Identification of differentially expressed genes in omental adipose tissues of obesity and type2diabetes: A meta-analysis of microarray datasets

Preethi.V

(15PBI001)

Thesis submitted to

DEPARTMENT OF BIOCHEMISTRY, BIOTECHNOLOGY AND BIOINFORMATICS

Avinashilingam Institute for Home Science and Higher

Education for Women, Coimbatore - 641043

In Partial Fulfilment of the Requirements for the

Degree of Master of Science in Bioinformatics

April, 2017

CERTIFICATE

Identification of differentially expressed genes in omental adipose tissues of obesity and type2diabetes: A meta-analysis of microarray datasets

Preethi.V

(15PBI001)

Thesis submitted to

DEPARTMENT OF BIOCHEMISTRY, BIOTECHNOLOGY AND BIOINFORMATICS

Avinashilingam Institute for Home Science and Higher

Education for Women, Coimbatore - 641043

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Bioinformatics

April, 2017

10/04/2017

Signature of the

Head of the Department

10/4/2017

Signature of the Supervisor

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

*Every work on its backdrop has the blessing of **GOD ALMIGHTY**. Therefore I submit my reverential gratitude at the feet of lord Almighty and thank him for the grace showered on me in making the study a successful one.*

*I am grateful to **Dr. (Thiru) P.R. Krishna Kumar,** Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for providing all facilities necessary for the study.*

*My special debt to **Dr. (Tmt) Premavathi Vijayan,**Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for providing the opportunity and exposure to the world of knowledge.*

*My special thanks to **Dr. (Tmt) S.Kowsalya,** Registrar (i/c), Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for administrative support and for providing adequate help required to carry out the work.*

*My sincere thanks to **Dr. (Tmt) A. Parvathi**, Dean, Faculty of Science, Avinashilingam Institute for Home Scienece and Higher Education for Women, Coimbatore for her guidance and affectionate support, expert suggestions throughout the project.*

*My sincere thanks to **Dr. (Mrs.) S. Annapoorani,** Professor and Head, Department of Biochemistry, Biotechnology and Bioinformatics, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for, her guidance and moral support, encouragement given throughout the study period.*

*I would like to express my immense gratitude to my guide **Dr.N.Santhi,** Assistant Professor, Department of Biochemistry, Biotechnology and Bioinformatics, Avinashilingam Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for the opulent guidance rendered at every stage of the dissertation. I feel immense pleasure in extending my deep sense of gratitude for her motivation, crucial support, guiding suggestions and encouragement throughout the work.*

*I heart fully thank **Dr.M.Rajeswari,** Lecturer, Department of Biochemistry, Biotechnology and Bioinformatics, Avinashilingam Institute for Home Science and Higher Education for Women,*

for inspiring thoughts, constant support and tremendous care rendered for carrying out my thesis successfully.

I owe my special thanks to ***Dr.A.Pushpa,*** *Assistant Professor, Department of Biochemistry, Biotechnology and Bioinformatics, Avinashilingam Institute for Home Science and Higher Education for Women, for her moral support and timely help during this project work.*

I express my sincere thanks to all the ***Staff members*** *of the Department of Biochemistry, Biotechnology and Bioinformatics, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for their support and encouragement during the period of work.*

I whole heartedly express my deep sense of gratitude to my ***beloved parents*** *and* ***my sister*** *for their emotional support to complete the research work successfully.*

Preethi.V

CONTENTS

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1. INTRODUCTION

Over the past two decades there has been a large increase in the prevalence of metabolic disorders such as obesity and diabetes in both developed and developing countries. Diabetes is a major cause of several disability worldwide. It is a chronic disease affecting 9% of the population worldwide, with the prevalence predicted to double in the year 2030 by World Health Organization (WHO). Disabilities resulting from diabetes has grown substantially since 1990, with large increase among people aged 15-69 years (Krug, 2016).

Diabetes is a group of disorders characterized by chronic high blood glucose levels (hyperglycemia) due to the body's inability to produce any or enough insulin, the hormone that regulates blood glucose levels. There are two main types of diabetes. Type I diabetes, which often occurs in children or adolescents, is caused by the body's inability to make insulin and Type II diabetes, which occurs as a result of the body's inability to react properly to the produced insulin (insulin resistance). Type II is more prevalent than type 1 diabetes and is therefore seen in roughly 90% of all diabetic cases. It has become a major global problem and during the past few decades, the genetic basis of various monogenic forms of the disease and their underlying molecular mechanisms have been studies (Ashcroft and Rorsman, 2012).

Equally challenging is the problem of weight gain over the past 50 years, associated in part with major worldwide changes in caloric intake and dietary composition. In India, the prevalence of obesity is been increasing at a rapid rate due to the increase in energy intake owing to availability of high fat and energy-dense foods along with reduced energy expenditure due to urbanization and mechanization. In parallel to the rise in overweight and obesity, the prevalence of metabolic syndrome like type 2 diabetes mellitus is also increasing in India and reached epidemic populations (Gulati and Misra, 2014). Recent studies have established that obesity is associated with systemic chronic inflammation and that this low-grade inflammation may play vital role obesity-associated insulin resistance leading to type 2 diabetes.

Adipose tissue acts as an endocrine organ secreting variety of substance like inflammatory mediators for immune response and also other hormones that helps in metabolic regulation. Obesity alters the metabolic and endocrine functions that leads to an increased release of fatty acids, hormones and proinflammatory molecules that contribute to obesity-induced complications.

The elevated inflammatory status appears to originate from infiltrated macrophages or other emerging immune cells in adipose tissue. (Weisberg *et al.,* 2003).

In recent years, a large number of human population studies have linked insulin resistance to systemic inflammation with growing evidences that has pointed out the correlative and causative relationship due to proinflammatory cytokines produced in the adipose tissues. Adipose tissue inflammation is shaped by the regulation of pro- and anti-inflammatory immune cell homeostasis, and obesity skews this balance towards a pro-inflammatory status. Thus indeed, the role of the immune system in adipose tissues has become an exciting new area of research in the field of obesity and metabolic regulation owing to mediation of pathogenesis causing development of obesity-induced insulin resistance (Cheol and Lee, 2014).

Microarray technology has permitted the interrogation of nearly all expressed genes under a wide range of pathological conditions. The patterns of gene expression in response to obesity and diabetes have yielded important knowledge into the pathogenesis of diabetes and its relationship to obesity. Microarray studies have motivated research in adipose tissue that provides clues to the importance of inflammation in obesity with new adipocyte-derived hormones involved in insulin resistance. Recent studies have indicated several inflammatory mediators pertaining to the cause of diabetes via insulin resistance in obese individuals. The most significant ones being the TNF-α and IL-6 in adipose tissues at elevated levels causing response to insulin in glucose uptake. These evidences has stressed the importance of inflammatory mediated insulin resistance that accompany obesity. These studies have greatly expanded the understanding of mechanisms underlying the pathogenesis of obesity-induced diabetes by constructing casual network models linking gene expression with disease (Keller and Atte, 2010).

As large number of gene expression data profiling obesity with diabetes studies has become available, the identification of intersections between the disorders has also become interestingly popular. Such studies although conceptually simple, but face a number of technical challenges like inaccessible data, managing the high number of data and replicate errors. (Griffith *et al.,* 2006). To overcome these challenges, this study included the strategy of meta-analysis for reporting the differentially expressed genes between the obese and obese diabetic individuals.

Combining information from multiple existing studies can increase the reliability of results. The use of statistical techniques to combine results from independently related studies is called

“meta-analysis”. Through meta-analysis, studies can increase the statistical power for obtaining more precise estimate of gene expression differentials and access the heterogeneity of the overall estimate. Several meta-analysis techniques have been proposed in the context of microarrays and its advantages in gene expression have gone high rate by researchers in various fields (Ramasamy *et al.*, 2008).

The objective of the present study was to identify the differentially expressed genes between the Metabolically Healthy Obese (MHO) and Metabolically Unhealthy Obese (MUO) individuals with the gene expression data and finding the striking evidences associated with omental adipose tissue inflammation of inulin resistant obese subjects. Furthermore, the study also tends to examine the relationship between inflammatory mediators in adipose tissue of obese individuals pertaining to the causative agent of type 2 diabetes.

Hence the specific aim of the study was:

- ✓ To identify the differentially expressed genes between metabolically healthy obese and diabetic obese.
- ✓ To identify the biological function and gene enrichment pathway of differentially expressed genes.

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

The review of literature pertaining to the present study "Identification of differentially expressed genes in omental adipose tissues of obesity and type2diabetes: A meta-analysis of microarray datasets" is discussed below.

2.1 DIABETES MELLITUS

Diabetes mellitus is a group of diseases associated with various metabolic disorders, being the main feature is the chronic hyperglycemia due to insufficient insulin action (Seino *et al.,* 2010). It is associated and characterized with the abnormal levels of glucose in the blood with the underlying cause being varied in types including genetics, environmental factors, history of gestation diabetes, excess weight or obesity and sedentary lifestyle (Dall *et al.,* 2010).

2.2 CLASSIFICATION AND PATHOPHYSIOLOGY

The classification of diabetes mellitus is principally derived from etiology and includes the stages of pathophysiology based on the degree of deficiency and action of insulin. These disorders are classified into four groups (Table 1).

Table 1

Types of clinical conditions in diabetes mellitus (Belle *et al.,* 2011)

Condition	Description
Type I diabetes	It is a condition due to autoimmune disorder that leads to beta-cell destruction in the pancreas and it is characterized by complete lack of insulin production
Type II diabetes	It is a condition due to abnormal increased resistance to the action of insulin and the body's inability to produce enough insulin to overcome the resistance
Gestational diabetes	It is a condition which is a form of glucose intolerance that affects some women during pregnancy
Miscellaneous types	Diabetes caused due to other pathological disorders and associated with other clinical conditions like pancreatitis, cystic fibrosis, genetic defects of beta-cell function and infections

2.2.1 Type I diabetes

Type I diabetes is a chronic autoimmune disorder that occurs in genetically susceptible individuals due to various factors. In this type of diabetes mellitus, the body's own immune system attacks the β cells in the islets of Langerhans of the pancreas, causing damage to them which eventually eliminates the production of insulin hormone (Belle *et al.,* 2011). This destruction progresses within the body over months or years until β cell mass decreases to a point that insulin concentrations are no longer adequate to control the plasma glucose levels. This form of the disease only accounts for 5-10% of diabetic people, commonly affecting children and adults, but can also occur at any age. (Kerner and Bruckel, 2014).

2.2.2 Type II diabetes

Type II diabetes is generally characterized by insulin insensitivity that occurs due to insulin resistance, which declines the insulin production and eventually leads to pancreatic β cell failure, which in turn causes the decrease in glucose transport into the liver, muscle cells as well as the fat cells. And therefore, there is an increase in the breakdown of fat with hyperglycemia (Kerner and Bruckel, 2014).

This form of diabetes generally occurs due to various lifestyle factors like physical inactivity, sedentary lifestyle, cigarette smoking and consumption of alcohol. Above all, obesity has been found to contribute about 55% of the cases in type II diabetes. Recent studies have shown that environmental toxins are also causing eventual increase in the rate of developing type II diabetes (Olokoba *et al.,* 2012).

2.2.3 Gestational diabetes

Gestational diabetes mellitus (GDM) is a common medical complication that occurs in women which is defined as degree of glucose intolerance with onset or first recognition during pregnancy. By the current definition, GDM is included as a subgroup with more severe hyperglycemia that presents special issues concerning the management during pregnancy. The issues raised by this subgroup with those of GDM are of greater concern because of the rising prevalence of obesity, type II diabetes and other metabolic disturbances among the adult as well as the younger age groups (Metzger *et al.,* 2010).

2.2.4 Miscellaneous types of diabetes

Miscellaneous cause of diabetes mellitus account only of a small proportion of cases that includes genetic defects affecting the beta-cell function or insulin action whereas the causes for the occurrence in most cases is due to other disorders like pancreatitis, cyctic fibrosis, endocrinopathies and even drug-induced diabetes, most notable drugs like glucocorticoids, beta-blockers, protease inhibitors and therapeutic doses of niacin (Belle *et al*., 2011).

2.3 MECHANISM OF REGULATION OF BLOOD GLUCOSE

Insulin and Glucagon are the two important hormones secreted by the islet cells within the pancreas in response to blood sugar levels, working in contrary fashion. Understanding the role of insulin in physiological processes and its synthesis and secretion, along with the actions of molecular mechanisms has significant implications for understanding chronic disease like diabetes mellitus (Wilcox, 2010).

Insulin is a peptide hormone secreted by the β cells of the pancreatic islets of Langerhans which helps to maintain the normal blood glucose levels by facilitating cellular glucose uptake and regulating carbohydrate metabolism (Wilcox, 2010). The stimulus of insulin secretion is at high level of blood glucose, similarly as blood glucose falls, the amount of insulin secretion goes down.

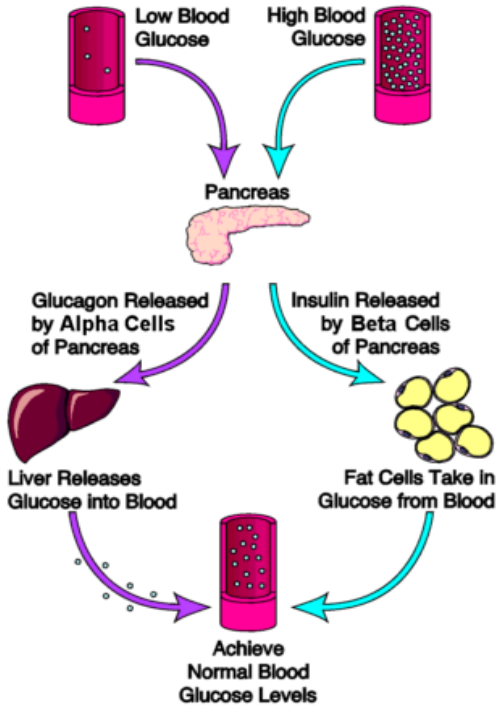


Figure 1. Mechanism of glucose regulation by the hormones in the pancreas
(Olokoba *et al.*, 2012)

As shown in Figure 1, insulin has an effect on number of cells - muscle cells, red blood cells as well as fat cells. In response to the hormone insulin, these cells absorb glucose from the blood, and shows the net effect of lowering the high blood glucose levels to the normal range.

Glucagon is secreted by the α cells of the pancreatic islets in the same manner as that of insulin except the mechanism of action being contrary to that of insulin. When the blood glucose is high, then glucagon is not secreted whereas when the blood glucose goes down (such as between meals or exercise) more of glucagon is secreted. Like insulin, glucagon also has more notable effects in all cells especially in the liver.

2.4 FACTORS CONTRIBUTING TO DIABETES

The pathogenesis of diabetes mellitus involves both genetic as well as environmental factors (Seino *et al.,* 2010). There are many risk factors for type II diabetes such as age, race, pregnancy, stress, certain medications, genetics, family history, cholesterol levels and obesity. However overweight and obesity are the crucial pathogenetic factors. Type II diabetes mellitus usually manifests only after the age of 40 years but it can also occur earlier, especially in populations with high prevalence of obesity (Ginter and Simko, 2012).

Although genotypic variants related to energy balance, the rapid emergence of type II diabetes in genetically diverse populations worldwide is most likely caused by environmental factors (Dyck *et al.,* 2010). Urbanization is associated with the change in lifestyle that in turn leads to lack of physical activity, unhealthy diet and obesity, all of which have been implicated as a contributing factor in developing diabetes (Yang *et al.,* 2010).

2.5 OBESITY AND DIABETES

Obesity is an important risk determinant of diabetes, with an increasing prevalence world-wide, mostly in developed countries (Vasan *et* al., 2014). Increasing evidences suggests that there is a complex relationship between obesity and diabetes. The influence of obesity on type II diabetes risk is not only determined by the degree of obesity but also by where the fat accumulates. Increased upper body fat including visceral adipose fat, as reflected in increased abdominal girth or waist-to-hip ratio, which is associated with the metabolic syndrome, type II diabetes. Beyond differences in body fat distribution, emerging evidence also suggests that different subtypes of adipose tissue may also affect glucose homeostasis differentially. Adult humans have limited and

variable number of brown fat cells, which play a role in the thermogenesis and potentially influence energy expenditure and obesity susceptibility (Eckel *et al*., 2011).

2.5.1 Obesity associated insulin resistance

Obesity has become a major contributing risk for developing type II diabetes, hypertension and nonalcoholic fatty liver disease (NAFLD) that reduces life expectancy leading to economic and social consequences. It is characterized by low-grade chronic inflammation that is evidenced by an increased systemic concentration of proinflammatory molecules which contribute substantially to insulin resistance (Bertola *et al.,* 2011).

Insulin resistivity fluctuates during different stages of life from puberty till the aging process, even based on lifestyle variations and physical activity. However, obesity is been considered to be the most important factor in case of development of metabolic disorders like diabetes. Adipose tissue increase in obese conditions forms the major factor affecting the glucose metabolism by certain hormone secretions like leptin, adiponectin, cytokines and other proinflammatory substance leading to the release of non-esterified fatty acids (NEFAs). The increased release of these NEFAs is commonly observed in obese individuals and this forms the association of obesity with insulin resistance (Goblan *et al.,* 2014).

Insulin resistance is defined as the inhibition of insulin stimulation of several metabolic pathways like glucose transport, glycogen synthesis and anti-lipolysis which is pathophysiologically linked to type II diabetes (T2DM). Insulin resistance can have many causes, being the most common case is obesity. The current defined mechanism of obesity associated insulin resistance is given as the possibilities of excessive nutrient intake that gets stored as the expanded adipose tissue. Adipose tissue being not only a storage site for excessive fat, it is also a metabolically active tissue that synthesizes and secretes a number of biologically active substances like proinflammatory cytokines, acute phase reactants, angiotensin II, leptin, resistin, adiponectin and others, some of these when secreted in large amounts can produce insulin resistance (Boden,2011).

2.5.2 Adipose tissue and inflammation

Adipose tissue is a complex form of tissue with adipocytes, immune cells and vasculature. Immune cells like macrophages are engaged to adipose tissues during obese conditions. Recent

researches have also proved that obesity is associated with progressive caspase-1 activation in adipose tissues. And also some functional consequences of inflammation of adipocytes as shown deficiency of certain signaling molecules in insulin signaling and other issues with insulin secretion. In adipose tissues of obese, the mechanisms like phosphorylation of AKT (protein kinase B), reduction of IL-1 (Interleukin-1) signaling, decreased glucose transport are more affected due to the inflammasome activation (Grant and Dixit, 2013).

Similarly, TNF-α is another form of active substance which is a cytokine but nor produced from macrophages. It has an important role to play in the body lipid and glucose metabolism. The production and expression of this cytokine in the adipose tissue is shown elevated in the models of obese subjects in certain studies. And recent studies has also indicated that the elevation of this cytokine in obesity is strongly correlated with hyperinsulinemia, which is a maker of insulin resistance (Hotamisligil *et al*., 2011).

The hormones and cytokines produced by adipose tissue are leptin, adiponectin, resistin and cytokines such as TNF-α and Ilβ. Leptin and adiponectin are considered to be true adipokines or primary adipokines because they are produced by adipocytes. The TNF-α, IL10 and most of the other inflammatory mediators are called as pseudo adipokines because expressed both in adipocytes as well as activated macrophages and other immune cell. Immune cells in adipose tissue especially macrophages correlates with obesity in terms of macrophage infiltration. Adipose tissue macrophages (ATMs) are a prominent source of proinflammatory cytokines, such as tumor necrosis factor (TNF) –α and interleukin (IL)-6 that can block insulin action in adipocytes. Thus, adipose tissue is a complex secretory active organ that modulate energy expenditure, appetite, insulin sensitivity, inflammation and immunity. (Shoelson *et al.,* 2007) (Ota, 2013).

2.5.3 Obesity – Proinflammatory site

Obesity is defined as the expansion of fat especially in abdominal fat depots leading to the risk of induction of metabolic diseases. Increasing evidence supports that obesity-induced inflammation is primarily mediated by immune cells such as macrophages in metabolic tissues in particular adipose tissue macrophages (ATMs) playing a critical role. The hypertrophic adipocytes produce inflammatory cytokines such as TNF-α, IL-6 and resistin. Growing evidences shows the increased additional acute phase reactants in obese subjects like TNF-α, IL-6 have direct effects on adipocyte insulin resistance (Shoelson *et al.*, 2007).

The altered production of proinflammatory molecule by adipose tissue has been implicated in the metabolic complications of obesity. In comparison with the lean individuals, the adipose tissue of obese expresses increased amount of proinflammatory proteins such as TNF-α, IL-6. The proinflammatory molecules have direct effect in the cellular metabolism. For instance, the TNF-α directly decreases insulin sensitivity and increases lipolysis in adipocytes (Weisberg *et al.*, 2003).

2.5.4 Inflammation – Connecting factor between obesity and insulin resistance

Obesity induced insulin resistance is associated with proinflammatory cytokines such as TNF-α, IL-6 and IL-1β. These cytokines activates diverse range of stress responsive and counter regulatory signaling pathway including activation of JNK, 1KKβ, mTOR, ERKs and SOCS proteins. These pathways collaborate to produce two metabolically important effects. First, each pathway converges upon inhibiting insulin signaling pathways through serine phosphorylation of IRs, which blunts insulin actin in stressed target tissues. Second, these signals converge on two inflammatory signaling pathways, JNK and 1KKβ initiating inflammatory response within metabolic tissues. In parallel with these actions, dysregulated nutrient intake also triggers innate immune receptors like Toll like receptor 4 and increased gut-derived lipopolysaccharide (LPS) translocation and intestinal dysbiosis.

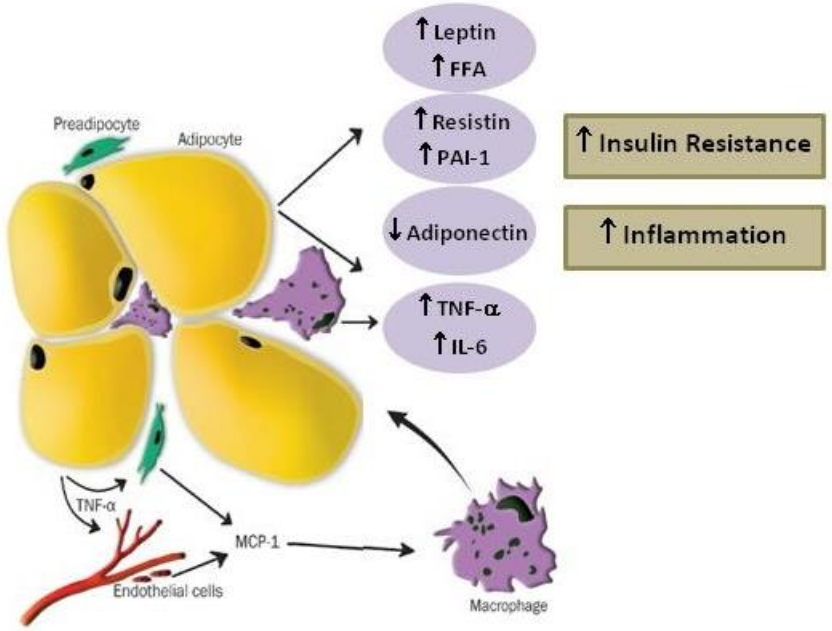


Figure 2. Tissue inflammation and insulin resistance (Xu *et al.*, 2003)

The chronic sub-acute inflammatory state that accompanies obesity is evident in affected tissues like adipose tissue, liver as well as in the vasculature. Moreover, the inflammatory changes associated with obesity can also be found in both immune and nonimmune cells. A large scale prospective studies have shown that markers of inflammation predict incident type II diabetes through resistance of insulin. However, the correlation between adipose tissue macrophages and body mass playing pathogenic roles in obesity induced insulin resistance is still an issue to be debated (Romeo *et al.,* 2012).

2.6 GLOBAL ESTIMATES AND PREVALENCE OF DIABETES AND OBESITY

Overweight and obesity is considered to be a global health burden in both developed and developing countries. The increased prevalence of obesity nowadays has drawn attention worldwide. Obesity is interlinked with many clinical and pathological outcomes the most devastating one being type 2 diabetes. The rising incidence of developing type 2 diabetes even among children and younger adults is also highly related to the epidemic of obesity. The prevalence is escalating with a pandemic need for stopping by considering the economic costs, social hazards and diseases (Goblan *et al.,* 2014).

Global estimates of diabetes prevalence have shown increases over the past 15 years. King *et al.,* 1998, projected that the number of adults with diabetes would reach about 300 million by 2025, and in 2004, the World Health Organization (WHO) estimated that by 2030 the number would even exceed 366 million. It also estimates that the countries large adult populations are with high susceptibility to diabetes. China has the highest number of people with diabetes with over 98.4 million adult population being affected ; which is followed closely by India with 65.1 million (Guariguata *et al.,* 2014).

On March 8, 2017, the International Diabetes Federation (IDF) announced that 199 million obese women are living with diabetes and this total is projected to increase to 313 million by 2040. Also reported that 415 million people were living with diabetes in the world. The World Health Organization (WHO) estimates that 90 per cent of people around the world have type II diabetes (International Diabetes Federation, 2017). The disease is responsible for 4.6 million deaths each year or even one every seven seconds. It affects ~12% of US adults and >2.5% of people over the age of 65. Diabetes is no longer restricted to the western world, the increased incidence of the

disease in next few decades are even expected to be in China and India. And these figures serve to emphasize that there is a fast-growing diabetes pandemic currently (Ashcroft and Rorsman, 2012).

Being the third leading cause of mortality, diabetes seriously threatens to human health worldwide and it has caused large disease burden to the patients, their families and the society, especially in developing countries (Liu *et al.,* 2016).

2.7 GENETIC FACTORS LINKING OBESITY AND DIABETES

Type II diabetes and obesity are complex diseases with sharing polygenic susceptibility. Recent genome wide association (GWA) studies have provided significant insight towards gene discovery and found about 30 common variant loci associating obesity with type II diabetes. Environmental determinants of these complex diseases have been well characterized by wealth of epidemiological literature but very few is been understood about the mechanism or interaction with genetic susceptibility. The epigenetic mechanisms regulate gene expression induced through somatic and other germline pathways (Bell *et al.,* 2010).

A few of Genome-wide association studies (GWAS) and candidate gene approaches have identified approximately 40 genes linking obesity and diabetes. Among which, most of the genes appear to be related with β-cell dysfunction, that are involved in pathways related to insulin resistance (Eckel *et al*., 2011). Recent discoveries in the analysis of gene expression levels of diabetic subjects has shown the following significant genes developing type II diabetes, includes TCF7L2, PPARG, FTO, KCNJ11, NOTCH2, WFS1, CDKAL1, IGF2BP2, SLC30A8, JAZF1 and HHEX associating obesity (Olakoba *et al.*,2012).

2.8 GENE EXPRESSION ANALYSIS-A KEY TO UNDERSTAND DISEASE BETTER

High-throughput sequencing strategies is been improving rapidly as the standard method for quantifying RNA expression levels and also with rapid sequencing technologies at reduced costs, studying detailed profile of gene expression levels has adopted a great significance in molecular biology and clinical research. One of the main aim of these experiments is to detect the differentially expressed genes in one or more conditions in terms of certain measures. These measures are fundamentally obtained by probe based methods like microarrays (Rapaport *et al*., 2013)

Microarray technologies are frequently used in molecular biology research to gain knowledge on transcriptional activity in various tissues or population of cells. These profiles of data can then be compared in order to identify the gene expression of a particular condition or phenotype. Such experiments has become the most powerful tool in functional genomics, giving insight about the normal cellular processes as well as disease pathogenesis (Ritchie *et al.*, 2015).

The analysis of gene expression patterns obtained from normal and pathological conditions is also valuable tool in the discovery of therapeutics and as diagnostic markers. At the moment, the measurement of gene expression using microarrays has been the sole approach for gene characterization and its requirement in functional genomics. Genome-wide association studies (GWAS) have revolutionized the ability to understand and identify the casual determinants for common human diseases in a different perspective and level (Zhong *et al.*, 2010).

Similarly Genome-wide transcriptome analysis using expression arrays have recently gained popularity to acquire insight into disease pathogenesis, molecular analysis and even for identification of biomarkers for progression as well as treatment responses (Woroniecka *et al.*, 2011). It is also becoming clear that most of the non-communicable diseases are the result of complex combination of genetic processes and related mechanisms. Despite the reason of both genetic and environmental factors contributing to diseases, many recent studies have been emphasized to show genetic components (Patel et al 2010).

2.8.1 Microarray data analysis

Microarrays are one of the successful techniques in the field of molecular biology that allow us to monitor the expression levels of even ten thousand of genes simultaneously. Arrays have been applied to studies in gene expression, genetic mapping, and discrimination of SNP, determining transcription factor activity and toxicity, pathogen identification and many other applications (Selvaraj and Natarajan, 2011).

Microarray technologies are in the forefront of managing huge amount of genomic data over several period and have also been evolved. Over the decades, several variants have been developed and sophisticated measurements have been improved to get valuable insights about the available data. Also much more choices in advances for handling such data is been implemented and improving with accessibility, quality and interpretation. Because the need for augmentation

and integration of microarray data is also rapidly increasing due to the churning data (Ruskin, 2016).

Over the last two years, the gene expression database contents has grown to rapid rate. Although high throughput sequencing (HTS)-based experiments account only for 6% of the entire content of the database, the proportion of new HTS submissions has been grown exponentially over the past few years from 2% in 2009 to 6% in 2010, 7% in 2011 and 15% in 2012. By considering the application of HTS data, 50% of the experiments used RNS-seq only, 32% CHIP-seq data and the remaining experiments uses DNA-seq for genotyping or copy number variations and others (Rustici *et al*, 2013).

Gene Expression Omnibus (GEO) database has made a priority to support the microarray community by switching to next generation sequence technologies. GEO accepts and stores sequence data for studies that examine gene expression (RNA-seq), gene regulation and epigenomics like CHIP-seq, methyl-seq or other studies including DNase hypersensitivity. GEO hosts processed data files together with sample and study metadata along with the raw data files containing the original sequence reads linked with National Centre for Biotechnology Information (NCBI). To date, GEO has loaded >44 terabases of read data with furthermore several thousand processed data files being incorporated into NCBI's epigenomics database which are curated and available for viewing by genome browsers (Barrett *et al.,* 2012).

A large number of experiments in Genome wide expression microarray measurements have been performed over the past few decade, however a very little have been successful in fully identifying the functionally important genes in the pathogenesis of diabetes. Because of the large number of genes often detected as significant in the microarray experiment, it is hard to subset the optimal candidate genes from individual studies. More recently, investigators have applied microarrays to genetics by considering gene expression levels for quantitative traits and also for finding relations between gene variants, with successful application to the identification of genes and targets for diabetes (Kodama, 2012).

2.9 HIGHER LEVEL DATA ANALYSIS

Modern molecular biology data presents major challenges for statistical methods that used to detect differential expression. Especially in case of microarrays the abundance of a particular transcript as a measure of fluorescence intensity is a continuous response, whereas for digital gene

expression (DGE) data the abundance is observed in count. Therefore, the procedures that are successful for microarray data cannot be applicable to DGE data. It is expected that emerging digital expression technologies will overtake microarray technologies in the near future for its application in many functional genomic approaches (Robinson 2010).

Microarray datasets are generally huge, and analytical precision is commonly influenced by a number of variables making it trivial for analysis. Thus it is extremely useful to reduce the dataset to those genes that are best distinguished between the two conditions or cases that is considered (eg: normal vs diseased). Such data analysis produce a list of genes whose expression is considered for the change and are known as differentially expressed genes. This identification of differentially expressed genes is the first task of an in depth microarray analysis and are done in two methods, i.e. clustering and comparison (Selvaraj and Natarajan, 2011).

The differential analysis of genes at the transcript level requires the use of several sophisticated statistical models in order to estimate, rather than to count the expression levels of the transcripts. These estimates can be processed using different normalization algorithms that can be made from the Bioconductor package in R, which proposes the detection of differential exon usage and applying DEseq normalizations. The result is usually a vector of ratios and their associated gene identifiers for series of samples forming a two way data matrix that can be taken for further analysis. At this stage, a variety of methods can be adopted to coax desired information from the data based on the nature of experiment. Typical goals include: (1) the identification of genes exhibiting up-regulation and down-regulation. (2) the clustering or classification of genes based on the expression in multiple samples, (3) the identification of genes that may be used as biological markers and (4) elucidation of gene functions and mechanism of interactions or pathways and gene networks (Dillies *et al.,* 2012).

2.9.1 Meta-analysis

A meta-analysis is the approach of combining data from several independent studies one primarily which addresses the similar query to produce a single estimate. It involves the statistical approach of collecting large amount of data and combining it into a single group and then retrieving related results from the individual studies for the purpose of useful integration of findings. Thus the term has been used to denote a complete range of quantitative approaches for review of researches. These are generally explained as studies of studies providing logical framework to a similar research review (Gopalakrishnan and Ganeshkumar, 2013).

Meta-analysis is a statistical approach of quantitative, formal and epidemiological study design that access the reports of primary research to derive conclusions on the body of research. The outcomes of meta-analysis may include more precise estimates of the effect of a particular disease, risk factors of it or other outcomes that relates the pooled analysis. Meta-analysis are the subsets of systematic review that attempts to collate evidences that firmly fits the eligibility criteria in answering the current research query. The key requirements are the clearly stated objectives with defined eligibility criteria, reproducible methodology, a search to identify the eligible studies relating the objective, assessment of validity of the research findings and systematic presentation of the findings of the study (Haidich, 2010).

The benefits in meta-analysis involves the ability to make improvisation in the power of minor or hypothetical studies to obtain proper conclusions. It also has the ability of identifying sources of various categories of studies. The concept of this pooled analysis also reveals the heterogeneity among populations and their effects to certain cases. Meta analysis also helps in detecting the bias in the design, conduct as well as interpretation of the research. On the contrary, meta-analysis is limited to certain cases due to its disadvantageous in reporting the quality of the original studies, method misapplications, mishandling of analysis and few unaddressed potential bias. But however, the scope and necessity of meta-analysis being required at an alarming rate, the advantages masks the demerits in its application (Ioannidis and Lau, 2002).

And thus with the advent of these technologies for understanding pathological conditions at molecular level, the meta-analysis of microarray data serves as an effective tool for interpreting the underlying differential gene expressions and variations in different samples. Similarly, a meta-analysis of these datasets with different grouping based on conditions of the sample will give a clear insight on the genetics of diabetes linked with obesity.

METHODOLOGY

3. METHODOLOGY

The experimental design adopted in the study for the analysis of microarray data is briefed in this section. The following databases and software tools were used to analyze the microarray datasets for differential expression of genes.

3.1 MICROARRAY DATA REPOSITORIES

Microarray data repositories are large collections of data that are implemented from different array experiments to serve the research community. There are several repositories maintaining huge amount of microarray expression data providing open source access to retrieve and evaluate gene expression.

The databases NCBI GEO (Gene Expression Omnibus) and EBI ArrayExpress were utilized in the present study to obtain gene expression profile form microarray experiments.

3.1.1 GENE EXPRESSION OMNIBUS (GEO)

Gene Expression Omnibus (GEO) database maintained at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/geo/) has been served as a central hub for microarray data. It has datasets and profiles obtained from high-throughput gene expression data, hybridization arrays, chips and microarrays with curated online resource for query and retrieval. As additional features, it is also provided with supporting tools for downloading experimental samples and analysis. Each dataset in GEO provides knowledge in four sections as the platform indicating the array template; sample record holding the hybridization data; series features that experimentally relates samples together and along with raw data supplied optional.

3.1.2 ARRAY EXPRESS

The ArrayExpress archive is another international public database developed and maintained by the European Molecular Biology Laboratory (www.ebi.ac.uk/arrayexpress/) with collection of functional genomics experimental expression data serving facilities to query and as well as download data. It serves in two parts, one as the ArrayExpress Repository (ArrayExpress Archive) with MIAME supported archive of microarray data and the ArrayExpress Data Warehouse (Gene Express Atlas) with expression profiles from microarray experiments that area consistently re-annotated. This archive is also integrated with other sequence databases at the

European Bioinformatics Institute. The query interface in ArrayExpress allows to find expression profiles with list of matches relating the query with much more informative than the other gene expression databases and advanced interface options showing sample attributes like disease conditions, cell types and anatomy along with several links to other integrated databases.

3.2 DATASET SERACH AND RETRIEVAL

A search was conducted in the above mentioned databases for microarray datasets using the keyword "Obesity and Diabetes". A list of 2259 hits were found related to the search, among which based on the organism type-*Homo sapiens,* 392 datasets were selected. Information based on age, clinical condition, pathology, tissue type were gathered from these datasets of *Homo sapiens* species and further scrutinized based on several inclusion and exclusion criteria as follows:

3.2.1 IDENTIFICATION OF ELIGIBLE DATASETS FROM THE SEARCH

3.2.1.1 *Inclusion criteria*

Among all the datasets from *Homo sapiens,* experiments performed only on obese subjects with/without diabetes were included in the study. And also studies carried only on adipose tissue origin was considered, which was further focused on Omental adipose tissues.

3.2.1.2 *Exclusion criteria*

Experiments carried out in obese subjects under pathological conditions other than diabetes like polycystic ovarian syndrome (PCOS) women, biatric surgery patients and patients undergoing drug treatment were excluded in this study. Further, datasets related with visceral and liver tissues, stem cells and blood cells and children were excluded.

3.2.3 RETRIEVAL OF DATASETS

A total of four datasets with the common array platform (Affymetrix) with the annotation type hgu133plus2 array was selected and downloaded from the NCBI Gene Expression Omnibus (GEO). Table 2 shows the sample details about the datasets collected for the present study. Data were extracted from the four studies and the raw data were downloaded.

Table 2: Datasets selected for the study

	Accession No	No of samples	Conditions	
			Obese	Diabetic Obese
1	GSE15773	10	5	5
2	GSE29410	3	3	-
3	GSE20950	20	10	10
4	GSE71416	20	6	14
Total			53	

3.3 META-ANALYSIS OF MICROARRAY DATASETS

A total of 53 samples from the selected datasets were pooled into a single file containing all the .cel files of raw data. The open source software AltAnalyze and R programming environment was used to carry out the different layers of microarray data analysis as shown in the Figure 4.

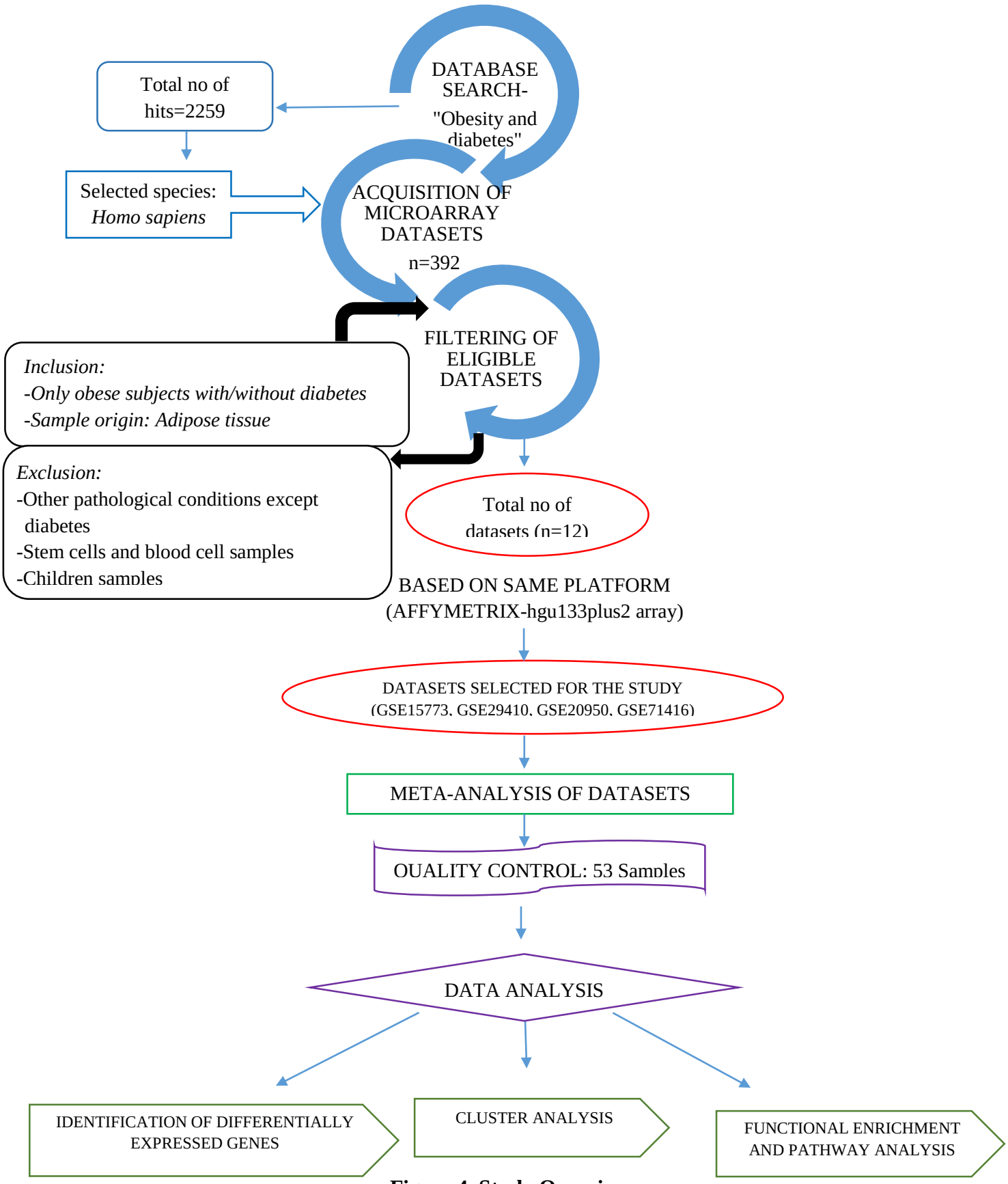


Figure 4. Study Overview

3.3.1 PROCESSING OF RAW DATA

In the analysis of microarray datasets, to improve the ability to detect outliers and defects in the sample it is important to measure the quality of the datasets before subjecting it to any analysis. Because samples pooled from different datasets tend to have replicates and also hybridized experiments may possess intensity errors in their design.

Thus the task of analyzing microarray data initially starts with the processing of the raw data in terms of quality control and normalization. The raw .CEL files of the datasets obtained for the study were processed using AltAnalyze. Quantile normalization was done to remove all the outliers. For each sample, the intensity values were log transformed and normalized to mean variance.

3.3.2 DIFERENTIAL GENE EXPRESSION ANALYSIS

To find the differentially expressed genes between the obese and diabetic obese, the data from CEL files were analyzed using AltAnalyze and the database version as EnsMart72. The two groups were compared using moderated t-test and linear model for microarray data (LIMMA) analysis with the cut-off of minimum fold change of 2.0 and maximum p-value of 0.05 (fold change>2.0 and p-value<0.05).

3.3.3 CLUSTER ANALYSIS

The differentially expressed genes among the groups were clustered and ordered by hierarchical clustering algorithm. The clustering was performed using the “ggplot” package in R and the heat map was constructed. Using this heat map, the expressed genes were represented graphically by coloring each cell on the basis of measured fluorescence ratio with fold change and p-value.

3.3.4 GENE ONTOLOGY AND PATHWAY ENRICHMENT ANALYSIS

The differentially expressed genes were further subjected to gene ontology studies using the online source program GENECODIS (www.genecodis.cnb.csic.es) and the pathway analysis was carried out using the program WebGestalt (www.webgestalt.org) to find the significant positive and negative regulating pathways of the DE genes.

RESULTS AND DISCUSSION

4. RESULTS AND DISCUSSION

The results obtained from the present study entitled "Identification of differentially expressed genes in omental adipose tissues of obesity and Type 2 diabetes: A meta-analysis of microarray datasets" are discussed in this section.

4.1 OVERVIEW OF THE DATASETS AND SOFTWARES INCLUDED IN THE STUDY

In this study, a total of 4 microarray datasets with 53 samples grouped under two conditions, among which 24 samples of obese and 2 samples of diabetic obese as summarized in Table:2 were taken for meta analysis. The study used the software AltAnalyze for carrying out the various layers of microarray data analysis and two more web tools WebGestalt (a Web based Gene Set Analysis Toolkit) and GENECODIS (Gene annotation co-occurrence discovery) were used for further analysis of the genes.

4.2 DATA PROCESSING AND NORMALIZATION

The preprocessing of microarray data includes the normalization of the data in which the expression ratios are log transformed into a reasonable measure prior to the detection of differentially expressed genes. It is the process of eliminating variations in the dataset that allows appropriate comparison of data. (Do and Choi, 2011). The average raw intensity signals in each array samples of the dataset against their prominent intensity based on the mean quantile variations were compared before and after normalization and represented in the form of a box plot **(**Figure 4 and 5)

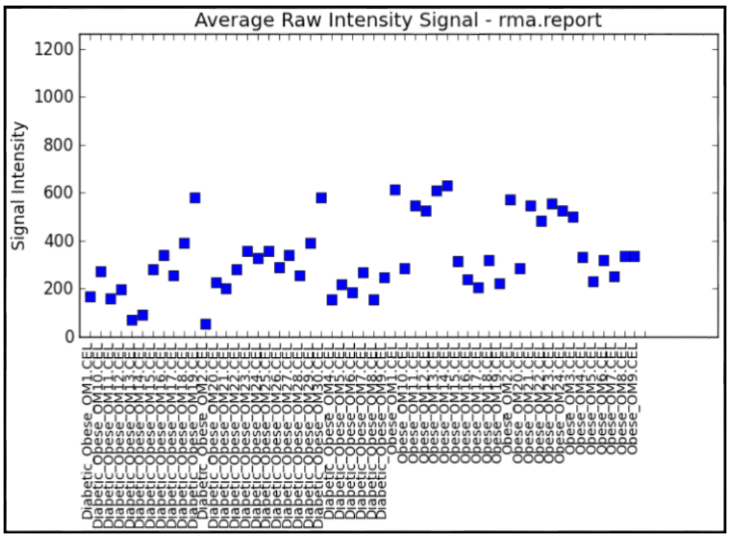


Figure 4. Plot showing the average raw intensities of each samples before normalization

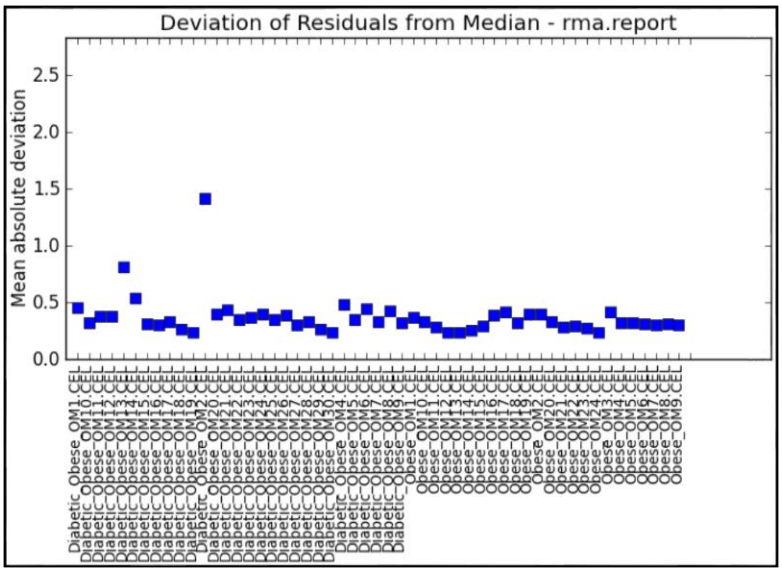


Figure 5. Plot showing the average raw intensities of each samples after normalization

The probe distribution of each array based on the log2 expressions were identified and measured (Figure 6)

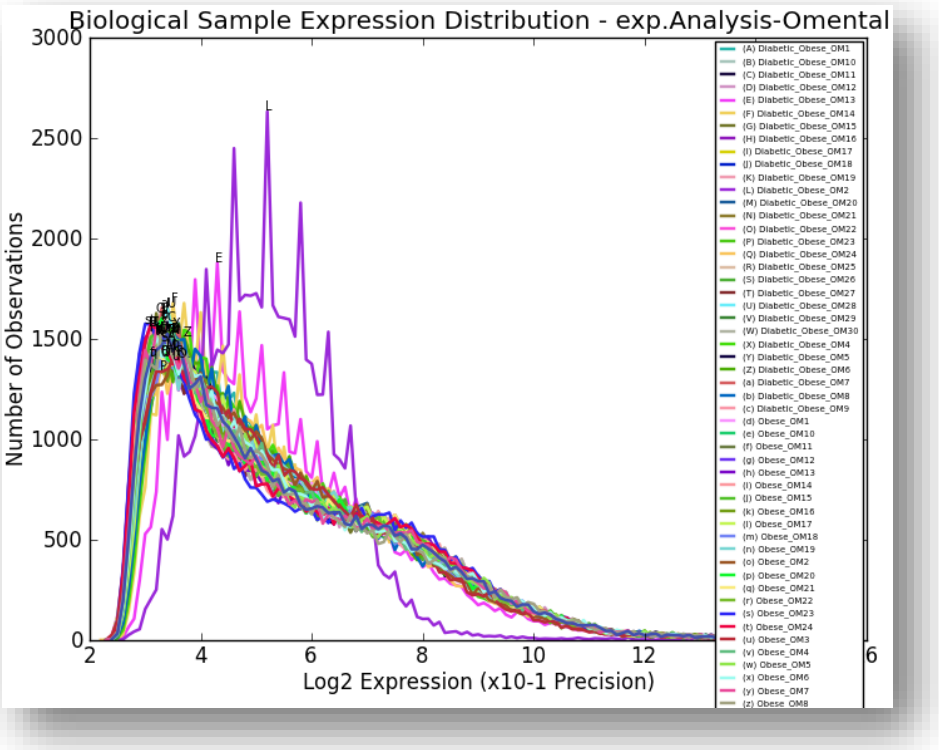


Figure 6. Expression level of each sample based on log transformation

4.3 DETECTION OF DIFFERENTIALLY EXPRESSED GENES

To identify the DE genes among the two conditions obese and diabetic obese, the .cel files of each samples were log transformed (base 2) and the filtered based on their raw p-value. Based on the linear model fit (LIMMA) analysis, the DE genes between the obese and diabetic obese conditions were observed.

A total of 145 genes were found to show altered expression in samples of diabetic obese with obese controls with the p-value < 0.05 and log fold change 2.0. The differentially expressed genes are listed in Table 3.

Table 3

List of differentially expressed genes in obese and diabetic obese in omental adipose tissue

S.no	GeneID	Gene symbol	Fold change	p-value	Description
1.	217427_s_at	HIRA	-1.07908	7.22E-08	histone cell cycle regulator
2.	214461_at	LBP	1.201253	3.66E-07	lipopolysaccharide binding protein
3.	214584_x_at	ACACB	-1.26223	9.30E-07	acetyl-CoA carboxylase beta
4.	238013_at	PLEKHA2	-1.04697	1.08E-06	pleckstrin homology domain containing A2
5.	235210_s_at	SBSPON	-1.16772	4.05E-06	somatomedin B and thrombospondin type 1 domain containing
6.	1558080_s_at	DNAJC3	-1.01318	6.93E-06	DnaJ heat shock protein family (Hsp40) member C3
7.	228948_at	EPHA4	-1.03936	7.23E-06	EPH receptor A4
8.	206423_at	ANGPTL7	-1.19089	1.02E-05	angiopoietin like 7

9.	204469_at	PTPRZ1	-1.11958	1.14E-05	protein tyrosine phosphatase, receptor type Z1
10.	228582_x_at	MALAT1	-1.43902	1.33E-05	Metastasis associated lung adenocarcinoma transcript 1 (non-protein coding)
11.	203413_at	NELL2	-1.19819	1.39E-05	neural EGFL like 2
12.	208612_at	PDIA3	-1.12742	1.64E-05	protein disulfide isomerase family A member 3
13.	209726_at	CA11	-1.0581	2.21E-05	carbonic anhydrase 11
14.	212249_at	PIK3R1	-1.04344	3.03E-05	phosphoinositide-3-kinase regulatory subunit 1
15.	219866_at	CLIC5	-1.23052	3.07E-05	chloride intracellular channel 5
16.	227510_x_at	NA	-1.71384	3.09E-05	Transcribed locus
17.	214693_x_at	NBPF10	-1.15786	3.34E-05	neuroblastoma breakpoint family, member 10
18.	219529_at	CLIC3	-1.21021	3.61E-05	chloride intracellular channel 3
19.	222562_s_at	TNKS2	-1.01812	4.03E-05	tankyrase 2
20.	223578_x_at	MALAT1	-1.24686	4.39E-05	Metastasis associated lung adenocarcinoma transcript 1 (non-protein coding)
21.	201430_s_at	DPYSL3	-1.05869	4.40E-05	dihydropyrimidinase like 3
22.	233314_at	PTEN	-1.13578	5.10E-05	phosphatase and tensin homolog
23.	223170_at	TMEM98	-1.08081	5.48E-05	transmembrane protein 98

24.	219791_s_at	HAND2-AS1	-1.09222	5.71E-05	HAND2 antisense RNA 1 (head to head)
25.	232449_at	BCO2	-1.01099	6.13E-05	beta-carotene oxygenase 2
26.	210092_at	MAGOH	-1.06064	6.74E-05	mago-nashi homolog, proliferation-associated (Drosophila)
27.	239893_at	NA	-1.23692	0.000102	Transcribed locus
28.	242541_at	ABCA9	-1.24635	0.000104	ATP binding cassette subfamily A member 9
29.	214587_at	COL8A1	-1.06308	0.000113	collagen type VIII alpha 1 chain
30.	229994_at	NFIA	-1.08604	0.000119	nuclear factor I/A
31.	237356_at	NA	1.072321	0.000123	Transcribed locus
32.	204364_s_at	REEP1	-1.28983	0.000133	receptor accessory protein 1
33.	224321_at	TUG1	1.257781	0.000136	taurine upregulated 1 (non-protein coding)
34.	215990_s_at	BCL6	-1.05596	0.000165	B-cell CLL/lymphoma 6
35.	232267_at	ADGRD1	-1.05903	0.000168	adhesion G protein-coupled receptor D1
36.	234103_at	KCNT2	-1.4528	0.000168	potassium channel, subfamily T, member 2
37.	224582_s_at	NUCKS1	-1.21041	0.000176	nuclear casein kinase and cyclin dependent kinase substrate 1
38.	219895_at	TMEM255A	-1.19822	0.00021	transmembrane protein 255A
39.	219869_s_at	SLC39A8	-1.23907	0.00022	solute carrier family 39 member 8

40.	1565717_s_at	FUS	-1.12945	0.00022	FUS RNA binding protein
41.	225768_at	NR1D2	-1.07875	0.000227	nuclear receptor subfamily 1 group D member 2
42.	206033_s_at	DSC3	-1.1232	0.000249	desmocollin 3
43.	1553171_x_at	LRRN4	-1.07218	0.000251	leucine rich repeat neuronal 4
44.	209264_s_at	TSPAN4	-1.13637	0.000267	tetraspanin 4
45.	219929_s_at	ZFYVE21	-1.08398	0.000282	zinc finger FYVE-type containing 21
46.	208937_s_at	ID1	-1.41504	0.000319	inhibitor of DNA binding 1, HLH protein
47.	206932_at	CH25H	1.753048	0.000322	cholesterol 25-hydroxylase
48.	204358_s_at	FLRT2	-1.15475	0.000374	fibronectin leucine rich transmembrane protein 2
49.	206658_at	UPK3B	-1.24682	0.000422	uroplakin 3B
50.	204004_at	PAWR	-1.02005	0.000445	pro-apoptotic WT1 regulator
51.	223940_x_at	MALAT1	-1.45132	0.000498	metastasis associated lung adenocarcinoma transcript 1 (non-protein coding)
52.	206404_at	FGF9	-1.10661	0.000526	fibroblast growth factor 9
53.	205018_s_at	MBNL2	-1.29986	0.000546	muscleblind like splicing regulator 2
54.	222853_at	FLRT3	-1.4561	0.000552	fibronectin leucine rich transmembrane protein 3

55.	213901_x_at	RBFOX2	-1.08521	0.000577	RNA binding protein, fox-1 homolog 2
56.	230040_at	ADAMTS18	1.129118	0.000629	ADAM metallopeptidase with thrombospondin type 1 motif 18
57.	216953_s_at	WT1	-1.01105	0.000663	Wilms tumor 1
58.	210655_s_at	FOXO3	-1.18308	0.00068	forkhead box O3
59.	244455_at	KCNT2	-1.02856	0.000708	potassium sodium-activated channel subfamily T member 2
60.	231576_at	NA	-1.02437	0.000719	Transcribed locus
61.	230165_at	SGO2	-1.12643	0.000729	shugoshin 2
62.	234074_at	NA	-1.10352	0.000776	CDNA FLJ10946 fis, clone PLACE1000005
63.	201551_s_at	LAMP1	-1.22821	0.000783	lysosomal associated membrane protein 1
64.	224856_at	FKBP5	-1.05915	0.000803	FK506 binding protein 5
65.	224589_at	XIST	-1.94048	0.000836	X inactive specific transcript (non-protein coding)
66.	205321_at	EIF2S3	-1.05901	0.000933	eukaryotic translation initiation factor 2 subunit gamma
67.	214040_s_at	GSN	-1.01459	0.00109	gelsolin
68.	212344_at	SULF1	-1.09357	0.001094	sulfatase 1
69.	208451_s_at	C4B	-1.28248	0.001095	complement component 4B (Chido blood group)

70.	205433_at	BCHE	-1.14942	0.00127	butyrylcholinesterase
71.	1552487_a_at	BNC1	-1.53023	0.001415	basonuclin 1
72.	226237_at	COL8A1	-1.07572	0.001571	collagen type VIII alpha 1 chain
73.	218182_s_at	CLDN1	-1.21503	0.001594	claudin 1
74.	205082_s_at	AOX1	-1.10488	0.001618	aldehyde oxidase 1
75.	201008_s_at	TXNIP	-1.2793	0.001739	thioredoxin interacting protein
76.	219250_s_at	FLRT3	-1.07085	0.001785	fibronectin leucine rich transmembrane protein 3
77.	217294_s_at	ENO1	-1.20698	0.001941	enolase 1
78.	206392_s_at	RARRES1	-1.19569	0.001969	retinoic acid receptor responder 1
79.	213872_at	NA	-1.01636	0.001972	Transcribed locus, strongly similar to NP_001017510.1 uncharacterized protein LOC498750
80.	1555564_a_at	CFI	-1.32135	0.002002	complement factor I
81.	201294_s_at	WSB1	-1.01109	0.002077	WD repeat and SOCS box containing 1
82.	1558214_s_at	CTNNA1	-1.37416	0.002143	catenin alpha 1
83.	229733_s_at	NA	-1.02497	0.002267	Transcribed locus
84.	204560_at	FKBP5	-1.2796	0.002686	FK506 binding protein 5
85.	204885_s_at	MSLN	-1.13857	0.002687	mesothelin
86.	1557910_at	HSP90AB1	-1.24806	0.002756	heat shock protein 90 alpha family class B member 1

87.	214279_s_at	NDRG2	-1.31925	0.002764	NDRG family member 2
88.	222330_at	NA	-1.05723	0.002796	Transcribed locus
89.	217023_x_at	TPSAB1	-1.18115	0.002855	tryptase alpha/beta 1
90.	1559954_s_at	DDX42	-1.19437	0.00289	DEAD-box helicase 42
91.	224568_x_at	MALAT1	-1.31194	0.002974	metastasis associated lung adenocarcinoma transcript 1 (non-protein coding)
92.	231181_at	NA	-1.28974	0.003117	Transcribed locus
93.	216474_x_at	TPSB2	-1.0653	0.003282	tryptase beta 2 (gene/pseudogene)
94.	205960_at	PDK4	-1.6032	0.003334	pyruvate dehydrogenase kinase 4
95.	213350_at	RPS11	1.04709	0.003544	Ribosomal protein S11
96.	1555229_a_at	C1S	-1.69568	0.00359	complement C1s
97.	214945_at	FAM153B	-1.06703	0.003618	family with sequence similarity 153, member B
98.	210317_s_at	YWHAE	-1.17764	0.003709	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon
99.	221321_s_at	KCNIP2	-1.0644	0.0038	potassium voltage-gated channel interacting protein 2
100.	1555230_a_at	KCNIP2	-1.43465	0.004256	potassium voltage-gated channel interacting protein 2
101.	204456_s_at	GAS1	-1.18439	0.004277	growth arrest specific 1

102.	203074_at	ANXA8	-1.04943	0.00439	annexin A8
103.	1558220_at	LINC00969	-1.07051	0.005296	Long Intergenic Non-Protein Coding RNA
104.	201295_s_at	WSB1	-1.31804	0.005564	WD repeat and SOCS box containing 1
105.	1553694_a_at	PIK3C2A	-1.15255	0.005607	phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 alpha
106.	215236_s_at	PICALM	-1.38014	0.005771	phosphatidylinositol binding clathrin assembly protein
107.	217202_s_at	GLUL	-1.09065	0.005905	glutamate-ammonia ligase
108.	230673_at	PKHD1L1	-1.14881	0.006065	polycystic kidney and hepatic disease 1 (autosomal recessive)-like 1
109.	226863_at	FAM110C	-1.0812	0.00692	family with sequence similarity 110 member C
110.	230733_at	NA	-1.00668	0.007106	Transcribed locus
111.	209156_s_at	COL6A2	-1.25261	0.007138	collagen type VI alpha 2 chain
112.	236571_at	SLC2A3	1.12577	0.007334	solute carrier family 2 member 3
113.	224590_at	XIST	-1.87404	0.007661	X inactive specific transcript (non-protein coding)
114.	227671_at	XIST	-2.18694	0.007948	X inactive specific transcript (non-protein coding)
115.	208621_s_at	EZR	-1.02767	0.008104	ezrin

116.	205700_at	HSD17B6	-1.01363	0.008479	hydroxysteroid 17-beta dehydrogenase 6
117.	202499_s_at	SLC2A3	1.31794	0.008501	solute carrier family 2 member 3
118.	1555730_a_at	CFL1	-1.09996	0.00899	cofilin 1
119.	201650_at	KRT19	-1.11275	0.009692	keratin 19
120.	209480_at	HLA-DQB1	1.311464	0.009978	major histocompatibility complex, class II, DQ beta 1
121.	214218_s_at	XIST	-1.90124	0.010493	X inactive specific transcript (non-protein coding)
122.	222549_at	CLDN1	-1.08141	0.010712	claudin 1
123.	213831_at	HLA-DQA1	1.376856	0.011036	major histocompatibility complex, class II, DQ alpha 1
124.	1555725_a_at	RGS5	-1.10898	0.01133	regulator of G-protein signaling 5
125.	223597_at	ITLN1	-1.37095	0.011634	intelectin 1
126.	1569106_s_at	SETD5	-1.07031	0.011637	SET domain containing 5
127.	235281_x_at	AHNAK	-1.05405	0.013411	AHNAK nucleoprotein
128.	221618_s_at	TAF9B	-1.09218	0.014823	TATA-box binding protein associated factor 9b
129.	221728_x_at	XIST	-1.76921	0.015011	X inactive specific transcript (non-protein coding)
130.	220037_s_at	LYVE1	-1.01358	0.016632	lymphatic vessel endothelial hyaluronan receptor 1

131.	224588_at	XIST	-2.18179	0.017504	X inactive specific transcript (non-protein coding)
132.	1553962_s_at	RHOB	-1.08055	0.018743	ras homolog family member B
133.	205808_at	ASPH	-1.06085	0.020474	aspartate beta-hydroxylase
134.	222830_at	GRHL1	1.019129	0.022761	grainyhead like transcription factor 1
135.	243509_at	NA	1.005059	0.023809	Transcribed locus
136.	205000_at	DDX3Y	1.666086	0.025541	DEAD-box helicase 3, Y-linked
137.	206359_at	SOCS3	1.012713	0.026839	suppressor of cytokine signaling 3
138.	227697_at	SOCS3	1.581777	0.027106	suppressor of cytokine signaling 3
139.	204409_s_at	EIF1AY	1.280034	0.027897	eukaryotic translation initiation factor 1A, Y-linked
140.	206700_s_at	KDM5D	1.108347	0.031911	lysine demethylase 5D
141.	1554574_a_at	CYB5R3	-1.03753	0.032195	cytochrome b5 reductase 3
142.	202768_at	FOSB	1.546339	0.033702	FosB proto-oncogene, AP-1 transcription factor subunit
143.	206211_at	SELE	1.517882	0.035704	selectin E
144.	204006_s_at	FCGR3A	-1.09402	0.037772	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)
145.	205001_s_at	DDX3Y	1.015187	0.039487	DEAD-box helicase 3, Y-linked

Among the 145 differentially expressed genes, 20 DE genes were found to be up-regulated and the remaining 125 genes were found to be down-regulated. A list of top 15 most significantly up and down-regulated genes are presented in Table 4 and 5.

Table 4

List of top 15 most significantly up-regulated DE genes

UP-REGULATED GENES					
S.no	**GeneID**	**Gene Symbol**	**p-value**	**Fold change**	**Description**
1	205000_at	DDX3Y	0.025540518	1.666085603	DEAD (Asp-Glu-Ala-Asp) box helicase 3, Y-linked
2	202499_s_at	SLC2A3	0.008501342	1.317940144	Solute carrier family 2 (facilitated glucose transporter), member 3
3	222830_at	GRHL1	0.022761324	1.019128764	Grainyhead-like transcription factor 1
4	206211_at	SELE	0.03570443	1.517881753	Selectin E
5	206359_at	SOCS3	0.026838617	1.012713204	Suppressor of cytokine signaling 3
6	206932_at	CH25H	0.000321919	1.753047974	Cholesterol 25-hydroxylase
7	209480_at	HLA-DQB1	0.009978465	1.311464052	Major histocompatibility complex, class II, DQ beta 1
8	224321_at	TMEFF2	0.000136017	1.257781164	Transmembrane protein with EGF-like and two follistatin-like domains 2
9	213350_at	RPS11	0.003544219	1.047089555	Ribosomal protein S11
10	214461_at	LBP	3.66E-07	1.201252629	Lipopolysaccharide binding protein
11	204409_s_at	EIF1AY	0.027896892	1.280033635	Eukaryotic translation initiation factor 1A, Y-linked
12	206700_s_at	KDM5D	0.031911492	1.108347414	Lysine (K)-specific demethylase 5D
13	202768_at	FOSB	0.033702334	1.546339468	FBJ murine osteosarcoma viral oncogene homolog B
14	230040_at	ADAMTS18	0.000628552	1.129118161	ADAM metallopeptidase with thrombospondin type 1 motif 18
15	213831_at	HLA-DQA1	0.011036303	1.376856451	Major histocompatibility complex, class II, DQ alpha 1

Table 5

List of top 15 most significantly down-regulated DE genes

DOWN-REGULATED GENES					
S.no	**GeneID**	**Gene Symbol**	**p-value**	**Fold change**	**Description**
1	204560_at	FKBP5	0.002686374	-1.279602126	FK506 binding protein 5
2	215990_s_at	BCL6	0.000164942	-1.055963606	B-cell CLL/lymphoma 6
3	203074_at	ANXA8	0.004389937	-1.049428951	Annexin A8
4	217202_s_at	GLUL	0.005905253	-1.090647399	Glutamate-ammonia ligase
5	214587_at	COL8A1	0.00011324	-1.063075115	Collagen type VIII alpha 1 chain
6	217023_x_at	TPSAB1	0.002854742	-1.18114921	Tryptase alpha/beta 1
7	210317_s_at	YWHAE	0.003708975	-1.177635876	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon
8	1565717_s_at	FUS	0.000220036	-1.129446408	FUS RNA binding protein
9	214218_s_at	XIST	0.010493486	-1.901237471	X inactive specific transcript (non-protein coding)
10	214693_x_at	NBPF10	3.34E-05	-1.157864095	Neuroblastoma breakpoint family, member 10
11	205960_at	PDK4	0.003334388	-1.603200115	Pyruvate dehydrogenase kinase 4
12	230165_at	SGO2	0.000729407	-1.126428319	Shugoshin 2
13	204004_at	PAWR	0.000445108	-1.020047701	Pro-apoptotic WT1 regulator
14	233314_at	PTEN	5.10E-05	-1.135778793	Phosphatase and tensin homolog
15	219529_at	CLIC3	3.61E-05	-1.210208836	Chloride intracellular channel 3

4.4 CLUSTER ANALYSIS OF THE DE GENES

In order to generate an overview of the differential gene expression profile, a hierarchical clustering was performed. The results of clustering of each gene was displayed using a heat map (Figure 7)

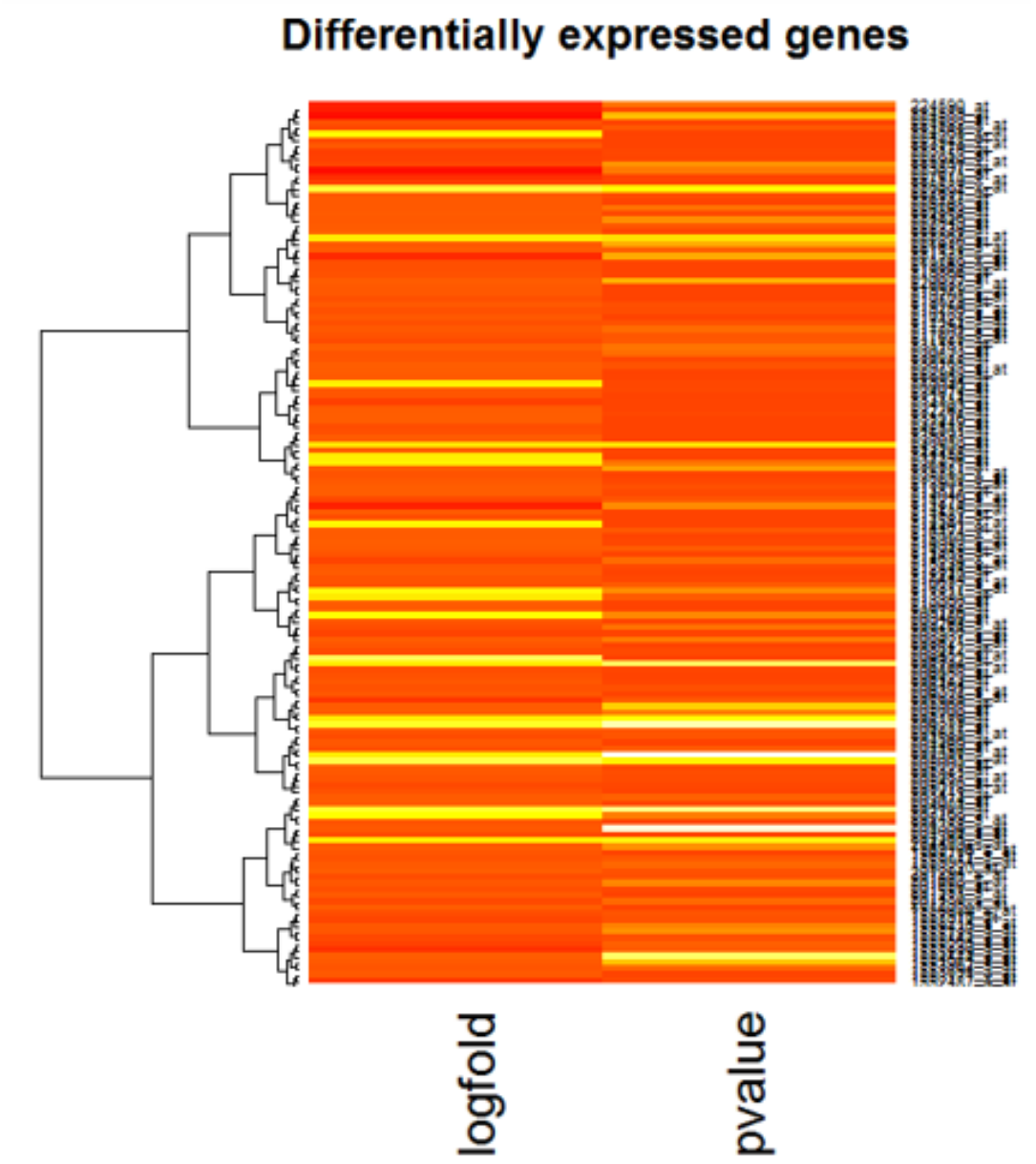


Figure 7. Heat map of 145 differentially expressed genes in obese compared with obese diabetic omental tissue samples. The log2 fold-change and p-value of each gene is used for plotting by hierarchical clustering. Gene IDs and the fold change, p-values are represented in rows and columns respectively.

4.5 FUNCTIONAL ENRICHMENT AND GENE ONTOLOGY

To gain insight about the biological significance of the differentially expressed genes, Gene Ontology (GO) enrichment analysis was performed using web based software GENECODIS. Gene ontology studies provide a brief descriptive framework on the functional annotation and biological classification of the gene sets in three different categories: biological process, cellular component and molecular functions.

4.5.1 Enrichment analysis of GO biological process

The initial process in functional enrichment analysis is the biological process of each differentially expressed genes. Figure 8 shows the differentially expressed genes with significantly enriched GO terms in biological process.

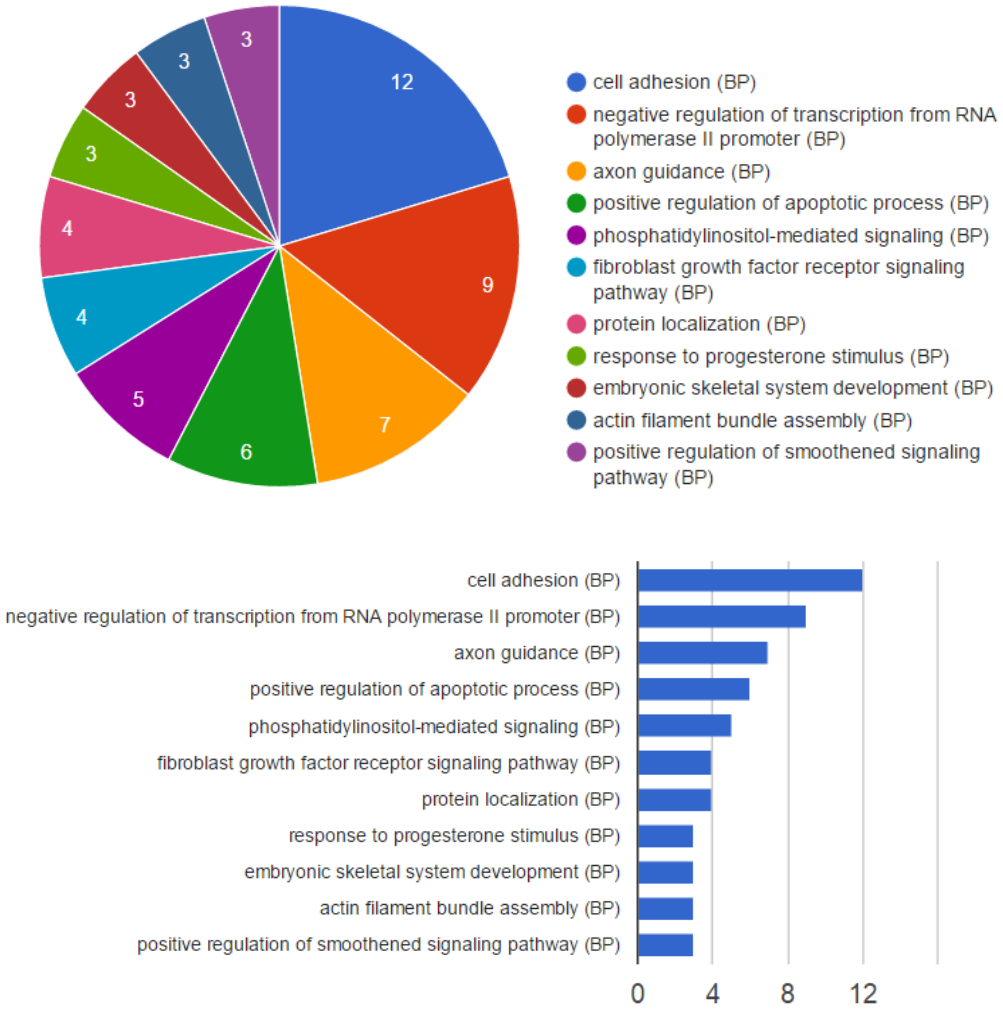


Figure 8.The hub genes depicted above represent the number of GO biological process categories. The proportion of hub genes falling into each category are indicated.

From the figure, the more significantly enriched biological processes were mainly associated with Cell adhesion (GO: 0007155, P=5.59749e-05) and negative regulation of transcription from RNA polymerase II promoter (GO: 0000122, P=0.000992776).

4.5.2 Enrichment analysis of GO cellular component

The cellular component ontology describes the locations of the differentially expressed genes at the levels of subcellular structures. This part of the annotation gave knowledge about the number of genes located at cellular level. Figure 9 depicts the number of differentially expressed genes enriched in terms of cellular location.

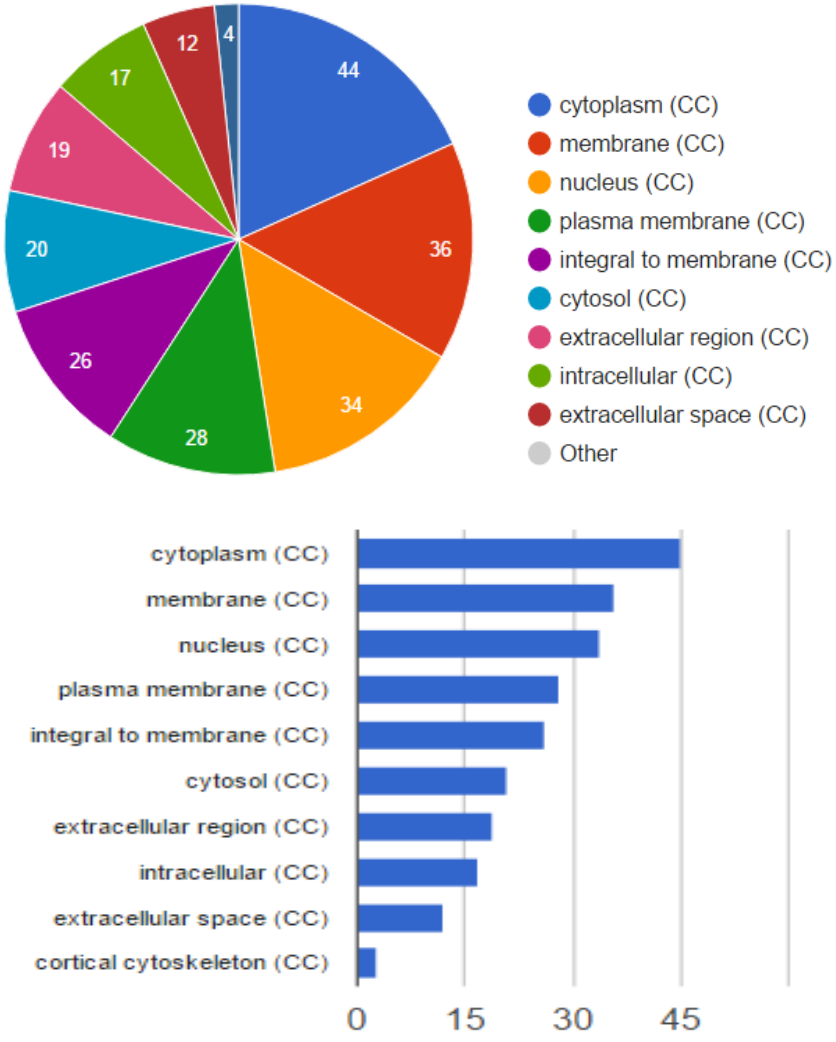


Figure 9. The hub genes depicted above represent the number of GO cellular component categories. The proportion of hub genes falling into each category are indicated

From the above figure, the significantly enriched GO terms for cellular component was higher in the cytoplasm (GO: 0005737, P=1.97097e-09) with the maximum number of genes-45 and in membranes (GO: 00160020, P=3.28193e-08) with 35 genes.

4.5.3 Enrichment analysis of GO molecular function

The molecular functional characterization of the significant differentially expressed genes were carried out to further elaborate the study in terms of functional categories (Figure 10).

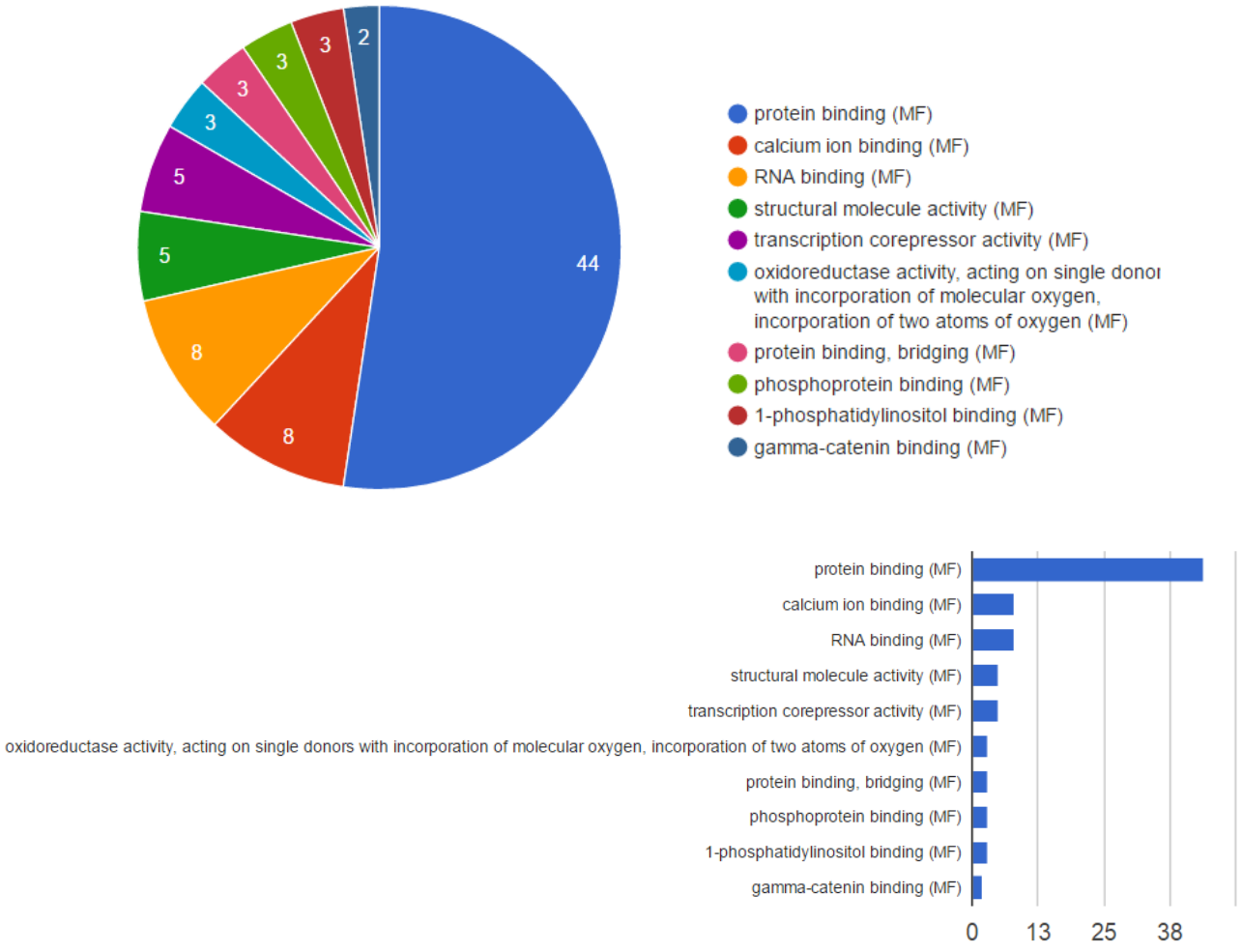


Figure 10. The hub genes depicted above represent the number of GO molecular function categories. The proportion of hub genes falling into each category are indicated

A list of 44 genes were significantly enriched with molecular function of protein binding (GO: 0005515, P=3.9281E-11) followed by calcium binding (8 genes) and RNA binding (8 genes).

4.6 PATHWAY ANALYSIS

To further evaluate the biological significance for the differentially expressed genes, KEGG pathway enrichment analysis was carried out using online software WebGestalt. Hypergeometric test with p-value <0.05 was used as the criteria for pathway detection. The DE genes were significantly eniched in 2 pathways with 10 positively regulated and 10 negatively regulated pathways (Table 6)

Table: 6

List of top 20 enriched KEGG pathway of differentially expressed genes

KEGG ID	KEGG name	No of genes	FDR value
Positively regulated			
hsa05164	Influenza A	5	2.49E-02
hsa04940	Type I diabetes mellitus	2	6.08E-02
hsa05152	Tuberculosis	4	6.27E-02
hsa05321	Inflammatory bowel disease (IBD)	2	6.41E-02
hsa05310	Asthma	2	6.57E-02
hsa04668	TNF signaling pathway	3	6.83E-02
hsa04672	Intestinal immune network for IgA production	2	6.88E-02
hsa05166	HTLV-I infection	3	7.06E-02
hsa05320	Autoimmune thyroid disease	2	7.11E-02
hsa05145	Toxoplasmosis	3	7.17E-02
Negatively regulated			
hsa04610	Complement and coagulation cascades	2	4.70E-01
hsa04390	Hippo signaling pathway	3	5.45E-01
hsa04621	NOD-like receptor signaling pathway	2	8.05E-01
hsa05133	Pertussis	2	9.49E-01
hsa01230	Biosynthesis of amino acids	1	9.50E-01

hsa05100	Bacterial invasion of epithelial cells	2	9.55E-01
hsa05200	Pathways in cancer	4	9.67E-01
hsa04915	Estrogen signaling pathway	3	9.70E-01
hsa04152	AMPK signaling pathway	3	9.73E-01
hsa01100	Metabolic pathways	5	9.74E-01

The pathways that are involved in immune response relating diabetes with obesity are highlighted (yellow) in the table 6.

4.7 DISCUSSION

The primary objective of the study was to identify differentially expressed genes among healthy obese and metabolically unhealthy obese (diabetic) with the gene expression data obtained from omental adipose tissues. Since the selection of data was solely based on obesity and adipose tissue (omental), there were many exclusions incorporated in retrieving the datasets for the study. But in spite of all limits, the study reveals novel features of obesity and its associated contribution to diabetes.

In this study, the microarray data obtained from databases pertaining to the conditions: obese and obese diabetic were integrated and successfully adopted for Meta-analysis. Firstly, the datasets were processed and normalized to gain knowledge about the quality of the data. The average intensities of each sample were viewed using a box plot before and after normalization (Figure 5 and 6). And the distribution of expression values in terms of conversion based on log2 transformation was also plotted (Figure 7) and visualized for each sample.

Primarily, the differential gene expression analysis was carried out by comparing the two conditions: obese and obese diabetic samples. The results obtained from the expression analysis suggested a list of 145 genes (Table 3) that were differentially expressed between the conditions with the regulation of fold change value >2.0 and p-value <0.05. Among these 145 genes, it was identified that 20 genes were found to be up-regulated and 125 genes were down-regulated.

The top 15 significantly up- and down- regulated genes showed relatedness with obesity induced diabetes in adipose tissue (Table 4).

Adipose tissue being an endocrine part secreting various hormones and cytokines like leptin, adiponectin, TNF-α, IL10 that leads to inflammatory mediation. Obesity associated with activation of these inflammatory mediators impairs the endocrine functions of adipose tissue leading to metabolic disorders like diabetes mellitus due to insulin resistance (Shoelson *et al.,* 2007 and Ota, 2013).

These altered production of proinflammatory molecules by adipose tissue have shown metabolic complications in obese state, because the adipose tissue of obese expresses increased amount of proteins like TNF-α, IL-6 and resistin which have direct effects on adipocyte insulin resistance and increased lipolysis (Weisberg *et al.,* 2003).

4.7.1 DE genes involved in the mechanism of inflammation and its pathways

The differentially expressed genes between the healthy obese and diabetic obese has shown a significant response in relating the secretion of the adipocytes contributing to inflammation leading to resistant insulin action or inactivation of important insulin signaling pathways. The significant relation of the up-regulated genes with inflammation and immune responses in adipose tissues that connects obesity mediated diabetes are highlighted in Table 4.

The up-regulated gene CH25H that codes for the protein cholesterol 5-hydroxylase is a major contributor of lipid metabolism and its regulation. It also acts as a corepressor, blocking sterol regulatory element binding protein. But studies have shown that CH25H has a broad activity against pathogenic virus and plays a beneficial role in promoting host immunity (Liu *et al.,* 2012). The activated macrophages that lacks CH25H were unable to produce cytokines during the type I IFN inflammatory conditions (Reboldi *et al.,* 2014).

Interestingly, another up-regulated gene LBP coding for the protein lipopolysaccharide binding protein plays an acute-phase immunological response binding to the bacterial lipopolysaccharide and LPS interacts with LBP which activates inflammatory changes through NF-kappa-B cells and MAPK signaling (Bochkov *et al.,* 2002). The MAPK (mitogen-activated protein kinases) pathway activates in response to insulin by the binding of insulin receptor substrate (IRS). Many studies have been implicated MAPKs in the development of insulin resistance. It is established that excessive activation of MAPKs is associated with detrimental effects in obesity and diabetes that contribute to disease progression (Gehart *et al.,* 2010).

And also the up-regulated gene FOSB is a member of family with four different isoforms-FOS, FOSB, FOSL1 and FOSL2 also has the immune response role of developing inflammation by the DAP12 receptors in the NK (Natural killer) cells. The protein selectin E coded by the gene SELE is a cell adhesion molecule found in cytokine-stimulated endothelial cells responsible for the accumulation of leukocytes at sites of inflammation.

Another significant gene expression was found to be the HLA-DQA1 gene which belongs to the Major histocompatibility complex, class II DQ alpha 1 that indeed triggers the immune system against the primary infection and also a few studies have investigated the susceptibility of the gene with genetic background of developing insulin dependent diabetes (Cucca *et al.,* 2000).

With consideration of the FDR value among the significant pathway, the most significant positively related is the Influenza-A pathway that is initiated via infection by the immune system. The infected pathway will lead to intracellular signaling cascades such as P13K pathway as well as MAPK signaling which plays an essential role in maintaining the insulin signaling. (Gaur *et al.*, 2011).

Significant insight was found in the pathway of TNF signaling which is a multi-functional cytokine that regulates cellular and molecular function in immune cells and also in energy metabolism. And also it has been stated that TNF-α is highly associated with obesity induced insulin resistance by its converging effects on inhibition of insulin signaling and blunting the insulin action in target tissues. (Cawthon and Sethi., 2007).

4.7.2 DE genes involved in the mechanism of glucose homeostasis

The up-regulated gene SLCA3 coding the protein solute carrier family 2 member 3 has significant pathway relation with the mechanism of glucose transport. It facilitates the transport of glucose that mediates the uptake of various monosaccharaides across the cell membrane.

SOCS3 also plays a vital role in cytokine signaling, in some cases aggravating a variety of diseases by interrupting the receptor and bringing inhibition of activation of the hormone signaling pathways. (Carow and Rottenberg, 2014). Also the gene SOCS3 has an exaggerated expression in adipose tissue of obese subjects. It has also shown a marked increase in the muscle of type 2 diabetic patients with reduced insulin resistance.SOCS3 is strongly related to inhibition of insulin signaling suggesting that elevated expression of SOCS3 in the muscle may contribute to insulin resistance (Rieusset *et al.*, 2004).

The complement and coagulation cascades are the most significant negatively-regulated pathway with the FDR score of 4.70E-01. The complement system is a main column of innate immunity and the coagulation system also a major column for massive activation of immune responses at early injury stages. In spite of the immune responses, these cascades also results in dysfunction of metabolic and endocrine organs due to the formation of certain inflammasome. The significant role of metabolic dysfunction is seen in insulin target organs such as the adipose tissue or even liver in the cases of obesity. On the other hand, the inflammasome at pancreas due to macrophages and IL-1 contributes directly to the islets dysfunction in the course of type 2 diabetes pathogenesis (Phieler *et al.,* 2013).

The pathway of NOD-like receptor has shown vital role in inflammasome. These NLRp3 (NOD-like receptors) causing inflammasome might actually inhibit glycolysis (Figure 10). The TLR (Toll like receptor) and NLPr3 produce the pro-inflammatory cytokine IL-1β. Hyperlipidemia will generate lipids such as palmitate and ceramides which activated NLRP3 that causes deposition of IAPP (islet amyloid polypeptide precursor) in the pancreas of type 2 diabetes.

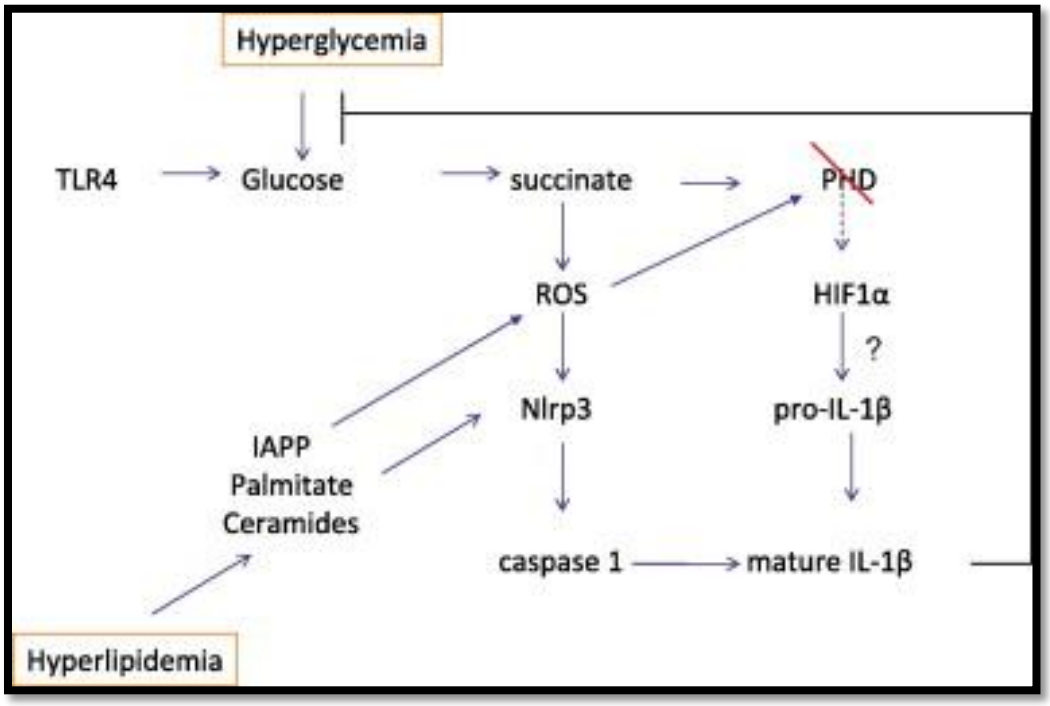


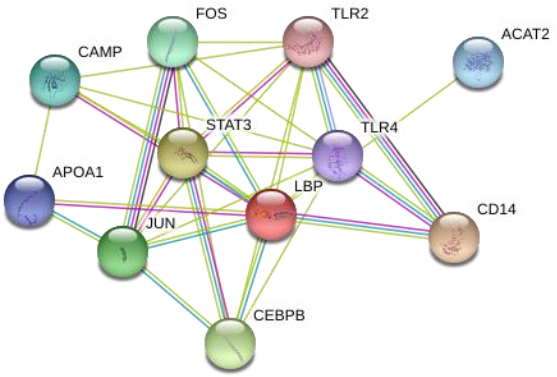
Figure 11. Role of TLR4, Nlrp3, hyperglycemia and hyperlipidemia in IL-1β production leading to negative feedback of glucose uptake and insulin

This NLRP3 activates caspase-1 leading to production of IL-1β, which has negative feedback to block insulin and glucose uptake. These events are critically identified in the pathogenesis of type 2 diabetes (Tannahill and O'Neill., 2011).

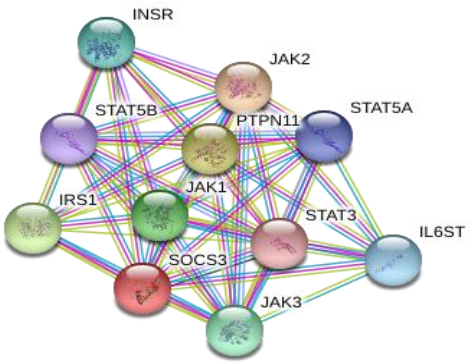
The major up-regulated gene has implications in causing inflammatory responses. Each gene associated with adipose inflammation links the suppression of secretion of insulin or increased resistant to insulin hormone via subsequent pathways. For instance, LBP gene has shown interactions with STAT3 (Figure 13a) which is a transcription factor that mediates cytokine signaling playing critical role in the pathogenesis of diabetes (Tc *et al.,* 2009). Signal transducer and activation of transcription 3(STAT3) involved in cytokine-induced insulin resistance and development of insulin resistance and type 2 diabetes is completely defined (Mashili *et al.,* 2013).

Interestingly, interaction of SOCS3 has also resulted in association of it with STAT3 (Figure 13b) inducing cytokine leading to leptin resistance (Hoene *et al.*, 2010).

The interaction study of FOSB gene (Figure 13c) has shown the incidence of its relation with the mitogen-activated kinase (MAPK) signaling pathway. In studies performed in identification of metabolic stress models has shown pronounced and transient up-regulation of the transcription factors encoded by FOSB gene.



a) STRING interaction network of LBP representing the association of STAT3 at one of it node



b) STRING interaction network of SOCS3 depicting the association of STAT3 at one of it node

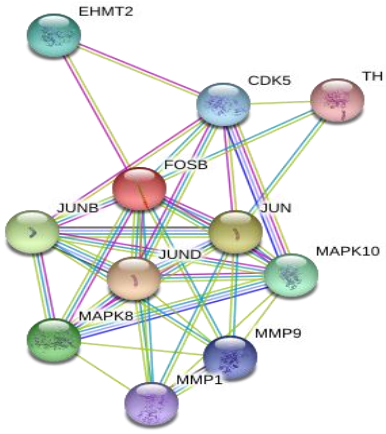


Figure 12: STRING interaction diagrams of most significantly genes involved in the inflammatory/immune responses and insulin resistant pathways.

a) STRING interaction of LBP, b) STRING interaction of SOCS3 and c) STRING interaction of FOSB

The inference of the identification of differentially expressed genes between obese and obese diabetic subjects of omental adipose tissue gave significant genes involving inflammation/immune response. Also the fact that a larger proportion of inflammation-related genes were upregulated (rather than down-regulated) in adipocytes of the obese individual samples further implies an active role for cells causing tissue inflammation. Some of these inflammation-related genes encodes chemokines, cytokines and other up-regulated genes encodes for cell adhesion molecules like fibronectin, integrin that helps to retain filtrating macrophages in the tissues leading to inflammation of adipocytes (Lan *et al.,* 2003).

From ontology studies of the DE genes represented in figure 9-11, it shown that the functional and molecular characterization of the DE genes has been widely distributed in different categories. Taking in account of the biological process, the enriched number of genes are found to be present in cell adhesion mechanisms which strongly interlink the up-regulated gene functions (LBP, SOCS3) respectively. Similarly, a drastic increase in the molecular function category was seen in protein binding with significantly enriched genes of 44 in number.

4.7.3 Comparison of DE genes between healthy obese vs unhealthy obese and their expression in other non-inflammatory pathways

The anti-inflammatory adipokines such as adiponectin, IL-4, IL-10, IL-13, IL-1 receptors antagonist and ILF-β are abundant in the adipose tissue of lean individuals. In contrary, the adipose tissue of obese is dominated by proinflammatory adipokines like leptin, TNF-α, IL-6, IL-18, retinol binding protein 4, angiopoietin-like protein 2, chemokine ligand 2, chemokine ligand 5 and nicotinamide phophoribosyl transferase.

However, there is an exception for the paradigm “more fat more metabolic disease”. Some of the obese individuals are healthy without any metabolic syndrome despite a high degree of obesity. These subjects were referred as “metabolically healthy obese (MHO)”. There is great inconsistency in the definition of MHO, Munoz-Garach et al defined that MHO, or “metabolically benign obese only refers to metabolic cardiovascular complications such as orthopedic problems, pulmonary complications or other physiological conditions.

The obese type 2 diabetes individuals have higher circulating level of IL-6, FFK and glycerol compared to MHO. It is also showed that MHO adipose tissue have low level of inflammatory marker as those of non-obese subjects.

The expansion capacity of adipose tissues were reported to be different between MHO and MUO. The various hypothesis namely “adipose tissue expandability hypothesis” and “fairly solid theory” were put forward to explain the adaptation of adipose tissue. The four factors playing major role in adipose tissue expansion and functionality are namely (1) lipogenesis, (2) adipogenesis via the newly formed progenitor cells, (3) apoptotic and antiapoptotic pathway and (4) angiogenesis.

The investigations towards further characterization of genetic contributions to the pathways of adipogenesis, apoptosis and angiogenesis that explains the difference between the adipose tissue in context of MHO individuals. The MHO were reported to expand the adipose tissue in healthier way by increasing the number and size of adipocytes whereas MUO interact with the cellularity of adipose tissue (Munoz-Garach *et al.,* 2016). There are also reports claiming that metabolic healthy obese have higher lipogenic and angiogenic capacity compare to MUO.

The gene ontology studies on the DE genes of MHO and MUO has shown an interesting insight of higher significance in positive regulation of apoptotic process. In type 2 diabetes mellitus, insulin resistance with obesity leads to glucose toxicity effect accelerating β-cell death by apoptosis. Specific proinflammatory cytokines cause β-cell death by the induction of mitochondrial stress and other responses because cytokines of immune cells that have infiltrated the pancreas are reported to be crucial mediators of β-cell destruction via the apoptotic pathway (Lin *et al.*, 2012).

Chronically exposed hyperglycemia leads to oxidative stress and inflammation that changes the regulation of gene expression which converges impaired insulin secretion and increased apoptosis (Gilbert and Liu, 2012). Oxidative stress leads to damage in organelles like mitochondria and endoplasmic reticulum as well as cellular proteins, lipids and nucleic acids. Among which the endoplasmic stress has shown to be strongly associated with the β-cell apoptosis in type 2 diabetes (Marchetti *et al.*, 2007).

The β-cell mass fluctuates according to the body need of insulin, β-cell deficiency correlates with glucose intolerance, β-cell death may directly lead to insulin deficiency when the loss is above 60% or more, it is accompanied by the presence of insulin resistance with obesity.

The results of this study has interestingly shown the differentially expressed genes that have strong associations linking obesity with insulin-resistant diabetes mediated by inflammatory responses and its related pathways. The study has also identified a number of novel genes, GO categories and its biological pathways that may be associated the pathobiology of obese and obese diabetic conditions. Of note, inflammasome formation in adipose tissue of obese individuals have been clearly established. And also its contributory roles to the etiology of diabetes has also been inferred significantly.

SUMMARY AND CONCLUSION

5. SUMMARY AND CONCLUSION

Obesity is a metabolic disorder of increasing prevalence worldwide and risk factor for the development of insulin resistance, metabolic syndrome type 2 diabetes. It is a state of low-grade systemic inflammation which is deeply involved in insulin resistance. The recent researches suggests that obese individuals are up to 80 times more likely to develop type 2 diabetes than those of normal people.

Adipose tissue is a regulator of various physiological process and its dysfunction in obese individuals greatly affect the metabolic homeostasis. Adiposity is the fraction of total body mass comprised of neutral lipid stored in adipose tissue, which is closely correlated with important physiological parameters such as increased cytokines production, synthesis of proinflammatory mediators and leptin concentrations. These changes in adipose tissue mass will strongly cause alterations in the endocrine and metabolic functions that links obesity with insulin sensitivity.

The deposition of fat under the skin causes the releases of proinflammatory chemical signals that makes the body sensitive to insulin or disrupts the insulin responsive cells and their ability to respond to insulin. Thus it has become a necessary measure to understand the differential gene expression pattern of adipose tissue of metabolically healthy obese subjects versus the metabolically unhealthy obese to get a brief insight on obesity and its association with diabetes.

In this study, the microarray datasets pertaining to metabolically healthy obese and diabetic obese subjects from omental adipose tissue origin was extracted from gene expression databases namely GEO and ArrayExpress. The search was scrutinized and focused based on several inclusion and exclusion criteria. The expression data of obese subjects with adipose tissue origin was considered whereas samples with clinical conditions other than diabetes, other tissue origin as well as children samples were excluded. Upon completion of extraction data with the selected criteria, the datasets were grouped into two as metabolically healthy obese (MHO) and metabolically unhealthy obese (MUO) in order to carry out the identification of differential pattern of gene expression in between both the conditions.

The significant up-regulated and down-regulated genes among the differentially expressed genes between the conditions were identified. The DE genes were further analyzed for their biological process, molecular functions and cellular components. To gain knowledge about

the mechanism or role of obesity causing insulin resistance, the genes hits were annotated for identifying the pathway using KEGG data source. Interesting inferences were made from the pathway analysis of DE genes which showed high degree of relation between insulin resistance mediated through obesity.

The differentially expressed genes that are involved in inflammatory insulin resistant pathways are CH25H, LBP, FOSB, SOCS3, SELE, HLA-DQA1 and HLA-DQA. In addition, the interaction of three genes namely LBP, FOSB and SOCS3 has showed strong interactions with significant pathways such as MAPK signaling pathway, STAT3 signaling which plays vital role in insulin signaling cascades. Along with the significant gene SOCS3 has efficiently shown improved role in cytokine signaling, which at elevated levels will strongly inhibit insulin signaling.

The altered production of proteins like TNF-α, IL-6 have major defects on adipocytes causing insulin resistance and increased lipolysis. This meta-analysis also confirmed the activation of inflammatory pathways like TNF signaling and NOD like receptor signaling which has molecular relation in the production of inflammatory mediators causing inflammation mediated insulin resistance in obesity. A positive regulation of apoptotic pathway was also observed. Several studies reported that adipocyte death in obese individuals has role in metabolic disorders including insulin resistance. The mechanism of caspase activation and adipocyte apoptosis has showed marked increase in adipose tissue of obese individuals. The genetic inactivation of Bid (BH3-nteracting domain) death agonist which is a ley pro-apoptotic molecule has been proved to serve as a link between the two cell death pathways, reduced caspase activation and adipocyte apoptosis that leads to the development of systemic insulin resistance. These evidences strongly suggest that adipocyte apoptosis is a key initial event that contributes to macrophage infiltration into adipose tissue causing insulin resistance.

The meta-analysis also further confirms the role of inflammatory mediators like TNF-α and IL-6 interfering with the insulin causing insulin resistance in obese.

In conclusion, the study examined the differential gene expression in adipose tissue of MHO and MUO subjects that showed great progress in terms of understanding of obesity-induced inflammation a causative factor for developing insulin resistance. It is also intriguing that an increase in inflammatory mediators predicts the future development of obesity and diabetes.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Ashcroft,F.M. and Rorsman,P. (2012) Diabetes mellitus and the β cell: the last ten years, *CellPress,*148:1160-1171.
- Barrett, T., Suzek, T.O., Troup, D.B., Wilhite, S.E., Ngau, W., Ledoux, P., Rudnev, D., Lash, A.E., Fujibuchi, W and Edgar, R. (2005) NCBI GEO: mining millions of expression profiles-database and tools, *Nucleic Acids Res,*33:D562-D566.
- Bell, C.G., Finer, S., Lindgren,C.M., Wilson,G.A., Rakyan,V.K., Teschendorff, A.E., Stupka, P.A.T., Down,T.A., Prokopenko, I., Morison,I.A., Mill, J. and Pidsley,R. (2010) Integrated Genetic and Epigenetic Analysis Identifies Haplotype-Specific Methylation in the FTO Type 2 Diabetes and Obesity Susceptibility Locus, *Plos One*,5(11).
- Belle, T.M., Coppieters, K.T. and Herrath, M.G.V. (2011) Type 1 Diabetes: Etiology, Immunology and Therapeutic Strategies, *Physiol Rev*,9:79–118.
- Bertola, A., Ciucci, T., Rousseau, D., Bourlier, V., Duffaut, C., Bonnafous, S., Wakkach,C.B., Anty, R., Iannelli,A. and Wakkach, A. (2012) Identification of Adipose Tissue Dendritic Cells Correlated With Obesity-Associated Insulin-Resistance and Inducing Th17 Responses in Mice and Patients diabetes, *Diabetesjournals,*61:2238–2247.
- Bochkov, V.N., Kadl, A., Huber, J., Gruber, F., Binder, B.R. and Leitinger, N. (2002) Protective role of phospholipid oxidation products in endotoxin-induced tissue damage, *Nature*,419:77-81.
- Boden, G. (2011) Obesity, Insulin Resistance and Free Fatty Acids, *CurrOpinEndocrinol Diabetes Obes,*18(2):139–143.
- Carow, B. and Rottenberg, M.E. (2014) SOCS3, a major regulator of infection and inflammation, *Front Immunol*,5:58.
- Cawthorn, W.P. and Sethi, J.K. (2007) TNF-α and adipocyte biology, *FBBS Letters*, 582(1):117-131.
- Cucca, F., Lampis, R., Congia, M., Angius, E., Nutland, S., Bain, C.S., Barnett, H.A., and Todd, A.J. (2001) A correlation between the relative predisposition of MHC class II alleles to type 1 diabetes and the structure of their proteins, *Hum Mol Genet*,10(19):2025-2037.
- Dall, T.M., Zhang, Y., Chen, Y.J., Quick, W.W., Yang, W.G. and Fogli, J. (2010) The Economic Burden Of Diabetes, *Health Affairs,*29,2:297-303.

- Dillies, M., Rau, A., Aubert,J., Antier,C.H., Jeanmougin,M., Servant,N., Keime, C., Marot,G., Castel,D., Estelle,J., Guerne,G., Jagla, B., Jouneau, L., Laloe,D., Gal,L., Crom,S.F., Guedj, M. and Jaffrezic, F. (2013) A comprehensive evaluation of normalization methods for illumine high-throughput RNA sequencing data analysis, *Briefings in Bioinformatics*,14: 6. 671- 683.
- Dyck, R., Osgood, N., Lin, T.H., Gao, A. and Stang,R. (2010) Epidemiology of diabetes mellitus among First Nations and non-First Nations adults, *CMAJ*,3,183.
- Eckel, R.H., Kahn, S.E., Ferrannini, E., Goldfine, A.B., Nathan, D.M., Schwartz, M.W., Smith, R.J. and Smith, S.R. (2011) Obesity and Type 2 Diabetes: What Can Be Unified and What Needs to Be Individualized?, *The Journal of Clinical Endocrinology and Metabolism*,96(6):1654-1663.
- Guar, P., Mujhal, A. and Lal, S.K. (2011) Influenza virus and cell signaling pathways, *Medical Science Monitor*,17(6):RA148-154.
- Gehart, H., Kumpf, S.,Ittner, A. and Ricci, R. (2010) MAPK signaling in cellular metabolism: stress or wellness?,*EMBO reports*,11: 834-840.
- Gilbert, R.E., and Liu, D. (2012), The missing link to understanding β-cell dysfunction in the pathogenesis of type 2 diabetes, *Epigenetics*,7:8, 841-850.
- Ginter, E. and Simko, V. (2012) Type 2 diabetes mellitus,pandemic in 21st century, *Advances in Experimental Medicine and Biology*,771:42-50.
- Goblan, A., Alfi,M.A. and Khan,Z.M. (2014) Mechanism Linking Diabetes Mellitus and Obesity, *Diabetes,Metabolic Syndrome and Obesity:Targets and Therapy*,7.
- Gopalakrishnan, S. and Ganeshkumar, P. (2013) Systemic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare, *Journal of Family Medicine and Primary Care*,2(1): 9-14.
- Grant, R.W. and Dixit.V.D. (2013) Mechanisms of disease: inflammasome Activation and the Development of type 2 diabetes, *Frontiers in Immunology*,4.
- Griffith, J., Barber, V.S., Morgan, L. and Young, D. (2006) Systematic review and meta-analysis of studies of the timing of tracheostomy in adult patients undergoing artificial ventilation, *BMJ*,330(7502):1243.

- Guariguata, L., Whiting, D.R., Hambleton, I., Beagley, J., Linnenkamp, U. and Shaw, J.E. (2014) Global estimates of diabetes prevalence for 2013 and projections for 2035, *Diabeteic Research and Clinical Practice*,103(2):137-49.
- Gulati, S. and Misra, A. (2014) Sugar Intake,Obesity, and Diabetes in India, *Nutrients*, 6(12):5955-5974.
- Haidich, A.B., (2010) Meta-analysis in medical research, *Hippokratia*,14:29-37.
- Hostamisligil, G.S., Arner, P., Caro, J.F., Atkinson, R.L. and Spiegeiman, B. (2011) Increased Adipose Tissue Expression of Tumor Necrosis Factore-α In Human Obesity and Insulin Resistance, *J. Clin. Invest*,95,2409-2415.
- Ioannidis, J.P.A. and Lau, J. (2002) 18F-FDG PET for the Diagnosis and Grading of Soft-Tissue Sarcoma: A Meta-Analysis, *Journal of Nuclear Medicine*,44:717-724.
- Keller, M.P. and Attie, A.D. (2010) Physiological insights gained from gene expression analysis in obesity and diabetes, *Annual Review of Nutrition*,30:341-64.
- Kerner, W. and Brukel, J. (2014) Definition,Classification and Diagnosis of Diabetes Mellitus, *Experimental and Clinical Endocrinology and Diabetes*,122(7):384-386.
- Kodamaa,K., Horikoshic, M., Todad, K., Yamadae,S., Harac,K., Iriee,J., Sirotaa, M., Morgana, A., Chena, M., Chena,R., Ohtsug,H., Maedah, S., Kadowakic,T.and Buttea,A.T. (2012) Expression-based genome-wide association study links the receptor CD44 in adipose tissue with type 2 diabetes,*PNAS*,109: 7049-7054.
- Krug, E.G. (2016) Trends in diabetes: sounding the alarm, *The Lancet*,387:1485-1486.
- Lan, H., Rabaglia, M.E., Stoehr, J.P., Nadler, S.T., Schueler, K.L, Zou, F., Yandell, B.S. and Attie, A.D. (2003) Gene Expression Profiles of Nondiabetic and Diabetic Obese Mice Suggest a Role of Hepatic Lipogenic Capacity in Diabetes Susceptibility, *Diabetes*,52(3):688-700.
- Lin, C.Y., Ni, C.C., Yin, M.C. and Lii, C.K. (2012) Flavanoids protect pancreatic beta-cells from cytokines mediated apoptosis through the activation of P13-kinase pathway, *Cytokine*,59:65-71.
- Liu, S.Y., Aliyari, R., Chikere, K., Li, G., Mardden, M.D., Smith, J.K., Oernet, O., Guo, H., Nusbaum, R., Zack, J.A., Freiberg, A.N., Su, L., Lee, B. and Cheng, G. (2012) Interferon-Inducible Cholersterol-25-Hydroxylase Broadly Inhibits Viral Entry by Production of 25-Hydroxycholesterol, *Immunity*,38:92-105.

- Liu, X., Li, Y., Li, L., Zhang, L., Ren, Y., Zhou, H., Cui,L., Mao, Z., Hu, D. and Wang, C. (2016) Prevalence, awareness, treatment, control of type 2 diabetes mellitus and risk factors in Chinese rural population: the RuralDiab study, *Scientific Reports*,6:31426.
- Marchetti, P., Bugliani, M., Lupi, R., Marselli, L., Masini, M., Boggi, U., Filipponi, F., Weir, G.c., Eiziriki, D.L. and Cnop, M. (2007) The endoplasmic reticulum in pancreatic beta cells of type 2 diabetes patients, *Diabetologia*,50:2486-2489.
- Mashili, F., Chibalin, A.V., Krook, A. and Zierath, J.R. (2013) Constitutive STAT3 Phosphorylation Contributes to Skeletal Muscle Insulin Resistance in Type 2 Diabetes, *Diabetes*,62(2):457-465.
- Metzger, B.E., Gabbe, S.G., Persson, B., Lowe, L.P., Dyer, A.R., Oats, J.J.N. and Buchanan, T.A. (2010) International Association of Diabetes and Pregnancy Study Groups Recommendations on the Diagnosis and Classification of Hyperglycemia in Pregnancy, *Diabetes Care*, 33(7):e98.
- Munoz-Garach, A., Cornejo-Pareja, I. and Tinahones, F.J. (2016) Does Metabolically Healthy Obese exist?, *Nutrients*, (6).
- Olokoba, A.B., Obateru, O.A. and Olokoba, L.B. (2012) Type 2 Diabetes Mellitus: A Review of Current Trends, *Oman Medical Journal*,27:269-273.
- Ota, T. (2013) Chemokine Systems Link Obesity to Insulin Resistance, *Diabetes & metabolism journal*,37(3):165-172.
- Patel, C.J., Bhattachary, J. and Butte, A. (2010) An Environment-Wide Association Study (EWAS) on Type 2 Diabetes Mellitus*),PLoS ONE* ,5(5): e10746.
- Phieler, J., Martin, R.G., Lambris, J.D. and Chavakis, T. (2013) The role of the complement system in metabolic organs and metabolic diseases, *Seminars in Immnunology*,25(1):47-53.
- Ramasamy, A., Mondry, A., Holmes, C.C. and Altman, D.G. (2008) Key issues in Conducting a Meta-Analysis of Gene Expression Microarray datasets, *PLoS Medicine*, 5:1320-1332.
- Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo,P., Mason, C., Socci,N.D. and Betel, D. (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq dat, *Genome Biology*,14:95.

- Reboldi, A., Dang, E.V., McDonald, J.G., Liang, G., Russell, D.W. and Cyster, J.G. (2014) Inflammation.25-Hydroxycholesterol suppresses interleukin-1-driven inflammation downstream of type I interferon, *Science*,345(6197):679-684.
- Rieusset, J., Bouzakri, K., Chevillotte, E., Ricard, N., Jacquet, D., Bastard, J.P., Laville, M. and Vidal, H. (2004) Suppressor of Cytokine Signaling 3 Expression and Insulin Resistance in Skeletal Muscle of Obese and Type 2 Diabetic Patients, *Diabetes*,53(9):2232-2241.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015) *limma* powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Research*,43(7):e47.
- Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics*,26(1):139-140.
- Romeo, R.G., Lee, J. and Shoelson, S.E. (2012) *Aterioscler Thromb Vasc Biol*,32:1771-1776.
- Ruskin, H.J. (2016) Computational Modeling and Analysis of Microarray Data: New Horizons, *Microarrays* ,5,26.
- Rustici, G., Kolesnikov, N., Brandizi, M., Burdett, T., Dylag, M., Emam, I., Farne, A., hastings, E., Ison, J., Keays, M., Kurbatova, N., Malone, J., Sharma, A., Tang, Y.A., Ternet, T., Tikhonov, A., Welter, D., Williams, E., Brazma, A., Parkinsons, H. and Sarkans, U. (2013) ArrayExpress update—trends in database growth and links to data analysis tools, *Nuclei Acids Research*,41:D987-990.
- Seino, Y., Nanjo, K., Tajima, K., Kadowaki, K., Kashiwagi, A., Araki, E., Ito, C., Inagaki, N., Iwamoto, Y., Kasuga, M., Hanafusa, T., Haneda, M. and Ueki, K. (2010) Report of the Committee on the classification and diagnostic criteria of diabetes mellitus, *Diabetol Int*,1:2-20.
- Selvaraj, S. and Natarajan, J. (2011) Microarray Data Analysis and Mining Tools, *Bioinformation*,6(3):95-99.
- Shoelson, S.E., Lee, J. and Goldfine, A.B. (2006) Inflammation and insulin resistance, *The journal of Clinical Investigation*,116:1793-1806.
- Tannhill, G.M. and O'Neill, L.A.J. (2011) The emerging role of metabolic regulation in the functioning of Toll-like receptors and the NOD-like receptor Nlrp3,*FEBS Letters*, 585:1568-1572.

- Tc, L., Wang, Z.H., Feng, X., Chuang, P.Y., Fang, W., Shen, Y., Levy, D.E., Xiong, H., Chen, N. and He, J.C. (2009) Knockdown of Stat3 activity in vivo prevents diabetic glomerulopathy, *Kidney International*,76(1):63-71.
- Vasan, N., Yelensky, R., Wang, K., Moulder, S., Dzimitrowicz, H., Avritscher, R., Wang, B., Wu, Y., Cronin, M.T., Palmer, G., Symmans, W.F., Miller, V.A, Stephens, P. and Pusztai, L. (2014) A Targeted Next-Generation Sequencing Assay Detects a High Frequency of Therapeutically Targetable Alterations in Primary and Metastatic Breast Cancers: Implications for Clinical Practice, *The Oncologist Express*,19(5):453-458.
- Weisberg, S.P., McCann. D., Desai, M., Leibel, R.L. and Ferrante, A.W. (2003) Obesity is associated with macrophage accumulation in adipose tissue, *The Journal of Clinical Investigation*,112(12):1796-1808.
- Wilcox, G. (2010), Insulin and Insulin Resistanc, *Clin Biochem Rev*, 26.
- Woroniecka, K.I., Park, A.S.D., Mohtat, D., Thomas, D.B., Pullman and Susztak, K. (2011) Transcriptome Analysis of Human Diabetic Kidney Disease, diabetes.*diabetesjournals.org*,60(9):2354-69..
- Xu, H., Barnes, G.T., Yang, Q., Tan, G., Yang, D., Chou, C.J., Sole, J., Nichols, A., Ross, J.S., Tartaglia, L.A. and Chen, H. (2003) Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance, *The Journal of Clinical Investigation*,112(12):1821-1830.
- Yang,W., Lu, J., Weng,J., Jia,W., Ji,L., Xiao,J., Shan,Z., Liu,J., Tian,,H., Ji,Q., Zhu,D., Ge,J., Lin, L., Chen,L., Guo,X., Zhao,Z., Li, Q., Zhou,Z., Shan,G.and He, J. (2017) Prevalence of Diabetes among Men and Women in China, *N Engl J Med*,362:1090-101.
- Zhang, Z.H., Jhaveri, D.J., Marshall, V.M., Bauer, D.C., Edson, J., Narayanan, R.K., Robinson, G.J., Lundberg, A.E., Barrlett, P.F., Wray, N.R. and Zhao, Q.Y.(2014) A Comparative Study of Techniques for Differential Expression Analysis on RNA-Seq Data, *PLOS/ONE*,9(8).

WEBOGRAPHY

- www.ncbi.nlm.nih.gov/geo/
- www.ebi.ac.uk/arrayexpress/
- www.genecodis.cnb.csic.es/
- www.webgestalt.org