-0.0025118	-0.0829423	0.0000022	0.0488485	15	OpenSea	ZNF710

Appendix Table 4.1 Model 1 CpG site information selected for pyrosequencing validation

CpG ID	Primer	Primer Sequence
cg01757206	Forward	TAGGTTTGGATATGTTGATTATTAGAGTTG
cg01757206	Reverse	CTAACCCAAACCTTCTCTAAA
cg01757206	Sequencing	ATGTTGATTATTAGAGTTGTTATAT
cg02459704	Forward	AAAATAGTTGTGGTTTGGGAGATT
cg02459704	Reverse	AACCCCTAAAAATAATAACTTCACC
cg02459704	Sequencing	AACTTCACCTCCCATCC
cg09183496	Forward	AGTTAAGATTTTTAATATTAGAGGGTATGG
cg09183496	Reverse	ATCACCAAAAATAACTCAATAAATACCA
cg09183496	Sequencing	ATTAGAGGGTATGGGTTAAT
cg12510923	Forward	GGAAYGGGGTAGTTTTTTGAGGAG
cg12510923	Reverse	CACTTAATCCACCCCTTATC
cg12510923	Sequencing	GGAGYGGGGTTGGTTGAAA
cg13482208	Forward	ATGTTTGGTAGGGAAATGAATTAGG
cg13482208	Reverse	AAACCCCTTCTCTCTCT
cg13482208	Sequencing	GGTATTTATTTTTTTGTTTTAG
cg16019875	Forward	ATTTTTTGGGGGGAGGGATAGT
cg16019875	Reverse	CCCCTCAACTCACTACATAAAAAATCAT
cg16019875	Sequencing	AACRACCRCATATATCTAATATCTA
cg24775300	Forward	GGGGTAGTTAGGTATTGGATAAG
cg24775300	Reverse	ACCTCCTAAAACTCTACTCAATATCC
cg24775300	Sequencing	CAATATCCRAACTCCACA

Appendix Table 4.2 Model 1 CpG site pyrosequencing primer information

cg Site	Meth Diff MCF10A	Meth Diff MCF12A	Value (Beta)	Std. Error	P-Value	FDR	Chromosome	Location	USCS Gene Name
cg11302943	0.0163398	0.6258736	-0.0643302	0.1696474	0.7081752	0.8498103	10	Island	CRTAC1
cg17263026	0.3416333	-0.0175997	-0.3465436	0.1672572	0.0501981	0.3011883	2	OpenSea	
cg18441082	-0.3846218	-0.3589245	0.2643760	0.2236327	0.2497512	0.7492536	5	Island	ZNF622
cg01447350	-0.6164187	-0.2948961	-0.0790644	0.1926719	0.6855125	0.8498103	16	OpenSea	IL34
cg19360943	0.1441783	0.3490847	0.0543944	0.1096048	0.6246207	0.8498103	12	OpenSea	ING4
cg12393503	-0.2875502	0.5555717	-0.0334012	0.2705381	0.9028621	0.9028621	3	Island	ECT2

Appendix Table 4.3 Model 2 CpG site information selected for pyrosequencing validation

CpG ID	Primer	Primer Sequence
cg11302943	Forward	GGGATTGYGGTGGTAGTAGT
cg11302943	Reverse	AAACRCCCTCTCCCCTCCTC
cg11302943	Sequencing	AGTYGGTAGTTAGGTTGGTT
cg17263026	Forward	TTGATTTTTGGGTATAGAAAGTTTT
cg17263026	Reverse	TCCACCCACTTCTTATTCCATTATCC
cg17263026	Sequencing	AAGAATATYGTAATTTTT
cg18441082	Forward	GTTGTGGTTGAAAGGGATTG
cg18441082	Reverse	AACRCCCCACTCCCCATT
cg18441082	Sequencing	ATTTGGGTTYGTTTAGTTAG
cg01447350	Forward	GGGTTTGAATAAAAGTAGGATTAAGG
cg01447350	Reverse	CCTTAACCTCCCAAATACTAAAATTACAA
cg01447350	Sequencing	CTAAAATTACAAACATAAACCATTA
cg19360943	Forward	AGATTTAGGAAGTTTATGGTAAGTGAAG
cg19360943	Reverse	AACCATACCATCTCATAAATCTACATAACA
cg19360943	Sequencing	GGTAAGTGTAAGGAATTGGT
cg12393503	Forward	GTGAGGGYGAAGATGTTTAGTAATATTGG
cg12393503	Reverse	CCCCAACTCAAACCTCCC
cg12393503	Sequencing	TGGGGGTTTYGGTTTTTAT

Appendix Table 4.2 Model 2 CpG site pyrosequencing primer information

cg Site	Average Diff MCF10A 2.5mM (Beta Value)	Average Diff MCF10A 5mM (Beta Value)	Average Diff MCF12A 2.5mM (Beta Value)	Average Diff MCF12A 5mM (Beta Value)	Chromosome	Location	USCS Gene Name
cg07372659	-0.1700667	-0.1883791	0.06951033	0.01042284	6	OpenSea	CDKAL1
cg14067425	-0.1118779	-0.1455202	0.00557528	0.00997604	7	OpenSea	PION

Appendix Table 4.5 Model 3 CpG site information selected for pyrosequencing validation

CpG ID	Primer	Primer Sequence
cg07372659	Forward	TGATGTAAAATGGTGTATTGTTATAGT
cg07372659	Reverse	TTTCTAACTCCTCAAATATTCACACAA
cg07372659	Sequencing	AATTATTATGTAGTTAGGAATT
cg14064105	Forward	GTGGGGTGGATATTTATTTTATAGT
cg14064105	Reverse	CTTAAATAATCCTTCCTACTTCTTCTATCT
cg14064105	Sequencing	TTTAAAAATATTTTAATAAGTAAG

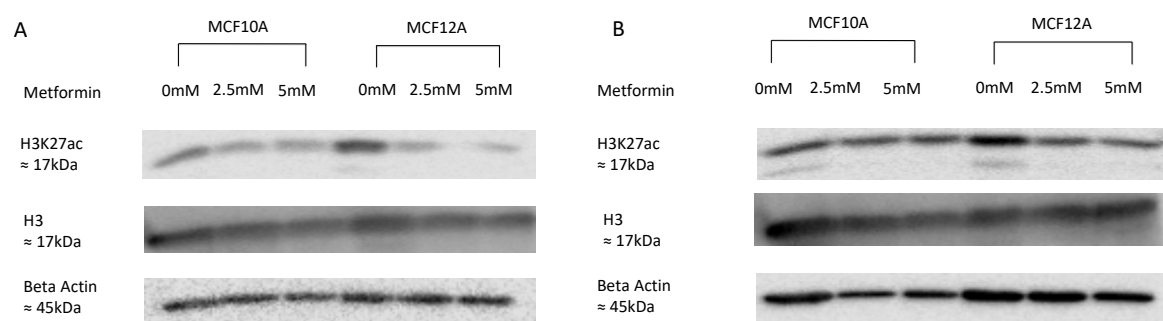
Appendix Table 4.6 Model 3 CpG site pyrosequencing primer information

Cell Line	Dose	Target ID	Ensemble ID	Gene Name	p-value	q_value	b
MCF12A	2.5mM	ENST00000306434	ENSG00000168906	MAT2A	1.01E-06	0.00013227	-3.4914615
MCF12A	2.5mM	ENST00000376583	ENSG00000177000	MTHFR	0.655789	0.97560936	-0.512225
MCF12A	2.5mM	ENST00000366577	ENSG00000116984	MTR	0.04292069	0.42555452	-2.5034362
MCF12A	2.5mM	ENST00000513439	ENSG00000124275	MTRR	0.35306769	0.86914062	-0.9936744
MCF12A	5mM	ENST00000306434	ENSG00000168906	MAT2A	1.03E-06	0.00010593	-3.4896413
MCF12A	5mM	ENST00000376583	ENSG00000177000	MTHFR	0.19894853	0.67202427	-1.4761705
MCF12A	5mM	ENST00000366577	ENSG00000116984	MTR	0.25341543	0.73055961	-1.4122912
MCF12A	5mM	ENST00000513439	ENSG00000124275	MTRR	0.00101125	0.03127688	-3.5175407
MCF10A	2.5mM	ENST00000306434	ENSG00000168906	MAT2A	0.42272519	0.99991568	-0.5204844
MCF10A	2.5mM	ENST00000376583	ENSG00000177000	MTHFR	0.7458158	0.99991568	0.43972786
MCF10A	2.5mM	ENST00000366577	ENSG00000116984	MTR	0.79715746	0.99991568	-0.0824651
MCF10A	2.5mM	ENST00000513439	ENSG00000124275	MTRR	0.85985232	0.99991568	0.25476478
MCF10A	5mM	ENST00000306434	ENSG00000168906	MAT2A	0.02885422	0.99973223	-1.4188613
MCF10A	5mM	ENST00000376583	ENSG00000177000	MTHFR	0.30509813	0.99973223	1.39118369
MCF10A	5mM	ENST00000366577	ENSG00000116984	MTR	0.9293416	0.99973223	-0.02845
MCF10A	5mM	ENST00000513439	ENSG00000124275	MTRR	0.8390453	0.99973223	-0.2930781

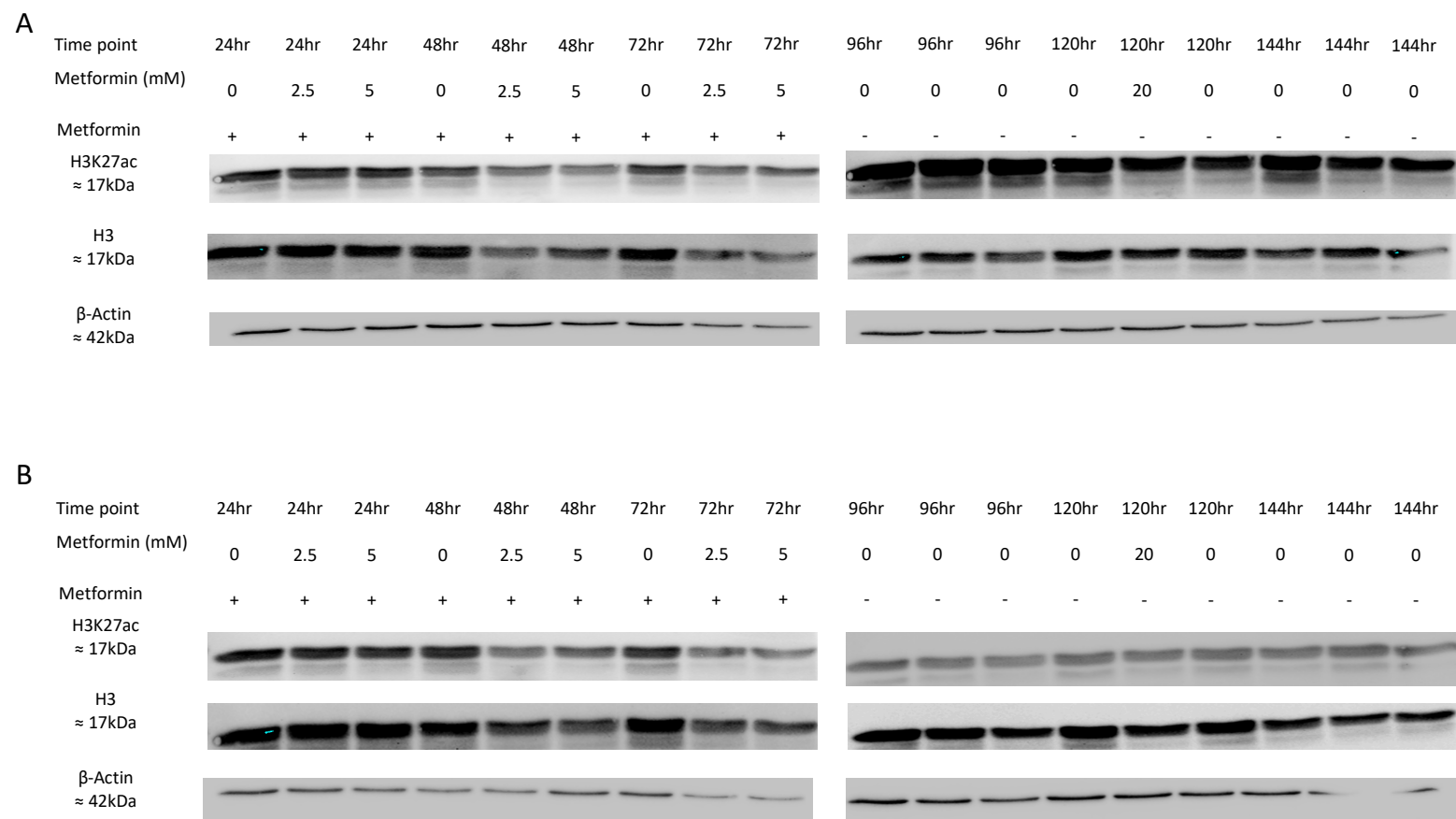
Appendix Table 4.7: Gene information obtained from Kallisto analysed data – one carbon metabolism pathway genes

Gene	Primer	Primer Sequence
MAT2A	Forward	ACCGTTCAGCTGCTTATGCT
MAT2A	Reverse	CCAACAAGTCTGGGGAAAAA
MTRR	Forward	GAAATCAACCGGCAGTGTTT
MTRR	Reverse	CGCAAGGTTGGTGGAAGT
MTHFR	Forward	GAAGTACGAGCTCCGGGTTA
MTHFR	Reverse	GCAGTTGTCCAGTGGGAAGT
MTR	Forward	GACCTCGGAAGTGTTCTTCG
MTR	Reverse	ACCTGTGGAGAGCACGTCTT

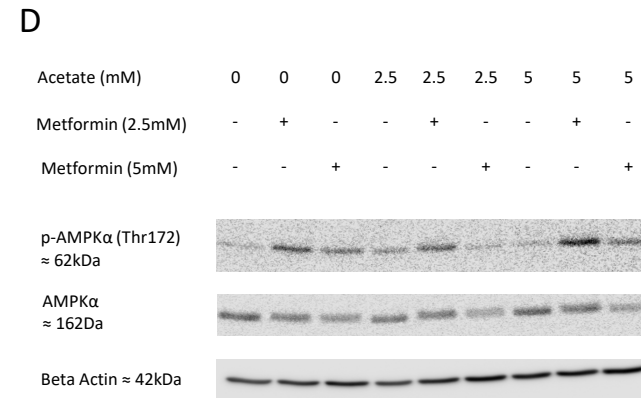
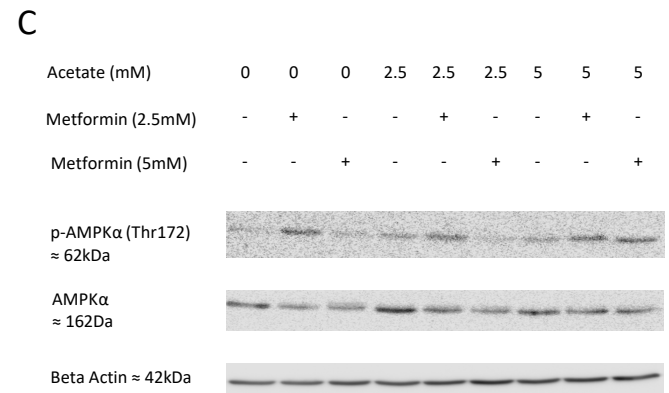
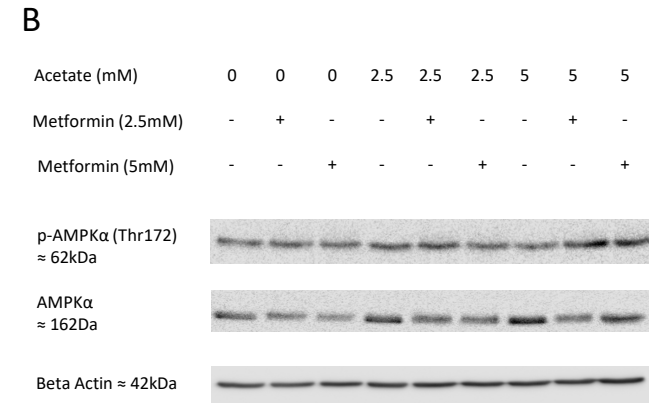
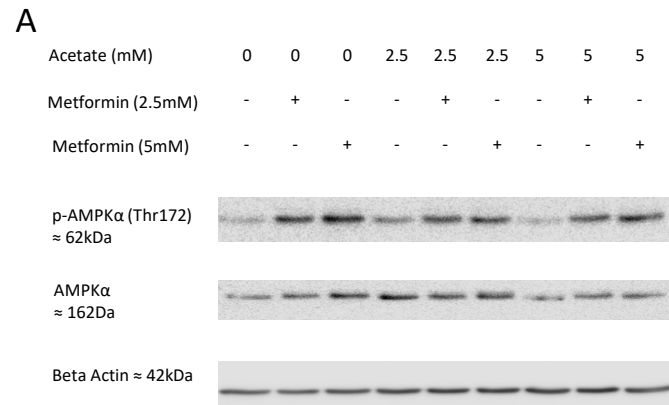
Appendix Table 4.8: Primer information genes one carbon metabolism pathway



Appendix Figure 5.1 Western blot of H3K27ac and H3 in non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin. (A) n=2, (B) n=3.

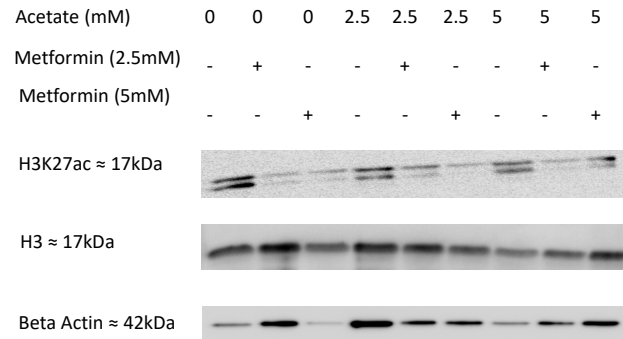


Appendix Figure 5.2 Western blot of H3K27ac and H3 in MCF12A non-cancerous breast epithelial cell lines treated 144 hours with 0mM, 2.5mM and 5mM metformin. (A) n=2 and (B) n=3

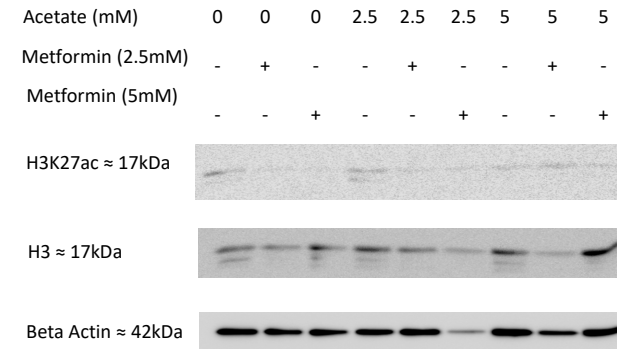


Appendix Figure 5.3 Western blot of p-AMPK and AMPK in non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. (A) MCF10A n=2, (B) MCF10A n=3, (C) MCF12A n=2, (D) MCF12A n=3.

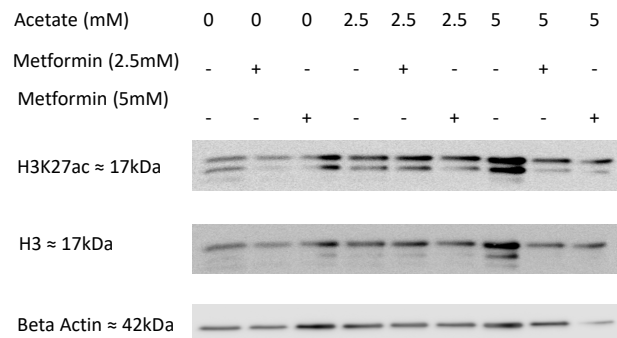
A



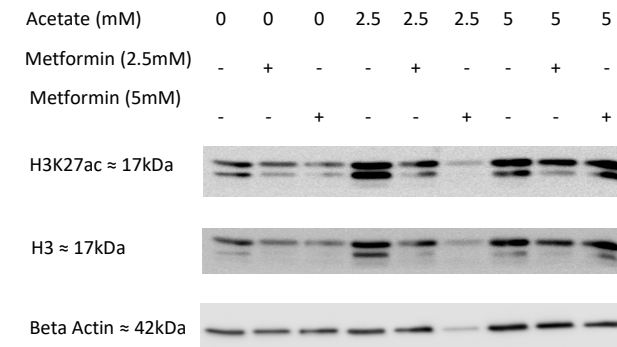
B



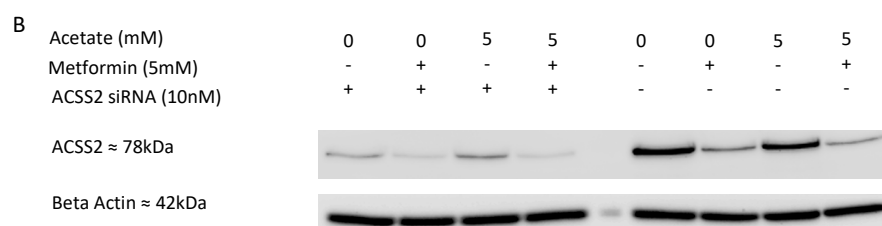
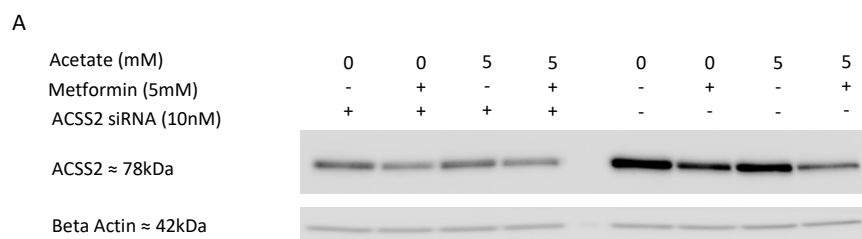
C



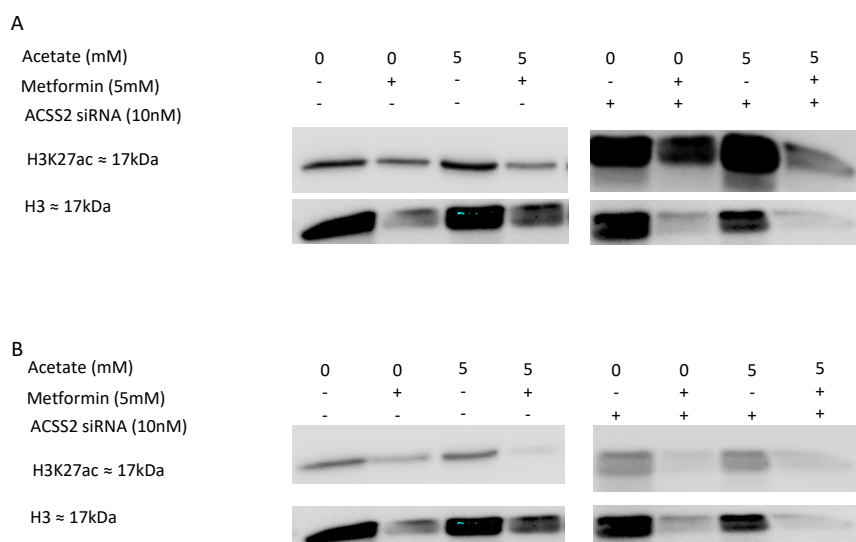
D



Appendix Figure 5.4 Western blot of H3K27ac in non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. (A) MCF10A n=2, (B) MCF10A n=3, (C) MCF12A n=2, (D) MCF12A n=3.



Appendix Figure 5.5 Western blot of ACSS2 in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Treated with 10nm siRNA for AceS1 (ACSS2). (A) n=2 and (B) n=3



Appendix Figure 5.6 Western blot of H3K27ac in MCF12A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin and 0mM, 2.5mM and 5mM acetate. Treated with 10nm siRNA for AceS1 (ACSS2). (A) n=2 and (B) n=3

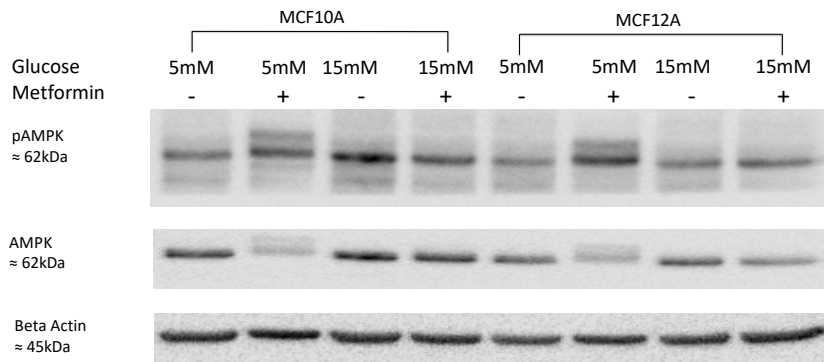
Cell Line	Dose	Target ID	Ensemble ID	Gene Name	p-value	q-value	b
MCF10A	2.5mM	ENST00000352035	ENSG00000131473	ACLY	0.63310239	0.99991568	-0.4196905
MCF10A	2.5mM	ENST00000360596	ENSG00000131069	ACSS2	0.72953656	0.99991568	0.5388811
MCF10A	2.5mM	ENST00000540249	ENSG00000131828	PDHA1	NA	NA	NA
MCF10A	5mM	ENST00000352035	ENSG00000131473	ACLY	0.59999164	0.99973223	-0.4610541
MCF10A	5mM	ENST00000360596	ENSG00000131069	ACSS2	0.55156359	0.99973223	-0.9280326
MCF10A	5mM	ENST00000540249	ENSG00000131828	PDHA1	NA	NA	NA
MCF12A	2.5mM	ENST00000352035	ENSG00000131473	ACLY	0.38077561	0.88781584	0.13651374
MCF12A	2.5mM	ENST00000360596	ENSG00000131069	ACSS2	0.4189231	0.90601221	1.0231564
MCF12A	2.5mM	ENST00000540249	ENSG00000131828	PDHA1	0.32641705	0.85475506	1.2811817
MCF12A	5mM	ENST00000352035	ENSG00000131473	ACLY	0.44442722	0.85194864	0.11911138
MCF12A	5mM	ENST00000360596	ENSG00000131069	ACSS2	0.40307176	0.82813919	-1.0584183
MCF12A	5mM	ENST00000540249	ENSG00000131828	PDHA1	0.09475533	0.4962187	-2.1813038

Appendix Table 5.1: Gene information obtained from Kallisto analysed data – acetyl co-A generation pathway genes

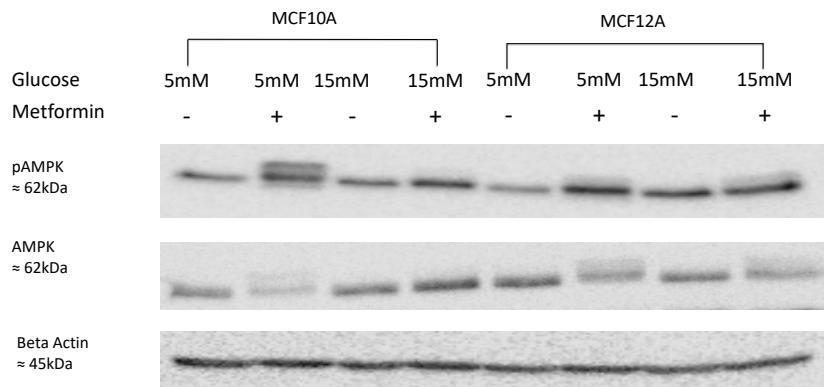
Gene	Primer	Primer Sequence
ACLY	Forward	TCATCGGAGTCGCATTGTGA
ACLY	Reverse	CAGTAGGGTTCCTGGCTCTG
ACSS2	Forward	GGGCATGGAGTCTTCCAGTA
ACSS2	Reverse	CGCGGTTCTAGGAAGTCTGAC
PDHA1	Forward	GCTGTTACCATCCTGTCCT
PDHA1	Reverse	GGTTTGCTGCTGCCTATTGT

Appendix Table 5.2: Primer information genes from acetyl co-A generation pathway

A

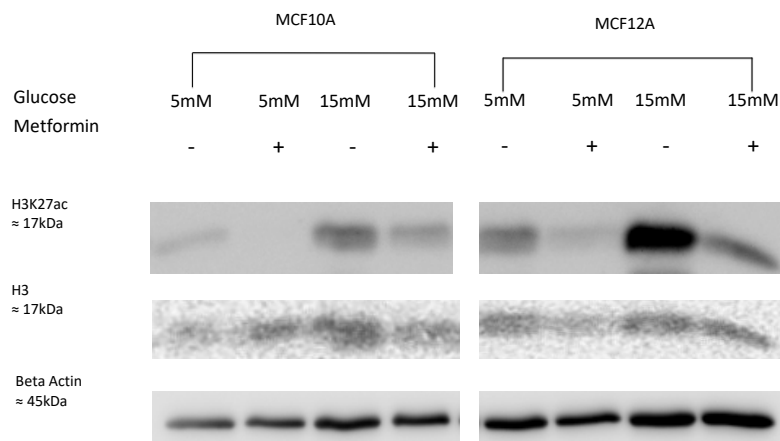


B

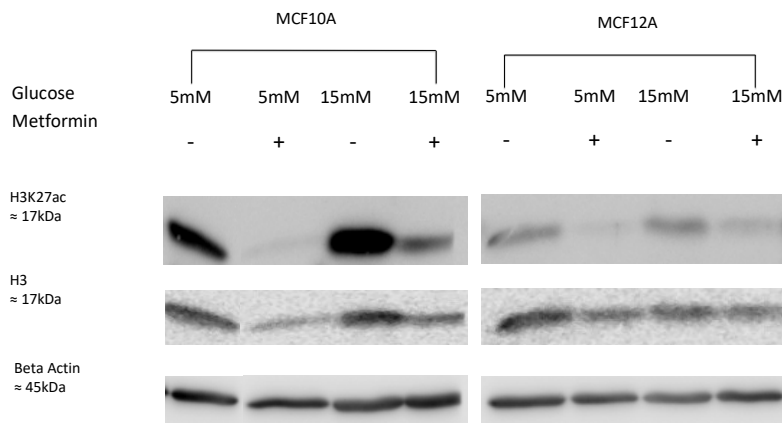


Appendix Figure 6.1 Additional two western blots for p-AMPK and AMPK for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Used for measurement of relative abundance. Biological replicates. Loading control beta actin.

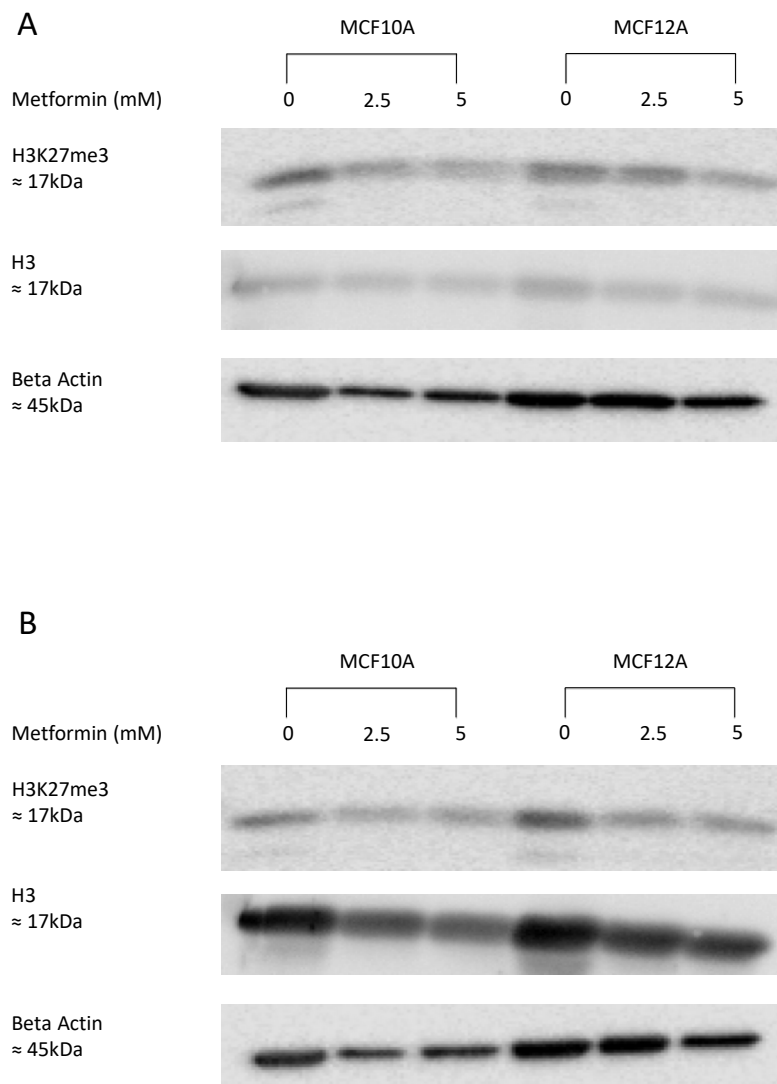
A



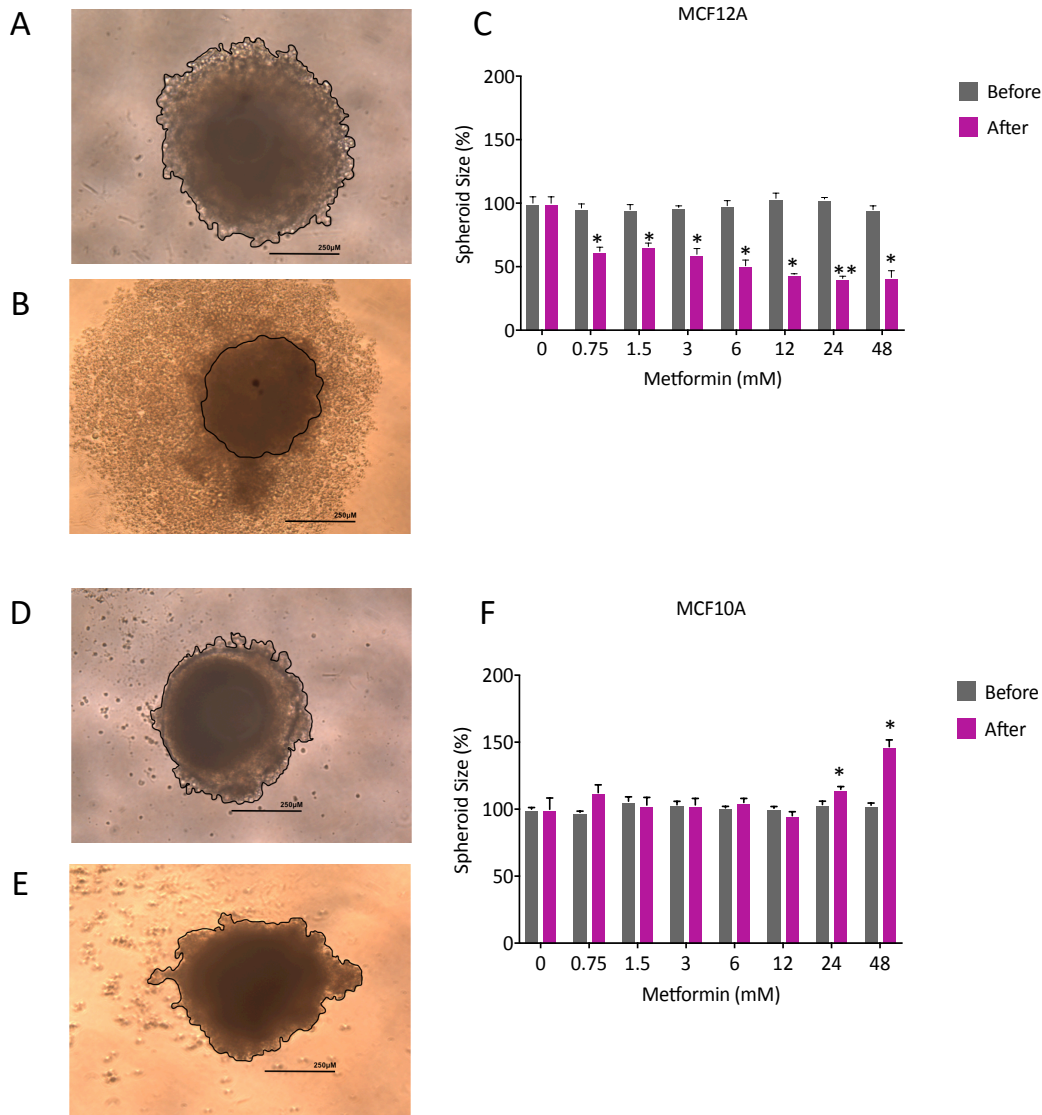
B



Appendix figure 6.2 Additional two western blot for H3K27ac and H3 for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM, 5mM and 15mM glucose. Used for measurement of relative abundance. Biological replicates. Loading control beta actin.



Appendix figure 6.3 Additional two western blot for H3K27me3 and H3 for MCF12A and MCF10A non-cancerous breast epithelial cell lines treated 72 hours with 0mM, 2.5mM and 5mM metformin in 0mM. Used for measurement of relative abundance. Biological replicates. Loading control beta actin.



Appendix figure 6.4 Analysis of spheroid formation in non-cancerous breast epithelial cell lines treated for 7-days with 24mM metformin. (A) Image of MCF12A spheroid grown for 4 days after formation before metformin treatment. (B) MCF12A following 7-day treatment with 24mM metformin treatment. Size was determined using the area of the spheroid (outlined in black). (C) Analysis of MCF12A comparing before and after metformin treatment (n = 3). (D) Image of MCF10A spheroid grown for 4 days after formation before metformin treatment. (E) MCF10A following 7-day treatment with 24mM metformin treatment. Size was determined using the area of the spheroid (outlined in black). (F) Analysis of MCF10A comparing before and after metformin treatment (n = 3). Student's t-test * ≤ 0.05 .