
# Exploring Molecular Links between Obesity and Breast Cancer

*Author:*
Rebekah CRAKE
(ID: 9854370)

*Supervisors:*
Dr. Logan WALKER
Dr. Margaret CURRIE
Dr. Elisabeth PHILLIPS

October 23, 2015

Bachelor of Biomedical Science Honours Degree

# Abstract

Obesity is associated with a high risk of incidence of, and mortality for, postmenopausal breast cancer. Despite this well-established link, the molecular and mechanistic basis of the obesity and breast cancer association still remains unclear. In obesity research, genetic variation due to copy number differences has become increasingly popular. The salivary amylase gene, *AMY1*, is well-known for its extensive copy number variation (CNV) in the human genome and has previously been correlated with a genetic predisposition toward obesity; however, research surrounding this association is controversial. Despite an established relationship between obesity and breast cancer risk, the recently reported genetic association between *AMY1* CNV and obesity has not yet been examined in normal and obese breast cancer patients. Furthermore, gene expression changes in breast tumours from obese women remain poorly characterised. We hypothesise that obese breast cancer patients are associated with (1) low *AMY1* copy number and (2) differential expression of candidate genes in the breast tumour.

This study included 55 post-menopausal breast cancer patients from The Cancer Society Tissue Bank, with a BMI (body mass index)> 30 (obese; n=28) or BMI < 25 (healthy; n=27). Quantitative PCR (qPCR) assessment of germline *AMY1* copy number status from blood showed that obese breast cancer patients have a lower average copy number of *AMY1* compared to normal weight patients. Examining breast tumour expression profiles of obese and non-obese patients from two published studies, identified four candidate genes (*GRIA2, DUSP4, NR2F1,* and *ADH1B*) shared between both studies. Analysis of gene expression data from The Cancer Genome Atlas (TCGA) indicated that these four genes are differentially expressed within clinically relevant breast tumour subtypes characterised by oestrogen receptor, progesterone receptor and HER2 status. qPCR analysis of each candidate gene within our study cohort showed that the average expression of *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* was lower in obese compared to healthy breast tumours, but these results were not statistically significant. My study indicated that

BMI may be associated with lower germline copy number of *AMY1* in post-menopausal breast cancer patients; however, further work with a larger cohort is needed to establish if *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* are associated with obesity related breast cancer.

# Acknowledgements

I would like to take the time to thank all the people that have supported me and my research throughout this *sigh* difficult year.

Firstly, to my supervisors Dr Logan Walker, Dr Margaret Currie and Dr Elisabeth Phillips. My primary supervisor Dr Logan Walker, for always offering his time, expertise, consistent encouragement and honest feedback. You have been a fantastic supervisor, who has developed my skills and knowledge as a scientist and I cannot thank you enough. Also to Dr Margaret Currie, thank you for offering professional advice both on my experimental and written research: it has been invaluable. Finally, thank you to Dr Elisabeth Phillips, for her everyday assistance in all the small but very meaningful details. Your bubbly personality made lab life such a pleasant experience as you are such a wonderful and caring person, whom I am glad to have had the pleasure of being one of my supervisors.

I would also like to say a massive, big, huge thank you to George Wiggins. Not only have you had to tolerate sitting beside me all year, and therefore listen to my incessant amount of sighing, you have continually supported, guided, encouraged and devoted endless hours to help me with my research. Honestly, this year would not have been the same without your happy-go-lucky self around.

Along with my supervisors and George, I want to acknowledge the rest of the Makenzie Cancer Research Group(ies) who have in some way influenced this year for me: Assoc. Prof. Gabi Dachs, Helen, (Dr) Lizzy, Vanessa, Cristina and Arthur. I have thoroughly enjoyed being a part of this group; each and every one of you has made an effort to get to know me and were always willing to help when it was needed. You are all exceptional, as scientists and as people; Thank You.

Helen Morrin (The Cancer Society Tissue Bank), I would also like to thank you for organising and providing the samples for my project, I truly appreciate it. Additionally, I would also like to acknowledge The Cancer Society for funding my project.

To the other Honours students who have endured this year with me, Millie, Sarah, Morgan, Anishah, Mark and Melanie. Thank you for your support throughout

the year's challenges, and good luck for both your thesis and all your future endeavours.

Obviously the BBiomedSci Honours course would not run this smoothly and be the success that it is without a few very important people, so thank you to the course co-ordinators Prof. Madhav Bhatia, Prof. Mark Hampton and Prof. Martin Kennedy. Also to Alice and Linda, our go to ladies.

Finally, just a few wee personal acknowledgments to some people who deserve a special mention and they are my amazing and always supportive parents (David and Julie), older brother (Matthew), and partner (Tian). With the encouragement and love I have from you I have been able to truly value this year's experience, without having to worry about other life stresses. It's something that I will never forget, Thank you.

# List of Abbreviations

| Abbreviation | Description |
|---|---|
| BCA | bicinchoninic acid |
| BSA | bovine serum albumin |
| BMI | body mass index |
| cDNA | complementary deoxyribonucleic acid |
| Chr | chromosome |
| CNV/s | copy number variation/s |
| CSC | cancer stem cell-like |
| CSTB | Cancer Society Tissue Bank |
| Ct | cycle threshold |
| DEPC | diethylpyrocarbonate |
| DNA | deoxyribonucleic acid |
| dNTP/s | deoxynucleotide triphosphate/s |
| DTT | dithiothreitol |
| E.coli | Escherichia coli |
| ELISA | enzyme-linked immunosorbent assay |
| ER | oestrogen receptor |
| ER- | oestrogen receptor negative |
| ER+ | oestrogen receptor positive |
| FFA | free fatty acids |
| GWAS | genome wide association studies |
| HapMap | haplotype map |
| HER2 | human epidermal growth factor receptor 2 |
| ID | identification |
| IDC | invasive ductal carcinoma |
| IDT | Integrated DNA Technologies |
| Kg | kilogram |
| Mg | microgram |
| $MgCl_2$ | magnesium chloride |
| min | minute |
| mL | millilitre |
| mM | millimolar |
| mRNA | messenger ribonucleic acid |
| n | sample size |
| NCBI | National Center for Biotechnology Information |
| ng | nanogram |

| | |
|---|---|
| nm | nanometre |
| PBS | phosphate-buffered saline |
| PCR | polymerase chain reaction |
| PR | progesterone receptor |
| PR- | progesterone receptor negative |
| PR+ | progesterone receptor positive |
| PVDF | polyvinylidene fluoride |
| qPCR | quantitative polymerase chain reaction |
| rcf | relative centrifugal force |
| RNA | ribonucleic acid |
| rpm | rotations per minute |
| RT | reverse transcription |
| RT-PCR | reverse transcription polymerase chain reaction |
| SDS | sodium dodecyl sulphate |
| sec | second |
| SNP/s | single nucleotide polymorphisms |
| TBS-T | tris-buffered saline and tween 20 |
| TCGA | The Cancer Genome Atlas |
| U | units |
| ™ | Trademark |
| WHO | World Health Organisation |
| ºC | degree Celsius |
| Δ | delta |
| µL | microliter |
| µM | micrometre |
| % | percentage |
| ® | registered trademark |
| 3D | 3-dimensional |

# List of Genes

| Gene | Gene Name |
|---|---|
| *ADH1B* | Alcohol Dehydrogenase 1B (Class I), Beta Polypeptide |
| *AMY1 (AMY1A, AMY1B, AMY1C)* | Salivary Amylase |
| *AMY2 (AMY2A, AMY2B)* | Pancreatic Amylase |
| *BDNF* | Brain-Derived Neurotrophic Factor |
| *COUP-TFI* | COUP Transcription Factor 1 |
| *DUSP4* | Dual Specificity Phosphatase 4 |
| *EEF1A1* | Eukaryotic Translation Elongation Factor 1 Alpha 1 |
| *ERBB2/HER2* | Erb-B2 Receptor Tyrosine Kinase 2 |
| ERK1 | Extracellular signal-regulated kinase 1 |
| ERK2 | Extracellular signal-regulated kinase 2 |
| *ESR1* | Oestrogen Receptor 1 |
| *FTO* | Fat Mass and Obesity Associated |
| *GALNT10* | Polypeptide N-Acetylgalactosaminyltransferase 10 |
| *GIPR* | Gastric Inhibitory Polypeptide Receptor |
| *GRIA2* | Glutamate Receptor, Ionotropic, AMPA 2 |
| *IGFBP* | Insulin-like Growth Factor-Binding Protein |
| *IGF-1* | Insulin-like Growth Factor 1 |
| JNK | C-Jun N-Terminal Kinase 1 |
| *LEP* | Leptin |
| *LEPR* | Leptin Receptor |
| MAPK | Mitogen-activated protein kinases |
| *MC4R* | Melanocortin 4 Receptor |
| *NR2F1* | Nuclear Receptor Subfamily 2, Group F, Member 1 |
| *NTRK2* | Neurotrophic Tyrosine Kinase, Receptor, Type 2 |
| *OLFM4* | Olfactomedin 4 |
| *PC1* | Prohormone Convertase 1 |
| RNase H | Ribonuclease H |
| RNase P | Ribonuclease P |
| *RPPH1* | Ribonuclease P RNA Component H1 |
| *SIM1* | Single-Minded Family BHLH Transcription Factor 1 |
| VEGF | Vascular Endothelial Growth Factor |
| β-actin | Beta-actin |

# List of Figures

# List of Tables

# Contents

# Chapter 1

# 1 Introduction

The prevalence of obesity is increasing worldwide at an alarming rate[1]. Increased weight gain and obesity are correlated with a number of adverse health effects including increased risk of developing a number of different cancers[2]. Obesity is associated with an increased risk and reduced survival rate of breast cancer in women worldwide, particularly in post-menopausal women[3, 4]. Despite the well-established association between obesity and increased risk of breast cancer, the precise nature in which obesity influences breast tumourigenesis still remains relatively unclear.

Differences in the sequences of an individual's genome is what contributes to our overall uniqueness. Any genomic variations such as large duplications or deletions (ie. copy number variation (CNV)) can influence our traits, and cause susceptibility to diseases such as obesity. Copy number variants are large structural and highly heritable germline variations extensive throughout the human genome (Figure 1.1), yet, non-recurrent *de novo* alterations can also cause the formation of unique germline CNV[5]. Research has confirmed that *AMY1* CNV, an extensively studied copy number variant in the human genome[6], has evolved as a response to strong positive selection imposed by starch intake in the human diet[7]. The *AMY1* gene is responsible for producing the salivary amylase protein active at the start of the human digestion process, breaking down large starch molecules into maltose as preparation for further catalysis in the stomach. The copy number of *AMY1* has recently been linked to a predisposition to obesity in which fewer copies of *AMY1* is a risk factor for the accumulation of excess fat mass[8]. Despite the already well supported link between obesity and breast cancer risk, the newly established genetic association between *AMY1* CNV and increased weight gain has not yet been examined in normal and obese breast cancer patients.

Therefore, this review will examine the potential impact that *AMY1* CNV may have on risk of breast cancer incidence, by initially providing an overview of obesity genetics, followed by examination of the obesity-breast cancer relationship and lastly assessing the impact of *AMY1* CNV on a predisposition to obesity.

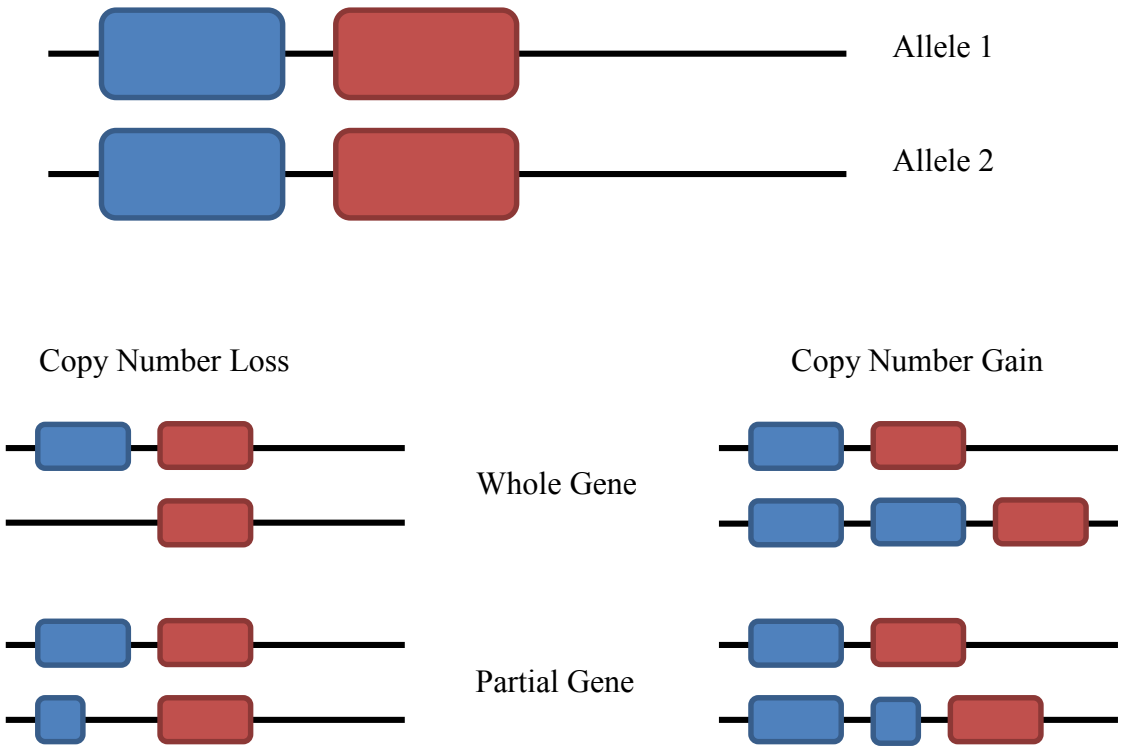



**Figure 1.1** Gene showing copy number variation. Copy number variation can be in the form of duplications and/or deletions seen here in the whole and partial gene loss and gain on the paternal allele.

## 1.1 Obesity

### 1.1.1 Obesity Epidemic

Obesity is a medical condition defined as the accumulation of excess fat mass to a point of significant increase in adverse health risks associated with the gain in weight[9]. The rising prevalence of obesity worldwide has given rise to the current term "obesity epidemic"[2, 10, 11]. In 2012 the World Health Organisation (WHO) calculated the prevalence of obesity worldwide to be 500 million, and by 2015 the number of cases was predicted to have increased to 700 million[9]. Obesity related literature presents confirmation that the incidence of obesity is increasing worldwide in both developing and developed countries and for people of all ages[1, 10, 12]. Generally the obesity epidemic is credited to an imbalance between energy consumption, energy expenditure and the consequent actions of the body's weight regulation systems producing a

consistent positive energy balance, in that dietary intake outweighs the energy output through exercise and metabolic activities[2, 10, 13].

Clinically a BMI (body mass index: weight in kilograms divided by the square of height in metres) above 30, categorises an individual obese[9]. BMI is a universal and efficient tool for measuring obesity[14].

### 1.1.2 Obesity and Predictive Ability of Genetics

Generally the obesity epidemic is credited largely to environmental influences; however, people subjected to comparable environments throughout their life can still have varied degrees of predisposition to obesity, and this is attributable to individuals having differential genetic makeup[15]. It is difficult to prove the precise underlying genetic cause of obesity, largely as a result of the complex mechanisms participating in the biological management of adiposity. Therefore, further research is required to establish the clinical significance of different genetic loci found to be associated with obesity. Following is a review of a number of studies that have attempted to uncover the genetic components affecting the incidence of obesity; split into monogenic and polygenic based analyses.

#### 1.1.2.1 Monogenic Obesity

In some cases the genetic component of obesity has been established as simply a result of rare single-gene mutations producing large effect, this is collectively defined as 'monogenic' obesity. Mouse studies in the mid-1990s revealed the first single gene mutation responsible for increased susceptibility to obesity and this was mapped to the human homologue *LEP* (leptin) gene[16]. Leptin functions as a cell signalling hormone in the regulation of food intake, body weight and appetite; where absence or resistance can lead to uncontrolled eating and weight gain. Since then research analysing single gene mutations have revealed numerous genes now known to play a role in Mendelian forms of obesity (Table 1.1). These single gene mutations have overtime lent insight into the genetic nature of this disease, along with uncovering a number of different biological pathways incorporating these genes that are now known to be implicated

during incidence of obesity. *MC4R* (Melanocortin 4 Receptor) is the most common single-gene form of obesity that is currently known, explaining about 5% of the severe early onset obesity cases[17]. *MC4R* codes for a membrane bound receptor protein that is mediated by G proteins and is reported to play an essential role in energy homeostasis and somatic growth[18].

Research surrounding monogenic obesity has become predominantly concerned with its relative importance in the development of personalised medicine[15]. For example, obese patients with *LEP* mutations can successfully overcome and reverse weight gain via leptin replacement[19, 20]. However, many studies have challenged the relative importance of monogenic forms of obesity in terms of their involvement in the obesity epidemic[21, 22], and are perhaps justified as these single-gene variants, although they produce a significant effect in the obese phenotype, only account for a minute proportion (~5%) of all severe obesity cases[17]. Moreover, medical advancements have been limited, and leptin deficiency is currently the only monogenic form of obesity with successful clinical treatments available[23].

**Table 1.1** Summary of genes associated with monogenic forms of obesity.

| Gene | Chr | Mutation Type | Obesity related characteristics | Study |
|---|---|---|---|---|
| ***LEP*** | 7 | Homozygous reading frameshift, Missense mutations | Severe, early onset obesity | Montague, C.T *et al* (1997)[24] Echwald, S.M. *et al*. (1997)[25] Oksanen, L. *et al.* (1997)[26] |
| ***LEPR*** | 1 | Skipping of exon16 producing truncated protein | Severe, early onset obesity | Clement, K. *et al.* (1998)[27] |
| ***POMC*** | 2 | Frameshift loss of function mutation | Severe, early onset obesity | Krude, H. *et al.* (1998)[28] |
| ***MC4R*** | 18 | Nonsense and missense mutations | Many obesity related traits | Farooqi, I.S. *et al.* (2003)[29] |
| ***BDNF*** | 11 | Deletions inducing haploinsufficient loss of function | Severe obesity | Han, J.C *et al.* (2008)[30] |
| ***NTRK2*** | 9 | Missense variant | Severe, early onset obesity | Yeo, G.S. *et al.* (2004)[31] |
| ***PC1*** | 5 | Compound heterozygote mutation | Severe, early onset obesity | Jackson, R.S. *et al.* (1997)[32] |
| ***SIM1*** | 1 and 6 | Haploinsufficiency due to balanced translocation between chromosomes | Severe, early onset obesity | Holder, J.L. *et al.* (2000)[33] |

Chr: chromosome

#### 1.1.2.2 Polygenic Obesity

Polygenic obesity is the term used in research to describe the much more common but complex forms of obesity resulting from the interaction of a multitude of different single nucleotide polymorphisms (SNPs), along with an interaction with environmental components. Studies exploring the genetic foundation of common obesity have been largely unsuccessful, that is until technological advancements unearthed novel genomic approaches such as genome wide association studies (GWAS). GWAS locate common variants associated with common phenotypes such as obesity. Association studies have identified many common SNPs associated with obesity and body mass index (Table 1.2). The *FTO* (fat mass and obesity associated) gene is known to contribute to the regulation of body size and fat accumulation, and even though its exact physiological function is not yet known, *FTO* dominates the genome wide association literature as the most common gene nearest obesity associated SNPs[34-36]. Similarly, over a range of selected GWAS, *FTO* was the most frequently connected loci to the strongest predictive SNP relating to obesity (Table 1.2). Additionally various loci, such as *MC4R*, that have been previously linked with monogenic obesity, also connect to common variants (SNPs) associated with complex obesity[22, 37]. It is important to note that although studies investigating obesity associated genes can be compared, such as in Table 2, they often vary in their definitions of obesity. For example one study[38] defined obesity by looking at the extremes (>3 SD) of the BMI normal distribution, whereas, another study[39] defined obesity by BMI class; BMI $\geq$ 30 kg/m$^2$ for obesity class I, BMI $\geq$ 35 kg/m$^2$ for obesity class II and BMI $\geq$ 40 kg/m$^2$ for obesity class III.

Although GWAS have been able to unearth many single nucleotide polymorphisms associated to increased adiposity, almost all of the single nucleotide base changes that have been associated with common obesity are not causal. Collectively these common variants have poor predictive power, only accounting for ~5% of the total heritability of BMI[21, 22, 40, 41]. Thus, research is still aiming to explain the remainder of body masses missing heritability and answers will require use of further technological advancements, such as next generation sequencing, in order to examine rarer variants whilst maintaining high predictive power[42].

**Table 1.2** Summary of the number and strongest associated SNPs with obesity or body mass index from different GWAS.*

| Study | Disease/Trait | Sample Size | Number of associated SNPs[a] | Strongest associated SNP (nearest gene ; Chr)[b] | Odds Ratio[c] or Beta |
|---|---|---|---|---|---|
| Wen et al. (2014)[43] | Body mass index | 86,739 | 14 | rs1558902 (FTO; 16) | 0.03-0.08 per BMI unit increase |
| Pei et al. (2013)[44] | Body mass index | 20,913 | 2 | rs6567160 (MC4R; 18) | 0.08-0.09 per kg/m2 increase |
| Graff et al. (2013)[45] | Body mass index | 13,627 | 6 | rs9940128 (FTO; 16) | 0.05-0.1 per BMI unit increase |
| Monda et al. (2013)[46] | Body mass index | 39,144 | 5 | rs7708584 (GALNT10; 5) | 0.02-0.07 per BMI unit increase |
| Berndt et al. (2013)[39] | Body mass index | 16,068 | 9 | rs11075990 (FTO;16) | 1.13-1.35[c] |
| Berndt et al. (2013)[39] | Obesity | 204,498 | 58 | rs7185735 (FTO; 16) | 1.04-1.45[c] |
| Wheeler et al. (2013)[38] | Obesity (early onset, extreme) | 6,889 | 8 | rs1421085 (FTO; 16) | 1.22-1.67[c] |
| Yang et al. (2012)[47] | Body mass index | 133,154 | 1 | rs7202116 (FTO; 16) | 0.04 per BMI unit increase |
| Bradfield et al. (2012)[48] | Obesity | 13848 | 2 | rs9568856 (OLFM4; 13) | 1.14-1.22[c] |
| Okada et al. (2012)[49] | Body mass index | 26,620 | 7 | rs12149832 (FTO; 16) | 0.04-0.07 per BMI unit increase |
| Wen et al. (2012)[50] | Body mass index | 27,715 | 5 | rs11671664 (GIPR; 19) | 2.55-4.22% increase per BMI unit |
| Jiao et al. (2011)[51] | Obesity | 327 | 2 | rs988712 (BDNF; 11) | 1.26-1.36[c] |
| Wang et al. (2011)[52] | Obesity | 1,060 | 2 | rs17817449 (FTO; 16) | NA |
| Speliotes et al. (2010)[53] | Body mass index | 123,865 | 31 | rs1558902 (FTO; 16) | 0.06-0.39 per kg/m2 increase |
| Scherag et al. (2010)[54] | Obesity (early onset extreme) | 2,258 | 2 | rs1558902 (FTO; 16) | 1.22-1.37[c] |
| Cotsapas et al. (2009)[55] | Obesity (extreme) | 3,972 | 1 | rs9941349 (MC4R; 18) | 1.48[c] |
| Meyre et al. (2009)[56] | Obesity | 2,796 | 3 | rs1421085 (FTO; 16) | 1.12-1.39[c] |
| Thorleifsson et al. (2008)[57] | Body mass index | 73,758 | 10 | rs8050136 (FTO; 16) | 3.63-8.04% standard deviation |
| Willer et al. (2008)[58] | Body mass index | 32,387 | 6 | rs9939609 (FTO; 16) | 0.07-0.33 per kg/m2 increase |
| Loos et al. (2008)[37] | Body mass index | 16,876 | 1 | rs17782313 (MC4R; 18) | 0.05 per unit increase in log(BMI) |
| Frayling et al. (2007)[36] | Body mass index | 10,657 | 1 | rs9939609 (FTO; 16) | 0.36 per kg/m2 increase |

Abbreviations; BMI- Body mass index, Chr- Chromosome, GWAS-Genome wide association studies, SNPs- Single nucleotide polymorphisms

[a] The number of associated SNPs with obesity or body mass index from different GWAS with a p-value threshold of $p < 10^{-8}$

[b] Strongest associated SNPs with obesity or body mass index from different GWAS with a p-value threshold of $p < 10^{-8}$

*****GWAS selected in: A Catalog of Published Genome-Wide Association Studies; https://www.genome.gov/page.cfm?pageid=26525384#searchForm

### 1.1.3 Obesity and Cancer Burden

It is well-known that excess weight gain is linked to an increased risk of many health related diseases, particularly heart disease, type 2 diabetes, obstructive sleep apnea, osteoarthritis and certain types of cancer[2]. Evidence suggests that a large proportion of all cancer related deaths in both sexes can be credited to patients carrying excess fat mass[59]. There is evidence for heterogenic qualities of obesity such as the existence and difference between white and brown adipose tissue[14], heritage of adipocyte cells[60, 61], and topographical location[62]. This heterogeneity has the potential to influence clinical outcome, particularly for obese cancer patients[14]. Thus, it is important to measure and examine obesity as a multifactorial disease as well as analysing the specific characteristics of adiposity, when carrying out obesity related research.

BMI is currently the best studied variable for the obesity and cancer relationship, most likely because BMI is a universal and well-established indicator of fat mass[59]. Numerous studies have confirmed that greater BMI is positively correlated with an increase in cancer risk and cancer mortality, most notably in cancers of the kidney, endometrium, colorectal, pancreas, post-menopausal breast and oesophageal adenocarcinoma[63-67]. Despite this well-studied and unequivocal BMI-cancer relationship, the multifactorial and heterogenic nature of both obesity and cancer means that these associations can only be assumed as correlative; the precise mechanisms acting during accumulation of excess fat mass causing an increased risk of cancer are less well understood[67, 68].

There are four distinct mechanisms that have emerged and currently dominate the literature in attempting to explain the pathophysiological epidemiology of obesity-related cancers. The four mechanisms are generally involved with both tumour initiation and tumour progression and include insulin resistance and hyperinsulinemia; adipokines expression such as leptin and adiponectin affecting cell growth, migration and invasion; obesity related inflammatory markers; and oestrogen producing proliferative effects[69-72]. Obesity is known to be associated with extended periods of increased circulating levels of insulin in the blood (reflecting insulin resistance), a condition known as hyperinsulinemia. Hyperinsulinemia, through the blocking of IGFBP (insulin-like growth factor-binding protein) production, triggers an increased

level of bio-available IGF-I (insulin-like growth factor 1), a potent mitogen and cell survival agent, thus acting to enhance and promote tumour initiation and tumour growth[73]. Interestingly, conditions that can promote insulin resistance include inflammation and cytokine secretion, which are themselves mechanisms that can alter the tumour microenvironment to promote tumour cell proliferation and angiogenesis[74]. Increased levels of obesity associated adipose tissue stimulates elevated levels of FFA, (free fatty acids) increasing the secretion of bioavailable adipokines (leptin), inflammatory cytokines and additional factors such as VEGF (vascular endothelial growth factor). Increased leptin levels are commonly observed in the serum of obese individuals, and this encourages further inflammatory responses in the adipose tissue, forming a proliferative positive feed-back loop[75]. Furthermore, leptin is known to contribute to the regulation of aromatase and oestrogen by adipose tissues[76], in which enhanced oestrogen signalling is linked to an increased cancer risk particularly in breast, endometrial and ovarian cancers[77]. However, the overall diversity in cancer development and progression means it is unlikely that one system is individually acting alone in obesity-related cancers[69], but it is important to understand these systems linking obesity and cancer, in order to improve procedures to prevent and treat obesity associated cancer.

## 1.2 Obesity and Breast Cancer

### 1.2.1 Susceptibility and Mortality Rates

It is well known that obesity can influence susceptibility and survival outcome in numerous human cancer types, including breast cancer. In 2011 breast cancer was rated the most frequent and prominent source of cancer death in women worldwide[78, 79]. Research has recognised that carrying excess fat mass increases susceptibility to breast cancer primarily in post-menopausal women[4, 80, 81]. Epidemiological evidence suggests that excess body mass can increase breast cancer risk in post-menopausal women by approximately 40%[63]. Importantly, excess BMI increases post-menopausal breast cancer risk in a non-linear fashion[82].

Generally increasing BMI is reported to act as a protective element for breast cancer risk in women pre-menopause[83, 84]. Although opposing evidence has been presented[4], a recent meta-analysis confirmed a clear inverse association between obesity and pre-menopausal breast cancer risk[85]. This study suggests that this link is heavily dependent on ethnic variation but suggests that body fat distribution may also play a role. The mechanisms responsible for the inversed relationship in pre-menopausal women are unclear. Metabolic and hormonal changes that are associated with abdominal adiposity have been linked with the risk of pre-menopausal breast cancer and are speculated to be causing this reversed relationship[86].

Obesity can also influence survival outcome in diagnosed breast cancer patients. A meta-analysis incorporating epidemiologic evidence from 43 publications assessed the relationship between breast cancer outcome and obesity at diagnosis, and determined that obese patients had a 66% reduced survival rate compared to non-obese counterparts[3]. Literature provides a somewhat diverse report on overall survival of obese patients compared to breast cancer specific survival rates of obese patients, where some evidence suggests that co-morbidity due to obesity in breast cancer patients causes more unfavourable overall survival rates[87, 88]. However, the above meta-analysis reported the consequence of obesity to be alike in both overall survival and breast cancer specific survival[3].

Four common hypotheses that may explain why obesity leads to reduced survival for breast cancer patients have been proposed. These include obese patients presenting with more advanced stages at time of diagnosis[89], having a higher risk of secondary breast cancer diagnosis[90], having naturally more aggressive tumours, and more likely getting lower doses of chemotherapy compared to non-obese equivalents. Research provides extensive evidence for the latter two hypotheses[91-94].

### 1.2.2 Prognosis and Pathogenesis

Obesity is associated with hormonal alterations that may be acting to promote breast tumourigenesis. There is strong evidence that increased concentrations of endogenous oestrogen can intensify the chances of developing post-menopausal breast cancer[95, 96]. Post-menopause, high levels of aromatase provided by adipose cells, converts androgens to oestrogens, therefore increasing the amount of serum oestrogen in obese

post-menopausal women, particularly in local breast tissue, compared to normal or underweight post-menopausal women[97, 98].

Research suggests the influence of obesity and weight gain on breast cancer susceptibility, and prognosis may fluctuate depending not only on menopausal status, mentioned earlier, but also oestrogen and progesterone receptor (ER and PR) expression[99-101]. Typically ER+PR+ breast cancer tumours are associated with an obesity mediated increased likelihood of recurrence and poorer prognosis in post-menopausal women[102, 103]. Conversely, obesity is reported to be associated with an absence of hormone dependence and reduced breast cancer risk in pre-menopausal women; however, excess fat mass in younger women often results in worse prognosis at the time of diagnosis[104, 105].

Obesity is significantly correlated with many traditional prognostic characteristics of breast carcinomas. Research confirms the predictive ability of increased BMI for presentation of larger tumour sizes, increased involvement of lymph nodes and greater chance of metastasis in post-menopausal oestrogen dependent breast tumours[102, 106].

Obesity exerts primary and secondary level effects on breast tumour pathogenesis. Weight gain in post-menopausal women parallels an increase in adipose tissue, in which adipose tissue is known to yield primary level effects in obesity-mediated breast tumourigenesis via secretion of inflammatory cytokines[102] and adipokines (leptin and adiponectin)[107, 108], which are associated with increased incidence of breast cancer pathogenesis. Furthermore, an increase in weight gain in post-menopausal women is indirectly associated with insulin resistance, demonstrated to exert secondary level consequences in breast tumourigenesis such as hyperinsulinemia and elevated bioavailability of oestrogen[109, 110]. Despite the wealth of literature dedicated to this area, the precise way in which obesity influences breast tumourigenesis still remains relatively unclear. Identification of possible mechanisms that are causing such a direct link have become centred on a dominant group of processes, including oestrogen signalling, hyperinsulinemia, adipokine expression (adiponectin, leptin) and inflammation[72, 110, 111]. General conclusions propose the relationship to be almost certainly occurring as a result of a complex interaction between these different mechanisms[109, 112].

### 1.2.3 Obese Breast Tumour Transcriptomic Signature

Genetic profiling is an informative tool enabling the genetic signatures of a number and/or different populations of cells to be identified and compared to each other. Analysis of tumour genomic DNA has classified a large cohort of breast tumours into four main breast tumour subtypes including, luminal A; luminal B; basal-like; and HER2 enriched[113, 114]. Currently, the expression of clinical biomarkers (ER, PR and HER2 expression) within a patients breast tumour, are utilised by pathologists to categorise which of these four established intrinsic subtypes of breast cancer the tumour represents. Most luminal A and B tumours are ER+ and/or PR+, HER2-enriched tumours are largely HER2+, and basal-like (triple negative) cancers are generally negative for all three clinical biomarkers[113, 114]. Based on a tumours ER, PR and HER2 status, breast cancer patients are offered clinical treatments intended to be characteristically effective for their individual tumour[114]. However, a number of ER+, PR+ and HER2+ tumours are still resistant to hormone blocking and anti-HER2+ therapies, which are current treatment strategies for targeting the growth of hormone dependent tumours[115]. The causal mechanisms responsible for this resistance are still unclear. Equally so, for tumours with non-dependence on all hormone receptors, giving it a basal-like (triple negative) phenotype, means, that not only is the tumour more aggressive, it cannot be treated effectively with either of the well-established endocrine and anti-HER2 clinical therapies. Consequently, triple negative tumour treatment relies predominantly on specifically designed chemotherapy regimens. Previous large scale studies have sought to identify novel co-expressing genes associated with ER, PR and HER2 breast tumour biomarkers. They suggest that revealing such genes may ultimately aid an explanation for the appearance of treatment resistance and/or elucidate novel therapeutic targets/treatments, which is particularly important in the case of aggressive basal-like tumours[116, 117].

Although, transcriptomic profiling of breast tumours has recognised these genomic patterns connected with clinically important pathological features [113, 114], little research has been carried out surrounding the genetic profiles of breast tumour genes in patients with differing BMI status, in fact only two studies have developed transcriptomic profiles for breast tumours from obese patients[118, 119].

The first, Creighton *et al.*, compared breast tumour signatures (n= 103) from normal and overweight patients collectively to obese tumour transcript patterns, from which they derived 662 genes ($p < 0.01$) that were being differentially expressed in obese versus non-obese tumours[118]. Of those genes differentially expressed in obese tumours, a high proportion (602 genes) were downregulated in obese tumours and linked with nucleus and transcription regulation systems. Similarly, Fuentes-Mattei *et al.* generated comparable transcriptomic data for 137 breast tumours, in which they identified 112 significantly differentially expressed genes, associated with 59 biological alterations, in tumours from obese oestrogen positive breast cancer patients[119]. Fuentes-Mattei *et al.* verified genes involved in the mechanistic pathways associated with adipokines as well as oestrogen, insulin and IGF-1 signalling in obesity-enhanced oestrogen dependent breast cancer development[119].

Interestingly, unpublished analysis of data derived from these studies highlights both *GRIA2* (glutamate receptor, ionotropic, AMPA 2) and *DUSP4* (dual specificity phosphatase 4) as potentially influential genomic locations in the obesity-breast cancer relationship, as both are significantly downregulated in obese patient tumours. *GRIA2* codes for one of four (*GRIA1-4*) glutamate receptor subunits that join to form a ligand-activated cation channel in the mammalian brain. The *DUSP4* gene product is expressed in a variety of different tissues where it is localised to the nucleus and functions by inactivating ERK1, ERK2 and JNK via phosphorylation.

Further analysis comparing transcriptomic signatures generated in these two studies will hopefully confirm an overlap between the genetic profiles, and inquiry into this overlap may introduce molecular links between obesity and breast cancer and therefore informative development of potential prevention and/or treatment targets.

## 1.3 Obesity and *AMY1* Copy Number Variation

### 1.3.1 Copy Number Variation

Differences in the DNA sequence of our genome contribute to our overall uniqueness, where no two people in the world are genetically identical, not even monozygotic

twins[120]. There are different forms of variation that appear in the human genome, ranging between gross structural alterations to single nucleotide polymorphisms (SNPs). Overall, SNPs are extremely common in the human genome but only account for a small proportion of the diversity currently acknowledged in human disease[121]. In contrast, structural variation such as copy number variants, inversions, insertions, deletions and duplications, that affect more than just a single nucleotide base, explain a much greater proportion of human genetic diversity in comparison to single nucleotide changes[122, 123].

Copy number variation (CNV) is a form of genetic variation present in the human genome. A copy number variant is a large structural duplication or deletion in which the consequence for the cell is an abnormal gene dosage of one or multiple genes. Copy number variants can range anywhere between 1000 nucleotide bases to numerous megabases in length. Duplication refers to extra copy/copies of a region and deletion refers to fewer than the normal number of copies (Figure 1.1). Research surrounding genomic variants has identified and catalogued greater than 100,000 copy number variants now regarded as a major source of genetic variation. Research indicates that CNV accounts for a significant percentage, approximately 12 %, of the variation currently recognised in the human genome[124]. The genome is susceptible to *de novo* copy number mutations during human development, but CNV can also be inherited[5]. CNV has potential clinical relevance in a number of diseases[125], due to the ability of copy number variants to influence our phenotypic traits. An example of a copy number variant in the human genome influencing susceptibility to common obesity is *AMY1*[8].

#### 1.3.1.1 Obesity and Copy Number Variants

Copy number variants are becoming increasingly popular in obesity research. Since SNPs classified by GWAS have only been able to partially account for the heritability of common obesity[15, 21, 22], there has been a shift toward more predictive power, and thus the motivation to assess the role of copy number variants in complex disease, including common obesity[121]. A large proportion of CNV studies interrogating obesity have discovered large rare structural variants linked to obesity[126-128]. Except for the well-studied 539 kb deletion at 16p11.2, replication of CNV analysis across different human populations is currently lacking in the literature. Based on this lack of

replication reports discussing correlations between obesity and copy number variants in the human genome are currently not as globally significant as well researched SNPs linked with common obesity[126, 128]. Research has also reported common biallelic copy number variants to be related to measures of obesity[129, 130], though common copy number variants are highly associated through linkage disequilibrium with SNPs. Polymorphisms yield small effect and restricted predictive power, thus the ability of common copy number variants in the genome will probably share a similar fate when it comes to contributing to the missing obesity heritability[131]. Not surprisingly, some copy number variants that have been identified to associate with the obesity phenotype are at loci also containing common polymorphisms that have been previously linked with BMI[57, 127]. Overall there is emphasis on the need for future research on complex multi-allelic copy number variants which are not as detectable in SNP-based GWAS and therefore may play an important role in understanding more about heritability of BMI and obesity[22, 132].

### *1.3.2 AMY1*

#### 1.3.2.1 Properties of Salivary Amylase

Numerous amylase genes cluster in the human genome on chromosome 1p21, *AMY1A*, *AMY1B* and *AMY1C* expressing in saliva and *AMY2A* and *AMY2B* in the pancreas. A large proportion of saliva proteins are produced by the salivary glands, in which salivary amylase is the most abundant of salivary proteins[133]. Salivary α-amylase is a monomeric protein secreted into saliva that begins the digestion of ingested starch molecules in the mouth and oral tract, by catalysing the hydrolysis of 1,4-alpha-glucoside bonds located throughout oligosaccharide and polysaccharide molecules[134]. Saliva amylase is the product of *AMY1*, well-known for its extensive copy number variation (CNV) in the human genome, ranging anywhere between 2 and 15 diploid copies per individual[6, 7].

The relative importance of salivary amylase in the digestion of large starchy molecules may have initially been underestimated, particularly when compared to pancreatic amylase enzymes. Studies have shown that salivary amylase not only has the ability to very rapidly modify the molecular structure of complex carbohydrates in the mouth and oral tract [134], but that salivary amylase exists and functions throughout the stomach

[135]. Moreover, once hydrolysis has begun, the somewhat metabolised starch acts to protect salivary amylase from stomach acid[136]. When collated these findings imply a significant contribution of salivary amylase to total amylase concentrations throughout the digestive system.

#### 1.3.2.2 Evolution of *AMY1* CNV

Variation in *AMY1* copy number at chromosome 1p21 is a stable and heritable genomic acquisition that arose during human evolution. Over time the human diet has become increasingly dependent on carbohydrate and starchy based foods as our agricultural industries have developed. Breakdown of these nutrients begins at the very start of the human digestion process where salivary amylase cleaves apart any large starchy molecules in preparation for further catalysis in the stomach[134]. Studies on salivary amylase have established and verified the functional, positive correlation between *AMY1* copy number and quantity of amylase protein found in saliva, thus influencing an increased rate of amylase enhanced starch breakdown[7, 137]. Moreover, research has confirmed that people coming from populations historically exposed to high starch diets have evolved, on average, more copies of *AMY1* compared to populations accustomed to reduced levels of starch intake[7]. Nutritional pressure exerted by high starch diets over time represents a strong positive selection on *AMY1* and subsequent salivary amylase production, causing the independent duplication of copy number in different populations around the world, thereby facilitating an increase in amylase induced starch metabolism[7, 138]. These findings indicate the critical role of salivary amylase in the digestion of large starchy molecules regularly consumed in the human diet.

#### 1.3.2.3 *AMY1* and Obesity

*AMY1* CNV is linked with a predisposition to obesity. A recent study by Falchi *et al.*[8] analysed gene expression data from adipose cells of 149 Swedish families with siblings that differed significantly in BMI using quantitative polymerase chain reaction (qPCR), and examined physiologically important copy number variants based on their ability to impact dosage of gene products; the study identified a single CNV region incorporating the amylase gene[8]. After replicating their findings in a total of just over 6,000 subjects, the researchers were able to conclude a dosage dependent affect associated with *AMY1*

CNV and a predisposition to obesity (Figure 1.2)[8]. The study argued that variation in *AMY1* copy number may explain somewhere between 2.5% and 20% of the total heritability associated with obesity[8], more than previously acknowledged by Mendelian variants and hundreds of common polymorphisms.

Falchi *et al.* findings[8] prompted Maria *et al.* in Mexico, where childhood obesity rates are of great concern, to assess the role of *AMY1* dosage in these children. Six hundred children from all around Mexico were sampled, with a ratio of 1:1, obese and normal weight according to BMI. *AMY1* copy number was assessed using digital polymerase chain reaction (digital PCR), and analogous to Falchi *et al.*, researchers concluded there was a notable dosage effect of *AMY1*; subjects with greater than 10 copies of *AMY1* were all normal weight according to BMI scoring[139]. Similarly, researchers in Finland have recently demonstrated that *AMY1* CNVs, specifically lower copy number, are associated with early-onset childhood obesity; however this association was only evident in young females[140].

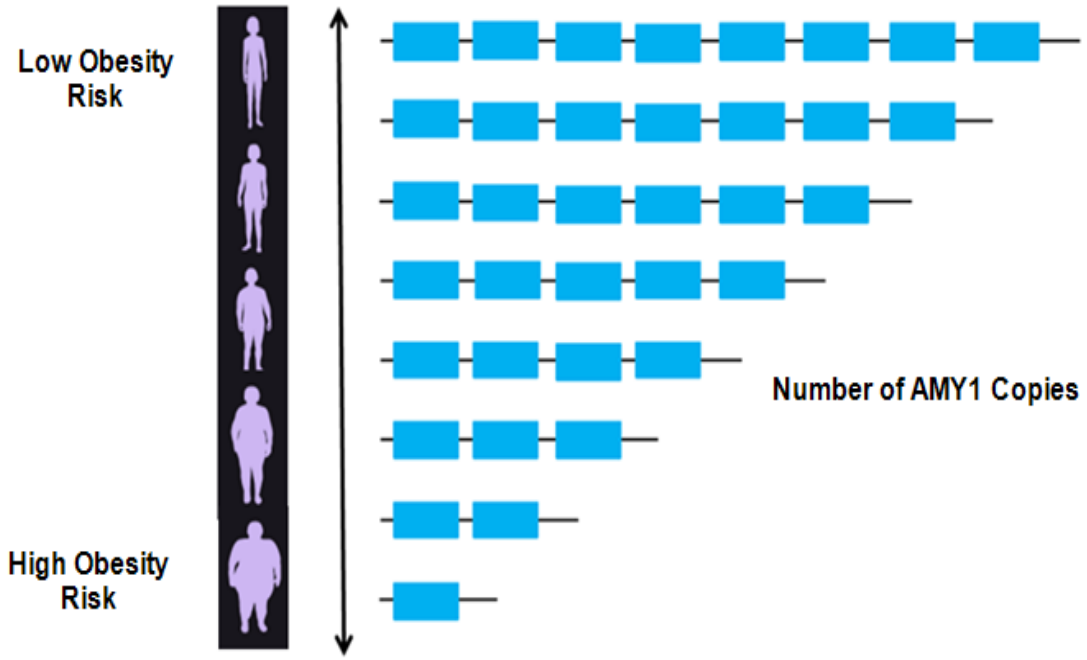



**Figure 1.2** Low dosage of *AMY1* has been credited as a significant risk factor toward accumulation of excess fat mass in people of European origin.

Despite the growing support in the literature for the inverse correlation between *AMY1* copy number and obesity, there is an undeniable discordance between what Falchi *et al.*[8] initially reported and what population based GWAS have revealed about the genetic contributions of common obesity. The amylase locus is structurally complex, making CNV not only especially difficult to calculate but challenging to derive connections with obesity susceptibility[141], questioning perhaps the choice of technique used in *AMY1* copy number derivation. A recently published study looking at *AMY1* copy number in a large multi-ethnic cohort, reported no association between *AMY1* copy number and BMI, and suggested that the disagreement is attributable to their use of higher resolution techniques, being digital PCR, for determining *AMY1* copy number[142]. Table 1.3 represents all of the current reports surrounding obesity and *AMY1* CNV, and the cohorts and techniques used to determine their findings.

**Table 1.3** Summary of the literature, to date, that has reported on *AMY1* copy number variation and its association with obesity, using either quantitative PCR or digital PCR to determine the status of the correlation. The study's authors, year of publication and the number of patient samples investigated are represented within each of the categories.

| ***AMY1* Copy Number Variation and Obesity** | | |
|---|---|---|
| | **Inverse Correlation** | **No Correlation** |
| **Quantitative PCR** | Falchi *et al*. 2014 (n=6,200) | |
| **Digital PCR** | Mejia-Benitez *et al*. 2014; (n=597)<br>Viljakainen *et al*. 2015 (n=132) | Usher *et al*. 2015 (n=3,500) |

### 1.3.3 Summary

There is clearly a need for research examining complex multi-allelic copy number variants to understand more about the heritability of obesity. *AMY1* represents the type of CNV that can fill this gap. The recently reported association between *AMY1* CNV

and a predisposition to obesity potentially explains some of the fairly large amount of missing heritability of human obesity.

Falchi *et al.* and Maria *et al.* are the first two studies of their kind, establishing a genetic link between digestion of starchy foods and predisposition to obesity[8, 139]. Furthermore, Falchi *et al.* reported a greater association between BMI and *AMY1* copy number, than between BMI and *FTO* SNPs, contradicting the currently dominant literature valuing *FTO* as the strongest associated loci with susceptibility to obesity in humans (Table 1.2). Interestingly, further investigation of the extensive genome wide meta-analysis carried out by Speliotes *et al.*, examining SNPs associated with measures of obesity, revealed a significant interruption of SNP reportage across the region incorporating *AMY1*[53]. Therefore, it is possible that increased structural variation such as large duplications or deletions, occurring in the amylase cluster, may obstruct the ability of GWAS locating SNPs in such genomic regions.

All of the above is evidence that CNV provides a relatively novel and largely unexplored component of the human genome. Further investigation of copy number variants, particularly those including *AMY1*, may answer some of the literature's pending questions, such as the missing obesity heritability, and the adverse health risks associated with weight gain, for example, obesity linked incidence and mortality of breast cancer.

## 1.4 Research Hypotheses

The prevalence of obesity is increasing at an alarming rate worldwide[1]. Obesity is recognised for its involvement in amplifying risk of adverse health effects, with particular concern focused on the increase in obesity associated breast cancer incidence and mortality worldwide[3, 4]. Regardless of past and current research supporting the positive correlation between obesity and breast cancer risk, the molecular basis and precise mechanistic properties of the association between obesity and increased breast cancer risk and development still remain unclear.

In the same way that past explorations have used genomic techniques to identify breast tumour subtypes, which can be treated with certain therapies for the most effective clinical outcome[114], it could be postulated that similar research may be able to determine genetic signatures and subsequent clinical therapies unique to obese breast tumours. It is conceivable that an obesity associated tumour initiation mechanism is not evident, or is transpiring at a less frequent rate in tumours arising in healthy weight women. Tumour gene expression analysis, of obese compared to healthy breast tumours, could reveal genes responsible for this suspected obesity associated breast tumour initiation system. Previous genetic profile studies have identified genes that may be associated with breast tumour phenotype in obese women[118, 119]. Although these intriguing findings suggest a potentially important relationship between genetic risk and molecular phenotype of breast tumours, this still requires further research with a well characterised cohort of breast cancer cases.

Additionally, recent research has demonstrated a clear inverse association between germline DNA copy number of the *AMY1* gene and the risk of obesity[8]. Conflict in the literature proposes that this inverse correlation is variable across different study cohorts. Therefore, although the association between *AMY1* copy number and obesity risk may be controversial[8, 142], the established genetic association between *AMY1* copy number variation and increased weight gain has not yet been examined in normal and obese breast cancer patients, despite the already well supported link between obesity and breast cancer.

**Hypothesis:** Based on the above it has been hypothesised for the current study that obese breast cancer patients are associated with (1) low germline *AMY1* copy number and (2) differential expression of other candidate genes in their breast tumours. In order to test the first section of this hypothesis, research was focused on determining whether lower *AMY1* copy number is inversely correlated with increased body mass index (BMI), in breast cancer patients. The second section of this hypothesis explored the molecular phenotype of breast tumours arising in women with differing BMI status, by investigating two research objectives. Initially, by selecting candidate gene markers for obesity associated breast cancer based on what is already known about the transcriptome of breast tumours arising in women with differing BMI status. Lastly, by analysing the gene expression profiles of these candidate gene markers in obese compared to healthy weight breast tumours.

For comparative purposes this project involved the recruitment of two cohorts of breast cancer patient samples; those considered healthy weight (BMI < 25) and those considered obese (BMI > 30) according to BMI scoring. Comparison between obese and healthy weight breast cancer patients was carried out on 55 blood and 40 overlapping tumour samples provided by The Cancer Society Tissue Bank Christhcurch.

# Chapter 2

# 2 Materials and Methods

## 2.1 Patient Samples and Tumour Specimens

A total of 16 pre-menopausal and 39 post-menopausal patients with histologically diagnosed breast cancer tumours available from The Cancer Society Tissue Bank were used in this study[143]. Sample selection was biased towards postmenopausal women (>55 years) and the patients recorded BMI at the time of sample collection. Twenty-seven patients with a BMI greater than 30 and twenty-eight patients with a BMI less than 25 were chosen. Available clinicopathological data was also obtained for all 55 samples, including age at onset; BMI; oestrogen and progesterone receptor status; *ERBB2* (HER2) expression; histological type and grade; and ethnicity (Table 2.1). Samples were approximately evenly split within each of these clinicopathological features among the obese and healthy cohorts.

The Cancer Society Tissue Bank only collects samples from which patients have given consent for banking and pertinent future research use. The University of Otago Ethics Committee approved the use of the banked samples and their available clinicopathological data (approval number H14/131).

The samples collected from The Cancer Society Tissue Bank were in the form of 55 patient FTA® blood cards (Whatman, Thermo Fisher Scientific, Albany, Auckland, New Zealand) and 40 overlapping fresh frozen tissue samples. The blood samples are taken from the consenting patient during pre-surgical blood collection and delivered along with notice of the patients' timetabled day of surgery. The tissue bank curator snap freezes all tissue samples in liquid nitrogen (-196 °C) as quickly as possible after collection and then stores them in secure -80 °C freezers[143].

**Table 2.1** Summary of the clinicopathological data for breast cancer patient samples used in this study.

| | Patient Blood Samples (n) | | | Patient Breast Tumours (n) | | |
|---|---|---|---|---|---|---|
| | BMI<25 | BMI>30 | **Total** | BMI<25 | BMI>30 | **Total** |
| **Obesity** | 27 | 28 | 55 | 18 | 22 | 40 |
| **Menopause** | | | | | | |
| Pre- | 8 | 8 | 16 | 0 | 2 | 2 |
| Post- | 19 | 20 | 39 | 19 | 19 | 38 |
| **Grade** | | | | | | |
| 1 | 2 | 1 | 3 | 2 | 1 | 3 |
| 2 | 9 | 7 | 16 | 6 | 4 | 10 |
| 3 | 16 | 20 | 36 | 11 | 16 | 27 |
| **ER** | | | | | | |
| Positive | 20 | 20 | 40 | 13 | 15 | 28 |
| Negative | 6 | 7 | 13 | 6 | 5 | 11 |
| Unknown | 1 | 1 | 2 | 0 | 1 | 1 |
| **PR** | | | | | | |
| Positive | 19 | 17 | 36 | 12 | 13 | 25 |
| Negative | 8 | 8 | 16 | 7 | 6 | 13 |
| Unknown | 0 | 3 | 3 | 0 | 2 | 2 |
| **HER2** | | | | | | |
| Positive | 4 | 6 | 10 | 3 | 6 | 9 |
| Negative | 18 | 18 | 36 | 10 | 12 | 22 |
| Unknown | 6 | 3 | 9 | 6 | 3 | 9 |
| **Histological Type** | | | | | | |
| IDC | 15 | 17 | 32 | 10 | 14 | 24 |
| Other | 12 | 11 | 23 | 9 | 7 | 16 |
| **Ethnicity** | | | | | | |
| NZ European | 26 | 24 | 50 | 19 | 18 | 37 |
| Maori | 0 | 3 | 3 | 0 | 2 | 2 |
| Other | 1 | 1 | 2 | 0 | 1 | 1 |

ER- Oestrogen Receptor; HER2- Human Epidermal Growth Factor Receptor 2; IDC- Invasive Ductal Carcinoma; PR- Progesterone Receptor

## 2.2 Nucleic Acid Extraction

### 2.2.1 DNA

#### 2.2.1.1 FTA® Extraction

Blood was extracted from FTA® cards using QIAamp DNA mini kit (Qiagen, Venlo, Linburg, Netherlands). FTA®cards trap lysed blood cells and provide a mechanism for germline DNA preservation. Punches (0.3 mm) were taken from blood stained FTA® cards for all 55 breast cancer patient samples. ATL buffer (tissue lysis buffer used for the purification of nucleic acids) and proteinase K were added to the FTA® punches in a microcentrifuge tube and incubated at 56 °C with shaking at 1,400 rpm for 1.5 hours. After incubation the extracted DNA was purified using the spin column/collection tubes and wash buffers from the QIAamp DNA mini kit (Qiagen, Venlo, Linburg, Netherlands). Optimisation required using 4 x 3 mm FTA® punches, 420 µL of ATL buffer and 60 µL of proteinase K to get sufficient DNA in the final 50 µL volume sample.

#### 2.2.1.2 Quantification of extracted DNA

DNA purity and yield, was quantified by measuring the A260/A280 absorbance ratio on the ND-8000 spectrophotometer (NanoDrop Technologies, Rockland, DE). Sample DNA at concentrations greater than 5 ng/µL were required, to ensure that after a 1:2 dilution there was still sufficient DNA from each patient sample for quantitative PCR (qPCR).

### 2.2.2 RNA Extraction and cDNA synthesis

Breast cancer tumour samples stored frozen at -80 ºC were collected from The Cancer Society Tissue Bank for the 40 breast cancer patients described above, matching with the blood samples used for germline DNA extraction. Frozen tumour samples were treated with liquid nitrogen and homogenised by crushing into a powder using a mortar and pestle held on dry ice.

The mRNA was then extracted from the frozen breast tumour powder using TRIzol Reagent® (Invitrogen Corporation, Carlsbad, California, USA) and chloroform (5:1)

and centrifuging for 15 minutes at 12,000rpm in 4 °C. The clear aqueous phase containing the RNA was collected, and the remaining DNA and protein phases were frozen at -80 °C. The aqueous RNA phase was purified using the RNA clean-up and concentration kit (Norgen Biotek Corporation, Thorold, Ontario, Canada; Cat# 23600). RNA yield was quantified using a Qubit™ RNA assay (Invitrogen Corporation, Carlsbad, California, USA).

## 2.3 Protein Extraction and Purification

The remaining DNA and protein phases frozen after RNA extraction for all 40 breast cancer tumours were treated with 300 µL of 100% ethanol and centrifuged at 2000 rcf for 5 minutes at 4 °C to pellet the DNA. The phenol-ethanol supernatant containing the tumour protein was collected and the remaining DNA pellet was frozen at -20 °C. The protein in the phenol-ethanol supernatant was collected and pelleted by centrifugation at 12,000 rcf for 10 minutes at 4 °C. The protein pellet was washed using 2000 µL of 0.3 M guanidine hydrochloride in 95% ethanol multiple times before being resuspended in 200 µL of 1% sodium dodecyl sulphate (SDS) by sonication. Insoluble material was removed and remaining soluble protein supernatant was used for Western blotting. Required protein quantities for Western blotting were attained for 36 of the 40 tumour samples.

## 2.4 Bioinformatics

The bioinformatic analysis focused on two different studies with publically available datasets. Searching “breast cancer gene expression signature obesity” in PubMed specifies 10 papers, of which two provide usable breast tumour gene expression data for the purpose of this study; one by Creighton *et al.*[118] and the other by Fuentes-Mattei *et al.*[119].

NextBio (https://www.nextbio.com) was used to gather breast cancer data generated by Creighton *et al.*[118] by searching for studies exploring 'obesity' and 'breast cancer'. A comparison of gene expression between 'obese (BMI > 30)' vs 'healthy (BMI < 25)' breast cancer patients according to BMI scoring was conducted. Creighton *et al.*[118] did not exclude the overweight group from their analysis; hence Nextbio was used to generate the 'obese' vs 'healthy' association. Only the top 100 of the most statistically differentially expressed genes from Creighton *et al*.[118] were selected for the current analysis.

The supplementary material from the Fuentes-Mattei *et al.*[119] study reported changes in gene expression after comparing ER+ breast tumours from obese (BMI > 30) and non-obese (BMI < 30) women. BMI data was not reported by Fuentes Mattei *et al.* (2014) and therefore non-obese patients could not be split between healthy and overweight. Thus, the non-obese category used by Fuentes-Mattei *et al.*[119] was treated as comparable to 'healthy + overweight' used by Creighton *et al.*[118].

Lastly, gene duplicates were removed for any genes that appeared more than once in a dataset due to multiple probe ID callings.

## 2.5 Quantitative Polymerase Chain Reaction (qPCR)

### 2.5.1 *AMY1* Copy Number Assay

#### 2.5.1.1 Primer/Probes

Three pre-designed different primer/probes were used for relative copy number analysis including, RNase P (ribonuclease P) and two different *AMY1* target primer/probes with varying amplicon length (Invitrogen Corporation, Carlsbad, California , USA) (Table. 2.2). These primer/probe assays were chosen because they were used successfully in a previous nature genetics paper that explored germline *AMY1* copy number variation (CNV) in the human genome[8]. RNase P was used as the reference assay which maps within the single exon *RPPH1* gene, the Ribonuclease P RNA component H1. Both *AMY1* primer/probes were used as a method of validation, both mapping the same exon boundaries of the *AMY1* target sequence.

**Table 2.2** Copy number assay primer/probes.

| Target | Assay ID/Catalog Number | Amplicon Length |
|---|---|---|
| *AMY1* | Hs 07226361_cn | 81 bp |
| *AMY1* | Hs 07226362_cn | 101 bp |
| RNase P | 4403328 | 87 bp |

#### *2.5.1.2* Quantitative PCR

Quantitative PCR was carried out using the Roche LightCycler® 480 II (Roche Applied Science, Indianapolis, IN, USA). The quantitative PCR protocol that was followed is the KAPA Probe Fast qPCR Kit Master Mix (2x) Universal for the 384 well plate. Assay optimisation indicated, due to *AMY1* having potentially a 7 fold increase in primer targets compared to RNase P, that primer limiting allowed RNase P to reach plateau, confirming the use of 10x *AMY1* primer/probe and 20x RNase P reference primer/probe reaction concentrations. Final reaction volumes included 1.5 µL of extracted target DNA, 1.5 µL of DEPC-treated water and 7 µL of the reaction Master Mix which included 5 µL of the 2x KAPA Probe Master Mix, 0.5 µL of 10x *AMY1* primer/probe (FAM™ dye–labelled)(Life Technologies, Carlsbad, CA, USA), 0.5 µL of 20x Ribonuclease P (RNase P) reference primer/probe (VIC® dye–labelled)(Life Technologies) and 1 µL of DEPC-treated water. Each PCR run had a 'no template' control (7 µL of Master Mix and 3 µL DEPC-treated water), HapMap samples NA18956 and NA18972 at 1 ng/µL with known copy number of 6 and 14, respectively, and a DNA sample of known *AMY1* copy number acting as an internal calibrator across the different reactions. Each sample was run in triplicate for both *AMY1* Hs07226361 and Hs07226362 assay ID primer/probes.

The PCR included an activation step at 95 °C for 10 minutes, followed by 40 cycles of denaturation (95 °C, 20 sec) and annealing (60 °C, 1 min) and finally cooling for 2

minutes at 37 °C. Fluorescence was measured once each cycle at 465-510 nm (FAM™ dye–labelled) and 533-580 nm (VIC® dye–labelled), after primer extension.

Relative expression values were calculated by the LightCycler® Software (Roche). Relative quantification was calculated using the comparative Ct model following the arithmetic formula $2^{(-\Delta Ct)(144)}$. ΔCt is the difference between the mean Ct values of the target gene and the assumed two copy number reference gene, RNase P. Mean Ct or the threshold cycle for each sample is inversely proportional to the amount of DNA present, and was calculated as the fractional cycle number averaged across the three replicates at which the fluorescence emitted exceeds the background threshold. The threshold for the detection of the sample during PCR was set above the background fluorescence and fell within the exponential phase of DNA amplification. Copy number of *AMY1* for each sample was calculated using two HapMap samples, NA18956 and NA18972, with known copy number of 6 and 14, respectively, to which each sample was normalised.

### 2.5.2 qPCR Analysis of Gene Expression Levels

#### 2.5.2.1 cDNA Synthesis

Reverse transcription of the extracted mRNA from each breast tumour to get first-strand cDNA was carried out immediately after extraction and purification, using Superscript® III First Strand Synthesis System for RT-PCR (First Strand Kit; Invitrogen, Carlsbad, California, USA). Each 20 μL reaction included 10 μL of 2x RT Reaction Mix (including, oligo$[dt]_{20}$ [2.5 μM], random hexamers [2.5 ng/μL], 10 mM $MgCl_2$ and dNTPs), 2 μL of RT Enzyme Mix (including, Superscript® III RT and RNaseOUT™) and 8 μL of extracted mRNA. Reactions were mixed together, centrifuged and placed under reverse transcription conditions, including incubation for 10 minutes at 25 °C, then 30 minutes at 50 °C and finally termination at 85 °C for 5 minutes. 1 μL of E.coli RNase H was added to each reaction, followed by 20 minutes incubation at 37 °C. Each reverse transcription run included a ‘no template’ control. The samples were frozen at -20 °C until used for quantitative PCR.

#### 2.5.2.2 Primer design

Primers were designed for 5 genes including, *GRIA2, DUSP4, NR2F1, ADH1B* and *EEF1A1 (*Table 2.3). Primers were designed using Primer3 version 0.4.0. All gene sequences were taken from the U.S. National Centre for Biotechnology Information (NCBI). The primers were selected to span adjacent exons acknowledged as reliable matches for the gene of interest. The National Centre for Biotechnology Information (NCBI) Primer-BLAST tool was used to validate the specificity of the primer sets. Primers were synthesised by Integrated DNA Technologies (IDT) and suspended in DEPC-treated water to a final concentration of 50 μM and stored at -20 °C.

**Table 2.3** SYBR® primer sequences.

| Target | Forward Primer | Tm | Reverse Primer | Tm | Amplicon Length |
|---|---|---|---|---|---|
| *GRIA2* | TCAATGGGACA AGTTTGCAT | 55.81 | TTGCCATTTCTT TTCAGCAG | 55.05 | 94 |
| *DUSP4* | CCTGGTTCATGG AAGCCATA | 57.27 | TAGGCCAGGCA GATGGTG | 58.68 | 112 |
| *NR2F1* | GCCTCAAGAAGT GCCTCAAA | 58.39 | AGTGCGTACTG GCCTGGAT | 60.99 | 94 |
| *ADH1B* | GCTACTGACTGG ACGCACCT | 61.88 | CGCATCCAGTG AAAACTTCTT | 57.10 | 112 |
| *EEF1A1* | AAAATGACCCA CCAATGGAA | 55.39 | CGTGTGGCAAT CCAATACAG | 57.17 | 107 |

#### 2.5.2.3 Quantitative PCR

Quantitative PCR was carried out using the Roche LightCycler® 480 II (Roche Applied Science, Indianapolis, IN, USA). For the purpose of this study the quantitative PCR protocol followed is the Platinum® SYBR® Green qPCR SuperMix-UDG for a 384 well plate. Each reaction 0.075 μL of target cDNA (reverse transcribed extracted mRNA), 2.925 μL of DEPC-treated water and 7 μL of the reaction Master Mix which included

3.75µL of SYBR Green Master Mix, 0.375 µL of both forward and reverse primers at 5 µM for the selected target and 2.5 µL of DEPC-treated water. Each PCR run contained both 'no template' controls (7 µL of reaction Master Mix and 3 µL DEPC-treated water) and reverse transcription negative controls (7 µL in reaction Master Mix and 3 µL of first-strand cDNA negative control). Each sample was run in triplicate.

The PCR included an activation step at 95 °C for 10 minutes, followed by 40 cycles of denaturation (95 °C, 15 sec) and annealing (60 °C, 1 min). Fluorescence was measured once each cycle at 465-510 nm after primer extension. After amplification a melting curve was commenced by lowering the temperature to 60 °C for 1 minute before continually increasing the temperature (11 °C per second) to 95 °C and measuring the fluorescence throughout melting. Finally, the plate was cooled for 30 seconds at 40 °C.

Relative expression values were calculated by the LightCycler® Software (Roche), where relative quantification was calculated using the comparative Ct model using the arithmetic formula $2^{(-\Delta Ct)(144)}$. ΔCt is the difference between the mean Ct values of the target genes (*GRIA2, DUSP4, NR2F1* and *ADH1B*) and the reference gene (*EEF1A1*). Mean Ct or the average threshold cycle for each sample is inversely proportional to the amount of cDNA present and was calculated as the fractional cycle number averaged across the three replicates at which the fluorescence emitted exceeds the background threshold.

## 2.6 Digital Polymerase Chain Reaction

### 2.6.1 Sample Selection

A recent publication has suggested that the sensitivity of quantitative PCR (qPCR) in insufficient for copy number calling and that digital PCR is more suitable[142]. For comparative purposes 3 of the 55 blood samples from the breast cancer patients used for *AMY1* copy number analysis using qPCR were selected for copy number investigation using digital PCR. The 3 samples chosen were all healthy weight patients and based on qPCR results, covered a range of *AMY1* copy number variants (6-15

*AMY1* copies). Patient germline DNA was extracted from FTA cards following the FTA extraction process described earlier. DNA sample concentrations were determined using Qubit™ DNA assay (Invitrogen Corporation, Carlsbad, California, USA) (Table 2.4).

**Table 2.4** DNA sample concentrations for the 3 patient samples selected for analysis using digital PCR.

| Patient Sample | BMI | DNA (ng/µL) |
|---|---|---|
| 5 | 19.34 | 0.40 |
| 8 | 24.34 | 0.49 |
| 22 | 18.02 | 0.48 |

### 2.6.2 DNA Digest

The QuantStudio 3D digital PCR system requires each target region (ie. each *AMY1* copy) to be in a discrete well, in the digital PCR chip, for accurate quantification. To ensure this a DNA digest was designed to cleave within *AMY1* but leaving the amplicon site uncut. The restriction enzyme used was *Hind*III, which cuts at 5'-A|AGCTT-3'. All *Hind*III restriction sites within *AMY1* were resolved using genome browser. The restriction enzyme was chosen because it was not effected by methylation and it did not cleave within either the *AMY1* amplicon or the RNase P reference gene (Supplementary Figure 6.1).

The controls used for this assay included a 'no template' control, a genomic DNA sample of known copy number (92.1 ng/µL) and the two HapMap DNA samples NA18956 (345 ng/µL) and NA18972 (331 ng/µL) with *AMY1* copy number of 6 and 14, respectively. Digests for controls were carried out in 50 µL reaction volumes including 3 µL of *Hind*III (20 U/mL) restriction enzyme (New England Biolabs, Ipswitch, Massachusetts, USA), 30 µL of 10X NEBuffer 2 (New England Biolabs, Ipswitch, Massachusetts, USA), 300 ng total DNA and remaining volume of water.

Due to the low DNA concentrations of the samples, digests were carried out in 30 µL reaction volumes. Reaction volumes included, 1.5 µL of *Hind*III (20 U/mL) restriction enzyme (New England Biolabs, Ipswitch, Massachusetts, USA), 5 µL of 10X NEBuffer 2 (New England Biolabs, Ipswitch, Massachusetts, USA) and 23.5 µL of sample DNA. DNA digests were incubated for 16 hours at 37 °C before increasing the temperature to 80 °C for 20 minutes to denature the enzyme.

### 2.6.3 Digital PCR

A digital PCR reaction involves a 20,000 well chip, PCR cycling and chip analysing on the Quantstudio™ 3D Digital PCR system (Thermo Fisher Scientific, Rockford, IL, USA). Each individual chip used in this experiment investigated *AMY1* copy number status for a single sample. The samples analysed included 3 patient germline DNA samples. Digital PCR chips were loaded as per manufacturer protocols. In brief, a total reaction volume of 16 µl containing, 7.25 µl of QuantStudio™ 3D Master Mix (Thermo Fisher Scientific, Rockford, IL, USA), 0.725 µL *AMY1* primer/probe (FAM™ dye–labeled)(Life Technologies, Carlsbad, CA, USA), 0.725 µL of 20x Ribonuclease P (RNase P) reference primer/probe (VIC® dye–labelled)(Life Technologies) and 7.3 µl of the digested DNA samples, was all loaded onto one chip. Loading was performed using the QuantStudio™ 3D Digital PCR Chip Loader, before chips were placed on a flat block thermocycler fixed at an 11° angle. The thermocycling reaction consisted of an initial activation step at 96 °C for 10 minutes, then 40 cycles of annealing (60 °C, 2 minutes) and denaturation (98 °C, 30 sec), and ending with cooling for 2 minutes at 60 °C. Chips were held at 10 °C until fluorescence was analysed. Fluorescence was measured once at the end of PCR on the QuantStudio™ 3D Digital PCR Instrument, to determine of the 20,000 wells per chip, the number of wells occupied by a target, reference, both or neither of these molecules. Fluorescence data was analysed using the online QuantStudio 3D AnalysisSuite Cloud Software, which calculates copies/µL of the *AMY1* and RNase P genes in each sample based on a Poisson distribution. Copy number is calculated by calibrating each *AMY1* target to its RNase P reference and lastly normalising to the HapMap samples NA18956 and NA18972, with known *AMY1* copy number of 6 and 14, respectively.

## 2.7 Western Blotting

### 2.7.1 Protein Quantification

Total protein concentrations for each tumour sample after protein extraction and purification were determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, Rockford, IL, USA) in order to ensure equal protein loading into the gel (n=36). A standard curve was constructed using BSA (0-1000 µg/µl) following microplate recommendations and reading absorbance at 562 nm on the wallac 1420 Victor$^3$ plate reader (PerkinElmer Life and Analytical Sciences, Massachusetts, USA) (Supplementary Figure 6.2). The standard curve was used to extrapolate from and calculate the extracted protein concentration for each tumour sample.

### 2.7.2 SDS-Page

Each protein sample was made up to a final volume of 200 µl at 1 µg/µl including, 100 µl of 2 µg/µl protein in PBS, 50µl of 4X NuPAGE® LDS sample buffer (Novex, Life Technologies, Carlsbad, CA, USA), 20 µl of 1M DTT (dithiothreitol) (Sigma-Aldrich, St. Louis, MO, USA) and 30 µl of water. The samples were heated to 70 °C for 10 minutes before loading into NuPAGE® 4-12% Bis-Tris Gel (Novex, Life Technologies). According to the protein loading optimisation Western blots for DUSP4 required 10 µg of total protein and 40 µg of total protein for GRIA2, loaded into each well. On each gel 4 µl of SeeBlue® Plus2 Prestained Standard (Novex, Life Technologies) was loaded and gels were run at 125V for 65-70 minutes using 1X NuPAGE® MOPS SDS Running Buffer (Novex, Life Technologies).

### 2.7.3 Transfer

PVDF membranes were incubated for 10 seconds in methanol for activation. The transfer setup (Supplementary Figure 6.3) was used to transfer protein from gel to PVDF (polyvinylidene fluoride) transfer membrane (Thermo Fisher Scientific, Rockford, IL, USA). Transfer used 1X Blot™ Transfer buffer (Novex, Life Technologies) with 10% methanol at 25V for 1 hour. GRIA2 and DUSP4 membranes

were blocked with 5% milk in TBS-tween 20 (Sigma-Aldrich, St. Louis, MO, USA) for 1 hour before antibodies were introduced.

### 2.7.4 Antibodies

The primary antibodies for GRIA2 and DUSP4 were analysed in breast tumour samples (n=36). The anti-ionotropic glutamate receptor 2 (GRIA2) rabbit monoclonal primary antibody (EP966Y) (Abcam, Cambridge, MA, USA) was used at an optimised concentration of 1/500 with 5% skimmed milk in TBS-T incubated overnight at 4 °C. The anti-DUSP4 C-terminal rabbit polyclonal primary antibody (Abcam, Cambridge, MA, USA) was used at manufacturers instructed concentration of 1/500 with 5% skimmed milk in TBS-T incubated overnight at 4 °C.

The DUSP4 blot was stripped with two 25 µL washes of mild stripping buffer (mild stripping buffer includes: 1.5 g glycine, 1 mL 10% SDS and 1 mL Tween20; made up to 100 µL and adjusted to pH 2.2). The stripped blot was blocked with 5% skimmed milk in TBS-T and re-probed for β-actin. The GRIA2 blots were cut to separate the GRIA2 and β-actin bands, before probing with primary antibodies. The β-actin primary antibody used was mouse monoclonal anti-β-actin antibody (Sigma-Aldrich, St. Louis, MO, USA) and was used at 1/10,000 with 5% milk in TBS-T and incubated for 1 hour.

Blots were washed with TBS-T and secondary antibodies were applied. The polyclonal goat anti-rabbit immunoglobulins/HRP secondary antibody (Dako, Glostrup, Copenhagen, Denmark) was used for both GRIA2 and DUSP4 blots at 1/5000 in TBS-T for 1 hour at room temperature. The secondary antibody used for β-actin was the goat anti-mouse polyclonal immunoglobulins/HRP antibody (Dako, Glostrup, Copenhagen, Denmark) used at 1/10,000 in TBS-T incubated for 1 hour. PVDF membranes were imaged by chemiluminescence on the Alliance 4.7 (Uvitec Cambridge, Cambridge, UK), after applying Amersham ECL primer Western blotting detection reagents (GE Healthcare, Little Chalfont, Buckinghamshire, UK).

Due to barely detectable GRIA2 expression and dense background disguising the bands, the GRIA2 blots were re-wet with methanol for 30 seconds, blocked and re-probed following the above protocol.

### 2.7.5 Relative expression

The Western blot bands were quantified using Image J software. The DUSP4 and GRIA2 densitometry values were compared to their respective β-actin bands to calculate relative expression values. Relative expression values for each sample were then normalised across gels using the positive control sample run on each gel.

## 2.8 Statistical Analysis

Welch T-tests were performed on mean expression data or copy number estimates between breast cancer patient groups stratified by clinicopathological features. A P-value less than 0.05 was considered as statistically significant.

In order to visualise the change in mean expression data or copy number estimates data was analysed and represented as scatterplots using R statistical software (R-Studio Version 3.2.2).

# Chapter 3

# 3 *AMY1* Copy Number Analysis

## 3.1 Introduction

There is a strong link between obesity and increased incidence and reduced survival rate in diagnosed breast cancer patients, particularly in post-menopausal women[3, 4]. Despite the vast array of evidence that supports the strong association between obesity and breast cancer, the precise underlying mechanisms responsible for such a strong link are not that well understood.

A form of germline alteration known as copy number variation (CNV) has become increasingly popular in molecular based obesity research. Copy number variants are large structural duplications or deletions which may overlap gene regions, resulting in variable gene dosage. These variants have been implicated in many different human diseases, including increasing susceptibility towards obesity[121], such as the CNV region overlapping the *AMY1* gene[7]. Recent research has demonstrated a clear inverse association between germline DNA copy number of the *AMY1* gene and the risk of obesity[8]. *AMY1* codes for the salivary amylase protein, functioning at the beginning of the human digestion process, which breaks down large starchy molecules in the diet, as preparation for further catalysis in the stomach. A case-control study in Mexican children also reported an overall dosage dependent association between *AMY1* copy number and risk of obesity, and that higher *AMY1* dosage is a protective factor against obesity[139]. A further study in Finland showed an increased risk of early-onset childhood obesity associated with lower *AMY1* copy number in females, but not in males[140]. Interestingly, however, a very recent publication investigated germline *AMY1* copy number in a large multi-ethnic cohort totalling approximately 3,500 individuals and reported no correlation between *AMY1* copy number and risk toward increased BMI[142].

Despite the now controversial but recently well supported link between obesity and *AMY1* CNV, the proposed genetic association between *AMY1* CNV and increased

weight gain has not yet been examined in normal and obese breast cancer patients. The aim of this study was to identify whether germline *AMY1* copy number has a molecular link with breast cancer through its association with obesity.

## 3.2 Experimental Design and Research Aims

In this chapter, DNA was analysed from 55 breast cancer patient blood samples sourced from The Cancer Society Tissue Bank, 27 of which were categorised as healthy weight (BMI < 25) and 28 as obese (BMI > 30) according to BMI scoring (see Table 2.1 for patient clinicopathological data). Quantitative PCR was used to determine germline *AMY1* copy number for comparison between healthy and obese BMI groups (section 2.5.1).

To examine *AMY1* copy number calculated using different experimental techniques, digital PCR was also utilised to measure *AMY1* copy number in three of the 55 blood samples (section 2.6).

## 3.3 Results

### 3.3.1 Germline *AMY1* Copy Number in Obese Relative to Healthy Breast Cancer Patients

The germline *AMY1* copy number in obese compared to healthy breast cancer patients shows a definite trend toward lower copy number of *AMY1* in obese breast cancer patients (Figure 3.1). Two different *AMY1* primer probe assays were investigated in this chapter, the Hs07226361_cn (AMY1-61) and the Hs07226362_cn (AMY1-62) *AMY1* assays. AMY1-61 showed a reduction in the average *AMY1* copy number from healthy (mean= 8.86) to obese (mean= 8.11) breast cancer patients. AMY1-62 also showed a reduction in the average *AMY1* copy number from healthy (mean= 9.69) to obese (mean= 8.06). Each assay provided evidence for an inverse correlation between *AMY1* copy number and obesity risk in breast cancer patients, although the difference in *AMY1* copy number between healthy and obese breast cancer patients was only

statistically significant using the AMY1-62 primer probe assay (p= 0.02). Overall the two different *AMY1* assays are supported by previous reports of a relationship between lower *AMY1* copy number and susceptibility toward obesity [8, 139, 140].

Germline *AMY1* copy number from the breast cancer study cohort was also compared to other pathological features; ER (positive vs negative); PR (positive vs negative); HER2 (positive vs negative); and Grade (1&2 vs 3) (Supplementary Table 6.1). These comparisons indicated that patient germline *AMY1* copy number was not significantly different between ER+ versus ER- (p= 0.67); PR+ versus PR- (p= 0.28); HER2+ versus HER2- (p= 0.76); and Grade 1-2 versus 3 (p= 0.15) breast tumours (AMY1-62 assay). Additional analysis of the association between BMI and pathological features of the patient's breast tumours, indicated that there was no significant difference within the pathological characteristics of the breast tumours between obese and healthy patients (Supplementary Table 6.2).

Overall, germline *AMY1* copy number did not differ within clinically important breast tumour pathological features, but does show an inverse correlation with BMI in breast cancer patients.

### 3.3.2 Comparing *AMY1* Copy Number Determined using Quantitative PCR and Digital PCR

The aim for this section was to determine if methodological technique plays a role in copy number determination. *AMY1* copy number in germline DNA, determined using quantitative PCR (qPCR) and digital PCR, was compared between three breast cancer patients and one healthy control sample (Figure 3.2). qPCR results gave the three samples (patients 5, 8, 22) copy number estimates of 8.7, 15.3 and 6.4 copies respectively. Alternatively, digital PCR predicted the same three samples to have copy number estimates of 12.3, 15.6 and 7.6, respectively. The germline *AMY1* copy number calculated by digital PCR was higher, compared to qPCR, for all three patient samples. The biggest difference in calculated copy number was patient Sample 5, with a difference of 3.6 copies, and the smallest difference was in Sample 8 with a difference of 0.3 copies. Digital PCR produced marginally higher *AMY1* copy number numbers compared to qPCR (p= 0.51).

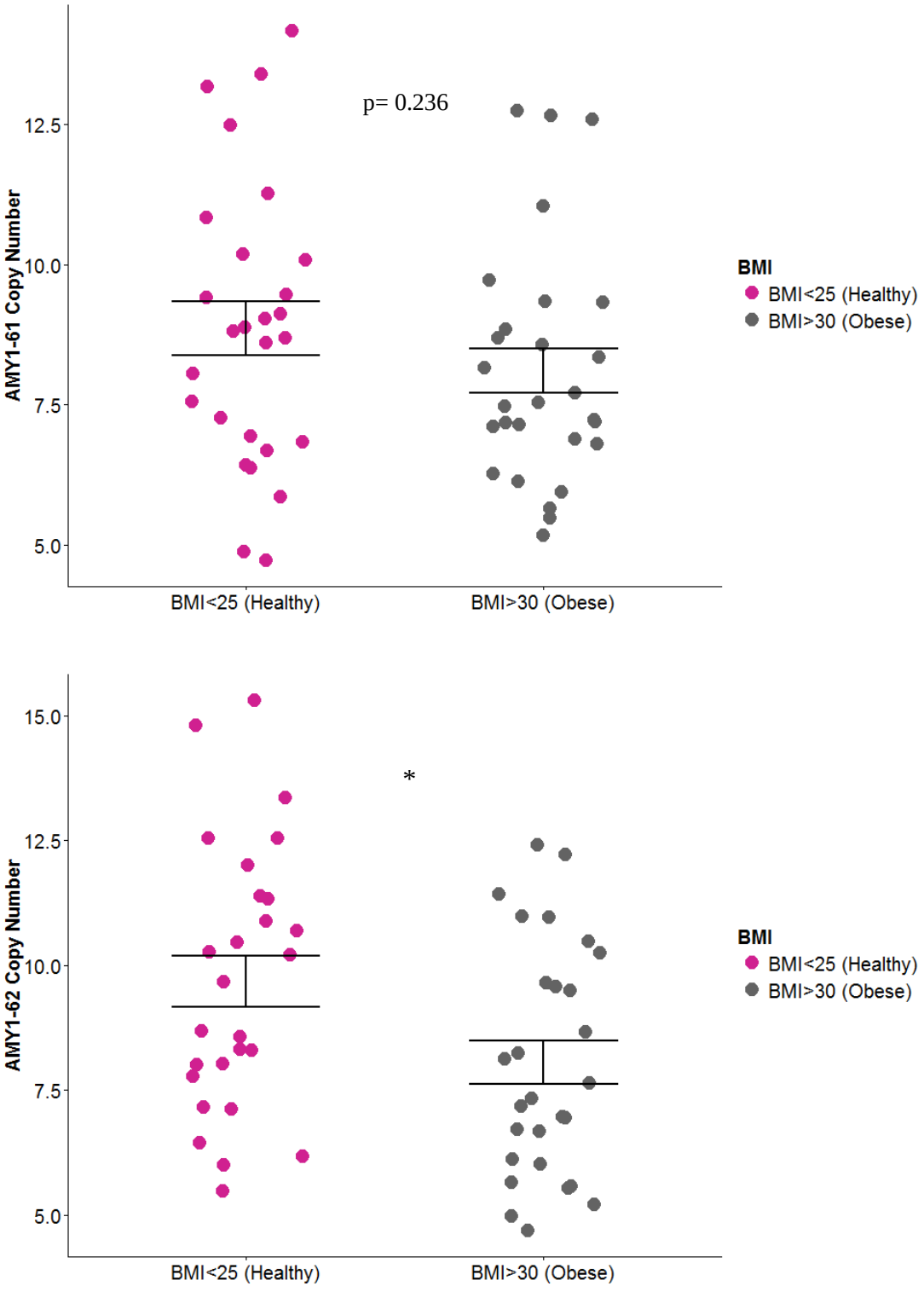



**Figure 3.1** *AMY1* copy number estimates for healthy (n= 27; pink) and obese (n= 28; grey) patients determined using quantitative PCR. Error bars represent the standard error of mean of each population. Two assays were used to determine sample copy number; Hs07226361_cn (AMY1-61) and Hs07226362_cn (AMY1-62). Only AMY1-62 showed significant difference between healthy and obese. *, P-values < 0.05

Restriction digestion is a required process when using digital PCR to determine genomic copy number. When investigating the digested DNA samples (n=3) using qPCR, the resulting *AMY1* copy numbers were more similar to the copy numbers calculated following digital PCR (p= 0.89), than between the undigested qPCR and digital PCR *AMY1* copy numbers mentioned above. The comparison of digested and undigested patient DNA sample copy numbers (p= 0.56), determined when using qPCR, suggested that digestion may affect copy number calling.

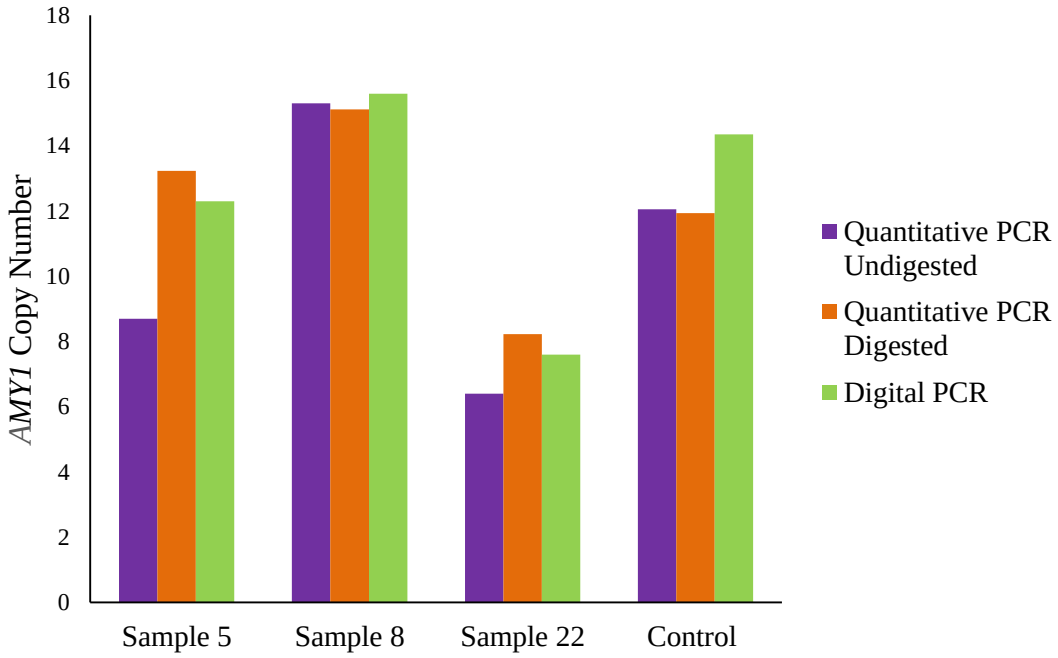



**Figure 3.2** *AMY1* copy number determined using quantitative PCR (undigested and digested genomic DNA) and digital PCR for a subset of breast cancer patient samples (n=3).

## 3.4 Discussion

In the present study, it was confirmed that lower germline *AMY1* copy number is associated with increased BMI in breast cancer patients. A significant inverse

correlation between *AMY1* copy number and BMI was observed in obese compared to healthy breast cancer patients. It was also established that both methodological process and the number of *AMY1* duplicates per genome may influence the calling accuracy of germline *AMY1* copy number in purified DNA samples from breast cancer patients. Using restriction digestion to individualise each *AMY1* duplicate for copy number determination using quantitative PCR (qPCR) is not currently recognised as a necessary process, yet when applied, was seen to alter the calculated copy number in samples with lower range *AMY1* duplication.

The current analysis has confirmed the inverse correlation between germline *AMY1* copy number and risk toward accumulation of excess fat, initially reported by Falchi *et al*.[8]. Additional studies investigating differing cohorts around the world have managed to replicate this reported inverse association[139, 140]. Yet despite the supporting evidence for this inverse association, the underlying physiological and/or metabolic functions that explain the apparent fitness advantage linked to the increased duplication of *AMY1* copy number is yet to be identified. Researchers have confirmed that variation in *AMY1* copy number is positively correlated with dosage of *AMY1* mRNA, salivary amylase protein and salivary amylase enzymatic activity rate[7, 137].

Based on the current findings, it would be logical to expect high *AMY1* copy number resulting in higher blood glucose levels, compared with low *AMY1* copy number, after an intake of starch. Based on this, many studies have focused on the ability of variation in salivary amylase to manipulate the glycemic load experienced by subjects after the ingestion of starch rich meals. The glycemic load is an estimate of how much a particular food would increase a person's blood glucose levels after consumption, in which lower GI foods are considered favourable to overall health. An informative *in vivo* study released starch samples straight into the small intestine, missing out the oral digestion process, and observed these subjects had substantially decreased starch digestion and glycemic load[145]. Interestingly however, when assessed in subjects with either high or low salivary amylase copies and therefore high or low enzymatic concentration and activity rate, the opposite effect was observed. Subjects with low salivary amylase concentrations had substantially higher blood glucose levels following starch ingestion[138], which was attributable to low salivary amylase lessening the pre-absorptive insulin release. Additionally, a clinical experiment inhibited the enzymatic activity of salivary amylase and therefore slowed down the digestion and absorption of

dietary starch and observed a reduced spike in blood glucose after a carbohydrate rich meal, favouring the expected response. However, also observed was a decrease in weight over a 12 week period of salivary amylase inhibition (on average, 2.9 kg), disputing the existing theory[146].

It has been suggested that the functional link between salivary amylase and increased weight gain may be connected with food perception, in which increased and/or reduced breakdown of starch influences the perceived sensation of oral food texture and flavour. When increasing concentrations of the amylase enzyme were added to custards, subjects reported reduced "food creaminess", "creamy mouth after sensations" and "flavour release", all of which are believed to be favourably desired food qualities[147, 148]. These results perhaps suggest that reduced amylase dosage attributable to lower *AMY1* copy number has the ability to positively influence one's liking of starchy foods. Based on this it could be hypothesised that individuals with lower amylase copy number are likely experiencing enhanced desirable sensations during oral ingestion of starch, and thus perhaps strengthening the chances of augmented dietary starch ingestion, compared to those with higher germline *AMY1* copy numbers. This theory would agree with the current study's observation where, in breast cancer patients, increased function of salivary amylase is correlated with decreased BMI.

The lack of validation for the functional link between salivary amylase dosage and increased BMI, despite the support for the inverse correlation, could be attributable to the variable study cohorts in which the inverse association has been investigated. Although the current study has observed a definite trend toward reduced *AMY1* dosage in obese breast cancer patients, only one of the assays used in the analysis reported a significant inverse relationship. The current lack of significance is possibly supported by a recent paper noting they observed no association between dosage of *AMY1* copy number and risk toward the accumulation of excess fat, in a large, multi-ethnic obese and lean cohort[142]. They suggest that the inconsistency between their findings and what has been previously reported is due to their use of higher resolution techniques, such as droplet digital PCR, to determine amylase copy number. Yet, of the previous publications reporting an inverse association between *AMY1* copy number and obesity, two utilised digital PCR to determine *AMY1* copy number within subject DNA samples[139, 140]. The current analysis comparing the use of qPCR and digital PCR to calculate *AMY1* copy number in germline DNA samples of breast cancer patients

suggests that digital PCR is calling marginally higher copy numbers compared to qPCR. Furthermore, the current evaluation has also considered that the lack of genomic DNA restriction digestion, that is required for copy number calculations using digital PCR, may be influencing the copy numbers that are calculated using qPCR. Overall, the comparison of quantitative PCR and digital PCR suggests that the germline *AMY1* copy number determined for an individual may vary depending partially on the experimental techniques that are used. Although contradictory, the findings by Usher *et al.*, validate the importance of investigating *AMY1* copy number, and its association with obesity risk, in breast cancer patients. Based on their findings, it cannot be assumed that obesity will be associated with low amylase copy number in different study cohorts.

A limitation of the current analysis comparing qPCR and digital PCR germline *AMY1* copy number determination is the small sample size being investigated (n=3). The differences seen during this analysis would need to be investigated in a much larger cohort of germline DNA samples if they were to be validated.

The current study supports a novel inverse correlation between *AMY1* copy number and increased BMI in breast cancer patients. This relationship may provide evidence for low *AMY1* copy number as a risk factor toward breast cancer development, through its association with obesity, but this inverse correlation still requires further validation. Firstly, the inverse association would need to be validated in a large cohort case-control investigation, and secondly, the fundamental functional connection between low salivary amylase dosage and increased fat accumulation would not only need to be identified but also confirmed to be biologically active in obese breast cancer patients. One thing is for certain, this is just the beginning of the *AMY1* copy number and obesity debate, and this intriguing human loci has a bright and excitingly unpredictable future.

# Chapter 4

# 4 Obese Breast Tumour Gene Expression Analysis

## 4.1 Introduction

The exact mechanisms responsible for causing the strong link between obesity and breast cancer are yet to be identified, despite extensive research that has provided confirmation for the strong association between them.

Currently, little research has been carried out surrounding the gene expression profiles of breast tumour genes in patients with differing BMI status, and only two studies have developed transcriptomic profiles for breast tumours from obese patients[118, 119]. Creighton *et al.* compared breast tumour gene expression patterns (n=103) from normal (BMI < 25) and overweight (BMI = 25-30) patients to obese (BMI > 30) tumour transcript patterns, from which they derived 662 genes significantly differentially expressed in obese breast tumours[118]. Similarly, Fuentes-Mattei *et al.* generated comparable transcriptomic data for oestrogen positive obese (BMI > 30) compared to non-obese (BMI < 30) breast tumours. They identified 112 genes, 62 that had significantly increased and 50 that had significantly decreased expression in obese compared to non-obese tumours (n=137)[119].

Although these intriguing findings suggest a potentially important relationship between gene expression patterns during breast tumourigenesis and the obese phenotype, further research is required with a well characterised cohort of breast cancer cases. The current study hypothesises that obese breast cancers are associated with differential expression of selected candidate gene markers.

## 4.2 Experimental Design and Research Aims

### 4.2.1 Identifying Candidate Genes

To identify genes associated with breast tumours from obese women, two publically available microarray expression datasets from previous publications were interrogated[118, 119]. The aim was to determine which genes differently regulated in obese breast tumours appear in both gene expression datasets (section 2.4).

Genes shared between the two datasets were further investigated using The Cancer Genome Atlas (TCGA) online repository of gene expression data from 825 breast tumours[114]. To determine which clinically important tumour phenotypes are associated with expression differences of these candidate genes, expression data were compared between clinically important breast tumour phenotypes (ER+ *vs* ER-, PR+ *vs* PR- and HER2+ *vs* HER2-).

### 4.2.2 mRNA and Protein Expression Analysis of Candidate Genes

This study analysed 40 post-menopausal fresh/frozen breast tumour samples collected from The Cancer Society Tissue Bank (n=18, BMI < 25; n=22, BMI > 30) (see Table 2.1 for patient clinicopathological data). The tumour samples were homogenised, RNA was collected, purified and reverse transcribed into complementary DNA (cDNA). Finally, quantitative PCR (qPCR) was used to determine the level of candidate gene expression in the breast tumours (section 2.5.2). Candidate gene expression was compared between obese and healthy breast tumours, with the aim of determining if the level of expression differed between these two BMI groups.

After RNA extraction, protein was also purified from 36 tumour homogenates. Protein expression levels of GRIA2 and DUSP4 were analysed using Western blotting to establish if the trend in *GRIA2* and *DUSP4* gene expression between the healthy and obese breast tumours was also evident at the protein level in the same tumours (section 2.7).

## 4.3 Results

### 4.3.1 Identifying Candidate Genes : Bioinformatic Analysis of Previous Microarray Analyses

Investigation of Creighton *et al.*'s[118] previous microarray dataset identified a total of 91 differentially regulated genes determined by probe IDs. They were presented as fold change in obese relative to healthy weight patients with both oestrogen receptor positive and negative (ER+ and ER-) tumours (Supplementary Table 6.3). Fuentes-Mattei *et al.*'s[119] dataset lists the top 110 most differentially regulated genes determined by probe IDs and presented as log ratio differences in obese compared to non-obese patients with ER+ breast tumours (supplementary Table 6.4).

Initially, the probe IDs mapping to genes that were significantly differentially regulated were compared across the Creighton *et al.*[118] and Fuentes-Mattei *et al.*[119] datasets, to identify which probe IDs were appearing in both datasets. This analysis identified five different probe IDs, three mapping to *GRIA2, DUSP4* and *NR2F1* (205358_at; 204014_at; 209505_at) and two mapping to *ADH1B* (209612_s_at; 209613_s_at). Overall, the four genes that intersected both datasets and were identified as commonly dysregulated in obese compared to healthy breast tumours were *GRIA2, DUSP4, NR2F1* and *ADH1B* (Figure 4.1).

Both microarray datasets report that *GRIA2, DUSP4* and *NR2F1* are significantly downregulated in obese breast tumours, however, the direction of regulation of *ADH1B* is controversial (Table 4.1). Creighton *et al.*[118] report *ADH1B* to be downregulated in obese tumours but Fuentes-Mattei *et al.*[119] suggest *ADH1B* was upregulated in breast tumours from obese patients.

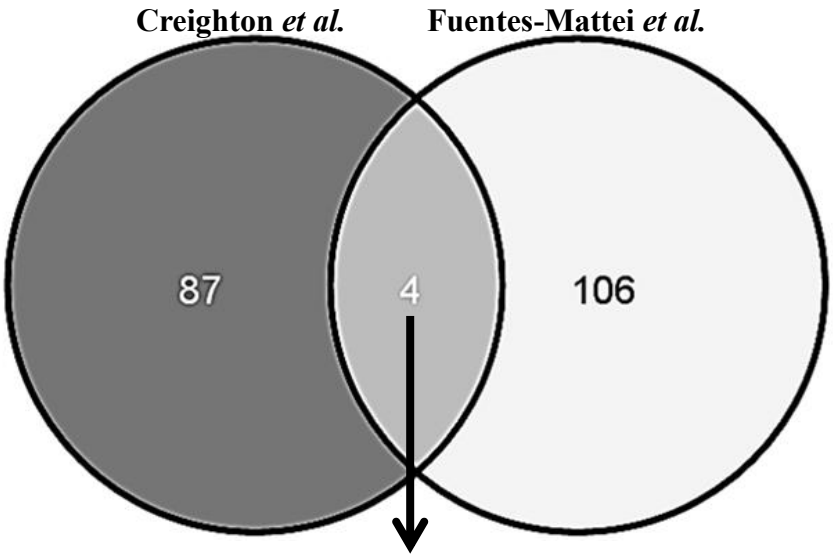



**Figure 4.1** Venn diagram of statistically differentially expressed genes shared between the Creighton *et al.*[118] and the Fuentes-Mattei *et al.*[119] breast tumour microarray expression datasets.

**Table 4.1** Differential expression data for the four genes reported to be significantly differentially expressed in breast tumours from obese women compared to healthy women by both Creighton *et al.*[118] (BMI < 25 vs BMI > 30) and Fuentes-Mattei *et al.*[119] (BMI < 30 vs BMI > 30).

| | **Creighton *et al.*** | | **Fuentes-Mattei *et al.*** | |
|---|---|---|---|---|
| **Genes** | **Fold Change (Log2)** | **P-Value** | **Fold Change (Log10)** | **P-Value** |
| *GRIA2* | -2.39 | 0.0082 | -0.31 | <0.001 |
| *DUSP4* | -1.80 | 0.0108 | -0.83 | <0.001 |
| *NR2F1* | -1.40 | 0.0322 | -0.31 | 0.004 |
| *ADH1B* | -2.30 | 0.0126 | 0.51 | <0.001 |

Based on their appearance in both microarray datasets these four significantly differentially regulated genes were selected to be further analysed as the candidate genes.

Gene expression data from The Cancer Genome Atlas (TCGA)[114] (http://www.cbio-portal.org/) was utilised to perform an analysis of the candidate gene markers in a large cohort of breast cancer patients. TCGA provides extensive gene expression and clinicopathological data for breast cancer tumours derived from 825 patients, including breast tumour ER (oestrogen receptor), PR (progesterone receptor) and HER2 (*ERBB2* expression) status. BMI information for patient tumour data in TCGA has not been recorded, thus no comparison could be made between breast tumours from obese and non-obese patients. The change in regulation of the four common genes was compared between the 825 breast cancer tumours based on presence or absence of the clinical biomarkers ER, PR and HER2.

Analysis suggests that candidate genes generally displayed significant differences in expression between ER(+/-), PR(+/-) and HER2(+/-) breast tumours. All genes showed statistically significant ($p < 0.05$) higher expression in ER+ breast tumours (Figure 4.2), while only *NR2F1* was not highly statistically significant ($p < 0.001$) in PR positive breast tumours (Figure 4.2 and 4.3). However, within the HER2 phenotype *GRIA2* and *ADH1B* are on average more highly expressed in HER2+, *NR2F1* more lowly expressed in HER2+, and *DUSP4* indicated no difference between HER2(+/-) breast tumours (Figure 4.4).

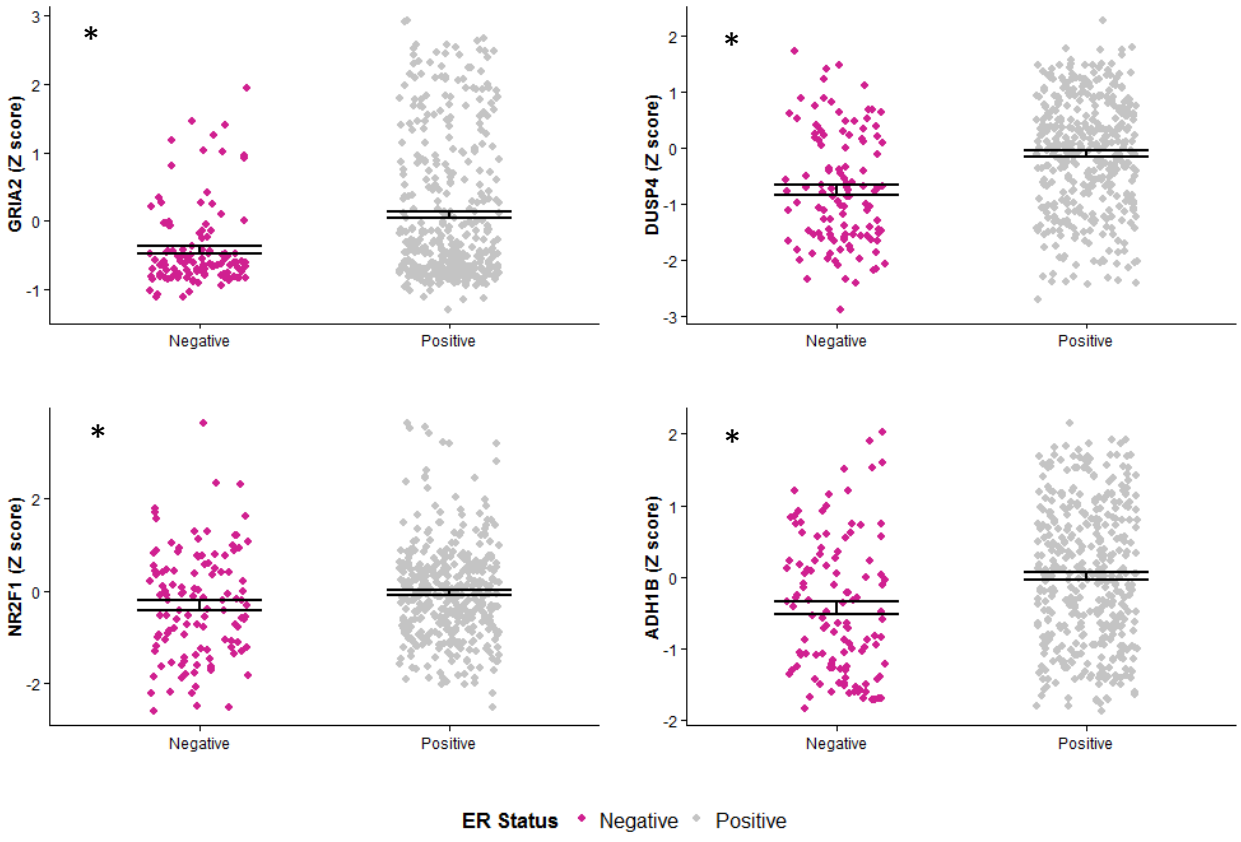



**Figure 4.2** Boxplots showing the difference in regulation of *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* between TCGA patients that have ER positive (n= 403) and ER negative (n= 118) breast cancer tumours. Error bars represent the standard error of the mean for each BMI cohort. *P-values < 0.05 are statistically significant differences in gene expression between ER positive and ER negative tumours (n=521).

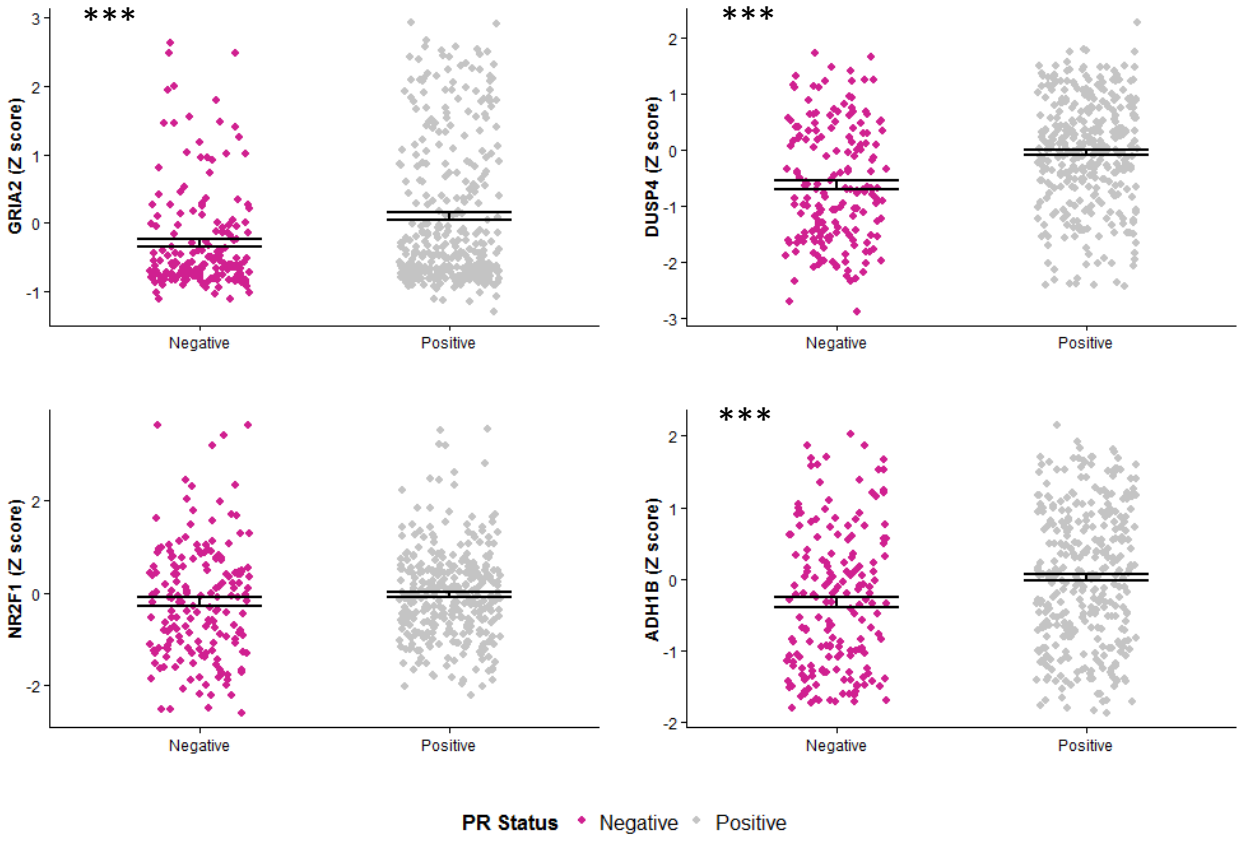



**Figure 4.3** Boxplots showing the difference in regulation of *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* between TCGA patients that have either PR positive (n= 341) or PR negative (n= 179) breast tumours. Error bars represent the standard error of the mean for each BMI cohort. ***P-values < 0.001 are statistically significant differences in gene expression between PR positive and PR negative tumours (n=520).

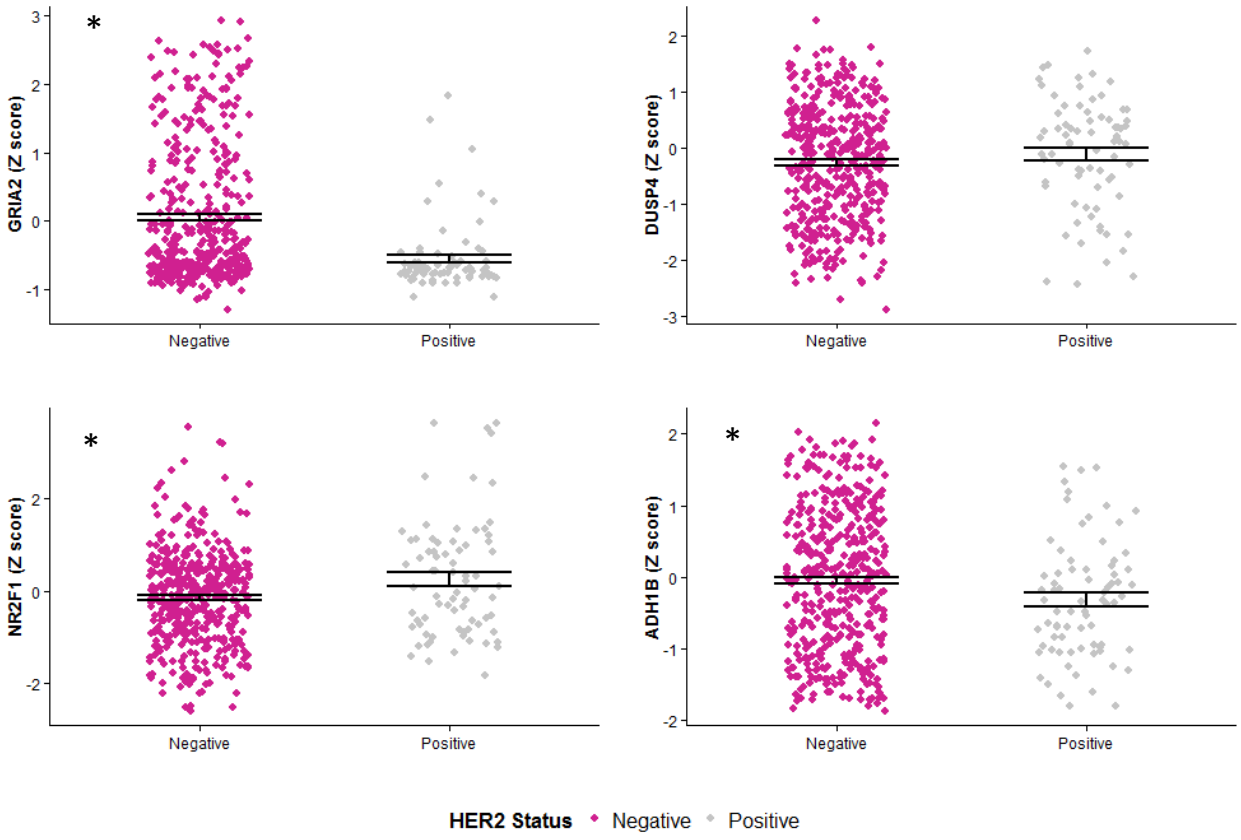



**Figure 4.4** Boxplots showing the difference in regulation of *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* between TCGA patients that have HER2 positive (n= 75) and HER2 negative (n= 433) breast cancer tumours. Error bars represent the standard error of the mean for each BMI cohort. *P-values < 0.05 are statistically significant differences in gene expression between HER2 positive and HER2 negative tumours (n=509).

### 4.3.2 Candidate Gene Expression in Breast Tumours from Obese and Healthy Weight Women

#### *4.3.2.1* *GRIA2, DUSP4, NR2F1* and *ADH1B* mRNA expression in obese breast tumours

All four of the candidate genes (*GRIA2, DUSP4, NR2F1* and *ADH1B*) exhibited decreased mean expression in obese relative to healthy weight breast tumours, although these differences were not statistically significant (Figure 4.5). The largest difference in expression was seen in *ADH1B*, expressing at least approximately 5X less in obese compared to healthy breast tumours. *GRIA2* had the smallest difference in expression in obese relative to healthy tumours, with half the amount in obese tumours. The healthy cohort had greater variation in candidate gene expression, whereas the obese tumours had more tightly clustered candidate gene expression. The coefficient of variation for *GRIA2, DUSP4, NR2F1* and *ADH1B* expression in the healthy group was 3.09, 2.36, 1.45 and 3.28, respectively. Similarly, the coefficient of variation for *GRIA2, DUSP4, NR2F1* and *ADH1B* expression for the obese group was 2.28, 0.98, 0.99 and 1.10, respectively. The increased variation in candidate gene expression within the healthy cohort may be explained by the unusually high expression of one influential tumour sample (Figure 4.6). The same influential tumour sample had unexpectedly high *GRIA2* and *DUSP4* expression. Similarly, a second tumour sample was expressed at an unexpectedly high level for *NR2F1* and *ADH1B*, influencing the analyses of these genes. Interestingly, when these influential points were removed from the analysis and fold change values re-calculated, the variation within the healthy cohort was substantially reduced across all four candidate genes (Supplementary Figure 6.4). However, the differences between the obese and healthy cohorts were still not statistically significant. The fold change in *DUSP4, NR2F1* and *ADH1B* expression for the obese group relative to the healthy group was also reduced by 0.40, 0.14 and 0.52, respectively, but remained downregulated in tumours from obese women. In contrast, the difference in *GRIA2* expression for obese relative to healthy tumours was increased, and was also reversed; *GRIA2* was being upregulated in obese breast tumours. Nevertheless, the influential values were not removed from the final analysis, as they were real expression values for these breast tumour samples.

Candidate gene expression was also compared within pathological features of the patients' breast tumours (Table 4.2). The expression of *GRIA2, DUSP4, NR2F1* and

*ADH1B* followed the same trends of expression between these pathological features as was seen in the analysis using TCGA data. However, only *GRIA2* showed a statistically significant association with ER+/- and PR+/- tumours. Lastly, the pathological status (ER, PR and HER2) of the tumours was compared to the patients' BMI (Table 4.3). There was no significant difference in the mean BMI scores between patients with ER, PR or HER2, positive and negative tumours.

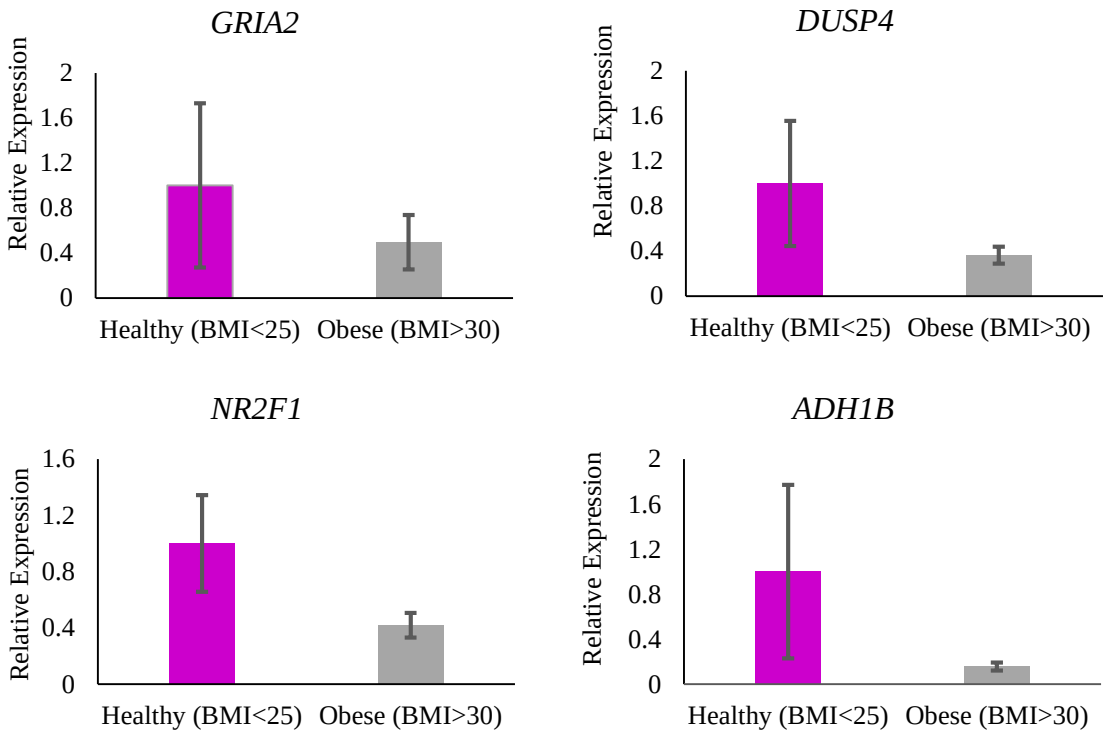



**Figure 4.5** The relative expression in obese (BMI > 30; n= 22) compared to the healthy (BMI < 25; n= 18) breast tumour cohorts for *GRIA2* (p= 0.52), *DUSP4* (p= 0.27), *NR2F1* (p= 0.12) and *ADH1B* (p= 0.29) expression. The healthy tumours are the pink bars and the obese the grey bars. The error bars are the standard error of the mean.

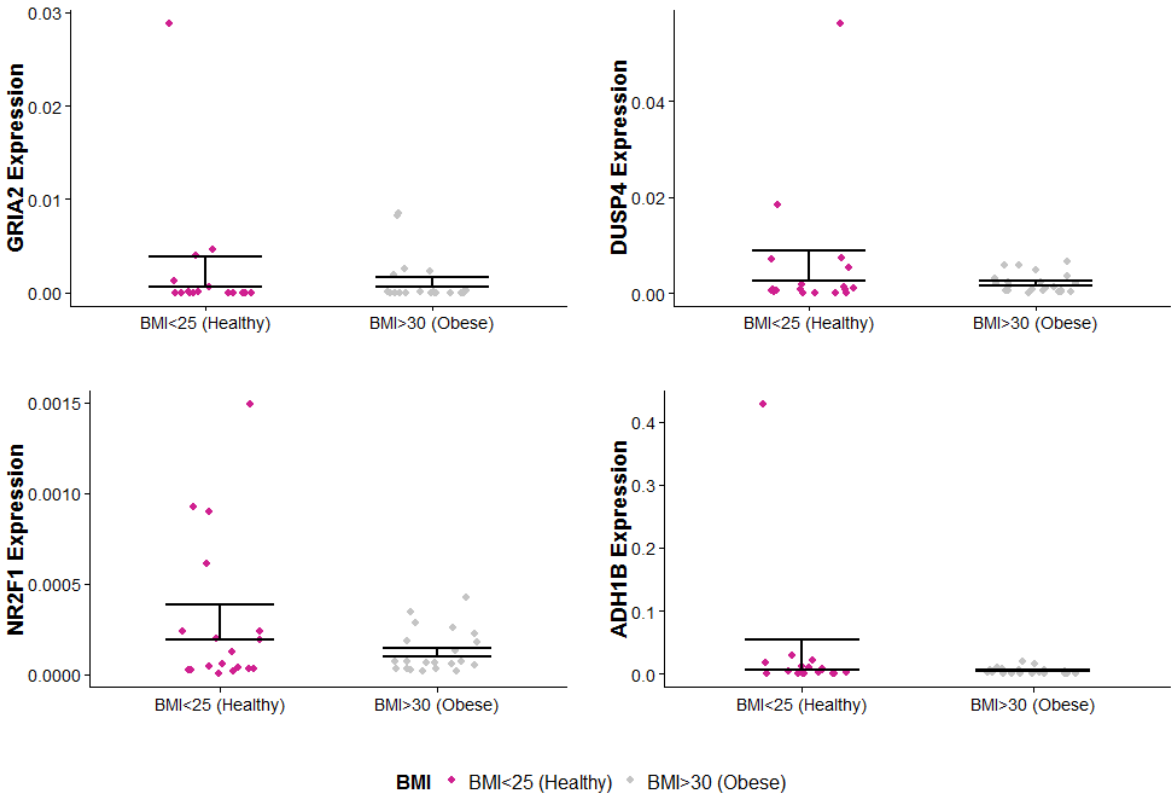



**Figure 4.6** The difference in gene expression (target/reference), for *GRIA2* (p= 0.52), *DUSP4* (p= 0.27), *NR2F1* (p= 0.12) and *ADH1B* (p= 0.29) in breast tumours (n=40). The target/reference ratios are the level of target gene expression normalised to the reference gene (*EFF1A1*) in that sample. Each circle represents an individual tumour sample, the pink cluster the healthy weight tumours and the grey, tumours from obese patients. The errors bars are the standard error of the mean for each BMI cohort.

**Table 4.2** Analysis of candidate gene expression within the pathological features of the patients' breast tumours.

| | Pathology Features | ER- (n= 11) | ER+ (n= 28) | PR- (n= 13) | PR+ (n=25) | HER2- (n= 22) | HER2+ (n= 9) |
|---|---|---|---|---|---|---|---|
| ***GRIA2*** | Fold Change[1] | 0.01 | | 0.005 | | 270.194 | |
| | P-Value | 0.047* | | 0.047* | | 0.054 | |
| ***DUSP4*** | Fold Change[1] | 0.185 | | 0.221 | | 3.681 | |
| | P-Value | 0.06 | | 0.091 | | 0.183 | |
| ***NR2F1*** | Fold Change[1] | 0.433 | | 0.394 | | 0.979 | |
| | P-Value | 0.066 | | 0.051 | | 0.960 | |
| ***ADH1B*** | Fold Change[1] | 0.180 | | 0.188 | | 1.492 | |
| | P-Value | 0.252 | | 0.275 | | 0.431 | |

[1]Difference in candidate gene expression in negative compared to positive breast tumours
*Significant difference in candidate gene expression ($P < 0.05$)

**Table 4.3** Comparison of the patients BMI within the pathological features of the breast tumours.

| Pathological Feature | | Sample Size | Average BMI | P-value |
|---|---|---|---|---|
| **ER Status** | Negative | 11 | 28.4 | 0.96 |
| | Positive | 28 | 28.3 | |
| **PR Status** | Negative | 13 | 28.8 | 0.93 |
| | Positive | 25 | 28.6 | |
| **HER2 Status** | Negative | 22 | 29.7 | 0.88 |
| | Positive | 9 | 29.2 | |

#### 4.3.2.2 DUSP4 and GRIA2 protein expression in obese breast tumours

GRIA2 and DUSP4 were chosen from the candidate genes as they have quality antibodies that are well validated, and are available for human sample analysis. Additionally, DUSP4 is well supported in the literature to be associated with breast cancer. The DUSP4 protein expression correlated with *DUSP4* mRNA expression in obese compared to healthy breast tumours (Figure 4.7). Densitometry analysis of DUSP4 protein expression exhibited, on average, a trend toward lower levels of DUSP4 in obese (mean= 0.37) compared to healthy (mean= 1.60) breast tumours; however, this difference was not statistically significant (Figure 4.7 (B)).

The healthy cohort of breast tumours displayed greater variation, than the obese group, in DUSP4 protein expression. When probed for β-actin (the loading control) tumour sample 1 belonging to a healthy weight patient consistently expressed β-actin at levels approximately 300x less than the average expression of β-actin in the healthy tumour cohort. However, the same breast tumour had DUSP4 levels that were only 10% lower than the average DUSP4 expression of this study cohort (Figure 4.8). Thus, after densitometry analysis normalising DUSP4 expression to its relative β-actin, sample 1 had an atypically high level of DUSP4 expression, causing a significant increase in the standard error of the mean in the healthy weight group. Interestingly, sample 1 was not the same tumour specimen that was expressing unusually high levels of *DUSP4* mRNA in the previous analysis (Figure 4.6); instead, sample 1 expressed relatively average levels of *DUSP4* mRNA compared with the healthy tumour cohort. Thus, all of the tumour samples remained in the analysis of DUSP4 protein expression.

Due to the low, barely detectable GRIA2 protein expression in the same tumour samples, expression data could not be derived for GRIA2 through Western blotting. Optimisation included increasing primary antibody concentration from 1/2000 to 1/500, increasing the total protein load from 10 µg to 40 µg per sample and re-wetting the membrane with methanol in case of accidental drying out during the transfer process. Despite making these changes to the original protocol, the GRIA2 bands were faint, and densitometry analysis was impeded by background (Supplementary Figure 6.5).

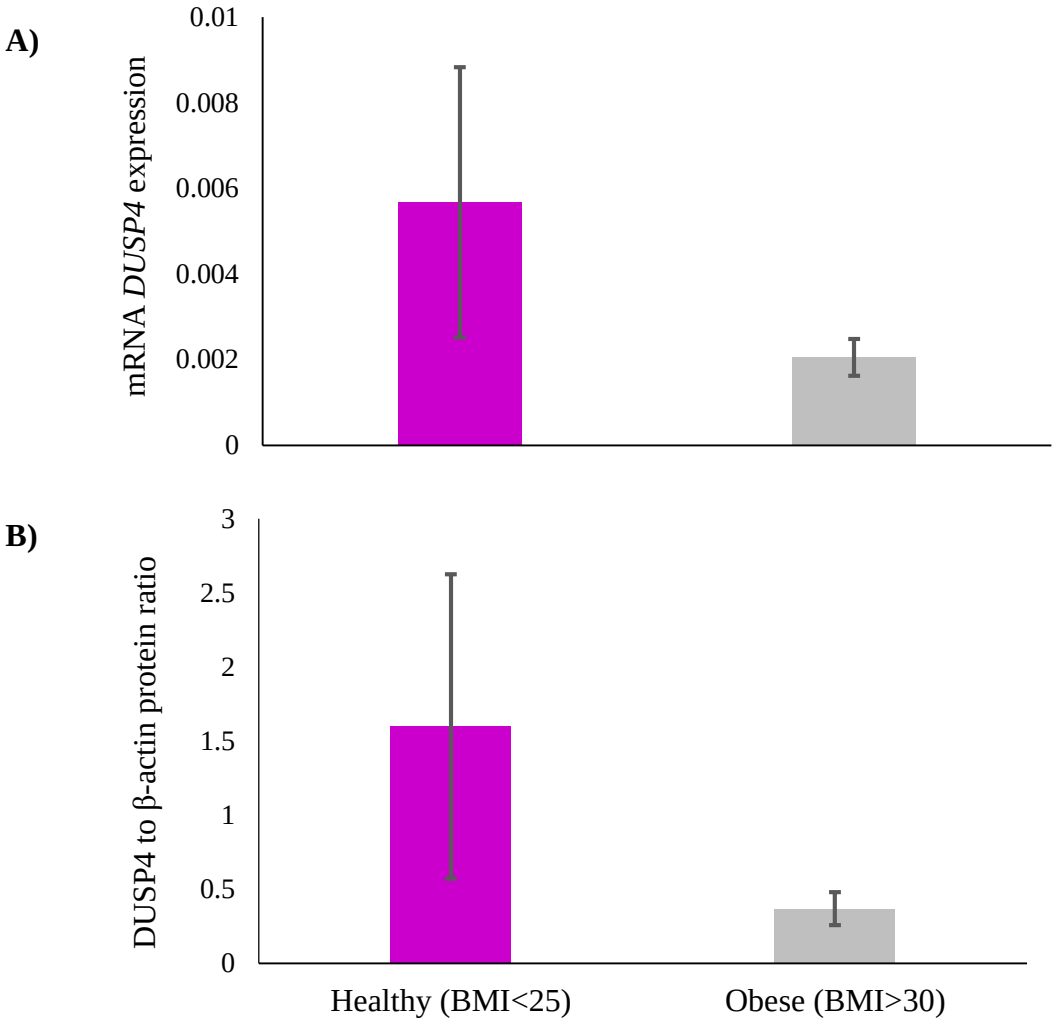



**Figure 4.7** A) The average difference in gene expression (target/reference), for *DUSP4*, in breast tumours from healthy (n= 18, pink) and obese (n= 22, grey) women (p= 0.27). The target/reference ratios are the level of target gene expression normalised to the reference gene (*EFF1A1*) in that sample. Error bars are the standard errors of the mean. B) The protein average expression of DUSP4 in obese (n= 20, grey) and healthy (n=16, pink) weight breast tumour samples (p= 0.25). Error bars are the standard errors of the mean.

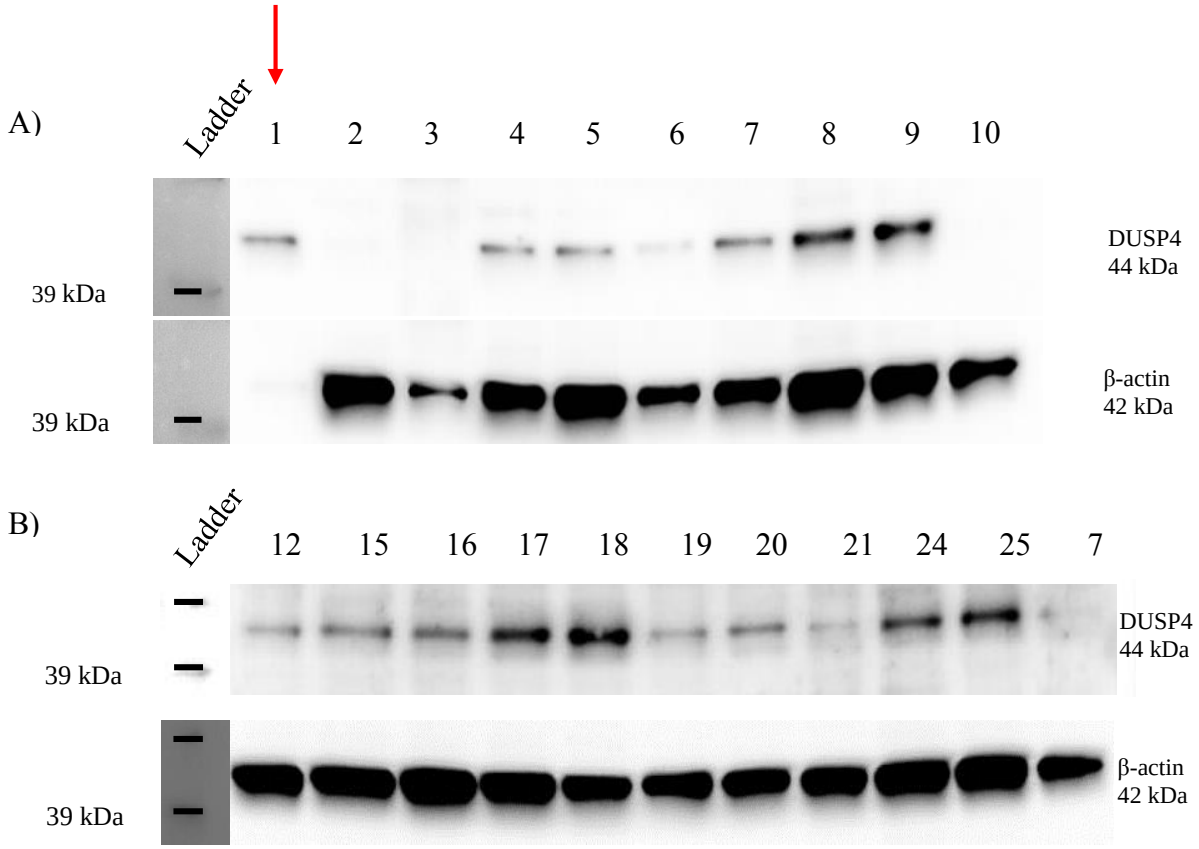



**Figure 4.8** Representative Western blots of DUSP4 with 3 minute exposures and 10 µg of total protein loaded per well. A) Tumour samples: 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. Red arrow- Sample 1 has low expression of β-actin and normal DUSP4 expression. B) Tumour samples: 12, 15, 16, 17, 18, 19, 20, 21, 24, 25, and 7. Sample 7 was used as a positive control to normalise across gels.

## 4.4 Discussion

The current analysis of candidate gene expression in breast tumours from obese women and women with a healthy weight determined that *DUSP4* was trending towards downregulation in obese compared to healthy breast tumours, and this correlated with DUSP4 protein expression in these tumours. Similarly, this study revealed that *GRIA2, NR2F1* and *ADH1B* mRNA showed a trend toward downregulation in obese breast tumours. However, no statistically significant differences were observed with mRNA and DUSP4 protein expression in the current study cohort of breast tumours from obese and healthy weight women. Despite using different techniques for resolving mRNA expression, the overall decrease observed for *GRIA2, DUSP4,* and *NR2F1* in tumours from obese women is supported by the previous microarray analyses. Similarly, the lower expression levels of *ADH1B* in breast tumours from obese women agrees with the earlier of the two genetic signatures[118, 119].

Initially, this chapter explored current gene expression profiles of obese breast tumours in order to identify and examine candidate genes that are differentially regulated in tumours derived from obese compared to healthy breast cancer patients. Considering that the underlying biological mechanisms responsible for the link between obesity and breast cancer are yet to be identified, it is surprising that only two previous analyses have reported transcriptomic profiles of obese breast tumours[118, 119]. A comparison of datasets derived from each of these studies identified four obesity associated candidate genes (*GRIA2*, *DUSP4*, *NR2F1* and *ADH1B*) that were significantly differentially expressed in breast tumours (Figure 4.1). The *GRIA2* (glutamate receptor, ionotropic, AMPA 2) gene codes for one of four glutamate receptor subunits (GRIA1-4) collectively forming a ligand-activated cation channel predominantly found in the mammalian brain[149]. The DUSP4 enzyme, encoded by *DUSP4* (dual specificity protein phosphatase 4), is localised to the cell nucleus and inactivates ERK1, ERK2 and JNK via phosphorylation[150]. *NR2F1* (nuclear receptor subfamily 2, group F, member 1), also known as COUP-TFI (COUP transcription factor 1), encodes a steroid hormone receptor protein, in which the targets of this transcriptional factor have not yet been fully identified[151]. Lastly, the ADH1B enzyme, encoded by *ADH1B* (alcohol dehydrogenase 1B), functions with high activity for ethanol oxidation and plays a main role in the catabolism of ethanol[152].

*GRIA2, DUSP4, NR2F1* and *ADH1B* are upregulated in ER+ and PR+ breast tumours, but show unique trends of up or downregulation within HER2+/- breast tumours (Figures 4.2-4.4). These results are somewhat supported by previous research that observed *GRIA2* expression to be associated with breast cancer through its co-clustering with the expression of *ESR1* (oestrogen receptor 1) in primary invasive breast tumour biopsies, expressing at higher levels in ER+ breast tumours[116]. However, other studies have not reported any statistical or molecular association between *GRIA2* expression and the ER, PR and HER2 status of breast tumours. Overall this investigation suggests that *GRIA2, DUSP2, NR2F1* and *ADH1B* may represent novel co-expressing ER, PR and/or HER2 related genes, but further investigation of their expression is required to determine whether these genes have the ability to aid in clinical applications.

The current study has shown that downregulation of *DUSP4* correlates with reduced levels of DUSP4 protein in obese breast tumours. Assuming that lower DUSP4 protein concentrations in the obese breast tumours are paralleled with a reduced rate in its enzymatic activity, it is reasonable to expect *DUSP4* mRNA to be expressed at lower levels in obese breast tumours. DUSP4 inactivates ERK1, ERK2 and JNK via phosphorylation[150]. Increased levels of JNK have been associated with the risk of obesity, obesity induced hyperinsulinemia, and more importantly, increased anti-apoptotic signalling[153-155]. Similarly, ERK1 and ERK2 are involved in a mitogen-activated protein kinase (MAPK) signalling cascade that is significantly upregulated in many different human cancers in which signalling promotes cellular migration, survival, proliferation and cell cycle progression[155, 156]. Additionally, highly aggressive basal-like breast cancers have been associated with a reduction in *DUSP4* mRNA expression. Evidence suggests that deficiencies in *DUSP4* regulate the resistance to anti-tumour chemotherapies in these aggressive breast tumour subtypes by promoting interleukin-6 and interleukin-8 expression to generate a cancer stem cell phenotype[150, 157]. Supplementary Figure 6.6 depicts a schematic of the *DUSP4* pathway model proposed by Balko *et al.*, in basal-like breast cancers[150]. Based on this, I hypothesise that downregulation of *DUSP4,* observed in breast tumours from obese women, is analogous to that shown in basal-like breast cancers, where increasing levels of ERK1, ERK2 and JNK, promote cellular mitosis and supress stress induced apoptosis. Such a

molecular phenotype in breast tumours from obese patients may contribute to poorer prognosis and poorer treatment responses.

Of all the candidate genes analysed in this study, *DUSP4* is the only gene consistently recognised in the literature for its association with breast cancer. However, the literature surrounding the association of *DUSP4* expression with breast tumourigenesis is still controversial. In support of the current study's findings, some researchers have provided evidence for increases in *DUSP4* acting in the suppression of tumour progression and metastasis[158-160], whereas opposing evidence suggests that higher *DUSP4* expression is playing a role in promoting breast cancer growth and proliferation[161-163]. These previous examinations of *DUSP4* expression have investigated differences only between malignant and benign breast tissues, thus the current study is the first to directly measure *DUSP4* mRNA in breast tumours from healthy and obese patients.

Aside from evidence proposing that adipokines are involved in the upstream regulation influencing the development of the overall obesity associated genetic signature[119], there is no current evidence describing the biological systems causing *DUSP4* to be uniquely downregulated in obese breast tumours. Investigation of the mechanisms responsible for lower *DUSP4* expression in breast cancer could reveal cellular systems linking obesity and the initiation of more aggressive breast tumours. For breast tumours developing in obese patients, better understanding of these functional mechanisms may elucidate potential targets for future therapies that can counteract the trend toward obese breast tumour *DUSP4* deficiency, observed in the current analysis.

Research investigating expression levels of *GRIA2* in different types of cancer cells, including liver, intestine, brain and uterus, has shown that downregulation of *GRIA2* both promotes and conversely supresses cancer cell replication[164-167], but only one study has provided evidence for an association between breast cancer and *GRIA2* expression[116]. The current study showed that the trend toward lower *GRIA2* expression in breast tumours from obese women was reversed when the single influential sample was removed from the analysis, suggesting that this trend is inconclusive (Figure 4.5, Supplementary Figure 6.4). This uncertainty could imply that the level of *GRIA2* expression in breast tumours is only associated with BMI by chance, or, that its link to obesity in previous genetic signatures[118, 119] is possibly due to its connection with

*ESR1*[116]. The gene expression results from the current study suggest that *GRIA2* was the only candidate gene significantly associated with ER status of the breast tumours (Table 4.2). This is supported by this previous link between *GRIA2* and *ESR1* expression[116]. Furthermore, the obese breast tumour genetic signature determined by Creighton *et al.* did not harbour ER, PR or any other oestrogen controlled genes, but did associate statistically with reduced ER levels[118]. The literature surrounding the association between obesity and ER status in breast cancer is still controversial[168]. This may explain why this study saw no association between BMI and the ER, PR and HER2 status of the breast tumours.

Investigations of ER+ breast cancer cell lines have reported an up regulation of *NR2F1* to be involved in the progression of breast cancer. Increased *NR2F1* expression during breast tumour development is suggested to modulate oestrogen signalling and thus is participating in enhancing tumour growth, invasion and hormone resistance[169, 170]. This is contrary to observations from the current study, in that breast tumours from obese women, that could be expected to have more progressive cancers, were expressing lower levels of *NR2F1* compared to tumours from healthy weight women. However, unlike the previous studies which showed *NR2F1* expression in predominantly ER+ breast cancer cell lines, the current study analysed *NR2F1* expression in both ER+ and ER- breast tumours, making comparison across these studies difficult.

The expression of *ADH1B* is not associated with the initiation or progression of breast cancer. On the other hand, the regulation of *ADH1B*, in another breast tissue (adipose), has been associated with risk factors strongly linked to breast cancer, such as obesity and insulin resistance[171].

Although the previous microarray studies[118, 119], by Creighton *et al.* and Fuentes-Mattei *et al.*, analysed gene expression in similar sized cohorts of pre-treatment breast tumour biopsies, the obesity-signatures from these studies were derived from slightly different tumour cohort pathologies and comparisons. Creighton *et al*. analysed mRNA in both ER positive and negative (ER+/-) tumours, and compared gene expression patterns from healthy (BMI < 25) and overweight (BMI = 25-30) tumours combined, to obese (BMI > 30) tumours[118]. In contrast, Fuentes-Mattei *et al.* analysed mRNA from ER+ breast tumours only, comparing gene expression patterns between healthy (BMI <

25) and obese (BMI > 30)[119]. Despite these inter-study differences conceivably limiting the likelihood of seeing any gene overlap between the studies, comparison of the datasets derived from each of these studies identified four common genes (*GRIA2, DUSP4, NR2F1* and *ADH1B*) differentially expressed in obese breast tumours.

The current study used a relatively small sample size for both mRNA (n=40) and protein (n= 36) expression analysis. If the trend seen in this study, downregulation of expression in breast tumours from obese women, is valid then it is likely that with a larger cohort the trend would become significant. The smaller sample numbers used in this study make the results vulnerable to outliers influencing the statistical analyses. The exclusion of the influential values contributed toward less variation in the healthy cohort and a reversal in *GRIA2* obese breast tumour expression (Supplementary Figure 6.4). However, results obtained during this study were from assays carried out in triplicate, therefore, all of the observed expression data remained in the final analysis.

It is well known that patient obesity is strongly associated with increasing the risk of breast cancer and developing more invasive and metastatic breast tumours, particularly in post-menopausal women[3, 4]. Better understanding of the obese breast tumour transcriptome may help identify genes involved in the underlying mechanisms responsible for these links between obesity and breast cancer. The current study suggests that low *GRIA2, DUSP4, NR2F1* and *ADH1B* expression in breast tumour biopsies may be related to patient obesity. However, these results are inconclusive, and further investigations with larger sample sizes would need to be done to better examine these preliminary findings.

# Chapter 5

# 5 Discussion

Obesity is a strong predictor for both increased prevalence and reduced survival rate in diagnosed breast cancer patients[3, 4]. Yet despite the strong link between obesity and breast cancer, the precise underlying mechanisms responsible for such a strong association are yet to be identified. Therefore, the overall aim of this study was to explore the molecular links between obesity and breast cancer, in order to identify genes potentially influencing obesity associated breast cancer incidence and mortality. This study explored two different forms of genomic changes within 55 breast cancer patients. Initially, a form of germline alteration known as copy number variation (CNV) was investigated in 55 breast cancer patient blood samples. Following this, changes in the level of candidate gene expression from 40 overlapping breast tumour biopsies were explored.

Germline alterations are commonly inherited changes, but can also be *de novo* mutations that arise during development. Advantageously, germline changes can be explored in human blood samples as they are evident in all cells throughout the body. The germline alteration of interest for the current analysis was a CNV at the *AMY1* locus, due to its proposed inversed relationship with obesity[8]. The *AMY1* locus (1p21.1) codes for the salivary α-amylase enzyme, the most abundant enzyme secreted from the salivary glands. Salivary α-amylase begins the oral digestion of large starch molecules, in which are large polymers of glucose monomers. Salivary amylase breaks apart these monomers via hydrolysis of the internal α-1,4 glycosidic bonds. The *AMY1* salivary amylase genes exhibit extensive CNV in the human genome ranging anywhere between 0 and 20 copies per individual. It has been suggested that selection pressure imposed on historical populations, by higher starch consumption, resulted in the evolution of greater *AMY1* copy numbers compared to populations adapted to low starch diets that have evolved lower *AMY1* copy numbers[7].

The aim of the first hypothesis was to identify whether germline *AMY1* copy number has a molecular link with breast cancer through its association with obesity. To test this hypothesis, I investigated the association between *AMY1* copy number and obesity in a cohort of well-characterised breast cancer patients, using quantitative PCR (qPCR). It was observed that low germline *AMY1* copy number was significantly associated with increased BMI in breast cancer patients. These findings are consistent with the previously reported inverse association between obesity and low germline *AMY1* copy number[8, 139, 140]. The results from this investigation suggest that low *AMY1* dosage could indirectly increase the risk of breast cancer, through promoting the development of an obese phenotype.

Salivary amylase is responsible for dietary starch breakdown into simple sugar molecules, as preparation for further digestion in the stomach. Based on this information, I propose that after starch ingestion, people possessing low *AMY1* copy number and therefore low amylase levels, would have lower and more prolonged blood glucose loads compared to individuals with higher amylase copy number (Figure 5.1). As blood glucose increases, so too do insulin concentrations in the blood, as insulin assists the entry of blood glucose into cells such as muscle[172]. Hyperinsulinemia is the extended increase in blood insulin, due to prolonged escalations of glucose. For individuals with low *AMY1* copy number ingesting starch rich diets, the prolonged elevation of blood glucose may be the physiological link to increased risk of obesity. Hyperinsulinemia mimics insulin resistance in the body and is commonly recognised as a symptom of obesity[73], however, the causal relationship between obesity and hyperinsulinemia remains to be resolved. It has been suggested that reducing circulating insulin may act as a preventative and/or treatment measure for mammalian obesity and its associated complications[173]. Based on this information and the current study's findings I hypothesise that low dosage of *AMY1* linking to hyperinsulinemia could initiate a proliferative feedback loop in which hyperinsulinemia is instigated in developing obesity, and obesity then heightens insulin hypersecretion (Figure 5.2). Obesity remains a complex and heterogenic disease. The model of obesity that emerges as a result of low dosage of salivary amylase cannot be assumed to be the only model of obesity that may be responsible for causing the increased incidence and reduced survival rate in diagnosed breast cancer patients.

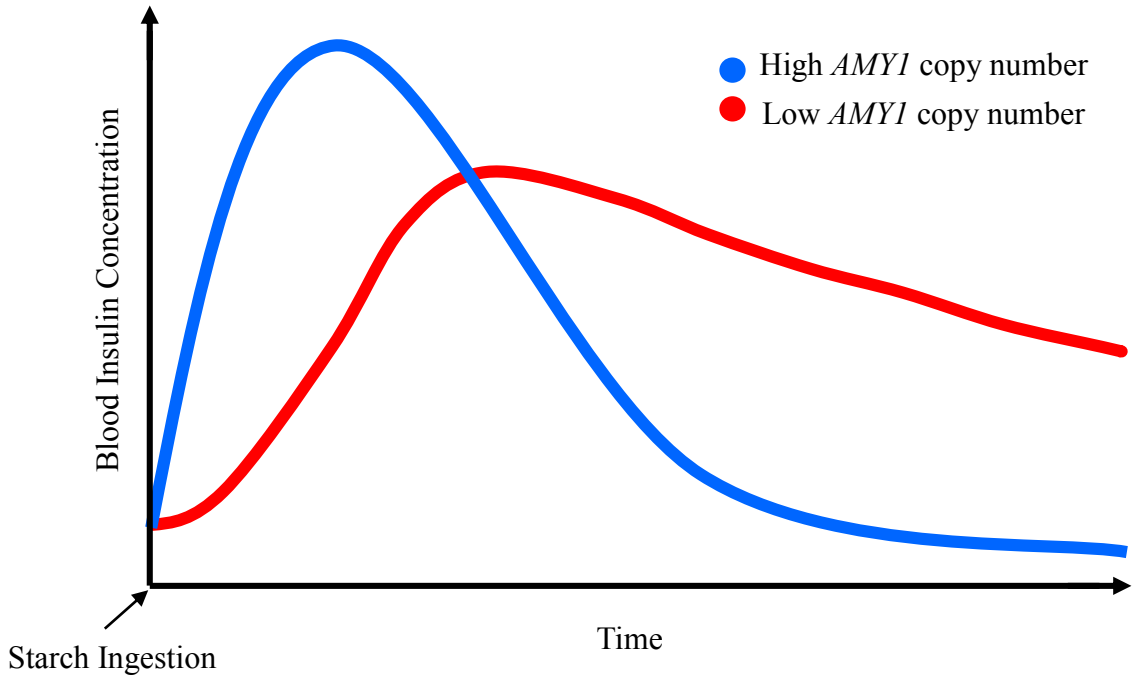



**Figure 5.1** This diagram represents this studies proposed bio-mechanism occurring when people possessing low *AMY1* copy number, and therefore low amylase levels, ingest starch rich meals compared to people with high *AMY1* copy number. People with lower *AMY1* copy number (red line) would have lower and more prolonged blood glucose loads that take longer to return to base-line, compared to individuals with higher *AMY1* copy number (blue line).

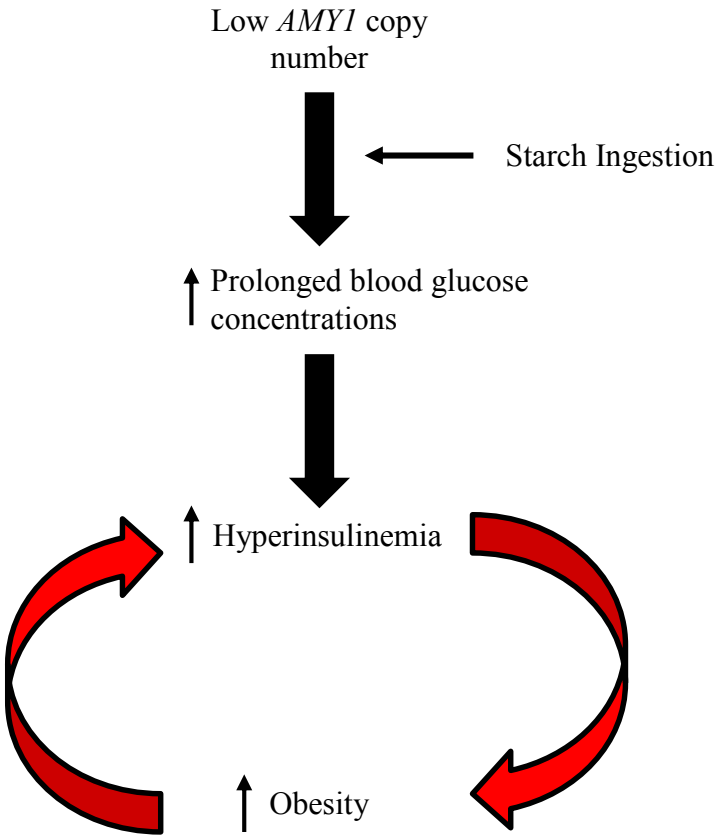



**Figure 5.2** Flow diagram representing the proposed model linking low *AMY1* copy number/dosage to the increased risk of obesity. After the ingestion of a starch rich meal, individuals with lower *AMY1* copy number undergo an increase in blood glucose that is slower and more prolonged than people that have higher *AMY1* copy number. This prolonged blood glucose is followed by elevation in blood insulin concentrations, in which prolonged increase in blood insulin mimics a condition known as hyperinsulinemia. The close link between obesity and hyperinsulinemia makes it hard to determine whether hyperinsulinemia causes obesity or obesity causes hyperinsulinemia, or both. In this model a proliferative feedback loop (red arrows) between hyperinsulinemia and obesity links low *AMY1* copy number and increased risk of obesity.

It is widely appreciated that sustained hyperinsulinemia is considered a possible mechanism involved in obesity associated breast tumour initiation and/or progression[109]. However, the precise way in which obesity associated hyperinsulinemia influences breast tumourigenesis remains unclear. Two previous microarray analyses have identified obese breast tumour genetic signatures, in which many genes were being either up or downregulated in obese compared to healthy breast tumours[118, 119]. Within these signatures there are specific genes that are differentially expressed in obese patient's breast cancers. The expression of these genes could provide a molecular link between obesity and obesity inducing co-morbidities, such as hyperinsulinemia, with the development and growth of breast carcinogenesis. The second hypothesis tested in this study was that obese breast cancers are associated with differential expression of candidate genes. This study identified four candidate genes (*GRIA2, DUSP4, NR2F1* and *ADH1B*), recognised to be commonly dysregulated in these previously established obese breast tumour transcriptomes[118, 119]. When directly measuring the level of *GRIA2, DUSP4, NR2F1* and *ADH1B* mRNA expression in breast cancer biopsy homogenates from obese and healthy women, these candidate genes showed a trend toward downregulation in obese compared to the healthy breast tumours, although this trend was not statistically significant. These results are consistent with the findings from the two previous microarray studies[118, 119]. Furthermore, it was determined that DUSP4 protein was also expressed at lower levels in obese compared to healthy breast tumours.

Previous transcriptomic analyses support the current study's observed downregulation of *DUSP4* in obese relative to healthy weight breast tumours[118, 119]. Based on previous research and this study's results it is hypothesised that low expression of *DUSP4* in obese breast tumours may play a role in aggressive cell phenotypes through its interaction with downstream MAPK enzymes. However, the upstream molecular mechanisms responsible for the likely downregulation of *DUSP4* within these obese patients have yet to be identified. Fuentes-Mattei *et al*.[119] suggest the obesity signature of ER+ breast tumours, in obese versus healthy weight patients, is most likely occurring as a result of upstream regulation by adipokines, insulin and IGF-I signalling. Low *DUSP4* expression is one of the genes within the obese breast tumour transcriptional signature, and therefore it is possible that prolonged insulin hypersecretion is acting to

negatively regulate *DUSP4* expression. This means that hyperinsulinemia may not only be increasing the risk of becoming obese, but also acting as a mechanism inducing tumour proliferation by causing *DUSP4* to be downregulated in obese breast tumours. Thus, lower *AMY1* copy number may cause prolonged insulin hypersecretion after ingestion of starch resulting in an upstream regulation and reduced *DUSP4* expression.

Low *AMY1* copy number has been speculated to be increasing the risk of obesity through its ability to cause prolonged increases in blood insulin concentrations following the consumption of starch rich meals. There may be a link between *AMY1* copy number and *ADH1B* expression in obese breast cancer through their proposed associations with hyperinsulinemia. *ADH1B* was observed to be downregulated in obese breast tumour samples, supported by the transcriptomic examination carried out by Creighton *et al*.[118]. A previous study has observed reduced *ADH1B* expression correlating strongly with hyperinsulinemia and obesity in adipose samples from obese relative to healthy weight subjects (according to BMI and waist circumference measures)[171]. This suggests that a reduction in *ADH1B* expression in obese breast cancers could be involved in inducing a degree of hyperinsulinemia within these tumours, triggering mitosis and increased tumour cell survival.

Lower germline copy number of *AMY1* may also be associated with a subsequent downregulation in *AMY1* mRNA within breast tumours of obese patients. Fuentes-Mattei *et al*.[119] recognised *AMY1* to be differentially regulated in ER+ breast tumour biopsies from obese relative to healthy weight patients (Supplementary Figure 6.4). Interestingly, *AMY1* was downregulated in obese tumours, which is perhaps what would be expected if obese breast cancer patients have significantly lower germline *AMY1* copy numbers. The lower expression of *AMY1* within the tumour of obese breast cancer patients could be influencing tumour hyperinsulinemia and perhaps more aggressive breast carcinogenesis.

Overall, this study has investigated potential molecular links (germline and tumour) that may be influencing the obesity associated increased risk and development of breast cancer. Results from this study indicated that obese breast cancer patients have both significantly lower germline *AMY1* copy number and tumours with a trend toward lower *GRIA2, DUSP4, NR2F1* and *ADH1B* expression. The link between lower *AMY1* copy number and reduced expression of these candidate genes, in obese women, may interact through their association with patient hyperinsulinemia to increase the risk of breast tumourigenesis and contribute to a more agressive tumour phenotype.

## 5.1 Future Research

To better understand the potential genetic link between *AMY1* and breast cancer, future research is required to investigate both the germline *AMY1* copy number and *AMY1* mRNA expression levels in healthy and obese breast cancer patients. The level of *AMY1* expression within a patient blood sample can be compared to their observed copy number status. These correlations have already been established in the saliva and blood of healthy subjects[137] but not in breast cancer patient blood samples. Additionally, to test whether low *AMY1* copy number is linked with hyperinsulinemia, and therefore obesity, it will be important to measure *AMY1* copy number and the degree of insulin resistance in obese and healthy breast cancer patients. These findings will determine whether there is a correlation between hyperinsulinemia and low *AMY1* copy number.

The current study has found a trend toward downregulation of *GRIA2, DUSP4, NR2F1* and *ADH1B* in obese breast tumours, but these associations were non-significant. If this study were to be continued, a larger sample cohort of breast cancer patients/tumours would need to be investigated in order to validate the observed molecular associations between obesity and breast cancer.

To better quantify the level of *GRIA2, DUSP4, NR2F1* and *ADH1B* protein expression in tumours, other technologies such as ELISA (enzyme-linked immunosorbent assay) could be used as they provide a more quantitative approach relative to Western blotting.

The DUSP4 ELSIA results could also be compared back to the DUSP4 Western blots carried out in the current study.

There is considerable evidence that dysregulation of *DUSP4* is playing a role in the basal-like breast cancer aggressive phenotype[150]. Whether the downregulation of *DUSP4* is associated with aggressiveness in breast tumours from obese patients is yet to be identified. The current study has been the first to show a trend of lower *DUSP4* mRNA and protein expression, with increased BMI in breast cancers. Previous studies have established that downstream effects of reduced *DUSP4* expression (ERK1, ERK2 and JNK activation), are influencing breast cancer cells to become more aggressive. Therefore, it would be of value to analyse the expression of ERK1, ERK2 and JNK in the breast tumours from the current study. Lower levels of DUSP4 protein may be associated with increased quantities of DUSP4 downstream targets in obese breast tumours.

A previous study suggests that that *DUSP4* may being downregulated in some breast tumours due to copy number loss or epigenetic downregulation[150]; however, the upstream regulation causing reduced *DUSP4* expression is yet to be identified. Further research investigating the mechanisms responsible for *DUSP4* downregulation in obese breast tumours would perhaps elucidate mechanisms tying in with obesity-associated patient hyperinsulinemia. The degree of insulin resistance of obese and healthy breast cancer patients could be compared to the level of DUSP4 expression within breast tumours. The influence of hyperinsulinemia on *DUSP4* regulation may provide both a mechanistic and molecular link between obesity and breast cancer. This could provide a potential therapeutic breakthrough, significant for patients who have *DUSP4* deficient obesity associated breast tumour phenotypes.

## 5.2 Conclusion

The overall aim for the current investigation was to explore potential molecular links between obesity and breast cancer, as the molecular and mechanistic basis of the obesity and breast cancer association still remains unclear. Furthermore, gene

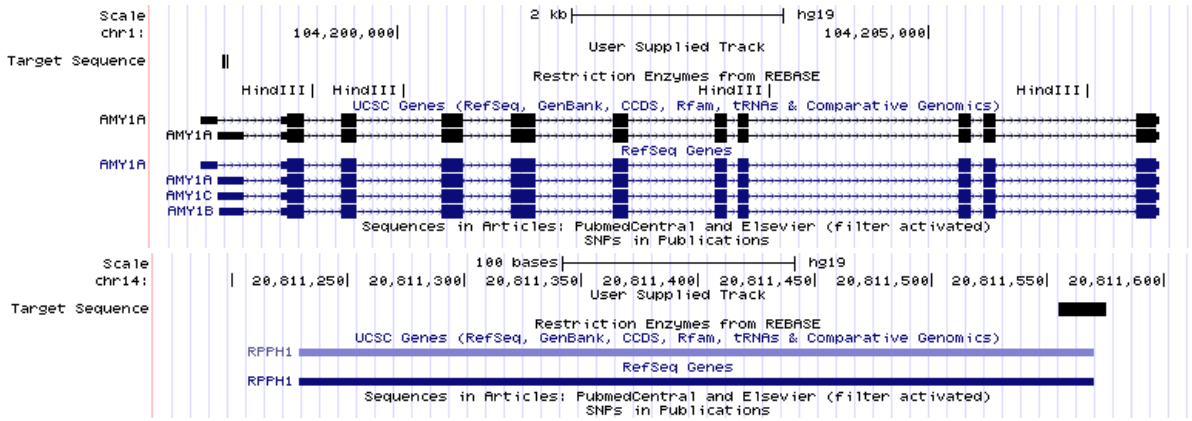
expression changes in breast tumours from obese women remain poorly characterised. To my knowledge, this study has been the first to investigate the relationship of germline *AMY1* copy number, and breast tumour *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* gene expression in obese and healthy breast cancer patients.

Results showed that obese breast cancer patients were associated with lower germline *AMY1* copy number. Thus, due to a relationship with obesity low *AMY1* copy number may be indirectly associated with a degree of genetic predisposition towards breast cancer. Additionally, the current study suggests that a trend toward lower *GRIA2*, *DUSP4*, *NR2F1* and *ADH1B* expression, observed in breast cancers from obese women, may be suggestive of an obese patient phenotype. However, future research investigating a larger sample cohort would be needed to determine if such a relationship exists.

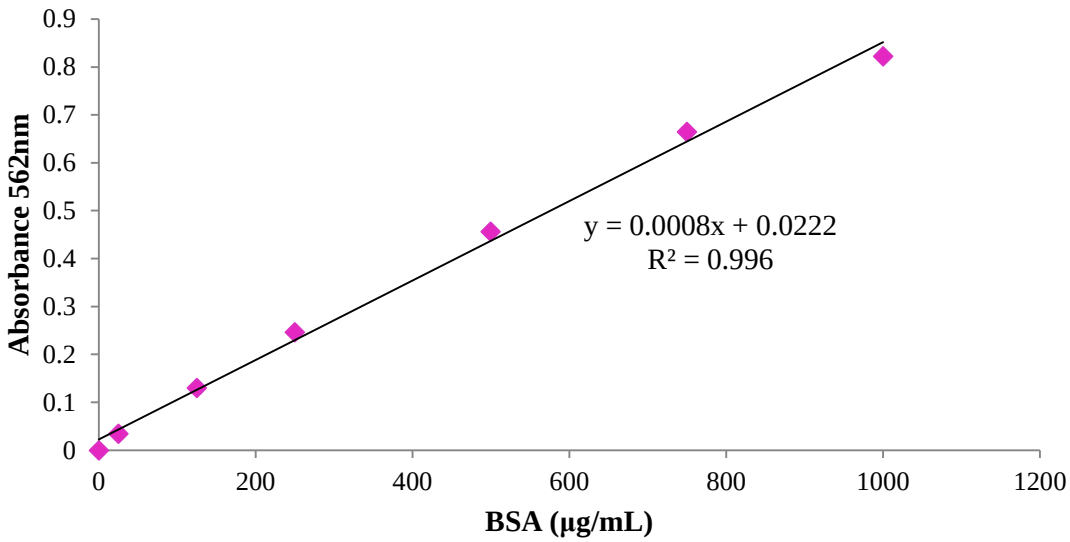
Hyperinsulinemia is considered an important mechanism for risk and poor outcome in breast cancer. This study proposes molecular mechanisms including, low *AMY1* copy number and negative regulation of *DUSP4*, that may underlie those links, and provide testable hypotheses for future research.

# Chapter 6

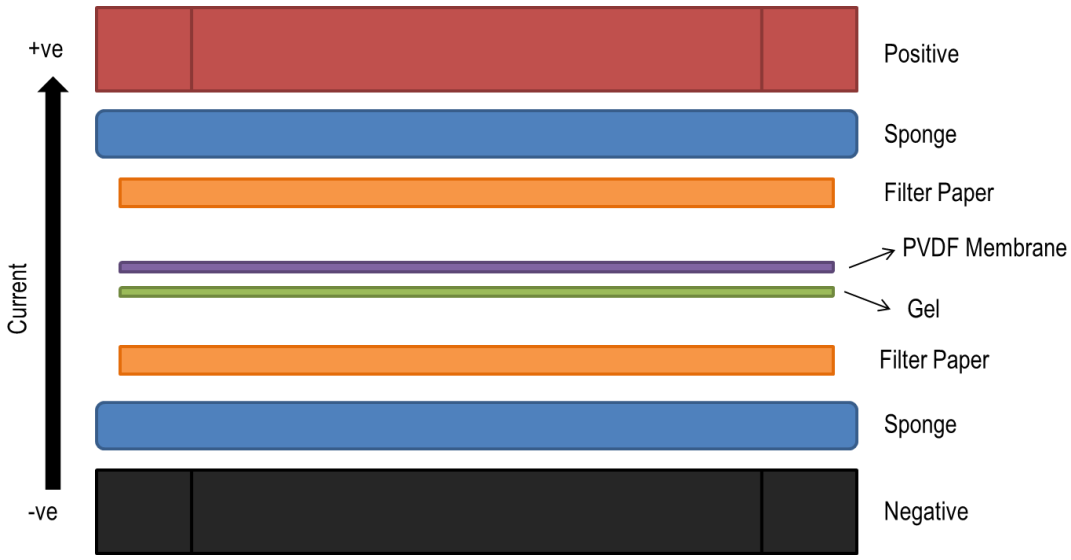
# 6 µSupplementary Data




**Supplementary Figure 6.1** Location of *Hind*III restriction digestion sites within the *AMY1* target gene and RNase P (*RPPH1*) reference genes and the location of the primer/probe target sequences within the *AMY1* and RNase P genes.




**Supplementary Figure 6.2** A representative BSA standard curve and best fit line equation using BCA assay (0-1000 µg/mL).




**Supplementary Figure 6.3** Schematic for the assembly of the transfer elements and the direction of the current used for Western blotting.

**Supplementary Table 6.1** Comparison of *AMY1* copy number, for both Hs07226361_cn (AMY1-61) and Hs07226362_cn (AMY1-62) primer/probe assays, to patient tumour pathological features.

| | | AMY1-61 | | | |
|---|---|---|---|---|---|
| **Pathological Feature** | | **Sample Size** | **Mean *AMY1* Copy Number** | **Standard Error** | **P-value** |
| **ER Status** | Negative | 13 | 8.24 | 0.51 | 0.56 |
| | Positive | 40 | 8.62 | 0.40 | |
| **PR Status** | Negative | 16 | 7.86 | 0.47 | 0.12 |
| | Positive | 36 | 8.86 | 0.42 | |
| **HER2 Status** | Negative | 36 | 8.23 | 0.37 | 0.37 |
| | Positive | 10 | 9.07 | 0.83 | |
| **Grade** | 1 & 2 | 17 | 9.12 | 0.60 | 0.29 |
| | 3 | 33 | 8.34 | 0.40 | |
| | | **AMY1-62** | | | |
| **Pathological Feature** | | **Sample Size** | **Mean *AMY1* Copy Number** | **Standard Error** | **P-value** |
| **ER Status** | Negative | 13 | 8.67 | 0.67 | 0.67 |
| | Positive | 40 | 9.02 | 0.43 | |
| **PR Status** | Negative | 16 | 8.42 | 0.57 | 0.28 |
| | Positive | 36 | 9.21 | 0.45 | |
| **HER2 Status** | Negative | 36 | 8.62 | 0.41 | 0.76 |
| | Positive | 10 | 8.94 | 0.92 | |
| **Grade** | 1 & 2 | 17 | 9.69 | 0.58 | 0.15 |
| | 3 | 33 | 8.61 | 0.45 | |

**Supplementary Table 6.2** Comparison of patient tumour pathological features and patient BMI.

| Pathological Feature | | Sample Size | Average BMI | P-value |
|---|---|---|---|---|
| **ER Status** | Negative | 11 | 29.17 | 0.84 |
| | Positive | 38 | 29.74 | |
| **PR Status** | Negative | 14 | 28.50 | 0.70 |
| | Positive | 34 | 29.44 | |
| **HER2 Status** | Negative | 33 | 30.36 | 0.71 |
| | Positive | 9 | 31.51 | |
| **Grade** | 1 and 2 | 17 | 29.75 | 0.95 |
| | 3 | 33 | 29.59 | |
| **Histological Type** | IDC | 29 | 30.11 | 0.69 |
| | Other | 22 | 29.17 | |

IDC= Invasive ductal carcinoma

**Supplementary Table 6.3** List of genes statistically significantly (fold change and p value < 0.05) differentially expressed in breast tumours from obese (BMI >30) compared with healthy patients (BMI <25) from data generated by Creighton *et al*.[118].

| Gene Symbol | Probe ID | Fold Change | P-Value |
|---|---|---|---|
| GRIA2 | 205358_at | -2.39 | 0.0082 |
| ADH1B | 209612_s_at | -2.3 | 0.0126 |
| AGTR1 | 205357_s_at | -1.99 | 0.0227 |
| DUSP4 | 204014_at | -1.8 | 0.0108 |
| GALNT7 | 218313_s_at | -1.69 | 0.0096 |
| MEST | 202016_at | -1.68 | 0.0125 |
| SCUBE2 | 219197_s_at | -1.68 | 0.0246 |
| FRY | 204072_s_at | -1.66 | 0.0017 |
| PRKAR2B | 203680_at | -1.65 | 0.008 |
| PLN | 204939_s_at | -1.6 | 0.0094 |
| SEMA3C | 203789_s_at | -1.59 | 0.0127 |
| TGFBR3 | 204731_at | -1.58 | 0.0355 |
| GYG2 | 215695_s_at | -1.57 | 0.032 |
| CCND1 | 208712_at | -1.56 | 0.0214 |
| FMO2 | 211726_s_at | -1.54 | 0.03 |
| IL6ST | 212195_at | -1.53 | 0.0075 |
| TFPI | 213258_at | -1.53 | 0.017 |
| ABAT | 209459_s_at | -1.52 | 0.0369 |
| RNASE4 | 213397_x_at | -1.49 | 0.011 |
| TNFSF10 | 202688_at | -1.48 | 0.024 |
| DUSP6 | 208892_s_at | -1.48 | 0.0207 |
| FLRT3 | 219250_s_at | -1.48 | 0.041 |
| ITM2A | 202746_at | -1.47 | 0.0463 |
| ITPR1 | 203710_at | -1.47 | 0.0095 |
| TMEM47 | 209656_s_at | -1.47 | 0.0122 |
| LRIG1 | 211596_s_at | -1.47 | 0.003 |
| PCLO | 213558_at | -1.47 | 0.0038 |
| CLDN8 | 214598_at | -1.46 | 0.0485 |
| BCL2 | 203685_at | -1.45 | 0.0241 |
| KAL1 | 205206_at | -1.45 | 0.0088 |
| BG251521 | 213156_at | -1.45 | 0.0271 |
| CTBP2 | 201218_at | -1.44 | 0.0004 |
| MAGI2 | 209737_at | -1.43 | 0.0061 |
| GPD1L | 212510_at | -1.43 | 0.0023 |
| CDKN1C | 213348_at | -1.43 | 0.0165 |
| ENOSF1 | 213645_at | -1.43 | 0.0006 |
| PLSCR4 | 218901_at | -1.43 | 0.0182 |
| SPARCL1 | 200795_at | -1.42 | 0.0416 |
| CYP1B1 | 202437_s_at | -1.42 | 0.0381 |
| XIST | 221728_x_at | -1.42 | 0.0265 |
| NRIP1 | 202599_s_at | -1.41 | 0.0307 |
| STS | 203767_s_at | -1.41 | 0.0209 |

| | | | |
|---|---|---|---|
| ZNF91 | 206059_at | -1.41 | 0.0032 |
| MEIS3P1 | 214077_x_at | -1.41 | 0.0003 |
| HIST1H2AC | 215071_s_at | -1.41 | 0.0411 |
| BCL6 | 203140_at | -1.4 | 0.0038 |
| 209505_at (NR2F1) | 209505_at | -1.4 | 0.0322 |
| LIMCH1 | 212327_at | -1.4 | 0.0211 |
| TNS1 | 221748_s_at | -1.4 | 0.0491 |
| PODXL | 201578_at | -1.4 | 0.0062 |
| FAM13A | 202973_x_at | -1.39 | 0.034 |
| PER2 | 205251_at | -1.39 | 0.000088 |
| ALG13 | 205583_s_at | -1.39 | 0.0002 |
| ZDHHC17 | 212982_at | -1.39 | 0.008 |
| ADI1 | 217761_at | -1.39 | 0.0025 |
| HSD17B11 | 217989_at | -1.39 | 0.0212 |
| CADPS2 | 219572_at | -1.39 | 0.0081 |
| NBEA | 221207_s_at | -1.39 | 0.0127 |
| TMX4 | 201581_at | -1.38 | 0.0169 |
| ACSL3 | 201661_s_at | -1.38 | 0.014 |
| PKP4 | 201928_at | -1.38 | 0.0002 |
| RHOBTB3 | 202976_s_at | -1.38 | 0.025 |
| RB1 | 203132_at | -1.38 | 0.0092 |
| PDZD2 | 209493_at | -1.38 | 0.0329 |
| PION | 213142_x_at | -1.38 | 0.0113 |
| PGRMC2 | 213227_at | -1.38 | 0.0069 |
| C1orf115 | 218546_at | -1.38 | 0.0064 |
| RYBP | 201845_s_at | -1.37 | 0.0051 |
| PCM1 | 202174_s_at | -1.37 | 0.0029 |
| KIF5C | 203130_s_at | -1.37 | 0.0334 |
| IRS2 | 209185_s_at | -1.37 | 0.0171 |
| CACNA1D | 210108_at | -1.37 | 0.0268 |
| GPR116 | 212950_at | -1.37 | 0.0275 |
| KIAA0485 | 214295_at | -1.37 | 0.0115 |
| SLTM | 217828_at | -1.37 | 0.0026 |
| ADD3 | 201034_at | -1.36 | 0.0416 |
| 214753_at | 214753_at | -1.36 | 0.0026 |
| LZTFL1 | 218437_s_at | -1.36 | 0.0075 |
| TRIM14 | 203148_s_at | -1.35 | 0.0105 |
| TSC22D3 | 208763_s_at | -1.35 | 0.0015 |
| NCOA1 | 209106_at | -1.35 | 0.0002 |
| DOCK9 | 212538_at | -1.35 | 0.0002 |
| ENO1 | 217294_s_at | 1.36 | 0.0312 |
| S100G | 207885_at | 1.4 | 0.0333 |
| ARNTL2 | 220658_s_at | 1.4 | 0.008 |
| NLRP2 | 221690_s_at | 1.43 | 0.027 |
| CCL18 | 209924_at | 1.46 | 0.0086 |
| KRT6B | 213680_at | 1.5 | 0.0415 |

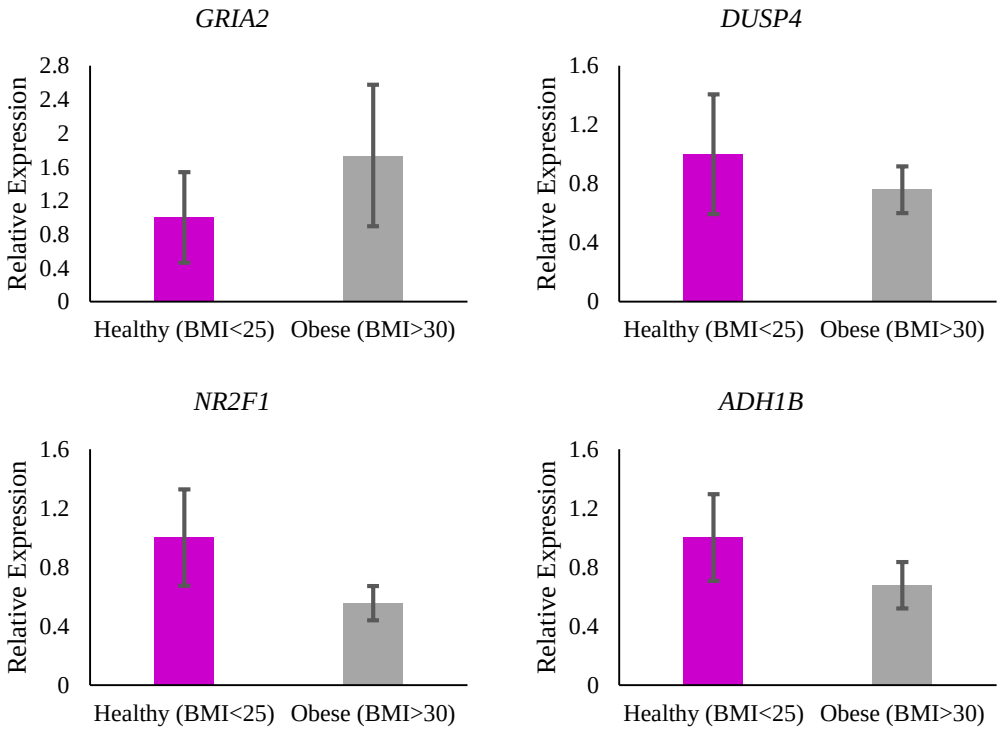
| | | | |
|---|---|---|---|
| MMP12 | 204580_at | 1.56 | 0.0323 |
| S100A2 | 204268_at | 1.58 | 0.0105 |
| MMP9 | 203936_s_at | 1.63 | 0.0015 |

**Supplementary Table 6.4** List of genes statistically significantly ($p \leq 0.01$ and absolute value of log-ratio > 0.1) differentially regulated in ER+ breast tumours in obese (BMI > 30) compared with non-obese (BMI < 30) patients from data generated by Fuentes-Mattei *et al*.[119].

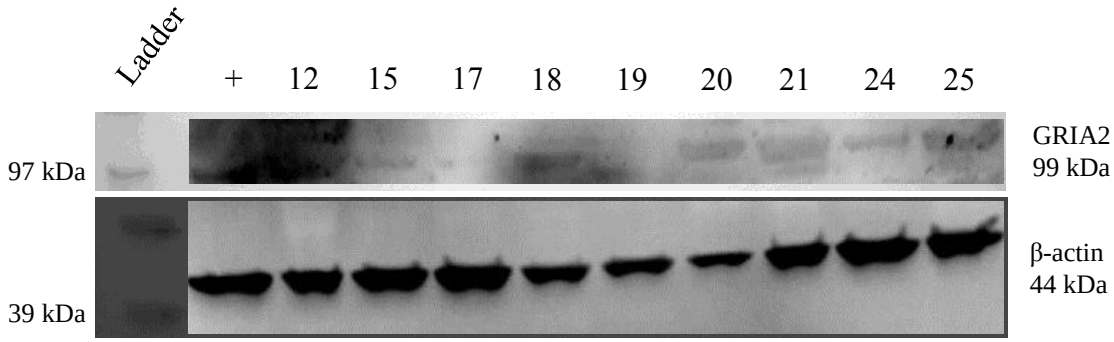
| Gene Symbol | Probe ID | Log ratio (Log10) | P-Value |
|---|---|---|---|
| PIP | 206509_at | -1.4353 | <0.001 |
| AQP3 | 39248_at | -1.2791 | <0.001 |
| GJA1 | 201667_at | -1.0564 | 0 |
| C8orf4 | 218541_s_at | -1.0396 | <0.001 |
| UGT2B28 | 211682_x_at | -0.9668 | <0.001 |
| MYBPC1 | 214087_s_at | -0.964 | <0.001 |
| NPY1R | 205440_s_at | -0.9324 | <0.001 |
| CXCL13 | 205242_at | -0.9132 | <0.001 |
| SGK3 | 220038_at | -0.8923 | <0.001 |
| HSPA2 | 211538_s_at | -0.8459 | <0.001 |
| DUSP4 | 204014_at | -0.8303 | <0.001 |
| AREG | 205239_at | -0.735 | <0.001 |
| PDLIM1 | 208690_s_at | -0.7316 | 0.005 |
| PRSS23 | 202458_at | -0.7164 | 0.004 |
| MUC1 | 207847_s_at | -0.7116 | 0.009 |
| SERHL | 214243_s_at | -0.6772 | <0.001 |
| CADM1 | 209031_at | -0.5888 | 0.001 |
| SULF1 | 212344_at | -0.5774 | <0.001 |
| IGFBP5 | 211958_at | -0.5225 | 0.007 |
| MSMB | 207430_s_at | -0.5198 | <0.001 |
| EHF | 219850_s_at | -0.4399 | 0.003 |
| ANXA3 | 209369_at | -0.4304 | 0.003 |
| SLC26A3 | 206143_at | -0.4272 | <0.001 |
| MMP10 | 205680_at | -0.3848 | <0.001 |
| FKBP5 | 204560_at | -0.3581 | <0.001 |
| IGHM | 211634_x_at | -0.3544 | 0.008 |
| SLC1A1 | 206396_at | -0.353 | <0.001 |
| TSPAN8 | 203824_at | -0.3404 | <0.001 |
| CITED1 | 207144_s_at | -0.334 | 0.002 |
| PGR | 208305_at | -0.3336 | <0.001 |
| RLN2 | 214519_s_at | -0.3178 | 0.001 |
| GRIA2 | 205358_at | -0.3115 | <0.001 |
| NR2F1 | 209505_at | -0.3106 | 0.004 |
| TFPI2 | 209277_at | -0.3104 | <0.001 |
| COL5A1 | 212489_at | -0.3099 | 0.008 |
| CRISP3 | 207802_at | -0.3012 | <0.001 |
| IL20RA | 219115_s_at | -0.292 | 0.003 |
| SMARCA1 | 215294_s_at | -0.2729 | 0.005 |
| NDP | 206022_at | -0.2659 | 0.004 |
| NPY5R | 207400_at | -0.2598 | 0.006 |
| GLRB | 205280_at | -0.2592 | <0.001 |
| DLX2 | 207147_at | -0.2522 | <0.001 |
| PLCL1 | 205934_at | -0.2443 | 0.002 |

| | | | |
|---|---|---|---|
| AMY1A | 208498_s_at | -0.2336 | <0.001 |
| MSLN | 204885_s_at | -0.227 | <0.001 |
| FLJ11184 | 218513_at | -0.216 | <0.001 |
| FAM5C | 217562_at | -0.2126 | <0.001 |
| UGT2B4 | 206505_at | -0.1888 | <0.001 |
| EREG | 205767_at | -0.1695 | 0.002 |
| IL33 | 209821_at | 0.1552 | 0.008 |
| ATRNL1 | 213745_at | 0.2 | <0.001 |
| C11orf30 | 219012_s_at | 0.208 | 0.004 |
| HBZ | 206647_at | 0.2222 | 0.007 |
| ALB | 214842_s_at | 0.2371 | <0.001 |
| PEG3 | 209243_s_at | 0.2394 | 0.001 |
| ITSN1 | 209298_s_at | 0.2481 | <0.001 |
| PIR | 207469_s_at | 0.2636 | 0.008 |
| CDSN | 206193_s_at | 0.2821 | <0.001 |
| GSTT2 | 205439_at | 0.2857 | <0.001 |
| TIMP4 | 206243_at | 0.2898 | <0.001 |
| PCOLCE2 | 219295_s_at | 0.2914 | <0.001 |
| FHL1 | 201539_s_at | 0.2978 | <0.001 |
| MRC1 | 204438_at | 0.3005 | 0.002 |
| IL8 | 211506_s_at | 0.3015 | <0.001 |
| DPT | 213068_at | 0.3045 | 0.001 |
| TRIM58 | 215047_at | 0.3177 | <0.001 |
| AFP | 204694_at | 0.3256 | <0.001 |
| TF | 203400_s_at | 0.3302 | <0.001 |
| CAV2 | 203323_at | 0.3331 | 0.002 |
| RSF1 | 218166_s_at | 0.3337 | <0.001 |
| CLEC3B | 205200_at | 0.3466 | <0.001 |
| AOC3 | 204894_s_at | 0.3483 | <0.001 |
| CHRDL1 | 209763_at | 0.3489 | 0.0002 |
| CRABP1 | 205350_at | 0.389 | <0.001 |
| CP | 204846_at | 0.4028 | 0.001 |
| LBP | 214461_at | 0.412 | <0.001 |
| ASCL1 | 209987_s_at | 0.4495 | <0.001 |
| TWIST1 | 213943_at | 0.4551 | <0.001 |
| RBP4 | 219140_s_at | 0.4764 | <0.001 |
| CD36 | 206488_s_at | 0.4912 | 0.005 |
| ADH1B | 209612_s_at | 0.5118 | <0.001 |
| LEP | 207092_at | 0.5158 | <0.001 |
| ACOX2 | 205364_at | 0.5172 | 0.002 |
| GRB14 | 206204_at | 0.5212 | <0.001 |
| CAV1 | 212097_at | 0.526 | 0.002 |
| TOP2A | 201292_at | 0.5498 | 0.003 |
| PLOD2 | 202620_s_at | 0.5637 | 0.002 |
| TSKU | 218245_at | 0.5767 | 0.001 |
| CLSTN2 | 219414_at | 0.5877 | <0.001 |

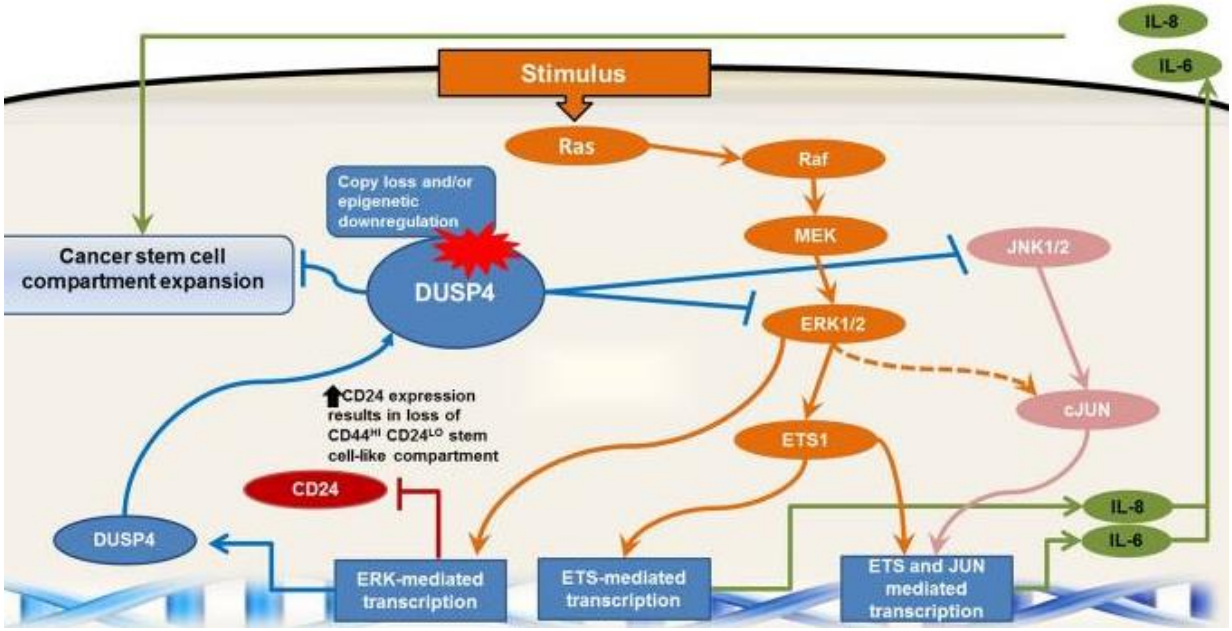
| | | | |
|---|---|---|---|
| ADIPOQ | 207175_at | 0.6023 | <0.001 |
| LPL | 203548_s_at | 0.6368 | <0.001 |
| GPX3 | 201348_at | 0.6369 | <0.001 |
| PPBP | 214146_s_at | 0.6429 | <0.001 |
| FABP5 | 202345_s_at | 0.6529 | <0.001 |
| SAA1 | 208607_s_at | 0.6929 | <0.001 |
| KCNK1 | 204679_at | 0.6946 | 0.003 |
| AMIGO2 | 222108_at | 0.6981 | 0.003 |
| PLIN | 205913_at | 0.7004 | <0.001 |
| AKR1C3 | 209160_at | 0.7246 | <0.001 |
| G0S2 | 213524_s_at | 0.7461 | <0.001 |
| HBG1 | 204848_x_at | 0.7619 | 0.002 |
| CFD | 205382_s_at | 0.7621 | <0.001 |
| HBG2 | 204419_x_at | 0.7872 | 0.001 |
| S100A8 | 202917_s_at | 0.7919 | <0.001 |
| HBD | 206834_at | 0.855 | <0.001 |
| AKR1C1 | 204151_x_at | 0.8679 | <0.001 |
| AKR1C2 | 209699_x_at | 0.8686 | <0.001 |
| FABP4 | 203980_at | 0.9573 | <0.001 |
| SNCA | 204466_s_at | 0.9795 | <0.001 |
| ALAS2 | 211560_s_at | 1.0281 | <0.001 |




**Supplementary Figure 6.4** The relative expression of the obese (BMI > 30; n= 22) compared to the healthy (BMI < 25; n= 17) breast tumour cohorts for *GRIA2* (p= 0.47), *DUSP4* (p= 0.60), *NR2F1* (p= 0.26) and *ADH1B* (p= 0.35) with the influential data points removed. The healthy tumours are the pink bars and the obese the grey bars. The error bars are the standard error of the mean.




**Supplementary Figure 6.5** Representative Western blot for GRIA2 with a 15 minute exposure and loading 40 µg of total protein per well. Tumour Samples: positive control, 12, 15, 17, 18, 19, 20, 21, 24 and 25.




**Supplementary Figure 6.6** The proposed pathway model for the function of *DUSP4* in basal-like breast cancers.

Schematic taken from: Balko JM, Schwarz LJ, Bhola NE, Kurupi R, Owens P, Miller TW, et al. Activation of MAPK Pathways due to DUSP4 Loss Promotes Cancer Stem Cell-like Phenotypes in Basal-like Breast Cancer. Cancer research. 2013;73(20):6346-58.[150]. Reprinted with permission from American Association for Cancer Research; License Number: 3727850326500.

## References


1. Marie NG, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet. 2014;384(9945):766-81.
2. Haslam DW, James WPT. Obesity. The Lancet. 2005;366(9492):1197-209.
3. Protani M, Coory M, Martin JH. Effect of obesity on survival of women with breast cancer: systematic review and meta-analysis. Breast cancer research and treatment. 2010;123(3):627-35.
4. Cheraghi Z, Poorolajal J, Hashem T, Esmailnasab N, Doosti Irani A. Effect of body mass index on breast cancer during premenopausal and postmenopausal periods: a meta-analysis. PLoS One. 2012;7(12):e51446.
5. Onishi-Seebacher M, Korbel JO. Challenges in studying genomic structural variant formation mechanisms: The short-read dilemma and beyond. Bioessays. 2011;33(11):840-50.
6. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nature genetics. 2004;36(9):949-51.
7. Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, et al. Diet and the evolution of human amylase gene copy number variation. Nature genetics. 2007;39(10):1256-60.
8. Falchi M, El-Sayed Moustafa JS, Takousis P, Pesce F, Bonnefond A, Andersson-Assarsson JC, et al. Low copy number of the salivary amylase gene predisposes to obesity. Nature genetics. 2014;46(5):492-7.
9. World Health Organisation 2012. [Obesity and Overweight]. Available from: http: //www.who.int/mediacentre/factsheets/fs311/en/.
10. Mitchell NS, Catenacci VA, Wyatt HR, Hill JO. Obesity: overview of an epidemic. Psychiatric Clinics of North America. 2011;34(4):717-32.
11. Kitzinger HB, Karle B. The epidemiology of obesity. European Surgery. 2013;45(2):80-2.
12. Kelly T, Yang W, Chen CS, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. International journal of obesity. 2008;32(9):1431-7.
13. Hill JO. Understanding and addressing the epidemic of obesity: an energy balance perspective. Endocrine reviews. 2006;27(7):750-61.
14. Berstein LM. Cancer and heterogeneity of obesity: a potential contribution of brown fat. Future Oncology. 2012;8(12):1537-48.
15. El-Sayed Moustafa JS, Froguel P. From obesity genetics to the future of personalized obesity therapy. Nature Reviews Endocrinology. 2013;9(7):402-13.
16. Zhang YY, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM. POSITIONAL CLONING OF THE MOUSE OBESE GENE AND ITS HUMAN HOMOLOG. Nature. 1994;372(6505):425-32.
17. Blakemore AI, Froguel P. Investigation of Mendelian forms of obesity holds out the prospect of personalized medicine. Annals of the New York Academy of Sciences. 2010;1214:180-9.

18. He S, Tao YX. Defect in MAPK Signaling As a Cause for Monogenic Obesity Caused By Inactivating Mutations in the Melanocortin-4 Receptor Gene. International Journal of Biological Sciences. 2014;10(10):1128-37.
19. Paz GJ, Babikian T, Asarnow R, Esposito K, Erol HK, Wong ML, et al. Leptin Replacement Improves Cognitive Development. PLoS One. 2008;3(8):7.
20. Dubern B, Clement K. Leptin and leptin receptor-related monogenic obesity. Biochimie. 2012;94(10):2111-5.
21. Xia QH, Grant SFA. The genetics of human obesity. AnnNY AcadSci. 2013;1281:178-90.
22. Apalasamy YD, Mohamed Z. Obesity and genomics: role of technology in unraveling the complex genetic architecture of obesity. Hum Genet. 2015;134(4):361-74.
23. Hainerova IA, Lebl J. Treatment Options for Children with Monogenic Forms of Obesity. In: Shamir R, Turck D, Phillip M, editors. Nutrition and Growth. World Review of Nutrition and Dietetics. 106. Basel: Karger; 2013. p. 105-12.
24. Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, et al. Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature. 1997;387(6636):903-8.
25. Echwald SM, Rasmussen SB, Sorensen TIA, Andersen T, TybjaergHansen A, Clausen JO, et al. Identification of two novel missense mutations in the human OB gene. International journal of obesity. 1997;21(4):321-6.
26. Oksanen L, Kainulainen K, Heiman M, Mustajoki P, KauppinenMakelin R, Kontula K. Novel polymorphism of the human ob gene promoter in lean and morbidly obese subjects. International journal of obesity. 1997;21(6):489-94.
27. Clement K, Vaisse C, Lahlou N, Cabrol S, Pelloux V, Cassuto D, et al. A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. Nature. 1998;392(6674):398-401.
28. Krude H, Biebermann H, Luck W, Horn R, Brabant G, Gruters A. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. Nature genetics. 1998;19(2):155-7.
29. Farooqi IS, Keogh JM, Yeo GSH, Lank EJ, Cheetham T, O'Rahilly S. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. N Engl J Med. 2003;348(12):1085-95.
30. Han JC, Liu QR, Jones M, Levinn RL, Menzie CM, Jefferson-George KS, et al. Brain-derived neurotrophic factor and obesity in the WAGR syndrome. N Engl J Med. 2008;359(9):918-27.
31. Yeo GSH, Hung CCC, Rochford J, Keogh J, Gray J, Sivaramakrishnan S, et al. A de novo mutation affecting human TrkB associated with severe obesity and developmental delay. Nat Neurosci. 2004;7(11):1187-9.
32. Jackson RS, Creemers JWM, Ohagi S, RaffinSanson ML, Sanders L, Montague CT, et al. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. Nature genetics. 1997;16(3):303-6.
33. Holder JL, Butte NF, Zinn AR. Profound obesity associated with a balanced translocation that disrupts the SIM1 gene. Hum Mol Genet. 2000;9(1):101-8.
34. Hinney A, Nguyen TT, Scherag A, Friedel S, Bronner G, Muller TD, et al. Genome Wide Association (GWA) Study for Early Onset Extreme Obesity Supports the Role of Fat Mass and Obesity Associated Gene (FTO) Variants. PLoS One. 2007;2(12):5.

35. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet. 2007;3(7):1200-10.
36. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316(5826):889-94.
37. Loos RJF, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nature genetics. 2008;40(6):768-75.
38. Wheeler E, Huang N, Bochukova EG, Keogh JM, Lindsay S, Garg S, et al. Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. Nature genetics. 2013;45(5):513-U76.
39. Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nature genetics. 2013;45(5):501-U69.
40. Hebebrand J, Volckmar AL, Knoll N, Hinney A. Chipping Away the 'Missing Heritability': GIANT Steps Forward in the Molecular Elucidation of Obesity - but Still Lots to Go. Obes Facts. 2010;3(5):294-303.
41. Hinney A, Herrfurth N, Schonnop L, Volckmar AL. Genetic and epigenetic mechanisms in obesity. Bundesgesundheitsblatt-Gesund. 2015;58(2):154-8.
42. Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, et al. Searching for missing heritability: Designing rare variant association studies. Proc Natl Acad Sci U S A. 2014;111(4):455-64.
43. Wen WQ, Zheng W, Okada Y, Takeuchi F, Tabara Y, Hwang JY, et al. Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. Hum Mol Genet. 2014;23(20):5492-504.
44. Pei YF, Zhang L, Liu YJ, Li J, Shen H, Liu YZ, et al. Meta-analysis of genome-wide association data identifies novel susceptibility loci for obesity. Hum Mol Genet. 2014;23(3):820-30.
45. Graff M, Ngwa JS, Workalemahu T, Homuth G, Schipf S, Teumer A, et al. Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. Hum Mol Genet. 2013;22(17):3597-607.
46. Monda KL, Chen GK, Taylor KC, Palmer C, Edwards TL, Lange LA, et al. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. Nature genetics. 2013;45(6):690-+.
47. Yang J, Loos RJF, Powell JE, Medland SE, Speliotes EK, Chasman DI, et al. FTO genotype is associated with phenotypic variability of body mass index. Nature. 2012;490(7419):267-+.
48. Bradfield JP, Taal HR, Timpson NJ, Scherag A, Lecoeur C, Warrington NM, et al. A genome-wide association meta-analysis identifies new childhood obesity loci. Nature genetics. 2012;44(5):526-+.
49. Okada Y, Kubo M, Ohmiya H, Takahashi A, Kumasaka N, Hosono N, et al. Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. Nature genetics. 2012;44(3):302-U105.

50. Wen WQ, Cho YS, Zheng W, Dorajoo R, Kato N, Qi L, et al. Meta-analysis identifies common variants associated with body mass index in east Asians. Nature genetics. 2012;44(3):307-U112.
51. Jiao H, Arner P, Hoffstedt J, Brodin D, Dubern B, Czernichow S, et al. Genome wide association study identifies KCNMA1 contributing to human obesity. BMC Med Genomics. 2011;4:10.
52. Wang K, Li WD, Zhang CK, Wang ZH, Glessner JT, Grant SFA, et al. A Genome-Wide Association Study on Obesity and Obesity-Related Traits. PLoS One. 2011;6(4):6.
53. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nature genetics. 2010;42(11):937-U53.
54. Scherag A, Dina C, Hinney A, Vatin V, Scherag S, Vogel CIG, et al. Two New Loci for Body-Weight Regulation Identified in a Joint Analysis of Genome-Wide Association Studies for Early-Onset Extreme Obesity in French and German Study Groups. PLoS Genet. 2010;6(4):10.
55. Cotsapas C, Speliotes EK, Hatoum IJ, Greenawalt DM, Dobrin R, Lum PY, et al. Common body mass index-associated variants confer risk of extreme obesity. Hum Mol Genet. 2009;18(18):3502-7.
56. Meyre D, Delplanque J, Chevre JC, Lecoeur C, Lobbens S, Gallina S, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nature genetics. 2009;41(2):157-9.
57. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nature genetics. 2009;41(1):18-24.
58. Willer CJ, Speliotes EK, Loos RJF, Li SX, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nature genetics. 2009;41(1):25-34.
59. Calle EE, Kaaks R. Overweight, obesity and cancer: Epidemiological evidence and proposed mechanisms. Nat Rev Cancer. 2004;4(8):579-91.
60. Majka SM, Barak Y, Klemm DJ. Concise review: adipocyte origins: weighing the possibilities. Stem cells. 2011;29(7):1034-40.
61. Esteve Ràfols M. Adipose tissue: Cell heterogeneity and functional diversity. Endocrinología y Nutrición (English Edition). 2014;61(2):100-12.
62. Boulet N, Esteve D, Bouloumie A, Galitzky J. Cellular heterogeneity in superficial and deep subcutaneous adipose tissues in overweight patients. Journal of physiology and biochemistry. 2013;69(3):575-83.
63. Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D, et al. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Bmj. 2007;335(7630):1134.
64. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. The Lancet. 2008;371(9612):569-78.
65. Renehan AG, Soerjomataram I, Tyson M, Egger M, Zwahlen M, Coebergh JW, et al. Incident cancer burden attributable to excess body mass index in 30 European countries. International Journal of Cancer. 2010;126(3):692-702.

66. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. The Lancet. 2014;384(9945):755-65.
67. Arnold M, Pandeya N, Byrnes G, Renehan AG, Stevens GA, Ezzati M, et al. Global burden of cancer attributable to high body-mass index in 2012: a population-based study. The Lancet Oncology. 2015;16(1):36-46.
68. Kaidar-Person O, Bar-Sela G, Person B. The two major epidemics of the twenty-first century: obesity and cancer. Obesity surgery. 2011;21(11):1792-7.
69. Renehan AG, Roberts DL, Dive C. Obesity and cancer: pathophysiological and biological mechanisms. Archives of physiology and biochemistry. 2008;114(1):71-83.
70. Godsland IF. Insulin resistance and hyperinsulinaemia in the development and progression of cancer. Clinical science. 2010;118(5):315-32.
71. King B, Jiang Y, Su X, Xu J, Xie L, Standard J, et al. Weight control, endocrine hormones and cancer prevention. Experimental biology and medicine. 2013;238(5):502-8.
72. Howe LR, Subbaramaiah K, Hudis CA, Dannenberg AJ. Molecular pathways: adipose inflammation as a mediator of obesity-associated cancer. Clinical Cancer Research. 2013;19(22):6074-83.
73. Renehan AG, Frystyk J, Flyvbjerg A. Obesity and cancer risk: the role of the insulin-IGF axis. Trends in Endocrinology and Metabolism. 2006;17(8):328-36.
74. Gilbert CA, Slingerland JM. Cytokines, Obesity, and Cancer: New Insights on Mechanisms Linking Obesity to Cancer Risk and Progression. Annual Review of Medicine. 2013;64:45-57.
75. VanSaun MN. Molecular Pathways: Adiponectin and Leptin Signaling in Cancer. Clinical Cancer Research. 2013;19(8):1926-32.
76. Geisler J, Haynes B, Ekse D, Dowsett M, Lonning PE. Total body aromatization in postmenopausal breast cancer patients is strongly correlated to plasma leptin levels. Journal of Steroid Biochemistry and Molecular Biology. 2007;104(1-2):27-34.
77. Eliassen AH, Hankinson SE. Endogenous hormone levels and risk of breast, endometrial and ovarian cancers: Prospective studies. Innovative Endocrinology of Cancer. Advances in Experimental Medicine and Biology. 630. Berlin2008. p. 148-65.
78. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: a cancer journal for clinicians. 2011;61(2):69-90.
79. Curado MP. Breast cancer in the world: Incidence and mortality. Salud Publica Mexico. 2011;53(5):372-84.
80. Modugno F, Kip KE, Cochrane B, Kuller L, Klug TL, Rohan TE, et al. Obesity, hormone therapy, estrogen metabolism and risk of postmenopausal breast cancer. International Journal of Cancer. 2006;118(5):1292-301.
81. Kawai M, Minami Y, Kuriyama S, Kakizaki M, Kakugawa Y, Nishino Y, et al. Adiposity, adult weight change and breast cancer risk in postmenopausal Japanese women: the Miyagi Cohort Study. Br J Cancer. 2010;103(9):1443-7.
82. Xia X, Chen W, Li J, Chen X, Rui R, Liu C, et al. Body mass index and risk of breast cancer: a nonlinear dose-response meta-analysis of prospective studies. Scientific reports. 2014;4:7480.
83. Loi S, Milne RL, Friedlander ML, McCredie MRE, Giles GG, Hopper JL, et al. Obesity and outcomes in premenopausal and postmenopausal breast cancer. Cancer Epidemiology Biomarkers and Prevention. 2005;14(7):1686-91.

84. Michels KB, Terry KL, Willett WC. Longitudinal study on the role of body size in premenopausal breast cancer. Arch Intern Med. 2006;166(21):2395-402.
85. Amadou A, Ferrari P, Muwonge R, Moskal A, Biessy C, Romieu I, et al. Overweight, obesity and risk of premenopausal breast cancer according to ethnicity: a systematic review and dose-response meta-analysis. Obesity reviews : an official journal of the International Association for the Study of Obesity. 2013;14(8):665-78.
86. Harris HR, Willett WC, Terry KL, Michels KB. Body Fat Distribution and Risk of Premenopausal Breast Cancer in the Nurses' Health Study II. Journal of the National Cancer Institute. 2011;103(3):273-8.
87. Dignam JJ. Obesity and breast cancer prognosis: an expanding body of evidence. Annals of Oncology. 2004;15(6):850-1.
88. Niraula S, Ocana A, Ennis M, Goodwin PJ. Body size and breast cancer prognosis in relation to hormone receptor and menopausal status: a meta-analysis. Breast cancer research and treatment. 2012;134(2):769-81.
89. Porter GA, Inglis KM, Wood LA, Veugelers PJ. Effect of obesity on presentation of breast cancer. Annals of surgical oncology. 2006;13(3):327-32.
90. Li CI, Daling JR, Porter PL, Tang MTC, Malone KE. Relationship Between Potentially Modifiable Lifestyle Factors and Risk of Second Primary Contralateral Breast Cancer Among Women Diagnosed With Estrogen Receptor-Positive Invasive Breast Cancer. J Clin Oncol. 2009;27(32):5312-8.
91. Daling JR, Malone KE, Doody DR, Johnson LG, Gralow JR, Porter PL. Relation of body mass index to tumor markers and survival among young women with invasive ductal breast carcinoma. Cancer. 2001;92(4):720-9.
92. Griggs JJ, Sorbero MES, Lyman GH. Undertreatment of obese women receiving breast cancer chemotherapy. Arch Intern Med. 2005;165(11):1267-73.
93. Ray A, Nkhata KJ, Cleary MP. Effects of leptin on human breast cancer cell lines in relationship to estrogen receptor and HER2 status. Int J Oncol. 2007;30(6):1499-509.
94. Hunter RJ, Navo MA, Thaker PH, Bodurka DC, Wolf JK, Smith JA. Dosing chemotherapy in obese patients: Actual versus assigned body surface area (BSA). Cancer Treat Rev. 2009;35(1):69-78.
95. Dorgan JF, Longcope C, Franz C, Stanczyk FZ, Chang LC, Stephenson HE, et al. Endogenous sex hormones and breast cancer in postmenopausal women: Reanalysis of nine prospective studies. Journal of the National Cancer Institute. 2002;94(8):606-16.
96. Folkerd EJ, Martin LA, Kendall A, Dowsett M. The relationship between factors affecting endogenous oestradiol levels in postmenopausal women and breast cancer. Journal Steroid Biochemical and Molecular Biology. 2006;102(1-5):250-5.
97. Baglietto L, English DR, Hopper JL, MacInnis RJ, Morris HA, Tilley WD, et al. Circulating steroid hormone concentrations in postmenopausal women in relation to body size and composition. Breast cancer research and treatment. 2009;115(1):171-9.
98. Brown KA, Simpson ER. Obesity and breast cancer: progress to understanding the relationship. Cancer research. 2010;70(1):4-7.
99. Suzuki R, Orsini N, Saji S, Key TJ, Wolk A. Body weight and incidence of breast cancer defined by estrogen and progesterone receptor status--a meta-analysis. International Journal of Cancer. 2009;124(3):698-712.

100. Vrieling A, Buck K, Kaaks R, Chang-Claude J. Adult weight gain in relation to breast cancer risk by estrogen and progesterone receptor status: a meta-analysis. Breast Cancer Research Treatment. 2010;123(3):641-9.
101. Suzuki R, Iwasaki M, Inoue M, Sasazuki S, Sawada N, Yamaji T, et al. Body weight at age 20 years, subsequent weight change and breast cancer risk defined by estrogen and progesterone receptor status--the Japan public health center-based prospective study. International Journal of Cancer. 2011;129(5):1214-24.
102. Madeddu C, Gramignano G, Floris C, Murenu G, Sollai G, Maccio A. Role of inflammation and oxidative stress in post-menopausal oestrogen-dependent breast cancer. Journal of Cellular and Molecular Medicine. 2014;18(12):2519-29.
103. Robinson PJ, Bell RJ, Davis SR. Obesity is associated with a poorer prognosis in women with hormone receptor positive breast cancer. Maturitas. 2014;79(3):279-86.
104. Rose DP, Vona-Davis L. Influence of obesity on breast cancer receptor status and prognosis. Expert Rev Anticancer Ther. 2009;9(8):1091-101.
105. Rose DP, Vona-Davis L. Interaction between menopausal status and obesity in affecting breast cancer risk. Maturitas. 2010;66(1):33-8.
106. Kaviani A, Neishaboury M, Mohammadzadeh N, Ansari-Damavandi M, Jamei K. Effects of Obesity on Presentation of Breast Cancer, Lymph Node Metastasis and Patient Survival: A Retrospective Review. Asian Pacific Journal of Cancer Prevention. 2013;14(4):2225-9.
107. Jarde T, Perrier S, Vasson MP, Caldefie-Chezet F. Molecular mechanisms of leptin and adiponectin in breast cancer. European journal of cancer. 2011;47(1):33-43.
108. Surmacz E. Leptin and adiponectin: emerging therapeutic targets in breast cancer. Journal of mammary gland biology and neoplasia. 2013;18(3-4):321-32.
109. Khandekar MJ, Cohen P, Spiegelman BM. Molecular mechanisms of cancer development in obesity. Nat Rev Cancer. 2011;11(12):886-95.
110. Belardi V, Gallagher EJ, Novosyadlyy R, LeRoith D. Insulin and IGFs in obesity-related breast cancer. Journal of mammary gland biology and neoplasia. 2013;18(3-4):277-89.
111. Grossmann ME, Ray A, Nkhata KJ, Malakhov DA, Rogozina OP, Dogan S, et al. Obesity and breast cancer: status of leptin and adiponectin in pathological processes. Cancer metastasis reviews. 2010;29(4):641-53.
112. Cleary MP, Grossmann ME, Ray A. Effect of obesity on breast cancer development. Vetinary Pathology. 2010;47(2):202-13.
113. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
114. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61-70.
115. Gadducci A, Biglia N, Sismondi P, Genazzani AR. Breast cancer and sex steroids: Critical review of epidemiological, experimental and clinical investigations on etiopathogenesis, chemoprevention and endocrine treatment of breast cancer. Gynecological Endocrinology. 2005;20(6):343-60.
116. Dressman MA, Walz TM, Lavedan C, Barnes L, Buchholtz S, Kwon I, et al. Genes that co-cluster with estrogen receptor alpha in microarray analysis of breast biopsies. Pharmacogenomics Journal. 2001;1(2):135-41.

117. Tozlu S, Girault I, Vacher S, Vendrell J, Andrieu C, Spyratos F, et al. Identification of novel genes that co-cluster with estrogen receptor alpha in breast tumor biopsy specimens, using a large-scale real-time reverse transcription-PCR approach. Endocrine-Related Cancer. 2006;13(4):1109-20.
118. Creighton CJ, Sada YH, Zhang Y, Tsimelzon A, Wong H, Dave B, et al. A gene transcription signature of obesity in breast cancer. Breast Cancer Research Treatment. 2012;132(3):993-1000.
119. Fuentes-Mattei E, Velazquez-Torres G, Phan L, Zhang F, Chou PC, Shin JH, et al. Effects of obesity on transcriptomic changes and cancer hallmarks in estrogen receptor-positive breast cancer. Journal National Cancer Institute. 2014;106(7).
120. Bruder CEG, Piotrowski A, Gijsbers A, Andersson R, Erickson S, de Stahl TD, et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. American journal of human genetics. 2008;82(3):763-71.
121. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.
122. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang YJ, et al. Origins and functional impact of copy number variation in the human genome. Nature. 2010;464(7289):704-12.
123. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, et al. Mapping copy number variation by population-scale genome sequencing. Nature. 2011;470(7332):59-65.
124. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature. 2006;444(7118):444-54.
125. A Database of Genomic Variants 2014. [A curated catalogue of human genomic structural variation]. Available from: http://dgv.tcag.ca/dgv/app/home.
126. Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature. 2010;463(7281):671-5.
127. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, et al. Large, rare chromosomal deletions associated with severe early-onset obesity. Nature. 2010;463(7281):666-70.
128. Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z, et al. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. Nature. 2011;478(7367):97-102.
129. Sha BY, Yang TL, Zhao LJ, Chen XD, Guo Y, Chen Y, et al. Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population. J Hum Genet. 2009;54(4):199-202.
130. Jarick I, Vogel CIG, Scherag S, Schafer H, Hebebrand J, Hinney A, et al. Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. Hum Mol Genet. 2011;20(4):840-52.
131. McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nature genetics. 2008;40(10):1166-74.
132. Locke DP, Sharp AJ, McCarroll SA, McGrath SD, Newman TL, Cheng Z, et al. Linkage disequilibrium and heritability of copy-number polymorphisms within

duplicated regions of the human genome. American journal of human genetics. 2006;79(2):275-90.
133. Oppenheim FG, Salih E, Siqueira WL, Zhang WM, Helmerhorst EJ. Salivary proteome and its genetic polymorphisms. Oral-Based Diagnostics. Annals of the New York Academy of Sciences. 10982007. p. 22-50.
134. Hoebler C, Karinthi A, Devaux MF, Guillon F, Gallant DJG, Bouchet B, et al. Physical and chemical transformations of cereal food during oral digestion in human subjects. Br J Nutr. 1998;80(5):429-36.
135. Fried M, Abramson S, Meyer JH. PASSAGE OF SALIVARY AMYLASE THROUGH THE STOMACH IN HUMANS. Dig Dis Sci. 1987;32(10):1097-103.
136. Rosenblum JL, Irwin CL, Alpers DH. STARCH AND GLUCOSE OLIGOSACCHARIDES PROTECT SALIVARY-TYPE AMYLASE ACTIVITY AT ACID PH. Am J Physiol. 1988;254(5):G775-G80.
137. Mandel AL, des Gachons CP, Plank KL, Alarcon S, Breslin PAS. Individual Differences in AMY1 Gene Copy Number, Salivary alpha-Amylase Levels, and the Perception of Oral Starch. PLoS One. 2010;5(10):9.
138. Mandel AL, Breslin PA. High endogenous salivary amylase activity is associated with improved glycemic homeostasis following starch ingestion in adults. Journal of Nutrition. 2012;142(5):853-8.
139. Mejia-Benitez MA, Bonnefond A, Yengo L, Huyvaert M, Dechaume A, Peralta-Romero J, et al. Beneficial effect of a high number of copies of salivary amylase AMY1 gene on obesity risk in Mexican children. Diabetologia. 2015;58(2):290-4.
140. Viljakainen H, Andersson-Assarsson JC, Armenio M, Pekkinen M, Pettersson M, Valta H, et al. Low Copy Number of the AMY1 Locus Is Associated with Early-Onset Female Obesity in Finland. PLoS One. 2015;10(7):e0131883.
141. Cantsilieris S, White SJ. Correlating Multiallelic Copy Number Polymorphisms with Disease Susceptibility. Human mutation. 2013;34(1):1-13.
142. Usher CL, Handsaker RE, Esko T, Tuke MA, Weedon MN, Hastie AR, et al. Structural forms of the human amylase locus and their relationships to SNPs, haplotypes and obesity. Nature genetics. 2015;47(8):921-+.
143. Morrin H, Gunningham S, Currie M, Dachs G, Fox S, Robinson B. The Christchurch Tissue Bank to support cancer research. The New Zealand Medical Journal. 2005;118(1225):U1735.
144. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001;29(9):6.
145. Fogel MR, Gray GM. STARCH HYDROLYSIS IN MAN - INTRALUMINAL PROCESS NOT REQUIRING MEMBRANE DIGESTION. Journal of Applied Physiology. 1973;35(2):263-7.
146. Barret ML, Udani JK. A proprietary alpha-amylase inhibitor from white bean (Phaseolus vulgaris): A review of clinical studies on weight loss and glycemic control. Nutrition Journal. 2011;10(24).
147. de Wijk RA, Prinz JF, Engelen L, Weenen H. The role of alpha-amylase in the perception of oral texture and flavour in custards. Physiology and Behavior. 2004;83(1):81-91.
148. Ferry AL, Hort J, Mitchell JR, Lagarrigue S, Pamies B. Effect of amylase activity on starch paste viscosity and its implications for flavor perception. Journal of Texture Studies. 2004;35(5):511-24.

149. Isaac JTR, Ashby M, McBain CJ. The role of the GluR2 subunit in AMPA receptor function and synaptic plasticity. Neuron. 2007;54(6):859-71.
150. Balko JM, Schwarz LJ, Bhola NE, Kurupi R, Owens P, Miller TW, et al. Activation of MAPK Pathways due to DUSP4 Loss Promotes Cancer Stem Cell-like Phenotypes in Basal-like Breast Cancer. Cancer research. 2013;73(20):6346-58.
151. Boudot A, Le Dily F, Pakdel F. Involvement of COUP-TFs in Cancer Progression. Cancers. 2011;3(1):700-15.
152. Smith M. GENETICS OF HUMAN ALCOHOL AND ALDEHYDE DEHYDROGENASES. Advances in Human Genetics. 1986;15:249-90.
153. Pearson G, Robinson F, Gibson TB, Xu BE, Karandikar M, Berman K, et al. Mitogen-activated protein (MAP) kinase pathways: Regulation and physiological functions. Endocrine reviews. 2001;22(2):153-83.
154. Hirosumi J, Tuncman G, Chang LF, Gorgun CZ, Uysal KT, Maeda K, et al. A central role for JNK in obesity and insulin resistance. Nature. 2002;420(6913):333-6.
155. Imajo M, Tsuchiya Y, Nishida E. Regulatory mechanisms and functions of MAP kinase signaling pathways. IUBMB Life. 2006;58(5-6):312-7.
156. R. RJ. ERK1/2 MAP kinases: Structure, function and regulation Pharmacological Research. 2012;66:105-43.
157. Balko JM, Cook RS, Vaught DB, Kuba MG, Miller TW, Bhola NE, et al. Profiling of residual breast cancers after neoadjuvant chemotherapy identifies DUSP4 deficiency as a mechanism of drug resistance. Nature Medicine. 2012;18(7):1052-+.
158. Armes JE, Hammet F, de Silva M, Ciciulla J, Ramus SJ, Soo WK, et al. Candidate tumor-suppressor genes on chromosome arm 8p in early-onset and high-grade breast cancers. Oncogene. 2004;23(33):5697-702.
159. Mazumdar A, Den Hollander P, Poage G, Hill J, Zhang Y, Tsimelzon A, et al. High expression of DUSP4 in ER-negative breast cancer cells suppresses growth and invasion. Cancer research. 2013;73(8, Suppl. 1):856.
160. Baglia ML, Cai QY, Zheng Y, Wu J, Su YH, Ye F, et al. Dual specificity phosphatase 4 gene expression in association with triple-negative breast cancer outcome. Breast cancer research and treatment. 2014;148(1):211-20.
161. Wang HY, Cheng ZY, Malbon CC. Overexpression of mitogen-activated protein kinase phosphatases MKP1, MKP2 in human breast cancer. Cancer Letters. 2003;191(2):229-37.
162. Liu Y, Du FY, Chen W, Yao MY, Lv KZ, Fu PF. Knockdown of dual specificity phosphatase 4 enhances the chemosensitivity of MCF-7 and MCF-7/ADR breast cancer cells to doxorubicin. Experimental Cell Research. 2013;319(20):3140-9.
163. Kim H, Jang SM, Ahn H, Sim J, Yi K, Chung Y, et al. Clinicopathological Significance of Dual-Specificity Protein Phosphatase 4 Expression in Invasive Ductal Carcinoma of the Breast. Journal of Breast Cancer. 2015;18(1):1-7.
164. Arslan AA, Gold LI, Mittal K, Suen TC, Belitskaya-Levy I, Tang MS, et al. Gene expression studies provide clues to the pathogenesis of uterine leiomyoma: new evidence and a systematic review. Human Reproduction. 2005;20(4):852-63.
165. Leja J, Essaghir A, Essand M, Wester K, Oberg K, Totterman TH, et al. Novel markers for enterochromaffin cells and gastrointestinal neuroendocrine carcinomas. Modern Pathology. 2009;22(2):261-72.

166. Beretta F, Bassani S, Binda E, Verpelli C, Bello L, Galli R, et al. The GluR2 subunit inhibits proliferation by inactivating Src-MAPK signalling and induces apoptosis by means of caspase 3/6-dependent activation in glioma cells. European Journal of Neuroscience. 2009;30(1):25-34.
167. Hu HX, Takano N, Xiang LS, Gilkes DM, Luo WB, Semenza GL. Hypoxia-inducible factors enhance glutamate signaling in cancer cells. Oncotarget. 2014;5(19):8853-68.
168. Pinheiro RL, Sarian LO, Pinto-Neto AM, Morais S, Costa-Paiva L. Relationship between body mass index, waist circumference and waist to hip ratio and the steroid hormone receptor status in breast carcinoma of pre- and postmenopausal women. The Breast. 2009;18(1):8-12.
169. Le Dily F, Metivier R, Gueguen MM, Christine LP, Flouriot G, Tas P, et al. COUP-TFI modulates estrogen signaling and influences proliferation, survival and migration of breast cancer cells. Breast cancer research and treatment. 2008;110(1):69-83.
170. Metivier R, Gay FA, Hubner MR, Flouriot G, Salbert G, Gannon F, et al. Formation of an hER alpha-COUP-TFI complex enhances hER alpha AF-1 through Ser118 phosphorylation by MAPK. The EMBO Journal. 2002;21(13):3443-53.
171. Winnier DA, Fourcaudot M, Norton L, Abdul-Ghani MA, Hu SL, Farook VS, et al. Transcriptomic Identification of ADH1B as a Novel Candidate Gene for Obesity and Insulin Resistance in Human Adipose Tissue in Mexican Americans from the Veterans Administration Genetic Epidemiology Study (VAGES). PLoS One. 2015;10(4):26.
172. Sonksen P, Sonksen J. Insulin: understanding its action in health and disease. British Journal of Anaesthesia. 2000;85(1):69-79.
173. Mehran AE, Templeman NM, Brigidi GS, Lim GE, Chu KY, Hu XK, et al. Hyperinsulinemia Drives Diet-Induced Obesity Independently of Brain Insulin Production. Cell Metabolism. 2012;16(6):723-37.