THE EPIGENOME OF HUMAN ADIPOCYTES

Hilal Varinli

This thesis is presented as partial fulfilment of the requirements

for the degree of Doctor of Philosophy

Department of Biological Sciences, Macquarie University

Food and Nutrition Flagship, CSIRO

Genomics and Epigenetics Division, Garvan Institute of Medical Research

January 2017

STATEMENT OF CANDIDATE

I certify that the work in this thesis entitled “The Epigenome of Human Adipocytes” has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree to any other university or institution other than Macquarie University.

I also certify that the thesis in an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself has been appropriately acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis.

Hilal Varinli

January 2017

ACKNOWLEDGEMENTS

There are many people who I would like to thank for their support and encouragement during the last five years.

First of all, I would like to thank my supervisors and mentors Dr. Peter Molloy, Professor Michael Gillings and Professor Susan Clark for their excellent guidance for research and their commitment to my development as a scientist. I feel very privileged to have worked with you and thank you sincerely for this opportunity and your friendship. I also thank you for your guidance, patience and kindness throughout my degree.

Most of my work would not have been possible without the expertise of Dr. Jason Ross, Dr. Tim Peters, Dr. Pascal Vallotton, Dr. Susan van Dijk, Dr. Stephen Bradford, Dr. Michael Buckley, Professor Reginald Lord, Professor Katherine Samaras and Dr. Musarat Ishaq. Thank you all for being generous with your time and advices.

I am grateful for the support and positive contribution of all of my colleagues in all three institutions: Glen Brown, Thu Ho Nguyen, Sue Mitchell, Yalchin Oytam, Anthony Rowe and Linda Lockett at the CSIRO Food and Nutrition Flagship; Marita Holley, Liette Waldron and Louise Chow the EMMA Lab at Macquarie University; and Aaron Statham, Dr. Shalima Nair, Wenjia Quo, Jenny Song and Dr. Clare Stirzaker in the Epigenetics Lab at the Garvan Institute of Medical Research.

I also thank Dr. Vijay Tiwari at the Institute of Molecular Biology, Mainz Germany for hosting my visit as part of the International Summer School Programme for PhD students. Additionally, participation of all subjects in the clinical study was greatly appreciated.

I am also grateful for the generous financial support provided by a number of sources: the Science and Industry Endowment Fund (Australia), grant RP03-064; the Macquarie University MQRES Scholarship; and the CSIRO OCE Postgraduate Top-up Scholarship.

This work would not have been possible without the full support and patience of my family. Mum, thank you for your patience and unconditional love. You are the source of my courage that led me to pursue my dreams. Dad, I know you are proud of me. I am proud of being your daughter and thank you for all the memorable moments you left to behind. Thank you to my sister and brothers for your endless support and love. I also want to specially thank my two adorable nieces (Mina and Alya) for bringing joy to all our lives.

I am blessed to have friends that became my family in Australia (Gulay & Selim, Nirvana & Goktug). Thank you for your incredible support and friendship. Thank you for allowing me to share anything and everything with you. Special mention to Yagiz Alp Aksoy, Sumit Salaria and my best friend Verena Sturm for always reminding me that learning is a life-long journey.

Finally, Ozenc Koyuncu, you have witnessed the late stages of my PhD journey. Thank you for your support, sarcastic jokes, proofreading and working with me on the weekends.

SUMMARY

The increasing incidence of obesity is a major public health crisis worldwide. Obesity is linked to a number of poor health outcomes such as Type 2 Diabetes (T2D), cardiovascular disease and cancer. Thus, the obesity epidemic threatens to lower life expectancy and reduce health-related life quality of current and future generations. It places a significant economic burden on healthcare systems. A better understanding of the factors that contribute to the development of obesity and how fat tissue contributes to these adverse effects are essential to improve prevention and treatment strategies.

In this thesis, I examined the potential role of epigenetics as a mediator of gene-environment interactions in manifesting functional changes in fat cells, and whether these processes could be targeted to prevent and treat obesity.

Firstly, I investigated whether exposure of human fat cells (adipocytes) to high glucose induces physiological and metabolic changes (Chapter 1). I then examined whether nuclear processes such as the genomic landscape of transcriptional and epigenetic regulation were involved in the manifestation of these changes. I used genome wide microarray analysis of both DNA methylation and transcription during the differentiation of human adipocytes *in vitro*. I showed that adipocytes grown in high glucose had significant changes in gene expression, particularly in mitochondrial and oxidative stress pathways, and that some of these changes were associated with epigenetic changes.

Secondly, increased lipid storage in adipocytes is a major contributor to the increase in fat in humans, and these lipids are stored as lipid droplets within adipocytes (Chapter 2). I developed a novel label-free lipid droplet quantification technique (LipiD-QuanT) to monitor fat accumulation in live adipocytes. The current techniques used to quantify fat accumulation during *in vitro* studies are based on either enzymatic digest or lipophilic staining. However, LipiD-Quant is label free and applicable to live cells. Additionally, I demonstrated the application of LipiD-QuanT to measure the effect of four potential pro- or anti-obesogenic

substances during human pre-adipocyte differentiation: Docosahexaenoic acid (DHA), rosiglitazone, D-glucose and Zinc oxide (ZnO) nanoparticles.

Thirdly, since I began my PhD, epigenetics analysis methods have continued to improve. While most common methods target locations of the genome where methylation occurs densely, there is growing evidence that methylation in less dense genomic locations is equally, if not more important. Hence, I developed a new genome wide reduced methylome method (COBRA-Seq) (Chapter 3). This method enriches methylated DNA fractions by digesting the genomic DNA with restriction enzymes that recognize potential methylation sites after bisulfite conversion. The genomic complexity is further reduced by removing DNA fragments without the enzyme recognition site using streptavidin coated magnetic beads. COBRA-seq samples the genome proportionally for the generation and analysis of genome-scale DNA methylation profiles with nucleotide resolution.

This proved an excellent method to study representative, genome wide DNA methylation profiles at reduced cost. I compared the functionality of our novel method with other methylome sampling methods and reviewed its other potential applications in the field.

I co-authored a systematic review on the most recent findings in the research area of human epigenetics and obesity (Appendix 1). I defined critical transcriptional and epigenetic signatures of human visceral adipose tissue (VAT) and purified visceral adipocytes (VA) associated with key parameters of obesity. For this part of my research, I coordinated a clinical study and used biological samples from 26 females across a range of BMIs (19-50 kg/m^2) (Chapter 4). In the context of a larger team (the EpiSCOPE project) I coordinated the collection of these samples. This research revealed that transcriptional profiles were best correlated with BMI, waist circumference and hip circumference among the range of anthropometric variables.

Additionally, I participated in the multi-institutional EpiSCOPE project to increase understanding of epigenetic changes in the emergence of obesity, and generated the first methylome maps of subcutaneous and visceral adipocytes in lean and obese individuals (Appendix 2). I contributed to two more scientific publications during my PhD that were not directly related to my PhD (Appendices 3 and 4).

TABLE OF CONTENTS

LIST OF PUBLICATIONS

Publications presented in this thesis

This thesis is based on the following publications which are ordered as the respective chapters of the thesis.

1. **Varinli H**, Peters TJ, Vallotton P, van Dijk SJ, Swarbrick MM, Gillings MR, Clark SJ and Molloy PL: Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis (Prepared to be submitted to *Diabetes*).
2. **Varinli H**, Osmond-McLeod MJ, Molloy PL, Vallotton P: LipiD-QuanT: a novel method to quantify lipid accumulation in live cells. *J Lipid Res* 2015;56:2206-2216.
3. **Varinli H**, Statham AL, Clark SJ, Molloy PL, Ross JP: COBRA-Seq: Sensitive and Quantitative Methylome Profiling. *Genes* 2015;6:1140-1163.
4. **Varinli H**, Nair S, Anwar F, van Dijk SJ, Luu P, Peters TJ, Lord RV, Samaras K, Bauer D, Buckley M, Ross JP, Molloy PL, Clark SJ and Bradford S. Transcriptome and epigenome profiling of visceral adipose tissue and visceral adipocytes from 26 females with a range of BMI (Prepared to be submitted to *Diabetes*).

Additional publications prepared during my PhD studies (Appendices)

1. van Dijk SJ, Molloy PL, **Varinli H**, Morrison JL, Muhlhausler BS, EpiSCOPE: Epigenetics and human obesity. *Int J Obesity* 2015;39:85-97 (Thompson ISI highly cited)
2. Bradford ST, … **Varinli H**, *et al*. Methylome of visceral and subcutaneous adipocytes defines a transcriptional factor cell type-specific module (*In preparation*)
3. Ishaq M, Kumar S, **Varinli H**, Han ZJ, Rider AE, Evans MDM, Murphy AB, Ostrikov K: Atmospheric gas plasma-induced ROS production activates TNF-ASK1 pathway for the induction of melanoma cancer cell apoptosis. *Mol Biol Cell* 2014;25:1523-1531
4. Brown C, Aksoy Y, **Varinli H,** Gillings M: Identification of the rainbowfish in Lake Eacham using DNA sequencing. *Aust J Zool* 2012;60:334-339

PRESENTATIONS ARISING FROM THE WORK IN THIS THESIS:

Oral presentations:

International Oral Presentations

Varinli et al. Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis. Presented at 48th Annual Australian Institute of Food Science and Technology Convention, Young Professionals in Science "*The Cost of High Sugar Diet to our Fat Cells*", Sydney, Australia. July, 2015. *Awarded "**Best Oral Presentation**".*

Varinli Presented at International Summer School for PhD students on "*Developing model systems to understand cardiomyogenesis and adipogenesis*", Mainz, Germany. August, 2013. *"**Invited/Sponsored PhD student**".*

Varinli Presented at International Summer School for PhD students on "*Effect of developmental stress on epigenome crosstalk during stem cell differentiation*", Mainz, Germany. July, 2013. *"**Invited/Sponsored PhD student**".*

National Oral Presentations

Varinli et al. Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis. Presented at Rotary Club of Ryde "*The Cost of High Sugar Diet to our Fat Cells*", Sydney, Australia. November, 2015. *"**Guest Speaker of the Month**".*

Varinli Presented and led an Advanced Biology Student Discussion Group at Macquarie University, Department of Biological Sciences on "*Clinical Epigenetics in Obesity Research and Advancements in Epigenetic Tools*", Sydney, Australia. October, 2015. *"**Guest Lecturer**".*

Varinli et al. Methylome of visceral and subcutaneous adipocytes defines a transcriptional factor cell type-specific module. Presented at Macquarie University Biofocus Conference

"*Analysis of the methylome of purified visceral and subcutaneous adipocytes, visceral adipose tissue and blood from three lean humans*", Sydney, Australia. September, 2014.

Varinli Presented and led an Advanced Biology Student Discussion Group at Macquarie University, Department of Biological Sciences on "*Epigenetics and Human Obesity*", Sydney, Australia. June, 2014. *"**Guest Lecturer**".*

Poster presentations:

International Poster Presentations

Varinli et al. Transcriptome and epigenome profiling of visceral adipose tissue and visceral adipocytes from 26 females with a range of BMI. Presented at Lorne Genome Conference, February, 2016.

Varinli et al. Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis. Presented at Lorne Genome Conference "*The epigenetic and transcriptional trajectories of human adipocytes differentiated in high glucose*", February, 2016.

Varinli et al. COBRA-Seq: Sensitive and Quantitative Methylome Profiling. Presented at EMBO: From Functional Genomics to Systems Biology "*Epigenomic methods for systems biology with large sample sets*", Heidelberg, Germany. November, 2014.

Varinli et al. Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis. Presented at Wellcome Trust Epigenomics of Human Disease, Cambridge, UK. October, 2014.

Varinli et al. Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis. Presented at EMBL: PhD Symposium- Inspired by Biology Exploring Nature's Toolbox, Heidelberg, Germany. October, 2014.

Varinli et al. COBRA-Seq: Sensitive and Quantitative Methylome Profiling. Presented at Epigenetics 2013, Australian Epigenetics Alliance "*Reduced Methylome Sequencing using GW-COBRA and LA-COBRA*", Shoal Bay, Australia. December, 2013.

National Poster Presentations

Varinli et al. LipiD-QuanT: a novel method to quantify lipid accumulation in live cells. Presented at Macquarie University Biofocus Conference "*Automated processing of label-free lipid dynamics in fat cell images*" September, 2014.

Varinli et al. LipiD-QuanT: a novel method to quantify lipid accumulation in live cells. Presented at CSIRO Process Science Conference "*Automated processing of label-free lipid dynamics in fat cell images*", Melbourne, Australia. May, 2013.

LIST OF ABBREVIATIONS

Abbreviations	Definitions
450K arrays	Illumina Infinium Humanmethylation450 Beadchip Array
5mC or 5	5-Methylcytosines
ac	Acetylation
ADIPOQ	Adiponectin
ADSCs	Adipose-Derived Stem Cells
AGPAT2	1-Acylglycerol-3-Phosphate Acyltransferase 2
AGRF	Australian Genome Research Facility
AMP	Adenosine Monophosphate
ATAC-seq	Transposase Accessible Chromatin With Sequencing
b	5' Double Biotin Groups
B-value	Regression Coeffcient
BAT	Brown Adipose Tissue
BMI	Body Mass Index
BMP	Bone Morphogenetic Protein
C/EBPα or CEBPA	CCAAT-Enhancer-Binding Proteins α
C/EBPβ or CEBPB	CCAAT-Enhancer-Binding Proteins β
C/EBPδ	CCAAT-Enhancer-Binding Proteins δ
CA	Component Analysis
Camp	Cyclic Adenosine Monophosphate
CARS	Coherent anti-Stokes Raman Scattering
CCL2	Cc-Chemokine Ligand 2
cDNA	Complementary DNA
CHARM	Comprehensive High-Throughput Arrays For Relative Methylation
CHeReL	Centre For Health Record Linkage
CHH	Cytosine Followed by Adenine, Thymine or Cytosine Dinucleotide
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation Followed By Whole Genome DNA Sequencing
COBRA	Combined Bisulphite Restriction Analysis
COBRA-seq	Combined Bisulphite Restriction Analysis for High-Throughput Genome Sequencing
CpA	Cytosine Followed By Adenine Dinucleotide
CpC	Cytosine Followed By Cytosine Dinucleotide
CpG or CG	Cytosine Followed By Guanosine Dinucleotide
CPM	Count Per Million
CpNpG or CNG	Cytosine Followed By Any Dinucleotide and Guanosine
CpT	Cytosine Followed By Thymine Dinucleotide
CT	Computerized Tomography
D0	Day 0
D10	Day 10

Abbreviations	Definitions
D14	Day 14
DAPI	4',6-Diamidino-2-Phenylindole
DE	Differentially Expressed
DHA	Docosahexaenoic Acid
DKK1	Dickkopf-1
DM	Differentially Methylated
DMP	Differentially Methylated Probes
DMR	Differentially Methylated Regions
DNMT1	DNA Methyltransferase 1
DNMT3A	DNA Methyltransferase 3A
DNMT3B	DNA Methyltransferase 3B
DPM	Disintegrations Per Minute
DREAM	Digital Restriction Enzyme Analysis Of Methylation
dRRBS	Dual-Enzyme Reduced Representation Bisulfite Sequencing
EPA	Eicosapentaenoic Acid
EPIC arrays	Methylationepic
EWAS	Epigenome-Wide Association Studies
FABP4	Fatty-Acid-Binding Protein
FC	Fold Change
FDR	False Discovery Rate
FFA	Fatty-Acid-Binding Protein
FFPE	Formalin-Fixed Paraffin Embedded
FGF	Fibroblast Growth Factor
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GC	Guanosine Followed By Cytosine Dinucleotide
GFP	Green Fluorescent Protein
GLUT4	Glucose Transporter Type 4
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
GW-COBRA	Genome-Wide Combined Bisulphite Restriction Analysis
H3K4me3	Histone-3-Lysine-4 Trimethylation
hADSCs	Human Adipose-Derived Stem Cells
hANT2	Human Adenine Nucleotide Translocase Isoforms
HbA1c	Hemoglobin A1c
HDL-c	High-Density Lipoprotein Cholesterol
HELP-Seq	HpaII-Tiny Fragment Enrichment By Ligation-Mediated PCR Sequencing
HG	High Glucose
HOMA-IR	Homeostatic model assessment and insulin resistance
HOX	Homeobox Genes
HPLC	High-Performance Liquid Chromatography
IGV	Integrated Genome Viewer
IL-1	Interleukin 1
IL-6	Interleukin 6

Abbreviations	Definitions
IRS	Insulin Receptor Substrate Proteins
JNK	c-Jun N-Terminal Kinases
KS.test	Kolmogorov-Smirnov Tests
LA-COBRA	Linear Amplification Combined Bisulphite Restriction Analysis
LADS	Linear Amplification For Deep Sequencing
LD	Lipid Droplet
LDL	Low-Density Lipoprotein
LEP	Leptin
LG	Low Glucose
lincRNA	Long Intergenic Non-Coding RNA
LINE-1	Long Interspersed Nuclear Elements-1
LipiD-QuanT	Lipid Droplet Quantification Technique
LMR	Low-Methylated Regions
lncRNA	Long Non-Coding RNA
LPL	Lipoprotein Lipase
LXR	Liver X Receptor
M	Methylated
MBDCap-Seq	Methyl-Cpg Binding Domain-Based Capture
me	Methylation
MeDIP-Seq	DNA Immunoprecipitation Sequencing
Methyl-Seq	Methyl-Sensitive Enzyme With Next-Generation (Next-Gen) DNA Sequencing Technology
MIRA-Seq	Methylated-Cpg Island Recovery Assay Sequencing
miRNA	Micro RNA
MRI	Magnetic Resonance Imaging
mRNA	Messenger RNA
ncRNA	Non-Protein Coding RNA
NF-κB	Nuclear Factor- κB
Nimblegen SeqCap	Nimblegen Seqcap Epi Target Enrichment
O.D.	Optical Density
p	5’ Phosphorylation
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PKA	Protein Kinase A
PLIN1	Perilipin
PPARγ or PPARG	Peroxisome Proliferator-Activated Receptor γ
PRDM16	PR Domain Containing 16
PTGS2	Prostaglandin-Endoperoxide Synthase 2
RADseq	Restriction-Site Associated DNA Sequencing
RNA-seq	RNA Sequencing
ROS	Reactive Oxygen Species
RRBS	Reduced Representation Bisulfite Sequencing

Abbreviations	Definitions
rRNAs	Ribosomal RNAs
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SAT	Subcutaneous Adipose Tissue
SCFA	Short Chain Fatty Acid
SGBS	Simpson-Behmel-Gobali Syndrome
siRNA	Short Interfering RNA
SNP	Single-Nucleotide Polymorphism
snRNAs	Small Nucleolar RNAs
SuBLiME	Streptavidin Bisulfite Ligand Methylation Enrichment
SVF	Stroma Vascular Fraction
T2D	Type 2 Diabetes
TBX	T-box Genes
TET	Ten-Eleven Translocation
TGFβ	Transforming Growth Factor β
TNC	Tenascin C
TNFα	Tumor Necrosis Factor α
TSS	Transcription Start Site
U	Unmethylated
VA	Visceral Adipocytes
VAT	Visceral Adipose Tissue
WAT	White Adipose Tissue
WGBS	Whole-Genome Bisulphite Sequencing
WHR	Waist To Hip Ratio
X-square	Chi-Square
ZnO	Zinc Oxide
Δ-me	Delta Methylation Values

LIST OF TABLES IN INTRODUCTION AND DISCUSSION

LIST OF FIGURES IN INTRODUCTION AND DISCUSSION

INTRODUCTION

In this Chapter, I present the function of adipose tissue in human body, the key differences between adipose tissue types, consequently their cellular compositions and functions. I then introduce key factors that regulate adipose tissue distribution, along with health risks and complications associated with adiposity. I also present some fundamental concepts of epigenetics and its potential role in mediating gene and environmental interactions.

1. Health impacts human adipose tissue

As the largest endocrine organ in the human body producing hormones, adipose tissue influences almost every cell, organ and function in our bodies. Most people view adipose tissue as a passive reservoir for energy storage therefore it is the only disliked endocrine organ (1). It constitutes less than 10% of the body weight in some adults and over 60% in others. It carries out significant functions such as glucose homeostasis, energy metabolism, regulation of food intake and immune functions (2). Adipose tissue has also a relatively high turn-over rate, 10% of its adipocytes are renewed annually (3). Generally it is to blame for suboptimal body types and shapes. Both excess and deficiency of adipose tissue have adverse health impacts.

Obesity is manifested by progressive increase in adiposity, progressive loss of adipose tissue is known as lipodystrophy. These conditions are characterized by gain or loss of fat in either all body parts or only in specific parts. It is interesting that obesity and lipodystrophy can occur simultaneously and their metabolic consequences are similar (4).

The World Health Organisation (WHO) estimates that 2.8 million people die each year as a result of being overweight or obese (5) and 17.5% of Australians are classified as obese (6). Obesity has reached alarming rates in many countries, including Australia. Accumulation of excess fat tissue is a risk factor for various metabolic conditions, particularly impaired glucose tolerance which may proceed to Type 2 Diabetes (T2D). Although not all obese individuals develop T2D, dysfunction and excess of adipose tissue plays a role in the

predisposition to T2D. Hence, it is clear that obtaining a better understanding of how adipose tissue contributes to adverse health effects is essential in improving prevention and treatment strategies for both obesity and T2D.

1.1.Obesity and associated complications

Obesity, a condition of excess adiposity, develops due to the imbalance between food intake and caloric utilization. According to the WHO, obesity is defined by a body mass index (BMI) of ≥ 30 kg/m^2 for both men and women (7). However, BMI is a measure of excess weight rather than excess adiposity. Therefore, two people with the same BMI may look different. On the other hand, fat distribution can be estimated by waist circumference alone or the ratio of waist to hip circumference (WHR). The waist circumference at which there is an increased metabolic complication risk is > 102 cm for men and >88 cm for women.

Body fat percentage is the most accurate measure of adiposity. Specialised imaging instruments such as Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI) can measure body fat distribution in specific depots, lean muscle mass and bone mass (8). However, these instruments are expensive to run, and therefore have not been used in large population studies.

The clinical factors associated with the progression of obesity are listed below (9):

- *Elevated fasting glucose:* higher than normal amount of circulating sugar in the blood before meals.
- *Glucose intolerance:* impaired ability to reduce blood glucose levels into the normal range ($\leq$5.5 mmol/L in the fasted state and under $\leq$7.8 mmol/L two hours after an oral glucose load) after eating; often known as prediabetes.
- *Insulin resistance:* a condition in which cells of the body have a lowered level of response to secreted insulin.
- *Hypertension:* high levels of blood pressure.

- *Increased serum triglyceride:* elevated levels of neutral fat in the blood. It is associated with heart disease risk.
- *Reduced high-density lipoprotein cholesterol (HDL-c):* HDL-c has protective effects such as decreased oxidation and low levels of HDL-c are associated with the progression of atherosclerosis.

1.2.Type 2 Diabetes and associated complications

Type 2 Diabetes (T2D) is a progressive disease in which individuals have high blood glucose levels due to impaired insulin sensitivity and secretion (10). It is diagnosed when blood glucose level is higher than the baseline of 7 mmol/L in the fasted state and 11.1 mmol/L two hours after an oral glucose load.

T2D was thought to only occur after age 40. However, the rising prevalence of obesity in younger generations contributes to the growth of T2D incidence in young people and in children (11). The progression of T2D is associated with multiple clinical factors listed below (12, 13):

- *Insulin resistance:* inability to reduce blood glucose levels into the normal range in the presence of insulin.
- *Hyperglycemia:* in pre-diabetes there is a higher than normal amount of circulating sugar in the blood, defined as a fasting blood glucose concentration > 5.5 mmol/L or > 7.8 mmol/L 2h after an oral glucose load.
- *Impaired beta-cell function:* at the pre-diabetic stage, an increased rate of insulin secretion and C-peptide production by pancreatic beta-cells to compensate insulin resistance. This causes hyperinsulinemia (≥25 mIU/L in the fasted state), which often leads to beta-cell damage. Therefore, rate of insulin secretion is lowered as T2D progresses.
- *Hypertension:* high levels of blood pressure.

- *Dyslipidemia:* abnormal concentrations of lipids and lipoproteins in the blood.

Both obesity and T2D have been associated with serious health complications such as cardiovascular diseases, stroke, cancer, neuropathy and eyesight and kidney failure (14). These complications are presented in Figure 1.

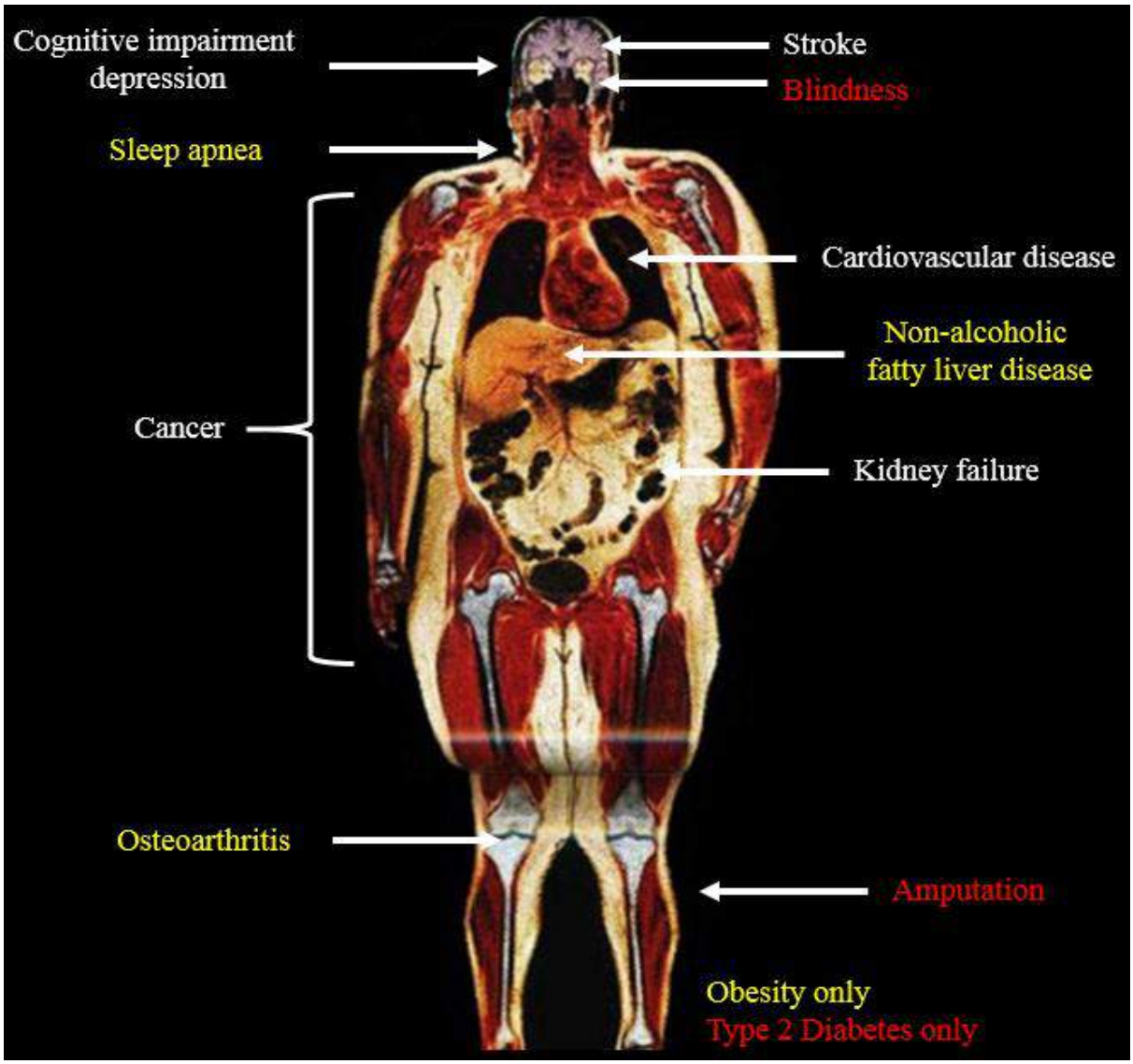


Figure 1: Health complications related to obesity and/or Type 2 Diabetes. The body scan image of woman (weight: 113 kg, height: 168 cm) is sourced from: (15). Obesity-only and Type 2 Diabetes-only associated health complications are shown in yellow and red respectively.

2. Cellular composition and classification of human adipose tissue

2.1.Cellular composition of human adipose tissue

Besides adipocytes, adipose tissue contains a variety of non-adipocyte cells including pre-adipocytes, fibroblasts, extracellular matrix, nerve cells, endothelial cells, blood cells, mesenchymal stem cells and immune cells collectively known as the Stroma Vascular Fraction (SVF) (Figure 2). The immune cell content of adipose tissue is highly variable, including pro-inflammatory immune cells (neutrophils, dendritic cells, M1 macrophages, Th1 cells, B cells, and mast cells) and anti-inflammatory immune cells (regulatory T cells, Th2 cells, M2 macrophages, and eosinophils) (16). The key functions of adipocytes and pre-adipocytes are discussed further below.

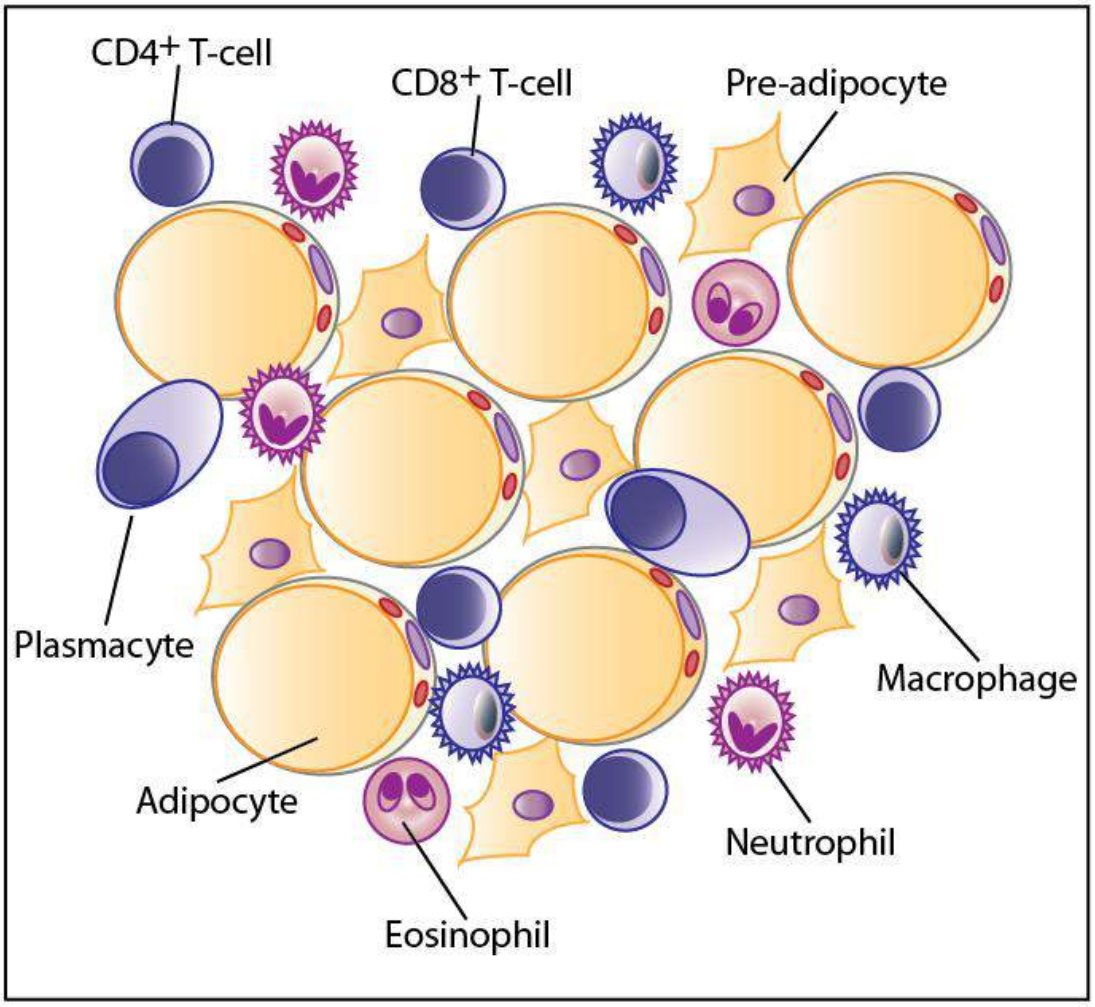


Figure 2: Schematic representation of human adipose tissue cellular composition.

2.1.1. Key functions of adipocytes

Triglyceride Storage and Release

Adipocytes have an important role in energy homeostasis by storing triglycerides and releasing fatty acids and glycerol as a source of energy as needed (2). About 90% of the volume of white adipocytes is taken up by a single fat droplet.

When the plasma glucose concentration is high, insulin regulates the transport of glucose into adipocytes by (17):

- Stimulating the translocation of the facilitative glucose transporter (GLUT4) through a process called exocytosis (18).
- Acting through its plasma membrane receptor leading to tyrosine phosphorylation of insulin receptor substrate proteins (IRS) and activation of well-documented PI3-kinase pathway that play an essential role in glucose uptake and GLUT4 translocation (19).

Once taken up by the cells, glucose is converted into acetyl-CoA in the mitochondria and then used as a substrate for fatty acid synthesis. A study showed that glucose is the main carbon precursor for *de novo* lipogenesis: isotopically (^{13}C)-labelled glucose contributed 70% of the carbons contained in triglycerides produced by human adipocyte cultures (20). *In vivo*, insulin also inhibits the breakdown of lipids in adipocytes (21). Hormones such as cortisol, atrial natriuretic peptide and growth hormone stimulate fat deposition (22).

When plasma glucose levels are low, lipolysis is activated by hormonal and biochemical signals that act on lipases and membrane proteins in adipocytes (23). For instance, glucagon and catecholamines (adrenaline and noradrenaline) act through G-coupled receptors to elevate cyclic AMP levels. cAMP stimulates protein kinase A (PKA) which leads the activation of hormone-sensitive lipase via phosphorylation (21). Triglycerides can be broken down by hormone-sensitive lipase into glycerol and free fatty acids and released to the circulation. These fatty acids then become substrate for oxidation in muscle and provide energy or

become substrate for re-esterification in liver and transform into lipoprotein (24). Dietary components such as calcium and caffeine have pro-lipolytic effects while alcohol has anti-lypolytic properties (23).

Adipokine Secretion

Adipocytes regulate food intake and appetite by secreting peptide hormones called cytokines. Adipocyte-secreted cytokines are known as adipokines, and maintain the crosstalk with other organs. The adipokines; resistin, retinol binding protein-4 (*RBP4*), tumor necrosis factor α (*TNFα*) and interleukin-6 (*IL-6*) raise blood glucose levels, while leptin, adiponectin, visfatin and omentin maintain homeostasis by reducing blood glucose levels through their action on organs of metabolic importance such as muscle, liver and adipose tissue (2, 25).

Adipokines affect lipid metabolism, fatty acid transportation, insulin responsiveness, angiogenesis, blood pressure and immune function, all of which can cause adverse health risks when perturbed (26). Although most adipokines have distinct steady state localization in adipocytes, they are released via vesicles through pathways that are not fully understood (27, 28).

Immune protection

Similar to macrophages, adipocytes express genes encoding inflammatory molecules and scavenger receptors that have important roles in both innate and adaptive immune responses (25). Adipocytes have intrinsic inflammatory properties. For instance, *TNFα* is an inflammatory cytokine overexpressed in the adipose tissue of obese individuals (29). Adipocytes secrete several other inflammation biomarkers such as *IL-6*, interleukin-1 (*IL-1*) and CC-chemokine ligand 2 (*CCL2*). The secretion of these inflammatory mediators is regulated by Nuclear Factor-κB (NF-Kb) signal transduction, c-Jun N-terminal kinases (JNK) signalling, and several other pro and anti-inflammatory signalling cascades (25).

2.1.2. Key functions of pre-adipocytes

Self-renewal

New adipocytes are generated through the proliferation and differentiation of adipocyte precursor cells known as pre-adipocytes. These are derived from adipose-derived mesenchymal stem cells (ADSC) through processes called *adipocyte lineage determination* and *commitment* (30).

ADSC have self-renewal and multipotency properties which make them an excellent source for regenerative medicine therapies (31, 32). ADSC can differentiate *in vitro* into adipocytes, but also into other cells such as osteoblasts, chondrocytes, neuronal-like cells and myocytes (33-35). This implies that *in vivo,* if ADSC can also be mobilized from adipose tissue to other organs, they may contribute to a pool for potential progenitor recruitment by other organs, perhaps following tissue injury.

The proliferation and differentiation capacity of ADSC greatly varies depending on their depots of origin. ADSC isolated from different depots have distinct patterns of gene expression, differentiation potential, and response to growth factors (36). Pre-adipocytes exhibit intrinsic depot-of-origin memory in their metabolic activities, from lipolysis to adipokine secretion (37). They also have antimicrobial and antiviral properties (38, 39) and they can express several inflammatory genes that play a role in immune responses (25).

Adipogenesis

Adipogenesis, the process of pre-adipocytes becoming adipocytes, is orchestrated by several transcription factors, regulating the expression of hundreds of genes that have diverse biological effects on adipocytes. Several biological pathways have direct effects on the adipogenic differentiation of ADSC (40, 41). The main regulators are listed below:

- Peroxisome proliferator-activated receptor γ (*PPARγ*), CCAAT-enhancer-binding proteins α (*C/EBPα*), *C/EBPβ*, and *C/EBPδ* are the main inducers of adipogenesis, with C/EBPβ and δ acting early in terminal differentiation.
- Insulin-like growth factor 1 (*IGF-1*) insulin signalling activates adipogenesis.
- The WNT and hedgehog (HH) pathways represses adipogenesis by inhibiting expression of pro-adipogenic transcription factors such as *PPARγ* and *C/EBPα*.
- Members of the transforming growth factor *β* (*TGFβ*), bone morphogenetic protein (*BMP*) superfamily, fibroblast growth factor (*FGF*) and Notch-signalling pathways have both pro- and anti-adipogenic functions.

Especially, *PPARγ* and *C/EBPα* are the master regulators of adipogenesis and terminal differentiation and they regulate the expression of genes involved in key adipocyte functions such as insulin sensitivity, lipogenesis and lipolysis (Figure 3) (42). Some key examples of these genes are:

- **Glucose transporter type 4(*GLUT4*):** is one of the 13 sugar transporters on the surface of human cells. It is the key regulator of whole-body glucose homeostasis by trafficking circulating blood sugar into adipocytes and muscle cells in the presence of insulin (43).
- **Fatty-acid-binding protein (*FABP4*):** is an intracellular protein affecting lipid fluxes, metabolism and signalling within cells. It circulates in the bloodstream, therefore having a critical hormonal role in whole-body metabolism (44).
- **Lipoprotein lipase (*LPL*):** plays a role in lipoprotein metabolism by its dual functions: breaking down triglycerides and acting as a coactivator for lipoprotein uptake. It is expressed in heart, muscle, and adipose tissue (45).
- **1-acylglycerol-3-phosphate acyltransferase 2 (*AGPAT2*):** is involved in the synthesis of phospholipid and triglyceride in the endoplasmic reticulum of adipocytes. Mutations in *AGPAT2* have been associated with lipodystrophy (46).

- **Leptin (*LEP*):** an adipokine, which plays a major role in the regulation of body weight. Leptin can trigger immune responses and has been linked to obesity and T2D (47).
- **Adiponectin (*ADIPOQ*):** is exclusively expressed in adipose tissue and encodes adiponectin, a molecule that regulates glucose levels and fatty acid breakdown and also has putative anti-inflammatory properties (48).
- **Perilipin (*PLIN1*):** perilipin is the phosphoprotein that coats lipid droplets in adipocytes (49).

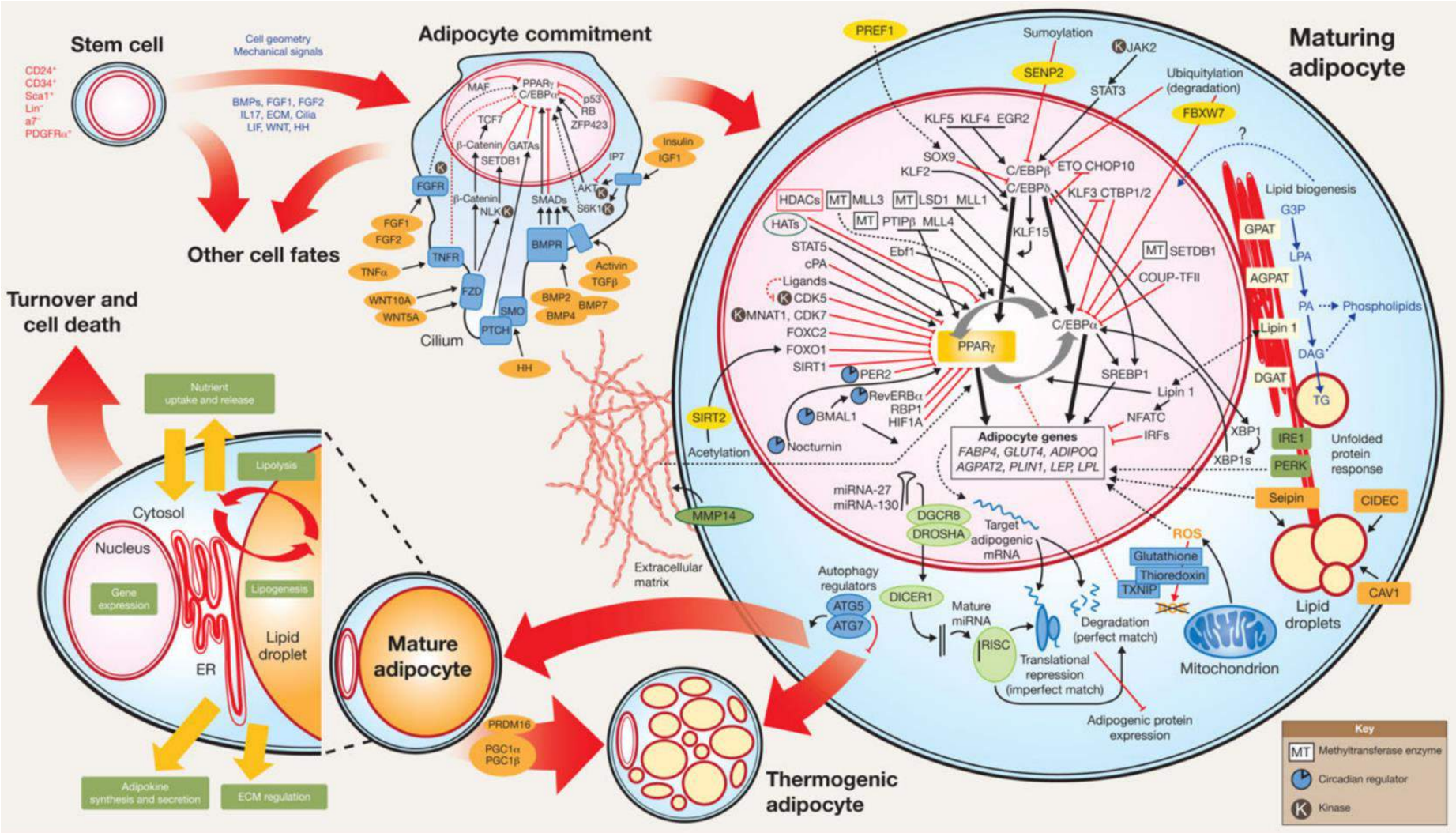


Figure 3: The basic network of transcription factors that regulate adipogenesis (30).

2.2.Classification of human adipose tissue

Adipose tissue is found all around the body. There are two main adipose depots in humans classified based on the coloration: Brown Adipose Tissue (BAT) and White Adipose Tissue (WAT) (Figure 4). They have antagonistic functions: BAT burns energy to produce heat and maintain body temperature while WAT is responsible for triglyceride storage as an energy reservoir. A new, third class of adipose tissue has been ascertained due to its colour: Beige Adipose Tissue. Its functions slightly differs from both classical BAT and WAT (50). The location of adipose tissue and the heterogeneous lineages of its adipocytes can dictate the degree of its association with adverse health, as will be further described below for each fat depot.

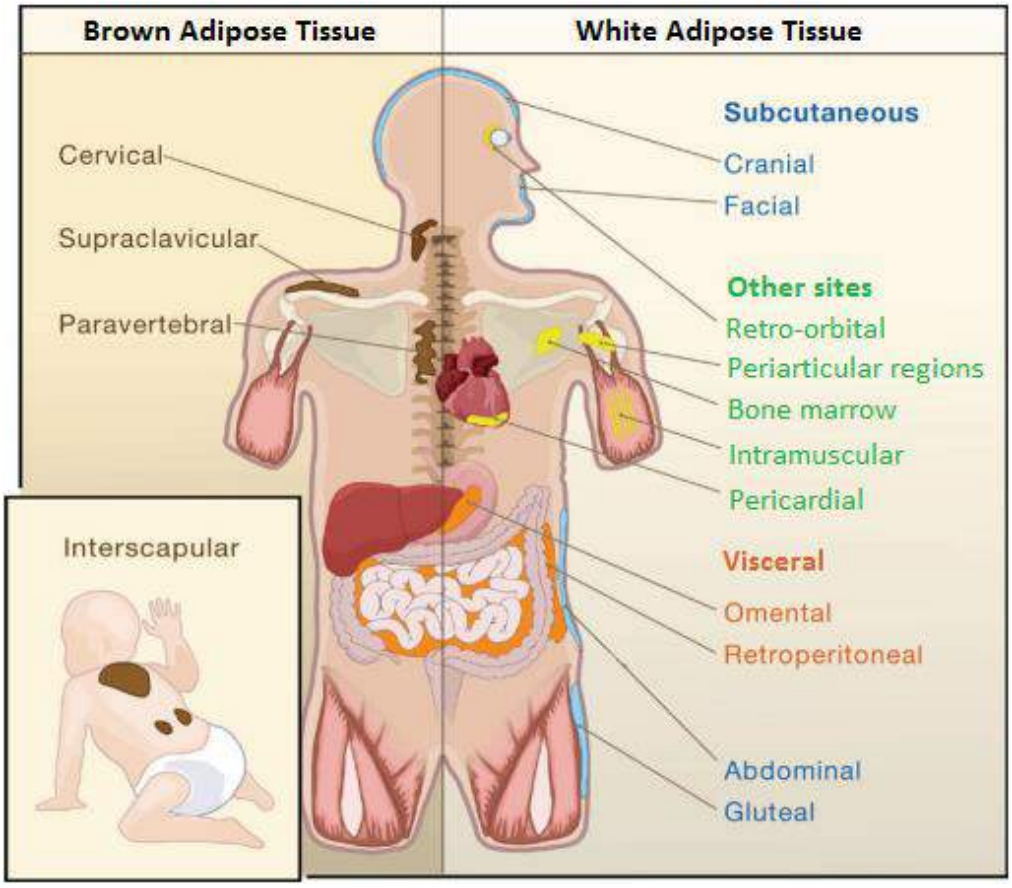


Figure 4: Adipose tissue depots in humans. White Adipose Tissue (WAT) is dispersed all over the body, with subcutaneous (blue) and visceral (orange) depots representing the main

compartments for fat storage. Other WAT sites are shown in green. Brown Adipose Tissue (BAT) (brown) is larger in infants than adults (adapted from (51)).

2.2.1. Brown adipose tissue (BAT)

BAT is characterized by rapid heat production due to its high mitochondrial content and expression of a unique protein: uncoupling protein-1 (*UCP1*). It is essential for non-shivering thermogenesis. In human foetuses and newborns, BAT is found largely around the blood vessels in the neck and in smaller quantity around kidneys (52). It has been hypothesized that this distribution maintains the body temperature by warming the blood in surrounding blood vessels before its delivery to the other organs, hence providing the optimal temperature for biochemical reactions (53). In rodents, the thermogenic potential of BAT has been estimated to be 300 times more than most other tissues (54). In fact in humans, 50 g of fully activated BAT can burn 20% of total daily energy expenditure (55). Therefore, BAT is an excellent resource to increase resting energy expenditure and maintain healthy weight.

Brown adipocytes, constituting the majority of BAT, are characterized by multilocular lipid droplets and an extraordinary number of mitochondria which gives the tissue its colour (Figure 5). Lineage tracing studies show that both brown adipocytes and skeletal muscles originate from *Myf5* (myogenic gene) expressing precursors and that browning is controlled by the expression of PR domain containing 16 *(PRDM16)* (56, 57), as shown in Figure 5. The key regulatory networks involved in browning are well explained in (58) and will not be discussed here.

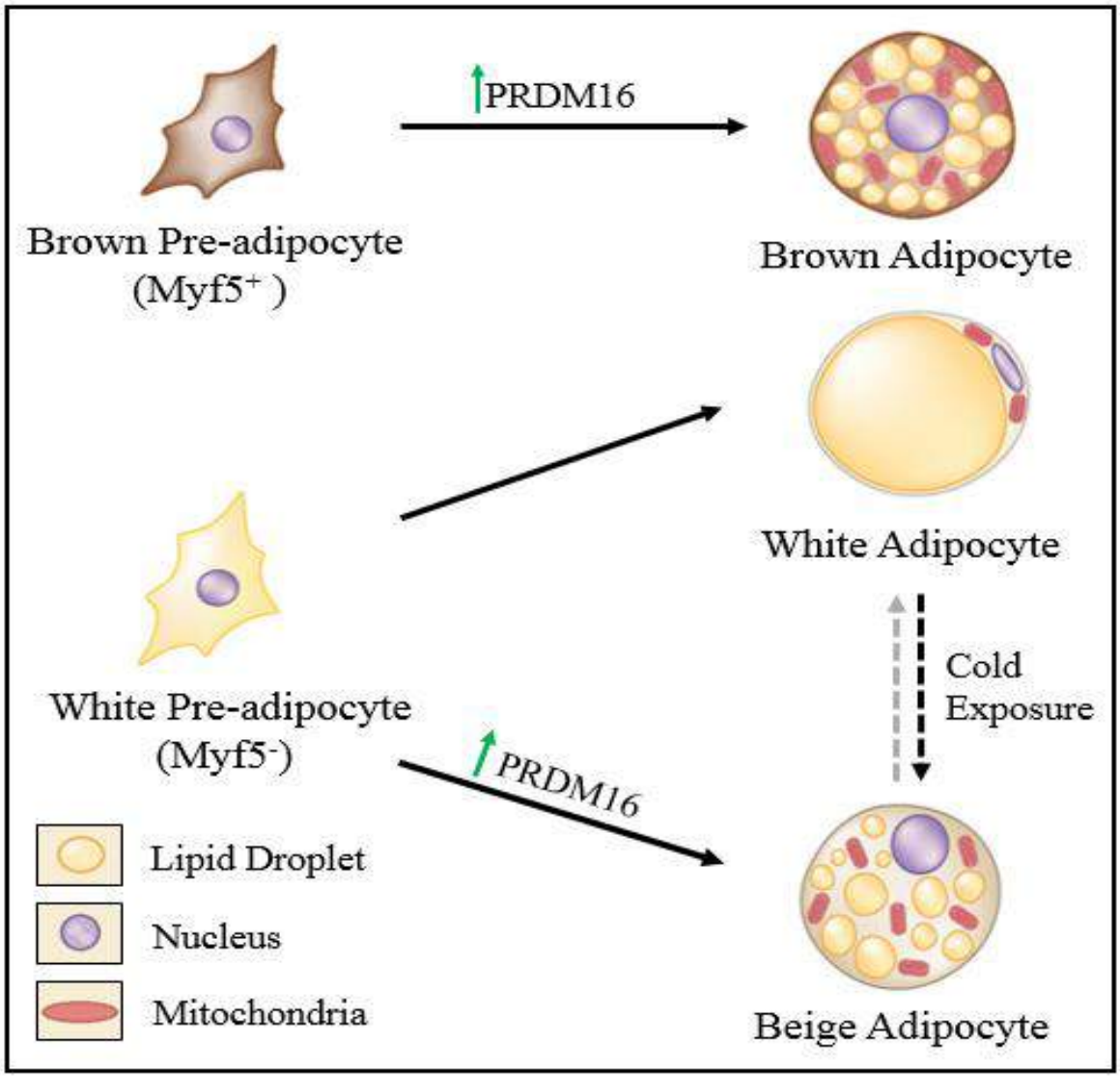


Figure 5: The origin and types of human adipocytes. Based on (59) where more detailed mechanisms of the development of human adipocytes can be found.

2.2.2. White adipose tissue (WAT)

WAT stores excess dietary fat in the form of triglycerides, to then release free fatty acids under energy demand. It is located in two main compartments of the body: Subcutaneous Adipose Tissue (SAT) forms a layer beneath the skin while Visceral Adipose Tissue (VAT) surrounds the vital organs (Figure 4). There are also marginal deposits of WAT in other sites such as around the face, in bone marrow and on the surface of the heart. WAT is non-homogenous in its cellular composition and metabolic characteristics. Unlike brown pre-adipocytes, white pre-adipocytes originate from adipogenic precursors lacking *Myf5*

expression. White adipocytes are commonly characterised by their unilocular lipid droplet morphology that maximizes their lipid storage capacity (Figure 5). They constitute the majority of the SAT and VAT and appear anatomically similar in both depots (60). Visceral white adipocytes are slightly smaller. Both SAT and VAT have been correlated with metabolic risk factors, but VAT is the more pathogenic fat depot (61). Additionally, VAT/SAT volume is highly sex-hormone dependent (62).

2.2.2.1. Subcutaneous adipose tissue (SAT)

SAT is found predominantly in the buttocks, thighs, and abdomen (Figure 4). Lower body SAT (in buttocks and thighs) is protective against metabolic risk factors (63). Tchoukalova *et al.* (64) showed that when both males and females gained weight, lower-body SAT increased adipocyte number (hyperplasia) while upper-body SAT increased adipocyte size (hypertrophy). Therefore, different SAT depots may function differently.

2.2.2.2. Visceral adipose tissue (VAT)

VAT is found in the abdominal cavity around the omentum, intestines, and perirenal areas (Figure 4). VAT is a more active secretory organ than SAT (37) and has a less capillary network density in lean individuals which is reversed in obesity (65). An excess of VAT is strongly associated with increased mortality. Moreover, the functional differences in VAT, especially within its immune cell component, is significantly associated with increased BMI (66). Post-menopausal women, aging men with lower testosterone levels and women in hyper-androgenic states all display a greater increase in VAT (reviewed in (62)), indicating that VAT is metabolically dynamic across the lifespan.

There are depot-specific characteristics differences between SAT and VAT which are summarised in Table 1.

Table 1: Functional differences between SAT and VAT.

	SAT	**VAT**	**Reference**
Adipocyte size	**Large**	Small	(60)
Lipolysis	Low	**High**	(67)
Insulin receptor affinity and ***IRS1*** **expression**	**High**	Low	(68)
Inflammation	Low	**High**	(51)
Cytokine Secretion	**Mostly Low** (Leptin ↑)	High	(37)
Turnover	**Fast**	Slow	(69)
Angiogenic capacity	**More**	Less	(65)
Adipose derived stem cell differentiation proliferation and differentiation capacity	**High**	Low	(70)
Developmental stage	**Early** (14 weeks)	Late	(71)
Developmental genes	*Tbx15*, *Shox2*, *En1* and *Sfrp2*	*Gpc4*, *Hoxc8*, *Hrmt1/2*, *Hoxa5*, *Thb2* and *Nr2f1*	(72)

2.2.3. Beige adipose tissue

The *Myf5*$^{-}$ precursor cells, giving rise to white adipocytes, have the potential to differentiate into brown-like adipocytes under the exposure to beta-adrenergic agonists. These cells are often called beige adipocytes and its depots are called beige adipose tissue. Beige adipocytes exhibit multilocular lipid droplets and high mitochondrial activity yet still maintain some characteristics of white adipocytes (73).

It is still unclear whether white adipocytes can transdifferentiate into beige adipocytes, especially under specific conditions such as exposure to cold. Cold exposure is a common method to activate non-shivering thermogenesis in BAT and it induces browning of WAT. There is a growing evidence for the trans-differentiation of white adipocytes into beige adipocytes in murine models (reviewed extensively in (60, 74)). In humans, the same phenomenon has been recently described with the possibility of having "inactive beige adipocytes" which have white adipocyte characteristics but can be re-activated under specific stimuli (40). Hence, the possibility of "trans-differentiation" or "re-activation" has been left open in the human adipose tissue context (Figure 5, dashed arrows).

Most recent studies have showed a contribution of the lymphoid cells in SAT to generate and activate adipocyte browning (75, 76). Although SAT is known to have more browning potential, adaptive thermogenesis of VAT has also been reported in mice and in humans (77). Beige adipose tissue is one of the most active research areas, as both BAT and browning of WAT provide great potential for weight loss treatments.

3. Use of *in vitro* systems to study human adipogenesis

Adipogenesis has been well-studied since the establishment of the murine pre-adipocyte cell line 3T3L-1 in 1975 (78). A human pre-adipocyte cell line, SGBS, was developed by the isolation of mesenchymal stem cells from the subcutaneous fat depot of a male infant with Simpson-Gobali-Behmel Syndrome in early 2000's (79). The SGBS cell line has a high proliferation and differentiation capacity, providing unique advantages in studying human

adipogenesis (Figure 6) (80). During adipogenesis, fibroblast-like precursor cells undergo significant morphological changes and become lipid storing globular adipocytes. Chapter 1 provides insight to the physiological, metabolic, transcriptional and epigenetic trajectories during human SGBS adipocyte differentiation, particularly under high glucose exposure.

Unlike the unilocular *in vivo* white cells, cultured white adipocytes contain multiple lipid droplets (Figure 6). Accurate measurement of lipid storage capacity of the white fat cells is important for exploring lipid dynamics *in vivo* and *in vitro*. Chapter 2 describes a novel label-free lipid droplet quantification technique (LipiD-QuanT) to monitor fat accumulation in live adipocytes.

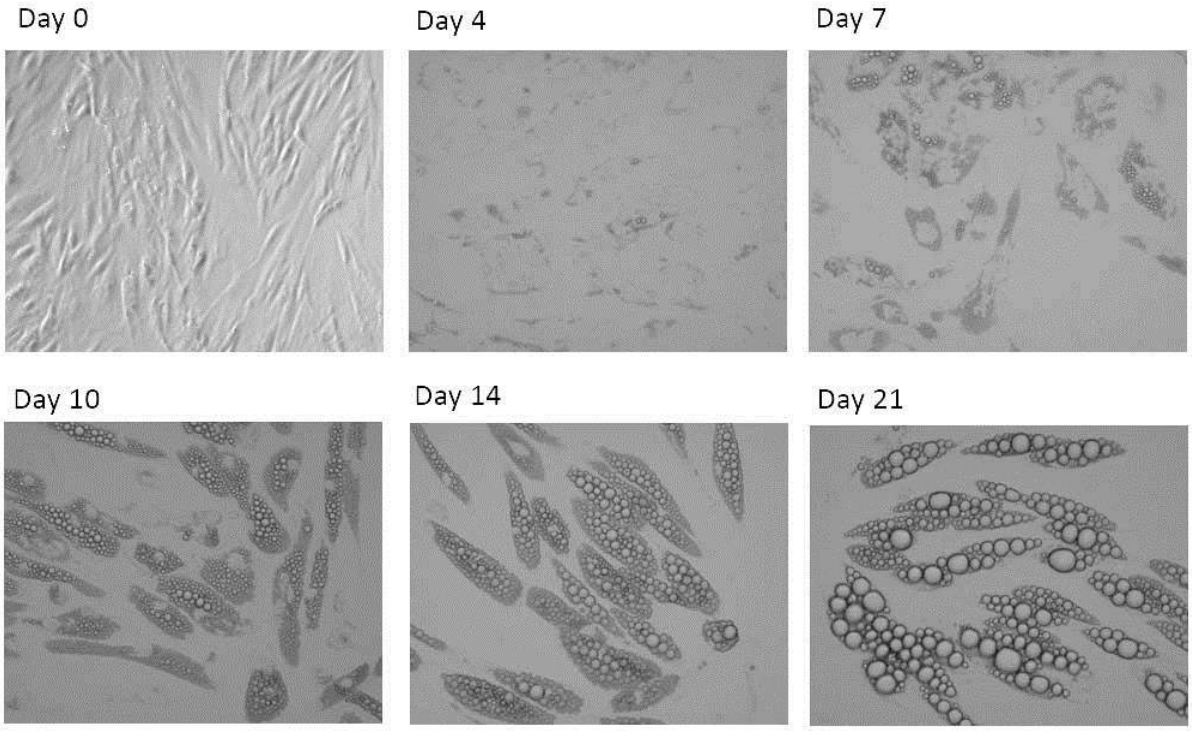


Figure 6: SGBS human adipogenesis in live cells. Lipids accumulate as adipocyte differentiation progresses.

4. *In vivo* adipogenesis and obesity

The increase in the mass of body fat can arise from an increase in the size of developed adipocytes (hypertrophy) or an increase in their total number (hyperplasia) by the generation of new adipocytes (81). Particularly in morbid obesity, when developed adipocytes reach their maximum capacity, new adipocytes are needed for the storage of excess energy intake in the form of fat. This raises two questions:

- **What is the origin of new adipocytes in hyperplastic obesity?**

To begin with, hyperplasia of the subcutaneous adipose tissue (SAT) is found in both sexes and across all BMI classes (82). Hence, there is inter-individual variability within the adipose tissue regardless of body size. It appears that the number of fat cells within a human body can increase throughout life (64). In murine models, when adipocytes reach a critical volume they recruit new adipocytes (83). On a high fat diet, both obesity-prone and –resistant mice recruited new adipocytes (84). This expansion of adipose tissue occurred in under 10 weeks.

Tang *et al.* (85) demonstrated that ADSC reside in the mural compartment of adipose tissue, which is involved in vascularisation. Using immunohistochemistry, Lin *et al.* (86) further narrowed the location of ADSC to the endothelium of capillaries and arterioles, excluding the endothelium of arteries surrounding the adipose tissue. Therefore, current understanding is that committed pre-adipocytes in the stromal vascular fraction of adipose tissue, and ADSC in the peri-vasculature are the main sources for new adipocytes. It is currently unknown whether ADSC give rise to other cell types *in vivo*. However, as discussed above, ADSC can differentiate into other cells such as osteoblasts, chondrocytes, neuronal-like cells and myocytes *in vitro*. However, are ADSCs the only pools? Until recently it was still unknown whether other sources of stem cells can infiltrate the adipose tissue and give rise to new adipocytes.

In a murine system, green fluorescent protein (GFP)-tagged bone marrow-derived cells were shown to contribute to adipose tissue with higher levels in visceral (gonadal) than subcutaneous fat (87).

Two recent studies in humans showed that bone marrow may also serve as a reservoir for adipocyte progenitors in the surplus of energy (88, 89). Genetic tracking of transplanted bone marrow stem cells showed that they could contribute up to 10% of the total adipocyte number in lean individuals and up to 25% in obese individuals. Hence, adipocytes within the same depot are not necessarily from the same developmental origin.

- **What is the consequence of having large adipocytes due to hypertrophic obesity?**

Hypertrophy without hyperplasia leads to metabolically unhealthy adipose tissue, with local inflammation, impaired lipolysis and insulin resistance in both obese (90) and lean individuals (91). Moreover, hypertrophic obese individuals have a reduced rate of adipocyte turnover, requiring their adipocytes to expand further under excess energy flux (92). Recently, in humans, ~100 Single-nucleotide polymorphisms (SNPs) were found to estimate the variance in BMI and WHR (93, 94) and expression levels of some of these genes (*CALCRL*, *CPEB4*, *FAM13A*, *HLA-DRA*, *MSC* and *PLCG2*) was correlated with adipocyte size in an independent cohort (95). Most recently, lower expression of Early B-Cell Factor 1 (*EBF1*) was closely linked to hypertrophy and consequently to altered lipid mobilisation and insulin resistance (96). Thus, adipocyte size is an important feature associated with pre-disposition to T2D and dysfunctional adipose tissue, independent of BMI (91).

5. Fat distribution in obesity

Fat distribution has a serious impact on metabolic risk. Particularly, the expansion in VAT and upper-body SAT compartments increases metabolic risk as they contribute to central obesity.

Central obesity is characterised by an apple-shape body and is defined by a WHR above 0.90. It has been associated with higher risk of Type 2 Diabetes (T2D) and metabolic diseases (Figure 7). On the other hand, obese individuals with low WHR have a pear shaped body and lower risk of metabolic complications (Figure 7). In fact pear shaped adiposity has been found to confer metabolic protection with improved metabolic profiles including lower fasting plasma glucose and triglycerides, and higher HDL-c (97, 98).

One of the most popular explanations for the strong association between abdominal adiposity and obesity has been that VAT is highly vascularised and drains into the hepatic portal vein. Hence this adipose tissue releases free fatty acids (FFAs) and inflammatory cytokines directly to the portal circulation, which provides blood flow to the liver (99).

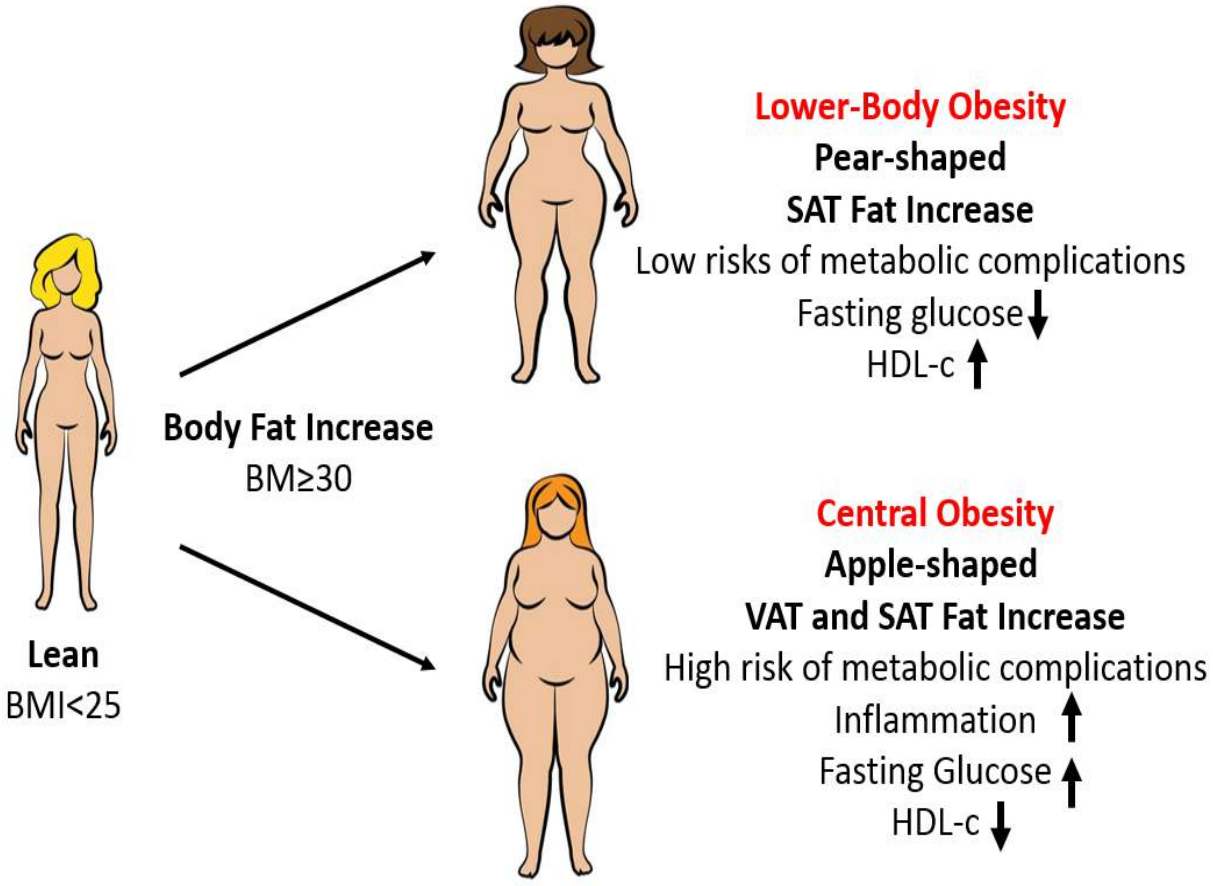


Figure 7: Human fat distribution and associated metabolic risk levels (After (51)).

6. Epidemiology and impact of obesity and Type 2 Diabetes

As of 2014, more than 1.9 billion adults aged 18+ years were overweight or obese, with over 600 million classified as obese (100). Currently, over 60% of Australian adults are overweight or obese (35% overweight, 28% obese) (101). This places Australia's rate of obesity as the fourth highest amongst the 34 countries members of Organisation for Economic and Co-operation and Development (OECD), behind the USA, Mexico and Hungary (102).

The global prevalence of diabetes is estimated to be 9% (103). The prevalence of Type 2 Diabetes (T2D) in Australia is about 4% according to 2011-2012 statistics (101). This places Australia's rate of T2D higher than western European countries but lower than especially Middle Eastern and Asian countries (104). Within the Australian context, the longitudinal "AusDiab" study has estimated that obesity increases incidences of T2D five-fold (105). Both diabetes and obesity are among the six Australian National Health Priority research areas listed in 1997 and 2008 (106, 107).

7. Drivers of obesity and Type 2 Diabetes

There are three main drivers of obesity and T2D: lifestyle, environment, and genetics. However, the severity of these conditions is highly contingent on the complex interplay of these drivers, and is unique to every individual.

Diet and physical activity greatly influence obesity and T2D (108). Frequent consumption of energy dense pre-processed foods and reduced levels of physical activity causes energy imbalance, resulting in increased storage of fat and increased risk of obesity and T2D.

The heredity of obesity and T2D has been estimated to be 40–70% (109, 110) and 40-50% (111) respectively. On the other hand, most recent studies showed that ~100 known obesity associated genomic loci explain only 5-11% of the inherited predisposition to obesity (93). Similarly in T2D, heredity of the known ~20 gene variants explain less than 10% of the predisposition (112, 113). Obesity co-occurs with T2D, and 18% of the T2D associated genetic traits are also correlated with increase in BMI, perhaps due to shared genetic

influences (114). Figure 8 represents a list of genomic loci associated with T2D, obesity and related metabolic measures.

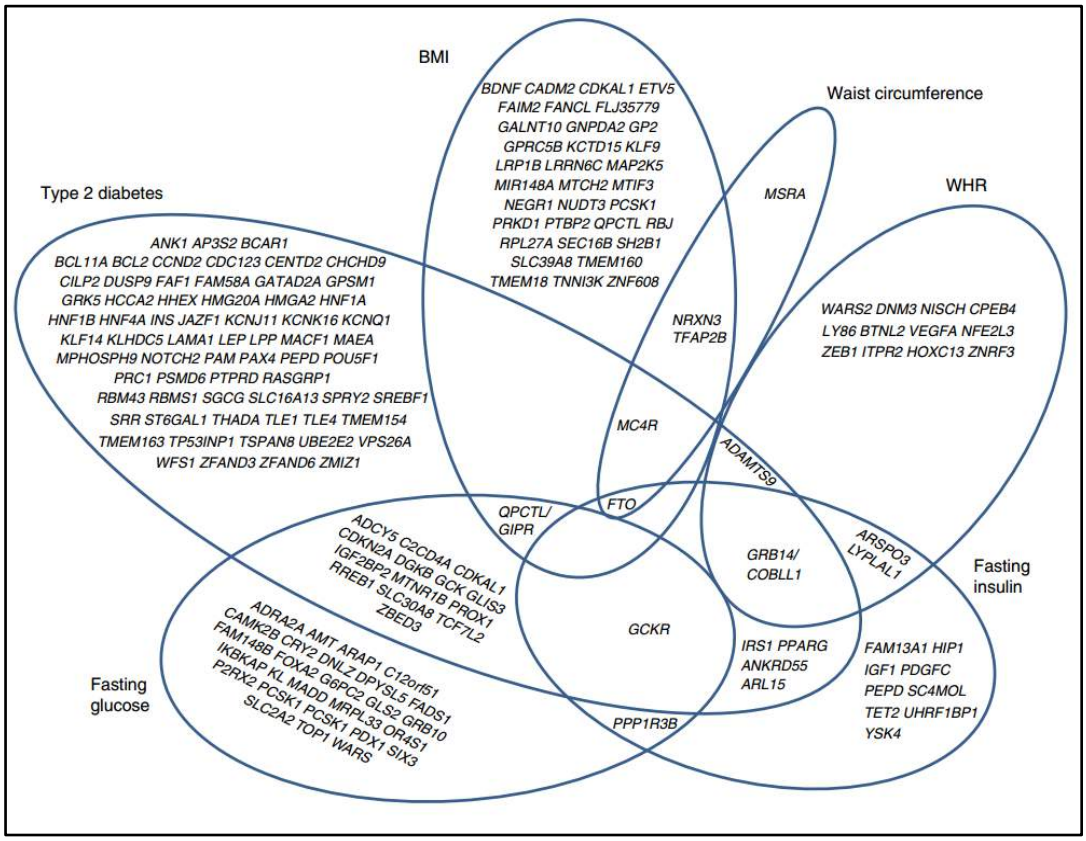


Figure 8: Venn diagram of genomic loci associated with T2D and obesity (110).

There is a great deal of effort to find the "missing link" connecting how lifestyle choices and environmental factors may be influencing the hereditary component of obesity and T2D. Epigenetic mechanisms regulate gene expression, and establish cell type specific gene expression profiles and epigenetic marks can be maintained through cell division. Hence, they may be involved in the regulation of depot-specific and obesity-associated gene expression profiles in the adipose tissue. Recent studies have shown that adipose tissue epigenetic signatures are altered with caloric restriction (115), exercise intervention (116) and following bariatric surgery (117-119). Hence its impact on the progression of obesity and T2D should be further investigated.

8. Key epigenetic mechanisms

Epigenetics is the study of mitotically heritable changes in gene expression that are not due to changes on the DNA sequence (120). Epigenetic modifications are potentially responsible for long lasting changes in gene expression in response to environmental factors.

There are four main mechanisms by which epigenetic signatures can be controlled (Figure 9):

- Chromatin remodelling via post-translational modifications of the histone proteins
- DNA methylation of the cytosine residues (CpGs in the mammalian context).
- Transcription factors (sequence-specific DNA binding proteins) by strategically localizing proteins that modify histones or DNA methylation at target regions via protein-protein interactions
- Functional elements, such as non-protein coding RNA (ncRNA) that may also be involved in targeting epigenetic modifications

Histone modifications and DNA methylation are relevant to the scope of this research, hence will be explained further below.

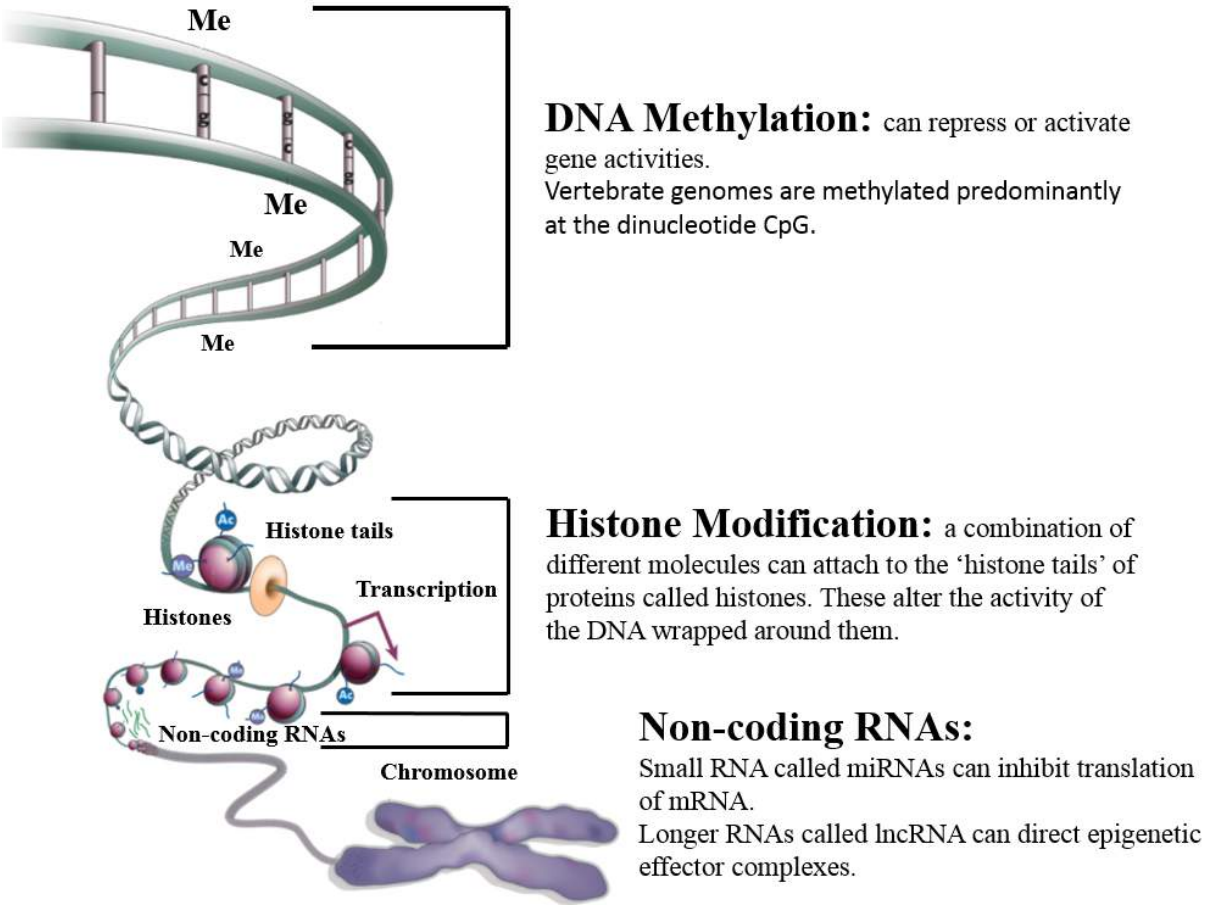


Figure 9: The main epigenetic mechanisms that control gene expression (Adapted from (121)).

8.1. Histone modification

Eukaryotic DNA is packed as chromatin, folded into repeating units of nucleosomes that are composed of roughly 147 bp of DNA wrapped around histone octamers. This structure allows the packaging of nearly 1.8 m of DNA into 1 micron of human nucleus. This compact packaging controls gene expression by the precise positioning of the nucleosomes. The tightly packed regions of the genome will be physically inaccessible while loosely packed regions will be accessible to transcriptional machinery (122). Also, covalent modifications of the histone tails and histone variants regulate the transcriptional machinery.

Some well characterised histone modifications, particularly acetylation (ac) and methylation (me) of lysine residues (K) on the tails of histone H3, and their known functions based on their location are presented in Figure 10.

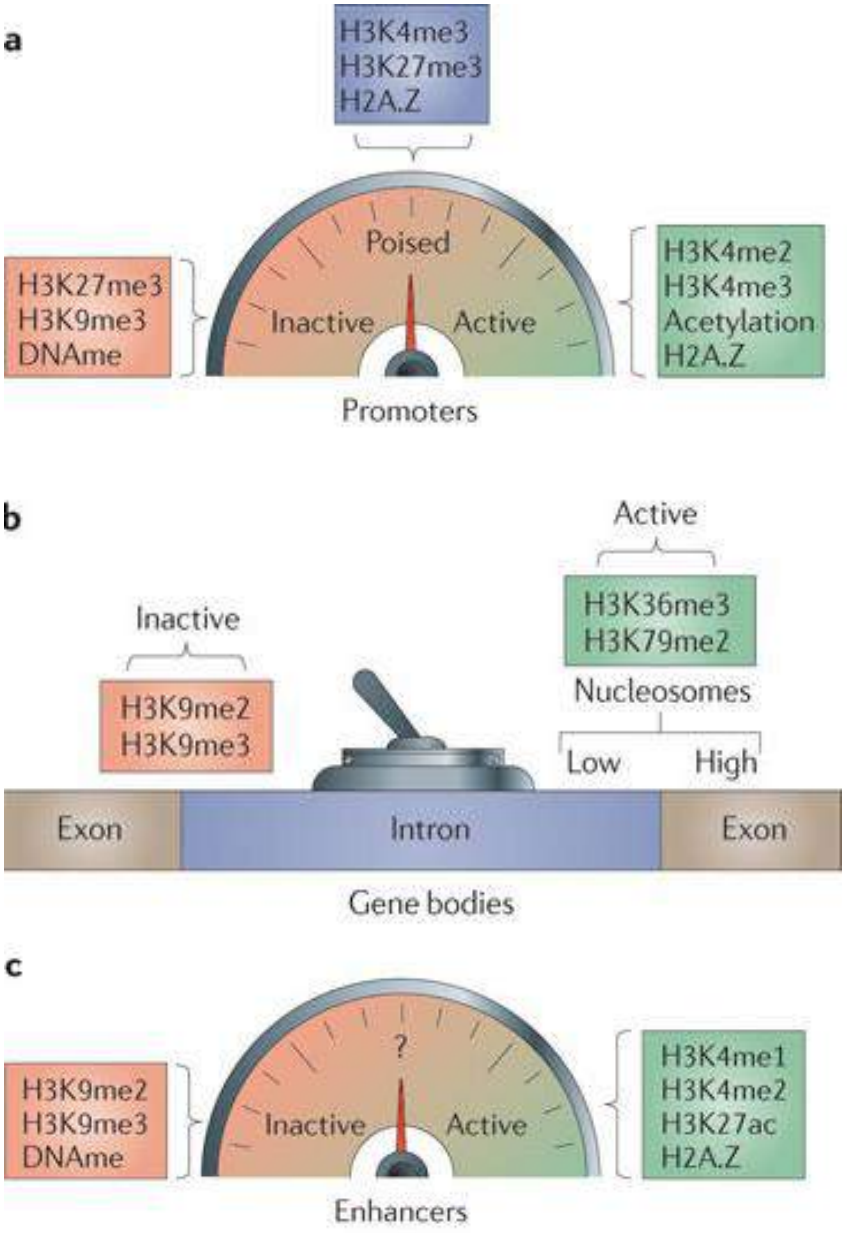


Figure 10: Known histone markers with known functions based on their location. A. At promoters histone modification can regulate the gene function from active to poised to an inactive state. **B.** At the gene body, histone modifications act as a switch from active to inactive states or vice-versa. **C**. At distal sites they correlate with enhancer activity (adapted from (123)).

8.2. DNA methylation

Methylation of cytosine by a covalent modification of DNA is the best characterised epigenetic modification. It occurs on CpG dinucleotides and is mediated by DNA methyltransferase enzymes in mammalian cells (124). While DNA methyltransferase 1 (DNMT1) is responsible for the maintenance of existing DNA methylation patterns, DNA methyltransferase 3A (DNMT3A) and DNA methyltransferase 3B (DNMT3B) are responsible for *de novo* DNA methylation, and they can repair errors made by DNMT1 (Figure 11) (125). During DNA replication, lack of DMNT1 and DNMT3A activity will lead to passive DNA demethylation. Recently, the Ten-eleven translocation (TET) family proteins have been found to be responsible from oxidation of 5-methylcytosine to 5-hydroxymethylcytosine and higher oxidative derivatives such as 5-formylcytosine and 5-carboxylcytosine (Figure 11) (126).

Therefore, this provided a mechanistic route for active demethylation with hydroxymethylation as an intermediate state between methylated and unmethylated cytosine. In the mammalian genome context, non-CpG methylation (127) and 5-hydroxymethylcytosines (128) have been observed in stem cells and may play a role in maintenance of the stem cell phenotype.

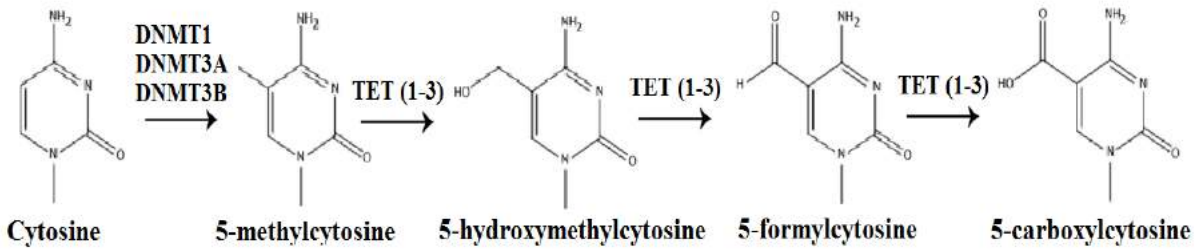


Figure 11: Overview of enzymatic cytosine modifications observed in mammalian genome context (126).

The human genome contains over 28 million CpG sites most of which are not evenly distributed across the genome (129). As a consequence of CpG to TpG mutation across mammalian evolution, CpGs are under-represented in the mammalian genome. The remaining CpGs are enriched in GC rich regions of the genome known as CpG islands (130), mostly found in gene promoters. Methylated CpG islands located in the promoter region are associated with repression of gene expression while unmethylated CpG islands in promoters are typically associated with activation of the gene expression (129). Region of lower CpG density flanking CpG islands are known as CpG shores (131). Differential methylation is frequently observed in CpG shores and its regulatory role appears to be region specific (121, 123).

With the development of whole genome methylation analysis, methylation in GC poor regions covering gene bodies, distant enhancers and GC poor promoters has recently become of more interest (132). Enhancer regions are several kilobases upstream or downstream of a gene promoter, yet they still influence tissue-specific gene expression (131). Methylation of CpGs within the gene body is associated with activation of gene expression and is often inversely related to promoter methylation (129).

9. Epigenetics and human obesity

Epigenetic dysregulation has long been identified as contributing to human disease phenotypes, particularly most types of cancer (133). Due to the growing obesity epidemic in the last decade, a great deal of attention has been directed towards understanding the role of epigenetics in the development of obesity (134).

Animal models provide unique opportunities for the assessment of epigenetic changes under specific dietary conditions. For instance, high fat diet induced changes in methylation pattern of genes that play important roles in body-weight regulation: the leptin gene (LEP) in rats (135) and melanocortin-4 receptor (MC4R) gene in B6 mice (136). Candidate gene approaches in humans have identified a large number of DNA methylation changes associated with obesity key parameters (109). Increased adiposity has been associated with

changes of the methylation pattern of metabolically important genes such as *HIF3A, LIPE, CETP, TIMP4* and *DOCK9* in the adipose tissue (137, 138).

Additionally, epigenetic modifications mediate genomic imprinting which determines expression of either maternal or paternal allele of a gene. Several of the genomic imprinted genes are associated with metabolic functions involved in the development of obesity and T2D (139). The role of epigenetics in obesity is discussed in more detail a systematic review on "Epigenetics and Human Obesity" which I co-authored (Appendix 1),

Chapter 4 provides insights into changes in the transcriptional signatures of human visceral adipose tissue (VAT) and purified visceral adipocytes (VA) across 26 females with a range of BMI (19-50 kg/m^2). This work will be coupled with currently on-going study of epigenome profiling of visceral adipocytes (VA) across the same set of individuals for future publication.

10. DNA methylome analysis methods

The invention of bisulphite sequencing was pivotal for the field of epigenetics as it allowed scientists to pinpoint the positions of 5' methyl cytosines. Bisulphite treatment replaces cytosines with uracil while 5' methylcytosines are preserved (140, 141). Thus, sequencing reveals only the methylated cytosine residues. The possibility of massively parallel sequencing together with improvements in the computational techniques allowed researchers to produce single resolution epigenome maps (142). The gold standard technique for quantifying methylation is whole-genome bisulphite shotgun sequencing (WGBS) (127).

At the time of starting my PhD, the cost of WGBS with 30X mean coverage was above 30,000 $AU. Although WGBS gives nucleotide-base resolution across a genome, it is still prohibitively expensive for most laboratories, particularly when many replicates are required. There were only a few affordable alternatives, which are further explained and quantitatively compared in Table 2.

Methods such as methylated DNA immunoprecipitation sequencing (MeDIP-Seq), methyl-CpG binding domain-based capture (MBDCap-Seq) are based on affinity enrichment of the

methylated regions using antibodies or methylated DNA binding proteins respectively, aiming to reduce the sequencing cost (143, 144). However, their enrichment is biased towards CpG rich regions of the genome. Another common approach was genomic complexity reduction with the use of either methylation sensitive or insensitive restriction enzymes. The most common method was reduced representation bisulphite sequencing (RRBS). It involved fragmentation of DNA using *MspI* endonuclease, then the fragmented DNA was separated according to size using electrophoresis and only appropriately sized DNA fragments sequenced after bisulphite treatment. There were also a few similar approaches such as Methyl-Seq (145) and CHARM (131, 146).

New epigenetic methods continue to be developed, particularly to improve genomic coverage in CpG-poor regions and also to reduce cost (DREAM (147) and SuBLiME (148)). More recently, methods that rely on capture of sequence-specific regions by hybridization have been developed (Agilent SureSelect (149) and Nimblegen SeqCap (150)).

Throughout my PhD, I have directly used four methylome methods. Three were already established methods: WGBS, Illumina Infinium HumanMethylation450 BeadChip array (450K array) (151, 152) and Nimblegen SeqCap (150) (Table 2, these are shown in bold). Additionally, I have developed a novel methylome method COBRAseq (Chapter 3).

Chapter 3 introduces a genome wide methylome method inspired by the original COBRA method that quantifies DNA methylation in a specific gene using restriction digest of the PCR amplicons generated from bisulfite treated template (153). Our method reduces genomic complexity by enriching methylated fragments containing potential CpG sites. It is compatible with high-throughput genome sequencing platforms and it has two versions called Genome-Wide COBRA (GW-COBRA) and Linear Amplification COBRA (LA-COBRA). They provide unique advantages and perform well compared to the other methods. This is further discussed in Chapter 3. Although I was developing COBRAseq to investigate methylation patterns in visceral adipocytes collected from the cohort described in Chapter 4 at

a reduced, additional methods were developed throughout my PhD. Particularly the development of customized Nimblegen SeqCap array allowed researchers to target specific regions of interest cell type and disease type specific manner (154).

As part of a larger group, we generated the first WGBS profiles of subcutaneous and visceral adipocytes in 3 lean and 3 obese individuals (Appendix 2). I was also involved in the design of customized Nimblegen SeqCap arrays specific to obesity-associated and adipose specific genomic locations based on our initial WGBS results on lean vs obese comparisons and publically available obesity genome-wide association studies (GWAS) and epigenome-wide association studies (EWAS). This method is currently being applied to the same set of patients for whom we investigated transcriptional profiles in Chapter 4.

The 450K array released just before the start of my PhD, has become the most popular and common method in human studies over the last five years. This method provides methylation quantification across ~480,000 (1.7%) CpG sites. The 450K array is known for robustness and has been a popular choice due to its affordable price and high precision. The majority of the publically-available methylation data on human cells has been generated using 450K arrays. Chapter 1 identifies the dynamics of changes in DNA methylation across human adipocyte differentiation *in vitro* and compares methylation profiles of *in vitro* grown human pre-adipocyte cell line to those of primary human pre-adipocytes and fibroblasts. Because of cost and the ability to compare directly with other datasets, I used 450K arrays for this work.

Table 2: Qualitative comparison of common methylome methods, chronologically arranged.

	Methylome Methods	Complexity Reduction Type	M or M+U Fraction	Methylome Sampling (Yes/No)	Favour of Enrichment Towards	Comments
Prior to the start of my PhD	**WGBS (127)**	**N/A**	**M + U**	**N/A**	**N/A**	**High cost, can detect non-CpG methylation, genomic input (0.05–0.1 μg).**
	MeDIP (143), MeDIP-Seq (144)	Affinity capture	M	No	CpG-rich	Bias towards 5mC-rich regions, Captures single-stranded DNA, prone to technical variability, coverage is read-depth dependent, input = 0.15–5 μg.
	MBDCap-Seq (155, 156), MIRA-Seq (157)	Affinity capture	M	No	CpG-rich	Dependent on CpG density, effected by salt concentration, covers about 18% of the CpGs (158), 28,500 CpG islands (156), DNA input = 0.2–1 μg.
	Methyl-Seq (145), HELP-Seq (159)	Restriction digest	M + U	Yes	CpG-rich	Assesses 0.25 to 1.3 million CCGG sites in human genome by difference in read fractions in *HpaII vs. MspI* libraries, input = 0.01–0.1 μg.
	CHARM (131, 146)	Restriction digest	M + U	Yes	CpG-rich and poor	Array-based and available for human, mice and rat, assesses 3.5 to 7 million CpG sites, input = 5 μg.
	RRBS (160, 161)	Restriction digest	M + U	Yes	CpG-rich/medium	Can detect non-CpG methylation, input = 0.1–0.3 μg.
	450K array (151, 152)	**Microarray**	**M + U**	**Yes**	**CpG-rich**	**Arrays comes in 12 per slide, available only for humans, not readily customized, input = 0.5–1 μg.**

Table 2. cont'd Qualitative comparison of common methylome methods, chronologically arranged.

	Methylome Methods	Complexity Reduction Type	M or M+U Fraction	Methylome Sampling (Yes/No)	Favour of Enrichment Towards	Comments
During my PhD	Agilent SureSelect (149)	Hybridization capture	M + U	Yes	CpG-rich and poor	Available for human genome only, covers 3.7 million CpG sites, input = 0.5 μg.
	DREAM (147)	Restriction digest	M + U	Yes	CpG-rich	Assesses methylation at ~0.15 million sites in human genome by sequential *SmaI/XmaI* digestion and library sequencing, input = 5 μg.
	SuBLiME (148)	Methylated cytosine capture	M	Yes	CpG-rich and poor	Substitutes biotin-14-dCTP or biotin-14-dGTP at the position of the 5mC in bisulfite treated DNA, input = 2 μg.
Now	**Nimblegen SeqCap (150)**	**Hybridization capture**	**M + U**	**Yes**	**CpG-rich and poor**	**"Off-the-shelf" version for human genome only/similar regions covered as 450K array, can be customized (150), can detect non-CpG methylation, input = 0.5–1 μg.**

CHAPTER 1: Hyperglycemia-induced epigenetic and transcriptional programming during human adipogenesis

Diet-induced hyperglycemia is a hallmark of Type 2 Diabetes (T2D) and obesity. However, it is not known whether hyperglycemia in itself contributes to the altered fat cell morphology, metabolism and gene expression that are mediated by epigenetic mechanisms in obesity. This project focused on the physiological, metabolic, transcriptional and epigenetic trajectories during human adipocyte differentiation *in vitro* under low and high glucose exposures.

One significant finding was that hyperglycemic exposure caused an increase in glucose uptake and in the rate of triglyceride accumulation during human adipogenesis. When grown in high glucose, adipocytes had significant changes in gene expression particularly in mitochondria and oxidative stress pathways and these changes were associated with perturbations in epigenetic signatures.

This chapter is the product of a working collaboration between myself and members of the CSIRO Food and Nutrition Flagship and Genomics and the Epigenetics Division in the Garvan Institute of Medical Research. I was predominantly involved in the design of the study, optimization and conducting of the experiments, interpretation of the data and writing the manuscript. The other co-authors: Molloy, Clark, Gillings, Swarbrick and van Dijk contributed to the design of the study and to the interpretation of the data. They also edited the manuscript. Peters analyzed the DNA methylation and transcriptome array data. Vallotton designed the lipid quantification software.

Detailed contributions:

	Varinli H	**Co-authors**
Experimental Design	90%	10%
Laboratory Experiments	100%	
Data Analysis	30%	70%
Data Interpretation	80%	20%
Writing	80%	20%

Hyperglycemia-induced epigenetic and transcriptional programming during *in vitro* human adipogenesis

Hilal Varinli[1,2,3], Timothy J. Peters[2], Pascal Vallotton[4], Susan J. van Dijk[1], Michael M. Swarbrick[5], Michael R. Gillings[3], Susan J. Clark[2] and Peter L. Molloy[1*]

[1] CSIRO Food and Nutrition Flagship, PO BOX: 52, North Ryde, New South Wales 1670, Australia

[2] Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales 2010, Australia

[3] Department of Biological Sciences, Macquarie University, North Ryde, New South Wales 2109, Australia

[4] Institute of Biochemistry, Department of Biology, ETH Zürich, Zürich 8093, Switzerland, formerly in CSIRO, Digital Productivity Flagship, PO BOX: 52, North Ryde, New South Wales 1670, Australia

[5] Centre for Diabetes, Obesity and Endocrinology, The Westmead Institute for Medical Research, The University of Sydney, New South Wales 2010, Australia

This manuscript is prepared to be submitted to *Diabetes*.

* To whom correspondence should be addressed:

Dr. Peter L. Molloy

CSIRO Food and Nutrition Flagship, Genomics and Cellular Sciences

Riverside Corporate Park, 11 Julius Avenue,

North Ryde NSW 2113, Australia

Phone: +61 2 9490 5168; Email: peter.molloy@csiro.au

Key words: obesity, diabetes, nutrition, human adipogenesis, DNA methylation

Abstract:

Hyperglycemia is a hallmark of Type 2 Diabetes and obesity. However, its effect on adipocyte morphology, metabolism, gene expression and epigenetic processes is not fully understood. We compared the effect of exposure to high (30 mM) and low (7.5 mM) glucose concentrations on *in vitro* adipogenesis of the SGBS cell line (derived from a patient with Simpson-Gobali-Behmel Syndrome). Across adipogenesis, there are extensive changes in lipid production, glucose uptake and transcriptional changes. We found 94 genes differentially expressed during SGBS adipogenesis that were proximal to differentially methylated regions. These genes provide a window into the epigenetic regulation of adipocyte function. The SGBS methylation pattern showed similarities to that of subcutaneous progenitors and skin primary fibroblasts.

Hyperglycemia led to an increase in glucose uptake and triglyceride accumulation, as observed in other studies. We identified 61 differentially methylated regions and 27 genes differentially expressed under hyperglycemic conditions; many of these are involved in mitochondrial function, adipose signalling, secretion, and inflammation. Some of these genes responsive to hyperglycemia might have health implications as evidenced by the gene enrichment analysis. In conclusion, this study combined genome-wide analysis of the transcriptome and epigenome of the human SGBS cell line under different glucose availability.

Introduction:

Hyperglycemia, defined as a fasting blood glucose concentration > 5.5 mmol/L or > 7.8 mmol/L at 2h after an oral glucose load, is a pre-diabetic condition associated with poor health outcomes. Hyperglycemia arises from either insufficient secretion of insulin or defects in insulin's actions, known as insulin resistance. Obesity is associated with fasting hyperglycemia (1), largely attributable to impaired insulin-mediated suppression of hepatic glucose production (insulin resistance). In the AusDiab longitudinal study, obesity (as measured by increased body mass index (BMI)) was associated with a five-fold increase in the incidence of Type 2 Diabetes (T2D) (2).

The function of adipocytes in the pathogenesis of T2D is of considerable interest. They play a pivotal role in energy homeostasis by taking up glucose under the influence of insulin when glucose levels in the blood are high, and by releasing fatty acids as a source of energy when blood glucose levels are low (3). Additionally, adipocytes have a lipid-buffering function, by suppressing the release of non-esterified fatty acids into the circulation and increasing triglyceride clearance (4). They also participate in endocrine signalling and regulate food intake and immune responses (5). While adipocytes are renewed in healthy adipose tissue through adipogenesis, in the surplus of energy, this process is also used for the expansion of adipose tissue (6). Understanding the process of human adipogenesis is of major importance for prevention and treatment of obesity and T2D.

Changes to DNA methylation in adipose tissue have been identified in obese subjects, and there is compelling evidence that epigenetic modifications are linked with metabolic health and obesity outcomes (reviewed in (7)). Recently, hyperglycemia has been found to induce inflammatory responses in cardiac skeletal muscle cells (8), endothelial cells (9), and adipose-derived mesenchymal stem cells (ADSCs) (10) and it has been proposed that this could be mediated by critical gene-activating epigenetic changes (9). We hypothesize that excess glucose availability due to hyperglycemia may modify the course of adipocyte differentiation,

which in turn may lead to perturbation in their epigenetic profile, and thus influence their phenotype and function.

Murine 3T3-L1 cells have been used extensively for mechanistic studies of adipogenesis. More recently, a human *in vitro* model with high proliferation and differentiation capacity (11), the SGBS cell line, was developed from the subcutaneous fat depot of a male infant with Simpson-Gobali-Behmel Syndrome (12). The differentiation of human ADSCs (hADSCs) has also been used in genome-wide analyses of transcription and chromatin regulation across adipogenesis (13). A small number of studies have examined the effects of hyperglycemia on murine 3T3-L1 adipocytes and found that they accumulated more lipids (14), contained higher levels of reactive oxygen species (ROS) (15), had defective insulin signalling (16) and developed insulin resistance (15, 17).

To investigate the potential hyperglycemia-induced adverse effects on human adipocytes, we differentiated SGBS pre-adipocytes into mature adipocytes under high (30 mM) and low (7.5 mM) glucose concentrations over 14 days. Our aim was to identify the dynamics of changes in DNA methylation across human adipocyte differentiation, its relationship to transcriptional changes, and how these might be affected by exposure to hyperglycemic conditions. We then investigated the physiological, metabolic, transcriptional and epigenetic trajectories during SGBS adipocyte differentiation under the low and high glucose exposure. We also compared transcriptional and DNA methylation profiles of SGBS cells with those of human primary adipocytes and hADSCs isolated from subcutaneous and visceral fat depots. In conjunction, we examined transcriptional regulation across SGBS adipogenesis by conducting genome-wide integration analysis of DNA methylation profiles with recently published ChIP-seq for the H3K4me3 promoter mark and three transcription factors; Peroxisome proliferator-activated receptor γ (*PPARγ*), CCAAT-enhancer-binding proteins α (*C/EBPα*), and liver X receptor (*LXR*) (18).

Methods:

SGBS Cell Culture Conditions & Treatments

SGBS cells were differentiated as previously described, with minor changes (19). The SGBS cells were differentiated in DMEM, low glucose (Life Technologies, Carlsbad, CA, USA). D-Glucose (Life Technologies, Carlsbad, CA, USA) was added to adjust the glucose concentrations to 7.5 mmol/L (LG0) and 30 mmol/L (HG). To test for a hyperosmolar effect, we also used 7.5 mmol/L glucose with 22.5 mmol/L D-Mannitol (Sigma, Saint Louis, MO, USA) (LG1). The LG concentration was chosen to be 7.5 mmol/L because the differentiation capacity of SGBS cells was reduced in lower (5 mmol/L) glucose concentrations (not shown). The progression of adipogenesis was compared on Day 0 (D0), Day 10 (D10 HG) and Day 14 (D14 HG or D14 LG).

The adipocyte differentiation rates for LG0 and LG1 were very similar and also no significant differences in any metabolic measures were observed between both conditions (data not shown). To ensure that any differences between the high and low glucose treatments were not due to differences in osmolarity, the HG treatment was compared to the LG1 treatment in all further analyses.

ii. Lipid Content and Lipid Droplet Size Measurements

The lipid content of differentiating adipocytes was measured on D0, D10 and D14 using a Triglyceride Accumulation Kit (Zen-Bio Inc., Durham, NC, USA). Lipid droplet (LD) diameter size density distribution was measured on D10 and D14 using LipiD-QuanT software (19).

iii. Glucose Uptake Measurement

Glucose uptake was measured on D0, D10 and D14 and biological duplicates were used per treatment group and day. One hour prior to measurement, SGBS cells were washed with PBS

and transferred to additive-free DMEM, low glucose (Life Technologies, Carlsbad, CA, USA). Cells were washed again and incubated in DMEM, low glucose containing 2 µCi/mL ^{14}C-Glucose (Perkin Elmer, Waltham, MA, USA) for 1h. Additionally, insulin stimulated glucose uptake was similarly measured in the presence of 20 nmol/L insulin (Sigma, cat#I2643).

After 1h, cells were washed with PBS and lysed in 200 µl RIPA buffer supplemented with protease and phosphatase inhibitors (Thermo Scientific, Carlsbad, CA, USA). The lysate was centrifuged at max speed for 15 min at 4^{0}C and then, 50 µl of the supernatant was solubilized in 6 ml scintillation fluid. Disintegrations per minute (DPM) from ^{14}C in the scintillation fluid were counted in a multi-purpose scintillation counter (Beckman Coulter Inc., LS650, Jersey City, NJ, USA). The readings were background corrected using the readings obtained from two empty wells treated with the ^{14}C-Glucose media. DPM were extrapolated to the whole sample and divided by the total amount of protein in the well which was determined using a BCA Protein Assay as per the manufacturer's instructions (Thermo Scientific, Carlsbad, CA, USA). Results were reported relative to D0 basal glucose uptake.

iii. Gene Expression Analysis

On D0, D10 and D14, total RNA was isolated using 500 µl TRI-reagent (Sigma, Saint Louis, MO, USA) as per the manufacturers' instructions. We examined the transcriptome profiles during adipogenesis on a total of 16 RNA samples (1 µg each) using Affymetrix GeneChip® Human Gene 2.0 ST Array cartridges (D0, D10 HG, D14 HG and D14 LG1 (n=4)). Sample labelling, hybridization and array processing were performed at the Ramaciotti Centre for Genomics (UNSW, Sydney, Australia).

Data were normalized using the Robust Multi-array Average (RMA) method from the R package *oligo* (20). Core probesets annotated as "main" were retained for analysis (44,629 probesets).

Differentially expressed (DE) genes were identified using the limma package (21). A linear model including all four sample types was constructed, and significant genes were extracted. Raw *p*-values were adjusted using the Benjamini-Hochberg method and significant DE probesets were identified based on a false discovery rate (FDR) *q*-value <0.05. Further interpretation of the data was done by Gene Set Enrichment Analysis (GSEA) (22) using an FDR *q*-value<0.05 for terms within $15 \leq x \leq 500$ range for term size *x*.

Transcriptional profiles of SGBS adipocytes were compared against those of purified human visceral and subcutaneous primary adipocytes from 3 lean individuals (Bradford et al., unpublished). These data were also generated using the Affymetrix GeneChip® Human Gene 2.0 ST Array.

iv. DNA Methylation Analysis

After cell lysis and proteinase K digestion, DNA was purified using a phenol-chloroform-isoamylalcohol method. We quantified hypomethylation levels in long interspersed nuclear elements-1 (LINE-1) using the end-specific PCR (ESPCR) assay (23). Samples were run in triplicate and normalised relative to a reference DNA sample from human blood (Roche Applied Sciences, Sydney, Australia). Hypomethylation levels were compared using one-way analysis of variance or *t*-test as appropriate.

We used Illumina Infinium HumanMethylation450 BeadChip (450k) arrays to map the genome-wide DNA methylation patterns. A total of 12 samples (D0 n=4, D10 HG n=2, and n=3 for both D14 HG and D14 LG1) were submitted to the Australian Genome Research Facility (Parkville, VIC, Australia), randomized across the array chip and processed following standard protocols. Three of the samples (one of D0, D14 HG and D14 LG1) did not meet quality criteria and were excluded from further analyses. There was high variation among the remaining duplicates of D14 HG. For the analysis of both DNA methylation and

gene expression changes during adipogenesis, irrespective of the treatment, all D14 data were combined and compared to D0.

Data were normalized using the *dasen* method from the Bioconductor software package wateRmelon (24). 211 probes that failed in one or more of the nine remaining samples based on a detection *p*-value>0.05 were excluded from the analysis. Mean beta values per sample were calculated for all probes on the array, as well as for subsets of probes based on their genomic annotation according to the 450k manifest file. The Welch *t*-test was used for comparison of mean methylation levels between the groups and the time points. Differentially methylated (DM) sites (CpG dinucleotides) and DM regions were identified using the Bioconductor packages limma (21) and DMRcate v 1.4.2 (25) respectively, using FDR<0.05.

To evaluate significant proximity of DM regions found in SGBS adipogenesis to publically available ChIP-seq data of three transcription factor binding sites on SGBS mature adipocytes (18), we applied the IntervalStats method (26), using an FDR threshold *q*-value<0.05 to evaluate individual significance for each DM regions, for each of the three transcription factor binding sites at D10 adipocytes and differential H3K4me3 peaks between D0 and D10.

The DNA methylation profiles of *in vitro* SGBS adipocytes were plotted against those of visceral and subcutaneous primary adipocytes as well as their cultured or uncultured progenitors. For this comparison, we used 450k DNA methylation profiles from 15 paired samples of human primary adipocytes and 2 paired samples of hADSCs all isolated from subcutaneous and visceral fat depots as well as 3 samples of visceral adipose tissue (Bradford et al., unpublished). Additionally, since SGBS cells have fibroblast like morphology, we also compared their methylation profile with those of fibroblast cell lines: IMR-90 (GSM999340, GSM868008) and MRC-5 (GSM999345, GSM868027) arising from published studies (27, 28).

Results:

Hyperglycemia leads to differential lipid accumulation in adipocytes

We hypothesized that excess glucose availability leads to altered adipocyte function. The SGBS pre-adipocyte differentiation was compared under two different conditions: osmolarity-adjusted Low Glucose (LG1) and High Glucose (HG) (Figure 1).

The differentiation rate was 95% under all treatment conditions based on counts of pre-adipocyte and mature adipocytes at D14. While adipogenesis appeared complete by D14 in both LG1 and HG (Figure 2A), a number of measures were indicative of differential lipid accumulation. The distribution of LD size (p=3.042e-11, Kolmogorov-Smirnov test) (Figure 2B), average total LD area per cell and average LD size (Figure 2C) were all higher under HG treatment. The majority of the lipid production occurred in the first 10 days of SGBS adipogenesis, with triglyceride content increasing a further ~25% by D14 (D10 HG vs D14 HG p=0.014, Welch t-test) (Figure 2D). HG treated mature adipocytes contained 1.7-fold more triglycerides than those grown in LG (p= 0.001, Welch t-test).

Both basal and insulin-stimulated glucose uptake were higher in HG treated mature adipocytes in comparison to LG1. Relative insulin responsiveness (ratio of basal to insulin-stimulated glucose uptake) was slightly lower under HG compared to LG1 (HG= 1.98 vs LG1=2.07) (Figure 2E).

Expression of metabolic pathway genes are up-regulated across SGBS adipogenesis

Gene expression profiles were measured at D0, D10 and D14. The pre-adipocyte (D0) transcriptional profiles showed a clear separation from adipocyte profiles of D10 and D14 in PC1 (Figure 3A). Nearly 25% of the genes on the Affymetrix Human Gene 2.0 ST array were differentially expressed (DE) between D0 and D14 (Table S1A). The number of DE genes and their direction of change are summarised in Table 1.

Gene set enrichment analysis (GSEA) of the D0 vs D14 comparison showed that cell cycle and nucleosome assembly pathways, that is functions associated with cell division, were among the most negatively enriched pathways. As expected, metabolic pathways such as oxidoreductase activity, lipid metabolic processes, fatty acid biosynthesis and gluconeogenesis were the most positively enriched pathways in mature adipocytes (Table S2A). Over 80% of the DE genes between D0 and D14 were also present between D0 and D10 (Table S1B) with similar enriched pathways (Table S2B) which indicated that the majority of the gene expression changes had occurred during terminal differentiation (from D0 to D10) and that these profiles were well-maintained during maturation (Figure 4A). From D10 to D14, transcriptional changes were minimal (Table S1C and S2C).

We assembled the DE genes that were significantly different from D0 to D10 and then to D14 into 4 modules based on their patterns of change and noticed that genes within each module functioned in similar pathways (Figure 5 and Table S3). Many of the genes in Figure 5A and B that show an overall increase in expression across time are central to adipocyte function. The drop in expression for some genes beyond D10 possibly relates to decreased need for on-going synthesis in mature adipocytes (Figure 5B). In contrast, *PTGS2* and *DKK1*, known to suppress adipogenesis and regulate self-renewal respectively, are among the genes that are expressed at D0 but continuously down-regulated across differentiation (Figure 5C). Across adipogenesis, expression of Tenascin C (*TNC*) was lowest at D10 (Figure 5D). It is associated with acute and chronic inflammation, known to be expressed in visceral adipose tissue of obese patients only (29).

Vesicle formation pathway genes are among those differentially methylated across SGBS adipogenesis

Next, we examined global levels of DNA methylation in pre-adipocytes and adipocytes at D10 and D14 using methylation of LINE-1 sequences as a surrogate. LINE-1 methylation levels showed a tendency to increase (decreased hypomethylation) across adipogenesis, with

this being more apparent in D10 HG than D14 HG (Figure 6A). Examination of methylation levels of the CpG sites on the 450k arrays showed that globally (all probes) and within the different genomic annotations the highest variation occurred in the body and 3' ends of genes (Figure 6C, *p*-values are provided in Figure 6B). Among the top 10% most variable probes per annotation, there was a decrease in average DNA methylation level from D0 to D14 and D10 to D14 when all the 450k probes were combined (Figure 6D), and this was evident across most genomic annotations.

The principal component analysis (PCA) of the 450k array DNA methylation data showed a similar trend to the transcription data, with pre-adipocytes separated from adipocytes in PC2 (Figure 3B). When comparing the methylation profiles of pre-adipocytes to D14 mature adipocytes, we found 1,090 individual DM CpG sites and 160 DM regions, with ~90% reflecting decreased methylation (Tables S4A and S5A respectively). One-third of the DM regions had a Δ-me (delta methylation values) of >0.2 and all of these regions showed a loss of methylation (Table 1). The same comparisons were also made with D0 vs D10 (Tables S4B, S5B and S6) and D10 vs D14 (Table S5C). Similar to the changes in gene expression, most changes in DNA methylation occurred from D0 to D10 (Figure 4B and C). The methylation profiles of selected DM regions identified between pre-adipocytes and D14 mature adipocytes are shown in Figure S1A. Gene ontology (GO) analysis of DM sites and their nearest protein-coding genes identified 6 significant terms, all related to regulation of extracellular membrane and vesicle formation. These terms included 15-19% (159-193) of DM sites (demethylated across adipogenesis) located in genes involved in production and transport of (extracellular) vesicles and exosomes such as: *ACACA, ALCAM, ANO6, AP1G1, AP2B1, FASN, LBP, RBP4, TOLLIP, FGFR1, LRBA, SAR1B, WLS, GPC6* and several of the solute carrier family members as well as vesicular "cargo" molecules – i.e. the adipokine *ADIPOQ* and growth factors such as *FGF1* and *PDGFD* (Table S6). Adiponectin encoded by *ADIPOQ* is exclusively secreted by adipocytes and regulates glucose levels and fatty acid breakdown and also has putative anti-inflammatory properties (30). Platelet-derived growth

factors (PDGFs) are autocrine growth factors known to regulate blood vessel development (31).

Table 1: Summary results of differential gene expression and DNA methylation.

Comparisons		Total number of DE genes (|FC|>32)[#]	Total number of DM sites (Δ-me of>0.2) [#]	Total number of DM regions (Δ-me of>0.2) [#]
Adipogenesis	**D14 vs D0**	**10,919**	**1,090**	**160**
	UP:	5,037 (33)	112	11
	DOWN:	5,882 (9)	978 (208)	149 (58)
Early Adipogenesis in High Glucose	**D10 HG vs D0**	**9,283**	**419**	**59**
	UP:	4,163 (30)	16	0
	DOWN:	5,120 (8)	403 (123)	59 (30)
Late Adipogenesis in High Glucose	**D14 vs D10 HG**	**422**	0	**14***
	UP:	184		10
	DOWN:	238		4
Adipogenesis in High Glucose	**D14 HG vs D0**	**10,017**	**694**	**96**
	UP:	4,508 (31)	47	7
	DOWN:	5,509 (10)	647 (200)	89(52)
Adipogenesis in Low Glucose	**D14 LG1 vs D0**	**9,598**	**843**	**117**
	UP:	4,371 (32)	69	6
	DOWN:	5,227 (9)	774 (214)	111 (58)
Effect of High Glucose Exposure	**D14 HG vs D14 LG1**	**27**	0	**61***
	UP:	10		32
	DOWN:	17		29

([#]) DE genes with magnitude of |Fold Change| >32, and DM sites and DM regions with magnitude of Δ-me (delta methylation values) of>0.2 are shown in brackets.

(*) DMRcate was run using relaxed parameters of all probes with Δ-me of> 0.1.

The relationship of DNA methylation and gene expression across adipogenesis

We first examined the overall relationship between the expression levels of individual genes and their profiles of DNA methylation. After binning genes into five equal-sized groups, ranked from low to high expression, we plotted mean methylation profiles across 10 kb flanking the transcription start site (TSS). For all sample groups (Figure 7A to D) and expression levels, methylation was lowest at the TSS and higher both upstream and in the gene body. There was a very clear gradation of decreasing methylation around the TSS (from

approximately -2 kb to + 2 kb) with increasing levels of expression with a converse increase in methylation further upstream or within the gene body.

We examined the degree of association between changes in DNA methylation across adipogenesis and the expression of nearby genes. Table S7A shows the distance from each of the 160 DM regions to the nearest differentially expressed gene. While representing only a small fraction of the 10,919 DE genes, 52% (83) of DM regions overlapped or lay within a DE gene, and 59% (94) within a 1 kb window either side of a DE gene. Of these 94 DM regions, 6 were hypermethylated regions and the transcription of their proximal gene was always repressed across adipogenesis. Of the 88 hypomethylated regions, 63 (72%) were associated with a proximal gene that was up-regulated across adipogenesis. Genes with critical involvement in adipocyte function, such as *ADIPOQ, CIDEC, LBP, FASN, FAR2* and *LIPE* were all up-regulated and contained a hypomethylated region within the gene (Figure 8A).

The relationship of DNA methylation and transcription factor binding across adipogenesis

Recently, Galhardo et al. (18) used ChIP-seq profiling for the binding sites for three transcription factors, *PPARγ*, *C/EBPα* and *LXR* in D10 SGBS adipocytes and for the H3K4me3 promoter mark between D0 and D10. We overlapped our identified 160 DM regions with this ChIP-seq data (Table 2). We found 16 DM regions, all hypomethylated across adipogenesis, overlapped with either *C/EBPα*, *LXR* or *PPARγ* high occupancy binding sites but there was no significant overlap of our DM regions and the altered H3K4me3 marks across the differentiation (18).

When hypomethylation of promoter regions overlapped with putative transcription factor binding site, it coincided with transcriptional activation of the gene in 8 of the 10 cases (Table 2). Some of these genes are known to have adipocyte specific functions; *THRSP, PFKFB1,*

CIDEC ADIPOQ and *ACSL1* expression increase by 28-fold or more across adipogenesis. Also, *ACSL1* contained both LXR and *PPARγ* binding sites (Table 2) and its methylation profile has been found to be highly responsive to weight loss in human subcutaneous adipose tissue (32). Similarly, *CORO2B*, is among the top responsive genes to weight loss in the same study (32), is hypomethylated across a *PPARγ* binding region about 4.3 kb upstream of its TSS (Table 2). The other genes in Table 2 (i.e. *TMEM184B, AHCYL1, ZNF22, PLEKHG6* and *HBP1*) show significant responses and should be checked in the context of obesity.

Table 2: The regions of overlap between the identified DM regions in our dataset with the binding sites of three transcription factors published in (18).

	Galhardo et al. (18)						
	C/EBPα **ChIP peaks (n=1,157)** [#]	***LXR*** **ChIP peaks (n=96)** [#]	***PPARγ*** **ChIP peaks (n=4,127)** [#]	**DM regions in D14 vs D0**		**DE genes in D14 vs D0**	
Gene Name				**Δ-me***	**FDR**	**FC***	**FDR**
CYP1B1	Gene Body			-0.18	1.30E-22	-1.84	3.92E-08
THRSP		Promoter		-0.28	8.98E-55	626.60	3.63E-23
ACSL1		Promoter	Promoter	-0.25	4.19E-77	28.84	1.12E-17
PFKFB1			Promoter	-0.23	1.25E-43	36.04	1.26E-14
HOXB2			Promoter	-0.09	2.14E-14	-1.95	3.99E-08
CIDEC			Promoter	-0.08	3.70E-24	221.22	4.69E-19
ADIPOQ			Promoter	-0.16	8.62E-32	248.87	2.55E-18
ZNF22			Promoter	-0.10	5.63E-16		
PLEKHG6			Promoter	-0.26	2.31E-39		
AHCYL1			Promoter	-0.12	8.36E-18	1.52	4.12E-06
TMEM184B			Promoter	-0.13	3.83E-16	-1.43	1.96E-05
HBP1			Promoter	-0.18	2.39E-18	1.35	0.0008
FGFR1			Gene Body	-0.08	2.46E-14		
GPAM			Intergenic	-0.28	8.98E-55	48.40	1.05E-19
CORO2B			Intergenic	-0.12	1.93E-15		
RRM1			Intergenic	-0.21	4.42E-54	-1.73	3.86E-07

(#) Transcription factor binding sites with high occupancy (> 30 ChIP peaks) were included to the analysis. (*) For genes with overlapping DM regions, the change in beta value is noted under Δ-me (delta methylation values) and for genes that are differentially expressed the magnitude of change is noted under FC (fold change).

SGBS adipocytes show transcriptional and epigenetic similarities to *in vivo* human adipocytes

We investigated whether DNA methylation profiles of *in vitro* grown SGBS cells were similar to those obtained from human adipose tissue: human primary adipocytes and hADSCs either before or after cell culture. Since SBGS cells have fibroblast-like features, we also compared their methylation profile with those of fibroblast cells lines IMR-90 and MCF-5. The PCA of the DNA methylation data showed a distinct separation in the first component of uncultured hADSCs and tissues from those of a range of cultured cells, including cultured hADSC and their *in vitro* differentiated adipocytes, SGBS and fibroblast cells (Figure 9). On the PC2, both SGBS pre-adipocytes and mature adipocytes, segregated with other subcutaneous adipose tissue samples – isolated adipocytes, SVF fraction and cultured progenitor cells, reflecting their depot of origin, subcutaneous fat (Figure 9).

We also compared the transcriptional profiles of SGBS cells to human primary adipocytes to ascertain whether *in vitro* grown SGBS cells were similar to those directly isolated from the human body. A heat map illustration of this comparison on selected DE genes is provided in Figure 10**.** Overall, some of the adipose-specific gene expression levels were consistent between SGBS adipocytes and *in vivo* subcutaneous adipocytes more than *in vivo* visceral adipocytes (Figure 10). For some genes clear differences were observed between SGBS and *in vivo* adipocytes (Figure 10). For instance, *GLYAT*, *LEP*, *IGF1* and *LPL* are higher expressed while *G6PD*, *FASN*, *RELA* and *FABP3* are lower expressed in SGBS adipocytes compared to *in vivo* adipocytes (Figure 10).

Additionally, our results had a significant overlap with publically available transcription data of differentiating SGBS and hADSCs. Our observed DE genes included 87% of the 272 unique DE genes associated with metabolic activity and 80% of the 75 unique DE genes associated with transcription factor binding sites identified across SGBS adipogenesis in (18)

as well as 69% of the 1670 differentially DE genes (with |Fold Change| >2) identified across adipogenesis of hADSCs (13).

Transcriptional and epigenetic changes induced by hyperglycemia

To investigate the extent of hyperglycemia-induced molecular changes in human adipocytes, we compared the transcription and methylation profiles of the mature adipocytes (D14) differentiated in the presence of 7.5 or 30 mmol/L glucose.

Within the overall context of gene expression changes during adipogenesis, the effect of elevated glucose was modest, with the transcriptional profiles of D14 HG separated from the D14 LG1 in PC5 of the PCA plot (Figure 3A). We found 27 significantly DE genes between D14 HG and D14 LG (Table S1D, summarized in Table 1). A schematic illustration of these genes is provided in Figure 8B. Some of the downregulated genes in HG vs LG1 have been associated with obesity, glucose sensitivity and body weight regulation such as *GLYAT*, *IGF1*, *ACAN*, *NDUFA9*, *LPL GREM1*, *MMP8* and *TNC*.

By using gene set enrichment analysis (GSEA) we were able to examine the impact of elevated glucose at a pathway level (where transcriptional changes in individual genes might not reach significance). This analysis identified negative enrichment in 50 biological pathways when comparing between D14 HG and D14 LG1 samples (GSEA, Table S2D). The overwhelming majority of these were mitochondrial pathways, while the rest included signal transduction and extracellular stress-activated oxidative stress pathways (Table S2D).

In the methylation dataset, separation between D14 HG and D14 LG1 samples was also evident in the fifth principal component. However, because of the higher variance between duplicates, no DM sites passed the 0.05 FDR threshold. Hence, we ran DMRcate using all probes with Δ-me of > 0.1. This identified 61 DM regions with a maximum methylation change of 23% identified between HG and LG treatment (Table S4D). None of the DM regions was associated with a DE gene, but multiple regions were associated with genes that

may have a role in glycemic control, transport, secretion and inflammation in adipocytes, eg. *DEPTOR*, *GNPNAT1*, *EIF4B*, *PRRC2A*, *RAB4A*, *PRRC2A*, *MTX3* and *EP400NL*. Figure 8B summarises the known function of hyperglycemia-responsive genes that were either differentially expressed or methylated in our study. Examples of methylation patterns across hyperglycemia responsive regions are shown in Figure S1B.

Discussion:

The adipogenic transcription program occurs with minimal change in DNA methylation

Widespread transcriptional changes were characterised across the process of differentiation of SGBS cells, encompassing 25% of genes measured. These changes were in good agreement with publically available transcription data, including over 70% of the previously identified genes in both differentiating SGBS cells (18) as well as *in vitro* differentiation of hADSCs (13). As expected, metabolic pathways were up-regulated while cell cycle pathways were down regulated during adipogenic conversion (13, 18). On the other hand, the methylation changes during adipogenesis were modest, with a bias toward hypomethylation changes. Therefore, extensive changes in gene expression across adipogenesis occur without DNA methylation change; this indicates that the role of DNA methylation in providing an epigenome state permissive for the adipocyte expression program is largely established by the pre-adipocyte state. This is consistent with previous data on promoter methylation on adipose mesenchymal stem cells, pre- and post- differentiation in culture (33). While the PCA plot of DNA methylation profiles shows the likely effect of cell culture in the first dimension, SGBS cells cluster closely with both primary mature and progenitor subcutaneous adipocytes in the second dimension.

During adipogenesis, genes associated with differentially methylated sites were enriched for pathways of extracellular vesicle formation and production of exosomes. This indicates a potential role for epigenetic signatures in regulation of extracellular vesicle formation in adipocytes and a role in intercellular communication between adipose tissue cell types. SGBS-secreted extracellular vesicles were previously reported to induce inflammation in macrophages, which reciprocally inhibits insulin signalling in SGBS cells (34). While small in number, nearly 60% of the DM regions were located within a 1 kb distance to a differentially expressed gene which indicates potential epigenetic regulation of gene expression. The majority of these regions were hypomethylated and resided in critical

adipogenic genes which were up-regulated. Additionally, all the DM regions that overlapped with either *C/EBPα*, *LXR* or *PPARγ* binding sites were hypomethylated across adipogenesis so that transcription could occur (Table 2). Across adipogenesis the DNA methylation changes occur in the absence of DNA replication; i.e. they must involve an active process. A recent study has shown that CCCTC-binding factor (CTCF) chromatin binding can promote *PPARγ* transcription binding and result in the activation of its target gene via active DNA demethylation pathway using TET methylcytosine dioxygenase (TET) enzymes (35). Hence, the transcription factor binding is either directly involved in the demethylation process or it is a subsequent event. In our data, 6 of the 10 regions where *PPARγ* binding sites overlapped with a demethylated promoter site of a gene, its expression was up-regulated across adipogenesis (Table 2). Additionally, differential methylation in a number of genes functioning in lipid metabolism coincided with *PPARγ* binding sites, and some of these overlaps (eg. *ACSL1*, GPAM, and *ADIPOQ*) were also observed across murine 3T3-L1 adipogenesis (36). A number of small nucleolar RNAs (snoRNAs) showed variable expression levels due to high glucose exposure in adipocytes (Figure 8A). Their function is largely unknown but a recent review has reported that they regulate alternative splicing and post-transcriptionally modify other RNAs like small nuclear RNAs (snRNAs) and ribosomal RNAs (rRNAs) (37).

Beige-like and other characteristics of SGBS adipocytes

Like Murlholm et al. (38), our transcription data suggested that SGBS cells could also be used to study 'beige-like' adipocyte formation (*UCP1* and *UCP2)*. Thus, SGBS cells show potential to be used to discover new metabolic targets for metabolic disease (39).

Simpson-Golabi-Behmel syndrome is commonly caused by mutational or deletion associated loss of activity of *GPC3* and *GPC4* genes (40). Glypicans promote growth factor signalling but mutation in these genes were not identified in SGBS cells. Our transcription data showed that both the basal level of expression and the fold change of glypicans across SGBS

adipogenesis are reduced compared to hADSCs adipogenesis (13). One conclusion can be that the reduction in expression can be partially mimicking the loss of glypican function in SGBS cells. Along with these genes, *GPC6* was slightly up-regulated during SGBS adipogenesis but down-regulated in hADSCs (13). Perhaps, mutations in *GPC6* could be the third contributor to the complex clinical phenotype of this condition.

Hyperglycemia treatment induced greater reliance to the breakdown of glucose for fatty acid synthesis rather than mitochondrial energy production

Glucose concentrations for this experiment were chosen to reflect normal and high serum levels observed physiologically. While 7.5 mM is at the upper limit of healthy human physiological glucose concentrations, it is well below the glucose concentration of 17.5 mM that the SGBS cells have been maintained in culture and 30 mM is considered as higher end of what may be observed clinically in diabetic patients.

SGBS adipocytes grown in high glucose were larger, had more lipids and reduced relative insulin responsiveness (but similar absolute responsiveness) compared to those grown in low glucose. During *in vitro* differentiation, fat can only be generated through *de novo* lipogenesis. Collins et al. (41) showed that isotopically (^{13}C)-labelled glucose contributed over 40% of the carbons contained in triglycerides produced by hASDC during adipogenesis. In line with these observations, in our transcription data, hyperglycemic exposure suppressed expression of mitochondrial genes. Decreased mitochondrial activity is consistent with a greater reliance on glycolysis for energy production and ultimately conveys an inefficient production of energy from higher levels of glucose due to the diversion of intermediate of substrates, such as citrate and pyruvate into fatty acid synthesis pathway. This could be achieved by the release of citrate to the cytosol; which is used as a substrate in *de novo* lipogenesis. Additionally, SGBS adipocytes might be metabolising excess glucose via anabolic pathways such as *de novo* palmitate and oleate syntheses and a combination of oxidative phosphorylation and aerobic respiration just as they do when exposed to excess

fructose (42). Consequently, lactate production would inhibit lipolysis in adipocytes (43) and thus will also induce fat accumulation. Translation of our results into *in vivo* situation, adipocytes synthesise higher levels of fatty acid from additional glucose rather than increasing mitochondrial energy production. However, there is need for detailed metabolic analysis, such as in (42) to characterize diversions of metabolites across pathways due to hyperglycemia treatment.

Hyperglycemia-induced DNA methylation changes (and expression) were modest but this is partly due to high variation among the duplicates of the HG samples at D14. Our results suggest that this variation is likely to be a biological effect of the high glucose exposure. Firstly, the quality of the methylation data and bisulfite conversion percentages were comparable among all the samples, and secondly, among HG treated samples, the variation in the metabolic measurements was also higher. Additionally, SGBS cells were cultured in high glucose concentration (17.5 mM) for number of passages. It would be an interest to culture primary cells in low glucose prior to differentiation under hyperglycemic conditions. Even so, the identified 27 DE genes and 61 DM regions between high and low glucose exposure provide a list of putative genes whose association with obesity and T2D should be explored. Several of the DE genes responsive to hyperglycemic exposure were previously associated with adverse health risks. Perhaps paradoxically, *IGF1* showed reduced expression under hyperglycemic conditions; however, its expression in adipocytes was also reduced in a mouse model of diet-induced obesity and it was suggested to play a homeostatic role against metabolic stress (44). Also, *TNC* was up-regulated along with *MMP8* under hyperglycemic conditions; both are involved in acute and chronic inflammation and the same changes have been observed in visceral adipose tissue of obese humans (29). Hyperglycemia-responsive DM regions were found in *DEPTOR*, which is expressed higher in obese mice (45), and in *PRRC2A* (*BAT2*), which is at the vicinity of *TNF-α* regulates inflammatory process of insulin.

In conclusion, hyperglycemia induced perturbations in the expression and epigenetic signatures of key adipose tissue functions such as mitochondrial activity, anabolic and catabolic metabolism as well as inflammation response. A number of key gene targets have been identified as having a central role in sequestration of fat into adipocytes under high glucose flux, which will potentially provide insight into the co-existence of obesity and T2D.

Acknowledgements:

The study is part of the EpiSCOPE project supported by the Science and Industry Endowment Fund (Australia), grant RP03-064 and Department of Biological Sciences in Faculty of Sciences, Macquarie University. In addition, HV is supported by iMQRES and CSIRO OCE Doctoral scholarships. We thank Stephen Bradford, CSIRO Food and Nutrition and David James, the Garvan Institute for providing the SGBS cell line, Thu Ho, CSIRO Food and Nutrition for helping with cell culture, Musarat Ishaq, Materials Science and Engineering for providing few reagents, and last but not least to Penny Bean, Meg Evans and Vijay Vaithilingam, Materials Science and Engineering for their assistance with the their microscopes. Yagiz A. Aksoy, Macquarie Universtiy, Keith N. Rand, CSIRO Food and Nutrition, Blake C. Ellis, CSIRO Food and Nutrition, Michael Buckley CSIRO Digital Productivity, and Andrew J. Beattie, Macquarie University for cultivating discussions.

H.V. designed and conducted the experiments, researched data and wrote the manuscript. P. L. M. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. P.L. M. designed the experiments and edited the manuscript. T. J. P. analysed the methylation and transcription array datasets and reviewed manuscript. P. V. conducted the adipocyte image analyses. S. J. D. and M.S contributed to discussion and reviewed manuscript. M. R. G. and S. J. C. reviewed the manuscript.

Conflict of interest:

The authors declare no conflict of interest.

References:

1. Tirosh A, Shai I, Afek A, Dubnov-Raz G, Ayalon N, Gordon B, et al. Adolescent BMI Trajectory and Risk of Diabetes versus Coronary Disease. N Engl J Med 2011;364(14):1315-1325
2. Tanamas SK, Magliano DJ, Lynch B, Seth IP, Willenberg L, Polkinghorne KR, et al. AusDiab Report 2012. The Australian Diabetes, Obesity and Lifestyle Study 2012.
3. Rosen ED, Spiegelman BM. Adipocytes as regulators of energy balance and glucose homeostasis. Nature 2006;444(7121):847-853
4. Frayn KN. Adipose tissue as a buffer for daily lipid flux. Diabetologia 2002;45(9):1201-1210
5. Wellen KE, Hotamisligil GS. Inflammation, stress, and diabetes. J Clin Invest 2005;115(5):1111-1119
6. Arner E, Westermark PO, Spalding KL, Britton T, Rydén M, Frisén J, et al. Adipocyte turnover: Relevance to human adipose tissue morphology. Diabetes 2010;59(1):105-109
7. van Dijk SJ, Molloy PL, Varinli H, Morrison JL, Muhlhausler BS. Epigenetics and human obesity. Int J Obesity 2015;39(1):85-97
8. Villeneuve LM, Reddy MA, Lanting LL, Wang M, Meng L, Natarajan R. Epigenetic histone H3 lysine 9 methylation in metabolic memory and inflammatory phenotype of vascular smooth muscle cells in diabetes. PNAS 2008;105(26):9047-9052
9. Cooper ME, El-Osta A. Epigenetics: mechanisms and implications for diabetic complications. Circ Res 2010;107(12):1403-1413
10. Ronningen T, Shah A, Reiner AH, Collas P, Moskaug JO. Epigenetic priming of inflammatory response genes by high glucose in adipose progenitor cells. Biochem Biophys Res Commun 2015;467(4):979-986
11. Fischer-Posovszky P, Newell FS, Wabitsch M, Tornqvist HE. Human SGBS cells - a unique tool for studies of human fat cell biology. Obes Facts 2008;1(4):184-189
12. Wabitsch M, Brenner RE, Melzner I, Braun M, Moller P, Heinze E, et al. Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. Int J Obesity 2001;25(1):8-15
13. Mikkelsen TS, Xu Z, Zhang X, Wang L, Gimble JM, Lander ES, et al. Comparative epigenomic analysis of murine and human adipogenesis. Cell 2010;143(1):156-169
14. Lin Y, Berg AH, Iyengar P, Lam TK, Giacca A, Combs TP, et al. The hyperglycemia-induced inflammatory response in adipocytes: the role of reactive oxygen species. J Biol Chem 2005;280(6):4617-4626
15. Lu B, Ennis D, Lai R, Bogdanovic E, Nikolov R, Salamon L, et al. Enhanced sensitivity of insulin-resistant adipocytes to vanadate is associated with oxidative stress and decreased reduction of vanadate (+5) to vanadyl (+4). J Biol Chem 2001;276(38):35589-35598
16. Gagnon AM, Sorisky A. The effect of glucose concentration on insulin-induced 3T3-L1 adipose cell differentiation. Obesity Res 1998;6(2):157-163
17. Tang S, Le-Tien H, Goldstein BJ, Shin P, Lai R, Fantus IG. Decreased in situ insulin receptor dephosphorylation in hyperglycemia-induced insulin resistance in rat adipocytes. Diabetes 2001;50(1):83-90
18. Galhardo M, Sinkkonen L, Berninger P, Lin J, Sauter T, Heinaniemi M. Integrated analysis of transcript-level regulation of metabolism reveals disease-relevant nodes of the human metabolic network. Nucleic Acids Res 2014;42(3):1474-1496
19. Varinli H, Osmond-McLeod MJ, Molloy PL, Vallotton P. LipiD-QuanT: A novel method to quantify lipid accumulation in live cells. J Lipid Res 2015;56(11):2206-22016
20. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics 2010;26(19):2363-2367

21. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47
22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. PNAS 2005;102(43):15545-15550
23. Rand KN, Molloy PL. Sensitive measurement of unmethylated repeat DNA sequences by end-specific PCR. Biotechniques 2010;49(4):13-17
24. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.
25. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, R VL, et al. De novo identification of differentially methylated regions in the human genome. Epigenet Chromatin 2015;8:6
26. Chikina MD, Troyanskaya OG. An effective statistical evaluation of ChIPseq dataset similarity. Bioinformatics 2012;28(5):607-613
27. Nazor KL, Altun G, Lynch C, Tran H, Harness JV, Slavin I, et al. Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. Cell Stem Cell 2012;10(5):620-634
28. The EPC. An Integrated Encyclopedia of DNA Elements in the Human Genome. Nature 2012;489(7414):57-74
29. Catalan V, Gomez-Ambrosi J, Rodriguez A, Ramirez B, Rotellar F, Valenti V, et al. Increased tenascin C and Toll-like receptor 4 levels in visceral adipose tissue as a link between inflammation and extracellular matrix remodeling in obesity. J Clin Endocrinol and Metab 2012;97(10):E1880- E1889
30. Hu E, Liang P, Spiegelman BM. AdipoQ is a novel adipose-specific gene dysregulated in obesity. J Biol Chem 1996;271(18):10697-10703
31. Jitariu AA, Cimpean AM, Kundnani NR, Raica M. Platelet-derived growth factors induced lymphangiogenesis: evidence, unanswered questions and upcoming challenges. AMS 2015;11(1):57-66
32. Benton MC, Johnstone A, Eccles D, Harmon B, Hayes MT, Lea RA, et al. An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. Genome Biol 2015;16(1):1-21
33. Noer A, Sorensen AL, Boquest AC, Collas P. Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured, and differentiated mesenchymal stem cells from adipose tissue. Mol Biol Cell 2006;17(8):3543-3556
34. Kranendonk ME, Visseren FL, van Balkom BW, Nolte-'t Hoen EN, van Herwaarden JA, de Jager W, et al. Human adipocyte extracellular vesicles in reciprocal signaling between adipocytes and macrophages. Obesity (Silver Spring) 2014;22(5):1296-1308
35. Dubois-Chevalier J, Oger F, Dehondt H, Firmin FF, Gheeraert C, Staels B, et al. A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and transcriptional induction of adipocyte differentiation. Nucleic Acids Res 2014;42(17):10943-10959
36. Oger F, Dubois-Chevalier J, Gheeraert C, Avner S, Durand E, Froguel P, et al. Peroxisome proliferator-activated receptor gamma regulates genes involved in insulin/insulin-like growth factor signaling and lipid metabolism during adipogenesis through functionally distinct enhancer classes. J Biol Chem 2014;289(2):708-722
37. Bhartiya D, Scaria V. Genomic variations in non-coding RNAs: Structure, function and regulation. Genomics 2016;107(2-3):59-68
38. Murholm M, Isidor MS, Basse AL, Winther S, Sorensen C, Skovgaard-Petersen J, et al. Retinoic acid has different effects on UCP1 expression in mouse and human adipocytes. BMC Cell Biol 2013;14:41

39. Tews D, Fischer-Posovszky P, Fromme T, Klingenspor M, Fischer J, Ruther U, et al. FTO deficiency induces UCP-1 expression and mitochondrial uncoupling in adipocytes. Endocrinology 2013;154(9):3141-3151
40. Cottereau E, Mortemousque I, Moizard MP, Burglen L, Lacombe D, Gilbert-Dussardier B, et al. Phenotypic spectrum of Simpson-Golabi-Behmel syndrome in a series of 42 cases with a mutation in GPC3 and review of the literature. Am J Med Genet 2013;163c(2):92-105
41. Collins JM, Neville MJ, Pinnick KE, Hodson L, Ruyter B, van Dijk TH, et al. De novo lipogenesis in the differentiating human adipocyte can provide all fatty acids necessary for maturation. J Lipid Res 2011;52(9):1683-1692
42. Varma V, Boros LG, Nolen GT, Chang C-W, Wabitsch M, Beger RD, et al. Metabolic fate of fructose in human adipocytes: a targeted (13)C tracer fate association study. Metabolomics 2015;11(3):529-544
43. Xu H-J. Role of lactate in lipid metabolism, just always inhibiting lipolysis? J Biol Chem 2009;284(31):le5
44. Chang HR, Kim HJ, Xu X, Ferrante AW, Jr. Macrophage and adipocyte IGF1 maintain adipose tissue homeostasis during metabolic stresses. Obesity (Silver Spring) 2016;24(1):172-183
45. Laplante M, Horvat S, Festuccia WT, Birsoy K, Prevorsek Z, Efeyan A, et al. DEPTOR cell-autonomously promotes adipogenesis, and its expression is associated with obesity. Cell Metab 2012;16(2):202-212

Figure 1: Schematic design of the experimental outline. SGBS cells plated on Day 0 and differentiated for 4 days in Quick differentiation medium. The cells maintained in Differentiation medium for the following 10 days.

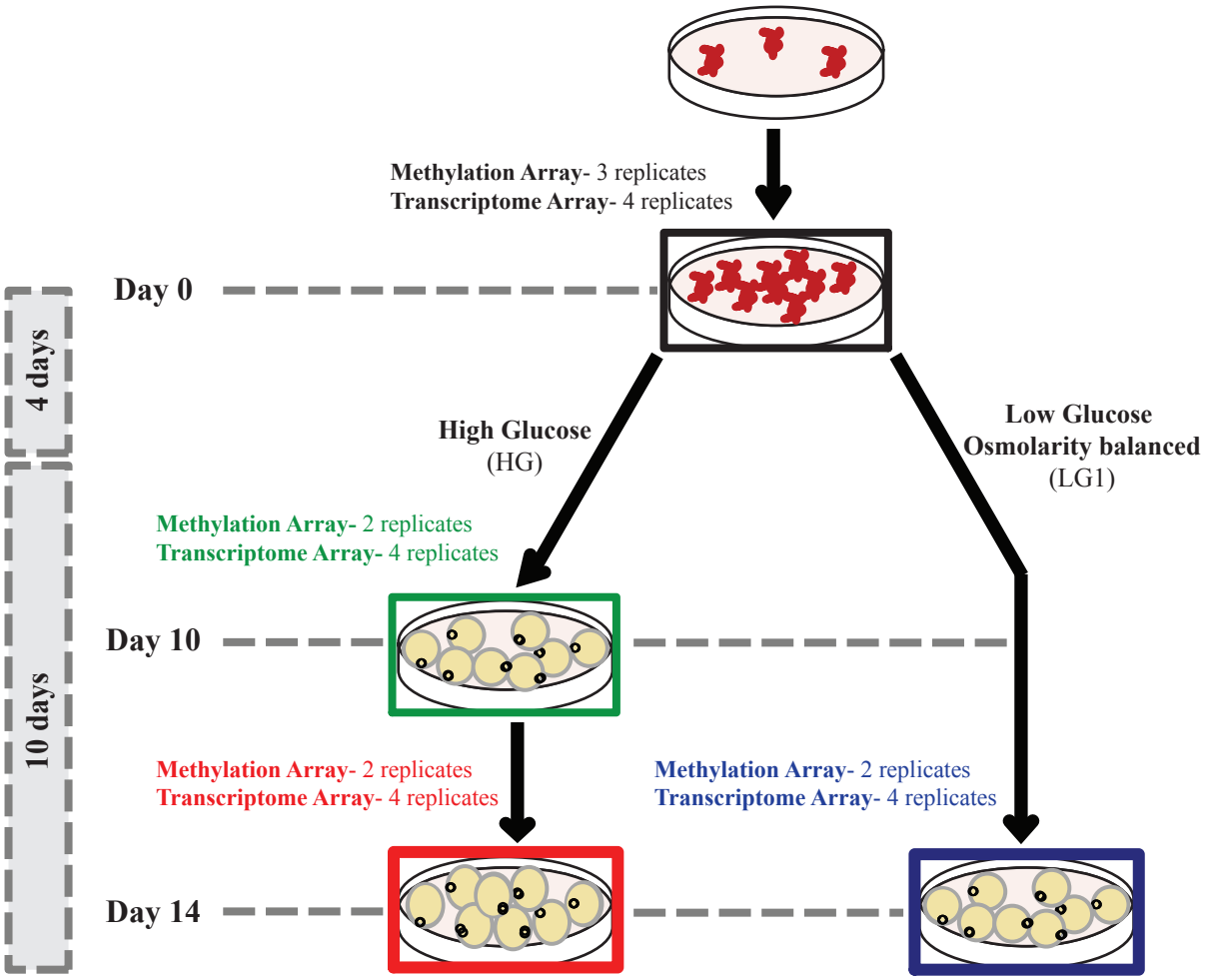
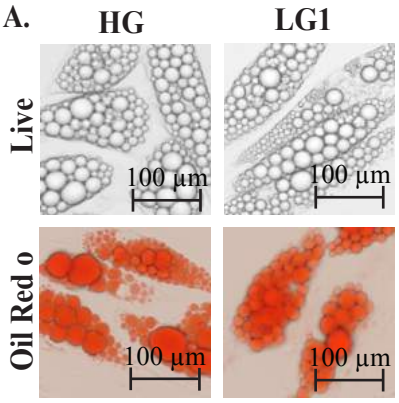
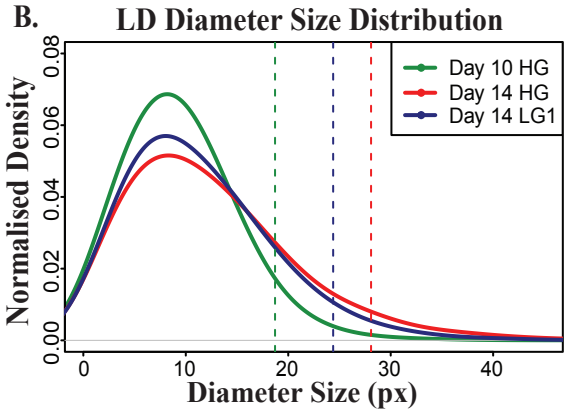


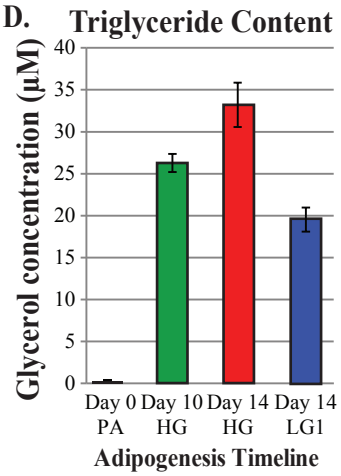
Figure 2: Physiological and metabolic measurements of SGBS adipogenesis progression.
A. Representative microscopy images live and Oil Red o stained SGBS adipocytes at D14.
B. Smoothed kernel distribution of LD diameter sizes of SGBS adipocytes treated with HG or LG1. The vertical lines indicate 95th percentile (Images n=6). There were a total of 98, 87 and 100 cells used in D10 HG, D14 HG and D14 LG1 groups respectively, **C.** Measurement of total LD area per cell and per lipid droplet, **D.** Quantification of total triglyceride content using enzymatic digest method. The total glycerol concentration is shown in µM. Data are shown as M±SD (n=3, technical n=2).
E. Basal and insulin stimulated glucose uptake normalized to total protein levels of the cells relative to pre-adipocytes. Data are shown as M±SD (n=2).





C.

Groups	LD Area in µm2	
	per Cell	per LD
D10 HG	960.98	8.29
D14 HG	1308.60	15.70
D14 LG1	1221.05	11.71



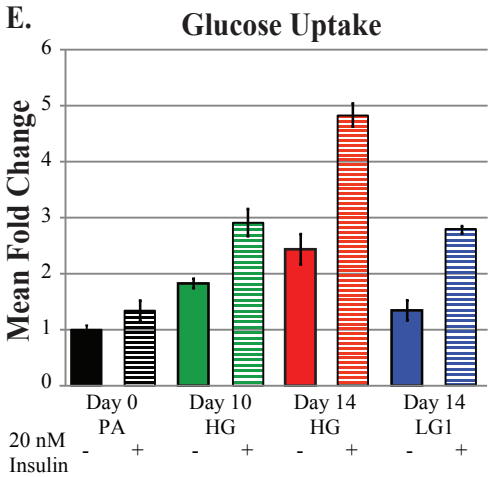


Figure 3: The first five principal components of **A.** transcription and **B.** methylation profiles of Day 0 pre-adipocytes, HG treated Day 10 mature adipocytes as well as HG and LG1 treated Day 14 mature adipocytes.

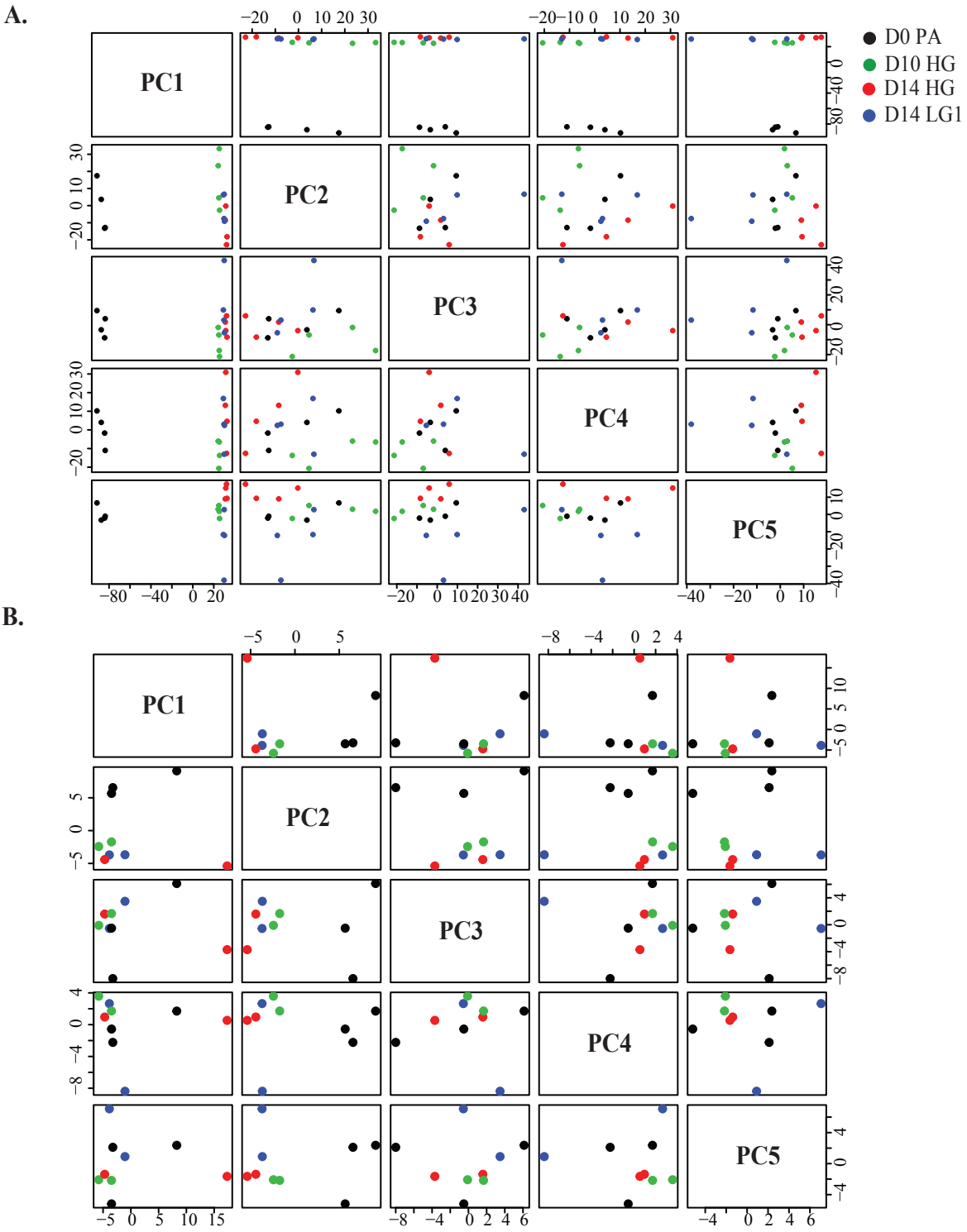


Figure 4: Venn diagram of **A.** and **D.** differentially expressed (DE) genes, **B.** and **E.** differentially methylated (DM) sites, and **C.** and **F.** DM regions in early vs complete SGBS adipogenesis and adipogenesis under HG vs LG1 conditions respectively.

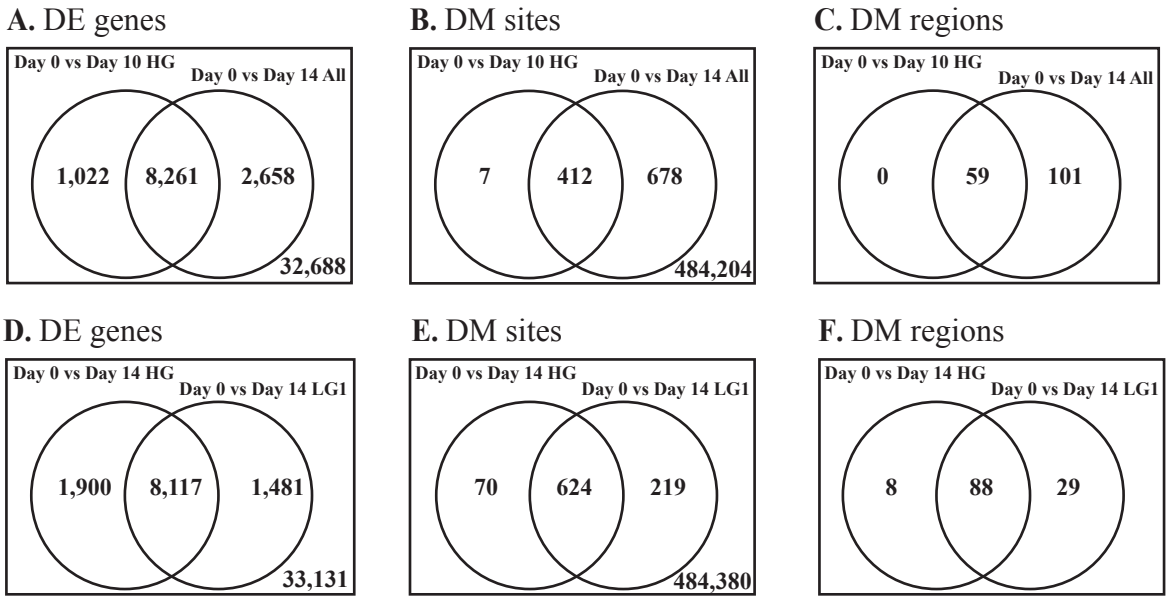


Figure 5: Gene expression profiles of four gene modules, clustered by their temporal expression pattern. Each line represents one gene probeset. **A.** Module 1- 89 genes are up-regulated at both time intervals, **B.** Module 2- 118 genes have a peak in their transcription at Day 10, **C.** Module 3- 122 genes are down-regulated at both time-intervalsand **D.** Module 4- 19 genes have a trough in their transcription at Day 10.

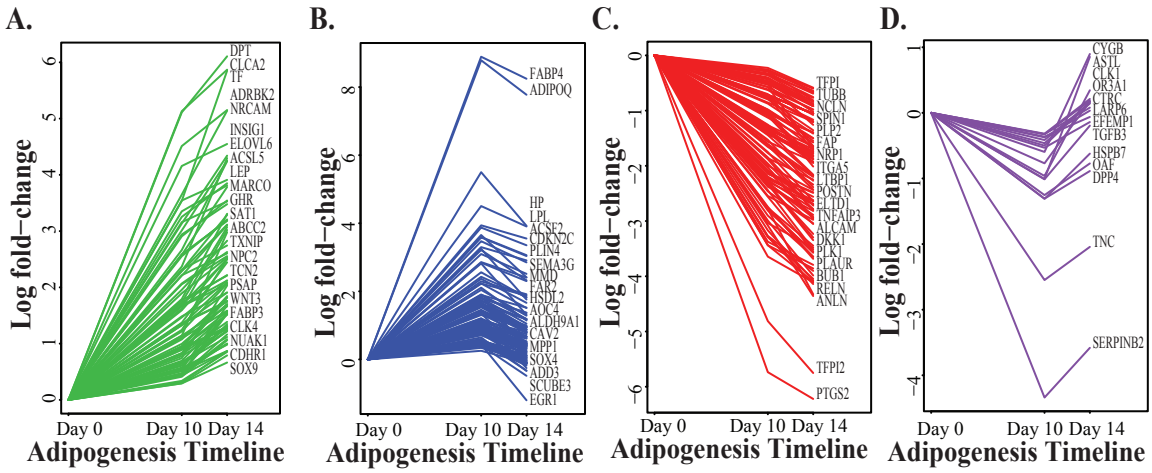
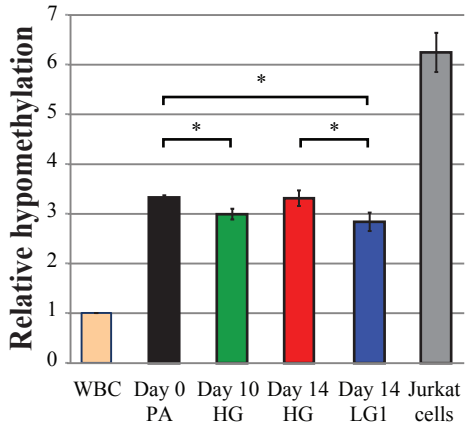


Figure 6: A. DNA methylation changes across SBGS adipogenesis. White blood cells (WBC) and Jurkat cells (T lymphocyte cells). **B.** T-test p-values for significant comparisons in Figure 7A and C, **C.** Global methylation levels from 450k array, **D.** Methylation levels in top 10% of the most variable probes per annotation.

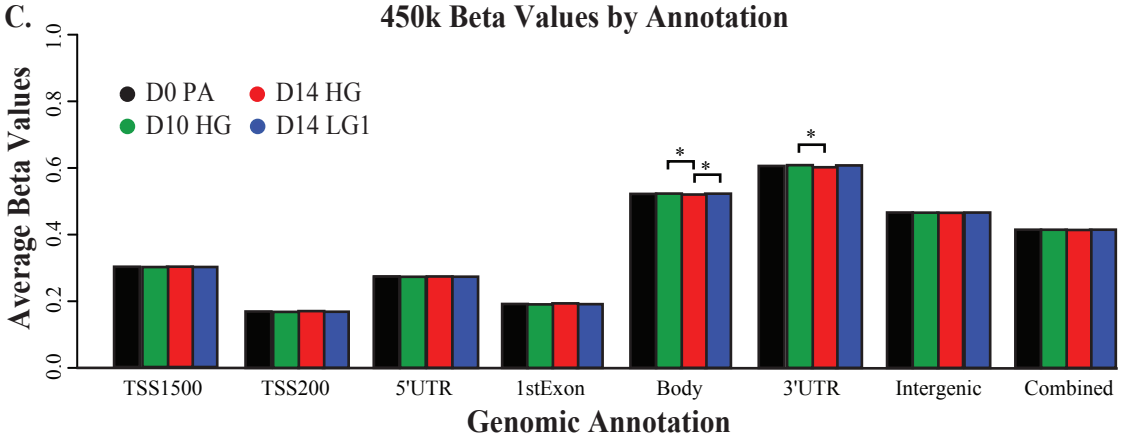
A.



B.

Comparisons:		p-value:
LINE1 Fig 7A	D0 PA vs D10 HG	0.0039
	D0 PA vs D14 LG1	0.0032
	D14 LG1 vs D14 HG	0.0276
450k Fig 7C	D10 HG vs D14 HG (Body)	0.0102
	D10 HG vs D14 HG (3'UTR)	0.0176
	D14 LG1 vs D14 HG (Body)	0.0199

C.



D.

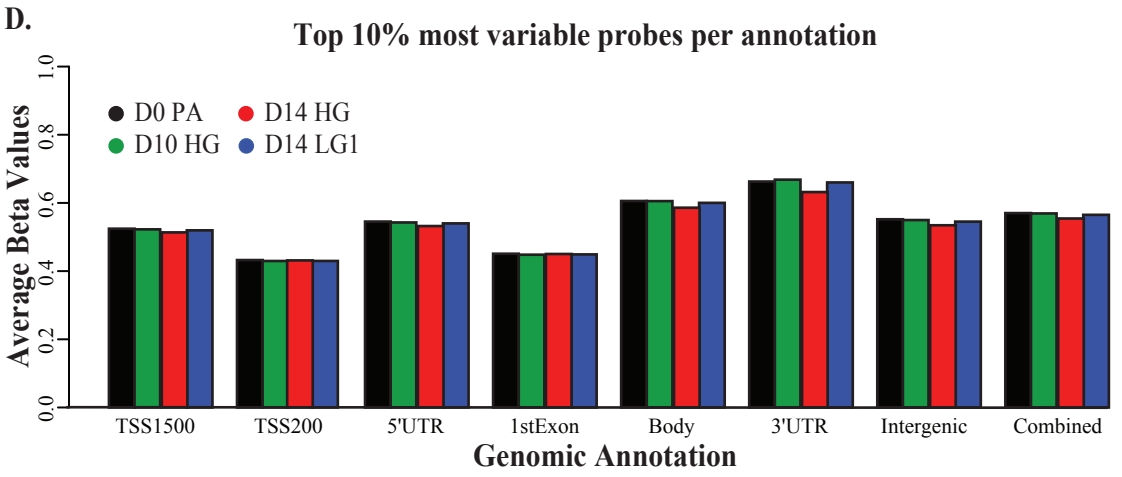


Figure 7: Mean methylation profiles flanking 5,000 bp either side of a transcriptional start site (TSS), after binning the corresponding genes into five equally-sized expression groups ranked on expression level. **A.** Day 0 pre-adipocytes, **B.** Day 10 HG adipocytes, **C.** Day 14 LG1 adipocytes and **D.** Day 14 HG adipocytes.

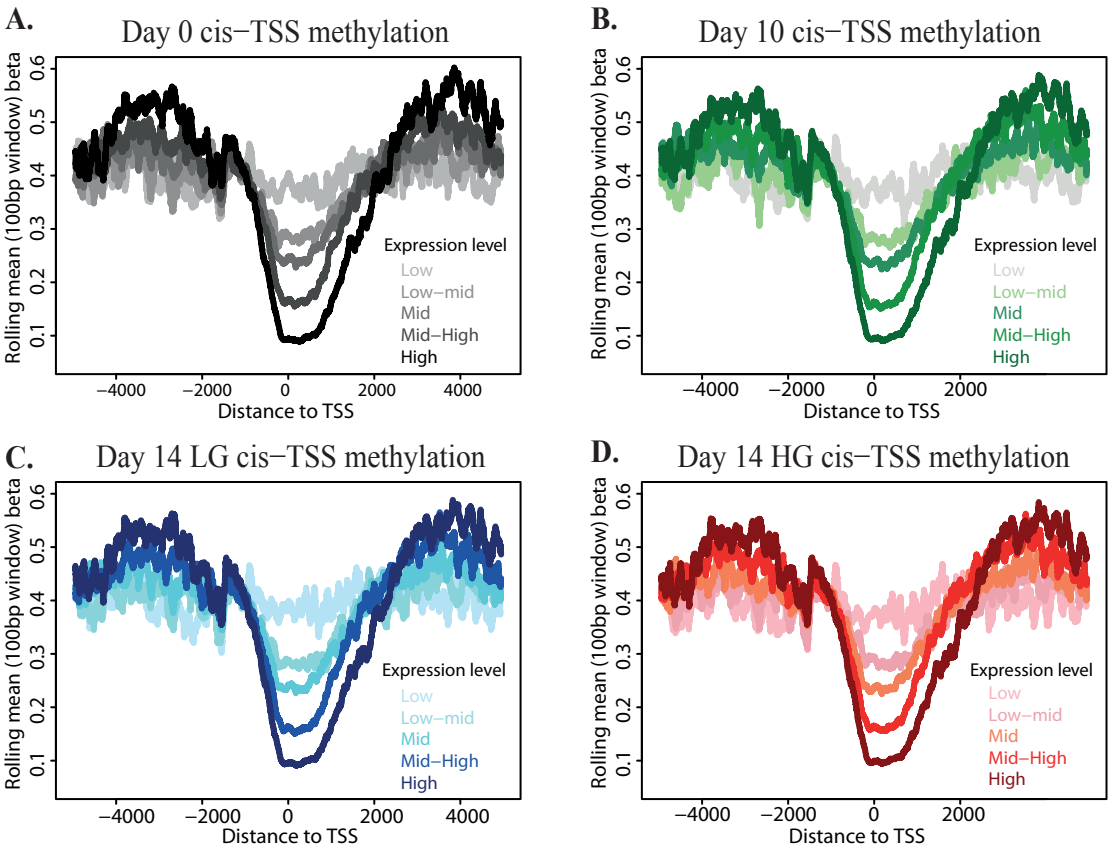


Figure 8A: Key transcriptional and DNA methylation changes during *in vitro* SGBS adipogenesis. Differential genes with **increased** and reduced expression levels from D0 to D14 are shown in black and red respectively. Underlines indicate presence of DMRs within a distance of 10 kb distance from that gene in the same comparison (promoter, gene body).

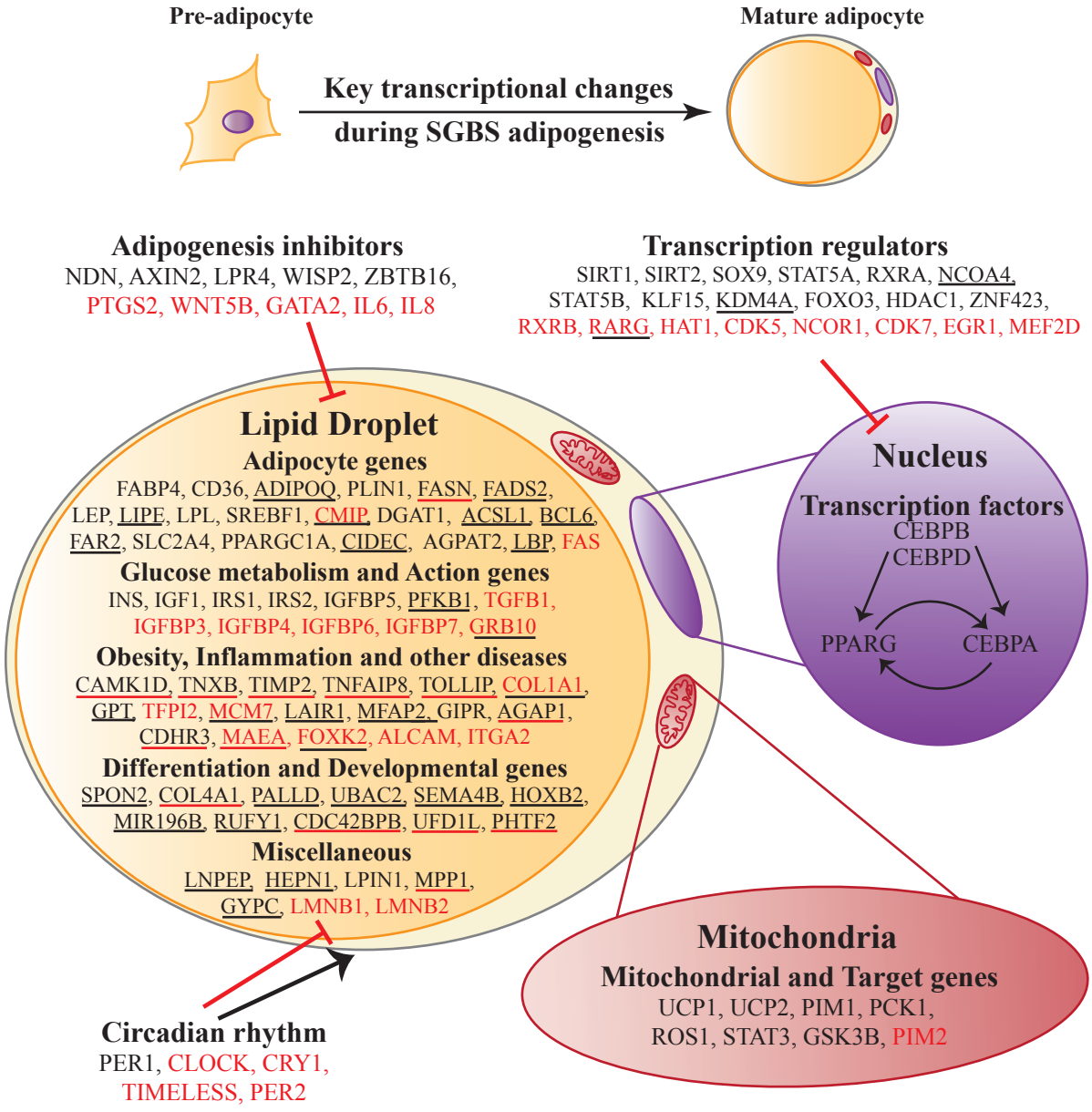


Figure 8B: Key transcriptional and DNA methylation changes between HG and LG1 treated adipocytes at D14. Differential genes with increased and reduced expression levels from D0 to D14 are shown in black and red respectively. Genes without a significant expression change is shown in gray. Underlines indicate presence of DMRs with a beta value difference more than 0.05 in the same comparison (promoter, gene body).

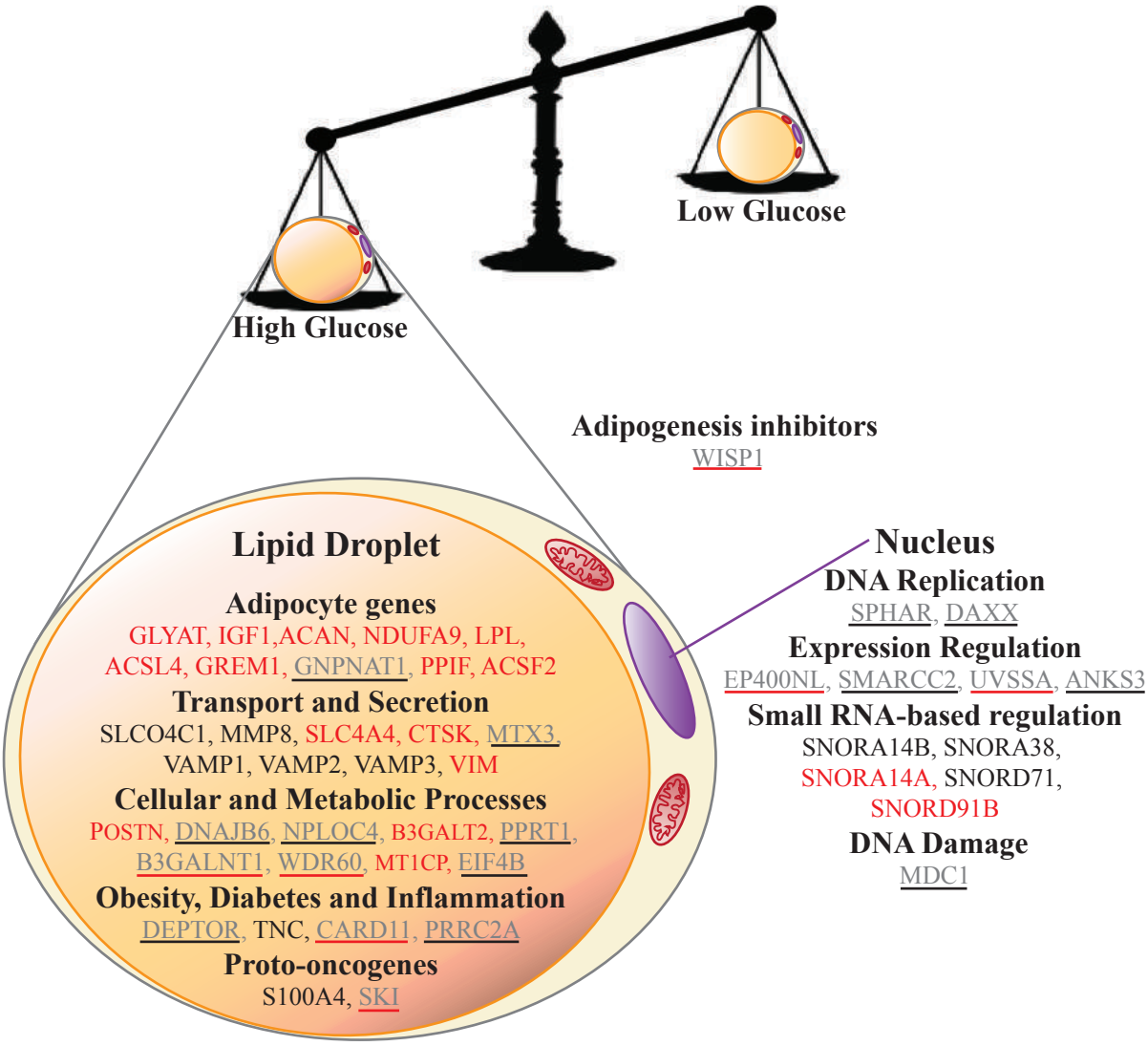


Figure 9: Two-dimensional principal component analysis of 450k methylation profiles of human visceral adipocytes, visceral adipose tissue, visceral stroma vascular fraction (SVF), cultured visceral primary cells, subcutaneous adipocytes, subcutaneous SVF, cultured subcutaneous primary cells, SGBS cells, IMR-90 cells and MRC-5 cells (n=54, # of probes=422,184, 62.24% variance is explained).

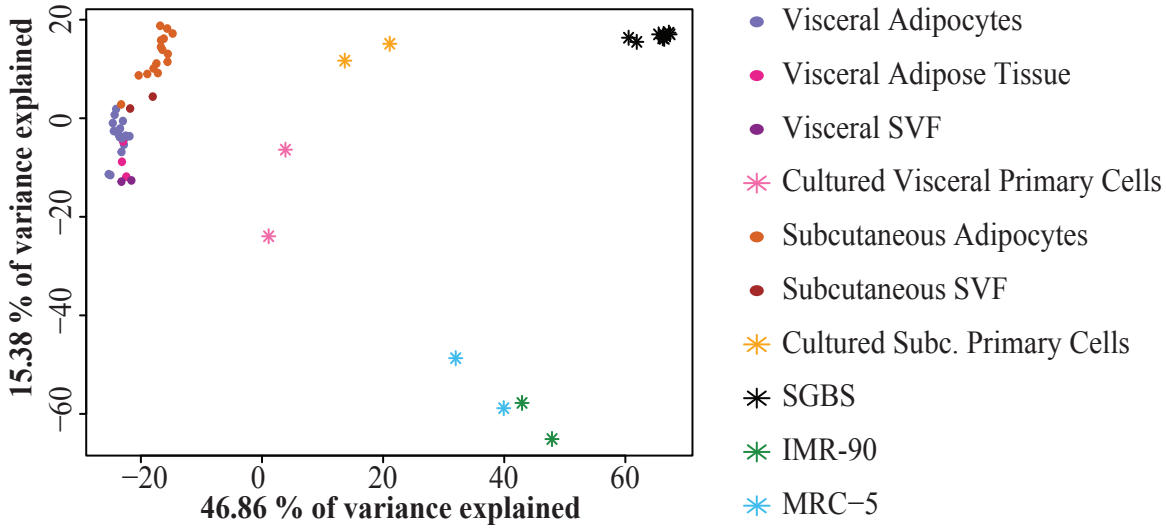


Figure 10: Heat map illustration of samples and selected DE genes. Selected genes (N = 78) differentially expressed (high expression: shades of yellow and low expression: shades of red) in SGBS cells (4 samples of D0-PA and 3 samples of D10-HG, D14-LG1 and D14-HG) and human uncultured primary purified visceral and subcutaneous adipocytes (isolated from 3 lean patients shown as VA, SA respectively).

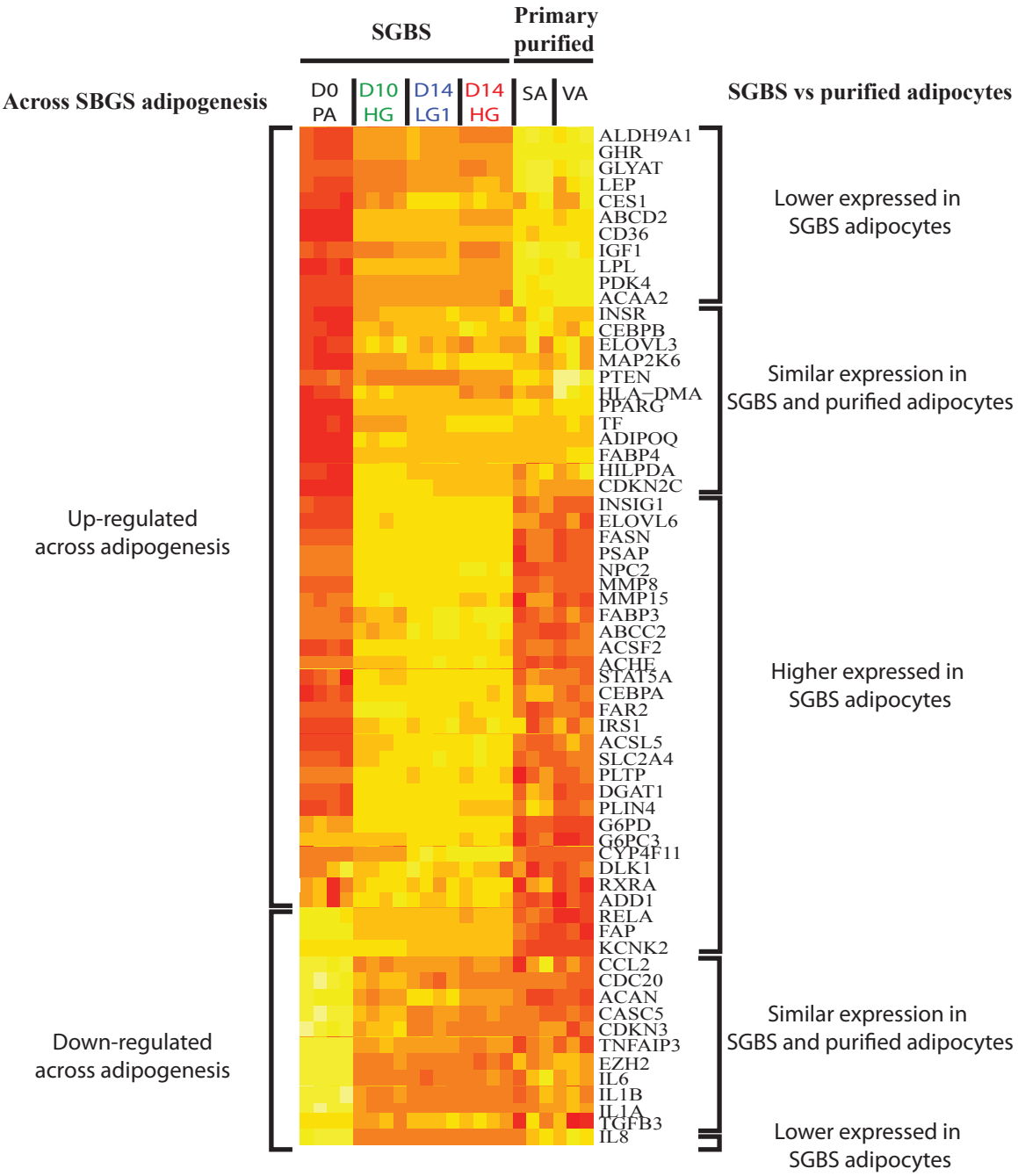


Figure S1A: Methylation profiles across DM regions between Day 0 pre-adipocytes and Day 14 mature adipocytes treated in HG or LG1 conditions. ● D0 PA ● D14 HG ● D14 LG1

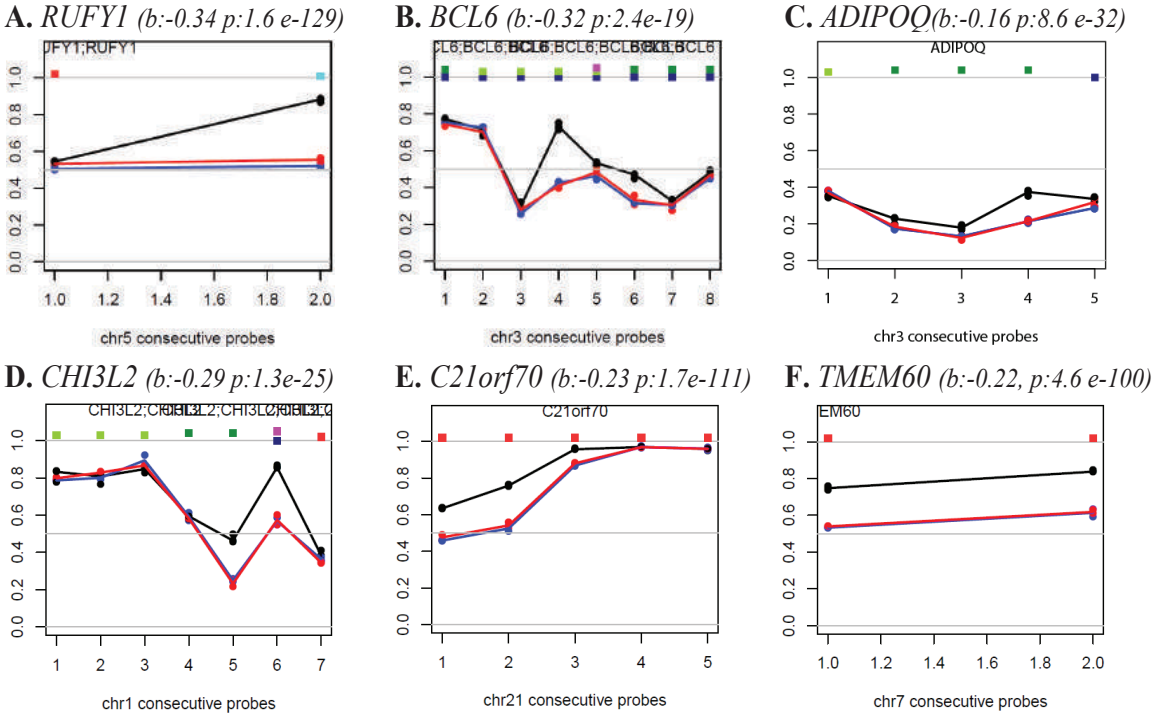
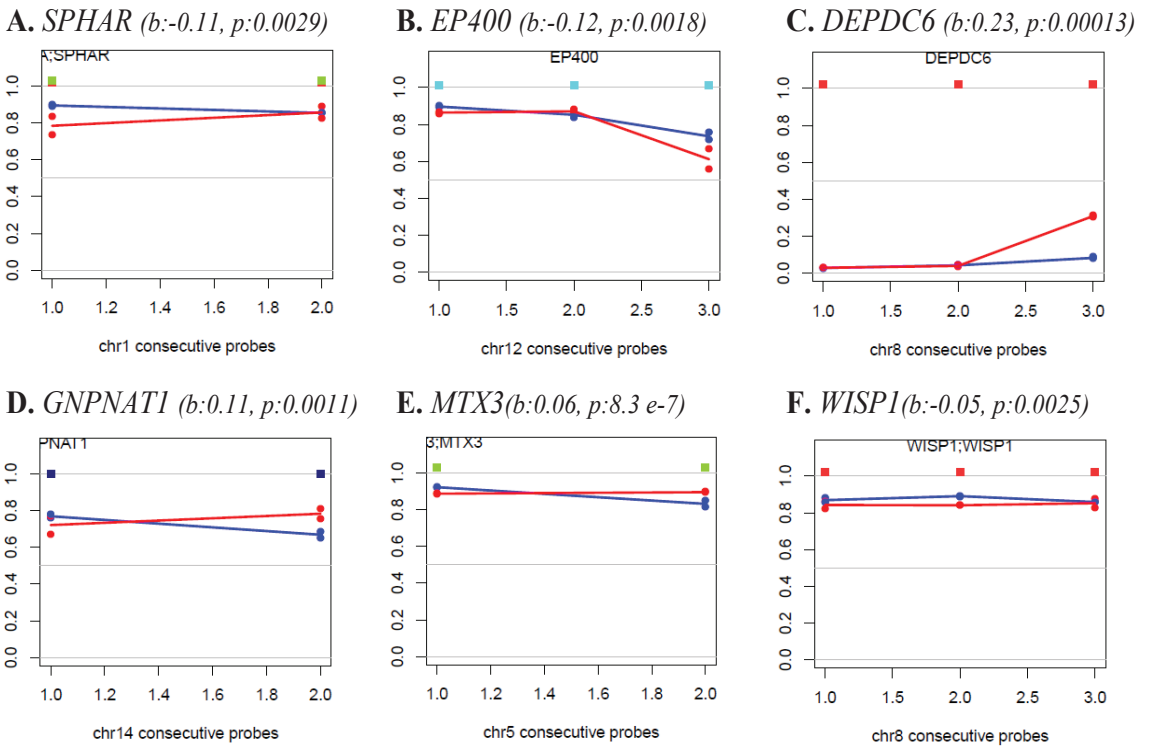


Figure S1B: Methylation profiles across DM regions between Day 14 mature adipocytes treated in HG vs LG1 conditions.

● D0 PA ● D14 HG ● D14 LG1



CHAPTER 2: LipiD-QuanT: a novel method to quantify lipid accumulation in live cells

In humans, increasing adiposity is directly proportional to increase in fat mass, which may lead to an increase in total number of fat cells and/or their lipid content. Assesment of lipid accumulation has been the most common practice to morphologically distinguish lipid storing fat cells from their fibroblast-like precursors *in vitro*. The most widely used lipid accumulation techniques are based on enzymatic digests or lipophilic staining, and they are either not quantitative or can be destructive to samples.

This chapter describes a novel label-free lipid droplet quantification technique (LipiD-QuanT) to monitor fat accumulation in live adipocytes. LipiD-QuanT is robust, time and cost effective compared to other triglyceride accumulation assays and it is the only tool available to study lipid droplet dynamics in live cells without the use of highly specialized equipment. We optimized the LipiD-QuanT method and validated the test using SBGS cells *in vitro*. In a LipiD-QuanT validation study to measure the effect of four potential pro- or anti- obesogenic substances, we found that Docosahexaenoic acid (DHA), D-glucose and Zinc oxide (ZnO) nanoparticles did not induce changes in the size of lipid droplets stored by adipocytes but prolonged exposure to rosiglitazone prevented the enlargement of lipid droplets.

This publication presented in this chapter is the product of a working collaboration between myself, Osmond-McLeod, Molloy and Vallotton. I was predominantly involved in the design of the study, interpretation of the data and drafting the manuscript. Vallotton was involved in the design of the software. Osmond-McLeod was involved in quality testing of the nanoparticles and provided reagents. Molloy was involved in the design of the study, interpretation of the data and the drafting of the manuscript.

This manuscript has been published in *Journal of Lipid Research,* November 2015.

Detailed contributions:

	Varinli H	**Co-authors**
Experimental Design	60%	40%
Laboratory Experiments	100%	
Software Design		100%
Data Analysis	40%	60%
Writing	90%	10%

Varinli, H., Osmond-McLeod, M. J., Molloy, P. L. & Vallotton, P. (2015). LipiD-QuanT: A novel method to quantify lipid accumulation in live cells. Journal of Lipid Research, 56, p. 2206-2216.

DOI: 10.1194/jlr.D059758

JOURNAL OF LIPID RESEARCH

LipiD-QuanT: A Novel Method to Quantify Lipid Accumulation in Live Cells

Hilal Varinli[1, 2, 3]*, Megan J Osmond-McLeod[1, 4], Peter L. Molloy[1] and Pascal Vallotton[5]

[1] CSIRO Food and Nutrition Flagship, North Ryde, New South Wales, Australia

[2] Department of Biological Sciences, Macquarie University, North Ryde, New South Wales, Australia

[3] Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia

[4] CSIRO Advanced Materials TCP (Nanosafety), North Ryde, New South Wales, Australia

[5] CSIRO Digital Productivity Flagship, North Ryde, New South Wales, Australia

* To whom correspondence should be addressed:

Hilal Varinli

CSIRO Food and Nutrition Flagship

Riverside Corporate Park, 11 Julius Avenue,

North Ryde NSW 2113, Australia

Phone: +61 2 94905136; Email: hilal.varinli@csiro.au

*Running title: LipiD-QuanT: Lipid Droplet Quantification Tool in Live Cells

Abbreviations: Lipid Droplet (LD), 3,3'-dioctadecyloxacarbocyanine perchlorate (DiO), Coherent anti-Stokes Raman Scattering (CARS), 1-methyl-3-isobutylxanthine (IBMX).

ABSTRACT

Lipid droplets (LD) are the main storage organelles for triglycerides. Elucidation of lipid accumulation mechanisms and metabolism are essential to understand obesity and associated diseases. Adipogenesis has been well-studied in murine 3T3-L1 and human SGBS pre-adipocyte cell lines. However, most techniques for measuring LD accumulation are either not quantitative or can be destructive to samples.

Here, we describe a novel label-free LD quantification technique (LipiD-QuanT) to monitor lipid dynamics, based on automated image analysis of phase contrast microscopy images acquired during *in vitro* human adipogenesis. We have applied LipiD-QuanT to measure LD accumulation during differentiation of SGBS cells. We demonstrate that LipiD-QuanT is a robust, non-destructive, time and cost effective method compared to other triglyceride accumulation assays based on enzymatic digest or lipophilic staining. Further we applied LipiD-QuanT to measure the effect of four potential pro- or anti-obesogenic substances: DHA, rosiglitazone, elevated levels of D-glucose, and zinc oxide nanoparticles. Our results revealed that 2 µmol/L rosiglitazone treatment during adipogenesis reduced lipid production and caused a negative shift in LD diameter size distribution, but the other treatments showed no effect under the conditions used here.

Keywords: adipocytes; obesity; lipid droplet; triglycerides; omega-3-fatty acids; glucose; SGBS; label-free image analysis; zinc oxide nanoparticles; rosiglitazone.

INTRODUCTION

Lipid droplets (LDs) are the main storage organelles for triglycerides in eukaryotic cells (1-3). They are found in several cell types, where they support membrane biosynthesis and aid lipid homeostasis. They also participate in the storage of toxic lipid species in hepatocytes, macrophages, cardiac myocytes, renal glomerular cells muscle and buccal cells (4, 5). Adipocytes are the main cell type that store the body's energy reserves as triglycerides, and they do so in larger LDs that serve as highly specialized lipid reservoirs. However, excessive storage of triglycerides leads to obesity and metabolic syndrome, posing serious risks of diet-related non-communicable diseases such as diabetes mellitus, hypertension, cardiovascular disease and stroke (6).

The use of animal model systems and comparison with human studies has contributed extensively to our understanding of processes of development of adipose tissue and obesity (reviewed (7)). Supplementing *in vivo* systems, i*n vitro* models are invaluable for dissecting molecular mechanisms, including the formation of LDs (8). Adipogenesis processes have been well-studied since the establishment of both murine (9) and human (10) pre-adipocyte cell lines. These models enable controlled investigation of the regulators of adipogenesis and also provide mechanistic insight into obesity via the measurement of physiological and molecular responses of adipocytes to specific therapeutic compounds. Since the establishment of human pre-adipocyte SGBS cell line in 2001 (10), it has been used in more than 100 research articles and proven to be a reproducible model to biological mechanisms in pre-adipocytes and mature adipocytes in context of human adiposity (11). SGBS cells, are neither transformed nor immortalized, have high proliferation and differentiation capacity and provides unique advantages to study human adipogenesis (12). SGBS cells are easier to maintain and less costly than human primary cell lines.

During adipogenesis, small LDs destabilize to increase storage capacity in two ways: coalescence by fusion of LDs or by Ostwald ripening, in which molecules from one LD diffuse to another (13). Quantification of LD accumulation is the most common measure of differentiation during *in vitro* adipogenesis. The three most widely used techniques are:

Lipid staining coupled with Optical Reading: Fixed cells are stained with lipophilic dyes, the most commonly used lipophilic dyes being Oil Red O, Sudan III, Nile Red, BODIPY 493/503 and 3,3'-dioctadecyloxacarbocyanine perchlorate (DiO) (14-16). The absorbed dye can be resolubilized and the intensity of the stain, measured via spectrophotometry, can be used as a proxy for cellular lipid content.

Biochemical Analysis: Following cell lysis, lipolytic enzymes are used to digest triglycerides to glycerol and free fatty acids. Additional enzymatic reactions turn glycerol into a colored product, the concentration of which may be measured using spectrophotometry.

Fluorescence microscopy: Monoclonal antibodies are available against LD-associated proteins such as perilipin, adipophilin, TIP47 and caveolin-1 (17). Fluorescence microscopy after antibody staining makes LDs visible, enabling measurement of LD size and ultimately lipid content.

These techniques are lengthy, require dedicated culture plates to measure the lipid content and do not permit reuse of cells for other biological or morphological measurements. Moreover, the staining-based methods comprise multiple steps including fixation, washing and dehydration, which occasionally result in detachment of cells and lysis of LDs. There is also growing evidence that staining conditions such as dye concentration and fixation time affect the fluorescence intensity of lipophilic dyes resulting in incorrect correlation between fluorescence intensity and actual total lipid content (18). Additionally, lipophilic dyes have high affinity towards hydrophobic surfaces which results in background staining when plastic culture dishes are used (19, 20). Finally, when using biochemical assays, a variable fraction of the lysed cells may be retained in the cell culture dishes, which may result in minor errors in the cellular lipid content calculation. Thus, existing assays tend to be more qualitative than quantitative.

More recently, Coherent anti-Stokes Raman Scattering (CARS) microscopy showed great promise as a technique to study LD dynamics (21). It has been used to identify LDs in *Caenorhabditis elegans* (22), *Drosophila melanogaster* (23), murine fibroblasts (24) and human adipose-derived stem cells (25, 26). However, it is currently only available in a handful of

research institutes with the appropriate equipment. Therefore a need still exists for a method of lipid accumulation quantification that is effective and reliable, suitable for use with live cells and uses common laboratory equipment.

In this contribution, we describe LipiD-QuanT, an automated image analysis tool, to quantify LD accumulation in live cells under phase contrast microscopy. The LipiD-QuanT algorithm depends on the ability of phase contrast microscopy to reveal even small refractive index differences between LDs and the surrounding cytosol by exploiting destructive interference effects. LDs have a well-defined spherical shape hence the measurements of LD volume and surface area may be easily obtained on a per-cell basis. Therefore, we utilized LipiD-QuanT on the human SGBS pre-adipocyte differentiation model which provides the ease to study the effect of any treatment of interest on human adipogenic differentiation in short time-frame.

We first evaluated the performance of LipiD-QuanT against benchmark Oil Red O staining and biochemical lipid accumulation techniques to monitor LD dynamics during *in vitro* differentiation of human SGBS pre-adipocytes over a 21 day timeframe. We monitored the gene expression changes in selected adipogenesis genes to assure the reproducibility of SGBS pre-adipocyte differentiation. We then utilized LipiD-QuanT to sensitively detect changes in LD growth in response to four potential pro- or anti-obesogenic treatments: Docosahexaenoic acid (DHA), high glucose, zinc oxide (ZnO) nanoparticles and rosiglitazone.

MATERIALS AND METHODS

Cell culture conditions

We used the SGBS cell line, a human-derived pre-adipocyte cell line isolated from the stromal cell fraction of subcutaneous adipose tissue from an infant with Simpson Behmel Gobali syndrome (10). Proliferation and differentiation media were prepared as previously described (10), with minor changes. In brief, cells were proliferated to 90% confluence in T75 or T150 flasks in DMEM/F-12, GlutaMAX™ (Life Technologies, cat#10565) supplemented with 10% FBS (Invitrogen, cat#10099-141), 10 U/ml penicillin-streptomycin (Invitrogen, cat#15070), 8 mg/L d-biotin (Sigma, cat#4639), and 4 mg/L d-panthothenic acid (Sigma, cat#P5155). Cells were then differentiated in 6-well plates at a concentration of 0.2 million cells/well. The experiments were undertaken with a minimum of three biological replicates at passage number 23. The cells were kept in serum free quick differentiation media, supplemented with 100 nmol/L cortisol (Sigma, cat#H0888), 0.01 mg/ml transferrin (Sigma, cat#T0665), 0.2 nmol/L triiodothyronine (Sigma, cat#T6397), 20 nmol/L insulin (Sigma, cat#I2643), 2 µmol/L rosiglitazone (Sigma, cat#2408), 25 nmol/L dexamethasone (Sigma, cat#D4902) and 0.5 mmol/L 1-methyl-3-isobutylxanthine (IBMX) (Sigma, cat#I5879) for the first four days. After four days, media was removed and replaced with differentiation media, further excluding rosiglitazone, dexamethasone and IBMX for 10 or more days. Media was changed every second day.

During adipogenesis, cells were exposed to the following pro- and anti-obesogenic treatments:

1- *DHA:* A single ml of DHA oil emulsion contained 125 mg of DHA (C22:6), 8.5 mg of eicosapentaenoic acid (C20:5, EPA), 9 mg of vitamin C and 0.19 mg of vitamin E (HiDHA™ oil emulsion: Nu-Mega Ingredients Pty. Ltd.). Across the 14 day course of SGBS differentiation 10 µmol/L DHA oil emulsion was added, at each change of media, a total of 7 times. LDs were monitored at the end of the differentiation period on Day 14.

2- *D-glucose:* SGBS cells were differentiated in 10 mM D-glucose (Sigma, cat#G7021), 10 mM D-glucose with 7.5 mM Sorbitol (Sigma, cat#S1876) for osmolarity control and 17.5 mM D-glucose. The low glucose concentrations were maintained by changing the medium to DMEM, low glucose, GlutaMAX™, pyruvate (Life Technologies, cat#11885) containing 5.5 mM D-glucose. The LDs were assessed at Day 7, 10 and 14.

3- *ZnO nanoparticles:* The differentiated SGBS cells were treated with 1 µg/mL or 10 µg/mL ZnO nanoparticles coated with a dimethoxydiphenylsilane/triethoxycaprylylsilane crosspolymer (Z-COTE MAX from BASF, batch# FCHE1301) on Day 6 for 48 hours. We have described the extensive physicochemical characterization of Z-COTE MAX elsewhere (27, 28); briefly, primary particle sizes were 36 ± 2 nm wide and 95 ± 5 nm length but formed larger aggregates in water and cell culture medium. Non-cytotoxic concentrations to SGBS cells were selected on the basis of previous work (27, 28). We monitored the change in LD size distribution immediately after the treatment period at Day 8, as well Days 10 and 14.

4- *Rosiglitazone treatment:* SGBS cells received an additional of 2 µmol/L rosiglitazone from Day 4. The LDs were measured on Day 10 and 14.

Biochemical measurement of lipid content with triglyceride accumulation assay

SGBS cells were lysed using 200 µl Triglyceride Accumulation Lysis Buffer per well in 6-well plates (Zen-Bio Inc., cat#TG-1-NC). Biological triplicates were included for each assay point and the lysates were stored at -80°C until the completion of the differentiation process. Triglyceride esters were converted to glycerol and glycerol concentration was quantitated enzymatically and measured as per the manufacturer's protocol, using 15 µL of lysate in technical duplicates.

Staining based measurement of lipid content using Oil Red O

Cells were fixed with 1 ml of 4% paraformaldehyde for 20 min in 6-well plates, washed with 1X PBS twice and stored at 4ºC in 1X PBS supplemented with 0.02% (w/v) Sodium azide until processing. The fixed cells were washed with 100% propylene glycol (Astral, cat#CSPL010) prior to 1 ml of Oil Red O staining for 15 min at room temperature (Sigma, cat#O1516). Cells were washed 3X with 1 ml of PBS to remove excess dye, and images were acquired using an inverted microscope (Nikon ECLIPSE 90i fitted with a high definition, cooled colour digital camera DXM1200C (Nikon, Japan). The Oil Red O dye was extracted from the stained cells using 500 µl of 100% isopropanol for each well and 2 aliquots of 200 µl transferred to black 96-well plates. The O.D. reading of Oil Red O was measured at 520 nm using a plate reader (POLARstar Omega microplate reader, BMG Labtech, UK).

Immunolabelling for LDs

Cells were differentiated on acid-treated coverslips for immunolabelling of the LD coating protein perilipin. The cells were fixed on coverslips with 1 ml of 4% paraformaldehyde for 20 min in 6-well plates, washed with 1X PBS twice before and after permeabilization in Triton X-100 solution (0.2 % Triton X-100 in PBS with 10% Goat serum (Cell Signaling Technology, cat#5425) for 10 min. Samples were blocked for 20 min in 10% Goat serum, 5% FCS and 0.5% BSA in PBS and stained with anti-Perilipin (D1D8) XP® Rabbit mAb (Cell Signaling Technology, cat#9349S) overnight at 4°C as per the manufacturer's recommendations. Samples were washed with 1X PBS twice and re-blocked. The secondary antibody staining was completed using Alexa Fluor® 488 Goat Anti-Rabbit IgG (H+L) as per the manufacturer's instructions (Invitrogen, cat#A-11034). A solution of 25 µg/ml DAPI (Sigma, cat#D9542) was added for nuclear staining. The coverslips were washed twice with 1X PBS and mounted on microscope slides using Fluoroshield mounting media (Sigma, cat#F6182).

RNA extraction

SGBS cells were lysed in 500 µl TRI-reagent (Sigma, cat#T9424) on Days 0, 4, 7, 14 and 21 and the lysates were stored at -80 °C until completion of sample collection. Total RNA was isolated as described by the manufacturer's manual. RNA concentration was determined by NanoDrop ND-1000 (USA) Spectrophotometer readings.

cDNA synthesis and quantitative Real-Time PCR

Each sample of first-strand cDNA was synthesized from 600 ng of total RNA using QuantiTect Reverse Transcription Kit (Qiagen, cat#205313). Gene-specific primers (Supplementary Table 1) were used to amplify target genes using 10 ng of first-strand cDNA as template in a 15 µl SYBR-green based quantitative RT-PCR reaction performed under the following conditions: 95ºC for 2 min, 45 cycles at 95ºC for 10 sec, 60ºC for 30 sec, 72ºC for 10 sec with a melting curve from 65 ºC to 95 ºC. The gene expression levels were normalized to *GAPDH* housekeeping gene expression in each sample.

Image acquisition

Images were acquired on an Olympus IX81 microscope equipped with a 20X/0.40 LCPLANFl Ph1 phase contrast objective (USA) and a Roper Scientific CoolSNAP FX monochrome camera (USA), aligned for positive mode phase

contrast microscopy. The image intensity was optimized to span the full camera dynamic range while the focus was adjusted to maximize the LD morphology displaying sharp edges. Care should be taken in the microscope set-up to ensure that any "halo" effect is homogeneous, as inhomogeneity may interfere with the image analysis. Also, without care in set-up, the smallest LDs may have dark boundaries that do not close entirely. A total of 6 images were taken per well. All the images were saved in uncompressed TIFF format. Pixel size in the object space was 0.3 µm. A comprehensive set of label-free images of human adipocytes were captured during adipogenesis (Supplementary File 1).

On average there were 25 cells in a single image with a variation of 2-5 cells between image fields. However, larger variation will lead to incorrect assumptions about total lipid quantity. Therefore, lipid quantities should preferably be normalized by the cell number. The cells were stained with the nuclear fluorescent dye DAPI, and imaged with a DAPI fluorescence filter set immediately before acquiring the phase contrast image (example images are provided in Figure 1A, 1B and 1C). Nuclei were counted using the Otsu threshold method available in Matlab™ (29). LD amounts per cell were calculated by matching nuclei numbers with the LD numbers produced by LipiD-QuanT (Figure 1). Although LipiD-QuanT is designed for LD quantification in living cells to allow their further use for other biological measurements, we used images of fixed cells to check the effect of specific stimulants which were utilized to compare our method with Oil Red O staining method on the same wells. More physiological nuclear stains may also be used such as Hoechst 33342 (available to stain live cells in media as NucBlue® by Life Technologies).

LipiD-QuanT image processing and analysis

LDs naturally adopt a spherical geometry -driven by their surface tension (30). LipiD-QuanT is based on the Laplacian edge detector, which we have previously employed to monitor bacteria under phase contrast microscopy (31). The Laplacian edge detector defines level sets (similar to level lines on topography maps), that automatically guarantees the formation of closed contours surrounding the target shapes (32). LipiD-QuanT detected contour lines closely corresponding to individual LDs in human adipocyte images. However, false positive contours are also formed in the background. These were eliminated on the basis that they are associated with low intensity variance along their contour.

Cytoplasmic spaces, surrounded by genuine LDs, tend to have a concave shape –being complementary to a set of limiting convex vesicles. We used the function TRACE_MooreNeighbourhood.m contributed by Adam H. Aitkenhead available on "Matlab Central" to trace the LD edges until it generated a closed contour circle (33). We then applied the LineCurvature2D.m function contributed by Dirk-Jan Kroon and available on "Matlab Central" to compute the curvature of the traces generated by the previous function (34).

Analysis of the curvature at the outer boundary of the segmented regions and subsequent removal of regions with negative average curvature allowed us to eliminate the false positive LDs (see Figure 1E, red arrow). LDs often occur as clusters. Occasionally, fused LDs form dumbbell shapes. We have incorporated the watershed transform algorithm to LipiD-QuanT to detect and divide dumbbell shaped LDs (35) and eliminated LDs external to cells. Additionally, a mask outlining the position of cells was constructed in order to avoid counting LDs in the background (36). The key steps of the LipiD-QuanT algorithm are presented in Figure 2.

LipiD-QuanT data analysis and output

LipiD-QuanT installer (Supplementary File 3) is available on the CSIRO Data Access Portal as well as on GitHub. The details are further explained in the instalment instructions included to Supplementary File 3. It will automatically process the images saved to the same directory, providing numeric outputs corresponding to LD diameter length in pixels as described in the LipiD-QuanT installer instructions. The script itself is provided in Supplementary File 4. LipiD-QuanT is fully automated and can be used even by potential users that are not familiar with Matlab™.

The result images produced by LipiD-QuanT correspond to binary images, where all the pixels of each identified LD are set to 1 and the background pixels are set to 0. We generated the size and shape information pertaining to LDs using the regionprops function in Matlab. LipiD-QuanT is applicable to both fixed and live cells because it uses "equivalent diameter" indicating the diameter of the disk that has the same area as the object as the absolute measure of LD diameter length in pixels. Based on the microscope set up, the measurements can be converted into µm.

We used the smoothed kernel density distributions to visualize the distribution of LD sizes among treatments and characterize potential effects. The complete script for producing smooth LD diameter size distributions from LipiD-

JOURNAL OF LIPID RESEARCH ASBMB

QuanT output is given in Supplementary File 5. The LipiD-QuanT output can also be used to calculate total LD area as an indicative of LD content. We advise to normalize the total LD area for number of cells to remove any technical or biological variance due to cell number among wells.

RESULTS

SGBS pre-adipocyte differentiation

SGBS pre-adipocyte cells were cultured and differentiated according to standard procedures (10). Undifferentiated SGBS cells exhibit a fibroblast-cell like structure. However, the cells initiate adipogenesis by pulling their longitudinal structures into more localized single globular structure with the influence of the adipocyte differentiation cocktail. By Day 4, tiny LDs, sparsely distributed within the cell, became more apparent. By Day 7, a high proportion of the cells appeared terminally differentiated containing multiple visible LDs. By Day 10, the majority of the differentiated cells matured and contained large quantities of lipids. Over the 14-day course of adipogenesis, 95% of the SGBS pre-adipocytes were differentiated into mature adipocytes. From Day 14 to 21, we observed LD enlargement by coalescence and growth.

We monitored changes in the cellular morphology and expression profiles of a set of key genes such as *CEBPB*, *PPARG*, *GLUT4*, *ADIPOQ*, *PLIN* and *FABP4* during adipogenesis. All of these genes except *ADIPOQ* had the highest relative expression on Day 10 (Supplementary Figure 1). While the expression of *CEBPB*, *PPARG* and *GLUT4* genes were higher in later stages of adipogenesis, *ADIPOQ*, *PLIN* and *FABP4* expression levels were reduced by at least 2-fold (Supplementary Figure 1). The gene expression changes were consistent with previously published data confirming that SGBS pre-adipocyte differentiation is a reproducible model for the study of human adipogenesis (12, 37-40).

Lipid-QuanT validation

i. **LipiD-QuanT algorithm validation against manually segmented LDs**

In Figure 1 we show the process of applying LipiD-QuanT. The first panel shows a phase contrast image of a field of differentiating adipocytes at Day 7, while panels B and C show the DAPI-stained image of the field and nuclei identification used to determine cell numbers per field. We present an example of a phase contrast image of a differentiated adipocyte in Figure 1D, and the processed LipiD-QuanT automated segmentation of LDs in Figure 1E. The shapes and sizes LDs measured automatically via LipiD-QuanT closely matched that in original images (compare Figure 1D and 1E).

We generated ground truth data by manually creating LD outlines in randomly selected 2 cells in each of a total of 6 test images taken on Day 10 of human adipogenesis, using the region of interest selection tool in Matlab™ (imellipse), which allows precisely overlaying an ellipse shape onto each individual LD, by adjusting its centre and main axes. An example image where these ellipses were fitted manually is shown in Figure 1F. The same LDs were also detected by executing LipiD-QuanT on the same 6 images. LipiD-QuanT detected over 95% of LDs as determined by manual segmentation (Figure 1E and 1F). In total 1,048 LDs were analyzed by the two methods (Figure 1G). While there was a minor shift of larger sizes in LD diameter distribution using manual segmentation, the LD size distributions obtained using the two methods were not statistically different according to Chi-Square (X-square) and Kolmogorov-Smirnov Tests (KS.test) ($p > 0.05$) (Figure 1G). The LipiD-QuanT algorithm requires an average of 1 min per image and detects LDs with a minimum diameter of 0.34 μm.

i. Anti-perilipin stained LDs structures corresponded accurately with stain-free LipiD-QuanT detection

In order to confirm correspondence of LDs detected by LipiD-QuanT with fluorescence stained images, we stained mature adipocytes using an antibody against perilipin – a protein known to decorate the outer surface of LDs (Figure 1H). The appearance and size of the fluorescently stained LDs corresponded to those detected under phase contrast. We also acquired 3D widefield stacks of anti-perilipin stained human SGBS adipocytes. LDs tended to be approximately spherical and predominantly arranged in a single horizontal plane (Supplementary File 2).

ii. LipiD-QuanT robustness against small variations in focal plane selection

LipiD-QuanT, developed for *in vitro* studies, relies on the acquisition of a single image for lipid quantification, rather than a full 3D stack of images spanning the entire 3 dimensional cell volume. We tested LipiD-QuanT's robustness against variations in focal plane selection by acquiring the same image field 5 times, each time independently refocusing to maximize the contrast of LD's. The resulting LD size distributions corresponding to individual images were almost indistinguishable (Figure 1I). Thus, as long as image contrast is maximized by the operator, LipiD-QuanT is resilient against focus variations. The choice of focal plane does not interfere with the results because adipocytes are predominantly arranged flat in horizontal plane as shown in Supplementary File 5.

iii. LipiD-QuanT consistency across image fields in the same well

While a certain biological variability in lipid content is expected for individual cells in the same well, cell specific variability is expected to be averaged out as images often contained multiple cells. We acquired 6 images from various locations of the same well in a total of 2 treatment groups. The 12 images were analyzed by LipiD-QuanT separately. The separate LD diameter size distribution was produced for each image taken from the same well. The results were highly consistent and robust in each treatment group (Supplementary Figure 2A and 2B).

Comparison of LipiD-QuanT with standard methods

We monitored lipid accumulation in differentiating SGBS cells on Days 0, 4, 7, 10, 14 and 21 via biochemical and lipid staining methods (Figure 3). We utilized LipiD-QuanT on images acquired from Day 7 and onwards as the LDs were too small to identify with a 20X objective before Day 7. A total of 5,670 LDs were characterized to assess the performance of LipiD-QuanT against staining and biochemical based triglyceride accumulation assays.

Over the process of well-orchestrated adipogenesis, some pre-adipocytes trigger differentiation more quickly, rapidly reaching their maximum LD storage capacity whereas others contain a plethora of differentially sized LDs. Similarly both live and Oil Red O stained adipocyte images confirmed the structural changes in LDs as adipogenesis progressed (Figure 3A). New LDs form while larger LDs merge and augment in size to increase lipid storage capacity until Day 21 (i.e. there are fewer but larger LDs as the cells mature) (Figure 3B). Figure 3A and 3B provide representative sections of images, more examples of label-free human adipocytes images during differentiation are provided in Supplementary File 1.

The overall trend for LD accumulation obtained via the Oil Red O staining method differed from both LipiD-QuanT and triglyceride accumulation results (Figure 3C, 3D and 3E). The biochemical assay demonstrated linear increase in lipid content (Figure 3D) consistent with the previously published results (41). LipiD-QuanT measurements demonstrated a sustained increase in lipid content across the entire differentiation time course (Figure 3E) while Oil Red O staining surprisingly indicated that total lipid content decreased after Day 10 (Figure 3C), possibly due to lysis or detachment during washing and staining procedures. Because LipiD-QuanT measurements are based on individual cells it is possible to discern features that are obscured in measurements of total lipid content.

The LD size distribution obtained by LipiD-QuanT reflected LD coalescence and ripening as the mean of LD diameter size distributions shifted towards larger LDs as the adipogenesis progressed (Figure 3F). The total LD area increased 10 times while the average LD diameter size increased 118% from Day 7 to Day 21 (Figure 3E and 3F). Median LD diameter size of SGBS adipocytes increased from 1.72 µm to 2.95 µm from Day 7 to 21. As the size range of LDs is similar to that of SGBS cells (42, 43). LipiD-QuanT coupled with nuclear staining also allowed us to calculate the ratio of differentiated cells during adipogenesis. The proportion of differentiated cells increased from 75% to 96% from Day 7 to Day 21.

LD size distribution in the presence of pro- and anti-obesogenic substances

Once we confirmed that LipiD-QuanT was accurately detecting LDs compared to the standard methods, we invoked LipiD-QuanT to measure the response to four potential pro- or anti-obesogenic interventions during adipogenesis:

1- *DHA:* An important omega-3 long-chain polyunsaturated fatty acid, reported to have anti-obesogenic effects in animal models (44); reduced plasma levels of omega-3 fatty acids have also been linked to obesity (45-47). Higher omega-3 fatty acid levels has been shown to reduce fat cell size in overweight and obese individuals (48).

2- *D-glucose:* Glucose has been shown to act through SREBP-1c to induce de novo lipogenesis *in vitro* in rat muscle satellite cells (49). Exposure of 3T3-L1 cells to high glucose conditions leads to insulin resistance (50, 51) and to accumulation of significantly increased amounts of lipid compared to low glucose (52).

3- *ZnO nanoparticles:* zinc homeostasis is involved in dysfunction of insulin metabolism (53). Low levels of erythrocytary zinc are linked to type-2 diabetes and metabolic syndrome outcomes (54-56). Since the therapeutic effects of ZnO nanoparticles in treating lipid associated diseases show promise (57, 58), but are under-explored, we tested their possible effect on LD dynamics in differentiating human pre-adipocytes.

4- *Rosiglitazone treatment:* The peroxisome proliferator-activated receptor gamma (PPARG) agonist is an insulin sensitizer and triggers new LD formation during adipogenesis (59). However, prolonged treatment is known to be anti-obesogenic, promote lipolysis and block LD fusion resulting adipocytes containing small LDs (60, 61).

In Figure 4, we show the smoothed distributions of LD sizes and total lipid quantification obtained using LipD-QuanT and Oil Red O staining methods respectively with representative images of adipocytes corresponding to each treatment and control group. The distributions of LD diameter results obtained from independent experiments indicated that measurement of LD accumulation during SGBS adipogenesis was very robust (Figure 4A, 4D, 4G and 4J). Chi-Square and Kolmogorov-Smirnov Tests indicated that DHA and ZnO nanoparticle treatments had neither positive nor negative effects on overall distribution of LD diameter size during adipogenesis (p-values > 0.05) (Figure 4A and 4G). While elevated glucose concentration also did not result in a statistically significant change in LD size distribution or total lipid content, a consistent slight shift to higher LD was seen (Figure 4D). The total lipid quantification using Oil Red O staining method was also unaffected in these treatments (t-test, p-values >0.05) (Figure 4B, 4E and 4H).

For example, under the effect of ZnO nanoparticles exposure (Figure 4G and 4H), the tail of LD diameter size distributions fattens indicating the increase in the proportion of larger LDs in adipocytes as the adipogenesis progress (Figure 3F, 4D, 4G and 4J). However these structural features are lost in measures of total lipid quantification is used (Figure 4E, 4H and 4K).

Rosiglitazone treatment is essential for the early stages of adipogenesis to promote LD formation and fusion however, prolonged stimulation is known to influence adipocyte morphology; leading to adipocytes containing small LDs scattered in the cytoplasm. We observed a statistically significant effect of rosiglitazone in LD diameter size distribution at both Day 10 (X-square p-value<2.2e-16 and KS.test p-value<2.2e-16) and Day 14 adipocytes (X-square p-value<8.62e-15 and KS.test p-value<2.2e-16). There was a negative shift in LD size distributions revealing an abundance of small LDs on both Day 10 and Day 14 (Figure 4J). The total lipid quantification using Oil Red O staining method also revealed that there was a significant reduction in the lipid content both in Day 10 and 14 in response to prolonged treatment with rosiglizatone (t-test, p-values= 1.6e-5 and 1.9e-5 respectively) (Figure 4K). Rosiglitazone treatment almost stopped LD fusion and enlargement while the increase in LD size was aberrant in the control group as there were fewer LDs under 3 µm diameter in Day 10 control group compared to Day 14 (Figure 4J). We observed the biggest effect of rosiglitazone on larger LDs. The 95th percentile LD diameter size was 0.6 and 1.8 µm smaller in rosiglitazone treatment compared to control group on

JOURNAL OF LIPID RESEARCH ASBMB

Day 10 and 14 respectively (vertical lines in Figure 4J). This indicated that prolonged treatment of rosiglitazone prevented LD enlargement during adipogenesis. The results are discussed further below.

DISCUSSION

There is growing public interest in pro- and anti-obesogenic substances to potentially combat the global obesity epidemic, yet little is known still about the effect of many substances on LD dynamics. Identification of specific interactions between fat storing cells and environmental signals will potentially delineate intrinsic mechanisms leading to fat depot expansion either through increase cell size or numbers (62). LipiD-QuanT provides a readily applicable method to monitor LD formation and growth during adipogeneisis and to study the impact of exogenous factors. We have demonstrated here that LipiD-QuanT is a suitable approach for measuring LD content and for determining whether a substance of interest can modulate LD accumulation, that it delivers a detailed picture of LD dynamics at the single LD/cell resolution and can measure LDs down to a size of a single pixel. LipiD-QuanT generates highly reproducible results, and is sensitive enough to determine small differences in LD sizes and size distributions during *in vitro* human adipogenesis. It compares favorably with biochemical or staining based techniques (Figure 3 and 4), while the results of biochemical and staining based techniques can be dependent on fixation or lysis conditions and lack information about structural features of LDs (Supplementary Figure 2C, 2D and 2E).

We have used LipiD-QuanT to monitor the impact of four treatments across adipogenesis. Anti-diabetic effects of ZnO nanoparticles had been observed in a diabetic rat model in which nanoparticles where they were absorbed by the fat tissue when consumed orally and reduced serum levels of free fatty acid and triglycerides over 40% (58). This prompted us to examine whether ZnO nanoparticles had a direct effect on LD formation and lipid accumulation. The lack of a significant effect, even at the relatively high dose used, suggests that the observed lowering serum free fatty acids and triglycerides is not due to a direct effect of ZnO nanoparticles on adipocytes.

In a number of systems, including stem cells from bone marrow, muscle and adipose tissue, high glucose exposure has been found to stimulate adipogenenic differentiation of the stem cells (63). *In vitro* studies in 3T3-L1 cells and rat muscle satellite cells have also shown that hyperglycemic conditions lead to increased lipogenesis and lipid accumulation (52). Across differentiation of SGBS cells in the presence of high (17.5 mM), compared with low (5 mM) glucose we did not see a significant change in lipid accumulation measured using either LipiD-QuanT or Oil Red O staining, though a consistent

slight shift to higher LD size was seen. It is possible that we may have seen greater effects, with a broader range of glycemic conditions, such as 4-25 mM glucose used in 3T3-L1 cells (52), but the lower concentration was chosen as SGBS cells do not reliably proliferate or differentiate below 5 mM glucose and 17.5 mM was chosen as being at the high end of the potential physiological range.

We studied the effect of DHA on differentiating SGBS cells at a 10 μM concentration that is relatively low compared with physiological plasma levels (64), but higher levels were toxic to SGBS cells. Neither LipiD-QuanT nor Oil Red O measurements showed any significant effect of DHA supplementation. Although our results contrasts with a DHA induced dose-dependent decrease in LD numbers and area in 3T3-L1 cells at higher (25-200 μM) levels of DHA (65), a more recent study also found no change in lipid accumulation of 3T3-L1 cells due to 2-day treatments of either 100 μM EPA or 50 μM DHA (66).

It has been shown in animal models that rosiglitazone lowers circulating triglyceride and free fatty acids hence improves insulin sensitivity (67, 68) as well as reducing adipocyte size and fat accumulation in muscle (59, 60). In cultured human subcutaneous adipocytes and in combination with insulin, rosiglitazone stimulated lipolysis *in vitro* (69). Moreover, on mature 3T3-L1 cells kept in high glucose, 2 day rosiglitazone treatment reduced the measured lipid content by 10% (70). Consistent with this, we have demonstrated that prolonged exposure to rosiglitazone has an anti-adipogenic effect during human adipogenesis, and LipiD-QuanT measurements demonstrate that it prevents enlargement of LDs leading to smaller LD size distribution (Figure 4J).

Although, the emphasis of this manuscript has been LD dynamics with the specific focus on human obesity, the application of Lipid-QuanT is not limited to a specific cell or organism. LDs contain well-conserved features among all organisms (71). Recently, commercial software has been used to measure LD dynamics in murine 3T3-L1 cells following fixation and Oil Red O staining (72). The size distribution of LDs in 3T3-L1 cells is very similar to that of SGBS cells measured using LipiD-QuanT. Thus Lipid-QuanT should readily be able to be applied to the murine 3T3-L1 adipocyte cell line. Additionally, HepG2 cells are reported to contain LDs in range from 0.2 to 1.4 μm (73). Visualization of the smaller lipid droplets (<0.3 μm) is feasible but would require capturing of images at higher magnification (2X) than the microscopy set-

up that is described. Thus, LipiD-QuanT adds further knowledge to previous research that investigated LD accumulation and lipid metabolism in the context of health and disease or due to the influence of various treatments during adipogenesis (74-76).

LipiD-QuanT complements staining and biochemical methods for measurement of lipid content and LD dynamics. It has the advantages that can provide measures of LD content and size distribution on a per cell basis, it is non-destructive to cells as they do not need to be fixed and it is user-friendly.

In addition, we believe LipiD-QuanT can be adapted to quantify lipid emulsions or lipid research in other systems where the contrast of LD contours is maximized and potentially be used for studies of LD biophysical characters such as stability, shape, and dynamics.

ACKNOWLEDGEMENTS

The study is part of the EpiSCOPE project supported by the Science and Industry Endowment Fund (Australia), grant RP03-064. In addition, HV is supported by iMQRES and CSIRO OCE Doctoral scholarships. We thank Stephen Bradford and David James' lab at the Garvan Institute of Medical Research for providing the SGBS cell line, Musarat Ishaq for providing PCR reagents, Susan van Dijk and Jing Zhou for providing HiDHA™ tuna oil and last but not least to Penny Bean, Meg Evans and Vijay Vaithilingam for their assistance with their microscopes. We also thank Michael Gillings, Kim Fung, Wayne Leifert and Lance Macaulay for critical reading the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Brown, D. A. 2001. Lipid droplets: Proteins floating on a pool of fat. *Curr Biol* **11**(11): R446-R449.

2. Haemmerle, G., Zimmermann R., and Zechner R. 2003. Letting lipids go: hormone-sensitive lipase. *Curr Opin Lipidol* **14**(3): 289-297.

3. Londos, C., Sztalryd C., Tansey J. T., and Kimmel A. R. 2005. Role of PAT proteins in lipid metabolism. *Biochimie* **87**(1): 45-49.

4. Francois, M., Leifert W., Hecker J., Faunt J., Martins R., Thomas P., and Fenech M. 2014. Altered cytological parameters in buccal cells from individuals with mild cognitive impairment and Alzheimer's disease. *Cytometry A* **85**(8): 698-708.

5. Fujimoto, T., and Parton R. G. 2011. Not Just Fat: The Structure and Function of the Lipid Droplet. *Cold Spring Harb Perspect Biol* **3**(3).

6. Miller, M., Stone N. J., Ballantyne C., Bittner V., Criqui M. H., Ginsberg H. N., Goldberg A. C., Howard W. J., Jacobson M. S., Kris-Etherton P. M., Lennie T. A., Levi M., Mazzone T., Pennathur S., Nursing C. C., and Dis C. K. C. 2011. Triglycerides and Cardiovascular Disease A Scientific Statement From the American Heart Association. *Circulation* **123**(20): 2292-2333.

7. Lutz, T. A., and Woods S. C. 2012. Overview of Animal Models of Obesity. *Current protocols in pharmacology / editorial board, SJ Enna (editor-in-chief) [et al]* **CHAPTER**: Unit5.61-Unit65.61.

8. Poulos, S. P., Dodson M. V., and Hausman G. J. 2010. Cell line models for differentiation: preadipocytes and adipocytes. *Exp Biol Med (Maywood)* **235**(10): 1185-1193.

9. Green, H., and Kehinde O. 1975. An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion. *Cell* **5**(1): 19-27.

10. Wabitsch, M., Brenner R. E., Melzner I., Braun M., Moller P., Heinze E., Debatin K. M., and Hauner H. 2001. Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. *Int J Obes* **25**(1): 8-15.

11. Allott, E. H., Oliver E., Lysaght J., Gray S. G., Reynolds J. V., Roche H. M., and Pidgeon G. P. 2012. The SGBS cell strain as a model for the in vitro study of obesity and cancer. *Clin Transl Oncol* **14**(10): 774-782.

12. Fischer-Posovszky, P., Newell F. S., Wabitsch M., and Tornqvist H. E. 2008. Human SGBS cells - a unique tool for studies of human fat cell biology. *Obesity facts* **1**(4): 184-189.

13. Thiam, A. R., Antonny B., Wang J., Delacotte J., Wilfling F., Walther T. C., Beck R., Rothman J. E., and Pincet F. 2013. COPI buds 60-nm lipid droplets from reconstituted water-phospholipid-triacylglyceride interfaces, suggesting a tension clamp function. *Proc Natl Acad Sci U S A* **110**(33): 13244-13249.

14. Fukumoto, S., and Fujimoto T. 2002. Deformation of lipid droplets in fixed samples. *Histochem Cell Biol* **118**(5): 423-428.

15. Gan, W. B., Grutzendler J., Wong W. T., Wong R. O. L., and Lichtman J. W. 2000. Multicolor "DiOlistic" labeling of the nervous system using lipophilic dye combinations. *Neuron* **27**(2): 219-225.

16. Listenberger, L. L., Ostermeyer-Fay A. G., Goldberg E. B., Brown W. J., and Brown D. A. 2007. Adipocyte differentiation-related protein reduces the lipid droplet association of adipose triglyceride lipase and slows triacylglycerol turnover. *J Lipid Res* **48**(12): 2751-2761.

17. Robenek, H., Robenek M. J., and Troyer D. 2005. PAT family proteins pervade lipid droplet cores. *J Lipid Res* **46**(6): 1331-1338.

18. Cirulis, J. T., Strasser B. C., Scott J. A., and Ross G. M. 2012. Optimization of staining conditions for microalgae with three lipophilic dyes to reduce precipitation and fluorescence variability. *Cytometry Part A* **81A**(7): 618-626.

19. Kacmar, J., Carlson R., Balogh S. J., and Srienc F. 2006. Staining and quantification of poly-3-hydroxybutyrate in Saccharomyces cerevisiae and Cupriavidus necator cell populations using automated flow cytometry. *Cytometry Part A* **69A**(1): 27-35.

20. Loudet, A., and Burgess K. 2007. BODIPY dyes and their derivatives: Syntheses and spectroscopic properties. *Chem Rev* **107**(11): 4891-4932.

21. Hellerer, T., Axang C., Brackmann C., Hillertz P., Pilon M., and Enejder A. 2007. Monitoring of lipid storage in Caenorhabditis elegans using coherent anti-Stokes Raman scattering (CARS) microscopy. *Proc Natl Acad Sci U S A* **104**(37): 14658-14663.

22. Le, T. T., Duren H. M., Slipchenko M. N., Hu C. D., and Cheng J. X. 2010. Label-free quantitative analysis of lipid metabolism in living Caenorhabditis elegans (vol 51, pg 672, 2010). *J Lipid Res* **51**(4): 875-875.

23. Dou, W., Zhang D. L., Jung Y., Cheng J. X., and Umulis D. M. 2012. Label-Free Imaging of Lipid-Droplet Intracellular Motion in Early Drosophila Embryos Using Femtosecond-Stimulated Raman Loss Microscopy. *Biophys J* **102**(7): 1666-1675.

24. Wong, C. S. Y., Robinson I., Ochsenkuhn M. A., Arlt J., Hossack W. J., and Crain J. 2011. Changes to lipid droplet configuration in mCMV-infected fibroblasts: live cell imaging with simultaneous CARS and two-photon fluorescence microscopy. *Biomedical Optics Express* **2**(9): 2504-2516.

25. Mouras, R., Bagnaninchi P. O., Downes A. R., and Elfick A. P. D. 2012. Label-free assessment of adipose-derived stem cell differentiation using coherent anti-Stokes Raman scattering and multiphoton microscopy. *Journal of Biomedical Optics* **17**(11).

26. Jungst, C., Klein M., and Zumbusch A. 2013. Long-term live cell microscopy studies of lipid droplet fusion dynamics in adipocytes. *J Lipid Res* **54**(12): 3419-3429.

27. Osmond-McLeod, M. J., Osmond R. I. W., Oytam Y., McCall M. J., Feltis B., Mackay-Sim A., Wood S. A., and Cook A. L. 2013. Surface coatings of ZnO nanoparticles mitigate differentially a host of transcriptional, protein and signalling responses in primary human olfactory cells. *Part Fibre Toxicol* **10**.

28. Osmond-McLeod MJ, O. Y., Osmond RIW, Sobhanmanesh F, McCall MJ 2014. Surface Coatings Protect against the In vitro Toxicity of Zinc Oxide Nanoparticles in Human Hepatic Stellate Cells. *J Nanomed Nanotechnol* **5**: 232.

29. Ostwald, W. 1897. Studien uber die Bildung und Umwandlung fester Korper. *Z Phys Chem* **22**: 289.

30. Ollila, O. H. S., Lamberg A., Lehtivaara M., Koivuniemi A., and Vattulainen I. 2012. Interfacial Tension and Surface Pressure of High Density Lipoprotein, Low Density Lipoprotein, and Related Lipid Droplets. *Biophys J* **103**(6): 1236-1244.

31. Vallotton, P. 2013. Size Matters: Filamentous Bacteria Drive Interstitial Vortex Formation and Colony Expansion in Paenibacillus vortex. *Cytometry Part A* **83**(12): 1105-1112.

32. Kimmel, R., and Bruckstein A. Regularized laplacian zero crossings as optimal edge integrators. 2001.

33. Aitkenhead, A. H. 2010. Boundary tracing using the Moore Neighbourhood. *Matlab Central File ExchangeRetrieved April 3, 2012*.

34. Kroon, D.-J. 2011. 2D Line Curvature and Normals. *Matlab Central File ExchangeRetrieved April 3, 2012*.

35. Soille, P. Morphological Image Analysis: Principles and Applications: Springer-Verlag New York, Inc.; 2003. 391 p.

36. Otsu, N. 1979. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* **9**(1): 62-66.

37. Lahnalampi, M., Heinäniemi M., Sinkkonen L., Wabitsch M., and Carlberg C. 2010. Time-Resolved Expression Profiling of the Nuclear Receptor Superfamily in Human Adipogenesis. *PLoS One* **5**(9): e12991.

38. Lindroos, J., Husa J., Mitterer G., Haschemi A., Rauscher S., Haas R., Gröger M., Loewe R., Kohrgruber N., Schrögendorfer K F., Prager G., Beck H., Pospisilik J A., Zeyda M., Stulnig T M., Patsch W., Wagner O., Esterbauer H., and Bilban M. 2013. Human but Not Mouse Adipogenesis Is Critically Dependent on LMO3. *Cell Metabolism* **18**(1): 62-74.

39. Murholm, M., Isidor M. S., Basse A. L., Winther S., Sorensen C., Skovgaard-Petersen J., Nielsen M. M., Hansen A. S., Quistorff B., and Hansen J. B. 2013. Retinoic acid has different effects on UCP1 expression in mouse and human adipocytes. *BMC Cell Biol* **14**: 41.

40. Weaver, R. E., Donnelly D., Wabitsch M., Grant P. J., and Balmforth A. J. 2008. Functional expression of glucose-dependent insulinotropic polypeptide receptors is coupled to differentiation in a human adipocyte model. *Int J Obes (Lond)* **32**(11): 1705-1711.

41. Calzadilla, P., Sapochnik D., Cosentino S., Diz V., Dicelio L., Calvo J. C., and Guerra L. N. 2011. N-Acetylcysteine Reduces Markers of Differentiation in 3T3-L1 Adipocytes. *International Journal of Molecular Sciences* **12**(10): 6936-6951.

42. Paar, M., Jungst C., Steiner N. A., Magnes C., Sinner F., Kolb D., Lass A., Zimmermann R., Zumbusch A., Kohlwein S. D., and Wolinski H. 2012. Remodeling of Lipid Droplets during Lipolysis and Growth in Adipocytes. *J Biol Chem* **287**(14): 11164-11173.

43. Suzuki, M., Shinohara Y., Ohsaki Y., and Fujimoto T. 2011. Lipid droplets: size matters. *J Electron Microsc (Tokyo)* **60**: S101-S116.

44. Buckley, J. D., and Howe P. R. C. 2009. Anti-obesity effects of long-chain omega-3 polyunsaturated fatty acids. *Obes Rev* **10**(6): 648-659.

45. Burrows, T., Collins C. E., and Garg M. L. 2011. Omega-3 index, obesity and insulin resistance in children. *Int J Pediatr Obes* **6**(2-2): e532-539.

46. Micallef, M., Munro I., Phang M., and Garg M. 2009. Plasma n-3 Polyunsaturated Fatty Acids are negatively associated with obesity. *Br J Nutr* **102**(9): 1370-1374.

47. Saito, E., Okada T., Abe Y., Kuromori Y., Miyashita M., Iwata F., Hara M., Ayusawa M., Mugishima H., and Kitamura Y. 2011. Docosahexaenoic acid content in plasma phospholipids and desaturase indices in obese children. *Journal of atherosclerosis and thrombosis* **18**(4): 345-350.

48. Garaulet, M., Hernandez-Morante J. J., Lujan J., Tebar F. J., and Zamora S. 2006. Relationship between fat cell size and number and fatty acid composition in adipose tissue from different fat depots in overweight/obese humans. *Int J Obes (Lond)* **30**(6): 899-905.

49. Guillet-Deniau, I., Pichard A. L., Kone A., Esnous C., Nieruchalski M., Girard J., and Prip-Buus C. 2004. Glucose induces de novo lipogenesis in rat muscle satellite cells through a sterol-regulatory-element-binding-protein-1c-dependent pathway. *J Cell Sci* **117**(10): 1937-1944.

50. Lu, B., Ennis D., Lai R., Bogdanovic E., Nikolov R., Salamon L., Fantus C., Le-Tien H., and Fantus I. G. 2001. Enhanced sensitivity of insulin-resistant adipocytes to vanadate is associated with oxidative stress and decreased reduction of vanadate (+5) to vanadyl (+4). *J Biol Chem* **276**(38): 35589-35598.

51. Tang, S., Le-Tien H., Goldstein B. J., Shin P., Lai R., and Fantus I. G. 2001. Decreased in situ insulin receptor dephosphorylation in hyperglycemia-induced insulin resistance in rat adipocytes. *Diabetes* **50**(1): 83-90.

52. Lin, Y., Berg A. H., Iyengar P., Lam T. K., Giacca A., Combs T. P., Rajala M. W., Du X., Rollman B., Li W., Hawkins M., Barzilai N., Rhodes C. J., Fantus I. G., Brownlee M., and Scherer P. E. 2005. The hyperglycemia-induced inflammatory response in adipocytes: the role of reactive oxygen species. *J Biol Chem* **280**(6): 4617-4626.

53. Saper, R. B., and Rash R. 2009. Zinc: An Essential Micronutrient. *Am Fam Physician* **79**(9): 768-772.

54. Mateo, M. C., Bustamante J. B., and Cantalapiedra M. A. 1978. Serum, zinc, copper and insulin in diabetes mellitus. *Biomedicine* **29**(2): 56-58.

55. Melchior, T., Simonsen K. W., Johannessen A. C., and Binder C. 1989. Plasma Zinc Concentrations during the 1st 2 Years after Diagnosis of Insulin-Dependent Diabetes-Mellitus - a Prospective-Study. *J Intern Med* **226**(1): 53-58.

56. Zargar, A. H., Shah N. A., Masoodi S. R., Laway B. A., Dar F. A., Khan A. R., Sofi F. A., and Wani A. I. 1998. Copper, zinc, and magnesium levels in non-insulin dependent diabetes mellitus. *Postgrad Med J* **74**(877): 665-668.

57. Alkaladi, A. A., A.M .; Afifi, M. 2014. Antidiabetic activity of zinc oxide and silver nanoparticles on streptozotocin-induced diabetic rats *Int J Mol Sci* **15**(2): 2015-2023.

58. Umrani, R. D., and Paknikar K. M. 2014. Zinc oxide nanoparticles show antidiabetic activity in streptozotocin-induced Type 1 and 2 diabetic rats. *Nanomedicine* **9**(1): 89-104.

59. Albrektsen, T., Frederiksen K. S., Holmes W. E., Boel E., Taylor K., and Fleckner J. 2002. Novel genes regulated by the insulin sensitizer rosiglitazone during adipocyte differentiation. *Diabetes* **51**(4): 1042-1051.

60. Johnson, J. A., Trasino S. E., Ferrante A. W., and Vasselli J. R. 2007. Prolonged decrease of adipocyte size after rosiglitazone treatment in high- and low-fat-fed rats. *Obesity* **15**(11): 2653-2663.

61. Molero, J. C., Lee S., Leizerman I., Chajut A., Cooper A., and Walder K. 2010. Effects of rosiglitazone on intramyocellular lipid accumulation in Psammomys obesus. *Biochim Biophys Acta* **1802**(2): 235-239.

62. Feve, B. 2005. Adipogenesis: cellular and molecular aspects. *Best Pract Res Clin Endocrinol Metab* **19**(4): 483-499.

63. nanoparticles, W. h. u. L.-Q. t. m. t. i. o. f. t. a. a. A.-d. e. o. Z. n. h. b. o. i. a. d. r. m. i. w., Leo S., Zavan B., Vindigni V., Rimessi A., Bianchi K., Franzin C., Cortivo R., Rossato M., Vettor R., Abatangelo G., Pozzan T., Pinton P., and Rizzuto R. 2008. High glucose induces adipogenic differentiation of muscle-derived stem cells. *Proc Natl Acad Sci U S A* **105**(4): 1226-1231.

64. Maki, K. C., Reeves M. S., Farmer M., Griinari M., Berge K., Vik H., Hubacher R., and Rains T. M. 2009. Krill oil supplementation increases plasma concentrations of eicosapentaenoic and docosahexaenoic acids in overweight and obese men and women. *Nutrition research (New York, NY)* **29**(9): 609-615.

65. Kim, H. K., Della-Fera M., Lin J., and Baile C. A. 2006. Docosahexaenoic acid inhibits adipocyte differentiation and induces apoptosis in 3T3-L1 preadipocytes. *J Nutr* **136**(12): 2965-2969.

66. Prostek, A., Gajewska M., Kamola D., and Balasinska B. 2014. The influence of EPA and DHA on markers of inflammation in 3T3-L1 cells at different stages of cellular maturation. *Lipids Health Dis* **13**: 3.

67. Oakes, N. D., Kennedy C. J., Jenkins A. B., Laybutt D. R., Chisholm D. J., and Kraegen E. W. 1994. A new antidiabetic agent, BRL 49653, reduces lipid availability and improves insulin action and glucoregulation in the rat. *Diabetes* **43**(10): 1203-1210.

68. Yki-Jarvinen, H. 2004. Thiazolidinediones. *N Engl J Med* **351**(11): 1106-1118.

69. McTernan, P. G., Harte A. L., Anderson L. A., Green A., Smith S. A., Holder J. C., Barnett A. H., Eggo M. C., and Kumar S. 2002. Insulin and rosiglitazone regulation of lipolysis and lipogenesis in human adipose tissue in vitro. *Diabetes* **51**(5): 1493-1498.

70. Wang, P., Renes J., Bouwman F., Bunschoten A., Mariman E., and Keijer J. 2007. Absence of an adipogenic effect of rosiglitazone on mature 3T3-L1 adipocytes: increase of lipid catabolism and reduction of adipokine expression. *Diabetologia* **50**(3): 654-665.

71. Thiel, K., Heier C., Haberl V., Thul P. J., Oberer M., Lass A., Jackle H., and Beller M. 2013. The evolutionarily conserved protein CG9186 is associated with lipid droplets, required for their positioning and for fat storage. *J Cell Sci* **126**(Pt 10): 2198-2212.

72. Rizzatti, V., Boschi F., Pedrotti M., Zoico E., Sbarbati A., and Zamboni M. 2013. Lipid Droplets Characterization in Adipocyte Differentiated 3T3-L1 Cells: Size and Optical Density Distribution. *European Journal of Histochemistry : EJH* **57**(3): e24.

73. Hur, W., Kim S. W., Lee Y. K., Choi J. E., Hong S. W., Song M. J., Bae S. H., Park T., Um S. J., and Yoon S. K. 2012. Oleuropein reduces free fatty acid-induced lipogenesis via lowered extracellular signal-regulated kinase activation in hepatocytes. *Nutrition research (New York, NY)* **32**(10): 778-786.

74. Koves, T. R., Sparks L. M., Kovalik J. P., Mosedale M., Arumugam R., DeBalsi K. L., Everingham K., Thorne L., Phielix E., Meex R. C., Kien C. L., Hesselink M. K., Schrauwen P., and Muoio D. M. 2013. PPARgamma coactivator-1alpha contributes to exercise-induced regulation of intramuscular lipid droplet programming in mice and humans. *J Lipid Res* **54**(2): 522-534.

75. Robciuc, A., Hyotylainen T., Jauhiainen M., and Holopainen J. M. 2012. Hyperosmolarity-induced lipid droplet formation depends on ceramide production by neutral sphingomyelinase 2. *J Lipid Res* **53**(11): 2286-2295.

76. Russell, T. D., Palmer C. A., Orlicky D. J., Fischer A., Rudolph M. C., Neville M. C., and McManaman J. L. 2007. Cytoplasmic lipid droplet accumulation in developing mammary epithelial cells: roles of adipophilin and lipid metabolism. *J Lipid Res* **48**(7): 1463-1475.

FIGURES LEGENDS AND FIGURES

Figure 1: LipiD-QuanT validation.

A- Image of 4% paraformaldehyde fixed SGBS human pre-adipocytes on Day 7, stained with DAPI to localize nuclei.

B- Nuclei in Figure 1A were segmented automatically via LipiD-QuanT in order to derive per-cell statistics for lipid accumulation.

C- The same image field as in Figure 1A was acquired under phase contrast microscopy.

D- A representative image of differentiated SGBS cells acquired under phase contrast microscopy showing well-contrasted LD boundaries.

E- A representative image of differentiated SGBS cells representing LDs segmented by LipiD-QuanT. A region of negative curvature cytoplasmic space between LDs is indicated by a red arrow.

F- A representative image of differentiated SGBS cells representing LDs segmentation by manual counting (The lower cell was not analyzed).

G- Comparison of smoothed distribution of LD diameter sizes obtained from automated LipiD-QuanT measurements and manual segmentation (n= 1,048 LDs). Pixel size in object space is 0.3 µm.

H- The LD identity was confirmed at the molecular level using anti-perilipin fluorescence staining of the protein (green channel) decorating their periphery.

I- Size distributions are not sensitive to focus choice. Smoothed distribution of LD diameter sizes of mature human adipocytes is reproduced reliably after re-setting the focus in independent trials (n=1,165 LDs).

Figure 2: Flowchart presenting the key steps of LipiD-QuanT algorithm

Figure 3: Performance of LipiD-QuanT in monitoring LD accumulation during adipogenesis, compared to Oil Red O staining and Triglyceride accumulation protocols

A- Representative images of Oil Red O stained SGBS cells over the adipogenesis time-course of 21 days. As expected Oil Red O is seen to partition predominantly in LDs.

B- Representative images of label-free differentiating SGBS cells as observed by phase contrast microscopy over a period of 21 days. The average LD's diameter is seen to increase spectacularly over this period.

C- The Oil Red O dye extractions were measured spectrophotometrically at 520 nm (OD_{520nm}) over a period of 21 days. Data are shown as mean±standard deviation (M±SD) (n= 3 biological replicates with 2 technical replicate per sample).

D- Quantification of total triglyceride accumulation using enzymatic digest method during adipogenesis. The total glycerol concentration is shown in µM. Data are shown as M±SD of experiment performed in triplicate wells (2 technical replicate per sample).

E- Total LD area calculated on Day 7, 10, 14 and 21 of human adipogenesis using LipiD-QuanT divided to total number of differentiated cells using per DAPI stained nuclei images coupled with phase contrast images. On day 7, 10, 14 and 21 there were a total of 99, 104, 96 and 132 cells respectively (6 images were used per differentiation time point).

F- Smoothed distribution of LD diameter sizes obtained using the LipiD-QuanT software during adipogenesis (n=5,670 LDs). A systematic shift towards larger LD sizes is measured.

Figure 4: The effect of pro- and anti-obesogenic treatments on LD formation and accumulation during adipogenesis

A- Smoothed distribution of LD diameter sizes in terminally differentiated SGBS cells in the presence and absence of 10 µmol/L DHA treatment over a 14 day period (n= 7,444 LDs in 8 images).

B- The Oil Red O dye extractions were measured spectrophotometrically at 520 nm (OD_{520nm}) in terminally differentiated SGBS cells in the presence and absence of 10 µmol/L DHA treatment over a 14 day period. Data are shown as M±SD (n= 3 biological replicates with 2 technical replicates per sample).

C- Representative microscopy images of control and 10 µmol/L DHA treated SGBS cells on day 14. For these and panels F, I and L, images are only representative of a small portion within a single image.

D- The change of smoothed distribution of LD diameter sizes in differentiating SGBS cells in low and high glucose concentrations for a total of 14 days (n=5,955 LDs).

JOURNAL OF LIPID RESEARCH ASBMB

E- The Oil Red O dye extractions were measured spectrophotometrically at 520 nm (OD_{520nm}) in differentiating SGBS cells in low and high glucose concentrations for a total of 14 days. Data are shown as M±SD (n= 4 biological replicates with 2 technical replicates per sample).

F- Representative microscopy images of SGBS cells fed in low and high glucose concentrated media up to day 14.

G- The change of smoothed distribution of LD diameter sizes in differentiating SGBS cells on day 8 and 12, following a 48h treatment with or without the 10 µg coated ZnO nanoparticles on day 6 (n=33,626 LDs).

H- The Oil Red O dye extractions were measured spectrophotometrically at 520 nm (OD_{520nm}) in differentiating SGBS cells on day 8 and 12, following a 48h treatment with or without the 10 µg coated ZnO nanoparticles on day 6. Data are shown as M±SD (n= 2 biological replicates with 2 technical replicates per sample).

I- Representative microscopy images of SGBS cells at day 14 following a 48h treatment with or without a 48h treatment the 10 µg coated ZnO nanoparticles on day 6.

J- Smoothed distribution of LD diameter sizes in SGBS cells differentiated in the presence and absence of 2 µmol/L Roziglitazone until day 14 (n=33,749 LDs). Vertical lines indicate 95th percentile; full lines, day 10 and dashed lines, day 14.

K- The Oil Red O dye extractions were measured spectrophotometrically at 520 nm (OD_{520nm}) in SGBS cells differentiated in the presence or absence of 2 µmol/L Roziglitazone until day 14. Data are shown as M±SD (n= 3 biological replicates with 2 technical replicates per sample).

L- Representative microscopy images of SGBS cells in the presence and absence of 2 µmol/L Roziglitazone for a total of 14 days.

JOURNAL OF LIPID RESEARCH

Figure 1: LipiD-QuanT validation

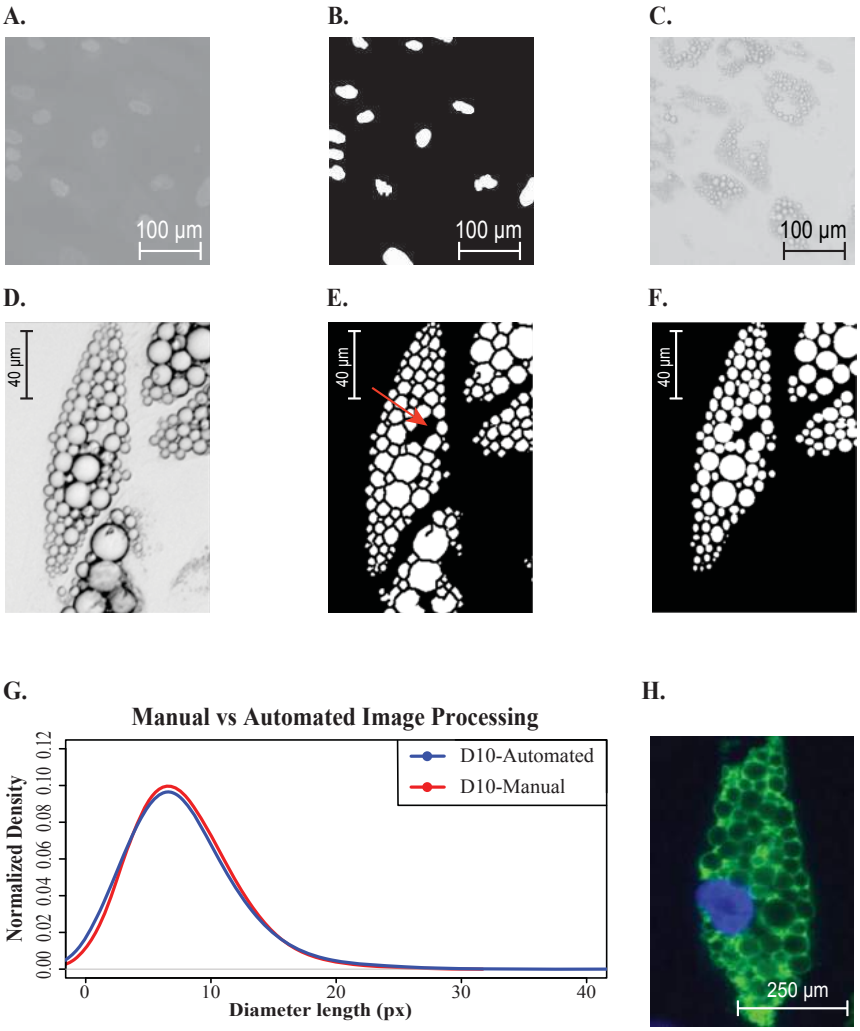


Figure 2: Flowchart presenting the key steps of LipiD-QuanT algorithm

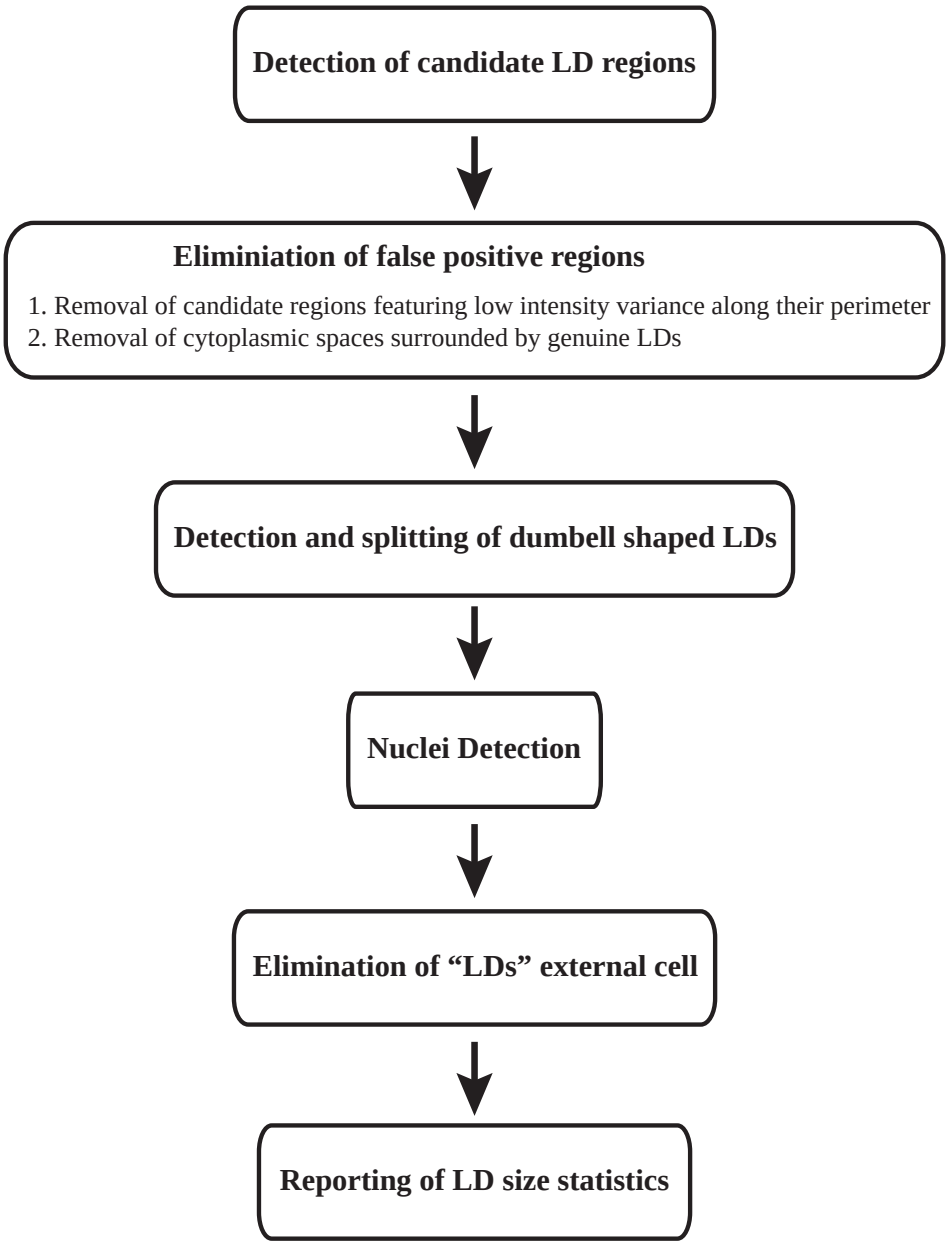


Figure 3: Comparison of LipiD-QuanT performance in monitoring LD accumulation during adipogenesis compared to widely used other protocols.

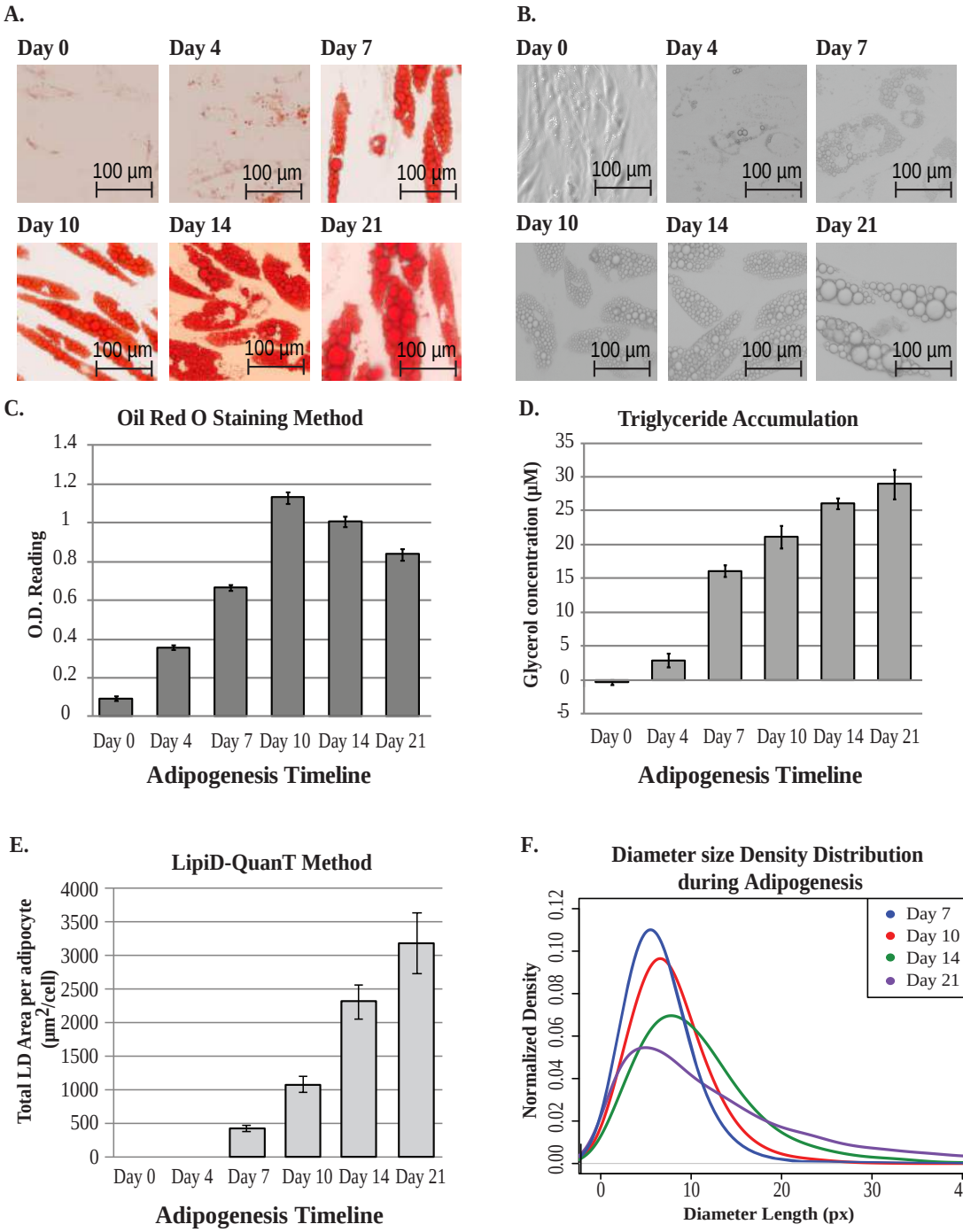
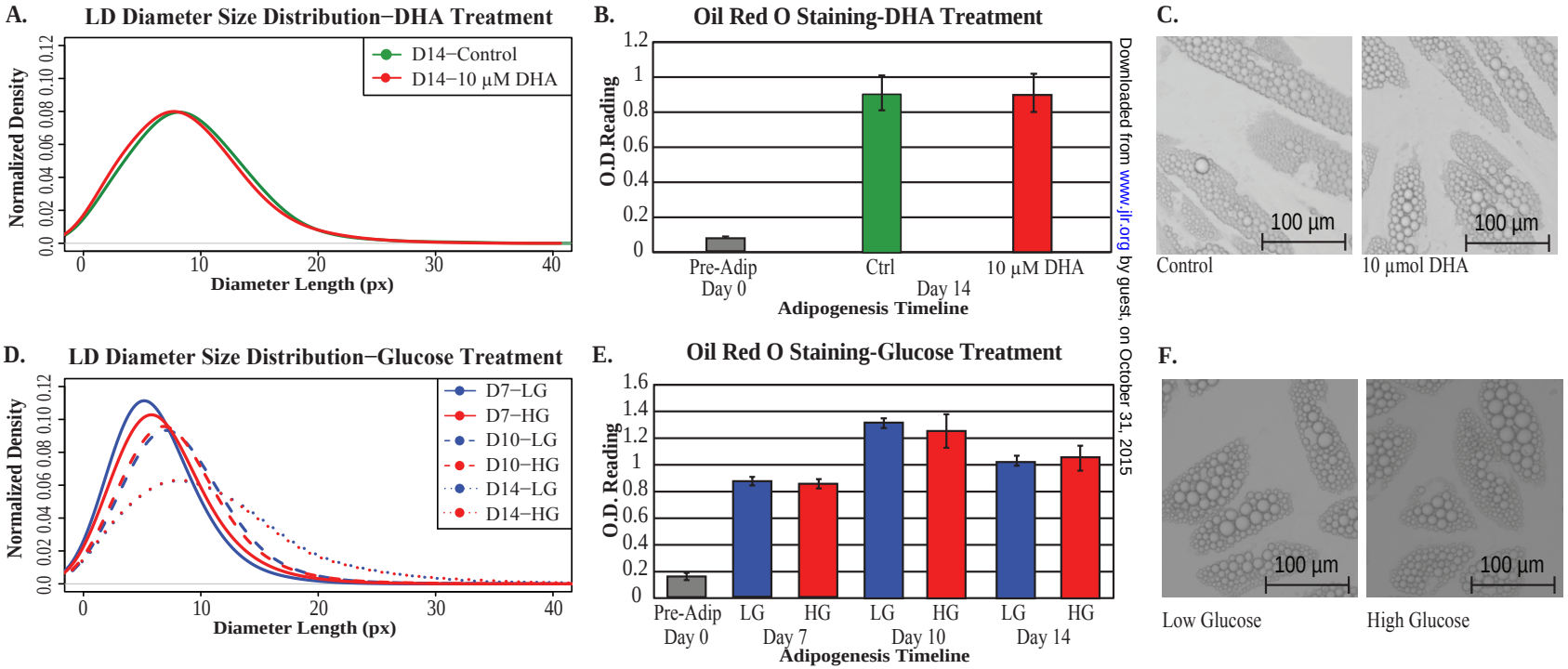
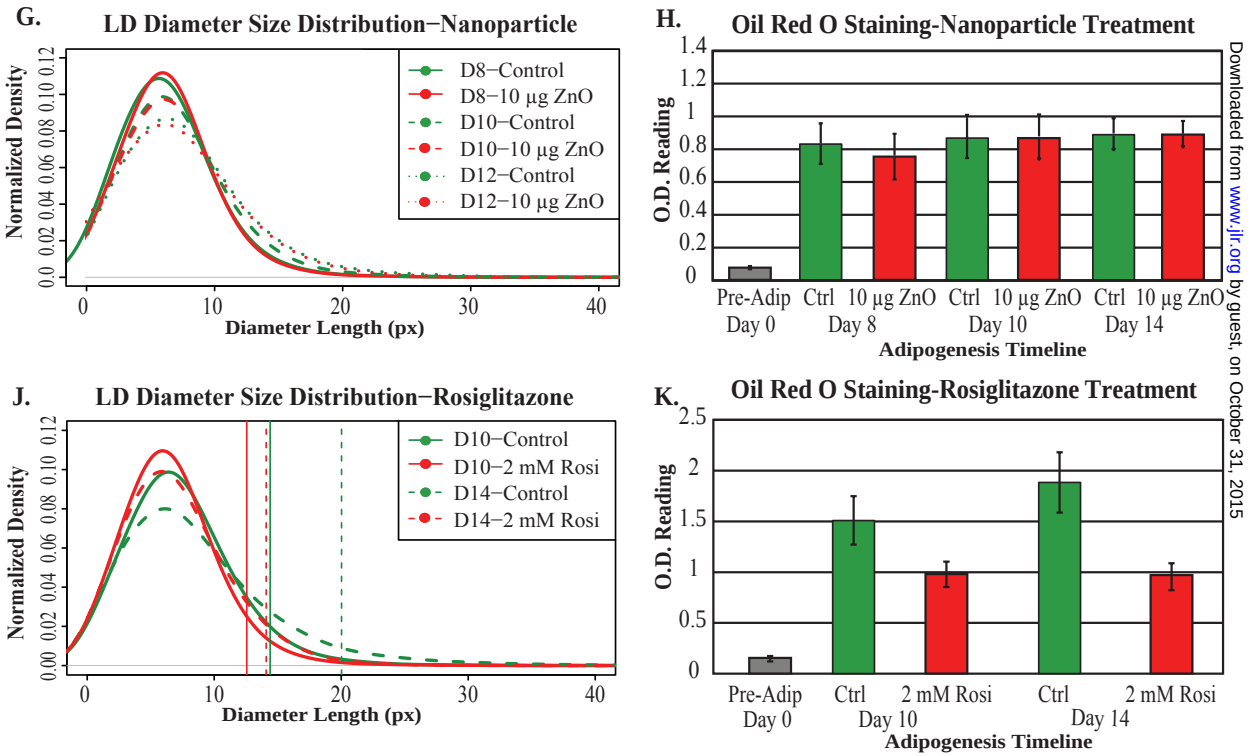


Figure 4: The effect of pro- and anti-obesogenic treatments on LD formation and accumulation during adipogenesis

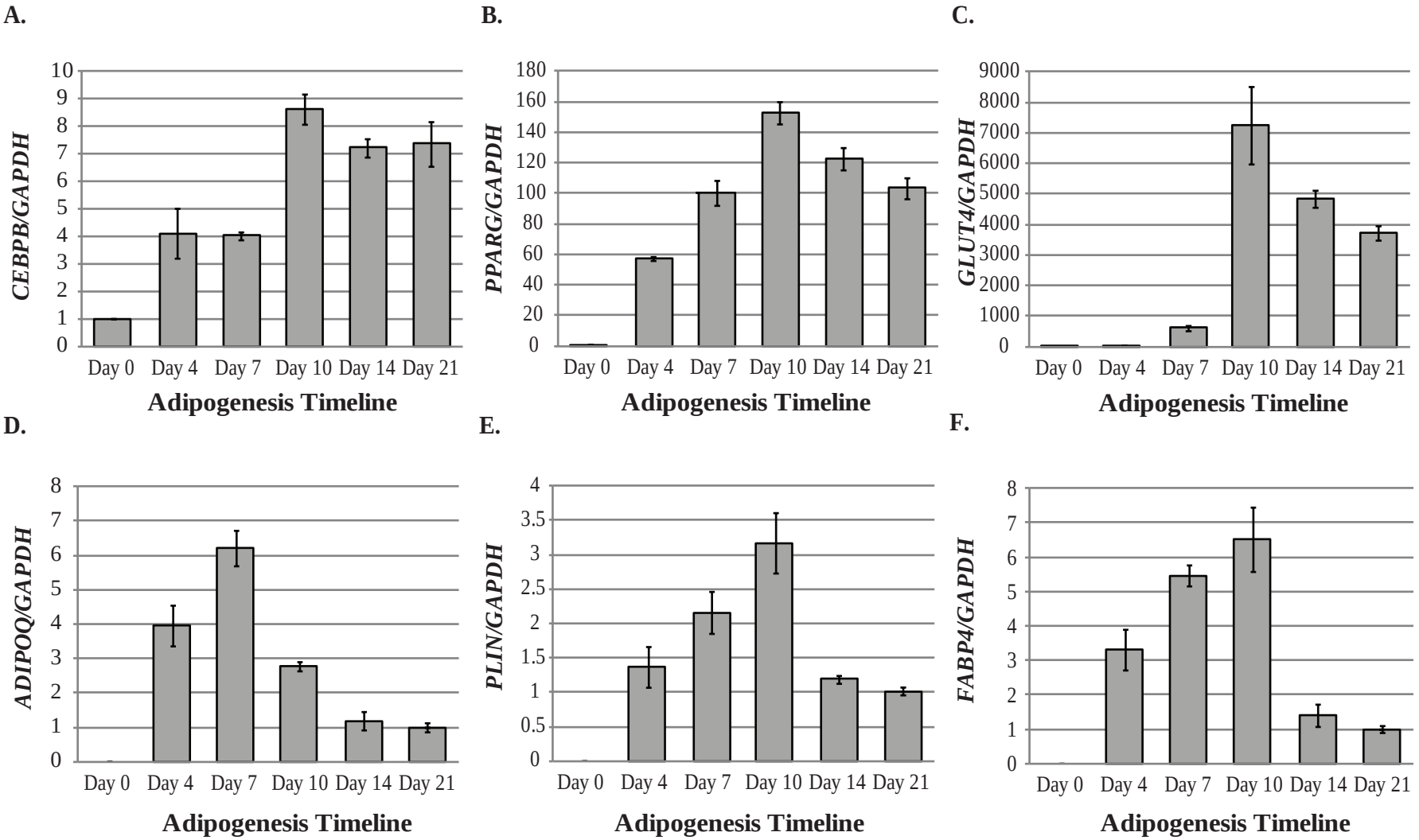


JOURNAL OF LIPID RESEARCH
ASBMB

Figure 4: The effect of pro- and anti-obesogenic treatments on LD formation and accumulation during adipogenesis



Supplementary Figure 1: Relative expression of *CEBPB*, *PPARG*, *GLUT4*, *ADIPOQ*, *PLIN* and *FABP4* during 21 days of adipogenesis in SGBS cells. Expression levels of target genes were determined by quantitative RT-PCR and are normalized to the control GAPDH gene.



Supplementary Figure 2: LipiD-QuanT performance on synthetic images with LDs and drawbacks of Oil Red o staining method for monitoring adipogenesis.

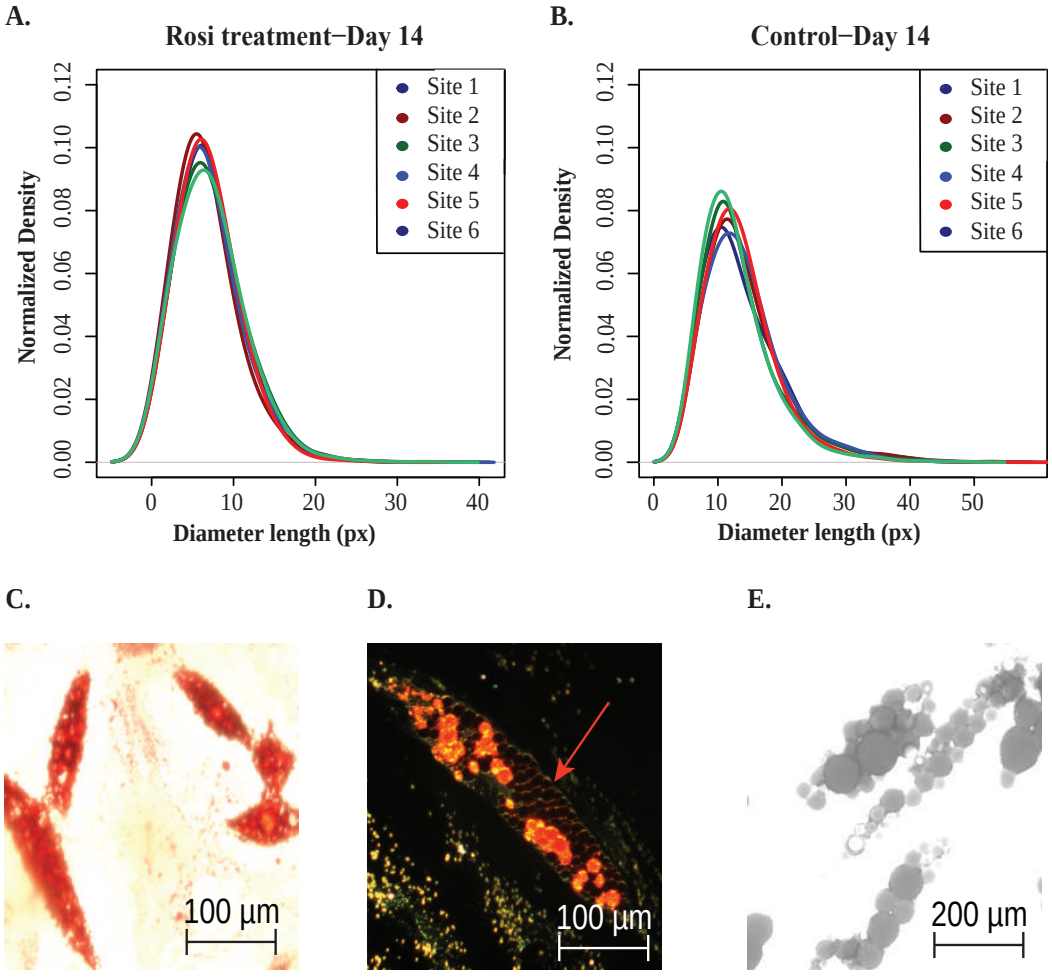
A- The LD diameter size density distribution is consistent between images taken from different parts of the same well containing 2 µmol/L Rosiglitazone treated cells on day 14 (n=4,572 LDs in 6 images).

B- The LD diameter size density distribution is consistent between images taken from different parts of the same well containing untreated cells on day 14 (n =1,679 LDs in 6 images).

C- Oil Red O stained SGBS adipocytes with a few burst LDs and high background staining under phase contrast microscopy.

D- Oil Red O stained SGBS adipocytes with a few burst LDs with three dimensional mapping using Cytoviva microscopy.

E- Phase contrast image of well-fixed and Oil Red O stained mature SGBS adipocytes. This illustrates the variability of Oil Red O staining, as apparent by the variability in staining intensities of LD's of similar sizes.



Supplementary Table 1: Sequence of PCR primers used in quantitative RT-PCR.

Gene Name (Human)	Gene Bank ID	Amplicon (nt)	Primer Name	Length	Sequence
Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*)	NM_002046.3	66	**GAPDH-Forward**	19	AGCCACATCGCTCAGACAC
			GAPDH-Reverse	19	GCCCAATACGACCAAATCC
CCAAT/enhancer binding protein (C/EBP), beta (*CEBPB*)	NM_005194.2	65	**CEBPB -Forward**	18	CGCTTACCTCGGCTACCA
			CEBPB -Reverse	19	ACGAGGAGGACGTGGAGAG
Peroxisome proliferator-activated receptor gamma (*PPARG*)	NM_138712.3	96	**PPARG-Forward**	24	GACAGGAAAGACAACAGACAAATC
			PPARG-Reverse	21	GGGGTGATGTGTTTGAACTTG
Adiponectin, C1Q and collagen domain containing (*ADIPOQ*)	NM_004797.3	61	**ADIPOQ- Forward**	21	GGTGAGAAGGGTGAGAAAGGA
			ADIPOQ- Reverse	21	TTTCACCGATGTCTCCCTTAG
Perilipin 1 (*PLIN*)	NM_002666.4	101	**PLIN-Forward**	24	ACATTAAAGGGAAGAAGTTGAAGC
			PLIN-Reverse	19	GCAGCACATTCTCCTGCTC
Solute carrier family 2 (facilitated glucose transporter), member 4 (*SLC2A4/GLUT4*)	NM_001042.2	62	**GLUT4-Forward**	20	CTGTGCCATCCTGATGACTG
			GLUT4-Reverse	20	CGTAGCTCATGGCTGGAACT
Fatty acid binding protein 4, adipocyte (*FABP4*)	NM_001442.2	70	**FABP4-Forward**	23	CCACCATAAAGAGAAAACGAGAG
			FABP4-Reverse	20	GTGGAAGTGACGCCTTTCAT

CHAPTER 3: COBRA-Seq: sensitive and quantitative methylome profiling

Epigenetics is among *the fastest* moving *areas in* molecular biology, shedding light on how environmental factors affect disease manifestation in humans. Most common methods for measuring DNA methylation target locations of the genome where methylation occurs densely, however there is growing evidence that less dense methylation in other genomic locations is equally, if not more important.

This paper describes a new genome-wide reduced methylome method (COBRA-Seq). The method enriches methylated DNA fractions by digesting the genomic DNA with restriction enzymes that recognize potential methylation sites after bisulfite conversion. The genomic complexity is further reduced by removing DNA fragments without the enzyme recognition site using streptavidin coated magnetic beads. COBRA-seq samples the genome proportionally for generation and analysis of genome-scale DNA methylation profiles with nucleotide resolution at a reduced cost. We compared the relative methylation measures generated with this novel method to other methylome sampling methods. This paper also reviews other widely used methylome methods.

The publication presented in this chapter is the product of a working collaboration between myself, Ross, Statham and Molloy. I designed the experiments, carried out all the laboratory experiments and wrote the manuscript. Ross and Statham analysed the genomic data. Ross, Molloy and Clark contributed to the experimental design and writing of the manuscript.

This manuscript has been published in *Genes*, October 2015. It was the 4th most downloaded manuscript in 2015 published by *Genes*.

Detailed contributions:

	Varinli H	**Co-authors**
Experimental Design	60%	40%
Laboratory Experiments	100%	
Data Analysis	10%	90%
Writing	70%	30%

Genes **2015**, *6*, 1140-1163; doi:10.3390/genes6041140

OPEN ACCESS

genes

ISSN 2073-4425

www.mdpi.com/journal/genes

Article

COBRA-Seq: Sensitive and Quantitative Methylome Profiling

Hilal Varinli [1,2,3,*], Aaron L. Statham [2], Susan J. Clark [2,4], Peter L. Molloy [1] and Jason P. Ross [1]

1 CSIRO Food and Nutrition Flagship, North Ryde, New South Wales 1670, Australia; E-Mails: peter.molloy@csiro.au (P.L.M.); jason.ross@csiro.au (J.P.R.)

2 Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales 2010, Australia; E-Mails: a.statham@garvan.org.au (A.L.S.); s.clark@garvan.org.au (S.J.C.)

3 Department of Biological Sciences, Macquarie University, North Ryde, New South Wales 2109, Australia

4 Vincent's Clinical School, Faculty of Medicine, UNSW, New South Wales 2010, Australia

* Author to whom correspondence should be addressed; E-Mail: hilal.varinli@csiro.au; Tel.: +61-2-9490-5136; Fax: +61-2-9490-5020.

Academic Editors: Jeffrey Craig and Thomas Mikeska

Received: 10 August 2015 / Accepted: 24 September 2015 / Published: 23 October 2015

Abstract: Combined Bisulfite Restriction Analysis (COBRA) quantifies DNA methylation at a specific locus. It does so via digestion of PCR amplicons produced from bisulfite-treated DNA, using a restriction enzyme that contains a cytosine within its recognition sequence, such as *TaqI*. Here, we introduce COBRA-seq, a genome wide reduced methylome method that requires minimal DNA input (0.1–1.0 μg) and can either use PCR or linear amplification to amplify the sequencing library. Variants of COBRA-seq can be used to explore CpG-depleted as well as CpG-rich regions in vertebrate DNA. The choice of enzyme influences enrichment for specific genomic features, such as CpG-rich promoters and CpG islands, or enrichment for less CpG dense regions such as enhancers. COBRA-seq coupled with linear amplification has the additional advantage of reduced PCR bias by producing full length fragments at high abundance. Unlike other reduced representative methylome methods, COBRA-seq has great flexibility in the choice of enzyme and can be multiplexed and tuned, to reduce sequencing costs and to interrogate different numbers of sites. Moreover, COBRA-seq is applicable to non-model organisms without the reference genome and compatible with the investigation of non-CpG methylation by using restriction enzymes containing CpA, CpT, and CpC in their recognition site.

Keywords: COBRA; DNA methylation; reduced representation; non CpG; non-model organism; restriction enzymes; next generation sequencing; enhancer; CHH

1. Introduction

Extensive effort has been devoted to mapping the methylome of various human tissues [1–3]. Genome-wide methylation is also widely studied in plants, *Arabidopsis thaliana* [4–9], *Oryza sativa* (rice) [5,8,10], and several model species ranging from fungi [8,11–13], *Drosophila melanogaster* (fruit fly) [14,15], *Apis mellifera* (honey bee) [16], *Danio rerio* (zebrafish) [5,17,18] and *Mus musculus* (mouse) [5,19,20]. Methylome analysis is also beginning to be applied to livestock species such as sheep [21]. Moreover, there is a growing interest in understanding the molecular bases of epigenetic inheritance in non-model organisms [22], including organisms of ecological and economic importance (*i.e.*, root-knot nematodes [23]).

The gold standard technique for quantifying methylomes is whole-genome bisulfite sequencing (WGBS) [1,6,24]. Bisulfite treatment converts cytosines to uracil while 5-methylcytosines (5mC) are preserved [25,26]. Thus, in subsequent sequencing of bisulfite-treated DNA the 5mC are read as cytosines, and unmethylated cytosines as thymines at a single base resolution. Although WGBS gives nucleotide-base resolution across the entire genome, in humans, only about 20% of the CpGs are reported as being dynamic (varying by more than 30%) among various human tissues [27]. WGBS is currently too expensive to generate large-scale data on multiple cell types or large sample sets. It is even more challenging in systems biology research that involves longitudinal sampling of multiple tissue and cell types in large cohorts, particularly in the case of biomarker discovery.

While methylation can occur in a variety of sequence contexts in plants and other organisms, vertebrate genomes are methylated predominantly at the dinucleotide CpG [28]. Empirically, it has been shown that 65% of 100 bp WGBS reads of mammalian DNA contain no CpG sites [27]. Consequently, much of the sequencing bandwidth of WGBS is wasted. Further, methylomes with low levels of 5mC, such as those from insect taxa [29] demonstrate an even more pronounced waste of sequencing bandwidth. For example, less than 1% of the CpGs are methylated in the honeybee genome [16]. There is a growing interest in profiling only a representative subset of cytosines across the genome to limit cost, and increase cohort size, regardless of the choice of organism.

A wide number of microarray and sequence-based DNA methylation detection and analysis methods have been developed that target portions of the genome and reduce genomic complexity [30–39]. Each method targets a different subset of the methylome, revealing methylation profiles across a different distribution of genomic features (*i.e.*, CpG rich and/or poor regions) [40]. There are three main approaches that exploit next-generation sequencing to derive methylome data: bisulfite conversion, affinity enrichment and enzymatic restriction using methylation sensitive endonucleases.

The traditional Combined Bisulfite Restriction Analysis (COBRA) method quantifies DNA methylation in a specific gene region by PCR amplification from bisulfite-treated DNA. This is followed by digestion of the PCR amplicon with a restriction enzyme such as *TaqI* (5'-T/CGA-3') that contains a CpG site in its recognition sequence. The methylation status of a particular CpG site within an enzyme recognition

site can then be determined. The proportions of cut *vs*. uncut DNA can be used to determine the level of methylation [41].

Here, we have optimized the original COBRA method for high-throughput genome sequencing platforms and refer to the method as COBRA-seq. COBRA-seq enriches methylated DNA fractions by digesting the genomic DNA with restriction enzymes recognizing potential methylation sites after bisulfite conversion. The genomic complexity is further reduced by removing DNA fragments without the enzyme recognition site using streptavidin coated magnetic beads. Therefore, COBRA-seq provides single base pair resolution data within multiple regions of interest containing methylated sites. COBRA-seq is compatible with various restriction enzymes allowing the user to explore 5mC in any sequence context (CpA, CpT, CpC as well as CpG).

We describe two versions of COBRA-seq: First, Genome-Wide COBRA (GW-COBRA) that uses PCR to exponentially amplify the sequencing library fragments after linker ligation. Second, Linear Amplification COBRA (LA-COBRA) that relies on T7 RNA polymerase-mediated transcription yielding many RNA transcripts of the library fragments generated in a linear manner which are subsequently converted to cDNA. The COBRA-seq library protocol is largely adapted to provide robust coverage of the human genome, however, all of the steps are applicable to both model and non-model organisms.

We prepared GW-COBRA and LA-COBRA sequencing libraries from the colon carcinoma (HCT116) cell line and analyzed the methylome with 100 bp single-end sequencing using Illumina HiSeq2000 chemistry. We also compared, *in silico*, the quantitative, qualitative and genome coverage specifications of COBRA-seq with other common DNA methylome technologies.

2. Materials and Methods

2.1. Cell Culture and Genomic DNA Isolation

HCT116 colon cancer cells were cultured in McCoy's 5A media (Life Technologies, cat#16600-082, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (Life Technologies, cat#10099-141). Genomic DNA was isolated using a Gentra Puregene Cell Kit (Qiagen, cat#158745, Redwood City, CA, USA) as per manufacturer's instructions. Purified genomic DNA was quantified with a Nanodrop ND-1000 (Thermo Scientific, Carlsbad, CA, USA).

2.2. Annealing Oligonucleotides to Construct COBRA-Seq Adapters

A final concentration of 50 μM annealed adapter stocks were prepared using the corresponding oligonucleotide pairs described below in 1× Quick Ligation™ Reaction Buffer (NEB, supplied in cat#M2200S, Ipswich, MA, USA). The reactions were heated to 95 °C for 5 min then cooled down gradually as follows: 72 °C for 5 min, 60 °C for 5 min, 50 °C for 3 min, 40 °C for 3 min, 30 °C for 3 min, 20 °C for 3 min, 10 °C for 3 min and 4 °C for forever. The annealed adapter stocks were stored in −20 °C.

GW-COBRA and LA-COBRA Adapter 2 (A2). A2, has a T-3' overhang, composed of A2-LowerStrand and A2-UpperStrand oligonucleotide pair provided in Table 1, were modified from the original TruSeq Illumina Y-shaped Adapter 2 oligonucleotides. All the cytosines were changed to 5mCs in the lower strand (A2-LowerStrand), hence the sequence were maintained post-bisulfite treatment. On the upper strand most cytosines in the flanking end were changed to 5mCs. Thus after bisulfite treatment, sequences in

the stem of the adapters are no longer complementary. After bisulfite treatment the sequence on the fragment ending with A2-LowerStrand oligonucleotide sequence is appropriate for next-generation sequencing flow cell amplification.

Table 1. Details of oligos used in constructing GW-COBRA and LA-COBRA methylome libraries. The less common sequence abbreviations are: 5 (5mC), b (5' two biotin groups) and p (5' phosphorylation).

Primer Name	Primer Sequence (5'–3')*	Purification Method
A2-UpperStrand	TGT 5A5 5G5 TGG T5A T5C GCT GCT CTT CCG ATC T	PAGE and HPLC
A2-LowerStrand	**p**GAT 5GG AAG AG5 T5G TAT G55 GT5 TT5 TG5 TTG	PAGE and HPLC
GW-A1-UpperStrand	CTA CAC TCT TTC CCT ACA CGA CGC TCT TCC GAT CT	PAGE and HPLC
GW-A1-LowerStrand	CGA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA G	PAGE and HPLC
LA-A1+P5-UpperStrand	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T	PAGE and HPLC
LA-A1+P5-LowerStrand	CGA GAT CGG AAG AGC GTC GTG TAG GGA AAG AGT GTA G AT CTC GGT GGT CGC CGT ATC ATT	PAGE and HPLC
GW-A2-FwdP	CAA GCA GAA GAC GGC ATA CGA GCT CTT CCG ATC T	PAGE
COBRA-A2-RevP	**b**TGT CAC CGC TGG TCA TCT GTT GTT T	HPLC
LA-A2+T7-FwdP	GAA TTT AAT ACG ACT CAC TAT AGG GAC AAG CAG AAG ACG GCA TAC GAG C	PAGE
FlowCell-FwdP	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T	HPLC
FlowCell-RevP	CAA GCA GAA GAC GGC ATA CGA GCT CTT CCG ATC T	HPLC
LADS P5	AAT GAT ACG GCG ACC ACC GA	HPLC
LADS P7	CAA GCA GAA GAC GGC ATA CGA	HPLC

GW-COBRA Adapter 1 (GW-A1). GW-A1, has a 5'-CG overhang, composed of GW-A1-UpperStrand and GW-A1-LowerStrand oligos in Table 1.

LA-COBRA Adapter-1 (LA-A1). LA-A1, composed of LA-A1+P5-UpperStrand and LA-A1+P5-LowerStrand oligos in Table 1, also has a 5'-CG overhang and contains an addition of the P5 primer region that allowed reverse transcription and cDNA synthesis. Both GW-A1 and LA-A1 contained the appropriate end for flow cell amplification when paired with A2.

2.3. Library Construction and Bisulfite Treatment

Fragment Preparation. The genomic DNA was resuspended in 300 µL low TE (10 mM Tris, 0.1 mM EDTA, pH 7.5) to a final concentration of 16.66 ng/µL and fragmented using a Bioruptor UCD-200 sonicator (Diagenode) at a power setting of "high" for sets of 30 cycles of 15 s on/off with 15 min intervals on ice in between each set. The fragmented DNA, 100–500 bp, was concentrated using ethanol precipitation. Each replicate containing 2 µg DNA was end repaired using the End-It™ DNA End-Repair Kit (Epicentre

Biotechnologies, cat#ER0720, Madison, WI, USA), repurified with the standard phenol:chloroform: isoamyl alcohol extraction, A-tailed using Klenow Exo- (NEB, cat#E6053S) as per manufacturer's protocol then the reaction was inactivated at 75 °C for 20 min.

Ligation of A2. Ligation was performed using a Quick Ligation™ kit (NEB, cat#M2200S) in the presence of a 10-fold excess of A2 (310 pmol for 2 μg of genomic library with average fragment size of 200 bp) as per the manufacturer's instructions. The DNA was cleaned up with a QiaQUICK PCR purification kit (Qiagen, cat#28104), eluted in 84 μL of elution buffer (EB).

Bisulfite Conversion. The bisulfite treatment of ligated genomic DNA was carried out with the EZ DNA Methylation Kit (Zymo Research, cat# D5001, Irvine, CA, USA) following the manufacturer's protocol with minor modifications. The bisulfite conversion reaction was incubated at 99 °C 5 min, 60 °C 25 min, 99 °C 5 min, 60 °C 85 min, 99 °C 5 min, 60 °C 175 min and 22 °C 5 min. Yeast tRNA were added to low levels of genomic DNA (<0.5 μg) to minimize sample loss.

Limited PCR Amplification. Single-stranded bisulfite DNAs were used as PCR templates at a final concentration of 1 ng/μL in 25 μL using GoTaq colorless master mix (Promega, cat#M7133, Madison, WI, USA) [25]. The reverse primer contained 5' double biotin (COBRA-A2-RevP in Table 1). In addition to GW-COBRA forward primer shown as GW-A2-FwdP in Table 1, LA-COBRA forward primer (LA-A2+T7-FwdP in Table 1) had an overhang of T7 promoter region that allowed linear amplification of the library fragments via *in vitro* transcription in the later steps. The temperature cycles for the PCR were: 98 °C 3 min; 98 °C for 15 s, 56 °C for 30 s and 72 °C for 1 min, for 6 cycles; a final extension of 72 °C for 2 min. A Wizard SV PCR Clean Up System (Promega, cat#A9281) was used to remove the enzymes and excess primers as per the manufacturer's instructions.

Restriction Digestion and Enrichment of Methylated DNA Fragments. PCR amplified and purified 2 μg library material was then digested overnight with 100 U of *TaqI* (NEB, cat#R0149S) in NEB buffer 4 in a final volume of 100 μL at 65 °C. After *TaqI* digest, there are three main fragment types in the libraries; uncut biotinylated fragments (no internal *TaqI* site containing 5' double biotin), the cut fragments containing 5' double biotin and the other part of the cut fragments which are non-biotinylated. Dynabeads® M-280 Streptavidin beads (Life Technologies, cat#11205D) were used to capture the biotinylated fragments as per manufacturer's instructions, hence enriching the non-biotinylated fragments in the eluate [42]. The eluate was ran through QIAquick PCR purification column (Qiagen, cat#28104) and resuspended in water for the following ligation step. The biotinylated fragments were released from the beads with an incubation step for 15 min in 30 mM d-biotin (Sigma, cat#47868, Saint Louis, MO, USA) then heating to 80 °C for 15 min. A similar approach was used to release biotinylated proteins previously [43]. We compared the non-biotinylated fragment and biotinylated fragments on an agarose gel for QC.

Ligation of GW-A1 or LA-A1 (5'-CG Overhang). The eluate containing 1 μg non-biotinylated, cut fragments were ligated to A1. The ligation was performed using a Quick Ligation™ (NEB, supplied in cat#M2200S) in the presence of 2-fold excess of GW-A1 or LA-A1 (16 pmol for 1 μg of genomic library with average fragment size of 200 bp), as per the manufacturer's instructions. The excess A1 and fragments less than 100 bp were removed with Agencourt AMPure XP Bead system (Beckman Coulter, cat# A63880, Brea, CA, USA) using the TruSeq DNA sample preparation guide with minor modifications: a ratio of 125 μL of well mixed beads with 135 μL of sample.

GW-COBRA Library PCR Amplification. Library fragments containing A1 and A2 were PCR amplified, with FlowCell-FwdP and FlowCell-RevP primers (Table 1), in a final concentration of 1 ng/μL in 25 μL using GoTaq colorless master mix (Promega, cat#M7133). The temperature cycles for the PCR were: 98 °C 3 min; 98 °C for 15 s, 65 °C for 30 s and 72 °C for 1 min, for 7 cycles; a final extension of 72 °C for 2 min.

LA-COBRA Library Linear Amplification. Ligated library fragments were *in vitro* transcribed to RNA using T7 RNA Polymerase Kit (NEB, cat#E2040S) at 37 °C for 16 h, then cleaned up with RNeasy MinElute Clean up kit (Qiagen, cat# 74204) and quantified using Quant-iT RNA assay (Life Technologies, cat#Q-33140) respectively as per manufacturers' protocol. The cDNA library was constructed using 600 ng RNA using the SuperscriptIII first-strand synthesis kit (Life Technologies, cat#18080-051) and the LADS P5 primer. Following treatment with RNase H, cDNA was made double stranded using the Klenow fragment of DNA Polymerase 1 and the P7 primer as described in [44].

Library Clean-up. Agencourt AMPure XP Beads (Beckman Coulter, cat# A63880) at a ratio of 125 μL of well mixed beads with 135 μL of sample were used to remove short fragments from both GW-COBRA and LA-COBRA libraries. Finally, the size distribution was visualized using Agilent DNA 1000 Assay in 2100 Bioanalyzer (Agilent Technologies, Los Angeles, CA, USA) using the manufacturer's protocol. Throughout the COBRA-seq library construction, appropriate products were ran on either 3% low range ultra-melting agarose gel (Bio-rad, cat# 161-3107, CA, USA) or a 4%–20% Criterion precast polyacrylamide TBE gel (Bio-rad, cat#345-0059) and stained with SYBR gold (Life Technologies, cat#S-11494).

Library Sequencing. GW- and LA-COBRA libraries of HCT116 DNA were sequenced with the 100 bp single-end Illumina HiSeq 2000 technology in a single lane each at the Australian Genome Research Facility. The COBRA library sequencing results are deposited in the CSIRO Data Access Portal which are accessible with the manuscript title search [45].

2.4. Bioinformatics and Statistics

Alignment. FASTQ reads were inspected for adaptor contamination and a set identified. This set, with the specification of at least an 8 base overlap, were subsequently removed using the fuzzy matching functionality of cutadapt [46]. Reads were also quality trimmed and length filtered by cutadapt using a quality setting of "-q 8" and minimal read length of 40 bases. The cutadapt processed reads were aligned with bwa-meth [47], a bisulfite-treated DNA tuned wrapper for the BWA-MEM aligner [48].

Clean up and Count Statistics. The alignments were further processed using a Python script. This script examined each alignment and excluded any alignment (by setting the alignment as unmapped; bit flag 4) if the alignment was identified as being a secondary alignment, did not have the expected three remaining bases of the *TaqI* restriction site, or did not align to an *in silico* identified *TaqI* site. Recognising the error-prone nature of sequencing, the three base *TaqI* site match was relaxed to a Levenshtein distance of 1. This fuzzy matching allowed a one base mismatch between the first three bases of a forward strand read and a "CGA" trimer, or the last three bases of a reverse strand read and the "TCG" trimer sequence. Forward alignments that started, or reverse alignemnts that ended on the exact genome coordinate of an *in silico* determined *TaqI* cut site were kept and tallied by *TaqI* site.

A record of the counts per *TaqI* site were exported as a bedGraph file. The file was used for visualisation in the IGV genome browser and was also imported into R for further analysis. The read cleaning procedure is provided in Figure S1.

In silico Analyses. Mapping of CpG sites, *in silico* bisulfite DNA treatment and *in silico* restriction enzyme digests on both strands of bisulfite-converted DNA were written in R using the functionality of the Bioconductor Biostrings and GenomicRanges libraries and the BSgenome.Hsapiens.UCSC.hg19 genome build library. Annotations were from the TxDb.Hsapiens.UCSC.hg19.knownGene library or downloaded from the UCSC web server via rtracklayer. Selection was further restricted to fragments greater than 70 bp as small fragments will be selected against through the library preparation process.

CpG island locations used were those in the "CpG Islands" UCSC table. CpG Shores were defined as the area flanking 2 kb of an island. There were 2,089,538 and 2,089,538 CpGs located in CpG Islands and shores respectively. The remaining CpGs were classified as CpG Ocean (24,105,864 CpGs). CpGs within 4 kb distance to transcription start sites were determined to be located in promoters (3,619,885 CpGs). The gene body CpGs was defined as those in the area between gene start and end coordinates (12,121,165 CpGs). Intergenic CpGs are those CpGs not within the genebody or TSS category (12,476,398 CpGs). Enhancer sites (205,740 CpGs) were defined as those within the start and end coordinates of FANTOM5 permissive enhancers [49].

Comparison to Other Data. The HCT-116 450K array was processed and beta values called using the R minfi library. All bioinformatics analysis scripts are deposited to the repository, along with a tutorial outlining their use [50].

3. Results and Discussion

3.1. COBRA-Seq Library Construction

The procedure for preparation of COBRA-seq libraries is outlined in Figure 1, described in detail in the Materials and Methods and Supplementary Materials and Methods. The oligonucleotides used are shown in Table 1 and Figure S2. We applied the method to prepare COBRA-seq libraries from HCT116 cell line DNA (Figure 2). Briefly, genomic DNA was fragmented and end-repaired as for normal genomic library construction. Sonicated HCT116 genomic DNA ranged between 150–500 bp (Figure 2A). After bisulfite treatment of the ligated DNA, the libraries were subjected to a minimal number of PCR cycles in order to replace uracils in the original DNA with thymines (some restriction enzymes do not cut uracil-containing DNA efficiently). The A2 reverse primer, used in the PCR step, contains two biotin groups sequentially placed on the 5' end. While the ligation of Adapter-2 did not generate a visual shift in the sonicated library size, ligation was verified by amplification with the GW-A2 Fwd and COBRA-A2-RevP or LA-A2+T7 Fwd and COBRA-A2-RevP primers pairs (Figure 2B).

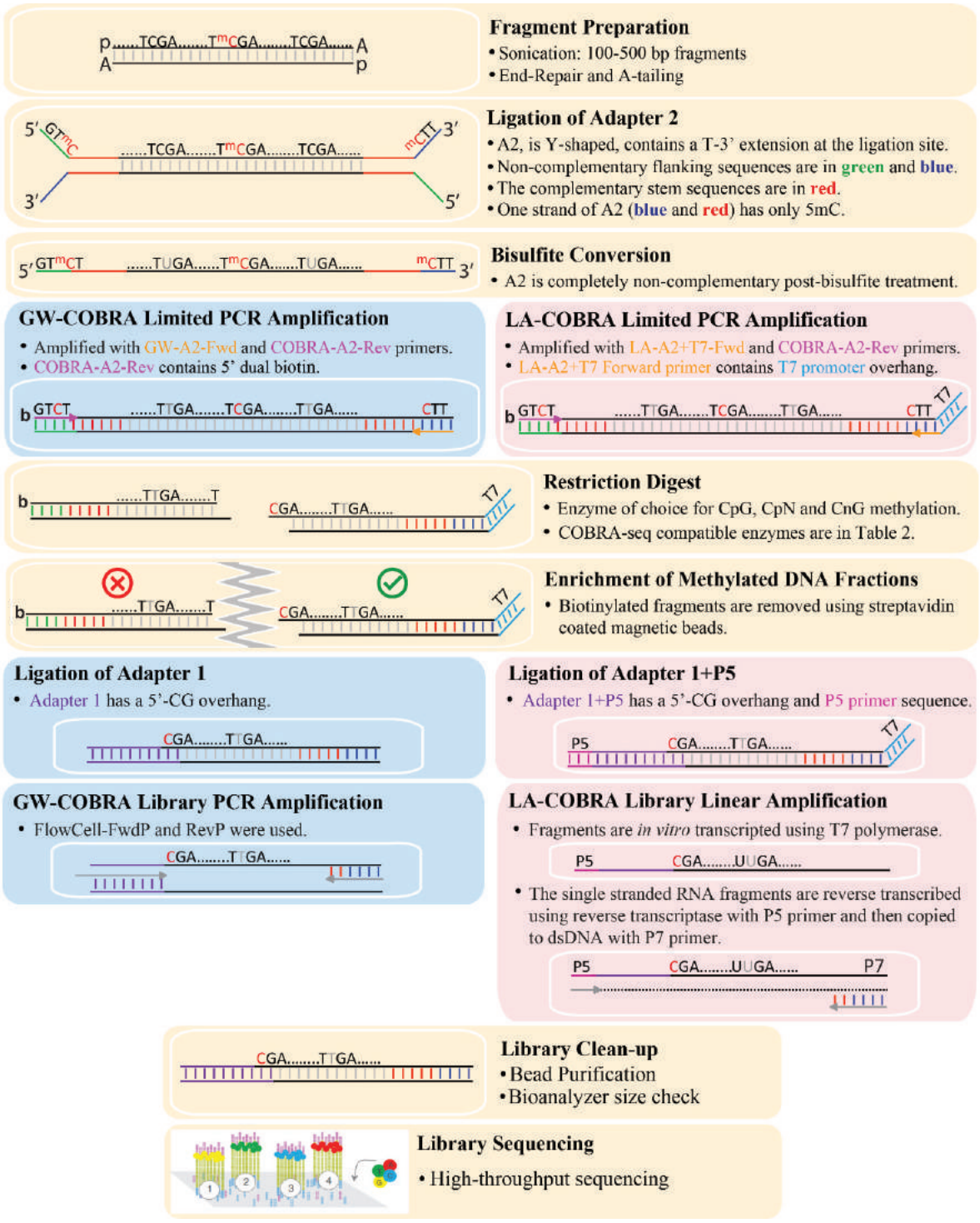


Figure 1. Flowchart of key steps in constructing COBRA-seq libraries with minor steps omitted for clarity. The common steps in both methods are shown on light yellow background. The GW-COBRA and LA-COBRA specific steps are on light blue and pink backgrounds respectively. The less common sequence abbreviations are: ^{m}C (5mC), b (two biotin groups sequentially placed on the 5'-end) and p (5' phosphorylation).

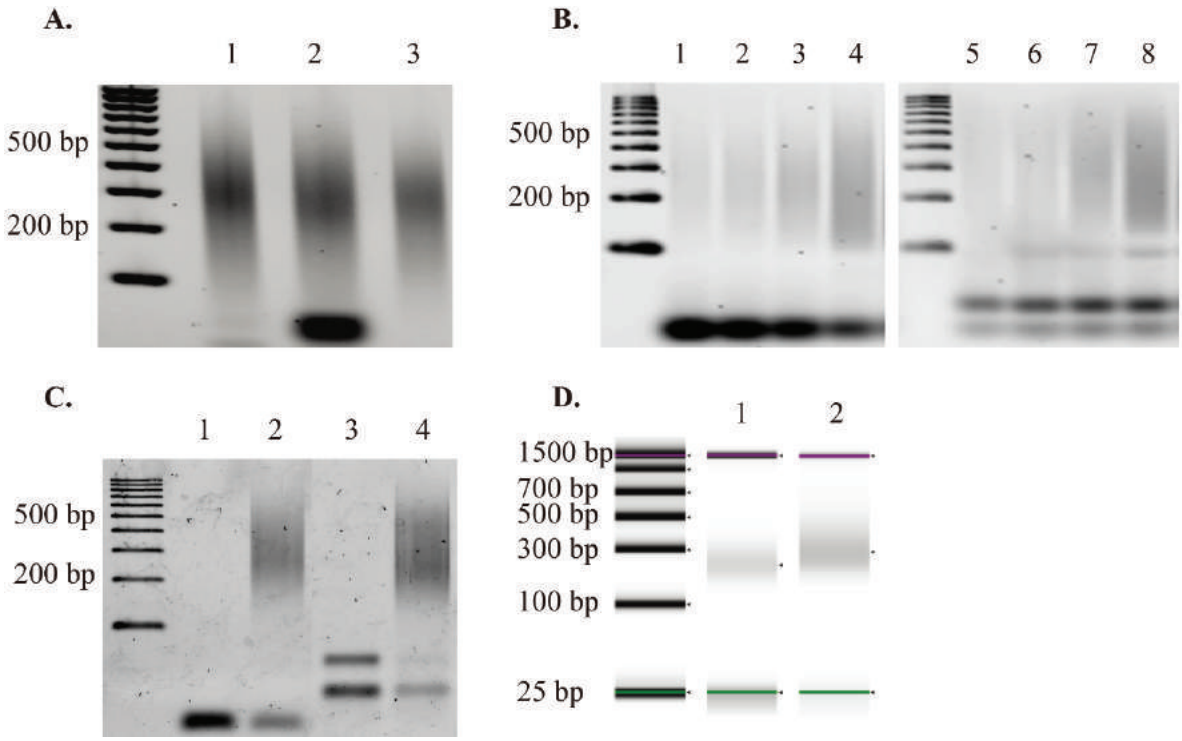


Figure 2. GW-COBRA and LA-COBRA library construction results. (**A**) Sonicated genomic DNA isolated from HCT116 cell line (lane 1), 50 ng of Adapter-2 ligated library material (lane 2). Column purified Adapter-2 ligated library material (lane 3); (**B**) PCR amplification test of bisulfite treated library materials. Lane 1, 2, 3 and 4 are produced by 6, 7, 8 and 9 cycles of PCR using GW-A2-FwdP and COBRA-A2-RevP primers from Table 1 with GW-COBRA library material as the template respectively. Lane 5, 6, 7 and 8 are produced by 6, 7, 8 and 9 cycles of PCR using LA-A2+T7-FwdP and COBRA-A2-RevP primers from Table 1 with LA-COBRA library material as the template, respectively; (**C**) Final library products: GW-COBRA and LA-COBRA methylome libraries of HCT116 DNA amplified using flowcell or LADS primer pairs respectively (Lane 2 and 4). Lane 1 and 3 are negative controls for GW-COBRA and LA-COBRA libraries, respectively; (**D**) Bioanalyzer results of GW-COBRA and LA-COBRA methylome sequencing libraries, respectively, prepared from HCT116 cell line DNA. The final library fragments ranged between 150–500 bp with an average size of 257 and 360 bp for GW-COBRA and LA-COBRA, respectively.

For the pilot study, we chose *TaqI* (5'-T/CGA-3') for complexity reduction as it has historically widespread usage in the traditional COBRA method. *TaqI* and a number of other enzymes contain CG within their recognition sequences and are therefore suitable for use with mammalian DNA that is principally methylated at CpG sites. COBRA-seq with *TaqI* covers nearly 16% of the CpG sites in the human genome. After digestion of the dual biotin-labelled material with *TaqI*, cut fragments will have one of the pair tagged with a dual biotin while the other cleaved fragment will not; all uncut molecules will contain a dual biotin tag at one end. Streptavidin-coated magnetic beads are then used to remove biotin-labelled material, the uncut molecules as well as one end of the cut fragments. The remaining fragments, containing a –CG overhang,

are ligated with Illumina Adapter 1 (GW-A1 and LA-A1) linkers, modified with a –CG extension. Therefore, COBRA-seq libraries contain only methylated CpG sites and give relative methylation level based on read counts.

For GW-COBRA, the ligated products are amplified by PCR and quantified prior to sequencing. In LA-COBRA, the Adapter 2 (A2) is modified to include a phage T7 RNA polymerase promoter site (right hand side of Figure 1). After ligation of LA-A1, T7 RNA polymerase is used to make RNA copies of the ligation products. Reverse transcriptase is used in combination with the P5 primer to copy these into cDNA. After RNaseH digestion, copies are made double-stranded using the P7 primer and the Klenow fragment of DNA polymerase 1. LA-COBRA produces full length fragments at high abundance via *in vitro* transcription and importantly removes PCR bias, which is a common side effect of amplifying bisulfite-treated DNA [51]. The final GW-COBRA and LA-COBRA library products are checked using Illumina Flowcell primer pairs or Linear Amplification for Deep Sequencing (LADS) primer pairs published in [44] (Figure 2C).

The final library fragments is ranged between 150–500 bp with an average size of 257 and 360 bp in GW-COBRA and LA-COBRA, respectively (Figure 2D). The amount of library material is recovered for each protocol averaged 450 ng for LA-COBRA and 50 ng for the PCR based GW-COBRA.

During COBRA-seq method development, we also examined the effect of limited amounts of starting material. Both GW-COBRA and LA-COBRA libraries were prepared with as little as 0.1 μg of genomic DNA of HCT116 cells which did not restrict the success of the library preparation protocol. The size distribution of the COBRA-seq library fragments was the same when the starting material was either 0.1 or 1.0 μg (Figure S3).

3.2. Number of GW-COBRA and LA-COBRA Reads and Mapping to Genome

Libraries were sequenced in a single lane each, on a 100 bp single-end in Illumina HiSeq2000 run, generating 115,097,029 and 142,245,797 million reads for GW-COBRA and LA-COBRA, respectively, with 83.1% and 84.1% of the reads mapped uniquely (Table S1). The bisulfite conversion rate was near complete (99.4%, FastQC) (Table S1). Consistent with T7 RNA polymerase having a higher nucleotide misincorporation rate than *Taq* polymerase [52], we observed LA-COBRA reads contained more sequence errors but this did not interfere with the mapping; 92.34% of the LA-COBRA library reads were mappable compared with 92.53% for GW-COBRA (Table S1). The FASTQC summary results are provided in Figure S4. Median read coverage for *TaqI* sites with at least one read were 8 and 9 for GW-COBRA and LA-COBRA, respectively. Density plots of the CpG site coverage in the HCT-116 GW-COBRA and LA-COBRA libraries and public RRBS and WGBS libraries are provided in Figure S5.

The empirical results demonstrate that representation of different DNA fragments was highly similar for LA-COBRA and GW-COBRA (Figures 3A and S4). Frequencies of read counts at specific sites were highly concordant between the two methods (Figure 3A, $R^2 = 0.905$). Because of this high concordance, we joined the datasets for comparisons with other methods (below).

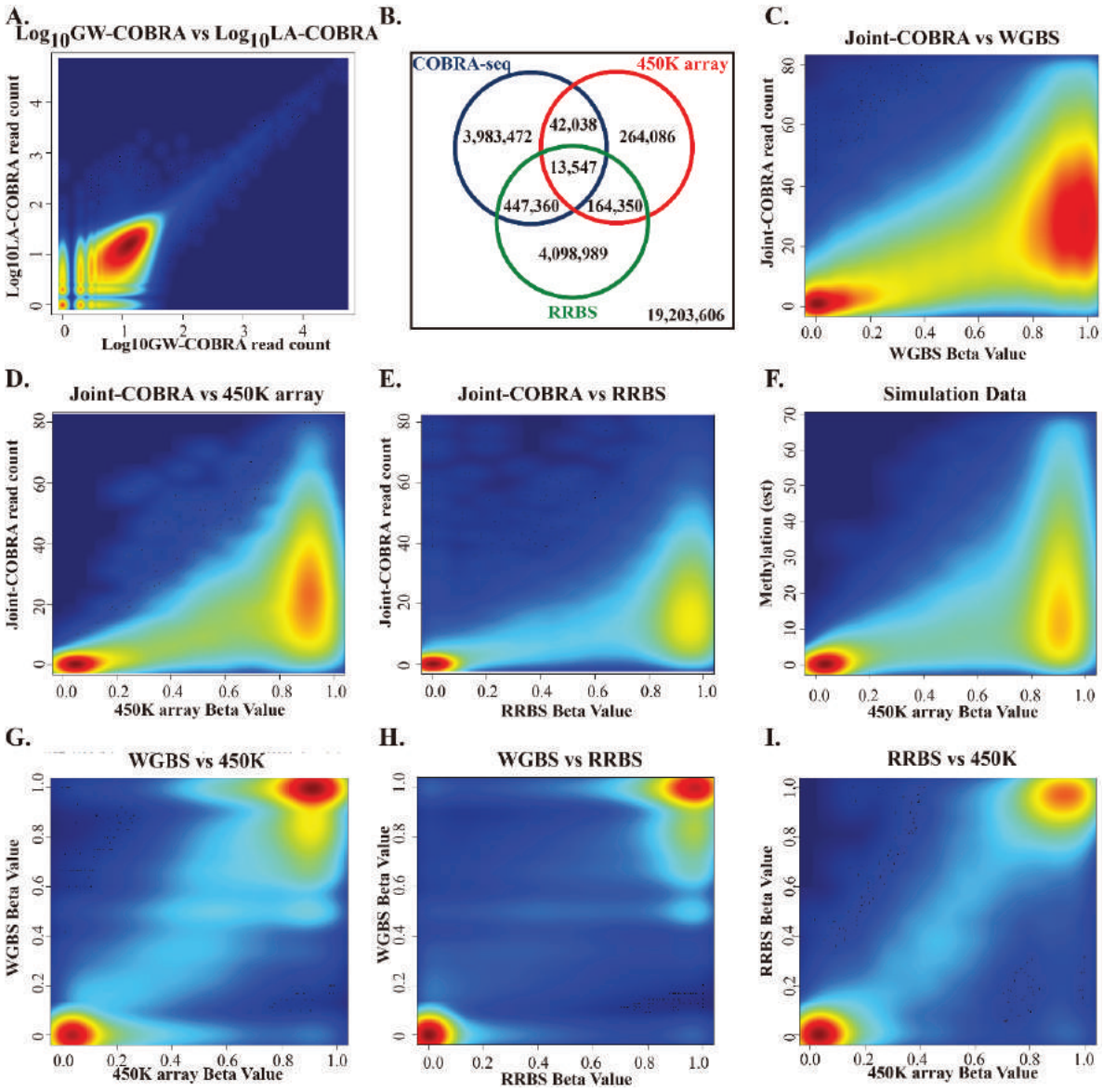


Figure 3. Empirical comparison of HCT116 COBRA-seq methylome with RRBS and 450K array data. (**A**) Comparison of the GW-COBRA and LA-COBRA log transformed read counts aligning to the same genome position (5,756,193 *TaqI* sites, $R^2 = 0.905$); (**B**) Venn diagram comparing coverage of CpG sites by the 3 methods; (**C**) Combined methylome sequencing data of GW-COBRA and LA-COBRA (Joint-COBRA; median coverage of 17 reads per strand) compared with HCT116 cell line WGBS data (902,990 shared CpG sites, $R^2 = 0.384$); (**D**) with HCT116 cell line 450K array data (31,988 shared CpG sites, $R^2 = 0.439$) and (**E**) with HCT116 cell line RRBS data (95,766 shared CpG sites; median coverage of 82.14 reads per strand, $R^2 = 0.443$); (**F**) Population mean (μ) coverage of 24.6 in M fraction only using simulated data; (**G**) Comparison of methylation scores of WGBS with 450K (463,300 shared CpG sites); (**H**) WGBS wish RRBS (978,735 shared CpG sites); (**I**) RRBS with 450K (46,374 shared CpG sites).

3.3. Comparison of HCT116 COBRA-Seq with RRBS and 450K Array Data

We compared our HCT116 COBRA-seq methylome data with WGBS data on HCT-116 cells and data obtained from two widely used methods that also sample a proportion of the genome: Illumina Infinium HumanMethylation450 BeadChip arrays (450K arrays) [30] and Reduced Representation Bisulfite Sequencing (RRBS) [53]. Each method targets a different subset of the methylome, with some overlapping CpG sites (Figure 3B). On microarrays, a very large number of molecules bind to each array feature, giving an analogue estimation of the population mean. In particular, the 450K array is known for robustness and has been a popular choice due to its affordable price and high precision. The 450K beta (proportion of methylation) values can be considered as an accurate quantification across ~480,000 (1.7%) CpG sites. HCT116 450K methylation data is made available in our repository [50].

RRBS is an efficient high throughput methylome method that samples a population of restriction fragments isolated following digestion with the methylation-insensitive enzyme, *MspI*. This DNA population, encompassing 16.7% of CpG sites, is then bisulfite-treated and sequenced, and so provides reads of both methylated and unmethylated sequences. We used HCT116 cell line WGBS and RRBS methylation data (GSM1465024 and GSM919980) arising from published studies [54,55].

We also considered simulated data to provide an overview of the expected data distribution and concordance at different levels of sequence coverage with two types of methylome methods: one that samples both methylated (M) and unmethylated (U) fractions of the genome such as RRBS (M + U method in Figure S6, left panel) and methods that enrich for methylated (M) fractions like COBRA-seq (M only method in Figure S6, right panel).

We used empirical COBRA-seq over-dispersion estimates and 450K CpG site beta distribution data to inform the simulation. Across different depths of hypothetical sequencing coverage, 200,000 sample read counts were modelled, conditional upon the distribution of beta values from an HCT-116 450K array (Figure S6). For M + U methods, as read coverage increases, the ratio of methylated to unmethylated read counts will become more accurate and estimated beta will converge to the true beta. As beta is a ratio, at low coverage, under- or over-sampling of the M or U fraction will result in high imprecision in beta value estimation near the top (1.0) and bottom (0.0) of the beta value range. For the M only method, the read count will increase conditionally on both coverage and methylation rate, except where the CpG site is completely unmethylated; in this instance, increased coverage will not yield increased read counts. So, instead of the beta-binomial distribution of M + U methods, a zero-inflated count-based distribution will be observed, which is analogous to RNA-seq count data. CpG site state is not independent of other sites, so the distribution is not Poisson, but a family of Poisson distributions (negative binomial). Details of the modeling process and data are discussed in more detail in Supplementary Methods and Materials and Figure S6. In summary, we find that M + U methods have poor estimates at betas near 0 or 1 with low read coverage, but estimated beta rapidly forms a good correlation with true beta with increased read coverage. For M only methods, read coverage and methylation rate are confounded so estimates of partial or complete methylation are difficult. Instead, M only methods are best suited to situations where there are biological replicates. This allows modelling the within-group variance and the effects of library size, allowing coverage and methylation rate to be decoupled. The simulation suggests M only methods, like COBRA-seq, may have advantages over M + U methods in detecting lack of methylation between replicated experimental groups at low read coverage.

When considering read coverage requirements, our simulations suggest a 10 to 15-fold coverage with M + U methods is sufficient for reasonable beta estimates. An empirical comparison mapping 200,000 HCT-116 450K array beta values to WGBS beta values gives much the same estimate (Supplementary Figure S7). For M-only methods, we suggest a 10 to 15-fold coverage is sufficient and for extra sensitivity the addition of extra samples should be considered over additional read coverage.

We also wished to make an empirical comparison between COBRA-seq and other methods. For this, we plotted data for CpG sites shared between Joint-COBRA and WGBS data (902,990 shared CpG sites with at least 10-fold WGBS coverage), 450K array data (31,988 shared CpG sites) and RRBS data (95,766 shared CpG sites) (Figure 3C–E). The empirical data was a good fit with our simulated expectation (Figure 3D compared with 3F). Higher COBRA-seq read counts were associated with higher methylation rate (as determined by the comparison method). In particular, it should be noted that sites determined to be unmethylated by WGBS, RRBS or 450K array did not yield COBRA-seq reads, demonstrating the underlying principle of our method is sound.

For a closer inspection across individual genome locations, we visualized the HCT116 cell line methylation profiles determined by GW-COBRA, LA-COBRA, WGBS, RRBS and 450K methods using the integrated genome viewer (IGV). Profiles of methylation across two genes, *BCAT1* and *EHD3*, which have been previously identified as biomarkers for colorectal cancer [42,56], are shown in Figure 4A,B. These regions are zoomed in to display a number of bases only in Figure S8. Additionally, more examples of the comparative methylation profiles across panel of genes (*SEPT9*, *MGMT*, *SLC6A15* and *FGF5*), highly methylated in colorectal cancer [57], are provided in Figure S9.

The methylation profiles generated by GW-COBRA and LA-COBRA have high concordance (Figure 4A,B; *ii*. and *iii*.), and also correlated highly with WGBS, RRBS and 450K array data (Figure 4A,B; *i*.) exemplified at the CpG rich promoter region of the *BCAT1* gene (Figure 4A). COBRA-seq provides a more even genomic distribution, compared to the bias towards CpG rich transcription start sites observed in 450K arrays and the RRBS method (Figures 4 and S8; *iv*). The difference in locality of cutting between *TaqI* (COBRA-seq) and *MspI* (RRBS) can be explained by the GC-richness of the restriction site.

3.4. COBRA-Seq Features

As COBRA-seq fragments should start with the 3' end of the enzyme recognition site and should align to restriction sites in the genome, these factors can be used to clean up the quality of the sequencing data and alignments. The read cleaning procedure is visualized in IGV and reported in Figure S1 with a brief description of the procedure in the Materials and Methods. The cleaning procedure was conservative, keeping only reads that aligned to *in silico* predicted *TaqI* sites and requiring the reads started with *TaqI* site nucleotides, with a one base mismatch tolerance. We saw evidence that around half of the discarded reads were of high quality and had a *TaqI* site start (Table S2), but they aligned to *TaqI* sites not predicted in the reference genome. We also only considered methylation of the CpG site at the start of the read although it is possible to consider methylation at CpG sites within the read.

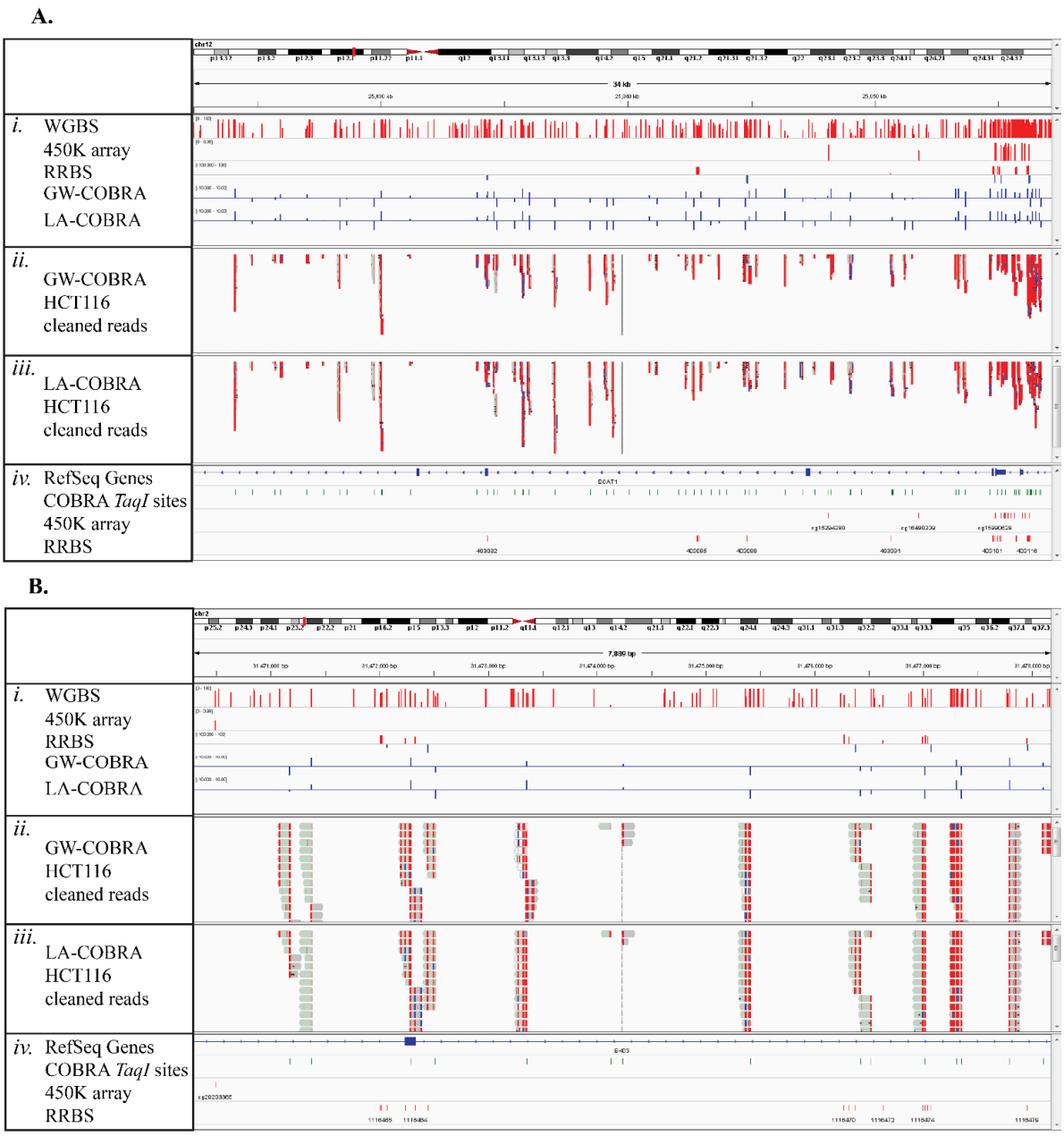


Figure 4. IGV screenshots on the selected colon cancer associated genes: (**A**) *BCAT1* and (**B**) *EHD3*. *i*. Methylation levels determined by WGBS, 450K arrays, RRBS GW-COBRA and LA-COBRA respectively; *ii*. Genome coverage and stacking of cleaned GW-COBRA reads; *iii*. Genome coverage and stacking of cleaned LA-COBRA reads; *iv*. Genomic locations of accessible COBRA *TaqI* sites, 450K array probes and accessible RRBS sites.

As COBRA-seq reads arise from bisulfite-treated DNA, it is possible to examine methylation at cytosines sites within a COBRA-seq read, however this has caveats. By design, reads depend of methylation of a cytosine within the enzyme recognition sequence. Therefore any observations at cytosines sites within reads are conditional upon methylation within the enzyme recognition sequence. While this

confounds conventional methylation estimation, it is useful for examination of detailed methylation profiles, heterogeneity and of local co-methylation, which is a well-documented property of DNA methylation. Also, understanding the methylation status of cytosines within reads may be used qualitatively as evidence for determining the best sites for biomarker assay (primer) design.

One of the features of COBRA-seq is that a variety of enzymes can be used to assess 5mC in different contexts. Table S2 provides a selection of restriction enzymes that are suitable for employing COBRA-seq in different sampling depths and contexts. This table includes examples of 4 and 6-base cutter enzymes that cleave to leave a CG 5' overhang and are particularly useful for studies of CpG methylation (*AclI*, *BstBI*, *ClaI*, *HpyCH4IV* and *TaqI*). Enzymes containing CNG within their recognition sequence such as *HpyCH4III* and *Hpy188I* are potentially useful for studies of plant CpNpG methylation.

Some enzymes contain a cytosine at the end of their recognition site and hence can be used to monitor methylation independent of the nature of the neighboring nucleotide. For instance, *Sau3AI*, with the 5'-/GATC-3' recognition site, is a 4-base cutter. However, in the context of COBRA-seq, *Sau3AI* acts like a 5-base cutter as the recognition site can only be preserved if the 3' cytosine is methylated. In vertebrate DNA, this would typically be in a CpG context, whereas in plants, insects or any other genome, it can identify adjacent methylation in any context (CpN methylation). Similarly, *CviQI* and *EcoRI*, with the 5'-G/TAC-3' and 5'-G/AATTC-3' recognition sites respectively, are good choices for CpN methylation profiling in organisms such as insects. There are also enzymes that contains multiple CpG sites at their recognition sites such as *HinP1I* (5'-G/CGC-3') and *Hpy99I* (5'-CGACG/-3'). Their recognition sites can only be preserved if both of the CpGs are methylated therefore the number of sites observed in the actual sequencing results can be lower than the *in silico* digest estimations. These enzymes would be a good selection for discovery of highly methylated regions, or for co-methylation studies.

The number of potentially addressable methylation sites within a genome is dependent on the cutting frequency of the enzyme(s) used. Estimates of addressable CpG sites with COBRA-seq ranges from 1.4% to 16% of the sites in the human genome are provided in Table S3 and explained further in Materials and Methods. Among COBRA-seq compatible enzymes; *HpyCH4IV* reaches the highest proportion of CpGs (Tables S2 and S3). It is important to note that linkers for ligation to the restriction enzyme-cut ends must be modified as necessary to provide complementarity to the fragment overhangs. Fortunately a number of useful enzymes provide a CG-5' overhang (Table S2). COBRA-seq libraries can be multiplexed to increase the percentage of accessible sites, or alternatively tuned to reduce sequencing costs with the use of 6-base cutters instead of 4-base cutters to interrogate fewer sites. A typical 6-base cutter will have around 1/8 of the sites of a 4-base cutter. It is possible to prepare libraries using more than one enzyme, but we recommend the users perform separate digestions for each enzyme of choice before pooling and ligating to generate libraries as concurrent multi-enzyme digests will result in small fragment sizes which may cause loss of potential CpG sites during purification steps targeting removal of excess Adapters.

Additionally, for organisms without a reference genome, COBRA-seq is complementary for use with Restriction-site associated DNA sequencing (RADseq), a restriction enzyme-based method which reduces genome complexity across target genomes for SNP discovery [58]. It is possible to use the same restriction enzyme sites, RADseq for SNP discovery and COBRA-seq for methylation quantitation, thereby providing a convenient platform for allele-specific methylation discovery on non-model organisms.

The COBRA-seq variation utilizing *TaqI* (5'-T/CGA-3') examined here has limited genomic bias due to the recognition site GC content and this in turn, yields relatively unbiased distribution across genomic

annotations in human DNA: 4% and 6% coverage of CpG islands and shores respectively (Figure 5B and Table S3). It covers approximately 11.9% of genome-wide CpG sites within promoters and also similar coverage for features with fewer CpGs, such as enhancers (12.3%) and DNaseI hypersensitive (12.9%) sites (Figure 5 and Table S3).

In the COBRA-seq protocol described here we have used the standard protocol of ligation of adapters to sheared double-stranded DNA. However, it is also possible, and may be advantageous, to adapt library preparation methods that used single-stranded bisulfite-treated DNA as a starting point, e.g., Swift Accel-NGS-Methyl-Seq or Illumina EpiGnome/Tru-seq Methylation libraries. This could be particularly the case for DNA such as isolated from formalin-fixed tissue samples where DNA may be damaged or isolated single-stranded. Pre-treatment to repair damaged DNA (Illumina Restoration kit, Illumina, San Diego, CA, USA) may also improve the quality of such libraries.

3.5. Comparison of COBRA-Seq Features with Other Methods of Methylome Sampling

A range of different technologies for studying DNA methylation on a genome scale have become available, each with inherent differences in resolution, coverage and biases [59,60]. COBRA-seq quality control, read alignment and data visualization are the broadly the same as any WGBS or RRBS protocol [61]. However, methylation scoring is by counts and the determination of differentially methylated cytosines bears strong resemblance to the analysis of RNA-seq data.

In Table 2, we provide comparative summaries of features of COBRA-seq and a number of other methylome sampling methods for DNA methylation analysis.

Specifications such as complexity reduction type and whether they enrich for methylated fractions were compared (Table 2). Genome complexity is reduced by affinity capture (MBDCap-Seq and MeDIP-Seq), restriction digest (COBRA-seq, RRBS, Methyl-Seq, HELP-Seq, CHARM and DREAM) and hybridization capture (Nimblegen SeqCap and Agilent SureSelect). MBDCap-Seq, MeDIP-Seq and COBRA-seq are the only methods that enrich for methylated fragments and hence reduce the sequencing cost, (by yielding a high ratio of CpG information per read sequenced), with the caveat that absolute methylation level estimation is traded for relative methylation level estimation. However, COBRA-seq has an advantage over these other enrichment methods, as the digestion-based enrichment step does not show the same dependence on methylation and CpG density. Therefore, COBRA-seq capture is more uniform across the genome, making it a suitable choice for interrogating regions or genomes of low methylation density. Another system for affinity-based capture of methylated DNA that we have described previously is SuBLiME [42]. Here capture is done after copying of bisulfite-converted DNA to incorporate biotin-dG opposite unconverted meC bases (or biotin-dC in a subsequent copying round). SuBLiME can efficiently capture methylated cytosines in different sequence contexts and can be tuned to different levels of genome coverage [42]. It should be noted that copy number variation in a sample has a multiplicative effect on input DNA. This is observed as a local change in the number of sequenced reads or probe signal intensity on a microarray. With methylation enrichment-based methods, methylation estimates are confounded with copy number and it is only possible to mathematically deconvolute methylation from copy number across large genomic regions.

Figure 5. Comparison of CpG coverage of COBRA-seq with WGBS and commonly used methylome sampling methods by genomic regions, as pie charts or bar diagram for proportion of enhancer sites (WGBS covers a total of 205,740 CpGs in enhancers). (**A**) WGBS covers 28,217,448 CpG sites; (**B**) COBRA-seq with *TaqI*, *HpyCH4IV*, *BstuI* and CviQ*I* covers 15.9%, 16.9%, 4.3%, and 5.2% of the total CpG sites respectively; **C.** 450K arrays, RRBS, SuBLiME and DREAM cover 1.7%, 16.7%, 75.8%, and 1.8% of the total CpG sites respectively.

Table 2. Qualitative comparison of selected methylome methods (*are the preferred genomic input, low input will increase variation due to sampling).

Methylome Methods	Complexity Reduction Type	M or M + U Fraction	Methylome Sampling (Yes/No)	Favour of Enrichment Towards	Comments
WGBS [1]	N/A	M + U	N/A	N/A	High cost, can detect non-CpG methylation, genomic input* (0.05–0.1 µg).
MBDCap-Seq [34,62], MIRA-Seq [38]	Affinity capture	M	No	CpG-rich	Dependent on CpG density, effected by salt concentration, covers about 18% of the CpGs [24], 28,500 CpG islands [34], DNA input = 0.2–1 µg.
MeDIP [63], MeDIP-Seq [36]	Affinity capture	M	No	CpG-rich	Bias towards 5mC-rich regions, Captures single-stranded DNA, prone to technical variability, coverage is read-depth dependent, input = 0.15–5 µg.
SuBLiME [42]	Methylated cytosine capture	M	Yes	CpG-rich and poor	Substitutes biotin-14-dCTP or biotin-14-dGTP at the position of the 5mC in bisulfite treated DNA, input = 2 µg.
COBRA-seq	Restriction digest	M	Yes	CpG-rich and poor	Provides relative methylation levels like MeDIP/MBD-Cap, 1%–17% of the CpGs for single digest, can detect non-CpG methylation, input = 0.1–1 µg.
Nimblegen SeqCap [39]	Hybridization capture	M + U	Yes	CpG-rich and poor	"Off-the-shelf" version for human genome only/similar regions covered as 450K array, can be customized [39], can detect non-CpG methylation, input = 0.5–1 µg.
Agilent SureSelect [64]	Hybridization capture	M + U	Yes	CpG-rich and poor	Available for human genome only, covers 3.7 million CpG sites, input = 0.5 µg.
450K array [30,65]	Microarray	M + U	Yes	CpG-rich	Arrays comes in 12 per slide, available only for humans, not readily customized, input = 0.5–1 µg.
RRBS [35,53]	Restriction digest	M + U	Yes	CpG-rich/medium	Can detect non-CpG methylation, input = 0.1–0.3 µg.
Methyl-Seq [33], HELP-Seq [66]	Restriction digest	M + U	Yes	CpG-rich	Assesses 0.25 to 1.3 million CCGG sites in human genome by difference in read fractions in *HpaII vs*. *MspI* libraries, input = 0.01–0.1 µg.
CHARM [32,67]	Restriction digest	M + U	Yes	CpG-rich and poor	Array-based and available for human, mice and rat, assesses 3.5 to 7 million CpG sites, input = 5 µg.
DREAM [37]	Restriction digest	M + U	Yes	CpG-rich	Assesses methylation at ~0.15 million sites in human genome by sequential *SmaI*/*XmaI* digestion and library sequencing, input = 5 µg.

In samples with a heterogenous cell population, for example, solid tumour tissue with a high degree of non-neoplastic cells, COBRA-seq presents the same concerns as other methods. For example, with an absolute methylation method, a 70% pure tumour will appear as having a maximum beta of 0.7 (given sufficient read coverage), while for COBRA-seq or other enrichment methods, this will translate to a reduced rate (lambda). Unmethylated regions will still have no reads, but methylated regions will have a reduced rate.

To sequence formalin-fixed paraffin embedded (FFPE) samples, we recommend that researchers optimize the initial sonication step to have fragments between 100–500 bp. FFPE-induced DNA adducts are not a concern for enzyme cutting due to the preceding PCR pre-amplification step. This is further explained in the step-by-step protocol submitted as Supplementary Methods.

The relative coverage of different sequence features by COBRA-seq and selected methylome sampling methods is shown in Figure 5 and Supplementary Table S3. It can be seen that COBRA-Seq using enzymes containing one CpG site provides relative coverage that is very similar to the overall distribution of features in the human genome—CpG islands, CpG shores, promoters, gene bodies and intergenic regions as well as enhancer regions. Use of enzymes containing two CpG sites such as BstUI focuses COBRA-Seq toward more CpG rich regions, islands, shores and promoters. RRBS and DREAM are both strongly biased toward coverage of CpG-rich regions and promoters, but notably, RRBS also covers a high proportion of enhancers. Recently a large *in silico* survey was used to characterise the properties of the RRBS method with other methylation insensitive enzymes [68]. Also, a dual-enzyme RRBS method (dRRBS) using *ApeKI* (5'-G/CWGC) as well as *MspI* has been described [69]. The addition of *ApeKI* offers more representative coverage in low-CG regions without overly fragmenting CpG-rich DNA. The increased coverage by dRRBS is characterized from an *in silico* and empirical standpoint. Because of their current designs, methods relying on hybridization selection of targets, such as 450K arrays and Agilent and Roche/Nimblegen capture systems, are strongly biased toward promoter and CpG-rich regions. However, such capture systems offer the opportunity for designs to evolve and to be specifically targeted to regions of high interest.

4. Conclusions

Building upon the previously described COBRA method, the COBRA-seq method described here makes it possible to measure DNA methylation at a large number of specific sites distributed across the genome. Methylome sampling methods broadly utilize two approaches to select the genomic fraction to be analysed. Some methods sample independently of methylation status and measure both unmethylated and methylated cytosines at each CpG site (e.g., 450K array, SeqCap, RRBS and others mentioned in Table 2). Alternatively, other methods specifically enrich the methylated fraction and determine the relative number of reads or level of methylation at different sites across the genome (e.g., MeDIP-Seq, MBDCap-Seq, COBRA-seq). This latter approach is advantageous to reduce cost with less wasted sequencing space and is suited to identification of differences between different samples. Its main drawback is that methylation levels determined are relative, rather than absolute, and can be influenced by the relative efficiency of capture and amplification of different target regions. Because capture of methylated DNA in COBRA-seq is based on restriction digestion and ligation at individual sites, it does not show the same dependence on methylation density as MeDIP-Seq and MBDCap-Seq. making it a suitable choice for interrogating regions or genomes of low methylation density.

In this study, we demonstrated the feasibility of COBRA-seq for generation and analysis of genome-scale DNA methylation profiles at nucleotide resolution and compared the relative methylation measures of COBRA-seq with other methylome sampling methods. COBRA-seq data shows high concordance with WGBS, RRBS and 450K array data, although, these methods vary in the proportions of the genome sampled. *TaqI*-based COBRA-seq does not disproportionally enrich specific genomic features such as promoters and CpG islands and compared to 450K probe density and RRBS *MspI* restriction sites, COBRA-seq *TaqI* sites are uniformly distributed across the genome. Moreover, COBRA-seq is highly adaptable as the principle is compatible with various enzymes, which provides choice for the efficient mapping of 5mCs in any genomic context, including low methylation genomes such as those of insects. The count data produced by COBRA-seq is not well suited for examining methylation within a single sample, instead it is best suited for examining the differences between groups of samples and with modelling of the variance and distribution, *p*-values may be generated in much the same manner as RNAseq count data.

In conclusion, COBRA-seq method has proven to be highly sensitive and samples uniformly across the genome giving better coverage for biomarker discovery studies. COBRA-seq is a unique alternative method to study methylation in low CpG regions such as in enhancers and CpG poor promoters providing unique advantage to the users to finetune the project costs according to budget.

Acknowledgments

The study is part of the EpiSCOPE project supported by the Science and Industry Endowment Fund (Australia), grant RP03-064 and NHMRC project grant #108814. Hilal Varinli is supported by iMQRES and CSIRO OCE Doctoral scholarships, Susan J. Clark is supported by a NHMRC Fellowship #1963559. We thank Glenn Brown, Sue Mitchell and Blake Ellis for their support. Last but not least we thank Michael Gillings, and David Midgley for their critical review of the manuscript.

Author Contributions

Jason P. Ross, Peter L. Molloy and Hilal Varinli designed the experiment. Hilal Varinli conducted the experiments. Hilal Varinli and Jason P. Ross wrote the manuscript. Jason P. Ross, Aaron L. Statham and Hilal Varinli analysed the sequencing results. Peter L. Molloy and Susan J. Clark reviewed/edited manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Lister, R.; Pelizzola, M.; Dowen, R.H.; Hawkins, R.D.; Hon, G.; Tonti-Filippini, J.; Nery, J.R.; Lee, L.; Ye, Z.; Ngo, Q.M.; *et al*. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **2009**, *462*, 315–322.
2. Novik, K.L.; Nimmrich, I.; Genc, B.; Maier, S.; Piepenbrock, C.; Olek, A.; Beck, S. Epigenomics: Genome-wide study of methylation phenomena. *Curr*. *Issues Mol*. *Biol*. **2002**, *4*, 111–128.
3. Rakyan, V.K.; Hildmann, T.; Novik, K.L.; Lewin, J.; Tost, J.; Cox, A.V.; Andrews, T.D.; Howe, K.L.; Otto, T.; Olek, A.; *et al*. DNA methylation profiling of the human major histocompatibility complex: A pilot study for the human epigenome project. *PLoS Biol*. **2004**, *2*, e405.

4. Cokus, S.J.; Feng, S.; Zhang, X.; Chen, Z.; Merriman, B.; Haudenschild, C.D.; Pradhan, S.; Nelson, S.F.; Pellegrini, M.; Jacobsen, S.E. Shotgun bisulphite sequencing of the arabidopsis genome reveals DNA methylation patterning. *Nature* **2008**, *452*, 215–219.
5. Feng, S.; Cokus, S.J.; Zhang, X.; Chen, P.Y.; Bostick, M.; Goll, M.G.; Hetzel, J.; Jain, J.; Strauss, S.H.; Halpern, M.E.; *et al*. Conservation and divergence of methylation patterning in plants and animals. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **2010**, *107*, 8689–8694.
6. Lister, R.; O'Malley, R.C.; Tonti-Filippini, J.; Gregory, B.D.; Berry, C.C.; Millar, A.H.; Ecker, J.R. Highly integrated single-base resolution maps of the epigenome in arabidopsis. *Cell* **2008**, *133*, 523–536.
7. Stroud, H.; Greenberg, M.V.; Feng, S.; Bernatavichute, Y.V.; Jacobsen, S.E. Comprehensive analysis of silencing mutants reveals complex regulation of the arabidopsis methylome. *Cell* **2013**, *152*, 352–364.
8. Zemach, A.; McDaniel, I.E.; Silva, P.; Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **2010**, *328*, 916–919.
9. Zilberman, D.; Gehring, M.; Tran, R.K.; Ballinger, T.; Henikoff, S. Genome-wide analysis of arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat*. *Genet*. **2007**, *39*, 61–69.
10. Li, X.; Zhu, J.; Hu, F.; Ge, S.; Ye, M.; Xiang, H.; Zhang, G.; Zheng, X.; Zhang, H.; Zhang, S.; *et al*. Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics* **2012**, doi:10.1186/1471-2164-13-300.
11. Jeon, J.; Choi, J.; Lee, G.W.; Park, S.Y.; Huh, A.; Dean, R.A.; Lee, Y.H. Genome-wide profiling of DNA methylation provides insights into epigenetic regulation of fungal development in a plant pathogenic fungus, magnaporthe oryzae. *Sci*. *Rep*. **2015**, doi:10.1038/srep08567.
12. Liu, S.Y.; Lin, J.Q.; Wu, H.L.; Wang, C.C.; Huang, S.J.; Luo, Y.F.; Sun, J.H.; Zhou, J.X.; Yan, S.J.; He, J.G.; *et al*. Bisulfite sequencing reveals that aspergillus flavus holds a hollow in DNA methylation. *PLoS ONE* **2012**, *7*, e30349.
13. Selker, E.U.; Tountas, N.A.; Cross, S.H.; Margolin, B.S.; Murphy, J.G.; Bird, A.P.; Freitag, M. The methylated component of the neurospora crassa genome. *Nature* **2003**, *422*, 893–897.
14. Salzberg, A.; Fisher, O.; Siman-Tov, R.; Ankri, S. Identification of methylated sequences in genomic DNA of adult drosophila melanogaster. *Biochem*. *Biophys*. *Res*. *Commun*. **2004**, *322*, 465–469.
15. Takayama, S.; Dhahbi, J.; Roberts, A.; Mao, G.; Heo, S.J.; Pachter, L.; Martin, D.I.; Boffelli, D. Genome methylation in D. melanogaster is found at specific short motifs and is independent of DNMT2 activity. *Genome Res*. **2014**, *24*, 821–830.
16. Lyko, F.; Foret, S.; Kucharski, R.; Wolf, S.; Falckenhayn, C.; Maleszka, R. The honey bee epigenomes: Differential methylation of brain DNA in queens and workers. *PLoS Biol*. **2010**, *8*, e1000506.
17. Jiang, L.; Zhang, J.; Wang, J.J.; Wang, L.; Zhang, L.; Li, G.; Yang, X.; Ma, X.; Sun, X.; Cai, J.; *et al*. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* **2013**, *153*, 773–784.
18. Potok, M.E.; Nix, D.A.; Parnell, T.J.; Cairns, B.R. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **2013**, *153*, 759–772.
19. Habibi, E.; Brinkman, A.B.; Arand, J.; Kroeze, L.I.; Kerstens, H.H.; Matarese, F.; Lepikhov, K.; Gut, M.; Brun-Heath, I.; Hubner, N.C.; *et al*. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **2013**, *13*, 360–369.
20. Meissner, A.; Mikkelsen, T.S.; Gu, H.; Wernig, M.; Hanna, J.; Sivachenko, A.; Zhang, X.; Bernstein, B.E.; Nusbaum, C.; Jaffe, D.B.; *et al*. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **2008**, *454*, 766–770.

21. Doherty, R.; Couldrey, C. Exploring genome wide bisulfite sequencing for DNA methylation analysis in livestock: A technical assessment. *Front*. *Genet*. **2014**, doi:10.3389/fgene.2014.00126.
22. Lloyd, V.K.; Brisson, J.A.; Fitzpatrick, K.A.; McEachern, L.A.; Verhulst, E.C. The epigenetics of emerging and nonmodel organisms. *Genet*. *Res*. *Int*. **2012**, doi:10.1155/2012/491204.
23. Perfus-Barbeoch, L.; Castagnone-Sereno, P.; Reichelt, M.; Fneich, S.; Roquis, D.; Pratx, L.; Cosseau, C.; Grunau, C.; Abad, P. Elucidating the molecular bases of epigenetic inheritance in non-model invertebrates: The case of the root-knot nematode meloidogyne incognita. *Front*. *Physiol*. **2014**, doi:10.3389/fphys.2014.00211.
24. Stirzaker, C.; Taberlay, P.C.; Statham, A.L.; Clark, S.J. Mining cancer methylomes: Prospects and challenges. *Trends Genet*. **2014**, *30*, 75–84.
25. Frommer, M.; Mcdonald, L.E.; Millar, D.S.; Collis, C.M.; Watt, F.; Grigg, G.W.; Molloy, P.L.; Paul, C.L. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **1992**, *89*, 1827–1831.
26. Clark, S.J.; Harrison, J.; Paul, C.L.; Frommer, M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*. **1994**, *22*, 2990–2997.
27. Ziller, M.J.; Gu, H.C.; Muller, F.; Donaghey, J.; Tsai, L.T.Y.; Kohlbacher, O.; de Jager, P.L.; Rosen, E.D.; Bennett, D.A.; Bernstein, B.E.; *et al*. Charting a dynamic DNA methylation landscape of the human genome. *Nature* **2013**, *500*, 477–481.
28. Deaton, A.M.; Bird, A. CpG islands and the regulation of transcription. *Genes Dev*. **2011**, *25*, 1010–1022.
29. Glastad, K.M.; Hunt, B.G.; Goodisman, M.A.D. Evolutionary insights into DNA methylation in insects. *Curr*. *Opin*. *Insect Sci*. **2014**, *1*, 25–30.
30. Sandoval, J.; Heyn, H.A.; Moran, S.; Serra-Musach, J.; Pujana, M.A.; Bibikova, M.; Esteller, M. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* **2011**, *6*, 692–702.
31. Bibikova, M.; Lin, Z.; Zhou, L.; Chudin, E.; Garcia, E.W.; Wu, B.; Doucet, D.; Thomas, N.J.; Wang, Y.; Vollmer, E.; *et al*. High-throughput DNA methylation profiling using universal bead arrays. *Genome Res*. **2006**, *16*, 383–393.
32. Irizarry, R.A.; Ladd-Acosta, C.; Carvalho, B.; Wu, H.; Brandenburg, S.A.; Jeddeloh, J.A.; Wen, B.; Feinberg, A.P. Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res*. **2008**, *18*, 780–790.
33. Brunner, A.L.; Johnson, D.S.; Kim, S.W.; Valouev, A.; Reddy, T.E.; Neff, N.F.; Anton, E.; Medina, C.; Nguyen, L.; Chiao, E.; *et al*. Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res*. **2009**, *19*, 1044–1056.
34. Serre, D.; Lee, B.H.; Ting, A.H. Mbd-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res*. **2010**, *38*, 391–399.
35. Ziller, M.J.; Muller, F.; Liao, J.; Zhang, Y.Y.; Gu, H.C.; Bock, C.; Boyle, P.; Epstein, C.B.; Bernstein, B.E.; Lengauer, T.; *et al*. Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet*. **2011**, *7*, e1002389.
36. Taiwo, O.; Wilson, G.A.; Morris, T.; Seisenberger, S.; Reik, W.; Pearce, D.; Beck, S.; Butcher, L.M. Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat*. *Protoc*. **2012**, *7*, 617–636.
37. Jelinek, J.; Liang, S.D.; Lu, Y.; He, R.; Ramagli, L.S.; Shpall, E.J.; Estecio, M.R.H.; Issa, J.P.J. Conserved DNA methylation patterns in healthy blood cells and extensive changes in leukemia measured by a new quantitative technique. *Epigenetics* **2012**, *7*, 1368–1378.

38. Jung, M.; Kadam, S.; Xiong, W.; Rauch, T.A.; Jin, S.G.; Pfeifer, G.P. MIRA-seq for DNA methylation analysis of CpG islands. *Epigenomics* **2015**, doi:10.2217/epi.15.33.
39. Li, Q.; Suzuki, M.; Wendt, J.; Patterson, N.; Eichten, S.R.; Hermanson, P.J.; Green, D.; Jeddeloh, J.; Richmond, T.; Rosenbaum, H.; *et al*. Post-conversion targeted capture of modified cytosines in mammalian and plant genomes. *Nucleic Acids Res*. **2015**, doi:10.1093/nar/gkv244.
40. Laird, P.W. Principles and challenges of genomewide DNA methylation analysis. *Nat*. *Rev*. *Genet*. **2010**, *11*, 191–203.
41. Xiong, Z.; Laird, P.W. COBRA: A sensitive and quantitative DNA methylation assay. *Nucleic Acids Res*. **1997**, *25*, 2532–2534.
42. Ross, J.P.; Shaw, J.M.; Molloy, P.L. Identification of differentially methylated regions using streptavidin bisulfite ligand methylation enrichment (SuBLiME), a new method to enrich for methylated DNA prior to deep bisulfite genomic sequencing. *Epigenetics* **2013**, *8*, 113–127.
43. Rybak, J.N.; Scheurer, S.B.; Neri, D.; Elia, G. Purification of biotinylated proteins on streptavidin resin: A protocol for quantitative elution. *Proteomics* **2004**, *4*, 2296–2299.
44. Hoeijmakers, W.A.; Bartfai, R.; Francoijs, K.J.; Stunnenberg, H.G. Linear amplification for deep sequencing. *Nat*. *Protoc*. **2011**, *6*, 1026–1036.
45. CSIRO Data Access Portal. Available online: https://data.Csiro.Au/dap/home?Execution=e1s1 (accessed on 30 september 2015).
46. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet*. *J*. **2011**, doi:10.14806/ej.17.1.200.
47. Pedersen, B.S.; Eyring, K.; De, S.; Yang, I.V.; Schwartz, D.A. Fast and Accurate Alignment of Long Bisulfite-Seq Reads. Available online: http://arxiv.org/ (accessed on 30 September 2015).
48. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with Bwa-Mem. Available online: http://arxiv.org/ (accessed on 30 September 2015).
49. Lizio, M.; Harshbarger, J.; Shimoji, H.; Severin, J.; Kasukawa, T.; Sahin, S.; Abugessaisa, I.; Fukuda, S.; Hori, F.; Ishikawa-Kato, S.; *et al*. Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol*. **2015**, doi:10.1186/s13059-014-0560-6.
50. COBRA-Seq: Sensitive and Quantitative Methylome Profiling. Available online: https://stash.Csiro.Au/users/ros259/repos/cobra-seq (accessed on 30 September 2015).
51. Ji, L.; Sasaki, T.; Sun, X.; Ma, P.; Lewis, Z.A.; Schmitz, R.J. Methylated DNA is over-represented in whole-genome bisulfite sequencing data. *Front*. *Genet*. **2014**, doi:10.3389/fgene.2014.00341.
52. Huang, J.; Brieba, L.G.; Sousa, R. Misincorporation by wild-type and mutant t7 RNA polymerases: Identification of interactions that reduce misincorporation rates by stabilizing the catalytically incompetent open conformation. *Biochemistry* **2000**, *39*, 11571–11580.
53. Smith, Z.D.; Gu, H.C.; Bock, C.; Gnirke, A.; Meissner, A. High-throughput bisulfite sequencing in mammalian genomes. *Methods* **2009**, *48*, 226–232.
54. Blattler, A.; Yao, L.; Witt, H.; Guo, Y.; Nicolet, C.M.; Berman, B.P.; Farnham, P.J. Global loss of DNA methylation uncovers intronic enhancers in genes showing expression changes. *Genome Biol*. **2014**, doi:10.1186/s13059-014-0469-0.
55. Akalin, A.; Garrett-Bakelman, F.E.; Kormaksson, M.; Busuttil, J.; Zhang, L.; Khrebtukova, I.; Milne, T.A.; Huang, Y.S.; Biswas, D.; Hess, J.L.; *et al*. Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet*. **2012**, *8*, e1002781.

56. Pedersen, S.K.; Baker, R.T.; McEvoy, A.; Murray, D.H.; Thomas, M.; Molloy, P.L.; Mitchell, S.; Lockett, T.; Young, G.P.; LaPointe, L.C. A two-gene blood test for methylated DNA sensitive for colorectal cancer. *PLoS ONE* **2015**, *10*, e0125041.
57. Mitchell, S.M.; Ross, J.P.; Drew, H.R.; Ho, T.; Brown, G.S.; Saunders, N.F.; Duesing, K.R.; Buckley, M.J.; Dunne, R.; Beetson, I.; *et al*. A panel of genes methylated with high frequency in colorectal cancer. *BMC Cancer* **2014**, doi:10.1186/1471-2407-14-54.
58. Davey, J.W.; Blaxter, M.L. RADSeq: Next-generation population genetics. *Brief. Funct. Genomics* **2010**, *9*, 416–423.
59. Ruike, Y.; Imanaka, Y.; Sato, F.; Shimizu, K.; Tsujimoto, G. Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. *BMC Genomics* **2010**, doi:10.1186/1471-2164-11-137.
60. Ulahannan, N.; Greally, J.M. Genome-wide assays that identify and quantify modified cytosines in human disease studies. *Epigenetics Chromatin* **2015**, doi:10.1186/1756-8935-8-5.
61. Adusumalli, S.; Mohd Omar, M.F.; Soong, R.; Benoukraf, T. Methodological aspects of whole-genome bisulfite sequencing analysis. *Brief. Bioinform.* **2015**, *16*, 369–379.
62. Robinson, M.D.; Stirzaker, C.; Statham, A.L.; Coolen, M.W.; Song, J.Z.; Nair, S.S.; Strbenac, D.; Speed, T.P.; Clark, S.J. Evaluation of affinity-based genome-wide DNA methylation data: Effects of CpG density, amplification bias, and copy number variation. *Genome Res*. **2010**, *20*, 1719–1729.
63. Weber, M.; Davies, J.J.; Wittig, D.; Oakeley, E.J.; Haase, M.; Lam, W.L.; Schubeler, D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **2005**, *37*, 853–862.
64. Gnirke, A.; Melnikov, A.; Maguire, J.; Rogov, P.; LeProust, E.M.; Brockman, W.; Fennell, T.; Giannoukos, G.; Fisher, S.; Russ, C.; *et al*. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol*. **2009**, *27*, 182–189.
65. Dedeurwaerder, S.; Defrance, M.; Calonne, E.; Denis, H.; Sotiriou, C.; Fuks, F. Evaluation of the infinium methylation 450K technology. *Epigenomics* **2011**, *3*, 771–784.
66. Oda, M.; Glass, J.L.; Thompson, R.F.; Mo, Y.; Olivier, E.N.; Figueroa, M.E.; Selzer, R.R.; Richmond, T.A.; Zhang, X.; Dannenberg, L.; *et al*. High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. *Nucleic Acids Res*. **2009**, *37*, 3829–3839.
67. Ordway, J.M.; Bedell, J.A.; Citek, R.W.; Nunberg, A.; Garrido, A.; Kendall, R.; Stevens, J.R.; Cao, D.; Doerge, R.W.; Korshunova, Y.; *et al*. Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis* **2006**, *27*, 2409–2423.
68. Martinez-Arguelles, D.B.; Lee, S.; Papadopoulos, V. *In silico* analysis identifies novel restriction enzyme combinations that expand reduced representation bisulfite sequencing CpG coverage. *BMC Res. Notes* **2014**, doi:10.1186/1756-0500-7-534.
69. Wang, J.; Xia, Y.; Li, L.; Gong, D.; Yao, Y.; Luo, H.; Lu, H.; Yi, N.; Wu, H.; Zhang, X.; *et al*. Double restriction-enzyme digestion improves the coverage and accuracy of genome-wide CpG methylation profiling by reduced representation bisulfite sequencing. *BMC Genomics* **2013**, doi:10.1186/1471-2164-14-11.

Supplementary Materials

1. Supplementary Figures

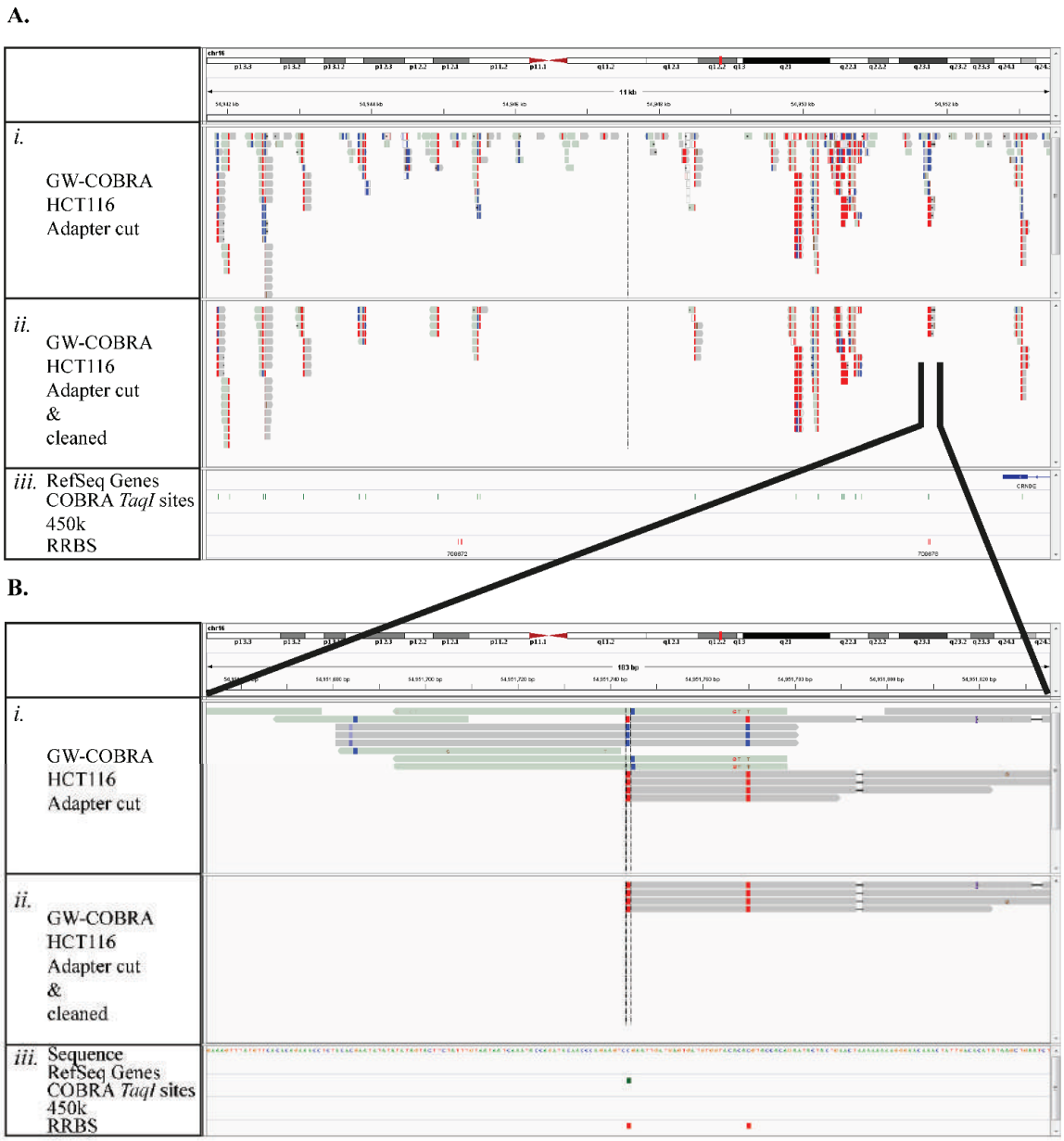


Figure S1. *Cont.*

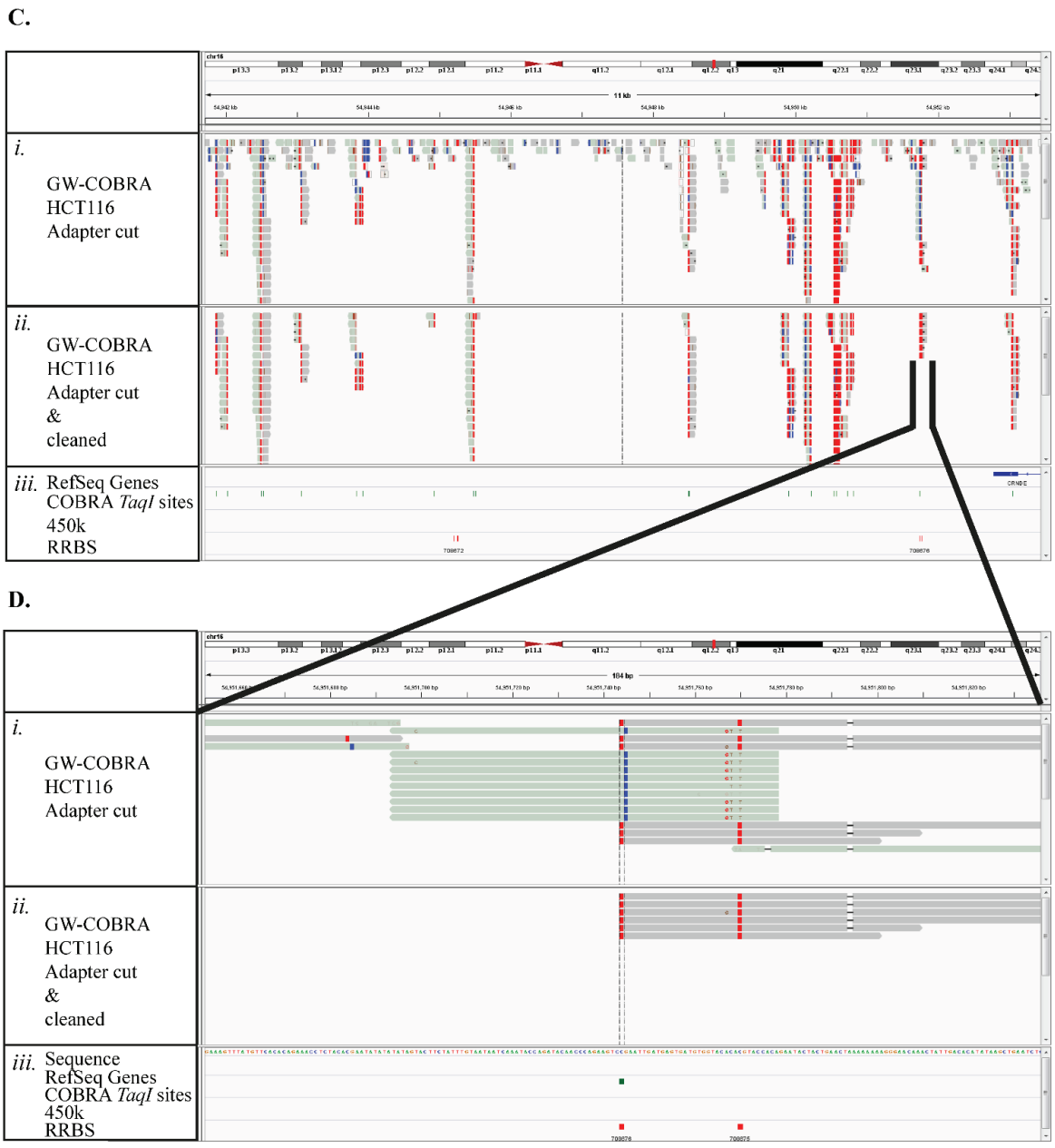


Figure S1. IGV screenshots of (**A**) GW-COBRA cleaning process presented in 11 kb window *CRNDE* gene; (**B**) GW-COBRA cleaning process presented in 183 bp window near *CRNDE* gene; (**C**) LA-COBRA cleaning process presented in 11 kb window *CRNDE* gene; (**D**) LA-COBRA cleaning process presented in 183 bp window *CRNDE* gene.

Fragment Preparation:
100 bp< fragments <500 bp, end polished, phosphorylated, Poly A-tailed DNA sample

```
                                                    5'-pNNNN.........NNNNA-3'
                                                    3'-ANNNN.........NNNNp-5'
```

Ligation of Adapter 2:
5-methylcytosines are shown as 5 whereas unmehtylated Cytosines are shown as C

```
                    5'-IGT5A55G5IGGT5AT5CGCT
                                            \
                                             GCTCTTCCGATCTNNNN.........NNNNAGAT5GGAAGAG5T5GTATG55GT5TT5TG5TTG-3'
                    3'-GTT5GT5TT5TG55GTATG5T5GAGAAGG5TAGANNNN.........NNNNTCTAGCCTTCTCG
                                                                                    \
                                                                                     TCGC5TA5TGGT5G55A5TGT-5'
```

Bisulfite Conversion:
Fragments are single stranded (cytosines were converted to uracils while5-methylcytosines were maintained as cytosines)

```
                    5'-TGTCACCGCTGGTCATCUGUTGUTUTTUUGATUTNNNN.........NNNNAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG-3'
```

GW-COBRA Limited PCR Amplification:
Cycle 1 (A single primer will be used in the first round of PCR as the upper and lower strands are same)

```
                    5'-TGTCACCGCTGGTCATCUGUTGUTUTTUUGATUTNNNN.........NNNNAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG-3'
                       ..................................................3'-TCTAGCCTTCTCGAGCATACGGCAGAAGACGAAC-5'
```

Initial Product (both same)

```
                    5'-TGTCACCGCTGGTCATCUGUTGUTUTTUUGATUTNNNN.........NNNNAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG-3'
                    3'-ACAGTGGCGACCAGTAGACAACAAAAAAACTAAANNNN.........NNNNTCTAGCCTTCTCGAGCATACGGCAGAAGACGAAC-5'
```

Cycles 2-6

```
                    5'-TGTCACCGCTGGTCATCUGUTGUTUTTUUGATUTNNNN.........NNNNAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG-3'
                       ..................................................3'-TCTAGCCTTCTCGAGCATACGGCAGAAGACGAAC-5'

                    5'-bTGTCACCGCTGGTCATCTGTTGTTT-3'.....................................................................
                    3'-ACAGTGGCGACCAGTAGACAACAAAAAAACTAAANNNN.........NNNNTCTAGCCTTCTCGAGCATACGGCAGAAGACGAAC-5'
```

Final Product of GW-COBRA Limited PCR amplification

```
                    5'-bTGTCACCGCTGGTCATCTGTTGTTTTTTTGATTTNNNN.........NNNNAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG-3'
                    3'-ACAGTGGCGACCAGTAGACAACAAAAAAACTAAANNNN.........NNNNTCTAGCCTTCTCGAGCATACGGCAGAAGACGAAC-5'
```

Figure S2. *Cont.*

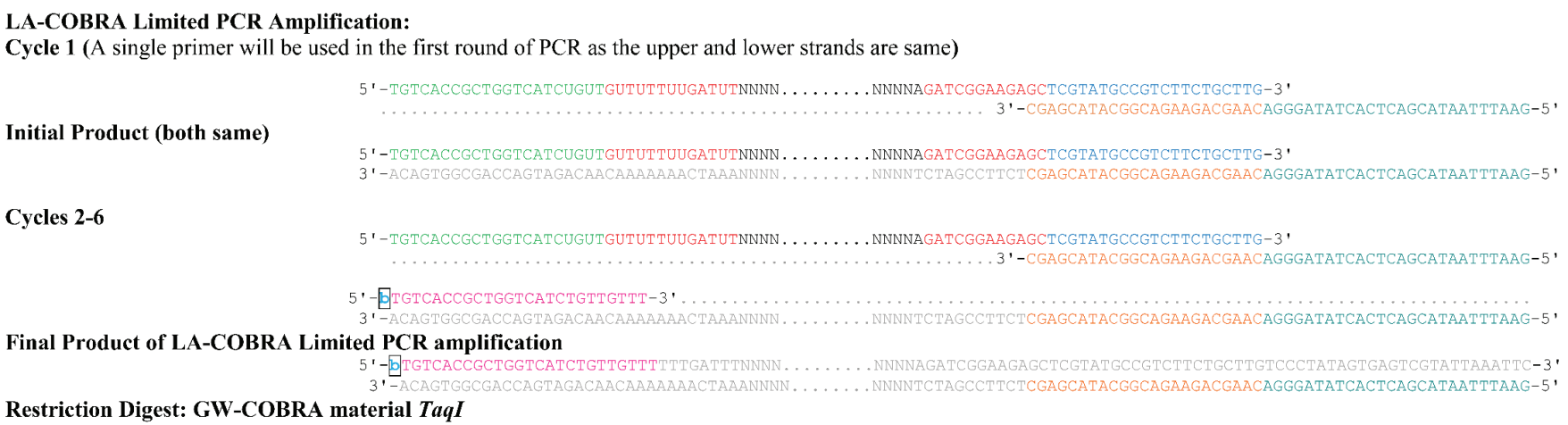


Figure S2. Extended COBRA-seq scheme for constructing both GW-COBRA and LA-COBRA sequencing libraries.

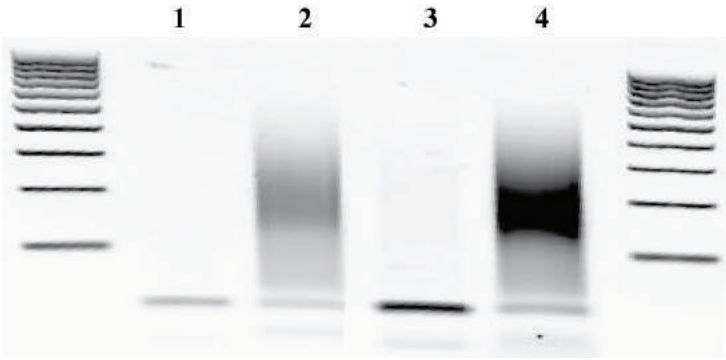


Figure S3. COBRA-seq library construction with minimal starting genomic DNA material. 3% low melting agarose gel and stained with SYBR gold. Adapter-2 ligated and bisulfite-treated genomic DNA amplified as described in limited PCR amplification step. 0.1 and 1 μg of starting genomic DNA material amplified using GW-A2 Fwd and Rev primers respectively (lane 2 and 4). Lane 1 and 3 are negative controls for the limited PCR amplification step.

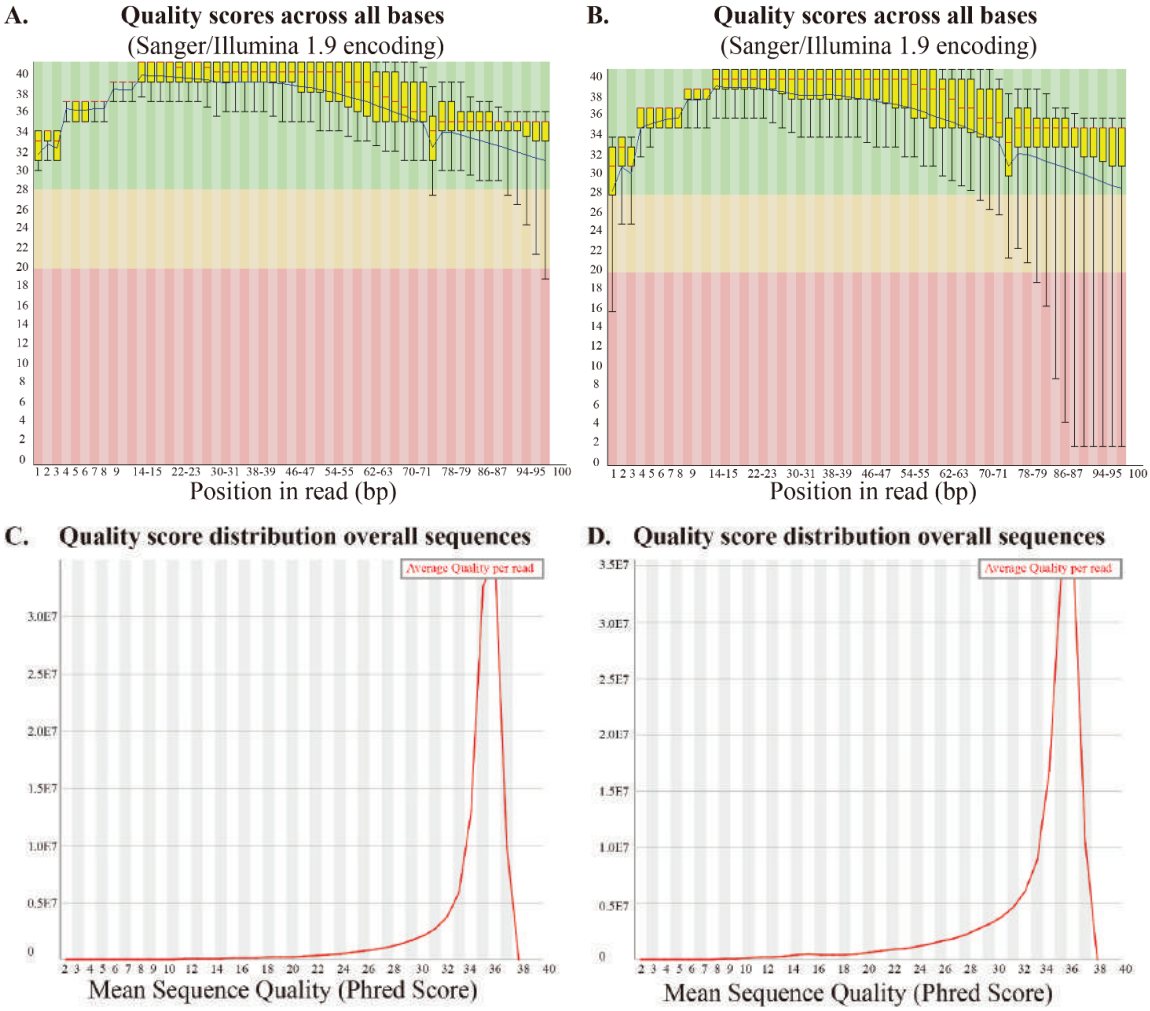


Figure S4. *Cont.*

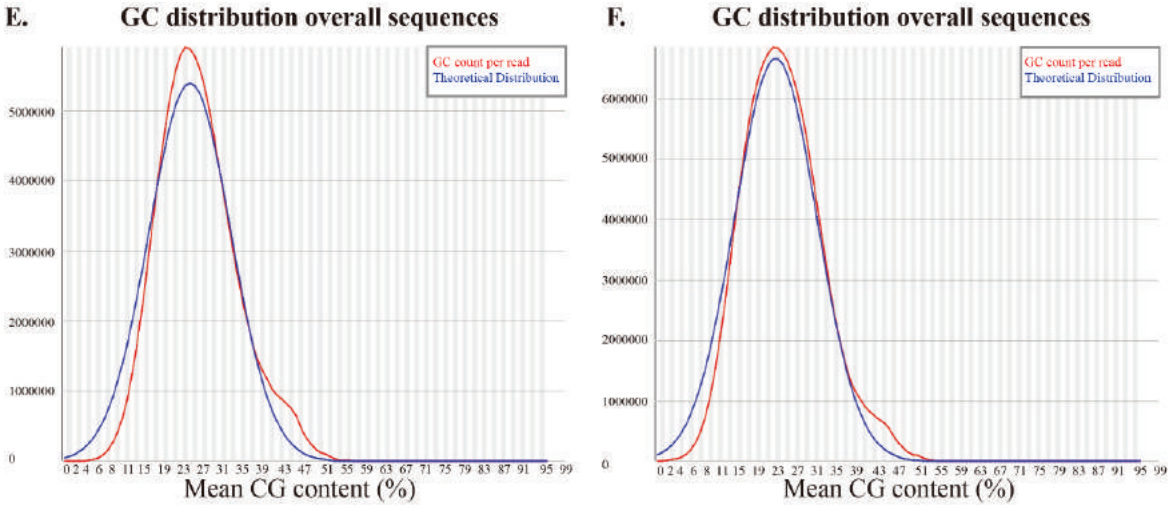


Figure S4. FastQC summary reports of sequencing reads. (**A**) and (**B**) Quality scores across all bases for GWCOBRA and LA-COBRA sequencing results respectively; **(C**) and (**D**) Quality score distribution of overall GW-COBRA and LA-COBRA sequences respectively; (**E**) and (**F**) GC distribution of overall GW-COBRA and LA-COBRA sequences respectively.

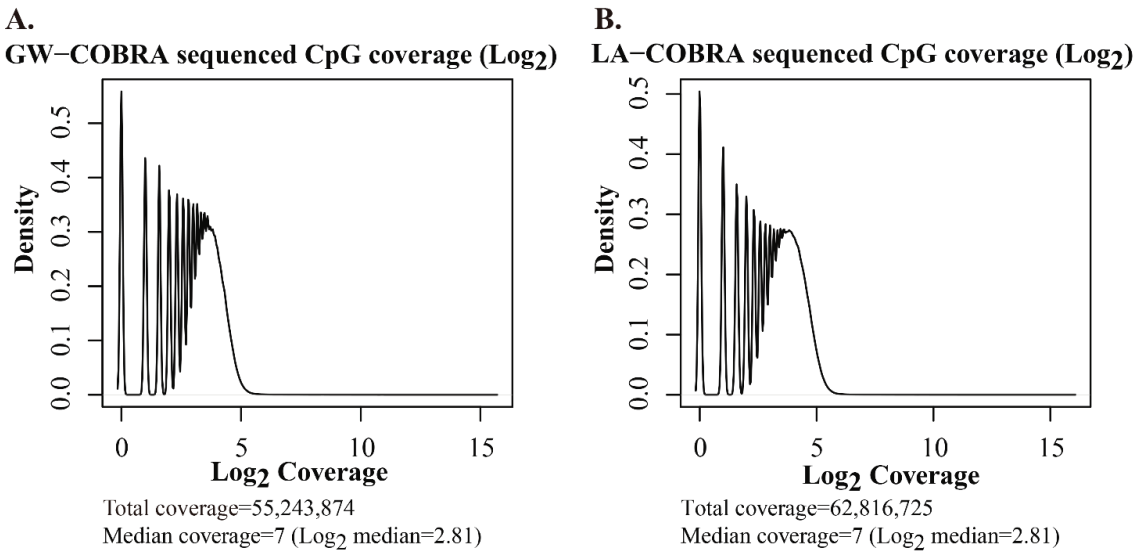


Figure S5. *Cont.*

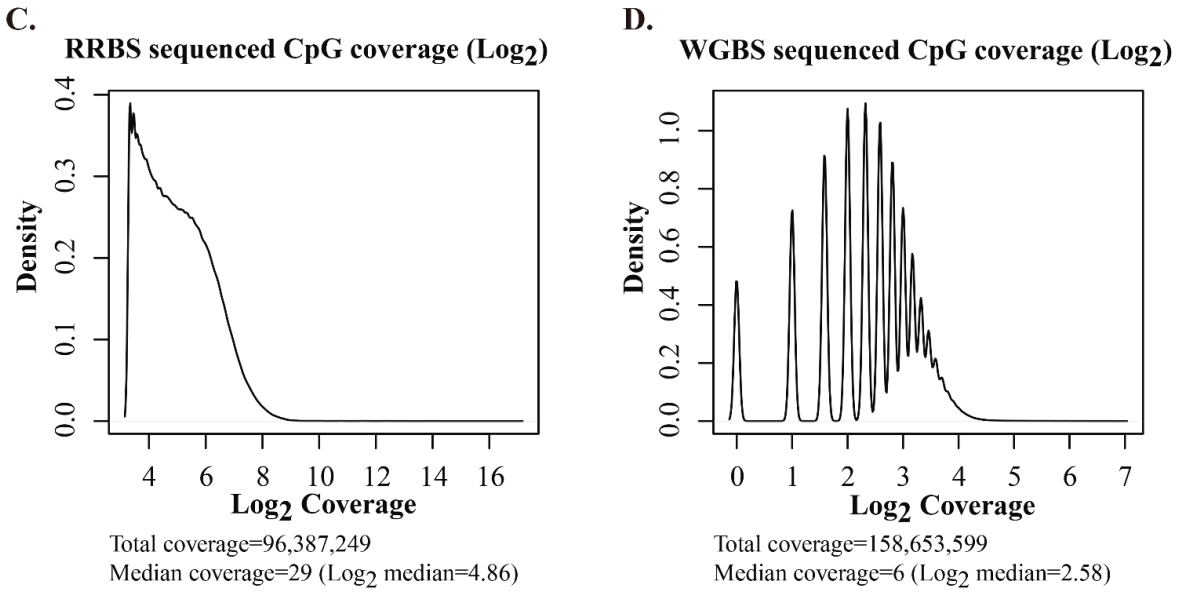


Figure S5. Density plots of the CpG site coverage in the HCT-116 GW- and LA-COBRA libraries and public RRBS and WGBS libraries. Coverage has been transformed by Log2 as each method has a small fraction of CpG sites with factors of magnitude more coverage than the median coverage. (**A**) GW-COBRA; (**B**) LA-COBRA; (**C**) RRBS and (**D**) WGBS.

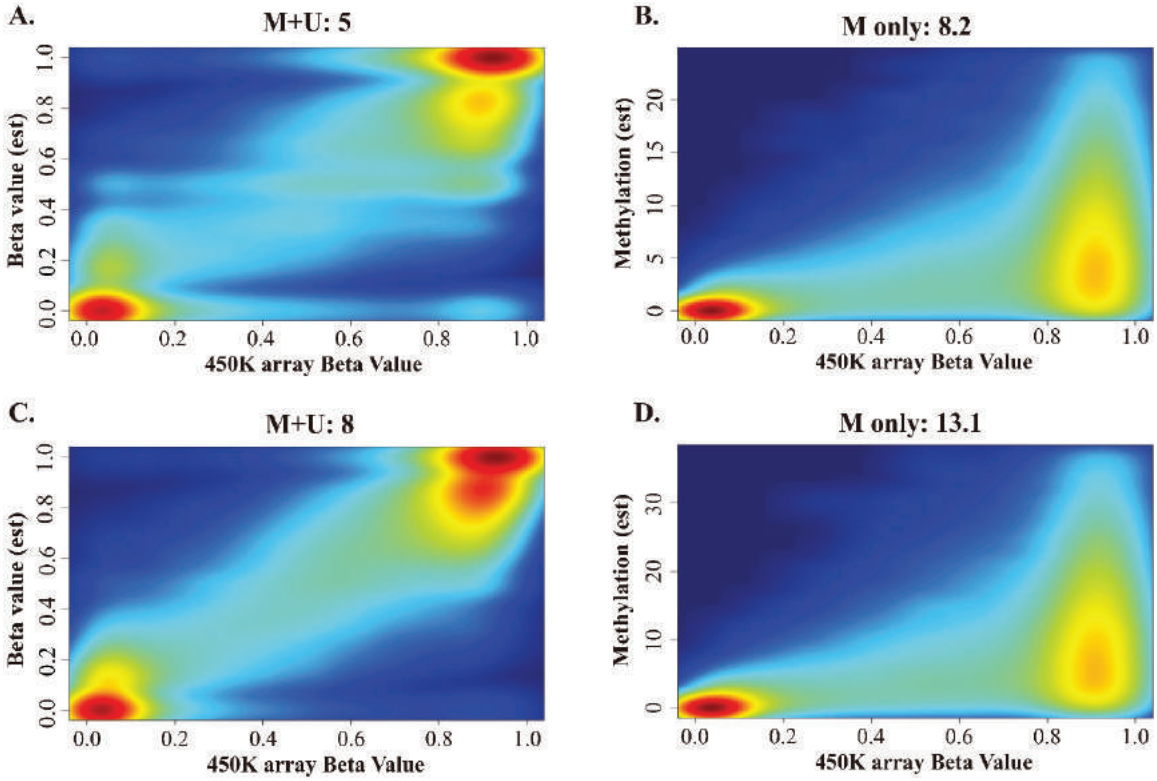


Figure S6. *Cont.*

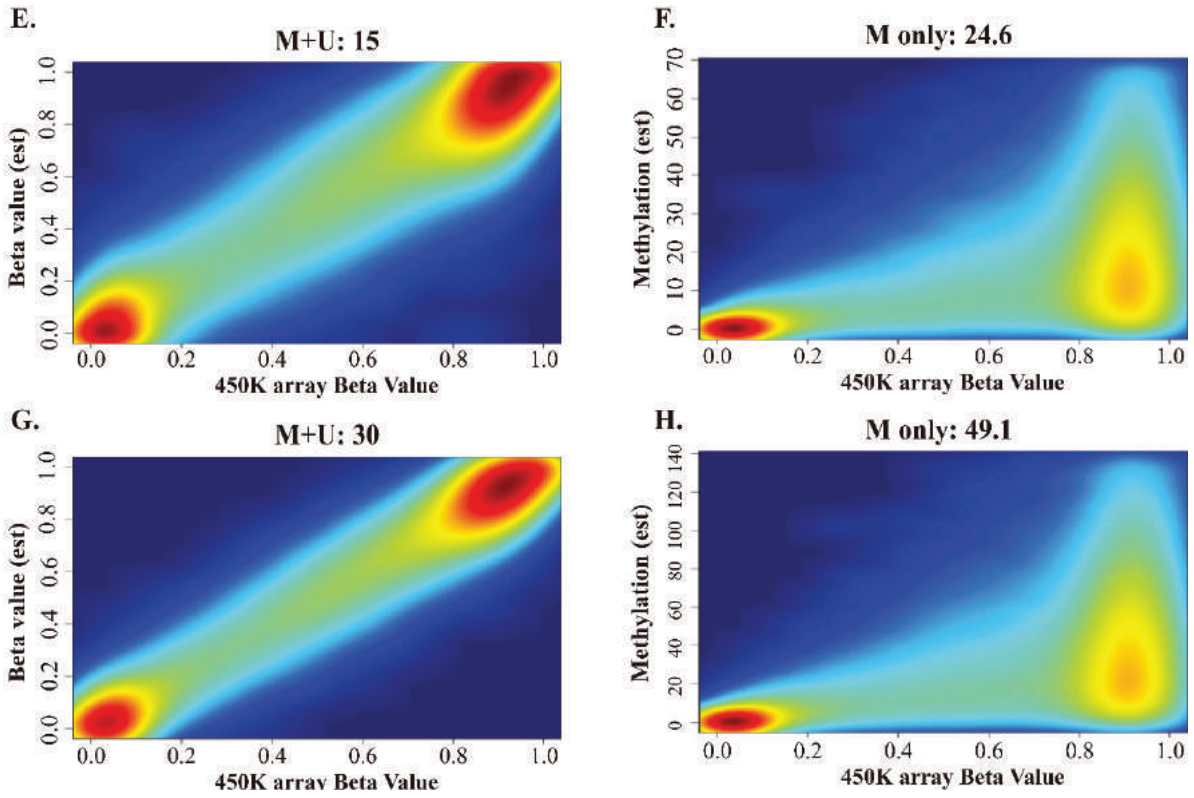


Figure S6. Simulation of methylome sampling methods that targets methylated (M) only or both M and unmethylated (U) fractions. (**A**) Population mean (λ) coverage of 5 in M+U fractions; (**B**) Population mean (μ) coverage of 8.2 in M fraction only (**C**) λ coverage of 8; (**D**) μ coverage of 13.1; (**E**) λ coverage of 15; (**F**) μ coverage of 24.6; (**G**) λ coverage of 30 and (**H**) μ coverage of 49.1.

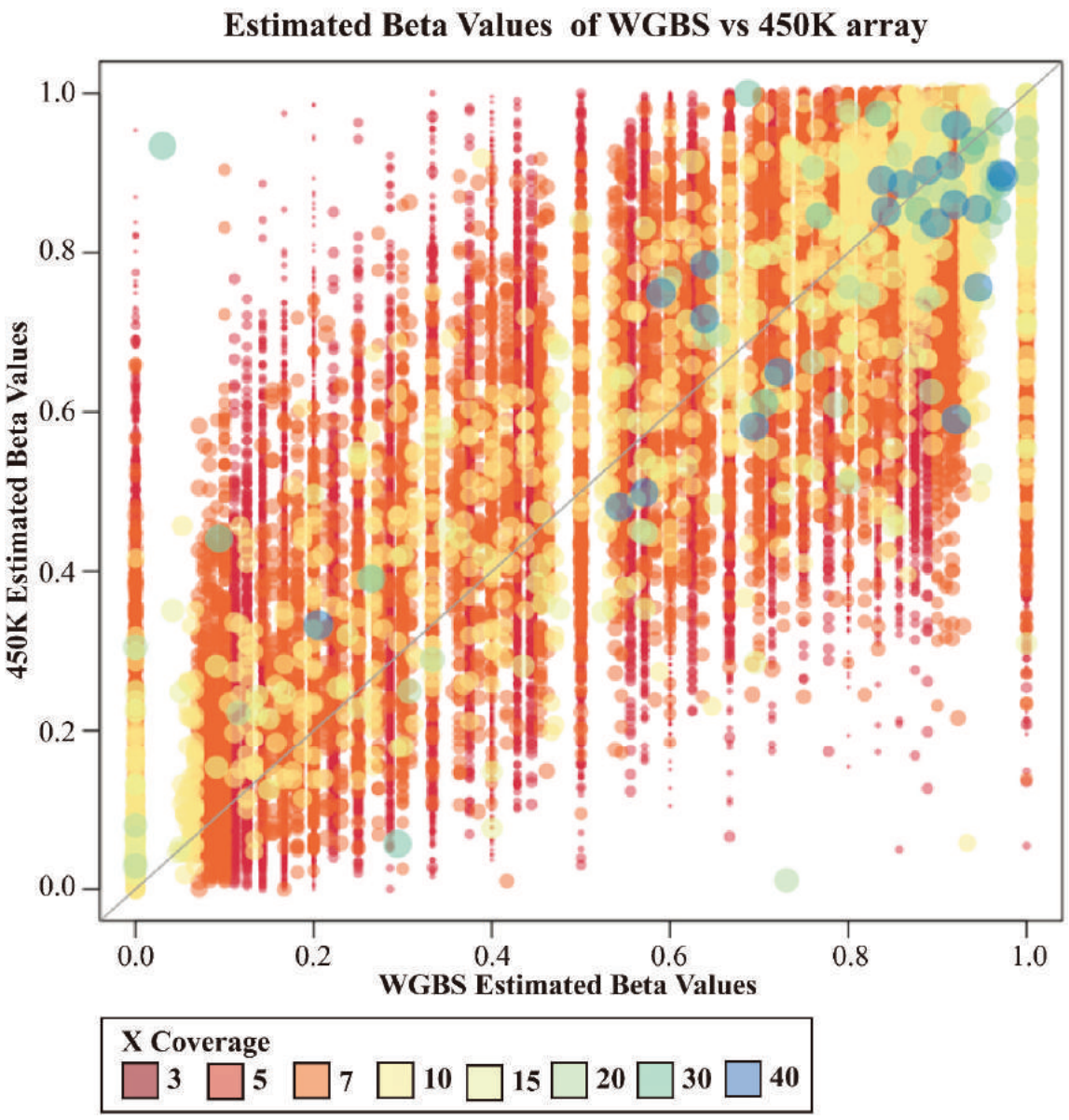


Figure S7. An empirical comparison mapping 200,000 HCT-116 450K array beta values to WGBS beta values.

A.

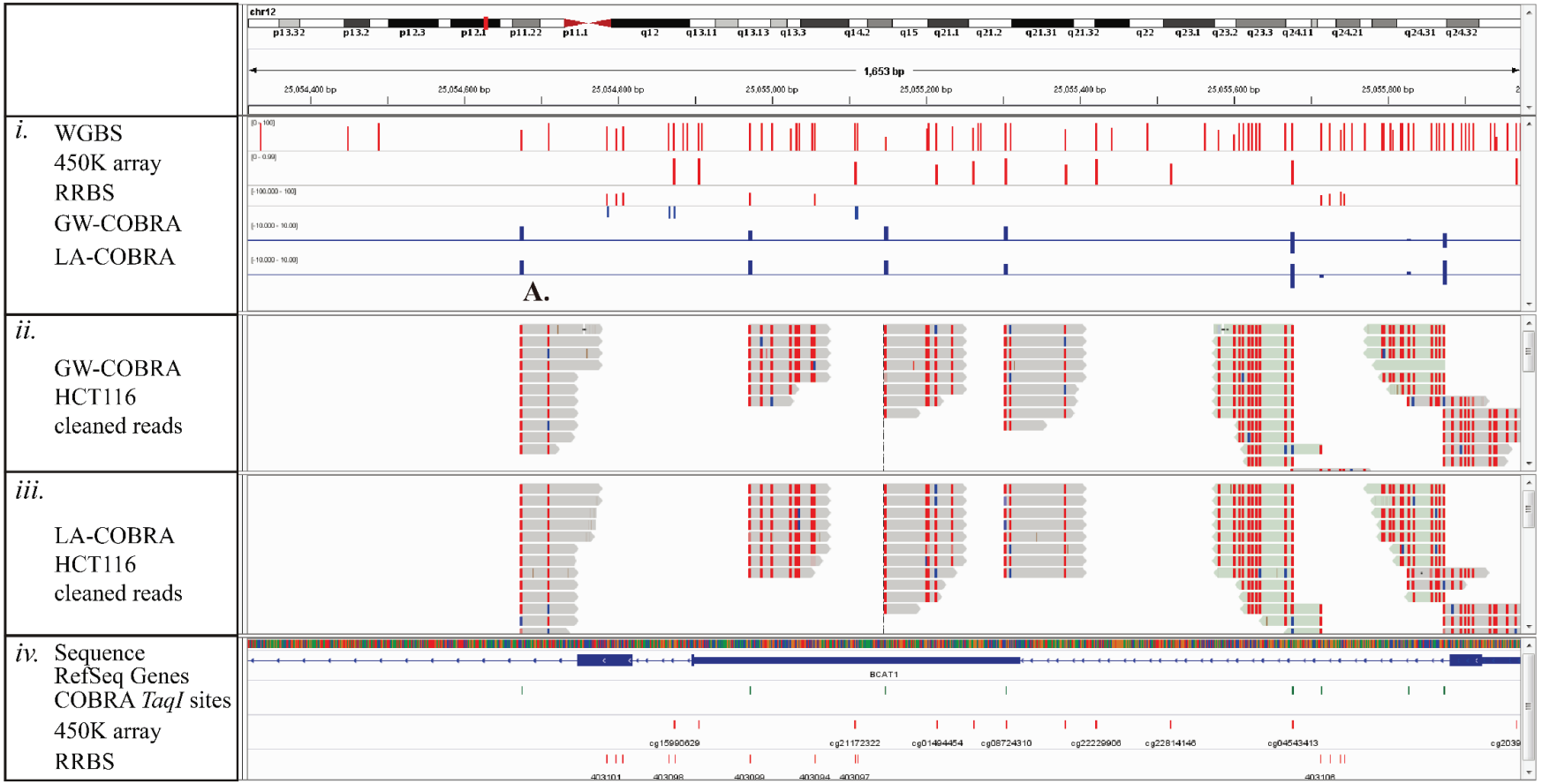


Figure S8. *Cont.*

B.

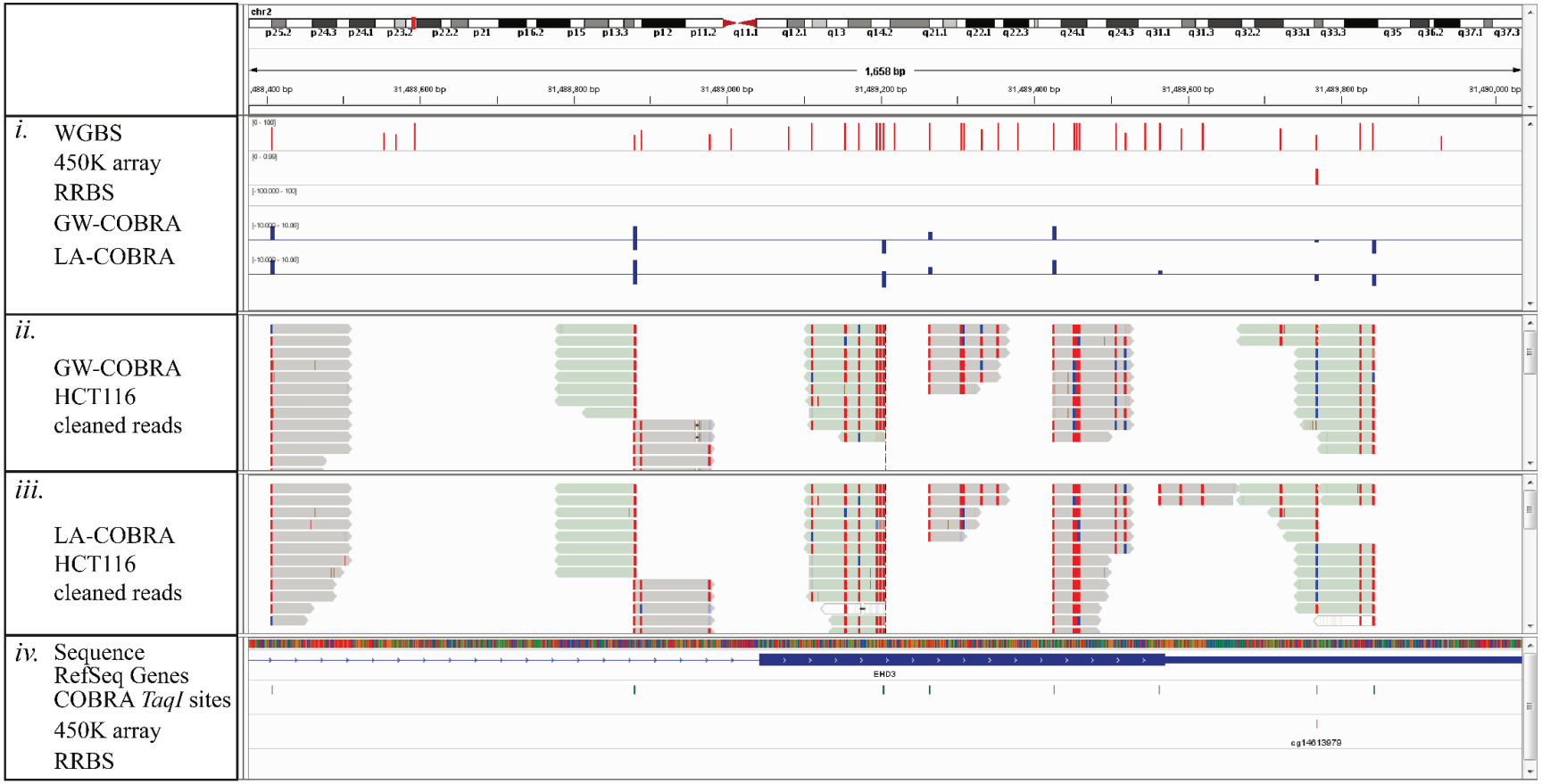


Figure S8. IGV screenshots are to the selected colon cancer associated genes to display read stacks: (**A**) *BCAT1*; (**B**) *EHD3*. *i*. Methylation levels determined by WGBS, 450K arrays, RRBS, GW-COBRA and LA-COBRA respectively; *ii*. Genome coverage and stacking of cleaned GW-COBRA reads; *iii*. Genome coverage and stacking of cleaned LA-COBRA reads; *iv*. Genomic locations of accessible COBRA *TaqI* sites, 450K array probes and accessible RRBS sites.

A.

Figure S9. *Cont.*

B.

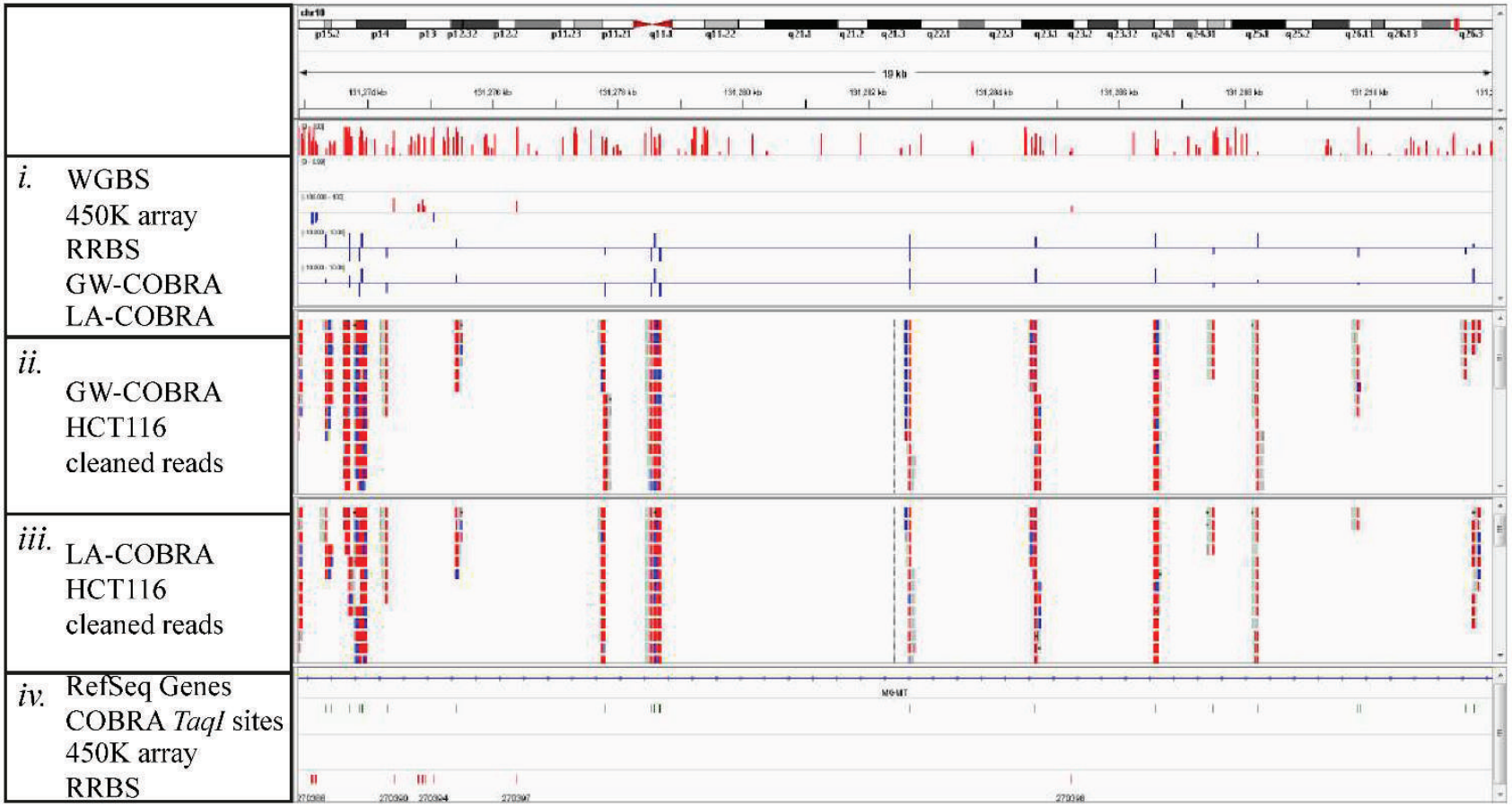


Figure S9. *Cont.*

C.

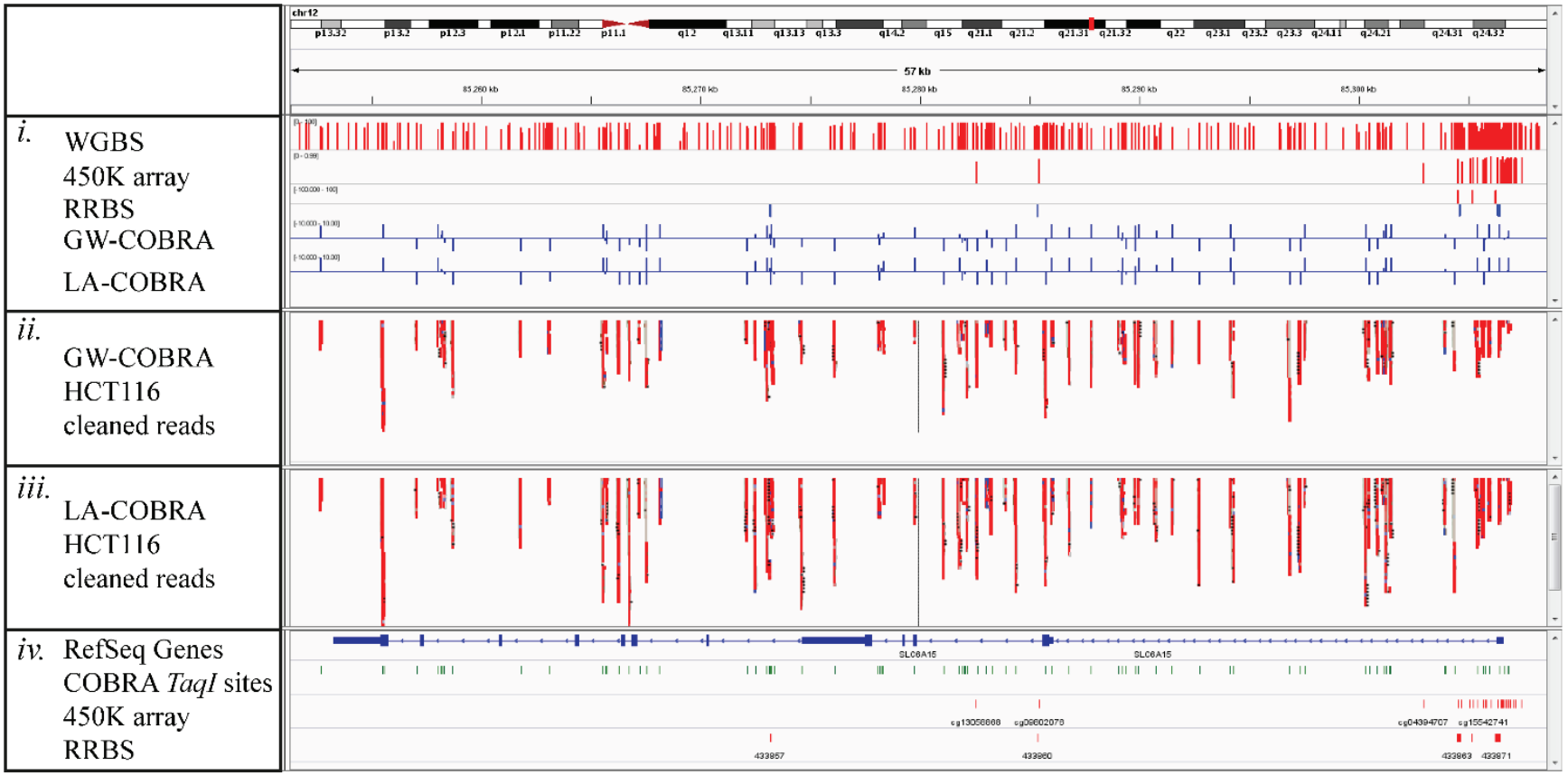


Figure S9. *Cont.*

D.

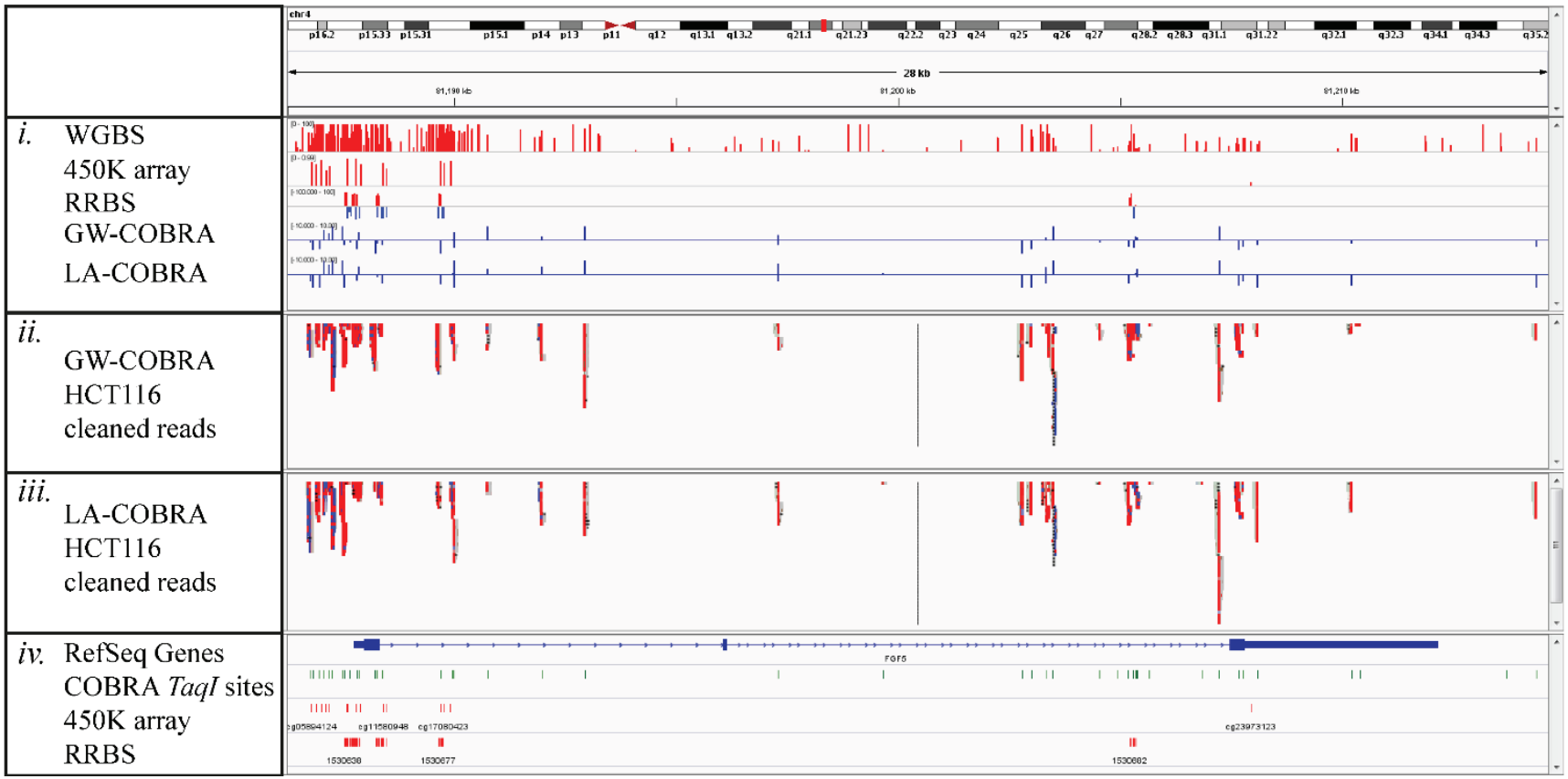


Figure S9. IGV screenshots of methylation profiles on selected colon cancer associated genes: (**A**) *SEPT9*; (**B**) *MGMT*; (**C**) *SLC6A15* and (**D**) *FGF5*. *i*. Methylation levels determined by WGBS, 450K arrays, RRBS, GW-COBRA and LA-COBRA respectively; *ii*. Genome coverage and stacking of cleaned GW-COBRA reads; *iii*. Genome coverage and stacking of cleaned LA-COBRA reads; *iv*. Genomic locations of accessible COBRA *TaqI* sites, 450K array probes and accessible RRBS sites.

2. Supplementary Tables

Table S1. GW-COBRA and LA-COBRA sequencing read statistics.

Stage	Statistic	GW-COBRA	LA-COBRA
Raw file	Read Counts	115,097.029	142,245,797
BAM file	Total BAM file entries	117,215,649	145,687,410
	Unique Read IDs	97,410,988	122,495,417
	%Unique entries	83.1	84.1
Mapping	Original Mapped	108,463,362	134,521,871
	Original Unmapped	8,752,287	11,165,539
	Forward	62,852,260	78,337,754
	Reverse	54,363,389	67,349,656
	% Mapped	92.53%	92.34%
Filtering	Converted to Unmapped Total	55,043,953	73,715,118
	Reason		
	Secondary alignment	11,052,374	12,026,454
	Start Sequence not *TaqI* site	27,499,711	44,507,174
	Not within *in silico* reference	53,219,488	71,705,146
	Not in reference, but could rescue		
	Not in ref, has *TaqI* site	31,633,261	33,795,306
	Not in ref, has *TaqI* site and primary	28,761,986	30,945,333

Table S2. Complexity reduction statistics.

Methylation Context	Enzyme	Recognition Site (5'-3')	Overhang	Restriction Sites	Total Fragments	Sequenceable Fragments	Total CpG Coverage
CpG	*HpyCH4IV*	A/CGT	CG-5'	5,192,248	7,343,344	6,506,928	4,776,485
	TaqI	T/CGA	CG-5'	4,977,110	6,489,977	5,756,193	4,486,417
	AclI	AA/CGTT	CG-5'	692,617	855,037	840,839	683,730
	ClaI	AT/CGAT	CG-5'	388,259	473,628	467,368	384,128
	BstBI	TT/CGAA	CG-5'	866,821	972,174	953,489	850,949
	PvuI	CGAT/CG	3'-AT	46,118	58,497	57,392	90,550
	BstuI	CG/CG	Blunt	725,727	1,451,454	1,058,938	1,209,854
	MluI	A/CGCGT	CGCG-5'	75,894	97,019	93,919	147,648
	BsiWI	C/GTACG	GTAC-5'	90,219	101,772	96,998	172,337
	Hpy99I	CGWCG/	3'-CGWCG	555,835	710,425	556,444	863,001
CpG and CpN	*HinP1I*	G/CGC	CG-5'	490,520	537,572	426,631	802,697
CpN	*CviQI*	G/TAC	TA-5'	1,493,098	1,504,651	1,444,102	1,476,629
	BfuCI/Sau3AI	/GATC	GATC-5'	1,295,990	1,308,369	1,269,237	1,294,261
	EcoRI	G/AATTC	AATT-5'	368,461	369,242	366,551	375,916
CnG	*HpyCH4III*	ACN/GT	3'-N	1,497,255	1,497,255	1,444,743	1,444,743
	Hpy188I	TCN/GA	3'-N	1,371,760	1,371,760	1,298,590	1,212,237
Other Methylome Methods	SuBLiME (*CviQI*)			5,047,091	10,094,182	8,702,060	21,388,229
	DREAM (*XmaI*)			374,921	749,842	51,046	497,412
	RRBS (*MspI*)			2,297,221	4,594,442	1,528,440	4,724,246
	450K array			-	-	-	484,022

Table S3. Complexity reduction statistics cont'd. The % of genome coverage per genomic location by COBRA-seq and other genome sampling methods.

	Transcript	Exon	TSS (4 kb)	Genebody	Intergenic	CpG Island	CpG Shore	CpG Ocean	5' UTR	3' UTR	DNaseI HS	Enhancer	Repeat	Total CpGs
Genome-wide	52.3	7.5	12.8	43.0	44.2	7.4	7.2	85.4	1.8	1.4	26.5	0.7	52.3	28,217,448
HpyCH4IV	50.4	4.8	8.2	44.6	47.3	2.5	5.7	91.8	0.8	1.3	21.8	0.7	52.0	4,776,485
TaqI	50.9	6.2	9.6	43.9	46.5	3.5	6.3	90.2	1.1	1.3	21.5	0.6	55.2	4,486,417
AclI	49.2	5.1	8.2	43.3	48.5	2.2	5.8	92.0	0.8	1.5	23.2	0.7	41.1	683,730
ClaI	48.4	6.5	7.8	42.6	49.7	2.3	5.1	92.6	0.9	1.4	20.9	0.5	43.2	384,128
BstBI	50.9	5.6	9.1	44.3	46.6	2.7	6.2	91.1	0.9	1.2	18.6	0.5	56.6	850,949
PvuI	57.9	17.7	28.6	36.4	35.0	24.8	9.8	65.4	5.7	1.8	44.4	1.1	36.2	90,550
BstuI	56.9	13.1	27.1	37.3	35.6	24.0	10.4	65.7	4.7	1.6	38.9	1.1	52.4	1,209,854
MluI	57.4	13.0	24.5	39.8	35.7	21.4	10.1	68.5	3.9	1.9	40.4	1.2	43.7	147,648
BsiWI	58.2	14.1	23.5	41.1	35.4	20.5	9.6	70.0	3.9	1.8	40.5	1.1	42.5	172,337
Hpy99I	59.6	20.7	31.9	35.6	32.5	30.2	10.6	59.3	6.6	2.1	49.9	1.3	33.0	863,001
HinP1I	56.8	12.3	28.6	36.5	34.8	25.6	9.6	64.8	5.1	1.3	35.9	1.0	61.4	802,697
CviQI	52.7	6.6	10.3	45.1	44.6	4.8	6.4	88.7	1.1	1.5	25.0	0.7	51.9	1,476,629
BfuCI/Sau3AI	52.1	6.7	11.0	44.0	45.0	5.0	6.6	88.4	1.4	1.3	21.9	0.6	59.3	1,294,261
EcoRI	51.0	4.3	9.2	44.5	46.3	3.0	6.4	90.6	0.8	1.0	14.6	0.4	72.9	375,916
HpyCH4III	51.9	6.7	9.3	44.9	45.7	3.5	6.5	90.1	1.0	1.5	25.7	0.7	47.2	1,444,743
Hpy188I	52.3	9.7	14.0	41.9	44.1	7.4	8.2	84.4	2.1	1.7	32.6	0.9	40.8	1,212,237
SuBLiME	52.5	7.4	11.7	43.8	44.4	6.1	6.6	87.2	1.6	1.5	24.8	0.7	53.0	21,388,229
DREAM	63.5	24.8	53.5	24.1	22.4	54.2	13.6	32.2	11.3	1.7	64.4	1.5	32.8	497,412
RRBS	59.4	17.8	33.4	34.7	31.9	30.4	12.7	56.9	6.4	1.9	46.7	1.2	44.9	4,724,246
450K array	61.6	23.1	51.2	29.8	19.0	31.1	23.2	45.6	8.4	4.6	66.8	1.7	15.8	484,022

3. Supplementary Materials and Methods

3.1. COBRA-Seq Library Preparation Protocol in Detailed

Part 1: Construction of Adapters

Each Adapter was made using two oligonucleotides dissolved in sterile water with a final concentration of 500 µM (Table 1). The Adapter-2 oligonucleotides partially complement each other therefore the adapter has a "Y" shape. Particularly after bisulfite treatment Adapter-2 upper strand remains unconverted and non-complementary to Adapter-2 lower strand On the other hand all the Adapter-1 combinations were designed using fully complementary oligonucleotides.

P1.1. Mix the matching oligonucleotides with 1× NEB Quick Ligation Buffer in a final concentration of 50 µM each and in a final volume of 100 µL.

P1.2. Hybridize the oligonucleotides with the following temperature cycles on a PCR machine: 95 °C 5 min, 72 °C 5 min, 60 °C 5 min, 50 °C 3 min, 40 °C 3 min, 30 °C 3 min, 20 °C 3 min and 10 °C 3 min. Upon hybridisation the adapters were stored at −20 °C.

P1.3. Store them in a cool rack at −20 °C.

Part 2: Sequencing Library Preparation Protocol

Part 2-1. DNA Fragmentation

P2-1.1. Resuspend genomic DNA in 300 µL low TE to a final concentration of 16.66 ng/µL.

P2-1.2. Fragment the genomic DNA using a Bioruptor UCD-200 sonicator (Diagenode, Belgium) at a power setting of "high" for four sets of 30 cycles of 15 s on/off with 15 min intervals on ice in between each set.

P2-1.3. For visualisation, run 5 µL of the sonicated material on 3% Biorad Low Range Ultra Agarose gel.

P2-1.4. The fragments range should be between 100–500 bp.

Part 2-2. DNA Precipitation

P2-2.1. To precipitate the fragmented DNA, add 600 µL volumes of cold, absolute ethanol, 30 µg of GlycoBlue and 100 µL of 3 M sodium acetate.

P2-2.2. Vortex the samples, quickly spin at full speed and incubate at −20 °C for 20 min.

P2-2.3. Centrifuge the samples at full speed (16,000 g) for 15 min at 4 °C.

P2-2.4. Decant the supernatant and wash the blue pellet with 70% cold ethanol.

P2-2.5. Air dry the residual ethanol for 5 min and samples.

P2-2.6. Resuspend the pellet in 36 µL low TE.

P2-2.7. Measure the DNA concentration was measured using 1 µL in NanoDrop 1000 Spectrophotometer.

Part 2-3. End Repair

The End-It Repair reaction kit (Epicentre, Wisconsin) was used to convert DNA with incompatible overhangs such as 5'-protruding and/or 3'-protruding ends to blunt-ended, 5'-phosphorylated DNA by exploiting the 5' to 3' polymerase and 3' to 5'exonuclease enzymes.

P2-3.1. The sonicated and precipitated library material was used with 5 µL 10× End-Repair Buffer, 5 µL dNTP mix, 5 µL ATP and 1 µL Enzyme mix, in a final volume of 50 µL.

P2-3.2. Incubate the reaction at room temperature for 45 min then heat activate the enzymes at 70 °C for 10 min.

Part 2-4. Reaction Clean up

P2-4.1. Increase the volume of the End Repair reaction to 200 uL using sterile water.

P2-4.2. Add an equal volume of phenol: chloroform: isoamyl alcohol.

P2-4.3. Vortex the samples for 15 s and spin at 16,000 g for 5 min.

P2-4.4. Transfer the upper aqueous phase to a new 2 mL microfuge tube.

P2-4.5. Add Sodium acetate to have a final salt concentration of 0.3 M.

P2-4.6. Add three volumes of cold, absolute ethanol, together with 2 µL of GlycoBlue.

P2-4.7. Invert the samples, quickly spin and incubate at −20 °C for 20 min.

P2-4.8. Centrifuge the samples at full speed (16,000 g) for 15 min at 4 °C.

P2-4.9. Decant the supernatant and wash the blue pellet with 70% cold ethanol.

P2-4.10. Air dry the residual ethanol for 5 min and samples.

P2-4.11. Resuspend the pellet in 42 µL low TE.

P2-4.12. Measure the DNA concentration was measured using 1 µL in NanoDrop 1000 Spectrophotometer.

Part 2-5. A-Tailing

A single adenine nucleotide was added to the 3' ends of the blunt fragments to facilitate ligation with adapters that have T overhang, and to limit ligating to one another during the adapter ligation reaction.

P2-5.1. Mix the end repaired library material with 5 µL of 10× NEBNext dA-Tailing reaction buffer (10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 0.2 mM dATP pH 7.9) and 3 µL of 3' to 5' exonuclease.

P2-5.2. Incubate at 37 °C for 30 min, then heat inactivate the enzyme at 75 °C for 20 min and cool on ice prior to ligation.

Part 2-6. Adapter-2 Ligation

Adapter-2 were designed to contain a corresponding single thymine nucleotide on the 3' end which provides a complementary overhang to ligate to the genomic library fragments (Figure S1).

P2-6.1. Prepare the Ligation reaction with 2 µg of fragmented DNA, 10 molar excess Adapter-2 in a 50 µL reaction containing 2000 U T4 Ligase Enzyme and 25 µL Quick Ligation Buffer.

P2-6.2. Incubate the reaction at room temperature for 20 min and cool the samples on ice.

Note: The ten molar excess can be calculated considering the library fragments average 200 bp and contain two ligatable ends.

There are **X** moles of ligatable ends in the sonicated material,

$$\mathbf{X} = 2\ \mu g\ DNA \times 2\ (two\ ends)/(650\ Da/bp \times 200\ bp) = 3.1 \times 10^{-5} = 31\ pmol$$

Adapter-2/Fragment = 10:1 $\qquad$ 31 pmol × 10 = 310 pmol

Therefore, 10 molar excess is 6.2 µL of the 50 µM Adapter-2 stock for 2 µg library material.

Part 2-7. Removal of Excess Adapter-2

P2-7.1. Mix one volume of ligation reaction with 5 volumes of Buffer PB supplied in Qiagen Quick PCR purification kit.
P2-7.2. Apply the mixture into a MinElute column and centrifuge at full speed for 1 min.
P2-7.3. Run the flow-through through the MinElute column again before discarding.
P2-7.4. Wash the column with 750 µL Buffer PE and centrifuged for 1 min.
P2-7.5. Wash the column again with 500 µL Buffer PE and centrifuge for 1 min.
P2-7.6. To remove the residual ethanol, place the column into an empty collection tube and spin when the lid is open at full speed for 1 min.
P2-7.7. Elute the library material twice in 42 µL Elution Buffer with 1 min centrifugation.

Part 2-8. Check point: PCR Reaction

Although this step was not necessary for the GW-COBRA and LA-COBRA library preparation, it was designed as a control point to determine the efficiency of Adapter-2 ligation step which is crucial to have a well distributed library in the range of 200–500 bp. The library material is expected to contain Adapter-2 in both ends, therefore the single primer was employed as a forward and reverse primer during this amplification step.

P2-8.1. Perform the PCRs in 25 µL volume using GoTaq white (Promega) in the buffer supplied with the enzyme, with the addition of 0.2 µM final concentration of a single primer targeting one strand of the Adapter-2.
P2-8.2. The temperature cycles for the PCR are: 95 °C for 3 min; 95 °C for 15 s, 60 °C for 30 s and 72 °C for 1 min, 8 cycles; a final extension of 72 °C for 2 min.
P2-8.3. Run the PCR product to access the success of Adapter-2 ligation on 3% Low Range Ultra Agarose (Biorad).

Note: Figurative representation of Check point PCR step is provided in Figure S10.

It is likely that, short fragments (100–200 bp) may form a hairpin structure or ligate to one another (Figure 2). Over amplification will create library artefact and resulting higher MW bands on an agarose gel. It is recommended to use maximum of 10 PCR cycles.

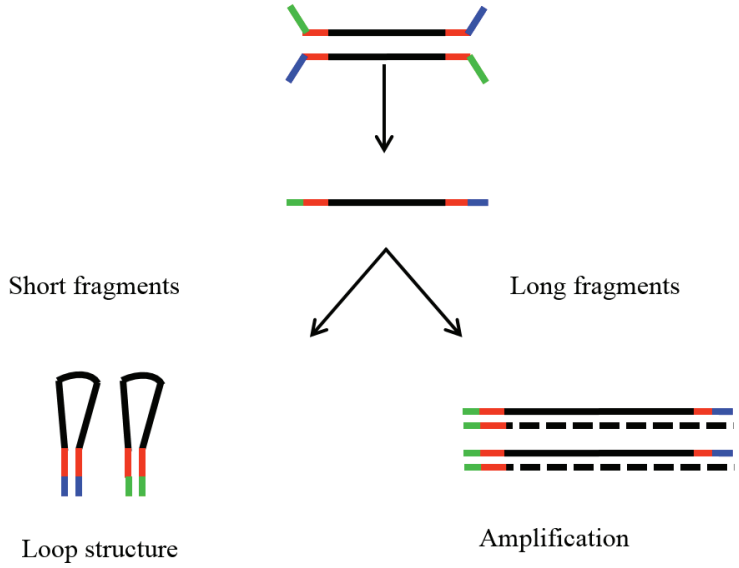


Figure S10. Schematic representation of products generated by the minimal PCR step.

Part 2-9. Bisulfite Treatment

Bisulfite treatment, also known as bisulfite conversion, is a gold standard method for DNA methylation analysis. Bisulfite deaminates unmethylated cytosine to produce uracil whereas methylated cytosines are protected. The locations of unmethylated cytosines and 5-methylcytosines can be determined at single-nucleotide resolution via sequencing.

P2-9.1. Ligated fragments containing Adapter-2 in both ends were bisulfite treated using an EZ DNA Methylation-Gold Kit (Zymo Research).

P2-9.2. Prepare aliquots of the 10 µg/µL library material in a final volume of 20 µL in PCR tubes and add 130 µL of the CT Conversion Reagent for the bisulfite treatment step.

P2-9.3. Incubate the bisulfite conversion reaction at 99 °C 5 min, 60 °C 25 min, 99 °C 5 min, 60 °C 85 min, 99 °C 5 min, 60 °C 175 min and 22 °C 5 min.

P2-9.4. Wash the reaction and desulphonate as per manufacturer's instruction.

P2-9.5. Elute the final product twice in 16 µL.

P2-9.6. Quantify Single stranded DNA concentration using 1 µL in NanoDrop 1000 Spectrophotometer with RNA settings.

Post-bisulfite treatment, the Adapter-2 oligonucleotides are not complementary. At this step, the library material is single stranded.

Part 2-10. Minimal PCR Amplification (Biotin Tagging)

Single stranded bisulfite treated library materials are used as templates with a final concentration of 1 ng/µL in 25 µL using GoTaq white (Promega, Madison, WI, USA) enzyme and buffer mixture. The primers are used in the final concentration of 0.1 µM. The reverse primer contains 5' double biotin whereas the LA-COBRA forward primer had an overhang of T7 promoter region that is used to generate RNA copies via *in vitro* transcription that are subsequently reverse-transcribed in the later steps.

P2-10.1. The temperature cycles for the PCR are: 98 °C 3 min; 98 °C for 15 s, 56 °C for 30 s and 72 °C for 1 min, for 6 cycles; a final extension of 72 °C for 2 min.

P2-10.2. Visualize the PCR products on a 3% Low Range Ultra Agarose Gel.

Part 2-11. Reaction Clean-Up

A Wizard SV PCR Clean Up System (Promega) was used to remove the enzymes and excess primers. The manufacturer's instructions were followed with minor modifications at the elution step. Samples were eluted twice in 44 µL of sterile water.

*The reaction clean up step is critical to remove the polymerase enzyme which may fill the enzyme restriction sites after digestion. Moreover, it is also needed to remove the excess the 5' double biotin primers as they would bind the streptavidin beads with high efficiency and interfere with the streptavidin purification step.

Part 2-12. Enzyme Digestion

P2-12.1. Digest the PCR amplified and purified library material overnight at 65 °C with 20–100 U Taq1, 100 µg/mL BSA and 1× NEB 4 Buffer in a final volume of 100 µL.

P2-12.2. Use a top layer of sterile mineral oil to prevent evaporation and maintain the enzyme concentration steady.

Part 2-13. Streptavidin Purification

The GW and LA-COBRA libraries are biotin tagged at the amplification step. After *Taq1* digest, there are three main fragment types in the libraries; uncut biotinylated fragments (no internal *Taq1* site), the cut fragments containing 5' double biotin and the other part of the cut fragments which are non-biotinylated. In this step, Dynabeads® M-280 Streptavidin beads, which are ideal for purification of biotinylated molecules, were used to capture the biotinylated fragments, hence enrich for the non-biotinylated fragments.

P2-13.1. Resuspend the Dynabeads to obtain a homogenous suspension by shaking the vial.

P2-13.2. Transfer 100 µL of the beads into non-sticky microfuge tubes for each sample.

P2-13.3. Place the aliquots on a magnetic rack for 2 min to ensure all the beads are collected on the tube wall and remove the supernatants.

P2-13.4. Wash the beads twice in 100 µL 2× Binding and washing (B&W) Buffer (2×) with addition of 0.2% Tween 20 (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2.0 M NaCl and 0.2% Tween 20).

P2-13.5. Resuspend the beads in 100 µL 2× B&W Buffer with 0.2% Tween 20 and mix with an equal volume of Taq1 digested biotinylated library material.

P2-13.6. Gently mix the samples with a pipette and agitate at 300 rpm for 20 min at room temperature.

P2-13.7. Separate the beads coated with biotinylated library fragments using a magnet for 2 min.

P2-13.8. Remove the supernatant containing enriched non-biotinylated fragments and transfer into another tube.

The streptavidin beads can be uncoated from the biotinylated fragments with few additional steps. Although this is not a necessary step for GW and LA-COBRA library preparations, it is a good practice to visualize the discarded fragment on an agarose gel.

P2-13.9. Wash the beads twice with 1× B&W Buffer with addition of 0.1% Tween 20 (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1.0 M NaCl and 0.1% Tween 20).

P2-13.10. Resuspend the beads in 1× B&W Buffer containing 30 mM d-biotin and incubate at room temperature with gently mixing for 15 min

P2-13.11. Heat treat the samples are heated to 80 °C for 15 min.

P2-13.12. Place the samples on a magnet for 2 min and aspirate the supernatant containing the biotinylated fragments as soon as the magnetic beads were cleared from solution.

Part 2-14. Adapter-1 Ligation

Following digestion, the cut ends of DNA fragments are ligated to Adapter-1. This adapter has a 5'-CG-3' overhang and is not phosphorylated to reduce Adapter-1 to Adapter-1 ligation. The LA-COBRA Adapter-1 sequence incorporated an addition of P5 primer region that allows priming and reverse transcription on the T7 RNA polymerase-generated RNA transcripts.

P2-14.1. Prepare the ligation reaction with 1 μg of fragmented DNA with 2-fold molar excess of Adapter-1 in a 50 μL reaction containing 2000 U T4 Ligase Enzyme and 25 μL Quick Ligation Buffer.

P2-14.2. Incubate the reaction mixture at room temperature for 20 min and cool on ice.

The two-fold molar excess is calculated considering the library fragments contain two ligatable ends with an acceptance that average library fragments are 200 bp.

There are **X** moles of ligatable ends in the sonicated material,

$$\mathbf{X} = 1\ \mu g\ DNA \times 1\ (single\ end)/(650\ Da/bp \times 200\ bp) = 7.8 \times 10^{-6} = 7.8\ pmol$$

Adapter-1/Fragment = 2:1 7.8 pmol × 10 = 15.6 pmol

Therefore, 0.32 μL of the 50 μM Adapter-1 stock is used for 1 μg library material.

Part 2-15. Adapter-1 Clean up

Adapter dimerization creates significant noise in the sequencing results, therefore purification steps are extremely important to remove any excess adapters. Ampure XP Bead Purification system is utilized to remove the fragments that have a size less than 100 bp using TruSeq DNA sample preparation guide with minor modifications.

P2-15.1. Cool down the Ampure XP beads to room temperature.

P2-15.2. Combine 25 μL of well mixed beads with 35 μL PCR water and 100 μL Sample.

P2-15.3. Gently mix the mixture via pipetting and incubate at room temperature for 15 min.

P2-15.4. Place the samples on a magnetic rack until the beads are collected to the tube wall.

P2-15.5. Remove the supernatant containing the fragments less than 100 bp.

P2-15.6. Wash the beads with 200 μL of 80% ethanol twice while the tubes are on the magnetic rack.

P2-15.7. Air dry the residual ethanol and resuspend the beads in 50–100 μL low TE for 2 min.

P2-15.8. Capture the beads on a magnetic rack.

P2-15.9. Transfer the solution containing the library material above 100 bp into a clean eppendorf.

Part 2-16. Check Point PCR

To test the presence of adapters in library fragments, perform PCRs in 25 µL volume using GoTaq white (Promega) in the buffer supplied with the enzyme, with the addition of 0.1 µM final concentration of Illumina flow cell primers on a fraction of the library.

P2-16.1. The temperature cycles for the PCR are: 98 °C for 5 min; 98 °C for 15 s, 65 °C for 30 s and 72 °C for 1 min, 8 cycles; a final extension of 72 °C for 2 min.

P2-16.2. Visualize the amplification products by electrophoresis on 3% Low Range Ultra Agarose (Biorad).

After the confirmation that library fragments contain Adapter-1 and Adapter-2 on the ends, GW-COBRA libraries are amplified using PCR whereas LA-COBRA libraries are amplified using *in vitro* transcription followed by cDNA synthesis.

Part 2-17. GW-COBRA: PCR Amplification

P2-17.1. Adapter 1 and 2 ligated GW-COBRA library fragments are used as templates with a final concentration of 1 ng/µL in 25 µL using GoTaq white (Promega) enzyme and buffer mixture.

P2-17.2. The Illumina flow cell primers are used in the final concentration of 0.1 µM.

P2-17.3. The temperature cycles for the PCR are: 98 °C 3 min; 98 °C for 15 s, 65 °C for 30 s and 72 °C for 1 min, for 7 cycles; a final extension of 72 °C for 2 min.

Note: It is important to use minimum number of PCR cycles to reduce PCR amplification bias towards CpG poor regions.

Part 2-18. LA-COBRA: *In vitro* Transcription

The LA-COBRA library is *in vitro* transcribed to RNA using T7 RNA Polymerase.

P2-18.1. Prepare the reaction buffer using 600 ng of library with 2 µL of 10 mM of ATP, 2 µL of 10 mM of GTP, 2 µL of 10 mM of UTP and 2 µL of 10 mM of CTP with 2 µL of T7 RNA Polymerase Mix in a total volume of 20 µL.

P2-18.2. Incubate the reaction in 37 °C overnight.

P2-18.3. Clean up using RNeasy MinElute Clean up kit.

P2-18.4. Quantify RNA concentration using NanoDrop 1000 Spectrophotometer with RNA settings as well as Quant-iT RNA assay as per manufacturer's protocol.

Part 2-19. cDNA Synthesis

P2-19.1. Mix approximately 600 ng (8 µL) of LA-COBRA transcripts with 1 µL of 100 µM P5 primer at 70 °C for 10 min then cool on ice for 2 min.

P2-19.2. Add the mixture to the SuperscriptIII first-strand synthesis reaction mix containing a final concentrations of 1× First strand buffer, 10 mM DTT, 1 mM dNTP mix with 40 U RNasin RNase inhibitor with 400 U SuperscriptIII enzyme in a final volume of 20 µL (QuantiTect® Reverse Transcription Kit).

P2-19.3. Incubate the reaction at 42 °C for an hour.

P2-19.4. Add RNase H (5 U) and RNase A (0.5 µg) into the reaction and further incubate at 37 °C for 15 min.

P2-19.5. Mix the first-strand synthesis reaction with a final concentration of 1× NEBuffer 2, 667 nM P7 primer and 0.3 mM dNTP mix in a final volume of 169 µL.

P2-19.6. Denature the final reaction mixture at 96 °C for 3 min and cool the reaction at 37 °C for 5 min.

P2-19.7. Add 5 units of Klenow fragment of DNA polymerase1 (exo-minus) and incubate the reaction at 37 °C for an hour.

Part 2-20. Size Selection by Ampure XP Bead Purification

The size selection is necessary to remove any adapter dimers and excess primers for two main reasons: the excess nucleotides will contribute to the library concentration quantifications and they may with the sequencing step, as the short fragments will amplify efficiently and populate the sequencing reads.

Use the Ampure XP Bead purification on both GW-COBRA and LA-COBRA final library materials as described earlier in Adapter-1 Clean Up step.

Part 2-21. Final Quantifications

DNA concentration in PCR amplified GW-COBRA and linear amplified LA-COBRA libraries post size exclusion is initially quantified using 1 µL of the samples on a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Carlsbad, CA, USA). Considering the libraries contain small amounts of DNA, for accurate measurement, Quant-iT Picogreen dsDNA Reagent was used and the DNA concentrations are calculated according to a standard curve made using the Kit's control sample. Finally, the size distribution is visualized using Agilent DNA 1000 Assay in 2100 Bioanalyzer (Agilent Technologies, Los Angles, CA, USA) using the manufacturer's protocol.

3.2. Supplementary Materials and Methods Cont'd

Simulation Data

For a reasonable simulation we based the model on estimates from empirical data. The distribution of COBRA-seq read count data is zero-inflated negative binomial; some CpG sites are fully unmethylated so no amount of additional sequencing will raise read counts above zero, while for the rest of the CpG sites, read counts are proportional to a combination of sequencing depth and methylation rate. To model this distribution, the negative binomial parameters, mu (µ) and size, were estimated from Joint-COBRA data for sites with counts. In addition, the relationship between Joint-COBRA sites with 0 counts and 450K array beta value was examined and the correlation was found to be negatively exponential in nature.

As beta values are bimodal (most are near 0 or 1) we used empirical data as a source of "true" beta values for the simulation. In total, 200,000 beta values were sampled from 450K array data of HCT116 and these were considered true population means for CpG sites. Mean population methylated read counts were modelled as beta × coverage per site and unmethylated read counts as (beta-1) × coverage per site. These per site read count population means (λ) were used to derive 200,000 random Poisson samples for the methylated (M) and unmethylated (U) fractions and beta was constructed as beta (estimated) = M/(M + U). The estimated betas were plotted against true beta. For the COBRA-seq like, M only simulation, the same set of 200,000 random true betas were used to derive random negative binomial read counts using the parameters estimated from the empirical COBRA-seq data. To represent the additional CpG site

coverage possible when sequencing only the methylated genome fraction, the read coverage was moderated by divided by the mean beta value. Next, some counts were converted to 0. A total of 15.75% of CpG sites shared between the 450K array and Joint-COBRA had 0 counts and we used this as the fraction of CpG sites which are unmethylated—so we would expect no COBRA-seq counts even with high coverage. So, 15.75% of the 200,000 simulated read counts were randomly converted to 0, with a chance weighted by $(\text{beta} + 0.02)^{-2}$. The addition of 0.02 was to moderate the chance of selection of beta values very near 0 and beyond the distribution peak at approximately 0.03.

3.3. Supplementary Results and Discussion

Simulation Data

For absolute DNA methylation estimation it is usual to quantitate the methylated (M) fraction and unmethylated (U) fraction of a biological sample and to divide the methylated fraction by the sum of both fractions (M/(M + U)), to derive the beta value—the proportion of M. With low coverage bisulfite sequencing, this estimate of beta is granular and inaccurate. For example, with a 5-fold coverage at a CpG site, it is only possible to estimate methylation in steps of 20% and often the sampling of methylated and unmethylated reads will highly diverge from the true rate. As coverage increases, precision, accuracy and trueness slowly converge from low to high. If the unmethylated fraction (U) is removed, as in COBRA-seq and other methylated DNA enrichment methods, we observe a count, with the count proportional to M and the library size. While the ability to estimate the true rate of absolute methylation is lost, the property of accuracy and precision increasing with coverage is retained and proportionally, far more reads (than with WGBS) contribute DNA methylation information. We wished to examine this trade-off between absolute methylation estimation (M + U scenario) and more reads contributing methylation information (M-only scenario).

As read coverage increases in the M + U scenario, accuracy in beta estimation (emergence of a diagonal trend between 0 and 1) appears early, with the precision (the decrease in the breath of the line) improving slowly. As beta is ratio-based, over- or under-sampling of the M or U fraction will yield this imprecision. In the M-only case, accuracy and precision are a function of the population methylation, with accuracy and precision decreasing as the methylation rate increases. Very high counts are likely to represent a population methylation rate of 0.8–1.0, however the converse statement cannot be made. High True betas are captured as a wide range of count values.

Collectively, the simulation suggests the M-only case has higher precision than the M + U case for hypomethylated CpG sites and considerably worse precision for hypermethylated sites. However, the increasing imprecision for the M-only case, a dependence of the variance on the mean, can be dealt with by using a variance stabilizing transformation (VST), or similar techniques like those for RNA-seq count data [54]. This suggests COBRA-seq is well suited for finding methylation differences between groups of replicates (where variance can be modelled).

Methylation beta and count estimates are analogous to the FPKM and count-based methods in RNA-seq. It is reasonable to use a RNA-seq count software such as EdgeR to examine differences in methylation between groups of samples. Increasing variance with count is recognized as a property of RNA-seq count data and software such as EdgeR, DESeq2 and Voom will model the distribution and

moderate the counts. Like RNA-seq count-based methods, COBRA-seq is best suited to inter-sample comparisons and not intra-sample comparisons.

CHAPTER 4: Transcriptome profiling of visceral adipose tissue and visceral adipocytes from 26 females with a range of BMI

Visceral adipose tissue (VAT) displays strong associations with metabolic risk factors. In this study, we characterized the transcriptional signatures of human visceral adipose tissue (VAT) and purified visceral adipocytes (VA) across 26 females with a range of BMI (19-50 kg/m^2).

We identified a set of genes and pathways associated with obesity. The main outcome of this study was that BMI, waist girth and hip girth were best correlated with transcriptional profiles. Additionally, we have shortlisted genomic regions of interest as potential biomarkers to improve obesity prevention and treatment strategies.

I was predominantly involved in sample collection, laboratory experiments, interpretation of data and writing the manuscript. Bradford, Clark, Molloy, Ross and myself designed the experiments. Lord, Samaras, Bradford and myself were involved in sample collection and patient selection. Molloy, Ross and myself designed the Roche Nimblegen's SeqCap Epi Target Enrichment Array. Nair undertook the DNA methylation library preparation and sequencing. Anwar, Luu, Ross, Peters, van Dijk, Buckley and Bauer analysed the genomic data.

Detailed contributions:

	Varinli H	**Co-authors**
Experimental Design	50%	50%
Sample Collection	60%	40%
Laboratory Experiments	70%	30%
Data Analysis	20%	80%
Data Interpretation	60%	40%
Writing	90%	10%

Notes: In this Chapter I present gene expression data (stranded RNA-seq) on purified VA and on VAT obtained on the cohort of 26 subjects. These data will be combined with analysis of DNA methylation profiles on the same set of subjects. As indicated in Figure 1, the DNA methylation data comprises targeted bisulphite sequencing of DNA from purified VA of regions covering about 1% of the genome. DNA methylation data using Illumina 450K arrays has also been obtained on peripheral blood leukocyte and buccal cell DNA from these subjects. Analysis of these data is ongoing.

My contributions to the epigenome and SNP profiling are stated below:

Epigenome Profiling:

- I contributed to the design of Roche Nimblegen's SeqCap Epi Target Enrichment.
- I isolated VA from VAT and isolated DNA from VA samples for the epigenome component of this project and planned the batching of samples to avoid technical biases in the analysis of Nimblegen's SeqCap Epi Target Enrichment.
- I isolated DNA from buccal cells for the epigenome component of this project and planned the batching of samples to avoid technical biases in the analysis of 450K array.

SNP profiling:

- I contributed to the optimization of laboratory protocols for DNA isolation from buffy coat samples.
- On surgery days, I pre-processed the buffy coat samples and documented them on a database.

Data Analysis and Interpretation:

- I contributed to the data collection, processing, analyses and discussion.
- I contributed to the manuscript planning and review.

Transcriptome profiling of visceral adipose tissue and isolated adipocytes from females across a range of BMI

Hilal Varinli[1,2,3], Shalima Nair[2], Firoz Anwar[7], Susan van Dijk[1], Phuc-Loi Luu[2], Timothy Peters[2], Reginald Lord[4], Katherine Samaras[5],Denis Bauer[6], Michael Buckley[7], Jason Ross[1], Peter Molloy[1], Susan Clark[2,8] and Stephen Bradford[1,2]

[1] CSIRO Health and Biosecurity, PO BOX: 52 North Ryde, New South Wales 1670, Australia

[2] Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales 2010, Australia

[3] Department of Biological Sciences, Macquarie University, North Ryde, New South Wales 2109, Australia

[4] School of Medicine, University of Notre Dame, Darlinghurst, New South Wales 2010, Australia

[5] St Vincent's Hospital, Darlinghurst, New South Wales 2010, Australia

[6] CSIRO Health and Biosecurity, Riverside Life Sciences Centre, 11 Julius Avenue, North Ryde, New South Wales 2113, Australia

[7] CSIRO Digital Productivity Flagship, Riverside Life Sciences Centre, 11 Julius Avenue, North Ryde, New South Wales 2113, Australia

[8] St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Darlinghurst, New South Wales 2010, Australia

This manuscript is under preparation for *Diabetes*.

* To whom correspondence should be addressed:
Dr. Peter L. Molloy
CSIRO Food and Nutrition Flagship, Genomics and Cellular Sciences
Riverside Corporate Park, 11 Julius Avenue, North Ryde NSW 2113, Australia
Phone: +61 2 9490 5168; Email: peter.molloy@csiro.au

Key words: obesity, BMI, metabolism, adipocytes, adipose tissue

ABSTRACT

Adipose tissue is a key driver of metabolism, energy homoeostasis and inflammation. Perturbations in the amount and distribution of adipose tissue have serious adverse health consequences. We characterised visceral adipose tissue (VAT) and purified adipocytes (VA) of 26 females across a range of BMI (19-50 kg/m^2) to understand at what point in the obesity spectrum molecular signatures change from healthy to unhealthy. Genome-wide transcriptional changes associated with indices of obesity were explored using linear regression analysis. We further examined how these changes relate to changes in DNA methylation in purified visceral adipocytes by undertaking targeted methylome mapping from the same set of subjects using Roche Nimblegen's SeqCap Epi Target Enrichment approach (covering 30 Mb). We observed a clear trend in gene expression in accordance with the severity of obesity. In particular, we noted a transition in the gene expression profiles of VA in individuals with a BMI of 32 kg/m^2 or higher. This might be indicative of a metabolic threshold for potential health risks. Greater numbers of differentially expressed (DE) genes were seen for all obesity measures in VA compared to VAT. This is perhaps due to heterogeneous cellular composition of VAT. Additionally, many of the previously identified genes in VAT are intrinsic to adipocytes. In VA, obesity DE genes were enriched in vesicle transport and cellular organisation pathways while, in VAT, DE genes were enriched in catabolic pathways and leukocyte proliferation among others. The gene lists provided in this study are useful resources for exploration of metabolic changes in obesity that may relate to adverse health outcomes.

1. INTRODUCTION

White adipose tissue (WAT) stores excess energy in the form of triglycerides. WAT is metabolically active and carries out significant functions such as glucose homeostasis, energy metabolism, regulation of food intake and immune function (1). WAT is divided into regional depots with differences in biological function. In humans, the main classes of WAT depots are subcutaneous adipose tissue (SAT) located beneath the skin and visceral adipose tissue (VAT) surrounding the inner organs (2). SAT and VAT have distinct features, and vary in their potential for metabolic risk. Specific differences in gene expression profiles (3) lead to differences in the release of fatty acids and secretion of adipokines, thereby affecting the function of other tissues (4). The functional differences between SAT and VAT are not limited to differences in their adipocytes.

WAT is a loose connective tissue containing mostly adipocytes but also non-adipocyte cells which are collectively termed the stromal vascular fraction (SVF). SVF includes mesenchymal stem cells, fibroblasts, immune cells, endothelial cells and red blood cells (5). SVF is a critical player in the adipose secretome (6) and displays differences in gene expression profiles compared to those of adipocytes from the same depot (7). Additionally, SVF derived from SAT display differences in their proteome (8), gene expression (9) and adipogenic capacity (10) compared to those derived from VAT.

During the development of obesity, adipose tissue expands through both increased cell size of adipocytes (hypertrophy) and increased numbers of adipocytes. Excessive enlargement of adipocytes is believed to contribute to adipose dysfunction through release of pro-inflammatory adipokines and fatty acids (11). Obesity is also associated with an increased number of adipocytes (12) and recent data have shown that this may occur through recruitment of stem cells from other sites, including bone marrow (13; 14).

With the global rise in obesity rates, the anatomical distributions of fat and unique functions of VAT and SAT have assumed considerable medical significance. There are substantial

differences in gene expression in both depots in obese compared with normal weight subjects (15-19). Some of the obesity associated changes are down-regulation of lipolysis genes and up-regulation mitogen-activated protein kinases and inflammation genes which implies that obesity causes dysfunction in adipose tissue. Additionally, some studies have identified associations between DNA methylation and parameters of obesity in functionally important genes mainly in more accessible sample types such as blood, SAT or both (20-27). Most studies to date have been on whole tissue, and mostly on SAT, because VAT is accessible only through surgery under general anesthesia. However, adipocytes from different depots themselves also have distinct characteristics.

Recently, we characterized the differences in gene expression (RNA-seq) and DNA methylation (Whole Genome Bisulfite Sequencing, WGBS) in purified subcutaneous and visceral adipocytes (SA and VA) from a core set of three lean (BMI 19.1 – 25.4) and three obese (BMI 37.7 – 42.1) females (Bradford et al., unpublished). Comparing SA and VA of lean and obese subjects, transcriptional differences were primarily associated with extracellular matrix components and cell migration, angiogenesis, immune/inflammatory function, cytokine production and lipid metabolic processes. Extensive differences in DNA methylation between depots were associated with metabolic functions and a subset of differentially-expressed (DE) genes, but the strongest differential methylation was found at genes encoding transcription factors, particularly of developmental genes such as members of the homeobox (HOX) and T-box (TBX) families. This reflects the recent finding that visceral adipocytes (VA) arise, at least in part, from different developmental origins than subcutaneous adipocytes (SA) (28).

Focusing on visceral tissue, we also identified major changes within the adipocyte populations of the core set of three obese compared with three lean subjects (Bradford et al., unpublished). The major transcriptional changes were associated with lipid metabolic processes, lipoprotein particle components, extracellular matrix, cell migration, angiogenesis

and growth factors. Again the strongest DNA methylation changes were associated with transcription factor and developmental genes.

While the lean vs. obese comparison is powerful for delineating significant changes between the two extreme groups, obesity is a progressive phenotype, with greater health risks as one heads towards the higher-end of the scale. This raises the interesting question: Where along the obesity distribution do molecular signatures change from healthy to unhealthy in purified adipocytes or in adipose tissue?

In this study, we explored genome-wide transcriptional changes associated with indices of obesity in VAT and VA of 26 females across a range of BMI (19-50 kg/m^2) using linear regression analysis. These data allowed us to identify extensive differences in expression profiles and to distinguish those intrinsic to adipocytes from those reflecting changes in other cells in the tissue, particularly immune cells. We further explored how these changes relate to changes in DNA methylation in purified visceral adipocytes by undertaking targeted methylome mapping from the same set of subjects using Roche Nimblegen's SeqCap Epi Target Enrichment approach (covering 30 Mb) (29). A schematic illustration of the study is provided in Figure 1.

2. METHODS

2.1.Subjects

26 female subjects meeting the selection criteria were recruited from February 2012 to November 2013 (Figure 2). All subjects underwent abdominal surgery at either St Vincent's Hospital or Macquarie University Hospital located in Sydney. This study was approved by the Human Research Ethics Committee at both St. Vincent's Hospital (H06/151 and SVH12/20) and Macquarie University Hospital (Ref: 5201300002). All subjects gave informed written consent to participate. The anthropometric and metabolic measurements of final subjects are summarized in Table S1. Because of potential impacts on the epigenome, we excluded subjects who were post-menopausal, current or recent smokers, those using anti-depressant medication or who had undergone previous weight loss surgery.

2.2.Sampling

Blood and saliva samples were collected at the beginning of the surgery while visceral (VAT) and subcutaneous (SAT) adipose tissue was collected from the greater omental region and from the site of surgical incision respectively and snap frozen in liquid nitrogen. Additional samples of SAT and VAT were weighed and minced with surgical scissors. For each gram of minced tissue, 2 mL of HEPES buffer containing 1.25 mg/ mL of Collagenase I (Sigma Aldrich, cat#C6885-1G) was added immediately. This was incubated for 30 – 45 min at 37ºC before being diluted 1:9 with additional HEPES buffer. The mixture was passed through a 250 micron mesh before being centrifuged at 300 rcf for 6 minutes at room temperature. The top lipid layer was removed and the purified visceral and subcutaneous adipocytes (VA and SA) were collected for snap-freezing in liquid nitrogen.

Blood was collected at the beginning of the surgery into 10 mL K2EDTA vacutainers (BD, cat#367525). Following centrifugation at 1,300 rcf for 10 minutes at room temperature, the top plasma layer was removed and the buffy coat layer collected. Cells were washed three times, with vigorous resuspension, in 10 mL Tris-EDTA buffer (100 mM Tris-HCl, 0.1 mM EDTA),

with pellets collected after centrifugation at 10,000 rcf for 10 minutes. The final pellet was resuspended in 500 µL of buffer and stored at -80ºC.

2.3.RNA isolation

VA or VAT samples (200-500 mg) were transferred into ice cold 3 mL of TRI-reagent (Sigma, cat#T9424), homogenized using a tissue disruptor (Kinematica AG, cat# PT120E) with 5mm probe (Kinematica AG, cat#PT-DA1207/2EC-E) at maximum speed. The homogenized samples were transferred to 2 mL tubes in aliquots of 1 mL and total RNA extraction was done as per the manufacturer's protocol. On-column DNase digestion was performed on the eluate using the RNeasy Lipid Tissue Mini Kit (Qiagen, cat#74804). RNA quantity was determined by the Qubit RNA HS Assay using a Qubit Flurometer (Thermo Scientific, Carlsbad, CA, USA). RNA quality was determined by the Agilent RNA Nano Assay using 2100 Bioanalyzer (Agilent Technologies, Los Angeles, CA, USA).

2.4.Whole Transcriptome Sequencing and Analysis

Per subject, 500 ng of RNA (per each VAT and VA samples) was used to prepare RNA-seq libraries with rRNA depletion using Illumina TruSeq Stranded Total RNA Library Prep Kit by the Australian Genome Research Facility (AGRF). In total 54 libraries were multiplexed then run on 11 lanes of paired-end Illumina HiSeq2500 chemistry4 by the AGRF. Two libraries prepared from VAT samples did not meet quality criteria and were excluded from further analyses. On average we obtained 25,362,601 and 19,744,199 reads per VA and VAT library respectively.

The fastq files of all samples were processed using the NGSANE RNA-seq pipeline (30). We first used trimmomatic to remove potential adapter contamination (31). We then mapped the read to the reference genome UCSC_hg19 using TopHat (32) and annotation of known genes from gencode.v14. Reads were then counted for each gene using HTSeq-count (33). The differential expression analysis was performed using edgeR analysis pipeline (34). Statistical analysis to identify differentially-expressed (DE) genes was performed using a linear

regression model. Genes with a count per million (CPM) value less than 1 in half or more of the samples were filtered from the analysis.

2.5.DNA isolation

DNA from the VAT samples was extracted using the DNeasy Blood and Tissue Kit (Qiagen, cat#69504) following an amended protocol for total DNA from animal tissue. Briefly, 30 mg pieces of tissue were homogenised in 360 µL of ALT buffer using a Mini- beadbeater (Bio Spec Products, OK, USA) set at full speed for 2 minutes with a 5 mm stainless steel bead (Qiagen, cat#69989). 40 µL of Proteinase K (20 mg/mL) was added and the solution incubated at 56ºC overnight with constant shaking (350 rpm). Then, 400 µL of Buffer AL was added and this mixture was passed through a DNeasy Mini spin column. The adherent DNA was then washed with Buffers AW1 and AW2 and then eluted in 200 µL of Buffer AE.

DNA from the VA samples was extracted using an in-house method. For every 200 – 500 µg of frozen sample, 1 mL of TP lysis buffer (Tris 50mM pH 7.5, NaCl 0.1M, SDS 0.5%, EDTA 5mM) and 100 µL of proteinase K (20 mg/ mL, Promega, cat#V3021) was applied. This solution was then incubated 1.5 hours at 55ºC with constant shaking (350 rpm). The next day, after centrifugation at 2,000 rpm, the lipid layer was discarded and an equal volume of phenol: chloroform: isoamyl alcohol was applied. After centrifugation the aqueous phase was collected, 50 µg Glycoblue (Life Tech, cat#AM9515) added into 1 ml and DNA ethanol precipitated. After washing in 75% ethanol DNA was resuspended in 100 µL milliQ water. DNA quantity was determined by the Qubit DNA Assay using Qubit Flurometer (Thermo Scientific, Carlsbad, CA, USA). DNA was extracted from buffy coat extracts following the Gentra Puregene Blood Cell Kit (Qiagen, cat#158445). No RNase digestion was performed on the samples prior to SNP array library preparation.

DNA was extracted from buccal cells following an in-house method. Approximately 3 ml of saliva was mixed with 7 ml of Listerine and dispersed via shaking. After centrifugation at 2,700 rcf for 15 min, the pellet of cells were lysed in 2 ml of TP lysis solution (Tris 50mM

pH 7.5, NaCl 0.1M, SDS 0.5%, EDTA 5mM) and 100 µL of proteinase K (20 mg/ mL, Promega, cat#V3021) was applied. From this point on DNA samples were processed exactly as DNA isolated from VA samples.

2.6.SeqCap Epi Target Enrichment Array Design

We designed a 30 Mb capture system targeting specific genomic regions, based on our depot or obesity-associated Differentially-Methylated Regions (DMRs) and publically available studies. SeqCap Epi Target Enrichment array probes (60 – 70 bases) were designed against a set of regions covering ~1% of the genome. The selection of regions is listed in Table 1 and a full list is available as separately downloadable supplementary file.

2.7.SeqCap Epi Target Enrichment Array Sequencing and Analysis

1 µg DNA derived from the VA samples was sheared to range between 180-220 bp. Barcoded genomic libraries were prepared using the Kapa LTP Library prep kit (KAPA Biosystems, cat#KK8230, Boston, MA, USA) from 21 subjects of varying BMI. After bisulphite-treatment of libraries with EZ DNA Methylation-Lightning Kit (Zymo Research, cat#D5030, Irvine, CA, USA), the single-stranded bisulfite DNAs were amplified in low-melting PCR reaction for 13 cycles using KAPA HiFi HotStart ReadyMix (provided in the library kit) as per manufacturer's instructions. SeqCap Epi Target Enrichment Array was carried out as per the Roche Nimblegen's SeqCap Epi Target Enrichment protocol. Along with 21 samples, DNA isolated from LNCaP and PreC cell lines were prepared as control sample libraries. The size distribution of the final libraries was visualized using Agilent DNA 1000 Assay in 2100 Bioanalyzer (Agilent Technologies, Los Angeles, CA, USA) using the manufacturer's protocol. Libraries were multiplexed in groups of 4 and sequenced as a 100bp paired end Rapid run on a V1 flow cell on HiSeq2500. On average 50% of reads obtained were on target with coverage ranging between 150-250x for different subjects.

Table 1: Summary of final SeqCap Epi Target Enrichment Array design for obesity associated Differentially-Methylated Regions (DMRs).

Description:	Min Length (bp):	Final Size (Mbp):	Relevant studies or databases:
DMRs identified using DNA methylation array (450K) data:			
1. VA vs SA in 3 lean individuals	500	2.89	(Bradford et al., unpublished)
2. Lean VA vs obese VA (5 vs 5 individuals)			(Bradford et al., unpublished)
3. D0 pre-adipocytes vs D14 mature adipocytes across SGBS adipogenesis			(Varinli et al 2017, unpublished)
4. BMI associated DMRs in buccal cells			(Varinli et al., unpublished)
DMRs identified using WGBS data:			
1. VA vs SA in 3 lean individuals: DMRs with a stringent t-stat=4 and Δ-me > 0.4	500	0.24	(Bradford et al., unpublished)
2. Lean VA vs obese VA: DMRs with t-stat=3 and Δ-me > 0.1	500	11.59	
Adipocyte distal enhancer elements as low-methylated regions (LMRs) identified using WGBS data:			
1. LMRs overlapping with FANTOM5 enhancers (35) or co-ordinates from open chromatin regions of Transposase Accessible Chromatin with Sequencing (ATAC-seq) performed on lean and obese visceral and subcutaneous progenitor cells	400	9.59	(Bradford et al., unpublished)
2. LMRs with no overlap	200	10.02	
Obesity associated Differentially Methylated Probes (DMPs) identified in adipose tissue	200	2.29	EWAS studies (36-39)
2,248 unique SNPs identified in GWAS associated with obesity parameters	200	0.45	NHGRI-EBI SNP Catalog (40)

2.8.450K Methylation Array Library Preparation and Analysis

We used Illumina Infinium HumanMethylation450 BeadChip (450k) arrays to map the genome-wide DNA methylation patterns of DNA from buccal cells or white blood cells. A total of 26 buffy coat and 27 buccal cell samples were submitted to the Australian Genome Research Facility (Parkville, VIC, Australia), randomized across the array chip and processed following standard protocols.

Data were normalized using the *dasen* method from the Bioconductor software package wateRmelon (41). Probes that failed in one or more of the nine remaining samples based on a detection *p*-value>0.05 were excluded from the analysis. Mean beta values per sample were calculated for all probes on the array, as well as for subsets of probes based on their genomic annotation according to the 450k manifest file. The Welch *t*-test was used for comparison of mean methylation levels between the groups. Differentially Methylated Probes (DMPs) and differentially Methylated Regions (DMRs) were identified using the Bioconductor packages limma (42) and DMRcate v 1.4.2 (43) respectively, using FDR<0.05.

2.9.SNP Array Analysis

Using DNA derived from the buffy coat samples of each subject, Genome-wide human SNP array 6.0 (Affymetrix, cat#901182) were carried out as per the Affymetrix protocol and run on GeneChip Fluidics Station 450.

3. RESULTS

3.1.Anthropometric and metabolic characteristics of the study cohort

Clinical characteristics, anthropometric or metabolic indices of obesity of the subjects included in the current study are shown in Table 2. Anthropometric measures included 4 measurements: BMI; waist circumference; hip circumference; and waist to hip ratio (WHR). Seven metabolic measures included: level of total cholesterol; triglyceride; fasting glucose; insulin; glycated hemoglobin (HbA1c); low-density lipoprotein cholesterol (LDL); and homeostatic model assessment of insulin resistance (HOMA-IR). The participants in the study were aged between 21 and 54 with broad ranges of anthropometric and metabolic indices (Table 2).

3.2. BMI associated gene expression differences in purified visceral adipocytes (VA)

Whole transcriptome RNA-seq (30-60 million reads per sample) was done on VA and VAT RNA from a total of 26 and 24 female subjects respectively across the BMI range (Table S2 and S3). Table 3 summarises the number of significant DE genes associated with anthropometric and metabolic measures. We first explored the overall gene expression changes associated with BMI in purified VA, using component analysis (CA). The CA plot showed that there was a strong association with BMI for gene expression in purified VA, with a clear trend in the first two CA dimensions (Figure 3A). Linear regression analysis showed that expression of a surprisingly high number of genes (5,984) was significantly associated with BMI (FDR<0.05) (Table S2) and nearly all were up-regulated with increasing BMI (Table 3). Comparison with the 402 DE genes identified in the core 6 subjects showed that 342 (85%) were validated in our expanded dataset. A plot of regression coefficients (B-value) of these overlapping genes against their log fold change (logFC) in the earlier comparison of VA of three normal weight and three obese subjects (Bradford et al., unpublished), shows high concordance with the expanded number of subjects (R^2= 0.89, Figure 3B).

A heat map illustration of top 40 genes with the highest change in expression in VA is shown in Figure 4A. The top differentially-expressed (DE) genes were involved in secretory functions (*GREM1*, *LEP*, *SCUBE1*, *CCND1*, *SELE* and *IGSF1*). While lean subjects have similar expression profiles, class I and II obese subjects (30<BMI<38) have more variable expression profiles (Figure 4A). At the upper end of the scale, severely obese subjects have similar expression profiles in VA (Figure 4A). Of the most differentially-expressed genes, Tenomodulin (*TNMD)*, expressing a putative angiogenesis inhibitor, had the highest change in gene expression (B-value=0.205, FDR=6.74e-9) (positively associated with BMI) across the cohort. The regression coefficient (B-value) of 0.205 translated to The regression coefficient (B-value) of 0.205 translated to a difference in gene expression between the top and bottom third ranging from 4.8 to 122.7 rpkm (reads per kilobase per million reads). To avoid the effect of outliers we report the range between 3rd lowest and 3rd highest expressing individuals. Notably a single gene in the top list of DE genes, carbonic anhydrase 3 (*CA3*), showed clearly reduced expression even in moderately obese subjects (B-value= -0.164, range= 16.8-976.7 FDR=2.47e-7) (Figure 4A).

Table 2: Obesity associated anthropometric and metabolic characteristics of the study participants.

Key Parameters	**Mean**	**Range**
Anthropometric indices (n=26):		
BMI (kg/m^2)	33.55	19.1-49.8
Waist circumference (cm)	111.58	77-149.5
Hip circumference (cm)	120.17	89.5-146.5
WHR	0.92	0.8-1.0
Metabolic indices (n=24):		
Total cholesterol (mmol/L)	4.38	3.4-6.7
Triglyceride (mmol/L)	1.15	0.4-2.3
Fasting Glucose (mmol/L)	5.05	3.8-8.1
Insulin (mU/L)	8.15	0.5-23.0
HbA1c (%)	5.20	4.4-6.6
LDL (mmol/L)	2.53	1.7-4.6
HOMA-IR	1.98	0.1-8.2
Other measures:		
Age (yrs)	41	21-54

Table 3: Summary results of differentially expressed genes associated with obesity parameters in VA and VAT. (*) Correlation is calculated as the slope of the regression fit per gene. It is reported as DE genes were significant at FDR<0.05.

Key Parameters	VA		VAT	
	# of Subjects	Total number of DE genes*	# of Subjects	Total number of DE genes*
BMI (kg/m^2)	26	**5,984**	24	**572**
UP:		5,836 (1)		443 (4)
DOWN:		148		129 (0)
Waist (cm)	26	**5,419**	24	**337**
UP:		5,292 (0)		245 (0)
DOWN:		127 (0)		92 (0)
Hip (cm)	26	**6,467**	24	**418**
UP:		6,362 (0)		328 (0)
DOWN:		105 (0)		90 (0)
WHR	26	**15**	24	**24**
UP:		9 (9)		13 (13)
DOWN:		6 (6)		11 (11)
Total cholesterol (mmol/L)	24	**1**	23	**2**
UP:		1 (1)		0
DOWN:		0		2 (2)
Triglyceride (mmol/L)	24	**1**	23	**0**
UP:		1 (1)		0
DOWN:		0		0
Fasting Glucose (mmol/L)	24	**46**	23	**12**
UP:		21 (21)		2 (2)
DOWN:		25 (25)		10 (10)
Insulin (mU/L)	24	**25**	23	**6**
UP:		24 (0)		2 (0)
DOWN:		1 (0)		4 (0)
HbA1c (%)	24	**78**	23	**5**
UP:		54 (54)		3 (3)
DOWN:		24 (24)		2 (2)
LDL (mmol/L)	23	**1**	22	**0**
UP:		1 (1)		0
DOWN:		0		0
HOMA-IR	24	**87**	23	**0**
UP:		75 (67)		0
DOWN:		12 (12)		0

3.3. General analysis of GO pathways in BMI related genes found in VA

In VA, Gene Ontology (GO) analysis found 653 terms associated with BMI using cut-off of FDR<0.05 for terms with a size range of 15≤x≤500 genes (Tables 4 and S4). Pathways involving vesicle formation, vesicle transport, cell migration and cellular organisation were among the most enriched terms associated with BMI in VA (Table S4). We then applied REViGO (44) to remove the redundant GO terms and summarise to major classes. The unique categories (with low dispensability and frequency) were antigen processing (log10p = -10.2), Golgi vesicle transport (log10p = -3.9), angiogenesis (log10p = -7.8), tissue migration (log10p = -7.8) and response to virus (log10p = -7.3) (Figure 5A). Multiple signalling and stress pathways such as Wnt, Notch, MAPK, NF-κB, oxidative stress and DNA damage pathways were categorised under actin cytoskeleton organization by REViGO because it relies on semantic relationship for clustering. These pathways have strong regulatory functions in adipocytes. Surprisingly, immune pathways such as antigen processing and T-cell mediated cytotoxicity were also enriched in isolated VA (Figure 5A). This might be indicative of reduced purity of the VA samples perhaps more so in those isolated from obese donors. In summary, these findings suggest an enrichment of GO processes involving transport of products, tissue functions and immune response and reflect our gene level finding that top BMI associated genes in VA are involved in secretory functions (Figure 5A).

Table 4: Summary results of GO terms associated with obesity parameters in VA and VAT. (*) FDR<0.05 was applied for terms with a size range of 15≤x≤500 genes. (**) GO terms unique to VAT were determined by using DE genes unique to VAT compared to VA.

Key Parameters	**Total number of GO terms***			
	VA	**VAT**	**VA specific***	**VAT specific****
BMI (kg/m^2)	**653**	**75**	**555**	**0**
Waist (cm)	**473**	**163**	**419**	**5**
Hip (cm)	**546**	**38**	**489**	**4**

3.4. Correlations and differences between genes associated with BMI and other obesity-related measures in VA

3.4.1. Other anthropometric measures

Anthropometric measures correlate with increased adiposity. Linear regression analysis showed that expression of high number of genes, were also significantly associated with waist (5,419) and hip circumference (6,467) in VA (Table 3 and S2). *SPP1* and *GREM1* were the top DE genes associated with hip and waist circumference in VA respectively. Over 96% of the genes associated with BMI were also associated either with waist or hip circumference (Figure 6A). In fact, the similarity was also prominent in the ranking of DE genes (Table S2). Additionally, significant overlap was observed between BMI associated GO-terms and their waist or hip circumference associated counterparts (473/653 and 546/653 respectively) (Table S4). The highest number of unique genes was associated with hip circumference (607) (Figure 6A).

A much smaller number of genes (15) were found to be associated with WHR in VA (Table 3 and S2). All of them were also correlated with the other anthropometric indices obtained in the study. Interestingly some were among the top genes in our study (i.e. *CISH, C6, ISL1, BMP* and *SELE*) and they were previously found to be associated with obesity related outcomes (Table 5).

3.4.1. Metabolic measures

Metabolic indices indicate elevated cardio-metabolic risk (45). Among the seven metabolic indices obtained in this study, four were directly related to glucose metabolism (HOMAIR, fasting glucose, insulin, HbA1c) while others were related to lipid metabolism (triglyceride, LDL and total cholesterol) (Table 3 and S2).

The highest number of genes (87 genes in VA) was associated with HOMAIR, an indicator for insulin resistance which included the 35 genes associated with either HbA1c or fasting glucose (Table 3 and Figure 6B). Most of these genes were clusters of differentiation molecules such as *CD14, CD163, CD28, CD33* and *CD248*. Three genes (*VIPR2* also known

as *VPAC2, MUC16* and *C4A*), expressed in VA, were associated with all four glucose metabolism indices. *VIPR2* encodes a receptor and involved in cAMP pathway was among the most highly expressed genes associated with metabolic indices.

There was a single gene associated with blood triglyceride levels: thyroid peroxidase (*TPO*), a membrane glycoprotein which plays a central role in thyroid gland function, and its expression (B-value= 1.30, FDR= 0.044) was positively associated with serum triglyceride levels. Additionally, *B4GALT6*, the only gene positively associated with increased blood levels of LDL and total cholesterol, encoded a membrane glycoprotein that is important for glycolipid biosynthesis.

Table 5: Genes that were associated with WHR and other anthropometric indices in VA which have previously been linked to obesity.

Gene ID:	B-value:	FDR	Associations:	Reference:
ITGA8	0.124	5.17 e4	Highly repressed in differentiated human adipocytes	(46)
CISH	-0.133	5.47e-7	Expression of *CISH* is decreased with diabetes in mice	(47)
C6	-0.129	7.66e-7	Up-regulated under very low caloric intake and down-regulated under fast food diet	(48)
ISL1	-0.088	9.15e-4	Inhibitory effect on *BMP4*	(49)
BMP3	-0.112	2.15e-3	*BMP3* expression was up-regulated in rat model of hyperplasic visceral adiposity	(50)
SELE	0.143	2.32e-4	In extreme obese individuals, *SELE* expression was found to be higher in SAT than VAT	(3)
IL-6	0.132	8.78e-5	Increased in obesity and strong predictor of abnormalities in adipocyte metabolism: i.e. Trans-signalling recruits macrophages to adipose tissue	(51; 52)
ICAM4	0.134	7.37e-5	Expression decreased with weight loss in women	(53)
HOXB9	0.115	1.41e-6	Used as a negative control chromosomal location for LMNA binding site study	(54)

3.5. Obesity associated gene expression differences in visceral adipose tissue (VAT)

This study comprises of transcriptional profiles of VA and VAT from the same individuals. This allowed us to examine the unique roles of adipose tissue as a whole and non-adipocyte component of VAT. The non-adipocyte component of VAT is dynamic including cells such as mesenchymal stem cells, fibroblasts, immune cells, endothelial cells and red blood cells. These cell types have significantly different gene expression profiles than adipocytes and both their expression profiles and their relative proportions in the tissue may vary with obesity.

Component analysis showed that there was an association with BMI for gene expression in VAT, with a detectible trend in the first two CA dimensions (Figure 3C). However, this was not as pronounced as for purified adipocytes (Figure 3A and 3C). Similar to VA, linear regression analysis showed the highest number of DE genes, over 500 genes, associated with BMI, waist and hip circumference in VAT (Table 3 and S3). There was a significant overlap (76%) between BMI and waist or hip circumference associated genes (Figure 6C). The majority of the DE genes in VAT (82-78%) associated with BMI, waist and hip circumference were in common with VA (Figure 7A, B and C). Therefore, adipocytes are the main contributor to the gene expression profiles of VAT.

A heat map illustration of top 40 genes with the highest change in expression in VAT shows more variable expression, especially among overweight individuals when compared with that of VA (Figure 4B). The transcriptional profile of visceral adipose tissue (VAT) varies greatly among individuals, probably due its heterogeneous cellular composition. For most genes, expression increased with BMI, with only 6 in the top list of DE genes (*ANKRD20A8P, C12orf39 (Spexin), RP11-61I13.3, SLC27A2, C6* and *CA3*) showing clearly reduced expression with increasing BMI (Figure 4B).

We compared the regression coefficients with respect to BMI of all genes that reached the minimum read count threshold in VA and VAT (Figure 3D). This plot shows that the majority of the genes showed similar BMI-related changes in tissue and isolated adipocytes;

notably, however, a subset of genes (shown in green) showed much greater changes in VAT. In VAT, GO analysis found 75 terms associated with BMI, 163 with waist circumference and 38 with hip circumference (Tables 4 and S4). There was a significant similarity between BMI associated GO-terms and its waist or hip circumference associated counterparts (Table S4). When the redundant GO-terms associated with BMI were reduced using REViGO, small molecule biosynthetic processes, response to substances such as steroid hormones and alcohol, circulatory system processes and leukocyte proliferation (oxidative stress, cell migration) were among the unique pathways associated with obesity measures in VAT (Figure 5B).

Among metabolic measures, the highest number of genes (12 genes in VAT) was associated with fasting glucose levels in the blood (Table 3 and S3). Of these genes, *SAA2* and *TUBB2A* had increased expression associated with increased levels of fasting glucose. *GSDMB* was the only gene associated with fasting glucose and HbA1c (Figure 6D). No genes were significantly associated with triglyceride, LDL and HOMA-IR levels in VAT (Table 3).

3.6. Unique signatures of VAT that may not be contributed by VA

To understand the relative significance of genes that are predominantly contributed by the non-adipocyte cells in VAT, we made two comparisons. Firstly, we applied a direct comparison of mean gene expression values between VA and VAT to identify genes that are expressed significantly higher in VAT. We have identified 641 genes that were enriched 5-fold or higher in VAT (Table S5). GO analysis of these highlights functions of the immune system (phagocytosis, B- and T-cells function), as well as extracellular matrix organisation, neurogenesis and circulatory system development Of the 641 genes, 38 were significantly associated with BMI in VAT (Table 6). These genes were heavily represented by genes of the immune system (monocyte and macrophage lineages).

Secondly, we compared DE genes identified in VAT to those found in VA. When DE genes associated with BMI, waist or hip circumference or WHR were cross-compared between the

two sample types, we found genes unique (varying between 19-128 genes per parameter) to VAT (Figure 7A, B, C and D). We conducted GO analysis on DE genes unique to VAT (Table 4 and S4). There were no significant GO pathways associated with BMI and specific to VAT. However, catabolic processes were associated with both waist and hip circumference and also specific to VAT (Table S4). To highlight differences in biological pathways associated with anthropometric measures of obesity, we summarised the genes expressed in VA and only in VAT using WebGestalt and illustrated the results in Figure 8A and 8B respectively. Most of the genes highlighted in these figures have been previously associated with obesity.

Table 6: List of 38 genes that have mean gene expression level (MGEL) 5-fold or higher in VAT compared to VA. This indicates a list of genes potentially expressed by non-adipocyte cells. Regression coefficient (B-value) and FDR are reported from BMI association in VAT.

Gene ID:	MGEL in VAT:	MGEL in VA:	Ratio:	BMI B-value:	FDR:
IGKV4-1	24.481	1.070	23.559	10.049	0.0002
IGHM	20.154	1.710	14.563	8.535	0.0007
IGHA1	32.340	2.781	14.161	6.909	0.0013
WIF1	18.004	1.685	13.209	11.381	0.0017
IGKV1-5	11.653	1.083	11.366	7.349	0.0011
IGJ	77.745	9.002	10.595	5.884	0.0068
IGLC2	20.370	2.274	10.233	6.638	0.0181
IGHG3	12.732	1.526	10.125	4.474	0.0223
IGHGP	6.616	0.813	9.708	6.406	0.0016
IGHA2	3.558	0.362	9.518	6.001	0.0166
IGKV3-11	9.109	0.927	9.409	5.327	0.0233
IGHG2	26.728	3.576	9.299	5.134	0.0100
IGHG1	108.443	14.837	9.242	4.355	0.0166
IGKV3-20	12.832	1.538	9.006	6.535	0.0035
IGLC3	15.809	2.073	8.733	5.889	0.0062
KCNH5	2.588	0.400	7.998	4.668	0.0438
IGHV1-18	3.034	0.180	7.502	9.773	0.0030
IGLV2-14	4.499	0.481	7.283	5.299	0.0160
IGHV4-39	2.956	0.232	7.170	7.094	0.0145
IGHV4-34	2.334	0.134	6.956	9.367	0.0026
IGKV1-27	3.013	0.237	6.916	7.219	0.0260
PRR15L	2.306	0.405	6.699	4.068	0.0475
IGHV1-2	3.467	0.413	6.653	10.307	0.0006
IGLV3-21	2.006	0.219	6.566	6.882	0.0168
IGHV3-7	5.191	0.768	6.544	6.485	0.0076
TRPA1	3.621	0.584	6.104	6.029	0.0004
DIAPH3	4.075	0.902	5.830	4.773	0.0389
FENDRR	1.498	0.236	5.634	3.445	0.0429
KRT7	23.179	5.295	5.567	3.531	0.0195
FAM196B	1.920	0.406	5.563	3.845	0.0318
IGLV1-44	2.217	0.283	5.552	5.829	0.0270
MZB1	2.083	0.342	5.471	5.024	0.0172
IGHV3-23	6.328	1.252	5.238	5.057	0.0082
IGLV3-19	2.834	0.413	5.235	5.688	0.0372
IGKV1-9	2.763	0.419	5.226	6.160	0.0164
NAPSB	2.247	0.471	5.132	3.071	0.0163
MPZL2	6.965	1.635	5.131	2.434	0.0456
COL17A1	2.476	0.523	5.070	5.049	0.0264

DISCUSSION

The majority of the obesity indices used in this study were highly associated with gene expression changes, particularly in purified VA. BMI, waist and hip circumference, WHR, HbA1c and HOMA-IR were all strongly associated with gene expression in VA. There were a surprisingly high number of genes associated with BMI in VA. This study showed that gene expression profiles of VAT and changes with obesity are mainly contributed by the adipocytes. Linear regression analysis of VAT gene expression profiles only identified $1/10^{th}$ of the number of statistically significant DE genes as in VA. The number of significant DE genes in VAT might be low due to greater variability between individuals. This might be indicative of heterogeneity in cellular composition of VAT (ratio of different cell types). Hence, purification of adipocytes provided significant advantages to the study outcomes. In particular, we noted a significant transition in the gene expression profiles of VA in individuals with a BMI of 32 kg/m^2 or higher. This might be indicative of a metabolic threshold for potential health risks.

Many of the differentially expressed genes, such as *GREM1, ISL1 LEP, TNMD* and *B4GALT6,* are already recognized in the literature. These genes are involved in adipogenesis, vesicle transport and signalling pathways (insulin, leptin, inflammatory and TNF) (Figure 8A). Previously, *TNMD* expression has been reported to be significantly higher in obese subjects compared to lean subjects in both subcutaneous and visceral adipose tissue (55). *PITX2* has been reported to be highly expressed in adipose stem cells of obese individuals (56). Additionally, expression of both *TNMD* and *B4GALT6* were previously found to be down-regulated with weight loss (57). Our study showed that both *TNMD* and *PITX2* expression in adipose tissue is largely contributed by adipocytes. We also found that *B4GALT6* expression in VA was positively associated with increased blood levels of LDL and total cholesterol.

Additionally, we identified some genes that have been recently discovered such as Spexin, a peptide hormone that is highly down-regulated in human visceral and subcutaneous fat in obesity (58). This study showed that its expression is mainly contributed by the adipocytes. These findings provide additional support for the potential role of adipocyte secretion and signalling in the pathogenesis of obesity, in agreement with previously published studies (59-61). In summary, our data are highly concordant with previously published studies. Therefore, it is worth examining the top genes that have not been previously considered in the context of obesity and metabolic diseases.

In conclusion, the gene expression data establish that purification of adipocytes allowed identification of adipocyte specific expression signatures. This provides a resource for better understanding metabolic changes in obesity that may relate to adverse health outcomes. Further studies on transcriptional regulation of the non-adipocyte component of VAT will be important for understanding the mechanism by which immune cells manifest metabolic alterations in adipose tissue.

DATA ACCESS

Data are available in the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo; Accession XXXXXX).

ACKNOWLEDGMENTS

The study is part of the EpiSCOPE project supported by the Science and Industry Endowment Fund (Australia), Grant RP03-064 and the Department of Biological Sciences in the Faculty of Science and Engineering, Macquarie University. In addition, HV is supported by iMQRES and CSIRO OCE Doctoral scholarships.

AUTHOR CONTRIBUTIONS

Bradford S, Clark SJ, Molloy PL, Ross JP and Varinli H designed the experiments. Lord RV, Samaras K, Bradford S and Varinli H were involved in sample collection and patient selection. Varinli H undertook the RNA transcription sample processing. Molloy PL, Ross JP and Varinli H designed the Roche Nimblegen's SeqCap Epi Target Enrichment Array. Nair S undertook the DNA methylation library preparation and sequencing. Anwar F, Luu P, Ross JP, Peters TJ, van Dijk SJ, Buckley M and Bauer D analysed the genomic data.

DISCLOSURE DECLARATION

Authors declare that there is no conflict of interest.

REFERENCES

1. Rosen ED, Spiegelman BM: Adipocytes as regulators of energy balance and glucose homeostasis. Nature 2006;444:847-853
2. Gesta S, Tseng YH, Kahn CR: Developmental origin of fat: tracking obesity to its source. Cell 2007;131:242-256
3. Gerhard GS, Styer AM, Strodel WE, Roesch SL, Yavorek A, Carey DJ, Wood GC, Petrick AT, Gabrielsen J, Ibele A, Benotti P, Rolston DD, Still CD, Argyropoulos G: Gene expression profiling in subcutaneous, visceral and epigastric adipose tissues of patients with extreme obesity. Int J of Obesity (2005) 2014;38:371-378
4. Bjorndal B, Burri L, Staalesen V, Skorve J, Berge RK: Different adipose depots: their role in the development of metabolic syndrome and mitochondrial response to hypolipidemic agents. J Obesity 2011;2011:490650
5. Zimmerlin L, Donnenberg VS, Pfeifer ME, Meyer EM, Péault B, Rubin JP, Donnenberg AD: Stromal vascular progenitors in adult human adipose tissue. Cytometry Part A 2010;77:22-30
6. Frayn KN, Karpe F, Fielding BA, Macdonald IA, Coppack SW: Integrative physiology of human adipose tissue. Int J Obes Relat Metab Disord 2003;27:875-888
7. Urs S, Smith C, Campbell B, Saxton AM, Taylor J, Zhang B, Snoddy J, Jones Voy B, Moustaid-Moussa N: Gene expression profiling in human preadipocytes and adipocytes by microarray analysis. J Nutr 2004;134:762-770
8. Peinado JR, Jimenez-Gomez Y, Pulido MR, Ortega-Bellido M, Diaz-Lopez C, Padillo FJ, Lopez-Miranda J, Vazquez-Martinez R, Malagon MM: The stromal-vascular fraction of adipose tissue contributes to major differences between subcutaneous and visceral fat depots. Proteomics 2010;10:3356-3366
9. Gesta S, Bluher M, Yamamoto Y, Norris AW, Berndt J, Kralisch S, Boucher J, Lewis C, Kahn CR: Evidence for a role of developmental genes in the origin of obesity and body fat distribution. PNAS 2006;103:6676-6681
10. Liu M, Guo L, Liu Y, Pei Y, Li N, Jin M, Ma L, Li Z, Sun B, Li C: Adipose stromal-vascular fraction-derived paracrine factors regulate adipogenesis. Mol Cell Biochem 2014;385:115-123
11. Exley MA, Hand L, O'Shea D, Lynch L: Interplay between the immune system and adipose tissue in obesity. J Endocrinol 2014;223:R41-48
12. Arner E, Westermark PO, Spalding KL, Britton T, Rydén M, Frisén J, Bernard S, Arner P: Adipocyte turnover: Relevance to human adipose tissue morphology. Diabetes 2010;59:105-109
13. Majka SM, Miller HL, Sullivan T, Erickson PF, Kong R, Weiser-Evans M, Nemenoff R, Moldovan R, Morandi SA, Davis JA, Klemm DJ: Adipose lineage specification of bone marrow-derived myeloid cells. Adipocyte 2012;1:215-229
14. Ryden M, Uzunel M, Hard JL, Borgstrom E, Mold JE, Arner E, Mejhert N, Andersson DP, Widlund Y, Hassan M, Jones CV, Spalding KL, Svahn BM, Ahmadian A, Frisen J, Bernard S, Mattsson J, Arner P: Transplanted bone marrow-derived cells contribute to human adipogenesis. Cell Metab 2015;22:408-417
15. Gomez-Ambrosi J, Catalan V, Diez-Caballero A, Martinez-Cruz LA, Gil MJ, Garcia-Foncillas J, Cienfuegos JA, Salvador J, Mato JM, Fruhbeck G: Gene expression profile of omental adipose tissue in human obesity. FASEB J 2004;18:215-217
16. del Pozo CH, Calvo RM, Vesperinas-Garcia G, Gomez-Ambrosi J, Fruhbeck G, Rubio MA, Obregon MJ: Expression profile in omental and subcutaneous adipose tissue from lean and obese subjects. Repression of lipolytic and lipogenic genes. Obes Surg 2011;21:633-643
17. Aguilera CM, Gomez-Llorente C, Tofe I, Gil-Campos M, Canete R, Gil A: Genome-wide expression in visceral adipose tissue from obese prepubertal children. Int J Mol Sci 2015;16:7723-7737

18. Baranova A, Collantes R, Gowder SJ, Elariny H, Schlauch K, Younoszai A, King S, Randhawa M, Pusulury S, Alsheddi T, Ong JP, Martin LM, Chandhoke V, Younossi ZM: Obesity-related differential gene expression in the visceral adipose tissue. Obes Surg 2005;15:758-765
19. Tam CS, Heilbronn LK, Henegar C, Wong M, Cowell CT, Cowley MJ, Kaplan W, Clement K, Baur LA: An early inflammatory gene profile in visceral adipose tissue in children. Int J Pediatr Obes 2011;6:e360-363
20. Aslibekyan S, Demerath EW, Mendelson M, Zhi D, Guan W, Liang L, Sha J, Pankow JS, Liu C, Irvin MR, Fornage M, Hidalgo B, Lin LA, Thibeault KS, Bressler J, Tsai MY, Grove ML, Hopkins PN, Boerwinkle E, Borecki IB, Ordovas JM, Levy D, Tiwari HK, Absher DM, Arnett DK: Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. Obesity (Silver Spring) 2015;23:1493-1501
21. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH, Hedman AK, Sandling JK, Li LA, Irvin MR, Zhi D, Deloukas P, Liang L, Liu C, Bressler J, Spector TD, North K, Li Y, Absher DM, Levy D, Arnett DK, Fornage M, Pankow JS, Boerwinkle E: Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Genet 2015;24:4464-4479
22. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aissi D, Wahl S, Meduri E, Morange PE, Gagnon F, Grallert H, Waldenberger M, Peters A, Erdmann J, Hengstenberg C, Cambien F, Goodall AH, Ouwehand WH, Schunkert H, Thompson JR, Spector TD, Gieger C, Tregouet DA, Deloukas P, Samani NJ: DNA methylation and body-mass index: a genome-wide analysis. Lancet 2014;383:1990-1998
23. Grundberg E, Meduri E, Sandling JK, Hedman AK, Keildson S, Buil A, Busche S, Yuan W, Nisbet J, Sekowska M, Wilk A, Barrett A, Small KS, Ge B, Caron M, Shin SY, Lathrop M, Dermitzakis ET, McCarthy MI, Spector TD, Bell JT, Deloukas P: Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. Am J Hum Genet 2013;93
24. Keller M, Kralisch S, Rohde K, Schleinitz D, Dietrich A, Schon MR, Gartner D, Lohmann T, Dressler M, Tonjes A, Stumvoll M, Kovacs P, Fasshauer M, Bluher M, Bottcher Y: Global DNA methylation levels in human adipose tissue are related to fat distribution and glucose homeostasis. Diabetologia 2014;57
25. Pietilainen KH, Ismail K, Jarvinen E, Heinonen S, Tummers M, Bollepalli S, Lyle R, Muniandy M, Moilanen E, Hakkarainen A, Lundbom J, Lundbom N, Rissanen A, Kaprio J, Ollikainen M: DNA methylation and gene expression patterns in adipose tissue differ significantly within young adult monozygotic BMI-discordant twin pairs. Int J of Obesity 2016;40(4):654-661
26. Ronn T, Volkov P, Gillberg L, Kokosar M, Perfilyev A, Jacobsen AL, Jorgensen SW, Brons C, Jansson PA, Eriksson KF, Pedersen O, Hansen T, Groop L, Stener-Victorin E, Vaag A, Nilsson E, Ling C: Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. Hum Mol Genet 2015;24:3792-3813
27. Wang X, Zhu H, Snieder H, Su S, Munn D, Harshfield G, Maria BL, Dong Y, Treiber F, Gutin B, Shi H: Obesity related methylation changes in DNA of peripheral blood leukocytes. BMC Med 2010;8
28. Chau YY, Bandiera R, Serrels A, Martinez-Estrada OM, Qing W, Lee M, Slight J, Thornburn A, Berry R, McHaffie S, Stimson RH, Walker BR, Chapuli RM, Schedl A, Hastie N: Visceral and subcutaneous fat have different origins and evidence supports a mesothelial source. Nat Cell Biol 2014;16:367-375
29. Li Q, Suzuki M, Wendt J, Patterson N, Eichten SR, Hermanson PJ, Green D, Jeddeloh J, Richmond T, Rosenbaum H, Burgess D, Springer NM, Greally JM: Post-conversion targeted capture of modified cytosines in mammalian and plant genomes. Nucleic Acids Res 2015;43(12):e81

30. Buske FA, French HJ, Smith MA, Clark SJ, Bauer DC: NGSANE: a lightweight production informatics framework for high-throughput data analysis. Bioinformatics 2014;30:1471-1472
31. Bolger AM, Lohse M, Usadel B: Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114-2120
32. Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009;25:1105-1111
33. Chandramohan R, Wu PY, Phan JH, Wang MD: Benchmarking RNA-Seq quantification tools. Conf Proc IEEE Eng Med Biol Soc 2013;2013:647-650
34. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139-140
35. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jorgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Baillie JK, Ishizu Y, Shimizu Y, Furuhata E, Maeda S, Negishi Y, Mungall CJ, Meehan TF, Lassmann T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Muller F, Consortium TF, Forrest ARR, Carninci P, Rehli M, Sandelin A: An atlas of active enhancers across human cell types and tissues. Nature 2014;507:455-461
36. Agha G, Houseman EA, Kelsey KT, Eaton CB, Buka SL, Loucks EB: Adiposity is associated with DNA methylation profile in adipose tissue. Int J Epidemiol 2015;44:1277-1287
37. Benton MC, Johnstone A, Eccles D, Harmon B, Hayes MT, Lea RA, Griffiths L, Hoffman EP, Stubbs RS, Macartney-Coxson D: An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. Genome Biol 2015;16:1-21
38. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, Bell JT, Yang T-P, Meduri E, Barrett A, Nisbett J, Sekowska M, Wilk A, Shin S-Y, Glass D, Travers M, Min JL, Ring S, Ho K, Thorleifsson G, Kong A, Thorsteindottir U, Ainali C, Dimas AS, Hassanali N, Ingle C, Knowles D, Krestyaninova M, Lowe CE, Di Meglio P, Montgomery SB, Parts L, Potter S, Surdulescu G, Tsaprouni L, Tsoka S, Bataille V, Durbin R, Nestle FO, O'Rahilly S, Soranzo N, Lindgren CM, Zondervan KT, Ahmadi KR, Schadt EE, Stefansson K, Smith GD, McCarthy MI, Deloukas P, Dermitzakis ET, Spector TD: Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet 2012;44:1084-1089
39. Multhaup ML, Seldin MM, Jaffe AE, Lei X, Kirchner H, Mondal P, Li Y, Rodriguez V, Drong A, Hussain M, Lindgren C, McCarthy M, Naslund E, Zierath JR, Wong GW, Feinberg AP: Mouse-human experimental epigenetic analysis unmasks dietary targets and genetic liability for diabetic phenotypes. Cell metabolism 2015;21:138-149
40. (EBI); BT, (NHGRI); HP, (EBI); HE, (NHGRI); HL, (NHGRI); JH, (NHGRI); KA, (EBI); MJ, (NHGRI); MT, (EBI); MJ, (EBI) PHEaWD: The NHGRI-EBI Catalog of published genome-wide association studies. Available at: www.ebi.ac.uk/gwas. Accessed 15/12/2014, version v1.0.
41. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC: A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics 2013;14:293
42. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47
43. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, R VL, Clark SJ, Molloy PL: De novo identification of differentially methylated regions in the human genome. Epigenet Chromatin 2015;8:6
44. Supek F, Bošnjak M, Škunca N, Šmuc T: REVIGO Summarizes and visualizes long lists of gene ontology terms. PLoS ONE 2011;6:e21800

45. Taylor AE, Ebrahim S, Ben-Shlomo Y, Martin RM, Whincup PH, Yarnell JW, Wannamethee SG, Lawlor DA: Comparison of the associations of body mass index and measures of central adiposity and fat mass with coronary heart disease, diabetes, and all-cause mortality: a study using data from 4 UK cohorts. Am J Clin Nutr 2010;91:547-556
46. Morandi EM, Verstappen R, Zwierzina ME, Geley S, Pierer G, Ploner C: ITGAV and ITGA5 diversely regulate proliferation and adipogenic differentiation of human adipose derived stem cells. Sci Rep 2016;6:28889
47. Nadler ST, Stoehr JP, Schueler KL, Tanimoto G, Yandell BS, Attie AD: The expression of adipogenic genes is decreased in obesity and diabetes mellitus. PNAS 2000;97:11371-11376
48. Franck N, Gummesson A, Jernas M, Glad C, Svensson PA, Guillot G, Rudemo M, Nystrom FH, Carlsson LM, Olsson B: Identification of adipocyte genes regulated by caloric intake. J Clin Endocrinol and Metab 2011;96:E413-418
49. Ma X, Yang P, Kaplan WH, Lee BH, Wu LE, Yang JY, Yasunaga M, Sato K, Chisholm DJ, James DE: ISL1 regulates peroxisome proliferator-activated receptor gamma activation and early adipogenesis via bone morphogenetic protein 4-dependent and -independent mechanisms. Mol Cell Biol 2014;34:3607-3617
50. Cernea M, Tang W, Guan H, Yang K: Wisp1 mediates Bmp3-stimulated mesenchymal stem cell proliferation. J Mol Endocrinol 2016;56:39-46
51. Kraakman MJ, Kammoun HL, Allen TL, Deswaerte V, Henstridge DC, Estevez E, Matthews VB, Neill B, White DA, Murphy AJ, Peijs L, Yang C, Risis S, Bruce CR, Du XJ, Bobik A, Lee-Young RS, Kingwell BA, Vasanthakumar A, Shi W, Kallies A, Lancaster GI, Rose-John S, Febbraio MA: Blocking IL-6 trans-signaling prevents high-fat diet-induced adipose tissue macrophage recruitment but does not improve insulin resistance. Cell Metab 2015;21:403-416
52. Trujillo ME, Sullivan S, Harten I, Schneider SH, Greenberg AS, Fried SK: Interleukin-6 regulates human adipose tissue lipid metabolism and leptin production in vitro. J Clin Endocrinol and Metab 2004;89:5577-5582
53. Campbell KL, Foster-Schubert KE, Makar KW, Kratz M, Hagman D, Schur EA, Habermann N, Horton M, Abbenhardt C, Kuan L-Y, Xiao L, Davison J, Morgan M, Wang C-Y, Duggan C, McTiernan A, Ulrich CM: Gene expression changes in adipose tissue with diet- and/or exercise-induced weight loss. Cancer Prev Res 2013;6:217-231
54. Lund E, Oldenburg AR, Delbarre E, Freberg CT, Duband-Goulet I, Eskeland R, Buendia B, Collas P: Lamin A/C-promoter interactions specify chromatin state–dependent transcription outcomes. Genome Research 2013;23:1580-1589
55. Saiki A, Olsson M, Jernas M, Gummesson A, McTernan PG, Andersson J, Jacobson P, Sjoholm K, Olsson B, Yamamura S, Walley A, Froguel P, Carlsson B, Sjostrom L, Svensson PA, Carlsson LM: Tenomodulin is highly expressed in adipose tissue, increased in obesity, and down-regulated during diet-induced weight loss. J Clin Endocrinol and Metab 2009;94:3987-3994
56. Oñate B, Vilahur G, Camino-López S, Díez-Caballero A, Ballesta-López C, Ybarra J, Moscatiello F, Herrero J, Badimon L: Stem cells isolated from adipose tissue of obese patients show changes in their transcriptomic profile that indicate loss in stemcellness and increased commitment to an adipocyte-like phenotype. BMC Genomics 2013;14:625
57. Johansson LE, Danielsson AP, Parikh H, Klintenberg M, Norstrom F, Groop L, Ridderstrale M: Differential gene expression in adipose tissue from obese human subjects during weight loss and weight maintenance. Am J Clin Nutr 2012;96
58. Walewski JL, Ge F, Lobdell H, Levin N, Schwartz GJ, Vasselli J, Pomp A, Dakin G, Berk PD: Spexin is a novel human peptide that reduces adipocyte uptake of long chain fatty acids and causes weight loss in rodents with diet-induced obesity. Obesity (Silver Spring) 2014;22:1643-1652

59. Berry DC, Stenesen D, Zeve D, Graff JM: The developmental origins of adipose tissue. Development 2013;140:3939-3949
60. Luo L, Liu M: Adipose tissue in control of metabolism. J Endocrinol 2016;231:R77-r99
61. Pellegrinelli V, Carobbio S, Vidal-Puig A: Adipose tissue plasticity: how fat depots respond differently to pathophysiological cues. Diabetologia 2016;59:1075-1088

Figure 1: Schematic design of the study. In total, 26 females across a range of BMI (19-50 kg/m2) who were undergoing abdominal surgery were recruited for the study. Visceral adipose tissue (VAT), visceral adipocytes (VA), peripheral blood and saliva samples were collected.

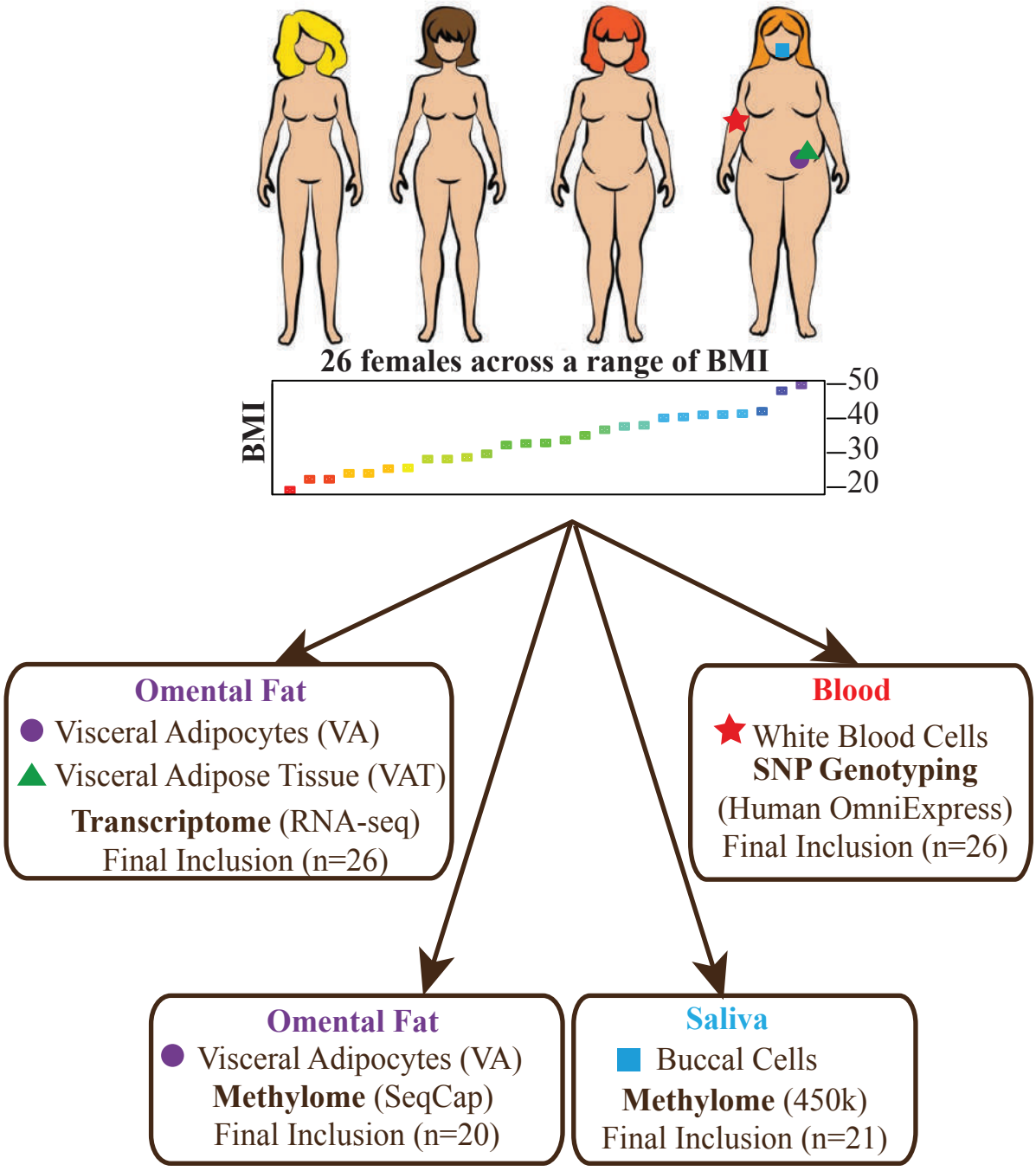


Figure 2: Selection process of individuals included to the study.

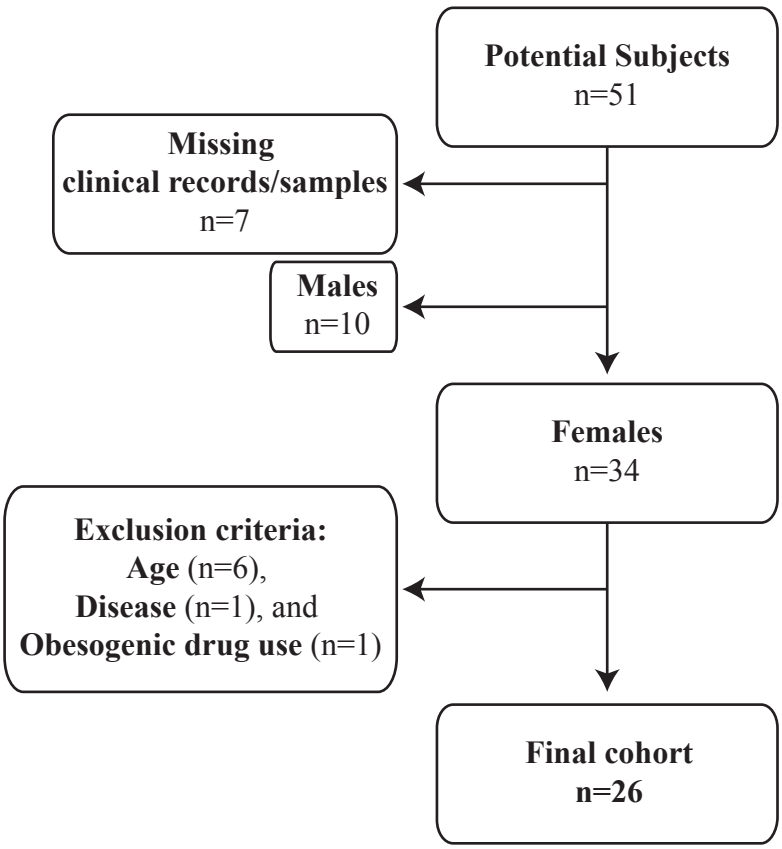


Figure 3: A. Component Analysis (CA) plot of RNA-seq transcription profiles of 26 VA from females across a range of BMI*, **B.** A plot of regression coefficients of the 342 DE genes against their logFC in the core 6 subjects (Bradford et al., unpublished) and our comparison of VA in 24 subjects**, C.** CA plot of RNA-seq transcription profiles of 24 VAT from females across a range of BMI* and **D.** A plot of regression coefficients of all the genes in our comparison of VA in 26 subjects and VAT in 24 subjects**.**
*The colour gradient from red to purple is indicative of BMI increase.

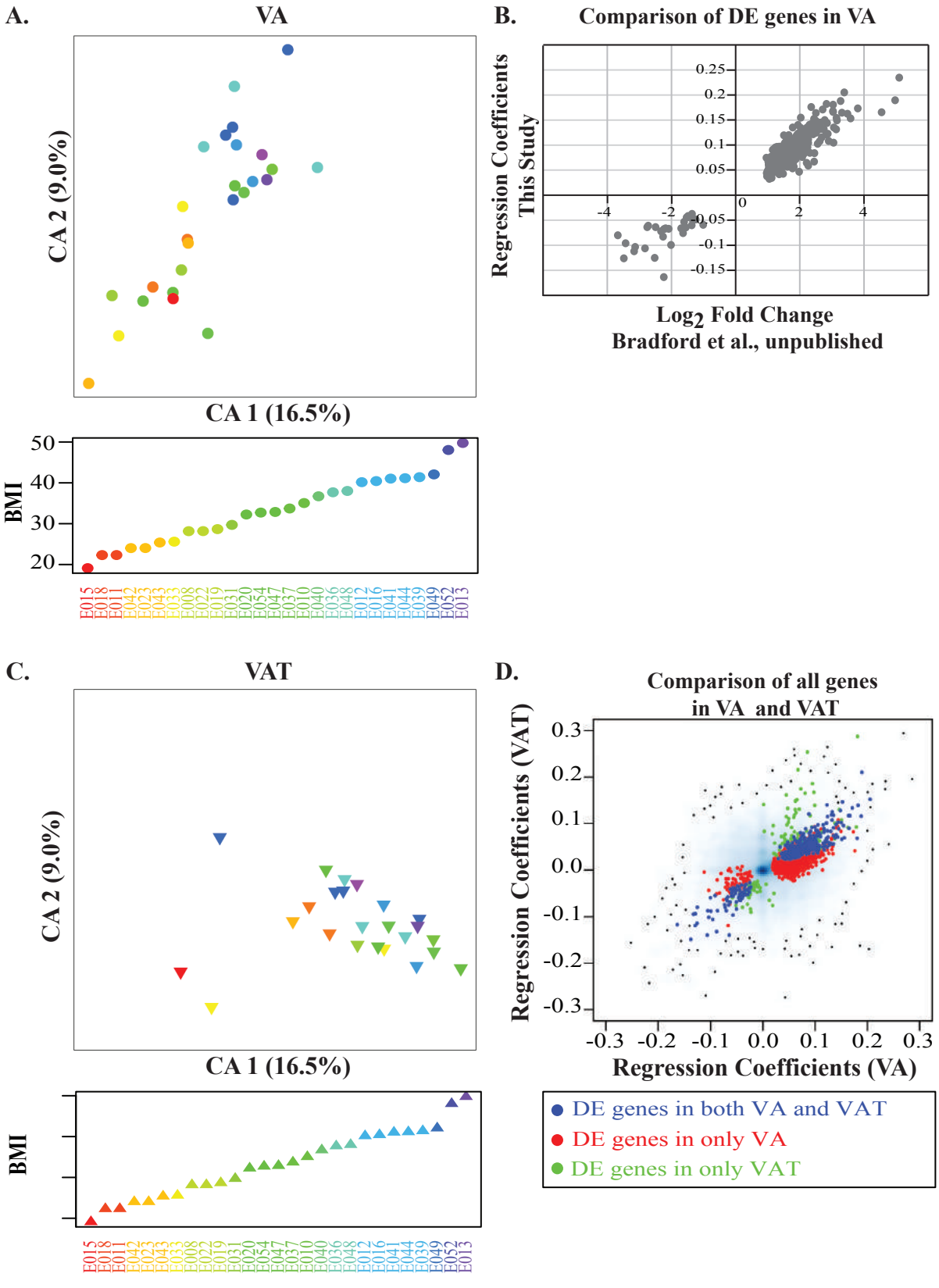


Figure 4: Heat map illustration of differences in gene expression of top 40 DE genes with highest change in **A.** 26 VA samples and **B.** 24 VAT in lean and obese donors (left to right respectively). Red/green scale is from highest to lowest expression scaled separately for each gene.

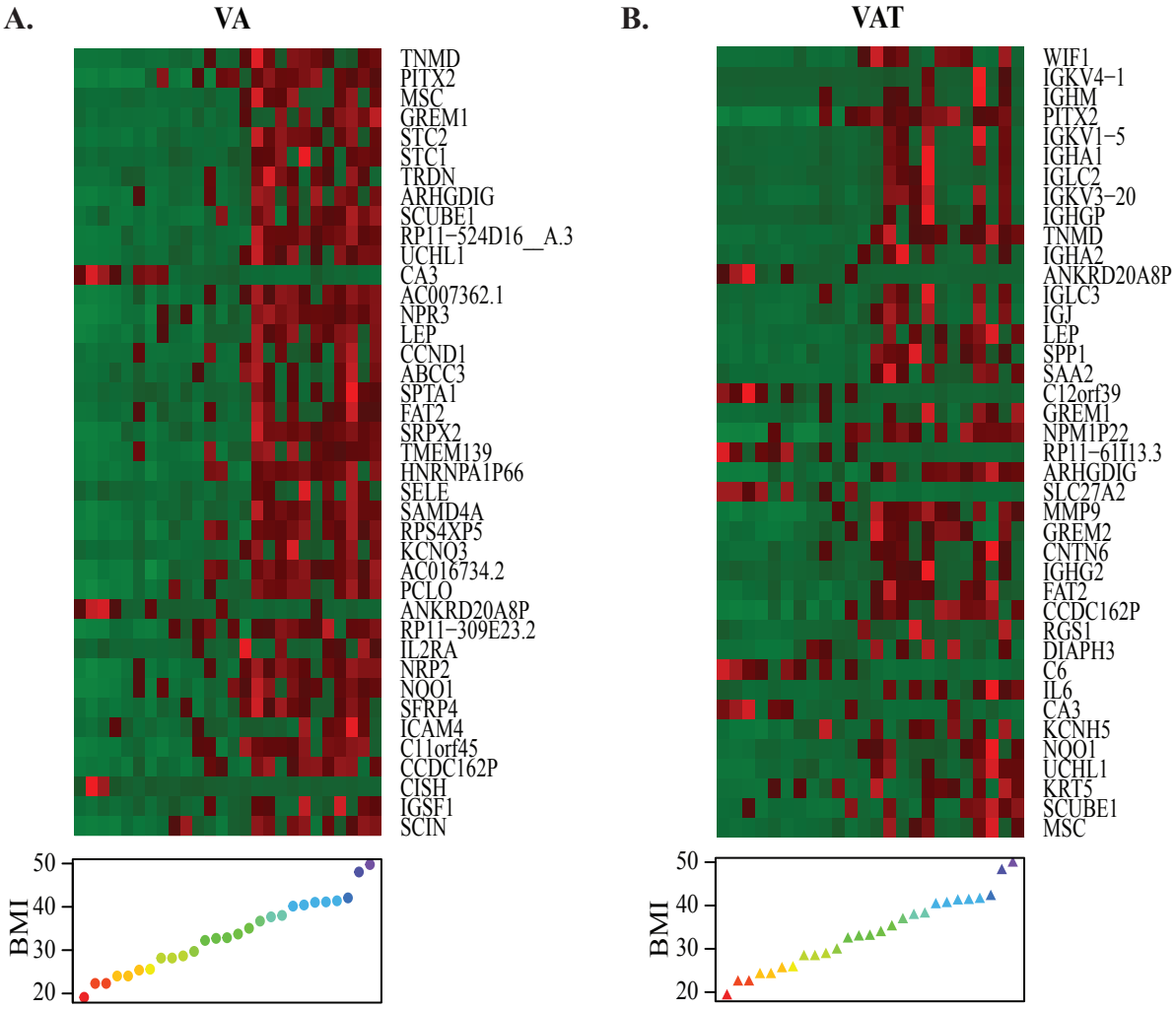


Figure 5A: REVIGO Treemap illustration of biological processes of GO terms significantly associated with BMI in VA. In total there were 653 GO terms significantly associated with BMI in VA. The distribution was as follows: 510 in Biological Processes, 88 in Cellular Component and 55 in Molecular Functions.

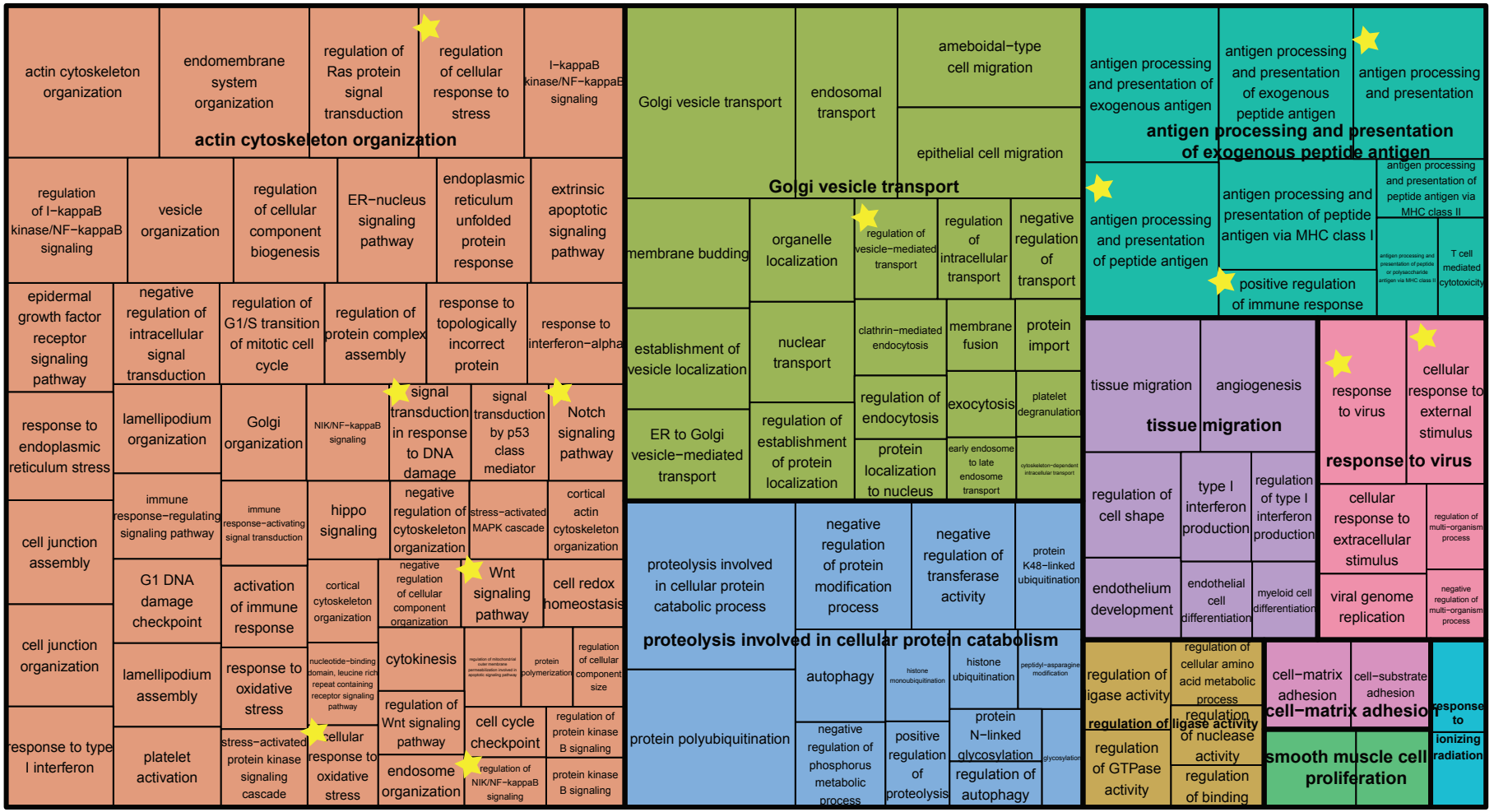


Figure 5B: REVIGO Treemap illustration of biological processes of GO terms significantly associated with BMI in VAT. In total, there were 75 GO terms significantly associated with BMI in VAT. The distribution was as follows: 67 in Biological Processes, 2 in Cellular Component and 6 in Molecular Functions.

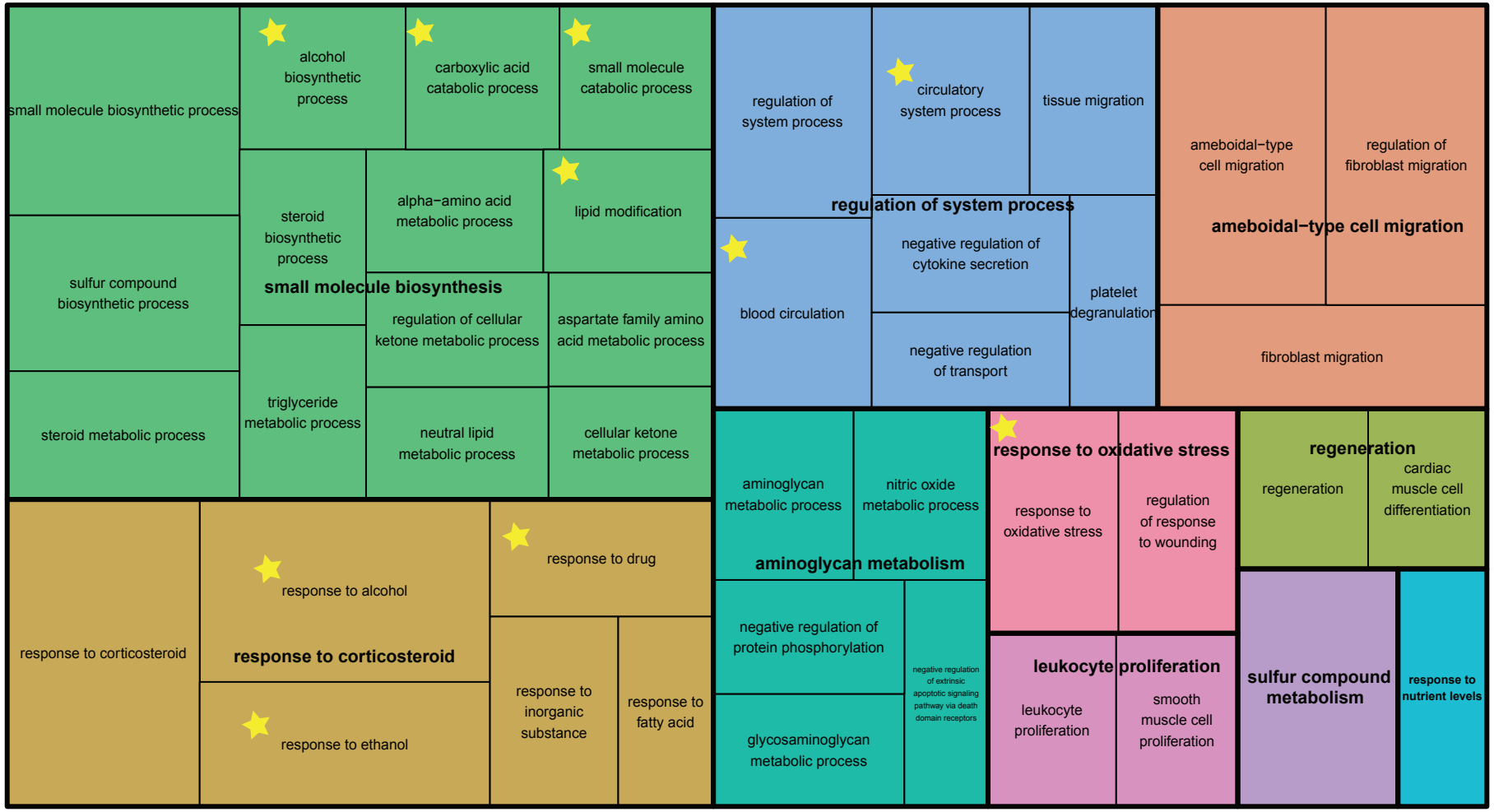


Figure 6: Venn diagrams of differentially expressed (DE) genes associated with **A.** BMI, waist and hip circumference in VA, **B.** HOMA-IR, HbA1c and fasting glucose in VA, **C.** BMI, waist and hip circumference in VAT and **D.** Insulin, HbA1c and fasting glucose in VAT. The number of genes that are not DE in any of the comparisons are indicated in right bottom corner.

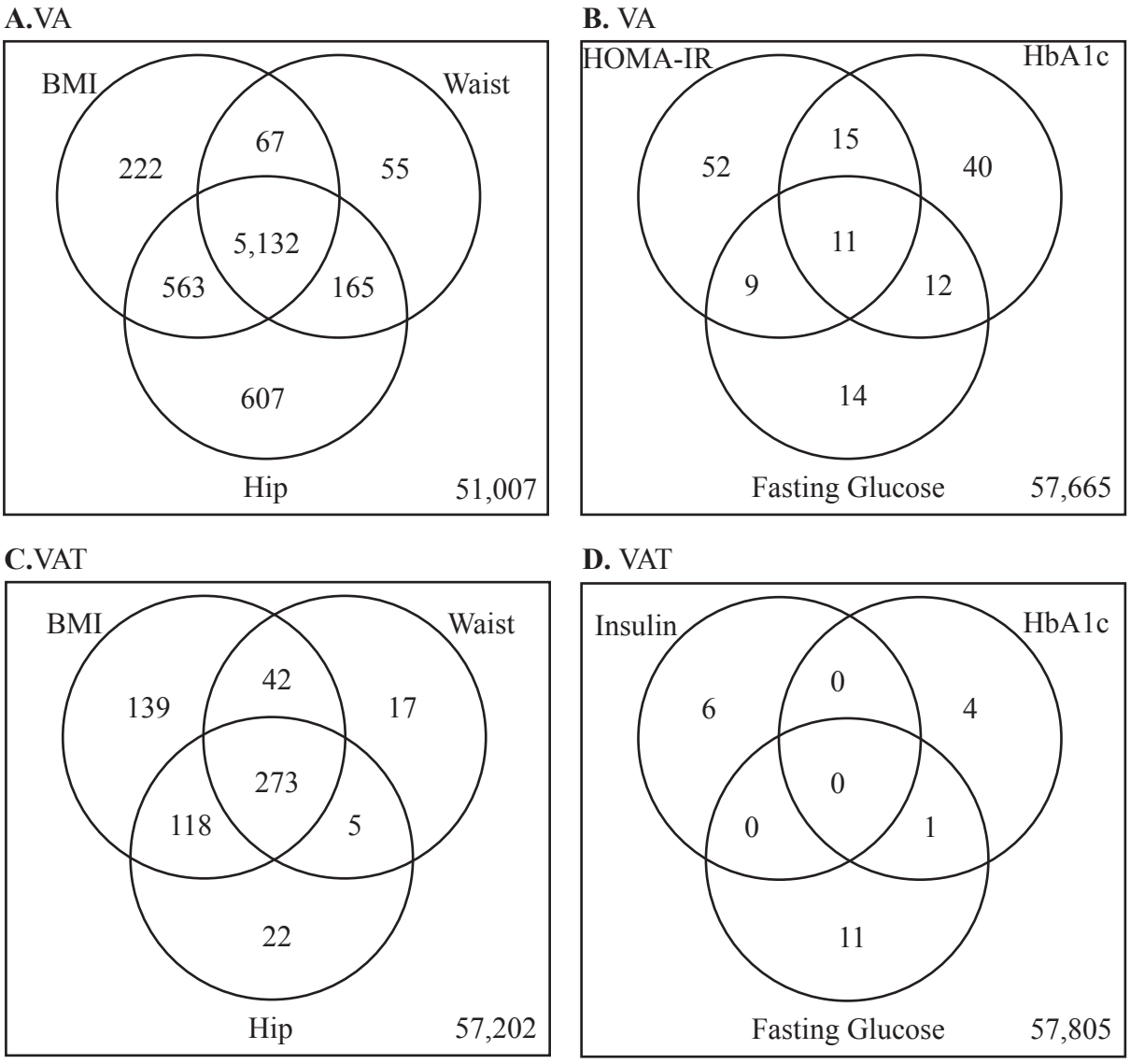


Figure 7: Venn diagrams of differentially expressed (DE) genes in VA and VAT, associated with **A.** BMI, **B.** waist circumference, **C.** hip circumference and **D.** WHR. The number of genes that are not DE in any of the comparisons are indicated in right bottom corner.

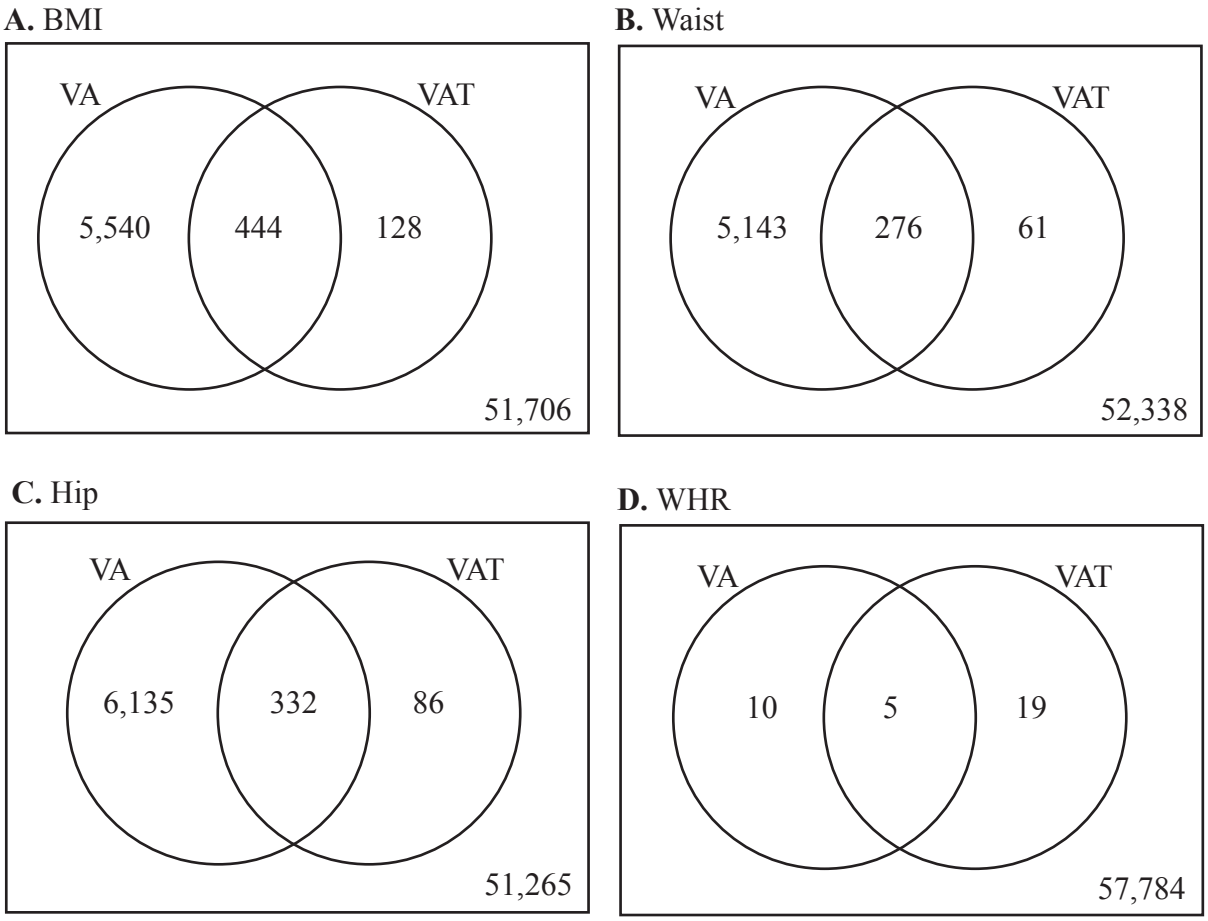
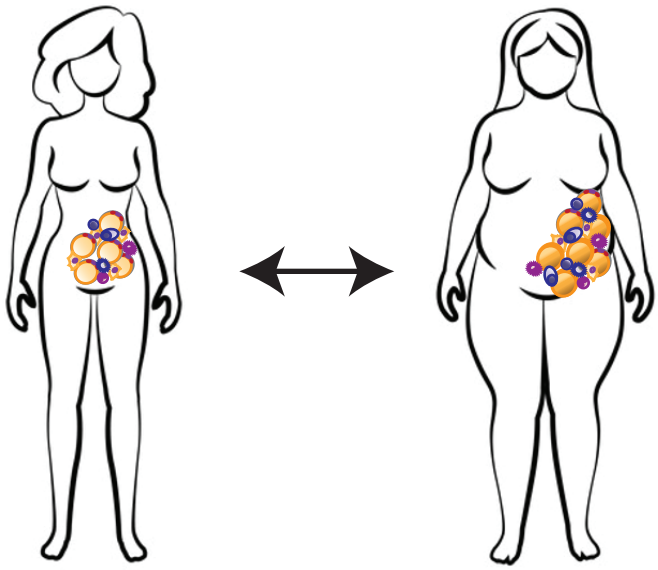


Figure 8A: Summary of 6,811 VA unique DE genes associated with anthropometric measures of obesity contributing to enriched pathways determined by WebGestalt Wikipathways analysis (FDR<0.05). In total 145 pathways were significantly enriched.



Adipogenesis

PPARD, RARA, CREB1, MEF2A, MEF2C, MBNL1, NCOA1, NRF1, NCOR1, GATA3, GTF3A, SLC2A, LIPE, LEP, IL6, AGT, BMP1, BMP2, BMP3, BMP4, KLF7, EGR2, EPAS1, NCOA2, PCK2, CNTFR, LIFR, IL6ST, CTNNB1, SMAD3, DVL1, E2F1, E2F4, RBL2, RB1, HMGA1, STAT1, STAT2, STAT3, STAT5SA/B, STAT6, HIF1A, CUGBP1, AHR, WWTR1, SERPINE1, SCD, SFRP4

Synaptic Vesicle Pathway

SYN2, UNC13B, STXB, STX2, STX3, NAPA, NSF, ATP1A2, CLN8, AP2A2, AP2B1, AP2M1, AP2S1, CLTC, CLTCL1, DNM1L, DNM2, DNM3

Insulin Signalling

SLC2A4, PIK3CB, PIK3CG, PIK3C2A, PIC3C3, PIK3R1, PIK3R3, PIK3RK4, PIK3, SGK, PDPK1, AKT1, GSK3A, GSK3B, FLOT1, FLOT2, STXBP3, STX4A, SNAP23, SORBS1, CAP1, CBL, CRK, RAPGEFF1, RHOQ, KIF5A, MYO1C, ARF1, ARF6, RHOJ, RAB4A, EHD2, PSCD3, LIPE, GYG, IGF1R, SHC1, TSC1, RHEB, FRAP2, RPS6KB1, LIPE, GYG, PTPN1, ENP, PTPN1, ENPP1, PTPN11, INPP4A, IKBKB, PRKCD, PRKCI, PRKAA1, HRAS, GRB14, SOS1, MAPK1, MAPK3, MAPK6, MAPK9, MAP2K1, MAP3K6, MAP4K2, MAP4K3, MAP2K6, MAP4K5, MINK1, MAP3K2, MAP3K3, MAP3K7, MAP3K10, MAP3K14, MAP2K4, MAP3K13

Glycogen Metabolism

PYGB, PHKA2, PHKB, CALM1, CALM3, UGP2, GYG, GSK3A/B, PPP2CA/B, PPP2R2A, PPP2R3A, PPP2R5B/C/E, PPP2R4

Leptin Signalling Pathway

LEP, LEPR, ERBB2, EGFR, PTPN1, SRC, FYN, CISH, JAK1, JAK2, PTPN11, RHOA, CDC42, RAC1, GRB2, SOS1, STAT5B, PIK3R1, HRAS,CFL2, PLCG1, AKT1, RAF1,MAP2K1, MAPK3, MAPK1, IKBKB, CHUK, IKBKG, CREB1, CCD1, RELA, NFKB1, REL, BCL2L1, STAT1. STAT3. NCOA1. IL1RN

Oxidative Stress

NFE2L2, NFKB1, NFIX,TXNRD1, GPX1, NQO1, NOX4, GSR, MGST1, SOD2, SOD3

Inflammatory Response Pathway

IL2RA, CD28, CD86, IL4R, TNFRSF1A/B, LAMA5, LAMB1, LAMB2, LAMC1,COL3A1

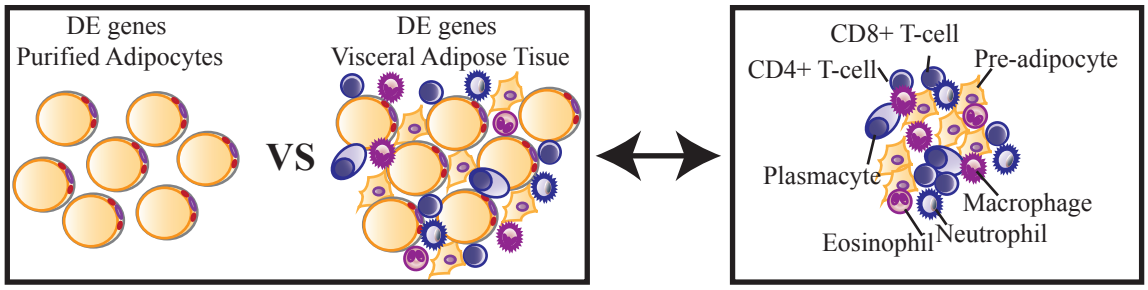
Angiogenesis

NOX4, ROBO4, FGF1, FGFR1, PLCG1, GRB2, SOS1, RAS, RAC1, MAP3K7, RAF1, MAP2K1, MAP2K4, MAP2K6, ANGPT1, TEK, GRB14, DAG1, SRC, PXN, SEMA4A, SEMA3A, PLXND1, NRP1, PXN, IQGAP1,DAG1, PDK1, MAPK1, CD42, NOS3, MTOR, ANGPT1, TEK, PTPN11,GRB14, TIMP4, TIMP2, MMP14, NCK1, CRK, ALT1, APC

TNF Signalling

TNFRSF1A/B, MADD, NSMAF, GRB2, TRADD, SOS1, RFK, CFLAR, TXN, TRAF1, FRADD, RFFL, RAC1, RAF1, TNFAIP3, IKBKB, TBK1, RIPK1, OTUD7B, TAB2, TAB3, TAB4, CASP8, CHUK, MAP3K7, MAP4K2, MAP3K3, MAP2K6, MAPK3, MAP2K4, APAF1, RELA, CASP3, PPP2CA, NFKB1A, NFKBIB/E, BIRC3, IL6, NFKBIE, BCL2L1, BID, RIPK3, PSMD2

Figure 8B: Summary of 170 VAT unique DE genes associated with anthropometric measures of obesity contributing to enriched pathways determined by WebGestalt Wikipathways analysis (FDR<0.05). In total 21 pathways were significantly enriched.



Adipogenesis
IGF1, PPARGC1A, IRS1, LPIN1, PCK2

Insulin Signalling
PIK3C2G, IRS1, MAP3K5

Fatty Acid Biosynthesis
PC, HADHSC

Lipid Metabolism
CYP4B1, ABCD2

Fatty Acid Beta Oxidation/ Tryptophan metabolism
GCDH, HADHSC

Benzo(a)pyrene metabolism
AKR1C2, AKR1C3

DNA Damage
PEPCK, WNT3, IRS1

Hypertrophy
GDF8, VEGF

Inflammatory Response Pathways
IGHA1, IGHA2

AGE/RAGE Signalling Pathway
MSR1, IRS1, MMP9

Osteopontin Signalling
SPP1, MMP9

Endochondrial Ossification
IGF1, VEGFA, MMP9

SIDS Susceptibility Pathways
IL10, IL1RN, VEGFA, PPARGC1A

DISCUSSION AND CONCLUSION

The worldwide prevalence of obesity and Type 2 Diabetes (TD2) has continued to rise rapidly. Obesity threatens to shorten the life expectancy and reduce health-related life quality of current and future generations. It places a significant economic burden on healthcare systems. A better understanding of factors contributing to the development, growth and regulation of fat cells may help us to understand triggers for the progression of obesity and regulation of lipid storage. These are also essential to improve prevention and treatment strategies of obesity.

Primary cause of obesity is excessive total body fat due to excessive lipid storage and altered lipid mobilization by fat cells. Accurate measurement of lipid storage capacity of fat cells is important for exploring lipid dynamics. My PhD thesis describes a novel label-free lipid droplet quantification technique to monitor fat accumulation in live adipocytes (Chapter 2). This tool also provides an ease to determine the effect of anti-obesogenic therapeutics on adipocytes.

Obesity is associated with hyperglycemia, likely due to impaired suppression of hepatic glucose production in response to insulin (insulin resistance). This thesis investigates the potential role of epigenetics as a mediator of gene-environment interactions, manifesting functional changes in fat cells that might be targeted to improve obesity prevention and treatment strategies. Chapter 1 provides insight to the physiological, metabolic, transcriptional and epigenetics trajectories during human adipocyte differentiation, particularly under high glucose exposure.

Adipose tissue is an accessible and rich source of multipotent stem cells. Adipose-derived stem cells (ADSCs) can be isolated from subcutaneous adipose tissue via minimally invasive procedures (32, 162). Therefore, they are good resources for studying adipocyte differentiation *in vitro*. In fact, some commercial companies provide vials of human ADSCs and proprietary medium. Commercially available ADSCs are, however expensive, and difficult to source visceral and subcutaneous ADSCs from the same individuals or even from

individuals with similar health records. Additionally, there is no standard *in vitro* culture for ADSCs and often they are exposed to various culture conditions with varying durations of exposure to adipogenic agents (163). Human ADSCs tend to lose proliferative response in the later passages (>14 passages) (164). When needed in large quantities (i.e. for drug testing), primary adipocytes do not provide a useful model system (165). Additionally, their abundance and capacity for differentiation greatly vary among fat depots and donors with different adiposity level and sex (62, 166, 167).

On the other hand, the human SGBS pre-adipocytes derived from subcutaneous adipose tissue are relatively homogeneous with high differentiation capacity yet they are not immortalized (80). Therefore I used these cells as an adipocyte differentiation model.

1. **Novel method to quantify lipid accumulation**

My initial work focused on measuring physiological response of SGBS human pre-adipocyte cells to external stimuli during *in vitro* differentiation. In particular, I investigated changes in lipid accumulation due to potential anti-obesogenic (docosahexaenoic acid (DHA) and rosiglitazone) and pro-obesogenic substances (D-glucose and zinc oxide nanoparticles). This was achieved by the development of a novel method named LipiD-QuanT published in Journal of Lipid Research (168).

Accurate measurement of lipid accumulation *in vitro* permits insights into physiological and pathological processes. LipiD-QuanT is a label-free lipid droplet (LD) quantification technique to monitor lipid dynamics, based on automated image analysis of phase contrast microscopy images. It does not require dedicated culture plates to measure lipid content and hence permits reuse of cells for other biological or morphological measurements. Furthermore, its application is not limited to a specific cell or organism and it can be adapted to quantify lipid content in tissue sections and emulsions.

In Chapter one, I demonstrated that LipiD-QuanT is a robust, non-destructive, time and cost effective method compared to other triglyceride accumulation assays used for *in vitro* differentiation studies. Therefore, LipiD-QuanT is an improvement for quantifying lipids, not

necessarily only in adipocytes. To date, this is the only label-free and freely available software to study lipid droplet dynamics. The major limitation of the current version of LipiD-QuanT is that it is applicable to quantify lipid droplet size in images of cells in culture, but not in images of tissue sections. It would be worthwhile to extend its application to tissue. This would require optimising parameters to determine the lipid droplet edges when their shapes are not round, but would allow droplet analysis on sections prepared for techniques such as immunofluorescence and immunohistochemistry.

2. Perspectives on *in vitro* human adipogenesis studies

2.1. SGBS pre-adipocyte cellular model

Chapter 1 also demonstrated that Simson-Gobali-Behmel Syndrome (SGBS) pre-adipocytes provide an excellent and easy-to-use *in vitro* model to study human adipogenesis. SGBS adipogenesis was highly reproducible and provided an inexpensive model to study the potential effects of substances with anti-obesogenic (docosahexaenoic acid (DHA) and rosiglitazone) and pro-obesogenic (D-glucose and zinc oxide nanoparticles) properties. I showed that 2 μmol/L rosiglitazone treatment during adipogenesis reduced lipid production and caused a negative shift in lipid droplet diameter size distribution, while other treatments showed no effect under the conditions used.

In addition to the easy-to-use properties of SGBS pre-adipocytes as a model, Chapter 2 demonstrated that human SGBS pre-adipocyte cells had similar gene expression characteristics to isolated primary human adipocytes. Chapter 2 was the first comprehensive study on human SGBS cell lines to date, combining comprehensive genome-wide analysis of both transcriptome and DNA methylome.

2.2. Other available human adipocyte cellular models

Experimental alternatives to SGBS pre-adipocytes are scarce. Unfortunately, there are only a handful of human cell lines that are available to study adipogenesis *in vitro* (Table 3). These cells are derived from either subcutaneous or brown adipose depots or cancerous tissue types. Among subcutaneous adipose tissue-derived human pre-adipocyte cell lines, the SGBS cell line is the only one that is neither transformed nor immortalized. Its capacity for sustained culture does suggest that the underlying mutation in the child from whom the cells were derived partially relieves growth constraints. There are a few connective tissue derived pre-adipocyte cells lines, however all these originated from adults with liposarcoma (LS14, Lisa-2 and SW872).

Table 3: Available models to study human adipogenesis *in vitro*.

Cell line name:	Depot origin:	Details:	Reference:
SGBS	Subcutaneous adipose tissue	Infant with Simpson-Gobali-Behmel Syndrome Pre- and post-natal overgrowth	(169)
Chub-S7	Subcutaneous adipose tissue	Immortalization via retroviral infection (*TERT* and *HPV E7*)	Darimont, 2003 #202}
FTO-deficient SGBS	Subcutaneous adipose tissue	Complete knockdown of *FTO* in SGBS cells	(170)
LS14 cells	Connective tissue	52-yr-old man with a recurrent liposarcoma	(171)
Lisa-2	Connective tissue	53-yr-old man with retroperitoneal liposarcoma	(172)
SW872	Connective tissue	36-yr-old man with a recurrent liposarcoma	ATCC (Rockville, MD)
HPB-AML-I	Peripheral blood mononuclear cells	A patient with acute myeloid leukaemia	(173)
PAZ6	Brown adipose tissue	From an infant Immortalization via simian virus 40 t and T antigen gene injection	(174)
Shingo	Brown adipose tissue	From supraclavicular of 30 and 43-yr-old lean subjects. Immortalized as for PAZ6.	(175)

2.3. Advantages of the establishment of SGBS pre-adipocyte like model derived from visceral adipocytes

While visceral adiposity is linked with metabolic alterations, including insulin resistance, subcutaneous adiposity is known to be benign and to protect against metabolic dysfunctions. Visceral and subcutaneous adipocytes, so-called "bad and good fat cells", have different developmental origins (71) and exhibit physiological and functional differences according to their depot of origin (176). As indicated in Table 3, there is no equivalent of SGBS pre-adipocytes derived from visceral adipose tissue.

Establishment of pre-adipocyte cell lines derived from human visceral tissue could become an essential tool to address these questions:

- What conditions could maintain/reverse the "bad and good" phenotype of adipocytes?
- What are the unique and similar responses of subcutaneous and visceral adipocytes to external stimuli?
- What signalling pathways are involved in "good" functions of adipocytes? Can these be stimulated?
- What factors stimulate/reduce nutrient catabolism by fat cells?
- Can adipocyte expansion be regulated under excessive nutrient availability?
- Are there therapies that can tackle obesity in a depot specific manner?

3. Hyperglycemic exposure conveys significant risks for adipocyte functions

Hyperglycemia, defined as elevated fasting blood glucose, is a pre-diabetic condition associated with poor health outcomes. Obesity is associated with fasting hyperglycemia (177), likely due to impaired suppression of hepatic glucose production in response to insulin (insulin resistance). Increased BMI was associated with a five-fold increase in the incidence of Type 2-Diabetes (T2D) in the AusDiab longitudinal study (178). In Chapter 2 of this PhD

thesis, I focused on the identification of the dynamics of changes in DNA methylation across human adipocyte differentiation, its relationship to transcriptional changes, and how these might be affected by elevated glucose exposure.

In response to hyperglycemia, SGBS adipocytes grew larger and had reduced relative insulin responsiveness. There were 27 genes differentially expressed and 61 differentially methylated regions due to high glucose exposure. Nearly all differentially methylated sites showed reduced methylation. Hyperglycemia-induced gene expression changes were modest when genes were assessed individually, however gene set enrichment analysis defined gene sets that were collectively affected by the glycemic exposure. In fact, the overwhelming majority of the down-regulated genes due to hyperglycemia were enriched in mitochondrial, signal transduction and extracellular stress-activated oxidative stress pathways. The results indicate reduced mitochondrial function and oxidative phosphorylation, and suggest a greater reliance on glycolysis in the presence of elevated glucose levels. Hence, produce energy inefficiently via a combination of oxidative phosphorylation and aerobic glycolysis. The recourse of aerobic glycolysis is a Warburg-like effect such as observed in fast-growing cancer cells and also in murine adipocytes exposed to hyperglycemia under abundant oxygen availability (179).

4. Epigenetic regulation of gene expression during human adipogenesis

Chapter 2 also demonstrated that the transcriptional profiles are distinctly different while the methylation profiles are profoundly similar between pre-adipocytes and mature adipocytes. A total of 83 genes had differential expression and overlapping differential methylation at their transcription start site. In fact, 67 of these were up-regulated during adipogenesis and had reduced methylation at their transcription start site. Some of these genes were key adipogenesis regulators, and obesity or inflammation associated genes such as *ADIPOQ*, *CIDEC*, *LIPE*, *FASN* and *CH13L1*. The integration of our work with a publically available SBGS study (180) established potential genomic locations where methylation signatures might be controlled by transcription factors during human adipogenesis.

5. Published studies on the SGBS pre-adipocyte model throughout my PhD

Since its establishment in 2001, SGBS has become a popular cellular model system to study various aspects of human adipogenesis *in vitro*. It has been used in nearly 100 research articles and proven to be a reproducible model to biological mechanisms in pre-adipocytes and mature adipocytes in context of human adiposity. These studies were examining biological processes in adipocytes or response of adipocytes to external stimuli. In order to summarise the use of these cells lines and the contribution of Chapter 2 in this field, I reviewed the findings of the research studies on SGBS published throughout my PhD. The PubMed database was searched for studies using the SGBS pre-adipocyte cellular model, that were published between 15 December 2011 and 15 October 2016, using search terms SGBS in title or abstract. From the 94 articles identified, 61 studies were included and reviewed. We also identified additional 6 studies on SGBS published within the period but missed out by the search criteria (165, 170, 181-185). These were included to the review. From the 94 articles identified by the search strategy, 33 were excluded; 13 were conducted on humans with Simpson-Gobali-Behmel Syndrome (SGBS), 19 used SGBS as an abbreviation for another term (i.e. stellate ganglion block) and one was not a full article. Of the 67 studies included and reviewed; 42 were on adipose biology, 2 were methodology papers and 23 studies assessed the effect of certain exposures such as diet (9 studies), potential drugs (6), environmental exposure (4) and hormones (4) in human adipocytes. In a nutshell, among the published studies that used SGBS model, few looked into adipocyte metabolism or gene regulation in SGBS adipocytes similar to Chapter 1. The two studies investigated the perturbations in glucose metabolism when SGBS adipocytes were exposed to low glucose conditions (≤1 mM for 4 days) and found reduction in inflammatory phenotypes (186, 187). Additionally, a comprehensive analysis of transcript-level metabolism regulation including genome-wide ChIP-seq profiles for PPARγ, CEBPα, LXR and H3K4me3 and micro RNA (miRNA) target identification for miR-27a, miR-29a and miR-222in SGBS adipocytes was published (180). In fact, Chapter 1 examined transcriptional regulation across

SGBS adipogenesis by conducting genome-wide integration analysis of DNA methylation profiles with recently the ChIP-seq for the H3K4me3 promoter mark and three transcription factors; PPARG, CEBPA and LXR published in this study (180). The key studies and their main findings are summarised in Table 4. This comparison provides a unique set of genes that their function should be studied in further in the context of obesity.

Table 4: Key studies on SGBS cells that are published during the course of my PhD.

Objective	Summary	Reference
Diet		
Effect of pre-exposure to oleic acid and hydroxytyrosol on TNFα-induced decrease of adiponectin secretion in adipocytes	TNFα induced reduction in the adiponectin protein expression and release in SGBS cells. However, when pre-treated with 10 µm oleic acid and 10 µm hydroxytyrosol the adiponectin expression level was not changed.	(188)
Effect of short chain fatty acid (SCFA) exposure on fat storage	SCFA, functions as PPARγ agonist thus stimulate *ANGPTL4* expression and secretion in colon adenocarcinoma cells but not in SGBS cells.	(189)
Identification of Vitamin D Receptor-binding sites	At the sites of *DUSP10*, *NRIP1*, and *THBD* a treatment of the cells for 1 and 2 h with active vitamin D metabolite led to a statistically significant increase of VDR binding over 2-fold in SGBS adipocytes.	(190)
investigated the effect of resveratrol on mitochondrial mass increase and remodelling during n murine and human adipogenesis	Resveratrol induced lipid accumulation (Oil Red o) and mitochondrial mass increase during murine and human adipogenesis (mitochondrial protein expression and citrate synthase activity).	(191)
Metabolism of fructose by SGBS cells	Fructose induced anabolic processes in differentiated SGBS cells: de novo palmitate and oleate syntheses and extracellular secretion of palmitate.	(192)
The effect of caloric restriction (96h low-glucose ≤1 mM) on secretion profile of SGBS adipocytes	Lipolysis was induced; SIRT1 expression was significantly upregulated via caloric restriction. Secretome changes associated with caloric restriction indicated reduced inflammatory phenotype in adipocytes.	(186)
The effect of resveratrol exposure on lipid mobilization in adipocytes	After 24 h of treatment with 100 µM of resveratrol, *ATGL* expression was enhanced in both 3T3-L1 and SGBS adipocytes.	(193)
The effect of resveratrol exposure on expression of the plasminogen activator inhibitor-1 (PAI-1) in inflammation-induced human adipocytes	Used the inflammation model descripted in (165). Resveratrol exposure (in concentrations ≥30 µM) significantly reduced *PAI-1* expression in inflammation-induced human adipocytes. Additionally, when SGBS cells were exposed to macrophage secreted factors, they didn't induce ROS generation.	(194)
The effect of resveratrol exposure on the secretion profile of SGBS adipocytes	Resveratrol exposure induced lipid accumulation, induced upregulation of *ADIPOQ* and *APOE* with downregulation of *PAI-1* and *PEDF*.	(195)

Table 4: cont'd Key studies on SGBS cells that are published during the course of my PhD.

Objective	Summary	Reference
Hormone		
Effect of hydroxycholesterol on lipid and glucose metabolism in metabolically active cells	Glucose uptake, de novo lipogenesis and also the expression of lipid metabolism genes (*SREBF1, FASN, ACACA* and *DGAT1*) were increased with hydroxycholesterol exposure in SGBS cells.	(196)
Effect of testosterone in adipogenesis	miR-375 expression was downregulated after testosterone (100 mM) treatment	(197)
Identification of how testosterone exposure induces phosphorylation of AMPK	Testosterone exposure for a day (10-1000 nM) had significantly increased the expression of *LKB1* in SGBS adipocytes.	(198)
Potential effects of fibroblast growth factor-1 (*FGF1*) during adipogenesis	*FGF1* treatment induced lipid accumulation, and increased the expression of *PPARγ, G3PDH, GLUT4* and *ADIPOQ*.	(199)
Potential Drug		
Characterisation of the inflammatory regulatory roles of aleglitazar in SGBS adipocytes	Aleglitazar inhibited the expression of TNFα mediated cytokines (*MCP1, CXC-L10* and *IL6*) and induced the expression of adiponectin.	(200)
Effect of antidepressant substances on SGBS pre-adipocytes and adipocytes	The therapeutic concentrations of imipramine and lithium stimulated triglyceride accumulation but did not affect gene expression in SGBS cells.	(185)
Effect of quercetin exposure on the expression of adipokine and glycolysis genes in human SGBS adipocytes	Under normoxic cultivation quercetin inhibited the expression of *ANGPTL4, Adipsin, PAI-1* and *PFKP*.	(201)
Elvitegravir is a integrase inhibitor used for HIV infection treatment; the study looks into characterising the role of elvitegravir in lipid metabolism	Elvitegravir impaired adipogenesis and adipocyte metabolism in human SGBS adipocytes in a concentration-dependent manner (delaying acquisition of adipocyte morphology and reducing the expression of adipogenesis marker genes such as PPARγ, glucose transporter GLUT4, lipoprotein lipase, and the adipokines adiponectin and leptin).	(202)
Retinoic acid treatment effects	*UCP1* expression increased in response to low concentrations of retinoic acid in 3T3-L1 but not in SGBS adipocytes. Study suggested that SGBS is brown-like and expresses *UCP1* in response to culture media containing rosiglitazone.	(184)

Table 4: cont'd Key studies on SGBS cells that are published during the course of my PhD.

Objective	Summary	Reference
Environment		
2-day co-culture of hypoxia-induced SGBS cells with breast tumour cells (MCF7)	Hypoxia was induced with the transfection of *HIF1α* siRNA in SGBS cells. When *HIF1α* transfected SGBS cells co-cultured with breast tumour cells, the expression and secretion of ESR1 were repressed in breast tumour cells.	(203)
Effect of CoCl2, a hypoxia mimetic, on the secretome of human SGBS cells	In SGBS adipocytes, CoCl2 resulted in an up-regulation of collagens and a down-regulation of specific functional classes of ECM proteins.	(204)
Identification of hypoxia-induced gene expression changes	Hypoxia-regulated genes are involved in the regulation of obesity, T2D, and metabolic syndrome (*ADM, ANKRD37, DDIT4, KDM3A, PFKFB4, PPP1R3C, VEGFA, WDR73* and *ZNF395*).	(183)
The influence of reduced oxygen supply on gene expression of SGBS adipocytes	When SGBS adipocytes were cultured with 1% O2 for 3h or more, glycolytic genes (*ENO2, PFKP, PFKFB4* and *ALDOC*) were significantly upregulated.	(205)
Methodology		
Pro- and anti-adipogenetic drug testing	Treatment with 2 μmol/L rosiglitazone reduced lipid droplet size while 10 μM DHA, 17.5 mM glucose and 10 μg zinc oxide nanoparticles treatments made no change.	(168)
Validation of SBGS human cell line against primary adipocytes and 3T3-L1 murine cell line	Adipogenenic genes are more expressed in SGBS compared to 3T3-L1. Additionally, SGBS had higher expression of *LPL, GLUT4, PPARg* and *FABP4* and lower expression of *LEP* compared to primary adipocytes.	(206)
Adipose Biology		
A comprehensive analysis of transcript-level metabolism regulation in adipocytes	Gene expression, genome-wide ChIP-seq profiles for PPARγ, CEBPα, LXR and H3K4me3 and miRNA target identification for miR-27a, miR-29a and miR-222 on SGBS cells.	(180)
Characterisation of the roles of all lipin family members in human adipose tissue and SGBS adipogenesis	Depletion of every lipin family member in SGBS pre-adipocytes, alters expression levels of adipogenic transcription factors and lipid biosynthesis genes in early stages of differentiation.	(207)
Effect of myostatin on energy metabolism in muscle cells and SGBS adipocytes	Myostatin expression induced increase in basal glucose update in primary human muscle cells but not in SGBS cells.	(208)

Table 4: cont'd Key studies on SGBS cells that are published during the course of my PhD.

Objective	Summary	Reference
Adipose Biology		
Expression of stress proteins during weight loss and maintenance and comparison to the findings with SGBS adipocytes	When SGBS adipocytes were starved for 4 days, in low glucose (0.55 mM). Levels of stress proteins were altered: significant increased expression was observed in β-actin and HSP60.	(187)
Function of Atrial natriuretic peptide (ANP) and insulin on lipolysis	ANP stimulated lipolysis in human adipocytes (SGBS and primary visceral adipocytes) which is inhibited by insulin in a glucose-dependent manner.	(209)
Generating cell line with browning capacity	SGBS induce expression of *UCP1* in response to genetic manipulation.	(170)
Identification of expressional and secretional differences between omental and subcutaneous adipose tissue	*GREM1*, *PTN* and *SLP1* are all expressed significantly higer in omental adipose tissue compared to subcutaneous adipose tissue. All three were also expressed by cultured SGBS adipocytes.	(210)
Immunomodulatory effect of adipocyte-derived extracellular vesicles	Both SGBS and *in vivo* secreted extracellular vesicles induced inflammation signal production in macrophages and reciprocally inhibit insulin signalling in SGBS cells.	(182)
Impact of TRAIL (TNF-related apoptosis-inducing ligand) on human adipogenesis.	Incubation of SGBS cells in TRAIL during adipogenesis, inhibited adipogenic differentiation in a dose-dependent manner. It activated the cleavage of caspase-8 and -3, which in turn resulted in a downregulation of the key adipogenic transcription factors *C/EBPα, C/EBPδ,* and *PPARγ.*	(211)
Model to study the inflammatory process in adipose tissue	SGBS cells are exposed to macrophage-secreted factors or co-cultured with THP-1 monocytic cells: this induced insulin resistance, inhibited insulin-stimulated Akt phosphorylation and induced apoptosis in SGBS cells.	(165, 181)
Role of defective *VPS13B* gene expression in SGBS cells	Designed as a as a model for Cohen Syndrome associated truncal obesity, siRNA silencing of *VPS13B* accelerated differentiation rate, increased expression of specific adipogenic and lipid metabolism genes.	(212)
Role of EP3-mediated signalling during adipogenessis under the influence of hypertrophy and hypoxia	Demonstrated that EP3 was up-regulated in the obese human primary visceral adipocytes isolated and human SGBS adipocytes during the development of adipocyte hypertrophy and hypoxia.	(213)

Table 4: cont'd Key studies on SGBS cells that are published during the course of my PhD.

Objective	Summary	Reference
Adipose Biology		
Role of expression of certain miRNAs in obesity	Among the identified obesity related miRNAs: miR-148a-3p and DNMT1 are up- and down-regulated respectively in SGBS adipocytes.	(214)
Role of human adenine nucleotide translocase isoforms (hANT2) in cell proliferation status	SGBS pre-adipocytes (are growth-arrest-sensitive cells), hANT2 mRNA levels decreased. This was not observed in HeLa or HepG2 cells.	(215)
role of miR-192 expression in triglyceride storage of obese VAT and SGBS	Overexpression of miR-192 in SGBS cells, induced 25% reduction in the cellular triglyceride content and reduction in *GLUT4* expression.	(216)
The involvement of invadolysin, a metalloprotease, in lipid storage	The level of invadolysin increased during human SGBS adipogenesis.	(217)

6. Gene expression changes in visceral adipocytes associated with obesity

Chapter 4 examined transcriptional profiles of visceral adipose tissue (VAT) and purified adipocytes (VA) of 26 females across a range of BMI (19-50 kg/m^2) to understand at what point in the obesity spectrum molecular signatures change from healthy to unhealthy. We observed a clear trend in gene expression in accordance with the severity of obesity. There were a surprisingly high number of genes associated with BMI in VA. Additionally, this study showed that gene expression profiles of VAT are mainly contributed by the adipocytes and there was a clear transition in the gene expression profiles of VA in individuals with a BMI of 32 kg/m^2 or higher. Many of the differentially expressed genes associated with obesity measures were involved in signalling, metabolic, cellular stress and inflammatory response pathways. The findings were highly concordant with previously published studies and also provided resource for further investigation. For instance, it is worth examining the top genes that have not been previously considered in the context of obesity and metabolic diseases.

The immediate future work will be to integrate gene expression data with DNA methylation profiles of visceral adipocytes in the same cohort. Additionally, we will investigate whether methylome data support the model, we developed using WGBS of VA in the core set of three lean and these obese individuals (Bradford et al., unpublished). This model predicts that obesity is accompanied by recruitment or differential expansion of population of progenitors of different developmental origin. This model is based on methylation signature of development-associated genes such as: Paired like Homeodomain (*PITX)* and T-box (*TBX)* families. DNA methylation data will allow interpretation of gene expression changes in light of this model.

7. Final remarks and future directions

The future of understanding the molecular nature of obesity is bright. Costs of genomic analysis (sequencing and library preparation) have reduced dramatically. Population level epigenomics studies have become feasible with technical advances including the Illumina Infinium HumanMethylation450 (450K arrays) and MethylationEPIC (EPIC arrays) Beadchip arrays. This thesis provides a novel methylome method COBRAseq with unique advantages and perform well compared to the other methods (Chapter 3). As part of the EpiSCOPE study (Chapter 4), I have contributed to the establishments of a cohort to study epigenetic mechanisms in human fat in Australia.

In recent years, there have been a few epigenome wide association studies (EWAS) aimed at understanding epigenetic modifications associated obesity. Most of these studies examined associations between the DNA methylation profiles of biologically relevant samples (i.e. blood and subcutaneous adipose tissue) and obesity. They used measures such as BMI, waist circumference, HOMA-IR and blood levels of triglyceride, cholesterol and insulin of in cohorts containing over 600 individuals (137, 218-224).

Additionally, in 2015, there was a significant initiative to integrate the analysis of 111 reference epigenomes including the first reference adipocyte epigenome (225).

7.1. International perspective:

There are number of global and country based initiatives to assemble cohorts for population level epigenetics studies. Recently, Athlome Consortium has been established that links 10 cohorts across the world to understand complex interactions between genome, epigenome and proteome in elite athletic performance (226). There are also number of cohorts initiated by countries that provides unique advantages to study epigenetics. These include:

- United Kingdom:
    - *MuTHER Consortium* reports an analysis of the genetics of gene expression in three tissues from approximately 850 mono- and dizygotic twins (227).

- The Newcastle Thousand Families 1947 birth cohort which is a transgenerational cohort (228).
- ARIES, a population-based resource of DNA methylation data from 1000 mother-child pairs at three time points (229).

- Netherlands:
  - Famine transgenerational cohort (n= ~1000) (230).
- United States of America:
  - Normative Aging Study (n= 700) (231).
- Mexico:
  - Early Life Exposure in Mexico to ENvironmental Toxicants (ELEMENT) longitudinal cohort in kids (n=250) (232).

7.2. Australian perspective:

Major initiatives are being taken in Australia through investments in infrastructure and establishment of high quality biobanks and study cohorts. For instance, in New South Wales, as part of the 2012-2022 Health and Medical Research Strategy, the Genomics Collaboration has been established in Garvan Institute which has 10 HiSeq X Ten sequencing platforms. There are also funding schemes to allow researchers across the country to access these sequencing platforms. The NSW Biobank Registry has been established to provide high quality biospecimens for research. This initiative has allowed standardisation of biospecimen storage across pathology units in public hospitals.

Population level studies were initiated such as the 45 and Up Study, led by the Sax Institute. It is the largest ongoing study for healthy ageing in southern Hemisphere with 250,000 participants (233). One of the priority areas of this study is: Health effects of obesity, overweight and physical activity. Moreover, there are also incentives to tie these individual initiatives to generate larger impacts in the genomics field. For instance, in 2015, the Garvan Institute, in collaboration with the Sax Institute, has committed to sequence 2000 of the 45 and Up study participants' blood samples, which will form the Medical Genome Reference

Bank in Australia. A major advantage of this study is that medical records of the participants can be utilized for data linkage by The Centre for Health Record Linkage (CHeReL) (234). This study will significantly contribute to our understanding of molecular mechanisms in the development of obesity.

Two Australian cohorts, DOMInO and the Barwon Infant Study, have already been used for epigenome-wide analyses: for example, the DOMInO cohort for analysis of the impact of ω-3 fatty acid supplementation on the epigenome of children at birth and 5 years (109, 134). In addition to cohort sampling in adults, a few cohorts have been established to sample newborns such as the NSW based BABY 1000 Study, the Victoria based Generation Victoria, the Queensland based MUSP project and the Western Australia based the ORIGINS project. These cohorts provide unique advantages for studying the effect of prenatal stressors on epigenetic mechanisms and their association with later life obesity.

Perhaps future challenges in the epigenomics field lie in the integration and interrogation of large datasets for possible discovery of biomarkers. Change in methylation levels detected in the tissues of lean and obese individuals are significantly less than those detected in the tissues of cancer vs normal individuals. When changes in methylation levels are subtle, verification of potential markers might be prohibitively expensive, and instead of a single biomarker approach, a set of biomarkers might be used for accurate measure.

Lastly, as part of the EpiSCOPE study, a list of candidate genes as biomarkers for obesity has been developed (Appendix 1 and 2). The team aims to develop this knowledge further in other cohorts with the view of developing an obesity biomarker panel in the near future.

REFERENCES

1. Kershaw EE, Flier JS. Adipose tissue as an endocrine organ. J Clin Endocrinol Metabol. 2004;89(6):2548-56.
2. Rosen ED, Spiegelman BM. Adipocytes as regulators of energy balance and glucose homeostasis. Nature. 2006;444(7121):847-53.
3. Spalding KL, Arner E, Westermark PO, Bernard S, Buchholz BA, Bergmann O, et al. Dynamics of fat cell turnover in humans. Nature. 2008;453(7196):783-7.
4. Rother KI, Brown RJ. Novel forms of lipodystrophy: why should we care? Diabetes Care. 2013;36(8):2142-5.
5. WHO. Global health risks: mortality and burden of disease attributable to selected major risks. 2009.
6. Economics A. The growing costs of obesity in 2008 – three years on. 2008.
7. WHO. Obesity: preventing and managing the global epidemic: report of a WHO consultation.2000.
8. Sun Q, van Dam RM, Spiegelman D, Heymsfield SB, Willett WC, Hu FB. Comparison of dual-energy x-ray absorptiometric and anthropometric measures of adiposity in relation to adiposity-related biologic factors. Am J Epidemiol. 2010;172(12):1442-54.
9. Redinger RN. The pathophysiology of obesity and its clinical manifestations. Gastroenterol Hepatol. 2007;3(11):856-63.
10. WHO. Definition, diagnosis, and classification of diabetes mellitus and its complications: report of a WHO consultation.1999.
11. McMahon SK, Haynes A, Ratnam N, Grant MT, Carne CL, Jones TW, et al. Increase in type 2 diabetes in children and adolescents in Western Australia. Med J Aust. 2004;180(9):459-61.
12. Fonseca VA. Defining and characterizing the progression of type 2 diabetes. Diabetes Care. 2009;32(Suppl 2):S151-S6.
13. Reinehr T. Clinical presentation of type 2 diabetes mellitus in children and adolescents. Int J Obes Relat Metab Disord. 2005;29(S2):S105-S10.
14. Kopelman P. Health risks associated with overweight and obesity. Obes Rev. 2007;8 Suppl 1:13-7.
15. Image available online: http://www.nutritionforlife.nz/effects-of-obesity/ (accessed on 23 February 2016).
16. Ghigliotti G, Barisione C, Garibaldi S, Fabbi P, Brunelli C, Spallarossa P, et al. Adipose tissue immune response: novel triggers and consequences for chronic inflammatory conditions. Inflammation. 2014;37(4):1337-53.
17. Chang L, Chiang S-H, Saltiel AR. Insulin signaling and the regulation of glucose transport. Mol Med. 2004;10(7-12):65-71.
18. Rea S, James DE. Moving GLUT4: the biogenesis and trafficking of GLUT4 storage vesicles. Diabetes. 1997;46(11):1667-77.
19. Leney SE, Tavare JM. The molecular basis of insulin-stimulated glucose uptake: signalling, trafficking and potential drug targets. J Endocrinol. 2009;203(1):1-18.
20. Collins JM, Neville MJ, Pinnick KE, Hodson L, Ruyter B, van Dijk TH, et al. De novo lipogenesis in the differentiating human adipocyte can provide all fatty acids necessary for maturation. J Lipid Res. 2011;52(9):1683-92.
21. Frayn KN, Karpe F, Fielding BA, Macdonald IA, Coppack SW. Integrative physiology of human adipose tissue. Int J Obes Relat Metab Disord. 2003;27(8):875-88.

22. Perea A, Clemente F, Martinell J, Villanueva-Penacarrillo ML, Valverde I. Physiological effect of glucagon in human isolated adipocytes. Horm Metab Res. 1995;27(8):372-5.
23. Duncan RE, Ahmadian M, Jaworski K, Sarkadi-Nagy E, Sul HS. Regulation of lipolysis in adipocytes. An Rev Nutr. 2007;27:79-101.
24. Karpe F, Dickmann JR, Frayn KN. Fatty acids, obesity, and insulin resistance: time for a reevaluation. Diabetes. 2011;60(10):2441-9.
25. Wellen KE, Hotamisligil GS. Inflammation, stress, and diabetes. J Clin Invest. 2005;115(5):1111-9.
26. Ronti T, Lupattelli G, Mannarino E. The endocrine function of adipose tissue: an update. Clin Endocrinol. 2006;64(4):355-65.
27. Xie L, O'Reilly CP, Chapes SK, Mora S. Adiponectin and leptin are secreted through distinct trafficking pathways in adipocytes. BBA. 2008;1782(2):99-108.
28. Ye F, Than A, Zhao Y, Goh KH, Chen P. Vesicular storage, vesicle trafficking, and secretion of leptin and resistin: the similarities, differences, and interplays. J Endocrinol. 2010;206(1):27-36.
29. Hotamisligil GS, Arner P, Caro JF, Atkinson RL, Spiegelman BM. Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance. J Clin Invest. 1995;95(5):2409-15.
30. Koutnikova H, Auwerx J. Regulation of adipocyte differentiation. Ann Med. 2001;33(8):556-61.
31. Bourin P, Bunnell BA, Casteilla L, Dominici M, Katz AJ, March KL, et al. Stromal cells from the adipose tissue-derived stromal vascular fraction and culture expanded adipose tissue-derived stromal/stem cells. Cytotherapy. 2013;15(6):641-8.
32. Boquest AC, Collas P. Obtaining freshly isolated and cultured mesenchymal stem cells from human adipose tissue. Methods Mol Biol. 2012;879:269-78.
33. Gentile P, Orlandi A, Scioli MG, Di Pasquali C, Bocchini I, Cervelli V. Concise review: Adipose-derived stromal vascular fraction cells and platelet-rich plasma: Basic and clinical implications for tissue engineering therapies in regenerative surgery. Stem Cells Transl Med. 2012;1(3):230-6.
34. Gaustad KG, Boquest AC, Anderson BE, Gerdes AM, Collas P. Differentiation of human adipose tissue stem cells using extracts of rat cardiomyocytes. Biochem Biophys Res Commun. 2004;314(2):420-7.
35. Boulland JL, Mastrangelopoulou M, Boquest AC, Jakobsen R, Noer A, Glover JC, et al. Epigenetic regulation of nestin expression during neurogenic differentiation of adipose tissue stem cells. Stem Cells Dev. 2013;22(7):1042-52.
36. Macotela Y, Emanuelli B, Mori MA, Gesta S, Schulz TJ, Tseng YH, et al. Intrinsic differences in adipocyte precursor cells from different white fat depots. Diabetes. 2012;61(7):1691-9.
37. Hocking SL, Wu LE, Guilhaus M, Chisholm DJ, James DE. Intrinsic depot-specific differences in the secretome of adipose tissue, preadipocytes, and adipose tissue-derived microvascular endothelial cells. Diabetes. 2010;59(12):3008-16.
38. Cousin B, Munoz O, Andre M, Fontanilles AM, Dani C, Cousin JL, et al. A role for preadipocytes as macrophage-like cells. FASEB. 1999;13(2):305-12.
39. Yu L, Yan K, Liu P, Li N, Liu Z, Zhu W, et al. Pattern recognition receptor-initiated innate antiviral response in mouse adipose cells. Immunol Cell Biol. 2014;92(2):105-15.
40. Rosen Evan D, Spiegelman Bruce M. What we talk about when we talk about fat. Cell. 2014;156(1–2):20-44.

41. Lowe CE, O'Rahilly S, Rochford JJ. Adipogenesis at a glance. J Cell Sci. 2011;124(16):2681-6.
42. Tontonoz P, Spiegelman BM. Fat and beyond: the diverse biology of PPARgamma. Annu Rev Biochem. 2008;77:289-312.
43. Huang S, Czech MP. The GLUT4 glucose transporter. Cell Metabolism. 2007;5(4):237-52.
44. Hotamisligil GS, Bernlohr DA. Metabolic functions of FABPs-mechanisms and therapeutic implications. Nat Rev Endocrinol. 2015;11(10):592-605.
45. Kobayashi J, Mabuchi H. Lipoprotein lipase and atherosclerosis. Ann Clin Biochem. 2015;52(6):632-7.
46. Agarwal AK, Sukumaran S, Cortes VA, Tunison K, Mizrachi D, Sankella S, et al. Human 1-acylglycerol-3-phosphate O-acyltransferase isoforms 1 and 2: biochemical characterization and inability to rescue hepatic steatosis in Agpat2(-/-) gene lipodystrophic mice. J Biol Chem. 2011;286(43):37676-91.
47. Brydon L. Adiposity, leptin and stress reactivity in humans. Biol Psychol. 2011;86(2):114-20.
48. Fernandez-Real JM, Vendrell J, Ricart W. Circulating adiponectin and plasma fatty acid profile. Clin Chem. 2005;51(3):603-9.
49. Beller M, Thiel K, Thul PJ, Jäckle H. Lipid droplets: A dynamic organelle moves into focus. FEBS Lett. 2010;584(11):2176-82.
50. Giralt M, Villarroya F. White, brown, beige/brite: different adipose cells for different functions? Endocrinology. 2013;154(9):2992-3000.
51. Gesta S, Tseng YH, Kahn CR. Developmental origin of fat: tracking obesity to its source. Cell. 2007;131(2):242-56.
52. Lidell ME, Betz MJ, Leinhard OD, Heglind M, Elander L, Slawik M, et al. Evidence for two types of brown adipose tissue in humans. Nat Med. 2013;19(5):631-4.
53. Cannon B, Nedergaard J. Brown adipose tissue: function and physiological significance. Physiol Rev. 2004;84(1):277-359.
54. Power G. Biology of temperature: the mammalian fetus. J Dev Physiol. 1989;12(6):295-304.
55. Tews D, Wabitsch M. Renaissance of brown adipose tissue. Horm Res Paediatr. 2011;75(4):231-9.
56. Seale P, Bjork B, Yang W, Kajimura S, Chin S, Kuang S, et al. PRDM16 controls a brown fat/skeletal muscle switch. Nature. 2008;454(7207):961-7.
57. Timmons JA, Wennmalm K, Larsson O, Walden TB, Lassmann T, Petrovic N, et al. Myogenic gene expression signature establishes that brown and white adipocytes originate from distinct cell lineages. PNAS. 2007;104(11):4401-6.
58. Xu S, Chen P, Sun L. Regulatory networks of non-coding RNAs in brown/beige adipogenesis. Biosci Rep. 2015;35(5).
59. Harms M, Seale P. Brown and beige fat: development, function and therapeutic potential. Nat Med. 2013;19(10):1252-63.
60. Hilton C, Karpe F, Pinnick KE. Role of developmental transcription factors in white, brown and beige adipose tissues. BBA. 2015;1851(5):686-96.
61. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu CY, et al. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. Circulation. 2007;116(1):39-48.
62. White UA, Tchoukalova YD. Sex dimorphism and depot differences in adipose tissue function. BBA. 2014;1842(3):377-92.

63. Camhi SM, Bray GA, Bouchard C, Greenway FL, Johnson WD, Newton RL, et al. The relationship of waist circumference and BMI to visceral, subcutaneous, and total body fat: sex and race differences. Obesity (Silver Spring). 2011;19(2):402-8.
64. Tchoukalova YD, Votruba SB, Tchkonia T, Giorgadze N, Kirkland JL, Jensen MD. Regional differences in cellular mechanisms of adipose tissue gain with overfeeding. PNAS. 2010;107(42):18226-31.
65. Gealekman O, Guseva N, Hartigan C, Apotheker S, Gorgoglione M, Gurav K, et al. Depot-specific differences and insufficient subcutaneous adipose tissue angiogenesis in human obesity. Circulation. 2011;123(2):186-94.
66. Curat CA, Wegner V, Sengenes C, Miranville A, Tonus C, Busse R, et al. Macrophages in human visceral adipose tissue: increased accumulation in obesity and a source of resistin and visfatin. Diabetologia. 2006;49(4):744-7.
67. Barakat H, Hickner RC, Privette J, Bower J, Hao E, Udupi V, et al. Differences in the lipolytic function of adipose tissue preparations from Black American and Caucasian women. Metabolism. 2002;51(11):1514-8.
68. Montague CT, O'Rahilly S. The perils of portliness: causes and consequences of visceral adiposity. Diabetes. 2000;49(6):883-8.
69. Berry DC, Stenesen D, Zeve D, Graff JM. The developmental origins of adipose tissue. Development. 2013;140(19):3939-49.
70. Baglioni S, Cantini G, Poli G, Francalanci M, Squecco R, Di Franco A, et al. Functional differences in visceral and subcutaneous fat pads originate from differences in the adipose stem cell. PloS One. 2012;7(5):e36569.
71. Chau Y-Y, Bandiera R, Serrels A, Martínez-Estrada OM, Qing W, Lee M, et al. Visceral and subcutaneous fat have different origins and evidence supports a mesothelial source. Nat Cell Biol. 2014;16(4):367-75.
72. Gesta S, Bluher M, Yamamoto Y, Norris AW, Berndt J, Kralisch S, et al. Evidence for a role of developmental genes in the origin of obesity and body fat distribution. PNAS. 2006;103(17):6676-81.
73. Petrovic N, Walden TB, Shabalina IG, Timmons JA, Cannon B, Nedergaard J. Chronic peroxisome proliferator-activated receptor gamma (PPARgamma) activation of epididymally derived white adipocyte cultures reveals a population of thermogenically competent, UCP1-containing adipocytes molecularly distinct from classic brown adipocytes. J Biol Chem. 2010;285(10):7153-64.
74. Park A, Kim WK, Bae K-H. Distinction of white, beige and brown adipocytes derived from mesenchymal stem cells. J Stem Cells. 2014;6(1):33-42.
75. Lee M-W, Odegaard Justin I, Mukundan L, Qiu Y, Molofsky Ari B, Nussbaum Jesse C, et al. Activated type 2 innate lymphoid cells regulate beige fat biogenesis. Cell. 2015;160(1–2):74-87.
76. Brestoff JR, Kim BS, Saenz SA, Stine RR, Monticelli LA, Sonnenberg GF, et al. Group 2 innate lymphoid cells promote beiging of white adipose tissue and limit obesity. Nature. 2015;519(7542):242-6.
77. Kiefer FW, Vernochet C, O'Brien P, Spoerl S, Brown JD, Nallamshetty S, et al. Retinaldehyde dehydrogenase 1 regulates a thermogenic program in white adipose tissue. Nat Med. 2012;18(6):918-25.
78. Green H, Kehinde O. An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion. Cell. 1975;5(1):19-27.
79. Wabitsch M, Brenner RE, Melzner I, Braun M, Moller P, Heinze E, et al. Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. Int J Obesity. 2001;25(1):8-15.

80. Fischer-Posovszky P, Newell FS, Wabitsch M, Tornqvist HE. Human SGBS cells - a unique tool for studies of human fat cell biology. Obes Facts. 2008;1(4):184-9.
81. Wang QA, Tao C, Gupta RK, Scherer PE. Tracking adipogenesis during white adipose tissue development, expansion and regeneration. Nat Med. 2013;19(10):1338-44.
82. Arner P. Differences in lipolysis between human subcutaneous and omental adipose tissues. Ann Med. 1995;27(4):435-8.
83. DiGirolamo M, Fine JB, Tagra K, Rossmanith R. Qualitative regional differences in adipose tissue growth and cellularity in male Wistar rats fed ad libitum. Am J Physiol. 1998;274(5 Pt 2):R1460-7.
84. Jo J, Gavrilova O, Pack S, Jou W, Mullen S, Sumner AE, et al. Hypertrophy and/or hyperplasia: Dynamics of adipose tissue growth. PLoS Comput Biol. 2009;5(3):e1000324.
85. Tang W, Zeve D, Suh JM, Bosnakovski D, Kyba M, Hammer RE, et al. White fat progenitor cells reside in the adipose vasculature. Science. 2008;322(5901):583-6.
86. Lin G, Garcia M, Ning H, Banie L, Guo Y-L, Lue TF, et al. Defining stem and progenitor cells within adipose tissue. Stem Cells and Development. 2008;17(6):1053-63.
87. Majka SM, Miller HL, Sullivan T, Erickson PF, Kong R, Weiser-Evans M, et al. Adipose lineage specification of bone marrow-derived myeloid cells. Adipocyte. 2012;1(4):215-29.
88. Ryden M, Uzunel M, Hard JL, Borgstrom E, Mold JE, Arner E, et al. Transplanted bone marrow-derived cells contribute to human adipogenesis. Cell Metab. 2015;22(3):408-17.
89. Gavin KM, Gutman JA, Kohrt WM, Wei Q, Shea KL, Miller HL, et al. De novo generation of adipocytes from circulating progenitor cells in mouse and human adipose tissue. FASEB. 2016;30(3):1096-108.
90. Stephens JM. The fat controller: Adipocyte development. PLoS Biology. 2012;10(11):e1001436.
91. Acosta JR, Douagi I, Andersson DP, Bäckdahl J, Rydén M, Arner P, et al. Increased fat cell size: a major phenotype of subcutaneous white adipose tissue in non-obese individuals with type 2 diabetes. Diabetologia. 2015:1-11.
92. Arner E, Westermark PO, Spalding KL, Britton T, Rydén M, Frisén J, et al. Adipocyte turnover: Relevance to human adipose tissue morphology. Diabetes. 2010;59(1):105-9.
93. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206.
94. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. 2015;518(7538):187-96.
95. Dahlman I, Ryden M, Brodin D, Grallert H, Strawbridge RJ, Arner P. Numerous genes in loci associated with body fat distribution are linked to adipose function. Diabetes. 2016;65(2):433-7.
96. Gao H, Mejhert N, Fretz Jackie A, Arner E, Lorente-Cebrián S, Ehrlund A, et al. Early B cell factor 1 regulates adipocyte morphology and lipolysis in white adipose tissue. Cell Metab.19(6):981-92.
97. Van Pelt RE, Jankowski CM, Gozansky WS, Schwartz RS, Kohrt WM. Lower-body adiposity and metabolic protection in postmenopausal women. J Clin Endocrinol Metab. 2005;90(8):4573.

98. Snijder MB, Dekker JM, Visser M, Yudkin JS, Stehouwer CD, Bouter LM, et al. Larger thigh and hip circumferences are associated with better glucose tolerance: the Hoorn study. Obes Res. 2003;11(1):104-11.
99. Jensen MD. Role of body fat distribution and the metabolic complications of obesity. J Clin Endocrinol Metab. 2008;93(11 Suppl 1):S57-63.
100. WHO. Obesity and overweight. Fact sheet No: 311 Available online: http://www.who.int/mediacentre/factsheets/fs311/en/ (accessed on 14 January 2016). 2013.
101. ABS. Australian Bureau of Statistics Profiles of Health, Australia, 2011-13 Cat. no: 4364.0.55.001 Available online:http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4338.0~2011-13~Main%20Features~Diabetes~10004 (Accessed on 14 January 2016). 2013.
102. OECD. Organisation for Economic Cooperation and Development: Health at a glance 2013: OECD indicators. Paris: OECD Publishing. 2013.
103. WHO. Diabetes Fact sheet: No: 312 Available online: http://www.who.int/mediacentre/factsheets/fs312/en/ (accessed on 14 January 2016). 2013.
104. Whiting DR, Guariguata L, Weil C, Shaw J. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Res Clin Pract. 2011;94(3):311-21.
105. Tanamas SKM, D. J.; Lynch, B.; Seth,i P.; Willenberg, L.; Polkinghorne, K. R.; Chadban, S.; Dunstan, D.; Shaw, J. E. . AusDiab report 2012: The Australian diabetes, obesity and lifestyle study. 2012.
106. AIHW. Diabetes as a cause of death, Australia, 1997 and 1998. Canberra: AIHW. 1997.
107. AIHW. Diabetes: Australian facts 2008. Canberra: AIHW. 2008.
108. Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. Circulation. 2012;126(1):126-32.
109. van Dijk SJ, Tellam RL, Morrison JL, Muhlhausler BS, L. MP. Recent developments on the role of epigenetics in obesity and metabolic disease. Clin Epigenet (2015). 2015;7:66.
110. Grarup N, Sandholt CH, Hansen T, Pedersen O. Genetic susceptibility to type 2 diabetes and obesity: from genome-wide association studies to rare variants and beyond. Diabetologia. 2014;57(8):1528-41.
111. Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, et al. Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. Diabetologia. 2011;54(11):2811-9.
112. McCarthy MI, Zeggini E. Genome-wide association studies in type 2 diabetes. Current diabetes reports. 2009;9(2):164-71.
113. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012;44(9):981-90.
114. Carlsson S, Ahlbom A, Lichtenstein P, Andersson T. Shared genetic influence of BMI, physical activity and type 2 diabetes: a twin study. Diabetologia. 2013;56(5):1031-5.
115. Bouchard L, Rabasa-Lhoret R, Faraj M, Lavoie ME, Mill J, Perusse L, et al. Differential epigenomic and transcriptomic responses in subcutaneous adipose tissue between low and high responders to caloric restriction. Am J Clin Nutr. 2010;91.
116. Ronn T, Volkov P, Davegardh C, Dayeh T, Hall E, Olsson AH, et al. A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. PLoS Genetics. 2013;9(6):e1003572.

117. Arner P, Sinha I, Thorell A, Rydén M, Dahlman-Wright K, Dahlman I. The epigenetic signature of subcutaneous fat cells is linked to altered expression of genes implicated in lipid metabolism in obese women. Clinical Epigenetics. 2015;7(1):93.
118. Dahlman I, Sinha I, Gao H, Brodin D, Thorell A, Ryden M, et al. The fat cell epigenetic signature in post-obese women is characterized by global hypomethylation and differential DNA methylation of adipogenesis genes. Int J Obesity. 2015;39(6):910-9.
119. Benton MC, Johnstone A, Eccles D, Harmon B, Hayes MT, Lea RA, et al. An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. Genome Biol. 2015;16(1):1-21.
120. Wolffe AP, Matzke MA. Epigenetics: regulation through repression. Science. 1999;286(5439):481-6.
121. Moving AHEAD with an international human epigenome project. Nature. 2008;454(7205):711-5.
122. Bai L, Morozov AV. Gene regulation by nucleosome positioning. Trends Genet. 2010;26(11):476-83.
123. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet. 2011;12(1):7-18.
124. Carén H, Pollard SM, Beck S. The good, the bad and the ugly: Epigenetic mechanisms in glioblastoma. Mol Aspects Med. 2013;34(4):849-62.
125. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11(10):726-34.
126. Hill PWS, Amouroux R, Hajkova P. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: An emerging complex story. Genomics. 2014;104(5):324-33.
127. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315-22.
128. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011;473(7347):398-402.
129. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: prospects and challenges. Trends Genet. 2014;30(2):75-84.
130. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol. 1987;196(2):261-82.
131. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet. 2009;41(2):178-86.
132. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484-92.
133. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. Nat Rev Genet. 2006;7(1):21-33.
134. van Dijk SJ, Molloy PL, Varinli H, Morrison JL, Muhlhausler BS, EpiSCOPE. Epigenetics and human obesity. Int J Obesity. 2015;39(1):85-97.
135. Milagro FI, Campion J, Garcia-Diaz DF, Goyenechea E, Paternain L, Martinez JA. High fat diet-induced obesity modifies the methylation pattern of leptin promoter in rats. J Physiol Biochem. 2009;65(1):1-9.

136. Widiker S, Karst S, Wagener A, Brockmann GA. High-fat diet leads to a decreased methylation of the Mc4r gene in the obese BFMI and the lean B6 mouse lines. J Appl Genet. 2010;51(2):193-7.
137. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aissi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet. 2014;383(9933):1990-8.
138. Agha G, Houseman EA, Kelsey KT, Eaton CB, Buka SL, Loucks EB. Adiposity is associated with DNA methylation profile in adipose tissue. Int J Epidemiol. 2015;44(4):1277-87.
139. Herrera BM, Keildson S, Lindgren CM. Genetics and epigenetics of obesity. Maturitas. 2011;69(1):41-9.
140. Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. Nucleic Acids Res. 1994;22(15):2990-7.
141. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. PNAS. 1992;89(5):1827-31.
142. Hirst M, Marra MA. Next generation sequencing based approaches to epigenomics. Brief Funct Genomics. 2010;9(5-6):455-65.
143. Weber M, Schubeler D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Curr Opin Cell Biol. 2007;19(3):273-80.
144. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, et al. Methylome analysis using MeDIP-seq with low DNA concentrations. Nature Protocols. 2012;7(4):617-36.
145. Brunner AL, Johnson DS, Kim SW, Valouev A, Reddy TE, Neff NF, et al. Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. Genome Res. 2009;19(6):1044-56.
146. Ordway JM, Bedell JA, Citek RW, Nunberg A, Garrido A, Kendall R, et al. Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. Carcinogenesis. 2006;27(12):2409-23.
147. Jelinek J, Liang SD, Lu Y, He R, Ramagli LS, Shpall EJ, et al. Conserved DNA methylation patterns in healthy blood cells and extensive changes in leukemia measured by a new quantitative technique. Epigenetics. 2012;7(12):1368-78.
148. Ross JP, Shaw JM, Molloy PL. Identification of differentially methylated regions using streptavidin bisulfite ligand methylation enrichment (SuBLiME), a new method to enrich for methylated DNA prior to deep bisulfite genomic sequencing. Epigenetics. 2013;8(1):113-27.
149. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nature Biotechnol. 2009;27(2):182-9.
150. Li Q, Suzuki M, Wendt J, Patterson N, Eichten SR, Hermanson PJ, et al. Post-conversion targeted capture of modified cytosines in mammalian and plant genomes. Nucleic Acids Res. 2015;43(12).
151. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011;3(6):771-84.
152. Sandoval J, Heyn HA, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics. 2011;6(6):692-702.
153. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res. 1997;25(12):2532-4.

154. Ulahannan N, Greally JM. Genome-wide assays that identify and quantify modified cytosines in human disease studies. Epigenet Chromatin. 2015;8:5.
155. Robinson MD, Stirzaker C, Statham AL, Coolen MW, Song JZ, Nair SS, et al. Evaluation of affinity-based genome-wide DNA methylation data: Effects of CpG density, amplification bias, and copy number variation. Genome Res. 2010;20(12):1719-29.
156. Serre D, Lee BH, Ting AH. MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. Nucleic Acids Res. 2010;38(2):391-9.
157. Jung M, Kadam S, Xiong W, Rauch TA, Jin SG, Pfeifer GP. MIRA-seq for DNA methylation analysis of CpG islands. Epigenomics. 2015:1-12.
158. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: prospects and challenges. Trends in Genet. 2014;30(2):75-84.
159. Oda M, Glass JL, Thompson RF, Mo Y, Olivier EN, Figueroa ME, et al. High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. Nucleic Acids Res. 2009;37(12):3829-39.
160. Smith ZD, Gu HC, Bock C, Gnirke A, Meissner A. High-throughput bisulfite sequencing in mammalian genomes. Methods. 2009;48(3):226-32.
161. Ziller MJ, Muller F, Liao J, Zhang YY, Gu HC, Bock C, et al. Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genetics. 2011;7(12).
162. Francis MP, Sachs PC, Elmore LW, E. Holt S. Isolating adipose-derived mesenchymal stem cells from lipoaspirate blood and saline fraction. Organogenesis. 2010;6(1):11-4.
163. Scott MA, Nguyen VT, Levi B, James AW. Current methods of adipogenic differentiation of mesenchymal stem cells. Stem Cells Dev. 2011;20(10):1793-804.
164. Legzdina D, Romanauska A, Nikulshin S, Kozlovska T, Berzins U. Characterization of senescence of culture-expanded human adipose-derived mesenchymal stem cells. Int J Stem Cells. 2016;9(1):124-36.
165. Keuper M, Dzyakanchuk A, Amrein KE, Wabitsch M, Fischer-Posovszky P. THP-1 Macrophages and SGBS Adipocytes – A New Human in vitro Model System of Inflamed Adipose Tissue. Front Endocrinol. 2011;2:89.
166. Tchkonia T, Giorgadze N, Pirtskhalava T, Thomou T, DePonte M, Koo A, et al. Fat depot-specific characteristics are retained in strains derived from single human preadipocytes. Diabetes. 2006;55(9):2571-8.
167. Maumus M, Sengenes C, Decaunes P, Zakaroff-Girard A, Bourlier V, Lafontan M, et al. Evidence of in situ proliferation of adult adipose tissue-derived progenitor cells: influence of fat mass microenvironment and growth. J Clin Endocrinol Metab. 2008;93(10):4098-106.
168. Varinli H, Osmond-McLeod MJ, Molloy PL, Vallotton P. LipiD-QuanT: a novel method to quantify lipid accumulation in live cells. J Lipid Res. 2015;56(11):2206-16.
169. Wabitsch M, Brenner RE, Melzner I, Braun M, Moller P, Heinze E, et al. Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. Int J Obes Relat Metab Disord. 2001;25:8-15.
170. Tews D, Fischer-Posovszky P, Fromme T, Klingenspor M, Fischer J, Ruther U, et al. FTO deficiency induces UCP-1 expression and mitochondrial uncoupling in adipocytes. Endocrinology. 2013;154(9):3141-51.
171. Hugo ER, Brandebourg TD, Comstock CE, Gersin KS, Sussman JJ, Ben-Jonathan N. LS14: a novel human adipocyte cell line that produces prolactin. Endocrinol. 2006;147(1):306-13.

172. Wabitsch M, Brüderlein S, Melzner I, Braun M, Mechtersheimer G, Möller P. LiSa-2, a novel human liposarcoma cell line with a high capacity for terminal adipose differentiation. Intl J Cancer. 2000;88(6):889-94.
173. Torii I, Morikawa S, Nakano A, Morikawa K. Establishment of a human preadipose cell line, HPB-AML-I: Refractory to PPARγ-mediated adipogenic stimulation. J Cell Physiol. 2003;197(1):42-52.
174. Zilberfarb V, Pietri-Rouxel F, Jockers R, Krief S, Delouis C, Issad T, et al. Human immortalized brown adipocytes express functional beta3-adrenoceptor coupled to lipolysis. J Cell Sci. 1997;110 (Pt 7):801-7.
175. Shinoda K, Luijten IH, Hasegawa Y, Hong H, Sonne SB, Kim M, et al. Genetic and functional characterization of clonally derived adult human brown adipocytes. Nat Med. 2015;21(4):389-94.
176. Jeffery E, Wing A, Holtrup B, Sebo Z, Kaplan Jennifer L, Saavedra-Peña R, et al. The adipose tissue microenvironment regulates depot-specific adipogenesis in obesity. Cell Metab. 2016;24(1):142-50.
177. Tirosh A, Shai I, Afek A, Dubnov-Raz G, Ayalon N, Gordon B, et al. Adolescent BMI trajectory and risk of diabetes versus coronary disease. N Engl J Med. 2011;364(14):1315-25.
178. Tanamas SK, Magliano DJ, Lynch B, Sethi P, Willenberg L, Polkinghorne KR, et al. AusDiab 2012: The Australian Diabetes, Obesity and Lifestyle Study Melbourne: Baker IDI Heart and Diabetes Institute 2013. 2012.
179. Sabater D, Arriaran S, Romero Mdel M, Agnelli S, Remesar X, Fernandez-Lopez JA, et al. Cultured 3T3L1 adipocytes dispose of excess medium glucose as lactate under abundant oxygen availability. Sci Rep. 2014;4:3663.
180. Galhardo M, Sinkkonen L, Berninger P, Lin J, Sauter T, Heinaniemi M. Integrated analysis of transcript-level regulation of metabolism reveals disease-relevant nodes of the human metabolic network. Nucleic Acids Res. 2014;42(3):1474-96.
181. Keuper M, Bluher M, Schon MR, Moller P, Dzyakanchuk A, Amrein K, et al. An inflammatory micro-environment promotes human adipocyte apoptosis. Mol Cell Endocrinol. 2011;339(1-2):105-13.
182. Kranendonk ME, Visseren FL, van Balkom BW, Nolte-'t Hoen EN, van Herwaarden JA, de Jager W, et al. Human adipocyte extracellular vesicles in reciprocal signaling between adipocytes and macrophages. Obesity (Silver Spring). 2014;22(5):1296-308.
183. Geiger K, Leiherer A, Muendlein A, Stark N, Geller-Rhomberg S, Saely CH, et al. Identification of hypoxia-induced genes in human SGBS adipocytes by microarray analysis. PloS One. 2011;6(10).
184. Murholm M, Isidor MS, Basse AL, Winther S, Sørensen C, Skovgaard-Petersen J, et al. Retinoic acid has different effects on UCP1 expression in mouse and human adipocytes. BMC Cell Biol. 2013;14:41-.
185. Loffler D, Landgraf K, Korner A, Kratzsch J, Kirkby KC, Himmerich H. Modulation of triglyceride accumulation in adipocytes by psychopharmacological agents in vitro. J Psychiatr Res. 2016;72:37-42.
186. Renes J, Rosenow A, Roumans N, Noben JP, Mariman EC. Calorie restriction-induced changes in the secretome of human adipocytes, comparison with resveratrol-induced secretome effects. BBA. 2014;1844(9):1511-22.
187. Roumans NJ, Camps SG, Renes J, Bouwman FG, Westerterp KR, Mariman EC. Weight loss-induced stress in subcutaneous adipose tissue is related to weight regain. Br J Nutr. 2016;115(5):913-20.
188. Scoditti E, Massaro M, Carluccio MA, Pellegrino M, Wabitsch M, Calabriso N, et al. Additive regulation of adiponectin expression by the mediterranean diet olive oil

components oleic Acid and hydroxytyrosol in human adipocytes. PloS One. 2015;10(6):e0128218.
189. Alex S, Lange K, Amolo T, Grinstead JS, Haakonsson AK, Szalowska E, et al. Short-chain fatty acids stimulate angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating peroxisome proliferator-activated receptor gamma. Mol Cell Biol. 2013;33(7):1303-16.
190. Ryynanen J, Neme A, Tuomainen TP, Virtanen JK, Voutilainen S, Nurmi T, et al. Changes in vitamin D target gene expression in adipose tissue monitor the vitamin D response of human individuals. Mol Nutr Food Res. 2014;58(10):2036-45.
191. Li S, Bouzar C, Cottet-Rousselle C, Zagotta I, Lamarche F, Wabitsch M, et al. Resveratrol inhibits lipogenesis of 3T3-L1 and SGBS cells by inhibition of insulin signaling and mitochondrial mass increase. BBA. 2016;1857(6):643-52.
192. Varma V, Boros LG, Nolen GT, Chang C-W, Wabitsch M, Beger RD, et al. Metabolic fate of fructose in human adipocytes: a targeted (13)C tracer fate association study. Metabolomics. 2015;11(3):529-44.
193. Lasa A, Schweiger M, Kotzbeck P, Churruca I, Simon E, Zechner R, et al. Resveratrol regulates lipolysis via adipose triglyceride lipase. J Nutr Biochem. 2012;23(4):379-84.
194. Zagotta I, Dimova EY, Funcke JB, Wabitsch M, Kietzmann T, Fischer-Posovszky P. Resveratrol suppresses PAI-1 gene expression in a human in vitro model of inflamed adipose tissue. Oxid Med Cell Longev. 2013;2013:793525.
195. Rosenow A, Noben JP, Jocken J, Kallendrusch S, Fischer-Posovszky P, Mariman EC, et al. Resveratrol-induced changes of the human adipocyte secretion profile. J Proteome Res. 2012;11(9):4733-43.
196. Hessvik NP, Bakke SS, Smith R, Ravna AW, Sylte I, Rustan AC, et al. The liver X receptor modulator 22(S)-hydroxycholesterol exerts cell-type specific effects on lipid and glucose metabolism. J Steroid Biochem Mol Biol. 2012;128(3-5):154-64.
197. Kraus M, Greither T, Wenzel C, Brauer-Hartmann D, Wabitsch M, Behre HM. Inhibition of adipogenic differentiation of human SGBS preadipocytes by androgen-regulated microRNA miR-375. Mol Cell Endocrinol. 2015;414:177-85.
198. McInnes KJ, Brown KA, Hunger NI, Simpson ER. Regulation of LKB1 expression by sex hormones in adipocytes. Int J Obesity. 2012;36(7):982-5.
199. Luo X, Jia R, Li K, Zhu X, Zhao D, J PW, et al. Fibroblast growth factor-1 inhibits Wnt/beta-catenin pathway during adipogenesis. J Centr South Uni Med Sci. 2015;40(8):843-50.
200. Massaro M, Scoditti E, Pellegrino M, Carluccio MA, Calabriso N, Wabitsch M, et al. Therapeutic potential of the dual peroxisome proliferator activated receptor (PPAR)alpha/gamma agonist aleglitazar in attenuating TNF-alpha-mediated inflammation and insulin resistance in human adipocytes. Pharmacol Res. 2016;107:125-36.
201. Leiherer A, Stoemmer K, Muendlein A, Saely CH, Kinz E, Brandtner EM, et al. Quercetin impacts expression of metabolism and obesity associated genes in SGBS adipocytes. Nutrients. 2016;8(5).
202. Moure R, Domingo P, Gallego-Escuredo JM, Villarroya J, Gutierrez Mdel M, Mateo MG, et al. Impact of elvitegravir on human adipocytes: Alterations in differentiation, gene expression and release of adipokines and cytokines. Antiviral Res. 2016;132:59-65.
203. Yao-Borengasser A, Monzavi-Karbassi B, Hedges RA, Rogers LJ, Kadlubar SA, Kieber-Emmons T. Adipocyte hypoxia promotes epithelial-mesenchymal transition-related gene expression and estrogen receptor-negative phenotype in breast cancer cells. Oncol Rep. 2015;33(6):2689-94.

204. Rosenow A, Noben JP, Bouwman FG, Mariman EC, Renes J. Hypoxia-mimetic effects in the secretome of human preadipocytes and adipocytes. BBA. 2013;1834(12):2761-71.
205. Leiherer A, Geiger K, Muendlein A, Drexel H. Hypoxia induces a HIF-1alpha dependent signaling cascade to make a complex metabolic switch in SGBS-adipocytes. Mol Cell Endocrinol. 2014;383(1-2):21-31.
206. Allott EH, Oliver E, Lysaght J, Gray SG, Reynolds JV, Roche HM, et al. The SGBS cell strain as a model for the in vitro study of obesity and cancer. Clin Transl Oncol. 2012;14(10):774-82.
207. Temprano A, Sembongi H, Han GS, Sebastian D, Capellades J, Moreno C, et al. Redundant roles of the phosphatidate phosphatase family in triacylglycerol synthesis in human adipocytes. Diabetologia. 2016;59(9):1985-94.
208. Hjorth M, Pourteymour S, Gorgens SW, Langleite TM, Lee S, Holen T, et al. Myostatin in relation to physical activity and dysglycemia and its effect on energy metabolism in human skeletal muscle cells. Acta Physiol. 2015;217(1):45-60.
209. Bordicchia M, Ceresiani M, Pavani M, Minardi D, Polito M, Wabitsch M, et al. Insulin/glucose induces natriuretic peptide clearance receptor in human adipocytes: a metabolic link with the cardiac natriuretic pathway. Am J Physiol. 2016;311(1):R104-14.
210. Hoggard N, Cruickshank M, Moar KM, Bashir S, Mayer CD. Using gene expression to predict differences in the secretome of human omental vs. subcutaneous adipose tissue. Obesity (Silver Spring). 2012;20(6):1158-67.
211. Zoller V, Funcke JB, Keuper M, Abd El Hay M, Debatin KM, Wabitsch M, et al. TRAIL (TNF-related apoptosis-inducing ligand) inhibits human adipocyte differentiation via caspase-mediated downregulation of adipogenic transcription factors. Cell Death Dis. 2016;7(10):e2412.
212. Limoge F, Faivre L, Gautier T, Petit JM, Gautier E, Masson D, et al. Insulin response dysregulation explains abnormal fat storage and increased risk of diabetes mellitus type 2 in Cohen Syndrome. Hum Mol Gen. 2015;24(23):6603-13.
213. Chan PC, Hsiao FC, Chang HM, Wabitsch M, Hsieh PS. Importance of adipocyte cyclooxygenase-2 and prostaglandin E2-prostaglandin E receptor 3 signaling in the development of obesity-induced adipose tissue inflammation and insulin resistance. FASEB. 2016;30(6):2282-97.
214. Sharma NK, Varma V, Ma L, Hasstedt SJ, Das SK. Obesity associated modulation of miRNA and co-regulated target transcripts in human adipose tissue of non-diabetic subjects. MicroRNA. 2015;4(3):194-204.
215. Gavalda-Navarro A, Mampel T, Vinas O. Changes in the expression of the human adenine nucleotide translocase isoforms condition cellular metabolic/proliferative status. Open Biol. 2016;6(2):150108.
216. Mysore R, Zhou Y, Sadevirta S, Savolainen-Peltonen H, Nidhina Haridas PA, Soronen J, et al. MicroRNA-192* impairs adipocyte triglyceride storage. BBA. 2016;1861(4):342-51.
217. Chang CW, Abhinav K, Di Cara F, Panagakou I, Vass S, Heck MM. A role for the metalloprotease invadolysin in insulin signaling and adipogenesis. Biol Chem. 2016.
218. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Gen. 2015;24(15):4464-79.

219. Su S, Zhu H, Xu X, Wang X, Dong Y, Kapuku G, et al. DNA methylation of the LY86 gene is associated with obesity, insulin resistance, and inflammation. Twin research and human genetics : Twin Res Hum Gene. 2014;17(3):183-91.
220. Grundberg E, Meduri E, Sandling JK, Hedman AK, Keildson S, Buil A, et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. Am J Hum Genet. 2013;93(5):876-90.
221. Irvin MR, Zhi D, Joehanes R, Mendelson M, Aslibekyan S, Claas SA, et al. Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. Circulation. 2014;130(7):565-72.
222. Frazier-Wood AC, Aslibekyan S, Absher DM, Hopkins PN, Sha J, Tsai MY, et al. Methylation at CPT1A locus is associated with lipoprotein subfraction profiles. J Lipid Res. 2014;55(7):1324-30.
223. Pfeiffer L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA methylation of lipid-related genes affects blood lipid levels. Circ Cardiovasc Gene. 2015;8(2):334-42.
224. Hidalgo B, Irvin MR, Sha J, Zhi D, Aslibekyan S, Absher D, et al. Epigenome-wide association study of fasting measures of glucose, insulin, and HOMA-IR in the genetics of lipid lowering drugs and diet network study. Diabetes. 2014;63(2):801-7.
225. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-30.
226. Pitsiladis YP, Tanaka M, Eynon N, Bouchard C, North KN, Williams AG, et al. Athlome Project Consortium: a concerted effort to discover genomic and other "omic" markers of athletic performance. Physiol Genomics. 2016;48(3):183-90.
227. Nica AC, Parts L, Glass D, Nisbet J, Barrett A, Sekowska M, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genetics. 2011;7(2):e1002003.
228. Pearce MS, Unwin NC, Parker L, Craft AW. Cohort profile: the Newcastle Thousand Families 1947 birth cohort. Int J Epidemiol. 2009;38(4):932-7.
229. Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H, et al. Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). Int J Epidemiol. 2015.
230. Lumey LH, Stein AD. Transgenerational effects of prenatal exposure to the Dutch famine. Int J Obstet Gynaecol. 2009;116(6):868.
231. Madrigano J, Baccarelli A, Mittleman MA, Wright RO, Sparrow D, Vokonas PS, et al. Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. Environ Health Perspect. 2011;119(7):977-82.
232. Zhang A, Hu H, Sánchez BN, Ettinger AS, Park SK, Cantonwine D, et al. Association between prenatal lead exposure and blood pressure in children. Environ Health Perspect. 2012;120(3):445-50.
233. Banks E, Redman S, Jorm L, Armstrong B, Bauman A, Beard J, et al. Cohort profile: the 45 and up study. International journal of epidemiology. 2008;37(5):941-7.
234. NSW Admitted Patient Data Collection (APDC) [Internet]. NSW Ministry of Health. Available: http://www.cherel.org.au/data-dictionaries. [cited 3/12/2016]

APPENDIX 1: Epigenetics and human obesity

Recent technological advances in epigenome profiling have led to an increasing number of studies investigating the role of the epigenome in obesity. An electronic literature search for relevant articles, published between September 2008 and September 2013 was performed. From the 319 articles identified, 46 studies were included and reviewed. The studies provided no consistent evidence for a relationship between global methylation and obesity.

I have summarized the main findings in the articles identified using the search criteria, and constructed the tables in the manuscript.

This manuscript has been published in *International Journal of Obesity,* January 2015, Cited 75 times (at November 2016) and listed as highly cited by Thompson ISI.

Detailed contributions:

	Varinli H	**Co-authors**
Literature Review	20%	80%
Data Interpretation	10%	90%
Writing	10%	90%

van Dijk, S. J., Molloy, P. L., Varinli, H., Morrison, J. L., Muhlhausler, B. S., & members of EpiSCOPE. (2015). Epigenetics and human obesity. *International Journal of Obesity*, 39, p. 85-97.

DOI: 10.1038/ijo.2014.34

Supplementary table 1: Methods of DNA methylation analysis, including those used in reviewed studies

	Sample preparation (a)	Analysis Method	Pros	Cons	Reference	Comments
Global methylation						
	Digestion of DNA to nucleotides	HPLC	Accurate measure of total fraction of 5meC	Requires larger amounts of DNA	(b)	Other quantification methods available (eg. mass spectrometry)
	Sodium bisulfite conversion, with specific PCR of amplicon	Pyrosequencing of repetitive elements (LINE, Alu, SAT)	High throughput, cost effective, low amount of DNA needed, targeting sites that are important for stability genome	Selected sites, surrogate of global, high apparent background of unmethylated cytosine due to CG to TG mutation.	(b)	
	DNA bound to microtiter plate	Antibody to 5meC	All sites (total 5 methylcytosine), high throughput	Requires accurate DNA concentration determination		Commercial kits available
	DNA digestion with methyl-sensitive enzymes	End-specific PCR for repeat sequence methylation	Low amount of DNA needed, no need for bisulfite conversion.	Measures sub-fraction of repetitive elements	(c)	
Candidate gene methylation						
	Sodium bisulfite conversion, with specific amplification of target amplicon	Mass spectrometry, Sequenom Epityper MassARRAY	High throughput and some scope for multiplexing, can measure CpGs in fragments to ~100 bp	Primer design constraints, requires in-house equipment (Sequenom) or use of service provider	(d)	
		Pyrosequencing	High throughput	Requires in-house equipment (Qiagen) or use of service provider	(e)	
		Bisulfite sequencing	Single base pair resolution. Can sequence clones or amplified population of molecules.	Need high number of reads to get accurate quantification	(a)	It is now reasonable to perform multiplexed amplicon sequencing of bisulfite-treated DNA on Illumina Miseq/NextSeq or Life Technologies' Ion Torrent platforms
		Single nucleotide extension on bisulfite converted DNA (MS-SNuPE)	Quantitative, capable of being multiplexed.	Oligonucleotide primer needed for each CpG site measured	(f)	Variation of original SNuPE method
	Sodium bisulfite conversion	Quantitative Methylation specific PCRs	Targeted, small quantity DNA needed	Primer design constraints, amplification of DNA	(a, g)	
	Affinity enrichment using methyl CpG binding protein	qPCR for target promoter regions	Data relatively easy to interpret	Capture efficiency is dependent on number and density of methylated CpG sites, so quantification non-linear	(h)	

Supplementary table 1 cont'd: Methods of DNA methylation analysis, including those used in reviewed studies

	Sample preparation (a)	Analysis Method	Pros	Cons	Reference	Comments
Genome wide methylation						
	Sodium bisulfite conversion	Array based Illumina Infinium 450K (development upon previous Golden gate and 27K array platforms)	Cost effective, high number of sites, robust and overall good accuracy	Selected sites (eg. enhancer regions not well covered), problems with some probes - cross hybridisation, presence of SNPs can impact on analysis and interpretation	(i)	
	Digestion with methylation-sensitive enzymes	Comprehensive high-throughput array-based relative methylation (CHARM) NimbleGen array	High coverage	Boutique method. Original array no longer available.	(j)	Isolation of the methylated (and/or unmethylated) fractions of the genome followed by array hybridization or sequencing can potentially provide broader coverage, but at greater cost, than a dedicated array such as the Illumina 450K Bead Chip
	Methyl-cytosine DNA immunoprecipitation	Promoter tiling array or deep sequencing	High coverage	Restricted to regions selected for array (this specific array no longer available)	(k)	
	Affinity enrichment using methyl CpG binding protein	Deep sequencing of captured DNA	High coverage	Cost	(h)	

Additional References
(a) General reference for bisulphite-based methods and consideration of issues such an incomplete conversion and amplification bias.
Clark et al. Nature Protocols 1:2353-64 (2006)
(b) Global DNA methylation comparative study.
Lisanti et al., PLoSOne 8:e79044 (2013)
(c) Rand & Molloy Biotechniques 49:xiii-xvii (2010)
(d) Coolen et al., Nucleic Acids Res. 35:e119 (2007)
(e) Bassil et al., Methods Mol. Biol. 1049:95-107 (2013)
(f) Gonzalgo & Liang, Nature Protocols 2:1931-36 (2007)
(g) Campan et al., Methods Mol. Biol. 507:325-37 (2009)
(h) Barres et al., Cell Rep. 3:1020-27 (2013)
(i) Sandoval et al., Epigenetics 6:692-702 (2011)
(j) Irizarry et al., Genome res. 18:780-90 (2008)

APPENDIX 2: Human adipocytes have unique epigenetic signatures based on their origin: an effect exacerbated by obesity

Adipose tissue plays an important role in metabolism and homoeostasis. Perturbations in the nature and amount of adipose tissue have serious health consequences for individuals and society. An increase in visceral adipose tissue (VAT) is associated with diseases ranging from Type 2 Diabetes to cancer, while subcutaneous adipose tissue (SAT) may have a protective effect. Several genes are differentially expressed between these two adipose depots, however the mechanisms controlling this differential expression remain poorly understood.

Our aims were to characterize potential functional and developmental differences between adipocytes from the two fat depots in lean individuals, then distinguish the methylation and transcriptional differences between visceral adipocytes of normal weight and obese individuals. We have achieved so in two stages. Firstly we analysed both the transcriptomes and methylomes of adipocyte cells purified from SAT and VAT of 3 lean female subjects. Then we extended our analysis of methylome and transcriptome of visceral adipocytes to a total of 14 subjects (7 lean and 7 obese).

Methylomes were mapped using both whole genome bisulfite sequencing (WGBS) and Illumina 450k arrays and transcriptomes were characterized using RNA-seq.

This manuscript is under preparation for Nature Genetics.

Detailed contributions:

	Varinli H	**Co-authors**
Experimental Design	10%	90%
Sample Collection	60%	40%
Laboratory Experiments	30%	70%
Data Analysis	10%	90%
Data Interpretation	20%	80%
Writing	10%	90%

To demonstrate the contribution of this study to the field, I have summarised the significant outcomes and my detailed contributions below:

1. This study provides first complete maps of primary adipocyte methylomes from both normal weight and obese subjects.
2. We found that methylation profiles of each sample type were highly consistent between lean individuals and clearly segregated visceral adipocytes (VA) from subcutaneous adipocytes (SA). The distinct methylation patterns in VA and SA are indicative of different developmental origins.
3. The transcriptional differences between SA and VA reflected known functional differences between their tissue origins. Particularly, gene ontology analyses revealed that the observed methylation differences signify the separate developmental origins of SA and VA, while differential expression of lipid metabolism and inflammatory response genes may reflect functional and environmental differences, respectively, between SAT and VAT. A total of 4,868 regions were identified as differentially methylated (Δ-beta methylation >0.1).
4. The methylome profiles of visceral adipose tissue (VAT) and purified adipocytes (VA) from lean individuals were highly concordant but the transcriptome profiles of the same samples types were more distinct. This indicates that non-adipocyte component of VAT contributes to the transcriptome profile of the tissue.
5. The methylation of visceral adipocytes (VA) from obese subjects largely differed from those of lean subjects.
6. In general, there was higher inter-individual variation in methylation profiles of visceral adipocytes (VA) from obese individuals.
7. HOX genes are transcription factors and important in cell type specification. There are clear patterns in the differential methylation of HOX genes which is indicative of expansion of a subpopulation of adipocytes of different developmental lineage accompanying depot expansion in response to body fat increase.
8. A total of 402 genes were differentially expressed in visceral adipocytes of lean vs obese individuals. In fact, nearly 50% of these changes were also identified between SA and VA of lean individuals.

9. Integration of transcriptome and methylome data identified key loci involved in adipocyte biology. For instance, we observed increase in expression of genes encoding pro-inflammatory cytokines in visceral adipocytes are associated with reduced methylation. We also identified key points of epigenetic regulation that are intrinsic to adipocyte specific functions such as lipid metabolism and inflammation related pathways.

My contributions were:

Donor Recruitment and Sample Collection:

- I conducted pre-surgery patient interviews. This involved documenting medical and family history and also taking the body measurement of the subjects.
- I attended most of the surgeries and collected biological samples (Blood, saliva and adipose tissue).

Sample Processing:

- I contributed to the optimization of laboratory protocols used in this study. Particularly to:
    - Human adipocyte isolation from adipose tissue
    - RNA isolation from human adipose tissue and human adipocytes
- I pre-processed the biological samples and documented them on a database. This specifically involved:
    - Human adipocyte purification from visceral and subcutaneous adipose tissue for most of the subjects.
    - White blood cell isolation and/or buccal cell isolation from blood and saliva samples respectively for some of the subjects.
- I isolated RNA for the transcriptome component of this project and planned the batching of samples to avoid technical biases in the analysis of RNA-seq results.

Project Operation and Manuscript:

- I contributed to the data collection, processing, analysis and discussions.
- I contributed to the manuscript planning and review.

APPENDIX 3: Atmospheric gas plasma–induced ROS production activates TNF-ASK1 pathway for the induction of melanoma cancer cell apoptosis

Atmospheric gas plasmas (AGPs) are able to selectively induce apoptosis in cancer cells, offering a promising alternative to conventional therapies that have unwanted side effects such as drug resistance and toxicity. However, the mechanism of AGP-induced cancer cell death is unknown.

In this manuscript, AGP is shown to up-regulate intracellular reactive oxygen species (ROS) levels and induce apoptosis in melanoma, but not normal melanocyte cells. By screening genes involved in apoptosis, we identified tumor necrosis factor (TNF)-family members as the most differentially expressed cellular genes upon AGP treatment of melanoma cells. TNF receptor 1 (TNFR1) antagonist-neutralizing antibody specifically inhibits the AGP-induced apoptosis signal, regulating apoptosis signal-regulating kinase 1 (ASK1) activity and subsequent ASK1-dependent apoptosis. Treatment of cells with the intracellular ROS scavenger N-acetyl-l-cysteine also inhibits AGP-induced activation of ASK1, as well as apoptosis. Moreover, depletion of intracellular ASK1 reduces the level of AGP-induced oxidative stress and apoptosis. The evidence for TNF-signaling dependence of ASK1-mediated apoptosis suggests possible mechanisms for AGP activation and regulation of apoptosis-signaling pathways in tumor cells.

I contributed to the experimental design of the molecular biology experiments and the interpretation of the results.

This manuscript has been published in *Molecular Biology of the Cell,* February 2014 and cited 42 times (at November 2016).

Detailed contributions:

	Varinli H	**Co-authors**
Laboratory Experiments	20%	80%
Data Analysis	20%	80%
Data Interpretation	20%	90%
Writing	10%	90%

MBoC | ARTICLE

Atmospheric gas plasma–induced ROS production activates TNF-ASK1 pathway for the induction of melanoma cancer cell apoptosis

Musarat Ishaq[a,b], Shailesh Kumar[b], Hilal Varinli[c], Zhao Jun Han[b], Amanda E. Rider[b], Margaret D. M. Evans[a], Anthony B. Murphy[b], and Kostya Ostrikov[b,d]
[a]CSIRO Materials Science and Engineering, North Ryde, NSW 1670, Australia; [b]Plasma Nanoscience Center, CSIRO Materials Science and Engineering, Lindfield, NSW 2070, Australia; [c]CSIRO Animal, Food and Health Sciences, North Ryde, NSW 1670, Australia; [d]School of Chemistry, Physics, and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD 4000, Australia

ABSTRACT **Atmospheric gas plasmas (AGPs) are able to selectively induce apoptosis in cancer cells, offering a promising alternative to conventional therapies that have unwanted side effects such as drug resistance and toxicity. However, the mechanism of AGP-induced cancer cell death is unknown. In this study, AGP is shown to up-regulate intracellular reactive oxygen species (ROS) levels and induce apoptosis in melanoma but not normal melanocyte cells. By screening genes involved in apoptosis, we identify tumor necrosis factor (TNF)–family members as the most differentially expressed cellular genes upon AGP treatment of melanoma cells. TNF receptor 1 (TNFR1) antagonist–neutralizing antibody specifically inhibits AGP-induced apoptosis signal, regulating apoptosis signal–regulating kinase 1 (ASK1) activity and subsequent ASK1-dependent apoptosis. Treatment of cells with intracellular ROS scavenger *N*-acetyl-L-cysteine also inhibits AGP-induced activation of ASK1, as well as apoptosis. Moreover, depletion of intracellular ASK1 reduces the level of AGP-induced oxidative stress and apoptosis. The evidence for TNF-signaling dependence of ASK1-mediated apoptosis suggests possible mechanisms for AGP activation and regulation of apoptosis-signaling pathways in tumor cells.**

Monitoring Editor
Carl-Henrik Heldin
Ludwig Institute for Cancer Research

Received: Oct 16, 2013
Revised: Feb 19, 2014
Accepted: Feb 19, 2014

INTRODUCTION

In recent decades, there has been significant progress in the development of new therapies to treat human cancers. However, fundamental problems related to chemotherapeutic drug delivery—resistance and toxicity to normal cells—remain. An ideal anticancer treatment should selectively kill cancer cells with limited side effects on normal cells and minimal drug resistance. Selective induction of apoptosis in target cancer cells would be an ideal treatment (Nagata, 1997; Kaelin, 1999).

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E13-10-0590) on February 26, 2014.

Address correspondence to: Musarat Ishaq (Musarat.Ishaq@csiro.au), Kostya Ostrikov (Kostya.Ostrikov@csiro.au).

Abbreviations used: AGP, atmospheric gas plasmas; ASK-1, apoptosis signal–regulating kinase 1; B2M, beta-2 microglobulin; BCA, bicinchoninic acid; CD95, cluster of differentiation 95; CM-H2DCFDA, 2′-,7′-dichlorofluorescein diacetate; CSIRO, Commonwealth Scientific and Industrial Research Organization; DAF-FM, 4-amino-5-methylamino-2′,7′-difluofluorescein; DPI, diphenyleneiodonium chloride; DR4/5, death receptor 4/5; FAS, TNF receptor superfamily, member 6; FBS, fetal bovine serum; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; GSH, glutathione; GSSG, glutathione disulfide; GUS-b, glucuronidase beta; H2AX, H2A histone family, member X; γ-H2AX, gamma-H2AX; HPRT1, hypoxanthine phosphoribosyltransferase 1; HRP, horseradish peroxidase; IgG, immunoglobulin G; JNK, c-Jun N-terminal kinase; MC, melanocytes; MKK3, dual specificity mitogen-activated protein kinase kinase 3; MKK4, mitogen-activated protein kinase kinase 4; MKK6, dual specificity mitogen-activated protein kinase kinase 6; MKK7, dual specificity mitogen-activated protein kinase kinase 7; MRC5, human fetal lung fibroblasts; MTS, (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium); NAC, *N*-acetyl-L-cysteine; OCE, Office of the Chief Executive; p38 MAPK, p38 mitogen-activated protein kinase; PMS, phenazine methosulfate; PPIA, peptidylprolyl isomerase A; qPCR, quantitative PCR; RF, radio frequency; RNAi, RNA interference; ROS, reactive oxygen species; RPL13A, ribosomal protein L13a; SAPK, stress-activated protein kinases; SYBR, *N*′,*N*′-dimethyl-*N*-[4-[(E)-(3-methyl-1,3-benzothiazol-2-ylidene)methyl]-1-phenylquinolin-1-ium-2-yl]-*N*-propylpropane-1,3-diamine; TNF, tumor necrosis factor; TNFα, tumor necrosis factor α; TNFR1, tumor necrosis factor receptor 1; Trx, thioredoxin; Z-VAD-FMK, 3-[2-(2-benzyloxycarbonylamino-3-methyl-butyrylamino)-propionylamino]-4-oxo-pentanoic acid.

Two pathways of apoptosis have been described for mammalian cells: the intrinsic mitochondrial pathway that is activated in response to cellular stress, and the extrinsic death receptor pathway that is activated at the cell surface by the binding of tumor necrosis factor (TNF)–family cytokines to their cognate death receptors (TNFR1, Fas/CD95, DR4/5). The extrinsic and the intrinsic pathways converge in a caspase cascade that results in cellular shrinkage and DNA fragmentation, culminating in apoptosis (Micheau and Tschopp, 2003; Kamata *et al.*, 2005).

Cellular oxidative stress can induce apoptosis by initiating the activation of a specialized group of mitogen-activated protein kinase (MAPK) cascades. Apoptosis signal–regulating kinase 1 (ASK1) is a MAPK kinase kinase (MAPKKK)–family member that plays a major role in stress-induced apoptosis. ASK1 is activated by various stress-related stimuli, including oxidative stress, reactive oxygen species (ROS), genotoxic agents, serum withdrawal, endoplasmic reticulum stress, and tumor necrosis factor (TNF; Ichijo *et al.*, 1997; Nishitoh *et al.*, 1998; Liu *et al.*, 2000). The activated ASK1 phosphorylates and activates the downstream kinases MKK4/MKK7 and MKK3/MKK6. These are in turn required to activate c-Jun N-terminal kinase (JNK) and p38 MAPK kinase, respectively, before caspase-3 activation and apoptosis. The knockdown of ASK1 inhibits ROS (H_2O_2)- and TNF-induced apoptosis (Noguchi *et al.*, 2005, 2008). Overexpression of a constitutively active form of ASK1 in cancer cells induces caspase-3–dependent apoptosis (Ichijo *et al.*, 1997). ROS such as H_2O_2 activate ASK1 by dissociation of thioredoxin (Trx), a reduction/oxidation (redox) regulatory protein that inhibits the kinase activity of ASK1 (Saitoh *et al.*, 1998). It has also been proposed that ROS generated by cytokines like TNFα or stress may oxidize and consequently dissociate Trx to activate ASK1 and subsequent ASK1-dependent apoptosis-signaling cascades (Saitoh *et al.*, 1998; Liu *et al.*, 2000; Nadeau *et al.*, 2009). These findings demonstrate that ASK1 plays an important role in stress-induced apoptosis.

Recently atmospheric gas plasmas (AGPs) delivered in a plume have shown the potential to be a safe anticancer therapy that can kill selectively a variety of cancer cells, such as melanoma (Keidar *et al.*, 2011; Kim *et al.*, 2011a), neuroblastoma (Walk *et al.*, 2013), glioma (Koeritzer *et al.*, 2013), and colorectal (Vandamme *et al.*, 2012), pancreatic (Partecke *et al.*, 2012), breast (Kalghatgi *et al.*, 2011), liver (Yan *et al.*, 2012), ovarian (Iseki *et al.*, 2012), leukemia (Thiyagarajan *et al.*, 2012), and lung (Huang *et al.*, 2011) cancers. AGP delivered as a plume has also shown significant effects in in vivo animal models by decreasing the size of tumors (Vandamme *et al.*, 2010, 2012; Keidar *et al.*, 2011; Partecke *et al.*, 2012; Walk *et al.*, 2013). It has been proven effective against drug-resistant brain tumor (glioblastoma) in combination with temozolomide (Koeritzer *et al.*, 2013). The AGP plume generated by an electrical discharge in a gas generates reactive oxygen species (e.g., H_2O_2, O_3, OH, NO), ions, and free electrons (Sensenig *et al.*, 2008; Liu *et al.*, 2010; Weltmann *et al.*, 2010; Keidar *et al.*, 2011; Graves, 2012; Ostrikov *et al.*, 2013). AGP leads to the formation of reactive oxygen species, including H_2O_2, in liquids that simulate cellular fluids (Winter *et al.*, 2013). Despite early promising results obtained using AGP as an antitumor treatment to kill cancer cells, little effort has been made to translate this into clinical applications. This is possibly due to a lack of understanding about the underlying cellular mechanisms and the means by which AGP can selectively kill cancer cells by inducing apoptosis without affecting normal cells.

In this study, we find that AGP treatment enhances the production of intracellular ROS. This elevated level of intracellular ROS induces oxidative stress, leading to apoptosis in melanoma cells without affecting normal control cells. Moreover, AGP treatment activates TNF and ASK1 pathways to increase downstream activity of JNK and p38 kinases and stimulate caspase-3/7–dependent apoptosis. Thus AGP shows selective induction of apoptosis in cancer cells by stimulating the oxidative stress–induced TNF-ASK1-JNK/p38–caspase-3/7 apoptotic pathway.

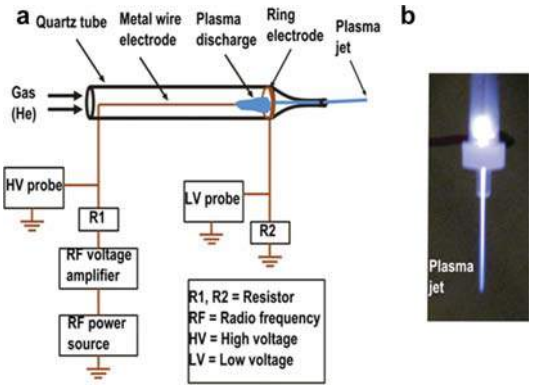


FIGURE 1: (a) Schematic of the atmospheric gas plasma jet device. (b) Digital photo of the generated gas plasma jet for cell treatment.

RESULTS

Selective apoptotic response to AGP is ROS dependent

We quantified the effect of AGP (presented and briefly described in *Materials and Methods*, Figure 1, and Supplemental Methods) on the cellular ROS level in melanoma cancer cells (Mel007 and Mel-RM) using the redox-sensitive fluorescent probe 2′-,7′-dichlorofluorescein diacetate (CM-H2DCFDA). Treatment with AGP caused a significant increase in ROS level in melanoma cells. AGP did not increase ROS level significantly in normal control cells (primary human epidermal melanocytes and human fetal lung fibroblasts (Figure 2a). Cotreatment with the intracellular ROS inhibitor *N*-acetyl-L-cysteine (NAC) fully reversed the AGP-induced increase in ROS in melanoma cells (Figure 2b). It has been shown that an increase in ROS production can lead to decrease in reduced glutathione (GSH) levels in cancer cells (Estrela *et al.*, 2006). GSH is an important intracellular antioxidant that protects cells from damage caused by ROS. It is able to remove O_2•– and provide electrons for glutathione peroxidase to reduce H_2O_2 to H_2O. Because AGP has been shown to induce ROS production and modulate redox homeostasis (Graves, 2012), we tested whether AGP can lead to a decrease in GSH levels. This was found to occur in melanoma cells. However AGP-mediated GSH depletion was inhibited by addition of the reducing agent NAC (Figure 2c). Parallel counting of viable cells showed that AGP-induced cell death in melanoma cells correlated with decrease in cellular GSH contents (Supplemental Table S1). We also found that nitric oxide was among the ROS species induced by AGP in melanoma cells but not in normal control cells (Figure 2d). Supplemental Figure S1 demonstrates the optical emission spectrum of the AGP plume over the range from 300 to 800 nm, further confirming that excited species of OH, N_2, N_2^+, He, and O exist in the AGP plume.

We next examined the effect of AGP on the viability of several cultured melanoma cells (Mel-RM, Mel007, and Mel-JD) and normal control cells (primary human epidermal melanocytes and

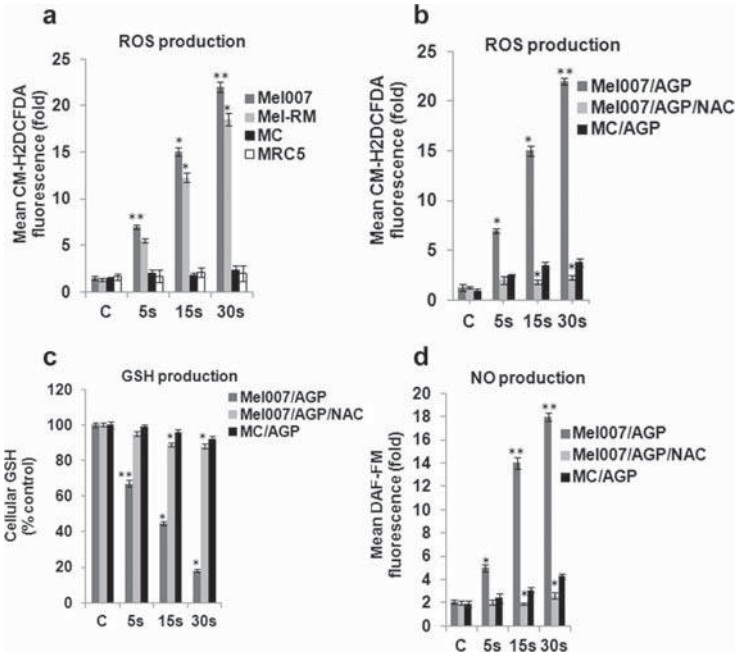


FIGURE 2: AGP enhances ROS accumulation in melanoma cells. (a) AGP enhances ROS levels in melanoma cells (Mel-RM, Mel007) but not normal cells (melanocytes, human fetal lung fibroblasts). Cells were treated with AGP (5, 15, 30 s). ROS level was quantified by the fluorescent dye CM-H2DCFDA and is shown as a fold change relative to control cells treated only with He gas flow. Data were normalized to ROS levels in control melanocyte cells. (b) Mel007 or melanocyte cells were treated with AGP (5, 15, 30s) or pretreated with NAC (1–2 h) followed by AGP treatment (5, 15, 30 s). ROS level was determined as in a. (c) Mel007 or melanocyte cells were treated with AGP or pretreated with NAC (1–2 h), followed by AGP treatment. GSH level was quantified as described in *Materials and Methods*. (d) Mel007 or melanocyte cells were treated with AGP (5, 15, 30 s) or pretreated with NAC (1–2 h), followed by AGP treatment (5, 15, 30 s). Nitric oxide was quantified after cells were labeled with DAF-FM diacetate. All values are mean ± SD of three independent experiments performed in triplicate. *$p \leq 0.01$, **$p \leq 0.001$; ANOVA.

human fetal lung fibroblasts). AGP treatment significantly induced apoptosis in the melanoma cells. When primary cultures of normal cells were treated with AGP, there was little effect on cell growth and proliferation. This indicated that AGP selectively induces cell death in melanoma cells (Figure 3a). As shown earlier (Figure 2, a and d), AGP induced intracellular ROS, so we next tested to see whether the selective cytotoxic effect of AGP on melanoma cells was ROS dependent or independent. Cotreatment of AGP and NAC completely reversed the toxic effects of AGP in melanoma cells (Figure 3b). AGP induced apoptosis and DNA damage in melanoma cells, as observed by increased activity of caspases 3/7 and stress response target protein γ-H2AX in AGP-treated melanoma cells, with no significant activity in normal cells (Figure 3, c and d). Taken together, this different response of cancer and normal cells to AGP treatment indicates that AGP targets cancer cell redox homeostasis, which results in both a stress response and DNA damage, leading to apoptosis in cancer cells. This selective induction of ROS in cancer cells indicates that the AGP-specific apoptotic response in melanoma cells is mediated by perturbation of cellular redox homeostasis.

TNF is involved in AGP-induced apoptosis

Next we examined the mechanism by which AGP specifically induces apoptosis in melanoma cells. To find the specific cellular factors involved in selective AGP-induced apoptosis, we screened >90 genes involved specifically in prosurvival or proapoptotic pathways by using real-time quantitative PCR (qPCR). We found that TNF family members are the cellular factors that most frequently showed differential expression in AGP-treated melanoma cells relative to control untreated cells (Supplemental Figure S2). We confirmed our qPCR gene expression screening data by measuring the gene expression of TNF-receptor family member 1 (TNFR1) in melanoma cells treated with AGP for different time intervals (5, 15, and 30 s) by qPCR and Western blotting. However, this AGP-induced TNFR1 expression was inhibited by the ROS scavenger NAC (Figure 4, a and b). This shows the involvement of ROS in AGP-induced TNF signaling. Moreover, we also observed increase in the production of TNF signaling ligand (TNFα) in AGP-treated melanoma (Mel007) cells but not in normal melanocytes (Figure 4c). Outcomes showed increased activity of stress response target protein γ-H2AX in the AGP-treated melanoma cells but no significant activity in melanoma cells pretreated with anti–TNFR1-neutralizing antibody (Figure 4d). Determination of cell viability and caspase 3/7 activity demonstrated that AGP-induced apoptosis and cytotoxicity was inhibited by cotreatment of AGP with antagonistic anti–TNFR1-neutralizing antibody, the caspase inhibitor Z-VAD-FMK, the inhibitor of nitric oxide synthetase diphenyleneiodonium chloride (DPI), or the H_2O_2 depleter catalase (Figure 4, e and f). These results indicate that selective AGP-induced apoptosis in melanoma cells is dependent on intracellular ROS production. Moreover, cell death induced by AGP is mediated by TNF-receptor pathways by activating caspase 3/7 activity.

ASK1 is required for AGP-induced apoptosis

We next sought to explore the downstream signaling factors involved in the AGP-induced TNF apoptosis pathway. It was shown that ROS (e.g., H_2O_2) induced by the TNF pathway plays a role in programmed cell death signaling mediated by ASK1 (Noguchi *et al.*, 2005). We examined whether ASK1 is involved in AGP-induced apoptosis in melanoma cells. AGP induced ASK1 activity at a similar level to H_2O_2. However, this ASK1 activation was inhibited by the ROS scavenger NAC (Figure 5a). This shows that AGP-induced ASK1 activation is ROS dependent. We further determined that AGP activated the ASK1-activity–dependent downstream-signaling kinases p38 and JNK at a level similar to H_2O_2, whereas this activation was strongly reduced by NAC (Figure 5, b and c).

ASK1 and JNK/p38 kinases play important roles in DNA damage responses by inducing cyclin-dependent kinase inhibitors, and

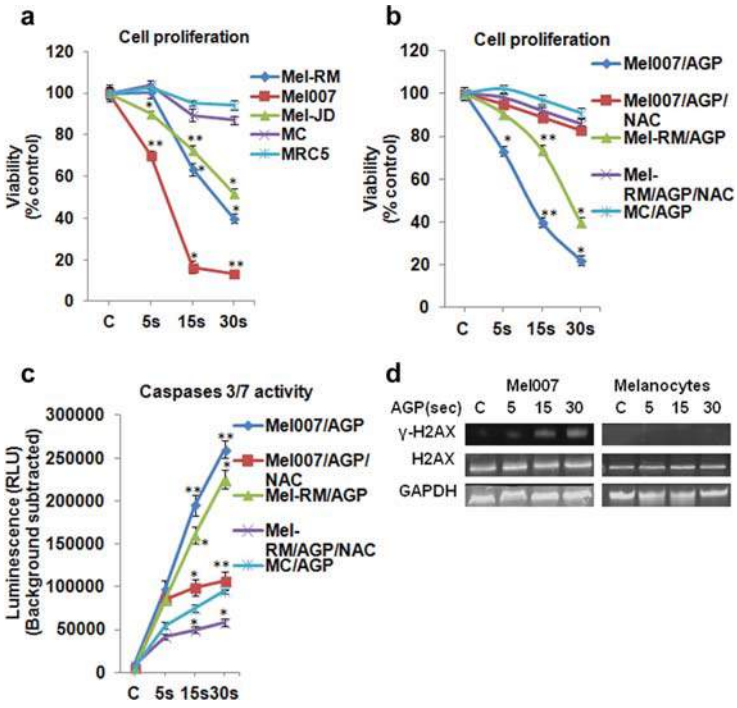


FIGURE 3: AGP selectively induces apoptosis in melanoma cells, and this is ROS dependent. (a) AGP treatment induced cell death in melanoma cells but not normal cells. Melanoma cells (Mel-RM, Mel007, Mel-JD), melanocytes (MC), and human fetal lung fibroblasts (MRC5) were cultured in 96-well plates overnight, treated with AGP for 5–30 s, and grown for 18–24 h before analysis. Cell viability was measured by cell titer nonradioactive cell proliferation assay. (b) AGP-induced cell death in melanoma cells was reversed by NAC. Mel007 or Mel-RM cancer cells or melanocytes were treated with AGP (5, 15, 30 s) or pretreated with NAC for 1–2 h, followed by AGP treatment (5, 15, 30 s) for 18–24 h. Cell viability was measured by cell titer nonradioactive cell proliferation assay. (c) Mel007, Mel-RM, and melanocytes were treated with AGP or pretreated with NAC. Caspase 3/7 activity was measured by Caspase-Glo 3/7 assay. (d) The effect of AGP on stress response targets was determined by Western blot analysis of H2AX and γ-H2AX protein in normal (melanocytes) and melanoma cells (Mel007). GAPDH expression was used as a loading control. In all experiments, control cells were mock treated with He gas flow only. All values are mean ± SD of three independent experiments performed in triplicate. *$p \leq 0.01$, **$p \leq 0.001$; ANOVA.

depletion of ASK1 significantly attenuates p38 activation and DNA damage response (Thornton and Rincon, 2009). ASK1 is also required for H_2O_2 (ROS)-induced DNA fragmentation and apoptosis. It was shown that depletion of the ROS scavenger NAC as well as of ASK1 strongly reduced DNA fragmentation and apoptosis (Noguchi *et al.*, 2008). To determine the specific role of ASK1 in AGP-induced apoptosis, we knocked down the ASK1 gene in melanoma cells by RNA interference (RNAi; Figure 5d). First, we tested the effect of AGP on caspase 3/7 activity in ASK1-knockdown melanoma cells. The results showed that activation of caspases 3/7 induced by AGP was inhibited in ASK1-depleted melanoma cells, as in cells pretreated with intracellular ROS scavenger NAC (Figure 5e). We also observed that pretreatment of melanoma cells with either NAC or knockdown of ASK1 by RNAi inhibited the toxic effects of AGP in melanoma cells (Figure 5f). Our results also showed that AGP-induced activation of stress response protein (H2AX) was inhibited in ASK1-knockdown or NAC-pretreated melanoma cells (Figure 5g). These results show the involvement of ASK1 and downstream kinases p38 and JNK in AGP-induced apoptosis in melanoma cells and clearly reveal that ASK1 plays a critical role in selective AGP-dependent signal transduction and apoptosis in melanoma cells.

TNF pathway involved in AGP-induced activation of ASK1 signaling and apoptosis

From the results already presented, it is plausible to conclude that ASK1 is an AGP effector that activates JNK and p38 kinases. However, the mechanism of ASK1 activation by AGP is still unclear. Previous studies show that TNFR family members are involved in the activation of stress-activated protein kinases (SAPKs) and MAPKs (Natoli *et al.*, 1997). The similarity between AGP- and H_2O_2-induced changes in the ASK1 signaling complex (Figure 5), together with the findings that TNF-induced activation of ASK1 depends largely on ROS, prompted us to examine whether TNF signaling may also be involved in AGP-induced activation of the ASK1 pathway (Natoli *et al.*, 1997). Tumor necrosis factor 1 (TNFR1) blocking antibody was used to neutralize the TNFR1 or inhibit the TNF signaling pathway (Defer *et al.*, 2007; Moh *et al.*, 2013).

AGP stimulates the production of ROS and TNF, and both ROS and TNF activation of ASK1 can be inhibited on depletion of ROS by using free radical scavenger NAC and the TNFR1-antagonist-neutralizing antibody (Figure 6a). Similarly, AGP-induced activation of p38 and JNK signaling were suppressed in melanoma cells pretreated with NAC or antagonist anti-TNFR1 antibody (Figure 6, b and c). ASK1, p38, and JNK kinases were found in the inactive form in normal melanocytes and melanoma (Mel007) cells at basal level. Moreover, AGP treatment did not activate ASK1, p38, and JNK kinases in melanocytes (Figure 6d)

These results suggest that the TNF pathway is required for AGP-induced activation of ASK1 signaling and apoptosis pathways.

DISCUSSION

Recent literature reviews show that AGPs are being studied with great interest for cancer treatment (Graves, 2012; Ishaq *et al.*, 2014). It has been shown that AGPs increase intracellular ROS production (Vandamme *et al.*, 2012), induce senescence (Arndt *et al.*, 2013), cause cell cycle arrest (Volotskova *et al.*, 2012), and induce oxidative stress, DNA damage, and apoptosis (Keidar *et al.*, 2011; Iseki *et al.*, 2012; Kalghatgi *et al.*, 2012). A major obstacle in the development of this tool for clinical application as an anticancer therapeutic is the lack of understanding of the mechanisms underlying the intercellular response to AGP exposure. For cancer treatment, knowledge of the specific intracellular signaling pathway factors involved in tumor sensitivity and resistance is critical to successful therapy. This requires exploration of the intracellular mechanisms involved in the

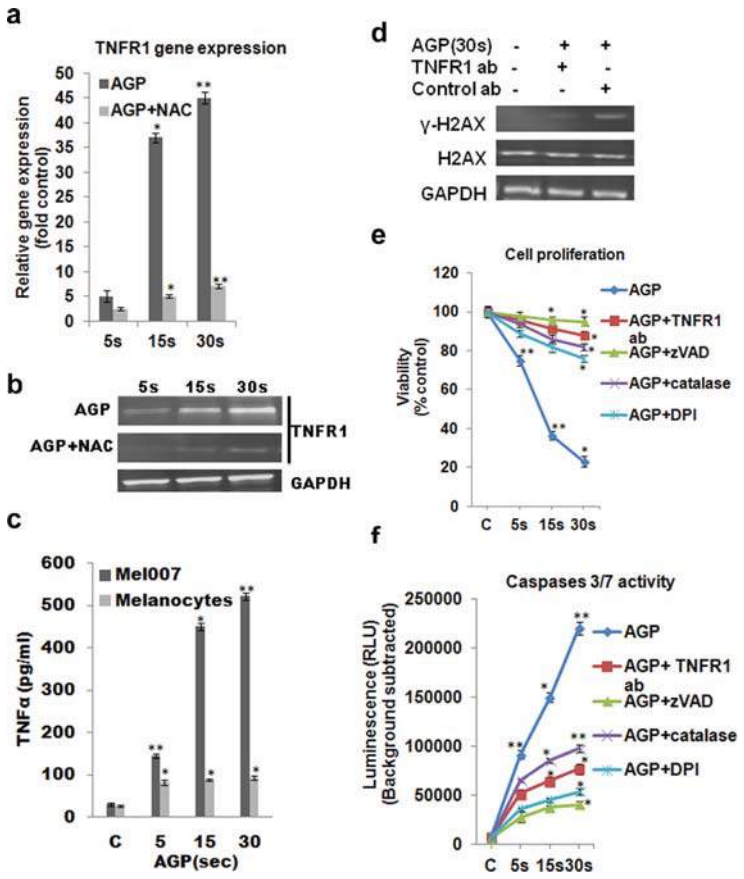


FIGURE 4: TNF-apoptotic pathway is involved for AGP-induced selective apoptosis in cancer cells. (a, b) Mel007 cells were treated with AGP (5, 15, 30 s), and gene expression for TNFR1 was measured by quantitative real-time PCR and Western blotting. (c) Effect of AGP on the activation of TNF ligand in Mel007 and melanocytes was determined by enzyme-linked immunosorbent assay. Cells were treated with AGP (5, 15, 30 s) for 48 h. The concentration of TNFα was measured in cell culture supernatant. (d) Effect of AGP pretreated with or without TNFR1-neutralizing antibody on stress response targets were determined by Western blot analysis of H2AX and γ-H2AX protein in melanoma cells (Mel007). Mouse IgG1 isotype antibody was used as negative control. GAPDH expression was used as loading control. (e, f) Mel007 cells were treated with AGP or pretreated with anti-TNFR1 antibody, caspase inhibitor zVAD, H_2O_2-depleting-agent catalase, and nitric oxide synthesis inhibitor DPI, followed by AGP treatment. Cell viability was measured using a cell titer nonradioactive cell proliferation assay, and caspases 3/7 activity was measured by Caspase-Glo 3/7 assay. In all experiments, control cells were mock treated with He gas flow only. All values are mean ± SD of three independent experiments performed in triplicate. *$p \le 0.01$, **$p \le 0.001$; ANOVA.

interaction of AGP with the localized cellular environment. This knowledge gap motivated our examination of the intracellular pathways involved in the sensitivity to AGP of cancer cell death, as shown in Figure 7.

The AGP-produced ROS, such as H_2O_2 and NO, are the main factors that form a system of oxidizing or nitrosylating species, which activate several signaling pathways. Excessive ROS inside cells results in oxidative stress, which leads to DNA damage and apoptosis by unknown mechanisms (Landino *et al.*, 1996; Kalghatgi *et al.*, 2011; Vandamme *et al.*, 2012). Here we found that AGP selectively induces increased ROS levels in melanoma cells compared with normal control cells, which are blocked by the endogenous ROS scavenger NAC. We also demonstrated that AGP-induced ROS are responsible for selective apoptosis of cancer cells but that this does not occur in normal control cells (Figure 3). This is the first report to show that AGPs induce ROS in cancer cells but not normal cells.

It has been suggested that ROS are involved in apoptosis of cancer cells by the TNF-dependent signaling pathway (Natoli *et al.*, 1997). We also found in our apoptosis-related gene screening study that TNF family members are differentially expressed in AGP-treated melanoma cells. We observed that blocking TNFR1 inhibits the activity of AGP-induced caspases 3/7 and reduces activation of stress response protein H2AX (Figure 4). This indicates the involvement of the TNF-apoptotic pathway in AGP-induced apoptosis in melanoma cells.

Several reports have shown that ROS (H_2O_2) induced by the TNF pathway participates in the apoptosis of cancer cells in ASK1-dependent signaling (Hatai *et al.*, 2000; Liu *et al.*, 2000; Nishitoh *et al.*, 2002; Kamata *et al.*, 2005). ASK1 is required for ROS-induced activation of JNK and p38. We also found that AGPs, similar to H_2O_2, activate the ASK1 pathway, which results in downstream activation of JNK and p38 signaling. The AGP-induced activation of endogenous JNK and p38 is lost by blocking with NAC. Moreover, intracellular depletion of ASK1 by RNAi inhibits AGP-induced caspase 3/7 activity, cell viability, and stress response (Figure 5). These results suggest the involvement of ASK1 in the AGP-induced apoptosis pathway.

Stimulation of the TNF pathway and inducement of ROS production for the activation of ASK1 by AGPs are believed to be involved in this selective cell death process. Although TNF is known to stimulate the production of ROS, and activation of the SAPKs can be partially reversed with free-radical scavengers (Natoli *et al.*, 1997), AGP-induced and TNF-mediated production of ROS had not been demonstrated. We found that AGP-induced ROS production and ASK1 activation was mediated by the TNF pathway. Pretreatment with antioxidant NAC and antagonist anti-TNFR1 antibody attenuated AGP-induced ASK1, p38, and JNK phosphorylation (Figure 6). These results reveal that AGP-induced production of cellular ROS is upstream of the ASK1 and p38/JNK signaling pathway.

Our results provide a plausible mechanisms for AGP-induced apoptosis in cancer cells. Extracellular ROS generated by AGP activates TNF signaling and TNF-dependent intracellular ROS generation. This high intracellular level of ROS activates caspases 3/7 via

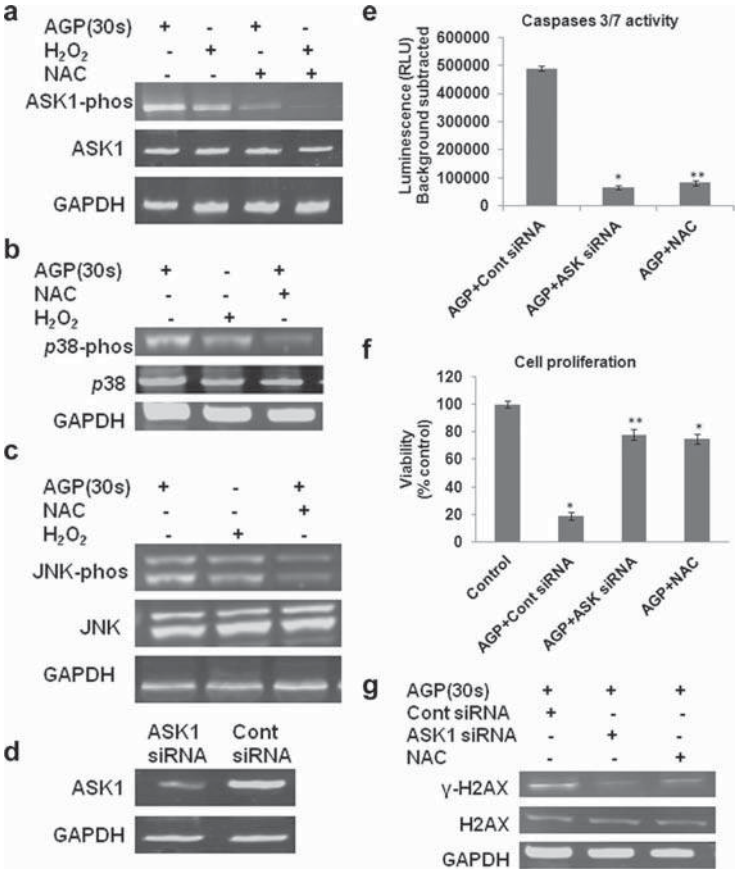


FIGURE 5: AGP stimulates ASK1-mediated p38, and JNK activation is ROS dependent and reversed with NAC. (a–c) Mel007 cancer cells were treated with AGP or H_2O_2 with or without pretreatment with NAC. Western blot analysis was performed on cell lysate using different antibodies as indicated (p-ASK1, S83 phosphospecific antibody for ASK1; p-JNK, phosphospecific antibody for JNK; p-p38, phosphospecific antibody for p-38). GAPDH antibody was used as a loading control. (d) Knockdown of ASK1 leads to decreased activation effect of AGP in melanoma cells. Mel007 cancer cells were transfected with small interfering RNA (siRNA) for ASK1 or control siRNA. The cells were analyzed 36–48 h after transfection to determine ASK1 protein levels. GAPDH was used as protein loading control. (e, f) At 24 h after treatment of ASK1 siRNA or control siRNA, Mel007 cells were treated with AGP (30 s) with or without pretreatment with NAC. Caspase 3/7 activity was measured by Caspase-Glo 3/7 assay, and cell viability was measured by cell titer nonradioactive cell proliferation assay. All values are mean ± SD of three independent experiments performed in triplicate. *$p \leq 0.01$, **$p \leq 0.001$; ANOVA. (g) 2At 4 h posttreatment of ASK1 siRNA or control siRNA, Mel007 cells were treated with AGP (30 s) with or without pretreatment with NAC. The effects of AGP on stress response targets were determined by Western blot analysis of H2AX and γ-H2AX protein levels in cancer cells. GAPDH expression was used as a loading control. In all experiments, control cells were mock treated with He gas flow only.

activation of ASK1 and subsequent p38/JNK pathway and eventually leads to apoptosis (Figure 7).

The precise mechanisms by which AGP increases intracellular ROS production and activate caspase-3–induced apoptosis need to be elucidated. It was reported that AGP induced apoptosis via caspase-3–dependent mechanisms (Kim *et al.*, 2011b; Sensenig *et al.*, 2011), but the upstream mechanisms remained unknown. Our results clearly demonstrate that ASK1 is strongly activated in AGP-treated cancer cells, similar to those treated with H_2O_2,and this may provide a useful model system to understand the AGP-ROS–dependent signaling pathway leading to apoptosis.

In summary, the present work demonstrates that activation of ASK1 through TNF signaling is required in AGP-induced apoptosis. This suggests the importance of TNF signaling and subsequent activation of ASK1 in AGP- and/or ROS-induced programmed cell death of melanoma cells. Further studies are required to elucidate the physiological functions of the AGP-TNF-ASK1 pathway in resistant tumor treatment.

MATERIALS AND METHODS

Atmospheric gas plasma jet device

All cells were treated with atmospheric gas plasma plume, which was generated using a custom-designed atmospheric gas plasma jet device (Figure 1). The device consists of a fused quartz tube equipped with two conducting electrodes. One electrode is a metal wire placed inside the tube along the axis and ~1.8 cm before the nozzle exit. The second electrode is a metal ring attached to the outer wall of the tube near the nozzle exit. The distance between the two electrodes is ~1.0 cm. On helium gas flow (2 l/min) through the quartz tube, a nonequilibrium plasma discharge is produced between the two electrodes by an applied AC high voltage. The high voltage was generated by RF power source coupled with RF voltage amplifier. The discharge voltage and operating frequency were kept at 1.1–1.8 kV and 230–270 kHz, respectively, during cell treatment. The plasma discharge produced at the end of the metal wire electrode was located in two different spatial regions. One part of the discharge was located entirely inside the tube, from the end of the metal wire electrode to the ring electrode, and part of the discharge from the ring electrode propagated along with the gas flow and extended out of the tube through the nozzle as a collimated plasma jet. The length of the plasma jet was >2.5 cm. The end diameter of this collimated plasma jet was ~500 μm and is referred as a cold plasma due to the measured low gas temperature (~35–40°C). During the discharge process, electrical parameters of the plasma, such as discharge voltage and current, were measured using high- and low-voltage probes, respectively. The discharge current between the metal wire and the ring electrodes was 10 mA. However, only 5–10% of this current is likely to extend out of the discharge tube, which was used for cell treatment (Keidar *et al.*, 2013).

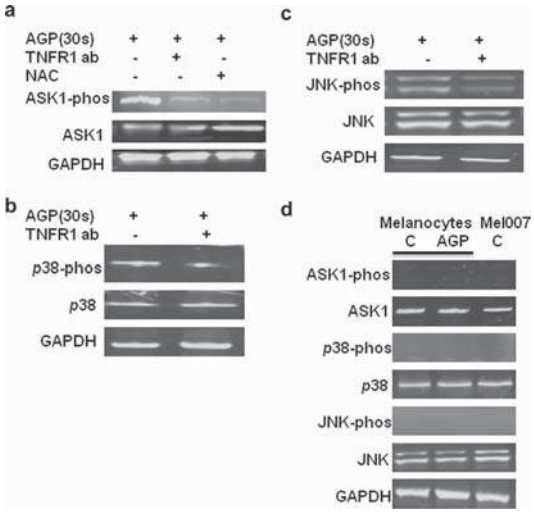


FIGURE 6: TNF signaling involved in the activation of ASK1, p38, and JNK by AGP. (a–c) Mel007 cancer cells were treated with AGP with or without pretreatment with TNFR1- neutralizing antibody or NAC. Western blot analysis was performed on cell lysate using different antibodies as indicated (p-ASK1, S83 phosphospecific antibody for ASK1; p-JNK, phosphospecific antibody for JNK; p-p38, phosphospecific antibody for p-38). GAPDH antibody was used as a loading control. (d) The basal protein expression levels for ASK1, p38, JNK, ASK1-phos, p38-phos, and JNK-phos were quantified in normal melanocytes and Mel007 cells by Western blot analysis. GAPDH antibody was used as a loading control.

Cell culture and reagents

The human melanoma cell lines Mel-RM, Mel-007, and Mel-JD and control human epidermal melanocytes were a kind gift of Peter Hershey, Melanoma Institute, University of Sydney, Sydney, Australia (Jiang *et al.*, 2010), and were maintained in DMEM (Invitrogen) plus 10% fetal bovine serum (FBS). Melanocytes were grown in special melanocyte growth medium 254 plus with human melanocyte growth supplement (S-002-05; Invitrogen, Grand Island, NY). Human fetal lung fibroblasts (MRC5) were purchased from the American Type Culture Collection (ATCC, Manassas, VA) and grown in MEM plus 10% FBS. H_2O_2 (216763), caspase inhibitor Z-VAD-FMK (G7231), NAC (antioxidant and ROS scavenger; A0737), DPI (inhibitor of nitric oxide synthetase; D2926), and catalase (H_2O_2 depleter; C-30) were purchased from Sigma-Aldrich (St. Loius, MO). TNFR1-neutralizing antibody (Defer *et al.*, 2007) and mouse immunoglobulin (Ig) G1 isotype control antibody were purchased from R&D Systems (Minneapolis, MN).

Cell viability assay

All cells were cultured in 96-well plates at 2×10^4 cells/well overnight, treated with AGP for indicated time periods, and incubated for 18–24 h. Cell viability was measured using CellTiter 96 Aqueous Non-Radioactive Cell Proliferation (MTS) Assay (G5421; Promega, Madison, WI) following the manufacturer's protocol. In some experiments the caspase inhibitor zVAD-FMK (50 μM), NAC (3 mM), TNFR1-neutralizing antibody (50 μg/ml), catalase (2000 U/ml), or DPI (10 μM) was added 1–2 h before AGP treatment. Assays were performed in triplicate.

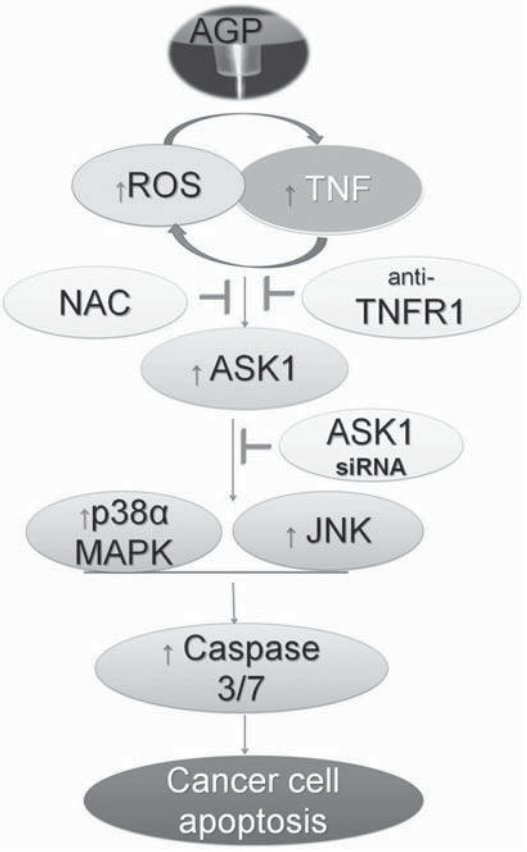


FIGURE 7: Model for the regulation of TNF-receptor–mediated apoptosis pathway induced by AGP. The induced production of ROS and TNF pathway activation by cross-talk (direct or indirect interaction), which leads to activation of ASK1 and stimulation of downstream p38α MAPK or JNK to induce caspase-3/7–dependent cancer cell apoptosis.

ROS detection assay

Cells were cultured in 96-well plates at 2×10^4 cells/well overnight and treated with AGP for the indicated time periods. ROS generation was detected with CM-H2DCFDA (C6827; Invitrogen, Grand Island, NY) following the manufacturer's instructions. Briefly, 1–2 h after AGP treatment, cells were incubated with 10 μM of CM-H2DCFDA for 30 min at 37°C in the dark. ROS were measured by a FLUOstar Omega fluorescence plate reader at excitation and emission wavelengths of 485 and 530 nm, respectively. Assays were performed in triplicate.

Caspase 3/7 activity assay

Cells were plated in 96-well plates at 2×10^4 cells/well, incubated for 24 h, and then treated with AGP for the indicated time periods. After 4–8 h of AGP treatment, caspase 3/7 activity was quantified by adding Caspase Glo 3/7 reagent (G8091; Promega). In some experiments the caspase inhibitor zVAD-FMK (50 μM), NAC (3 mM), TNFR1-neutralizing antibody (50 μg/ml), catalase (2000 U/ml),

or DPI (10 μM) was added 1–2 h before AGP treatment. Triplicate samples were run in standard 96-well plates to quantify the apoptotic response. Luminescence values were then determined using a LUMIstar Omega plate reader.

Total cellular glutathione assay

Cells were cultured in six-well plates overnight and treated with AGP, in some cases after pretreatment with NAC, for the indicated time periods. A total of 1×10^6 cells were collected and lysed in 100 μl of ice-cold lysis buffer for 10 min. At 1–2 h after the AGP treatment, supernatant was collected by centrifugation of lysate for a glutathione assay performed with a glutathione (GSH/GSSG/Total) assay kit (K264-100; BioVision Research Products, Milpitas, CA) following the manufacturer's instructions. The total amount of GSH was measured using a FLUOstar Omega fluorescence plate reader at excitation and emission wavelengths of 340 and 420 nm, respectively. Trypan blue stain was used in parallel, and the viable cells were counted.

Nitric oxide measurement

Cells were cultures in 96-well plates at 2×10^4 cells/well, incubated for 24 h, treated with AGP, and in some cases pretreated with NAC (3 mM), for the indicated time periods. Nitric oxide production was measured with 4-amino-5-methylamino-2′,7′-difluofluorescein diacetate (DAF-FM diacetate) following the manufacturer's protocols (D23844; Invitrogen). At 1–2 h after AGP treatment, cells were labeled with 10 μM DAF-FM diacetate for 30 min at 37°C. Nitric oxide radical was measured using a FLUOstar Omega fluorescence plate reader at excitation and emission wavelengths of 485 and 520 nm, respectively.

Quantitative real-time PCR

Cells were cultured in six-well plates at 5×10^5 cells/well overnight and treated with AGP for the specified time periods. At 18–24 h after AGP treatment, total cell RNA was extracted using TRIzol reagent (15596-026; Invitrogen) and reverse transcribed into single-stranded cDNA as explained previously (Ishaq *et al.*, 2008). Quantitative relative gene expression was determined by using SsoAdvanced SYBR Green SuperMix (172-5265; Bio-Rad, Hercules, CA) with a Roche LightCycler 480 qPCR system. Data were analyzed with LightCycler 480 software. The human apoptosis PCR array primers library (HPA-1; RealTimePCR.com, Elkins Park, PA) was used to quantify relative gene expression of >90 genes involved in intracellular prosurvival and proapoptotic pathways. glyceraldehyde-3-phosphate dehydrogenase (GAPDH), actin-b, GUS-b, B2M, HPRT1, PPIA, and RPL13A were used as internal controls to normalize data. Relative quantification of gene expression after AGP treatment was obtained by using the ΔC_T method compared with fold change of untreated (He gas flow only) control cells.

Enzyme-linked immunosorbent assay

Cells were cultured in six-well plates and either untreated or treated with AGP for the specified time period. After 48 h of treatment, cell culture supernatants were harvested, and the concentration of TNFα was measured by using the TNFα human enzyme-linked immunosorbent assay kit (ab46087) from Abcam (Cambridge, United Kingdom) as per the manufacturer's instructions.

RNA interference and Western blot

Cells grown to a density of ~70% confluence with antibiotic-free media were transfected with the ASK1 short hairpin RNA (shRNA; sc-29748; Santa Cruz Biotechnology, Santa Cruz, CA) and negative control shRNA expression vector by Lipofectamine (Invitrogen) according to the manufacturers' instructions. After 24 h of shRNA transfection, cells were treated with AGP. Cells were harvested and lysed in RIPA lysis and extraction buffer (89901), halt protease inhibitor cocktail (87786), and halt phosphatase inhibitor cocktail (78420; Thermo Scientific, Pittsburgh, PA) by incubating on ice for 30 min, as described previously (Ishaq *et al.*, 2009). Protein concentration was determined by BCA Protein Assay Kit (23227; Thermo Scientific). After adding 2× SDS loading buffer, the samples were subjected to SDS–PAGE (MP TGX 4–20%; 4561094; Bio-Rad). Protein was then transferred onto immunoblot polyvinylidene fluoride membrane (Millipore, Billerica, MA) and probed with the primary antibodies as specified and horseradish peroxidase (HRP)–conjugated secondary antibodies. The bonded proteins were visualized with a chemiluminescence detection kit (Clarity Western ECL substrate; Bio-Rad, 170-5060) using ImageQuant LAS400 (GE Technology, Little Chalfon, United Kingdom). Antibodies used included phospho-p38 MAPK (Thr-180/Tyr-182; 4511), p38 MAPK (8690), phospho-SAPK/JNK (Thr-183/Tyr-185; 4668), SAPK/JNK (9258), phospho-ASK1 (Thr-845; 3765), ASK1 (8662), H2AX (2595), phospho-histone H2A.X (Ser-139; 2577) and GAPDH from Cell Signaling Technology (Danvers, MA) and goat anti-rabbit IgG HRP (170-6515) and goat anti-mouse IgG HRP (170-6516) from Bio-Rad. In some experiments the antioxidant NAC (3 mM), TNFR1-neutralizing antibody (50 μg/ml), or mouse IgG1 antibody as negative isotype control were added 1–2 h before AGP treatment.

Statistical analysis

All values are presented as mean ± SD of three independent experiments. Statistical differences between controls and treated groups were determined by one-way analysis of variance (ANOVA) or Student's *t* test where applicable. Differences were considered statistically significant for $p \leq 0.05$ ($^*p \leq 0.01$, $^{**}p \leq 0.001$).

ACKNOWLEDGMENTS

We thank Peter Hershey for providing melanoma cell lines and melanocytes. We also thank Penny Bean, Vijay Vaithilingam, Gail Wertz, Denise Lewy, Bernie Tuch, Peter Molloy, Sumeet Bal, Glenn Brown, and Thu Ho for useful discussions and support for the experimental work and equipment. This study was supported by the CSIRO OCE postdoctoral fellowship program, the CSIRO Science Leadership program, CSIRO Transformational Biology and Advanced Materials Transformational Capability Platforms, and the Australian Research Council.

REFERENCES

Arndt S, Wacker E, Li YF, Shimizu T, Thomas HM, Morfill GE, Karrer S, Zimmermann JL, Bosserhoff AK (2013). Cold atmospheric plasma, a new strategy to induce senescence in melanoma cells. Exp Dermatol 22, 284–289.

Defer N, Azroyan A, Pecker F, Pavoine C (2007). TNFR1 and TNFR2 signaling interplay in cardiac myocytes. J Biol Chem 282, 35564–35573.

Estrela JM, Ortega A, Obrador E (2006). Glutathione in cancer biology and therapy. Crit Rev Clin Lab Sci 43, 143–181.

Graves DB (2012). The emerging role of reactive oxygen and nitrogen species in redox biology and some implications for plasma applications to medicine and biology. J Phys D Appl Phys 45, 263001–263042.

Hatai T, Matsuzawa A, Inoshita S, Mochida Y, Kuroda T, Sakamaki K, Kuida K, Yonehara S, Ichijo H, Takeda K (2000). Execution of apoptosis signal-regulating kinase 1 (ASK1)-induced apoptosis by the mitochondria-dependent caspase activation. J Biol Chem 275, 26576–26581.

Huang J, Li H, Chen W, Lv G-H, Wang X-Q, Zhang G-P, Ostrikov K, Wang P-Y, Yang S-Z (2011). Dielectric barrier discharge plasma in Ar/O_2 promoting apoptosis behavior in A549 cancer cells. Appl Phys Lett 99, 253701.
Ichijo H, Nishida E, Irie K, tenDijke P, Saitoh M, Moriguchi T, Takagi M, Matsumoto K, Miyazono K, Gotoh Y (1997). Induction of apoptosis by ASK1, a mammalian MAPKKK that activates SAPK/JNK and p38 signaling pathways. Science 275, 90–94.
Iseki S, Nakamura K, Hayashi M, Tanaka H, Kondo H, Kajiyama H, Kano H, Kikkawa F, Hori M (2012). Selective killing of ovarian cancer cells through induction of apoptosis by nonequilibrium atmospheric pressure plasma. Appl Phys Lett 100, 113702.
Ishaq M, Evans MM, Ostrikov KK (2014). Effect of atmospheric gas plasmas on cancer cell signaling. Int J Cancer 134, 1517–1528.
Ishaq M, Hu J, Wu X, Fu Q, Yang Y, Liu Q, Guo D (2008). Knockdown of cellular RNA helicase DDX3 by short hairpin RNAs suppresses HIV-1 viral replication without inducing apoptosis. Mol Biotechnol 39, 231–238.
Ishaq M, Ma L, Wu X, Mu Y, Pan J, Hu J, Hu T, Fu Q, Guo D (2009). The DEAD-box RNA helicase DDX1 interacts with RelA and enhances nuclear factor kappaB-mediated transcription. J Cell Biochem 106, 296–305.
Jiang CC *et al.* (2010). Apoptosis of human melanoma cells induced by inhibition of B-RAF(V600E) involves preferential splicing of bim(S). Cell Death Dis 1, e69.
Kaelin WG (1999). Choosing anticancer drug targets in the postgenomic era. J Clin Invest 104, 1503–1506.
Kalghatgi S, Fridman A, Azizkhan-Clifford J, Friedman G (2012). DNA damage in mammalian cells by non-thermal atmospheric pressure microsecond pulsed dielectric barrier discharge plasma is not mediated by ozone. Plasma Processes Polymers 9, 726–732.
Kalghatgi S, Kelly CM, Cerchar E, Torabi B, Alekseev O, Fridman A, Friedman G, Azizkhan-Clifford J (2011). Effects of non-thermal plasma on mammalian cells. PLoS One 6, e16270.
Kamata H, Honda S, Maeda S, Chang LF, Hirata H, Karin M (2005). Reactive oxygen species promote TNF alpha-induced death and sustained JNK activation by inhibiting MAP kinase phosphatases. Cell 120, 649–661.
Keidar M, Shashurin A, Volotskova O, Stepp MA, Srinivasan P, Sandler A, Trink B (2013). Cold atmospheric plasma in cancer therapy. Phys Plasmas 20, 057101.
Keidar M, Walk R, Shashurin A, Srinivasan P, Sandler A, Dasgupta S, Ravi R, Guerrero-Preston R, Trink B (2011). Cold plasma selectivity and the possibility of a paradigm shift in cancer therapy. Br J Cancer 105, 1295–1301.
Kim K, Choi JD, Hong YC, Kim G, Noh EJ, Lee J-S, Yang SS (2011b). Atmospheric-pressure plasma-jet from micronozzle array and its biological effects on living cells for cancer therapy. Appl Phys Lett 98, 073701.
Kim JY, Wei Y, Li J, Foy P, Hawkins T, Ballato J, Kim S-O (2011a). Single-cell-level microplasma cancer therapy. Small 7, 2291–2295.
Koeritzer J *et al.* (2013). Restoration of sensitivity in chemo-resistant glioma cells by cold atmospheric plasma. PLoS One 8, e64498.
Landino LM, Crews BC, Timmons MD, Morrow JD, Marnett LJ (1996). Peroxynitrite, the coupling product of nitric oxide and superoxide, activates prostaglandin biosynthesis. Proc Natl Acad Sci USA 93, 15069–15074.
Liu DX, Bruggeman P, Iza F, Rong MZ, Kong MG (2010). Global model of low-temperature atmospheric-pressure He + H2O plasmas. Plasma Sources Sci Technol 19, 025018.
Liu H, Nishitoh H, Ichijo H, Kyriakis JM (2000). Activation of apoptosis signal-regulating kinase 1 (ASK1) by tumor necrosis factor receptor-associated factor 2 requires prior dissociation of the ASK1 inhibitor thioredoxin. Mol Cell Biol 20, 2198–2208.
Micheau O, Tschopp J (2003). Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. Cell 114, 181–190.
Moh MC, Lorenzini PA, Gullo C, Schwarz H (2013). Tumor necrosis factor receptor 1 associates with CD137 ligand and mediates its reverse signaling. FASEB J 27, 2957–2966.
Nadeau PJ, Charette SJ, Landry J (2009). REDOX reaction at ASK1-Cys250 is essential for activation of JNK and induction of apoptosis. Mol Biol Cell 20, 3628–3637.
Nagata S (1997). Apoptosis by death factor. Cell 88, 355–365.
Natoli G, Costanzo A, Ianni A, Templeton DJ, Woodgett JR, Balsano C, Levrero M (1997). Activation of SAPK/JNK by TNF receptor 1 through a noncytotoxic TRAF2-dependent pathway. Science 275, 200–203.
Nishitoh H, Matsuzawa A, Tobiume K, Saegusa K, Takeda K, Inoue K, Hori S, Kakizuka A, Ichijo H (2002). ASK1 is essential for endoplasmic reticulum stress-induced neuronal cell death triggered by expanded polyglutamine repeats. Gene Dev 16, 1345–1355.
Nishitoh H, Saitoh M, Mochida Y, Takeda K, Nakano H, Rothe M, Miyazono K, Ichijo H (1998). ASK1 is essential for JNK/SAPK activation by TRAF2. Mol Cell 2, 389–395.
Noguchi T, Ishii K, Fukutomi H, Naguro I, Matsuzawa A, Takeda K, Ichijo H (2008). Requirement of reactive oxygen species-dependent activation of ASK1-p38 MAPK pathway for extracellular ATP-induced apoptosis in macrophage. J Biol Chem 283, 7657–7665.
Noguchi T, Takeda K, Matsuzawa A, Saegusa K, Nakano H, Gohda J, Inoue J, Ichijo H (2005). Recruitment of tumor necrosis factor receptor-associated factor family proteins to apoptosis signal-regulating kinase 1 signalosome is essential for oxidative stress-induced cell death. J Biol Chem 280, 37033–37040.
Ostrikov K, Neyts EC, Meyyappan M (2013). Plasma nanoscience: from nano-solids in plasmas to nano-plasmas in solids. Adv Phys 62, 113–224.
Partecke LI *et al.* (2012). Tissue tolerable plasma (TTP) induces apoptosis in pancreatic cancer cells in vitro and in vivo. BMC Cancer 12, 473.
Saitoh M, Nishitoh H, Fujii M, Takeda K, Tobiume K, Sawada Y, Kawabata M, Miyazono K, Ichijo H (1998). Mammalian thioredoxin is a direct inhibitor of apoptosis signal-regulating kinase (ASK) 1. EMBO J 17, 2596–2606.
Sensenig R *et al.* (2011). Non-thermal plasma induces apoptosis in melanoma cells via production of intracellular reactive oxygen species. Ann Biomed Eng 39, 674–687.
Sensenig R, Kalghatgi S, Goldstein A, Friedman G, Friedman G, Brooks A (2008). Induction of apoptosis in melanoma cells by non-thermal atmospheric plasma discharge. Ann Surg Oncol 15, 65–65.
Thiyagarajan M, Waldbeser L, Whitmill A (2012). THP-1 leukemia cancer treatment using a portable plasma device. Stud Health Technol Inform 173, 515–517.
Thornton TM, Rincon M (2009). Non-classical p38 map kinase functions: cell cycle checkpoints and survival. Int J Biol Sci 5, 44–51.
Vandamme M *et al.* (2012). ROS implication in a new antitumor strategy based on non-thermal plasma. Int J Cancer 130, 2185–2194.
Vandamme M, Robert E, Pesnel S, Barbosa E, Dozias S, Sobilo J, Lerondel S, Le Pape A, Pouvesle JM (2010). Antitumor effect of plasma treatment on U87 glioma xenografts: preliminary results. Plasma Processes Polymers 7, 264–273.
Volotskova O, Hawley TS, Stepp MA, Keidar M (2012). Targeting the cancer cell cycle by cold atmospheric plasma. Sci Rep 2, 636–610.
Walk RM, Snyder JA, Srinivasan P, Kirsch J, Diaz SO, Blanco FC, Shashurin A, Keidar M, Sandler AD (2013). Cold atmospheric plasma for the ablative treatment of neuroblastoma. J Pediatr Surg 48, 67–73.
Weltmann KD, Kindel E, von Woedtke T, Hahnel M, Stieber M, Brandenburg R (2010). Atmospheric-pressure plasma sources: prospective tools for plasma medicine. Pure Appl Chem 82, 1223–1237.
Winter J, Wende K, Masur K, Iseni S, Dunnbier M, Hammer MU, Tresp H, Weltmann KD, Reuter S (2013). Feed gas humidity: a vital parameter affecting a cold atmospheric-pressure plasma jet and plasma-treated human skin cells. J Phys D Appl Phys 46, 295401.
Yan X, Xiong Z, Zou F, Zhao S, Lu X, Yang G, He G, Ostrikov K (2012). Plasma-induced death of HepG2 cancer cells: intracellular effects of reactive species. Plasma Processes Polymers 9, 59–66.

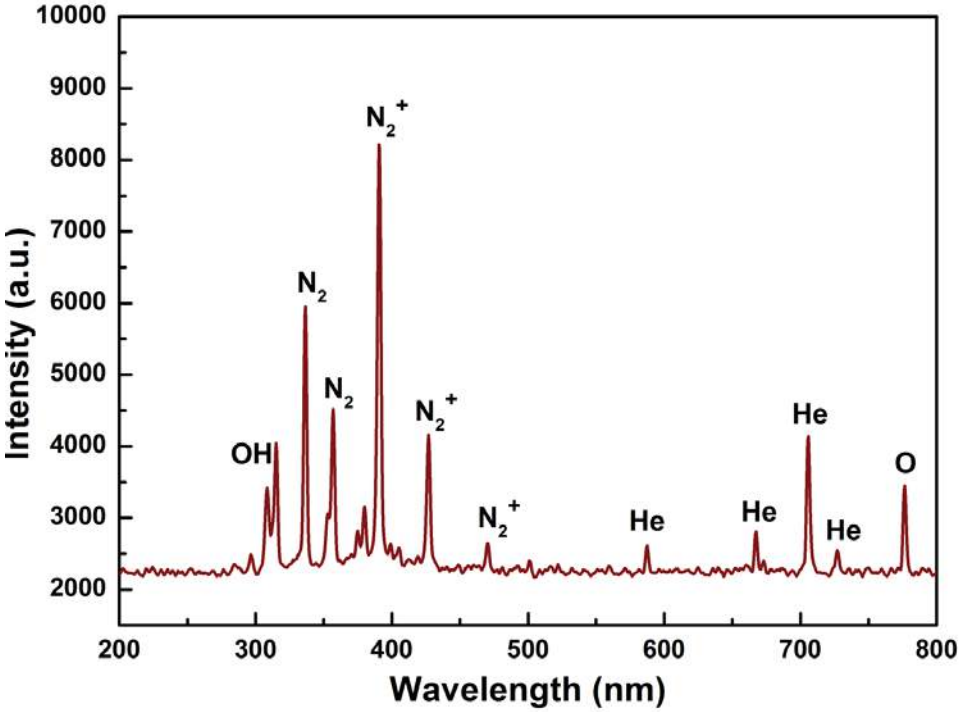
Supplementary data

Supplementary Methods

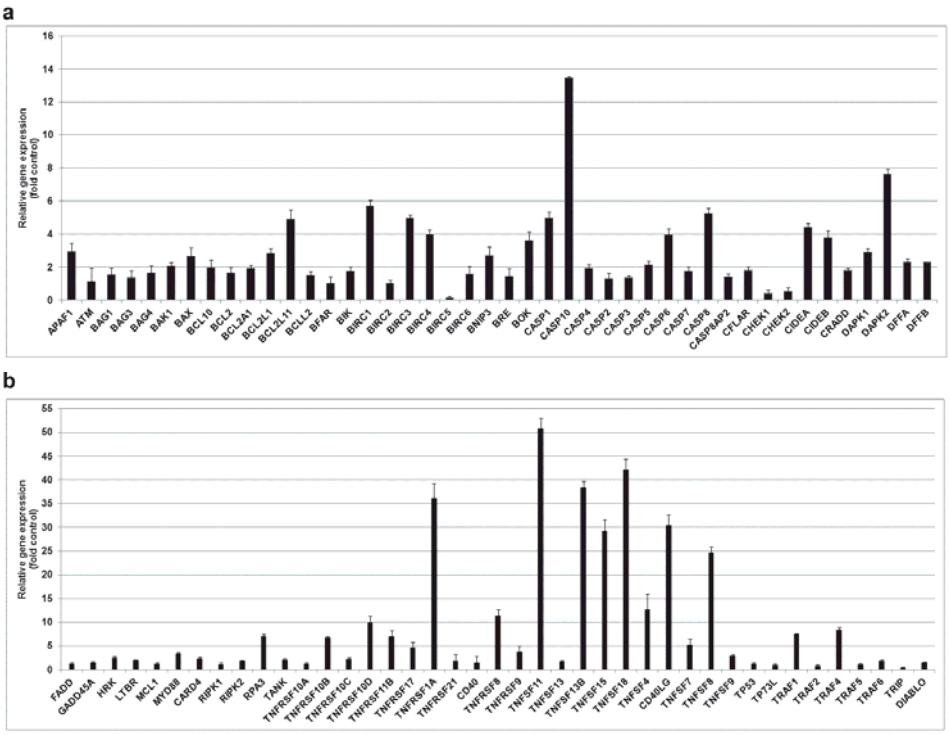
Identification of reactive species

The reactive species produced by the plasma jet were identified by collecting the emission spectra using a high-resolution emission spectroscopy (Princeton Instruments, Acton SP 2500, equipped with Pixis 256 CCD detector). The spectra (one of them is shown in Sup. Fig. 2) were collected during the cells treatment at the applied discharge voltage, Helium flow rate and RF frequency of 1.8kV, 2 L/min and 230 kHz, respectively. The wavelength calibration was performed with Hg and Ne/Ar light source prior to the collection of spectra. The spectroscope with the grating (1200 g/mm) and slit width (100 μm) were used in our experiments. Spectrum in Sup. Fig. 1 reveals that the excited species such as O, OH, He, N_2 and N_2^+ were generated in the plasma jet.

Supplementary Figures



Supplementary Fig. 1. An optical emission spectrum of the plasma jet in the wavelength range of 200-800 nm.



Supplementary Fig. 2. TNF signaling pathway family members up-regulated after AGP treatment. (a, b) Apoptosis-related intracellular differential gene expression pattern was quantified by using >90 different genes involved in apoptosis. The Mel007 cells were treated with AGP (30 sec) and relative gene expression was measured by quantitative real-time PCR after 24 hrs following the AGP treatment. Gene expression was normalized by internal genes (Actin-b, GAPDH and GUSB) and quantified as fold change compared to He gas-treated control cells. All values are mean ± s.d. of three independent experiments.

Table-1: Intracellular glutathione content of melanocytes and melanoma cells

	Melanocytes				Mel007			
	Control	AGP			Control	AGP		
		5s	15s	30s		5s	15s	30s
GSH (% of control)	100	93.12 ± 1.1	96.1 ± .23	92.8 ± 1.8	100	67.2 ± .23*	45 ± 1.8**	18.93 ± 2*
% Viability	91.2 ± .93	88.34 ± 2.2	90.1 ± 2.33	86.6 ± 3.1	94.2 ± 1.3	70.2 ± 1.2*	42.1 ± 1.2*	17.51± 2.4*
Total viable cells (10^5)	8.87 ± 2.0	8.67 ± 1.61	9.03 ± 1.53	8.03 ± 2.41	9.12 ± 2.02	6.8 ± 3.14*	4.36 ± .69*	2.17 ± 1.56*

Melanocytes and melanoma (Mel007) cells were treated with AGP (5, 15 and 30 sec). Total glutathione content was determined as described in materials and methods. The viable cells were counted by staining with trypan blue. The results represent the means ± s.d. of three independent experiments performed in triplicate. *$p \leq 0.01$, **$p \leq 0.001$; ANOVA.

APPENDIX 4: Identification of the rainbowfish in Lake Eacham using DNA sequencing

The Lake Eacham rainbowfish (*Melanotaenia eachamensis*) was once thought to be confined to its type locality within the Lake Eacham World Heritage National Park. *M. eachamensis* disappeared from the lake following the translocation of several species into the lake and the species was pronounced extinct in the wild in 1987.

In a 2007 survey we noticed that rainbowfish were present in the lake once again. We used a molecular marker to identify these fish and the likely source population. Analysis of the D-loop region of mitochondrial DNA revealed that the species now present in the lake is *Melanotaenia splendida*, and is most closely related to several *M. splendida* populations in the immediate vicinity.

Detailed contributions:

	Varinli H	**Co-authors**
Laboratory Experiments	80%	20%
Data Analysis	25%	75%
Data Interpretation	25%	75%
Writing	10%	90%

Brown, C., Aksoy, Y., Varinli, H. & Gillings, M. (2013). Identification of the rainbowfish in Lake Eacham using DNA sequencing. *Australian Journal of Zoology*, 60(5), p. 334-339.

DOI: 10.1071/ZO12106