Transcriptomic insights into the role of *F13A1* in human weight gain

Mansi Arora

Division of Experimental Medicine, Department of Medicine

Faculty of Medicine

McGill University, Montreal, Canada

April 2021

A thesis submitted to McGill University as partial fulfilment of requirements of the degree of

Master of Science -Thesis in Experimental Medicine

CONTENTS **Page Number**

Abstract

The prevalence of obesity − where excess body fat accumulates − has doubled worldwide in the past 30 years. Currently 25% of Canadians are obese and at a risk of developing several obesity-linked comorbidities including developing type 2 diabetes, cardiovascular diseases, hypertension, respiratory diseases, osteoarthritis, and several types of cancers which all lead to reduced life expectancy. It is now considered that the adipose tissue expansion in overnutrition results in a low-grade systemic inflammation which initiates many of the pathologies leading to the disturbed metabolism and comorbidities.

Many studies have searched 'obesity genes' specifically in white adipose tissue (WAT). Twin studies have proven to be outstanding approach for finding causative and reactive genes by controlling the genetic background. A recent study used on monozygotic (MZ) twins that are discordant in weight, i.e. lean-obese twin pairs to study genes that are linked to obesity. WAT gene expression comparison between the lean and heavy co-twins identified significant changes in 28 novel genes that significantly associated to weight. The gene with highest association to obesity in this study was *F13A1*/FXIII-A transglutaminase which has been found linked to adipogenesis in cells and to obesity in humans and mice in our laboratory. It's role in human acquired excess weight or subsequent metabolic consequences is not known.

In this thesis, I have used the GWA/Affymetrix mRNA data from adipose tissue and adipocyte fraction as well as metabolic and clinical data of the weight discordant MZ twins where *F13A1* was identified associated to obesity. The aim of this thesis was to examine the correlation of *F13A1* mRNA to human adiposity-related parameters (weight, fat%, circulating adiponectin, leptin, and adipocyte size and number) as well as to explore the associated genes and potential pathways in a full transcriptome association study (TWAS). I report that differential expression of *F13A1* in adipocyte fraction (Heavy-Lean Δ*F13A1)* shows significant positive correlation with ΔWeight (*P*=0.034) and ΔBodyFat (0.044) and ΔAdipocyte volume (*P*=0.012). Adipose tissue Δ*F13A1 associates* negatively to circulating adiponectin at tissue level (ΔAdiponectin) (*P*=0.0195). Subsequent TWAS on Δ*F13A1* vs weight correlated Δtranscriptome identified 182 genes which showed r>0.7, r^2>0.5, *P*=0.05 to Δ*F13A1* and significantly altered expression in heavier co-twin (*P*<.05). Gene Ontology analysis for all the genes revealed over-represented terms for biological functions corresponding to cell stress, inflammatory response, activation of

cells/leukocytes, extracellular matrix organization and supramolecular fibre organization, hemostasis and angiogenesis supporting a role for *F13A1* in obese adipose tissue inflammation and tissue remodelling. In a separate analysis on liver-fat discordant sets of MZ twins, *F13A1* did not show association with liver fat accumulation.

Understanding the molecular underpinnings. i.e., physiological and genetic factors contributing to adipose tissue expansion and inflammation can provide valuable information of etiology of obesity linked comorbidities and allow the identification of factors and biological events that can be excellent targets for designing new therapeutics to metabolic disturbances.

Résumé

La prévalence de l'obésité – où un excès de graisse corporelle s'accumule – a doublé dans le monde au cours des 30 dernières années. Actuellement, 25% des Canadiens sont obèses et à risque de développer plusieurs comorbidités liées à l'obésités, notamment le diabète de type 2, les maladies cardiovasculaires, l'hypertension, les maladies respiratoires, l'arthrose, et plusieurs types des cancers qui entraînent tous une réduction de l'espérance de vie. On considère maintenant que l'expansion du tissu adipeux avec la suralimentation aboutit à une légère inflammation systémique qui initie beaucoup de pathologies menant à un métabolisme perturbé et à des comorbidités.

De nombreuses études ont cherché des « gènes d'obésité » dans le tissu adipeux blanc en particulier. Les études sur les jumeaux se sont avérées être une approche remarquable pour trouver des gènes responsables et réactifs, tout en contrôlant le contexte génétique. Récemment, une étude a utilisé des jumeaux monozygotes différant en poids, c.-à-d. des paires de jumeaux minces-obèses, afin d'étudier les gènes liés à l'obésité. La comparaison de l'expression des gènes dans le tissu adipeux blanc entre les co-jumeaux minces et obèses a permis d'identifier des différences significatives entre 28 nouveaux gènes qui étaient significativement associés au poids. Le gène le plus fortement associé à l'obésité dans cette étude était *F13A1*/FXIII-A transglutaminase, qui a été lié à l'adipogenèse dans les cellules, et à l'obésité dans les humains et les souris dans notre laboratoire. Son rôle dans la prise de poids excessive chez les humains, ou ses conséquences métaboliques ultérieures n'est pas connu.

Dans cette thèse, j'ai utilisé les données GWA/Affymetrix ARNs provenant de tissu adipeux et de la fraction d'adipocytes, ainsi que des données métaboliques et cliniques sur des jumeaux monozygotes différant en poids chez qui *F13A1* avait été associé avec l'obésité. L'objectif de cette thèse était d'examiner la corrélation entre l'ARNm *F13A1* et l'adiposité humaine (poids, % de gras, adiponectine en circulation, leptine, taille et nombre d'adipocytes) ainsi qu'explorer ses gènes associés et mécanismes possibles dans une étude du transcriptome (TWAS). Je rapporte que l'expression différentielle de *F13A1* dans la fraction d'adipocytes (Obèse-Mince Δ*F13A1*) révèle une corrélation positive significative avec ΔPoids (p=0.034), ΔGraisseCorporelle (p=0.044) et ΔVolume d'Adipocyte (p=0.012). Le Δ*F13A1* dans le tissu adipeux est

négativement associé à l'adiponectine en circulation au niveau des tissus (ΔAdiponectine) (p=0.0195). Des TWAS suivantes sur Δ*F13A1* vs Δtranscriptome corrélé au poids ont identifié 182 gènes qui ont démontré r>0.7, r2>0.5, p=0.05 avec Δ*F13A1*, et une expression significativement altérée chez le/la co-jumeau/elle obèse (p<0.05). Une analyse de l'ontologie des gènes pour tous les gènes a révélé une surreprésentation des termes portant sur les fonctions biologiques qui correspondent au stress cellulaire, la réponse inflammatoire, l'activation des cellules/leucocytes, l'organisation de la matrice extracellulaire et des fibres supramoléculaires, l'hémostase et l'angiogenèse soutenant un rôle pour *F13A1* dans l'inflammation du tissu adipeux et le remodelage des tissus. Dans une analyse séparée étudiant des jumeaux monozygotes différant en graisse hépatique, *F13A1* n'a pas démontré d'association avec l'accumulation de graisse dans le foie.

Avoir une compréhension des mécanismes moléculaires, c.-à-d. les facteurs physiologiques et génétiques qui contribuent à l'expansion et l'inflammation du tissu adipeux peut nous apporter des informations précieuses sur l'étiologie des comorbidités liées à l'obésité, et peut nous permettre d'identifier des facteurs et des évènements biologiques qui peuvent être d'excellentes cibles pour la conception de nouvelles thérapies visant les perturbations métaboliques.

Acknowledgements

Scientific publications are never a one man show. I very much appreciate the chance to work with so many bright and supportive people. First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Mari Kaartinen, for her continuous help, motivation, encouragement and supportive attitude throughout my research. Her guidance and understanding helped me elevate and promote my work to a level which would not have been otherwise possible. It is because of her faith in me that kept me moving. I would also like to thank her for providing me opportunities being a part of such a great CIHR project, to present my work at multiple National and International conferences. It helped me develop my all-round profile. I am deeply grateful for her patience and willingness to devote her time and expertise throughout this dissertation, and my graduate education.

I would like to extend my heartiest gratitude towards Faculty of Medicine, McGill University for awarding me with Nyman Award and Internal Studentship 2019-20, and Division of Experimental Medicine for Entrance and Recruitment Awards 2018-19, and for supporting me during my graduate studies.

I would also like to extend my appreciation to Dr. Kirsi Pietilainen, University of Helsinki, our first collaborator, and members of her research group to provide me with such an incredible data set to work on, without which the objectives of this study would not exist. I benefitted a lot from her insights towards our publications. I would also like to address special thanks to the members of advisory committee, Dr. Jacek Majewski, Dr. Rob Sladek and my Graduate Advisor, Dr. Stephane Richard, for their valuable comments and suggestions for my project.

I am also indebted to all my colleagues and current members of Kaartinen lab with special thanks to Dr. Aida Kasaei-Roodsari and Dr. Huifang Sun for keeping me motivated and always involved, and to Ms. Sahar Ebrahimi Samani, and Ms. Preeti Bhatt for making the working environment around me so pleasurable. I would also like to thank my friend, Ms. Elisa Cohen for the French translation of the abstract presented here.

I owe a special appreciation to my undergraduate mentor and advisor, Dr. Ragothaman Yennamalli for introducing me to research during my undergraduate studies in India, his words always ignited a new fury in me; my dearest family, to my parents who have never stopped believing in me and have always encouraged me to keep going; my sister for unknowingly

teaching me so much about being happy even in the worst times; my uncle, aunt and brother in the UK for all the guidance and love; all my friends in Canada and the US, for joining me, supporting me in this adventure and reminding me every day that I could accomplish this goal when often I would forget; my friends in India, for never letting me miss home in a foreign country.

I dedicate this thesis to my grandparents without whom I would not be what I am today.

Mansi Arora

Contribution of Authors

This thesis is designed and written by Mansi Arora and reviewed by Dr. Mari Kaartinen.

Work presented in this thesis has been published as a manuscript. Thus, some textual overlap exists between these.

Kaartinen, M.T., Arora, M., Heinonen, S. *et al. F13A1* transglutaminase expression in human adipose tissue increases in acquired excess weight and associates with inflammatory status of adipocytes. *Int J Obes* **45,** 577–587 (2021). https://doi.org/10.1038/s41366-020-00722-0

Contribution of authors in this thesis is listed below:

- Kaartinen MT supervised the study and reviewed the thesis.
- **Arora M** designed and conducted the Bioinformatics analysis in R Statistical Software, GraphPad Prism, SPSS, and drafted the thesis.
- Pietiläinen group has collected all the clinical and transcriptomics data and shared it with Kaartinen group.

Work presented in this thesis is related to another study which is also published, for which the contribution of authors is same as above.

Kaartinen, M.T., Arora, M., Heinonen, S. *et al*. Transglutaminases and Obesity in Humans: Association of *F13A1* to Adipocyte Hypertrophy and Adipose Tissue Immune Response. *Int. J. Mol. Sci.* **21**, 8289 (2020). https://doi.org/10.3390/ijms21218289

List of Figures

List of Tables

List of Abbreviations

WHO – World Health Organization

BMI – Body Mass Index

AMDR- Acceptable Micronutrient Distribution Range

T2D – Type 2 Diabetes

AT- Adipose Tissue

MHO- Metabolically Healthy Obesity

MUO- Metabolically Unhealthy Obesity

TEE- Total Energy Expenditure

REE- Resting Energy Expenditure

DEE- Diet-induced Energy Expenditure

SVF- Stromal Vascular Fraction

ECM- Extracellular Matrix

WAT- White Adipose Tissue

sWAT- Subcutaneous White Adipose Tissue

vWAT- Visceral White Adipose Tissue

BAT- Brown Adipose Tissue

BMF- Bone Marrow Fat

DM- Diabetes Mellitus

UCP1- Uncoupling Protein 1

TG- Transglutaminase

WHR- Waist-to-hip Ratio

SAT- Subcutaneous Adipose Tissue

FDR- False Discovery Rate

1. Introduction

1.1 Obesity

Obesity - where excess body fat accumulates - is a global epidemic. According to the World Health Organization (WHO), its prevalence has nearly tripled worldwide, since 1975. Individuals with Body Mass Index (BMI) above 30 are classified as obese. The fundamental cause of gaining weight is the energy imbalance between calories consumed and calories spent. Multiple complex systems are responsible for maintaining the energy balance in the body. Each of these systems sense nutrition and energy levels and respond to changes in these levels by modifying the fuel availability. Globally, it has been found that there is an increased intake of energy-dense foods such as the ones high in fat and sugars. Another reason for obesity is the increase in physical inactivity due to modern day forms of work, varying modes of transportation and increase in urbanization.

Obesity is associated with clinical comorbidities that include type 2 diabetes (T2D), cardiovascular diseases, and several types of cancers. It is understood now that many of the comorbidities arise from failure of adipose tissue (AT) to function normally. Obesity has a strong genetic component which is estimated to be 40% to 70% (1, 2). It is considered that approximately 35-40% of children's BMI, how heavy or they are, is inherited from their parents (3). The obesity phenotype and comorbidities show great individual variation indicating that numerous genetic traits are behind weight gain and resulting clinical issues. A few studies indicate that children of obese parents are at a greater risk of developing obesity throughout their life and that such trends run in families (4). It can be expected that a heritable element in the etiology of adiposity. may be responsible for the synergistic relationship between the genetic make-up of an individual and the environmental factors (5).

1.1.1 Energy metabolism

Constant supply of energy is required for the survival of an organism. Energy homeostasis, or energy balance in simple words, is maintained by the physiological system by responding to changes in nutrient availability. Energy balance is affected by three parameters, energy intake, energy expenditure and energy storage. Body composition, especially weight, is one of important

features to assess energy balance. Based on the law of thermodynamics, energy storage in the body is determined by the disequilibrium between energy intake and energy expenditure, as follows:

- Increase in energy storage because energy intake is higher than the energy expenditure, also known as positive energy balance; causes increase in body mass
- Decrease in energy storage because energy expenditure is higher than energy intake, also known as negative energy balance; causes weight loss
- Stable body mass because energy intake equals energy expenditure; also known as normal energy balance (6-8)

Energy intake in humans is primarily associated with macronutrient profiles in various diets. These include majorly carbohydrates, protein, fats and alcohol (9). A major source of dietary energy is carbohydrates (50% of dietary intake), which are eventually converted into glucose. Glucose is used as energy or is stored in liver as glycogen. Excess glucose gets converted to triglycerides to be stored in AT (10). Apart from this, proteins contribute to nearly 15 % of the dietary energy. Amino acids have multiple forms of storage in the body, such as structural proteins, enzymes and other cellular proteins. Gluconeogenesis is a process by which non-carbohydrate material, such as amino acids, gets converted to glucose. This glucose is further utilized as energy or gets converted to glycogen by glycogenesis, or triglycerides by lipogenesis (6, 7). In addition, alcohol can also contribute to substantial amount of calories to a diet and contribute to obesity.

To study the conditions of long-lasting imbalance of energy intake and energy expenditure, it is important to consider the daily energy and nutrient balance of a healthy body with respect to macronutrient energy stores, intake and oxidation. According to Health Canada, nearly 25% of males and 23% of females in Canada, aged 19 years and older, have fat intakes above the Acceptable Macronutrient Distribution Range (AMDR - for fats, 25-35% of total energy intake). Fat stores inside body, also known as energy buffers, are large as compared to stores of other nutrients. Daily fat intake represents less than 1% of total fat stores and has very little impact on fat oxidation. The intakes and oxidation rates of carbohydrates, protein and alcohol are tightly controlled on a regular basis. Therefore, any major difference in energy intake and expenditure, impacts fat stores predominantly. Fat is one of the main nutrients causing energy imbalance as most of the energy is stored in AT as fats. Other nutrients contribute to energy imbalance indirectly by modulating this fat deposition in AT (6).

Maintenance of whole-body homeostasis is important, and it is achieved by efficient communication and coordination between the central (brain) and peripheral organs (stomach and intestine, liver, muscle, AT, and pancreas) of the body. A strong regulation exists between the energy intake and energy expenditure which directly affects the biochemical processes required to maintain cellular viability, physical activity and adaptive thermogenesis (11). Basal metabolic rate accounts for most of the energy needs of the body. It is the largest contributor to Total Energy Expenditure (TEE) (12). TEE can be divided into three components: maintenance or resting energy expenditure (REE), energy expenditure for the processing of ingested food or diet-induced energy expenditure (DEE), and activity-induced energy expenditure (AEE) (13). Fasting is one of the examples of food intake regulation. It triggers the increase of food intake. Food intake induces changes in energy expenditure as a function of changes in body size and body composition. In addition, energy restriction induces an adaptive reduction of energy expenditure through a lowering of tissue metabolism and a reduction of body movement (13).

There are complex interactions between regulatory centers in the Central Nervous System (CNS) in hypothalamus and regulated storage, and fat stores mobilization. A homeostatic network like this, maintains all energy or fat stores of the body. Hence, genes involved in activities of the molecular components of this system, may underpin obesity and other metabolic disorders (2).

1.1.2 Physiological Effects of Insulin

Made by the beta cells of pancreas, insulin is an anabolic hormone responsible majorly for the body to use glucose from carbohydrates in the food, in order to obtain energy or to store glucose for future use. Insulin has profound effects on both carbohydrate and lipid metabolism, and significant influences on protein and mineral metabolism.

Maintenance of **glucose homeostasis** primarily results from regulated insulin secretion. In coordination with its counter-regulatory hormone, glucagon, insulin prevents the blood sugar level from getting too high (hyperglycemia) or too low (hypoglycemia). Due to high postprandial blood glucose levels, beta cells in the pancreas are signaled to release insulin into the bloodstream (**Figure 1**). Insulin then binds to insulin receptors at the cell surface, thus promoting tyrosine kinase activity, and signals for glucose uptake into insulin-dependent muscle and adipose tissues. It also promotes **glycogenesis** (glucose → glycogen) in the liver.

Figure 1

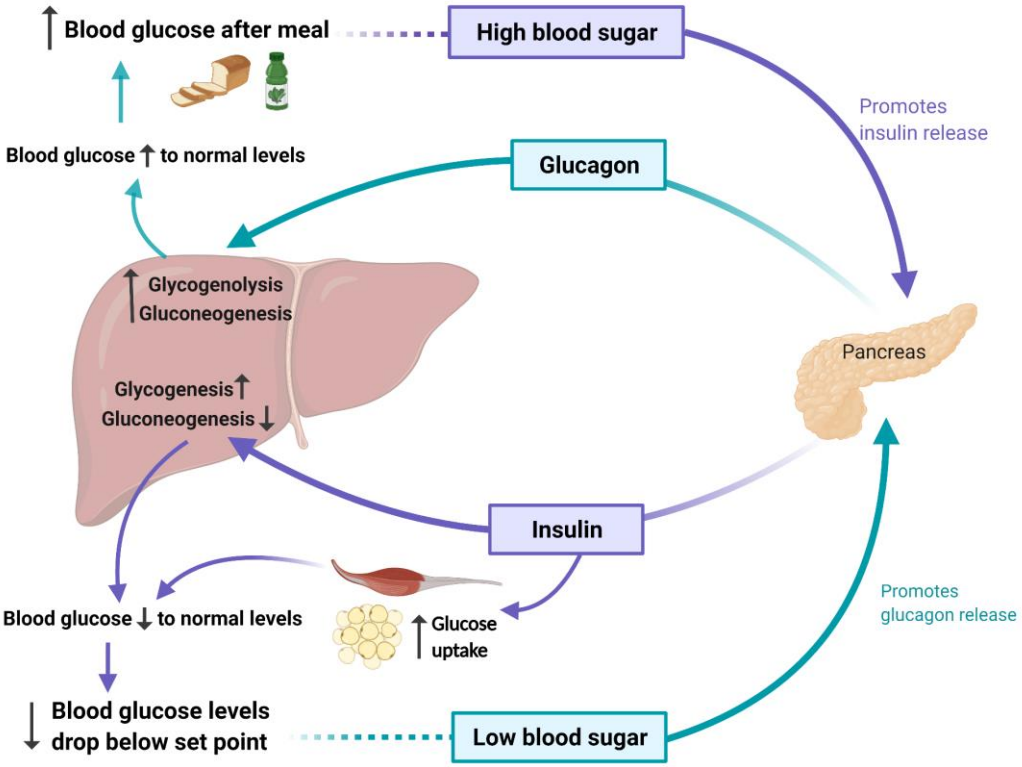

FIGURE 1 : Metabolism of blood glucose showing the coordinated release of insulin and glucagon, with corresponding change in levels of blood glucose levels. The mechanism works like a feedback loop. High blood glucose levels after meal cause pancreas to release insulin. This further signals glucose uptake into insulin-dependent muscle and adipose tissues, and promotes glycogenesis. When blood glucose concentration drops, pancreas secrete glucagon, which increases endogenous blood glucose levels through hepatic glycogenolysis. (Figure created using template from Biorender)

When blood glucose concentration drops, such as during exercise, the alpha cells of pancreas secrete glucagon, which increases endogenous blood glucose levels through hepatic **glycogenolysis** (glycogen → glucose). Additionally, during prolonged fasting, glucagon also stimulates hepatic and renal **gluconeogenesis** (de novo glucose synthesis). It also inhibits glycolysis and glycogenesis (15, 16).

Insulin is an important hormone which aids in lipogenesis and inhibits lipolysis, leading to storage of free fatty acids (FFA) in the form of triacylglycerols in the AT (17). Once liver is saturated with glycogen storage, glucose is then used to create fatty acids which are converted into lipoproteins and released into blood stream. These further breakdown into FFA to be used by other tissues. In the fat cells, glucose is used to form **glycerol**, which combines with excess FFA from liver and forms triglycerides. Insulin prevents the breakdown of fat in AT and triglycerides into fatty acids, hence causing triglycerides build up in fat cells.

Another important contribution of insulin is to increase the uptake of amino acids by tissues and enhance protein synthesis thereby reducing proteolysis. It also decreases the rate of protein degradation in muscle (17, 18).

1.1.3 Adipose tissue and adipocytes

AT is one of the important endocrine organs. It produces multiple hormones for regulation of metabolism. It is a very specialized loose connective tissue and is made up of cellular components like adipocytes, preadipocytes, macrophages (innate immunity), fibroblasts, endothelial cells, pericytes, and T cells, NK cells and mast cells (adaptive immunity). It also contains multi-potent stem cells with the support of vascularized loose connective tissue (19-22).

Almost 90% of AT by volume, consists of adipocytes and the rest 10% is stromal vascular fraction (SVF). Adipocytes are approximately 30-150μm in diameter. 85% of adipocyte volume is constituted by a large lipid droplet and the remaining 15% accounts for nucleus and cytoplasm displaced to the periphery. The SVF contains other cells like macrophages, endothelial cells, preadipocytes and multipotent stem cells. 15-50% of the SVF is composed of preadipocytes (22-24).

AT also comprises non-cellular component which is composed by extracellular matrix (ECM). ECM maintains the structural and functional integrity of AT. It also acts as a protection against

mechanical forces. Adipocytes and SVF contribute to the production of ECM components. ECM is constituted by collagen type I, III, IV and type VI, fibronectin, laminins, nidogens, decorin, tenascin C, osteopontin, matrix metalloproteinases, secreted protein acidic and rich in cysteine (SPARC) and many other components (22, 25, 26)

Obesity corresponds to increase in the fat cell mass and can lead to hormonal imbalances and has various metabolic effects. Metabolic syndromes or the metabolic complications due to obesity consist of insulin resistance often culminating in β-cell failure, impaired glucose tolerance and type 2 diabetes, dyslipidemia, hypertension, and premature heart disease. Other complications include abdominal obesity, ectopic lipid accumulation, hepatic steatosis, and sleep apnea (27). White adipose tissue is the main depot responsible for energy storage. It is a source of a variety of hormones which are responsible for regulation of body weight, such as leptin, visfatin, apelin, resistin, and adiponectin. Adipocytes have also been shown to secrete many types of biologically active molecules, called adipocytokines (28).

1.1.4 Adipose tissue depots

Development and progression of metabolic diseases associates more with fat distribution rather than the fat mass alone. Therefore, it is important to study the fat depots present in the body (22, 29).

White Adipose Tissue (WAT) is the most abundant AT in the body. It is pale in color with a glistening surface. Based on anatomical distribution, WAT is rather classified into two subtypes, subcutaneous WAT (sWAT) and visceral WAT (vWAT). Due to genetic differences in preadipocyte differentiation and epigenetic influences within environment, these two subtypes of WAT have morphological, physiological and functional differences (30, 31) such as adipokine secretion, rate of lipolysis and triglyceride synthesis. sWAT depots are present under the skin whereas vWAT depots mainly comprise the mediastinum and abdominal cavity around body's internal organs.

Brown Adipose Tissue (BAT) distinctively appears brown in color because of the high content of mitochondria and rich vascularity. It differs from white adipocyte in being multilocular due to presence of multiple small droplets of lipid. Relative to white adipocytes, brown adipocytes are

smaller with a diameter of approximately 60µm. These adipocytes contribute to nearly 50% of the total cells (24, 32). BAT is present in many locations of the body and it might increase, or decrease based on age, species, nutrition conditions and environmental factors. These cells are rich in mitochondria which contain uncoupling protein 1 (UCP1). UCP1 is involved in uncoupling of oxidative phosphorylation for generation of heat for non-shivering thermogenesis (32, 33).

Bone Marrow Fat (BMF) results from an accumulation of fat cells within the bone marrow. BMF has a distinctive phenotype and is different from other fat depots as in subcutaneous or visceral tissues. BMF is not correlated with weight, BMI or body fat (34, 35).

1.1.5 Adipose tissue under weight gain

White Adipose Tissue maintains healthy energy metabolism in obesity. Weight gain is characterized by AT expansion where adipocytes undergo hyperplasia (proliferation) and hypertrophy (adipocyte enlargement), along with cellular mitochondrial dysfunction and tissue ECM remodeling, inflammation, and altered angiogenesis.

1.1.5.1 Modes of growth – hyperplasia and hypertrophy

Adipocytes respond to energy surplus by two modes: via growing in size, i.e. Hypertrophy and via increasing in numbers i.e., hyperplasia **(Figure 2)** , the latter has been linked to metabolically healthier outcome. (36).

Nutrient excess maintains the buffering capacity of the AT. As an adaptive response to this excess of nutrients, the adipocytes show hypertrophy. The relationship between metabolic disorders and adipocyte hypertrophy is contextual. It has been suggested that there is a need for larger adipocytes for nutrient buffering capacity. Lean humans with a smaller adipocyte size have a worse metabolic outcome to overfeeding; this clearly indicates the importance of larger cells for nutrient buffering capacity (36). In some obese patients, this buffering capacity of adipocytes is exceeded because of expansion of the cells beyond a certain threshold, leading to ectopic lipid deposition in peripheral tissues and also impairing adipocyte metabolism (36).

Figure 2

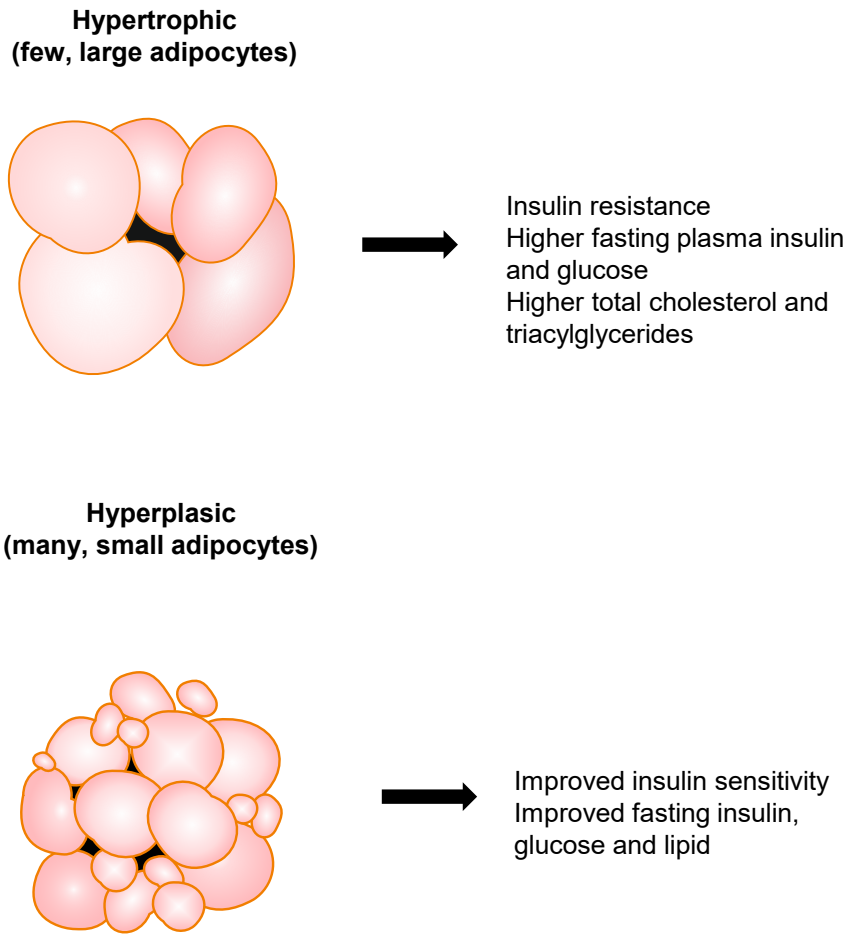

FIGURE 2 : Hypertrophy and Hyperplasia Symbolic representation of two ways of cell expansion. Hypertrophy corresponding to few large adipocytes and hyperplasia corresponding to multiple small adipocytes.

Studies suggest that AT fibrosis is a probable adaptive response to limit the adipocytes to expand beyond threshold, tending to reduce cell dysfunction. Differences in this AT fibrotic response among humans may therefore be a contributing factor to differing susceptibilities to metabolic disease (36, 37).

On similar lines, studies in the past have proven that adipocyte hypertrophy affects BMI strongly and is related to metabolic disorders in humans and mice (38-43). Local hypoxia causes AT accumulation due to hypertrophic or hyperplasic growth of the cells. It is a major trigger for tissue remodeling, adipocyte death, inflammation, tissue fibrosis and most importantly, many metabolic dysfunctions in the AT (44).

Based on the fat distribution, obesity can be classified as 'android' or 'gynoid'. Android obesity is characterized by fat storage around abdominal region and leads to individuals having an apple-like body shape. Fat can also be manifested in other areas of the upper trunk like the upper chest (front or back) nape area of the neck, and even the shoulders. Gynoid obesity on the other hand has excess fat being deposited around hip and thigh areas, leading the body form to look like a pear fruit. Studies suggest that the early obesity tends to have subcutaneous deposition of fat, and that visceral deposition of fat happens only after subcutaneous capacity has been reached. Both these fat depots show hypertrophy and hyperplasia, although, as obesity increases in severity, hyperplasia becomes most evident. Hyperplasia is believed to occur during early onset of obesity as well. However, it is also suggested that reaching specific mean fat cell size triggers a consequent increase in the fat cell number. Hence, it is challenging to infer what exactly causes the average increase in adipocyte number in severely obese individuals, adult adipocyte recruitment or predisposal from pre-adulthood to become obese and/or severely obese (45, 46).

1.1.5.2 Adipose tissue remodeling

AT is known for its unique property of unlimited capacity to expand without a transformation state. The cell alterations correspond to nutrition deficiency or excess, to which adipocytes adapt by hypertrophy or hyperplasia in order to maintain the whole-body homeostasis. One of the mechanisms involved in AT remodeling is hypoxia which is a limiting factor for healthy expansion of the cells. Other effects of tissue expansion include adipocyte cell death, enhanced chemokine secretion, and dysregulation in fatty acid fluxes (47).

The healthy expansion of AT is a result of adequate angiogenic response, proper remodeling of ECM and is efficiently coordinated within multiple cell types such as endothelial precursor cells, immune cells, preadipocytes and most importantly macrophages (47). Macrophages are involved in inflammatory response and their related genes have been identified to be significantly altered in obese AT (48). Unhealthy expansion of cells on contrary consists of humongous enlargement of adipocytes accompanied by limited angiogenesis and subsequent hypoxia **(Figure 3)**. AT remodeling is an ongoing process that is pathologically accelerated in the obese state, although, all states of cell expansion might not be associated with changes in pathological conditions. It is expected that adipogenesis and coordinated vascularization of adipose depot allow for preserved function of WAT in obesity (49). Therefore, 'metabolically healthy obese' individuals bypass these pathological changes corresponding to obesity by preserving insulin sensitivity due to healthy tissue expansion. This way the lipotoxic side effects associated with obesity are also avoided (47). Tissue remodeling involves processes such as ECM over-production, reduction in angiogenic remodeling, aggravation of immune cell infiltration and consequential proinflammatory response (47).

1.1.6 Insulin Resistance

Insulin resistance is pathological state where normal or elevated insulin level produces attenuated biological response. People with insulin resistance, also known as impaired insulin sensitivity, have built up a tolerance to insulin, making the hormone less effective.

Caloric excess triggers WAT structural and cellular remodeling during obesity. AT expands by increasing adipocyte number as well as their capacity to accumulate lipid. It undergoes remodeling of vasculature and ECM to allow tissue expansion, nutrient mobilization and oxygenation. However, sustained obesity and inflammation leads to failure of these adaptive mechanisms. This results in AT dysfunction, ectopic lipid accumulation, and insulin resistance, which in turn are tightly linked to the development of T2D (50, 51).

Multiple studies involving mouse models, have shown that adipogenesis allows AT to expand while limiting hypoxia, chronic inflammation, and fibrosis, and that healthy fat mass induces improved metabolic health (52-54). These findings indicate a possibility that "obesity-associated metabolic decline is not due to adiposity, per se, but rather is a result of an insufficient capacity of AT to further expand" and accommodate energy surplus (49).

Figure 3

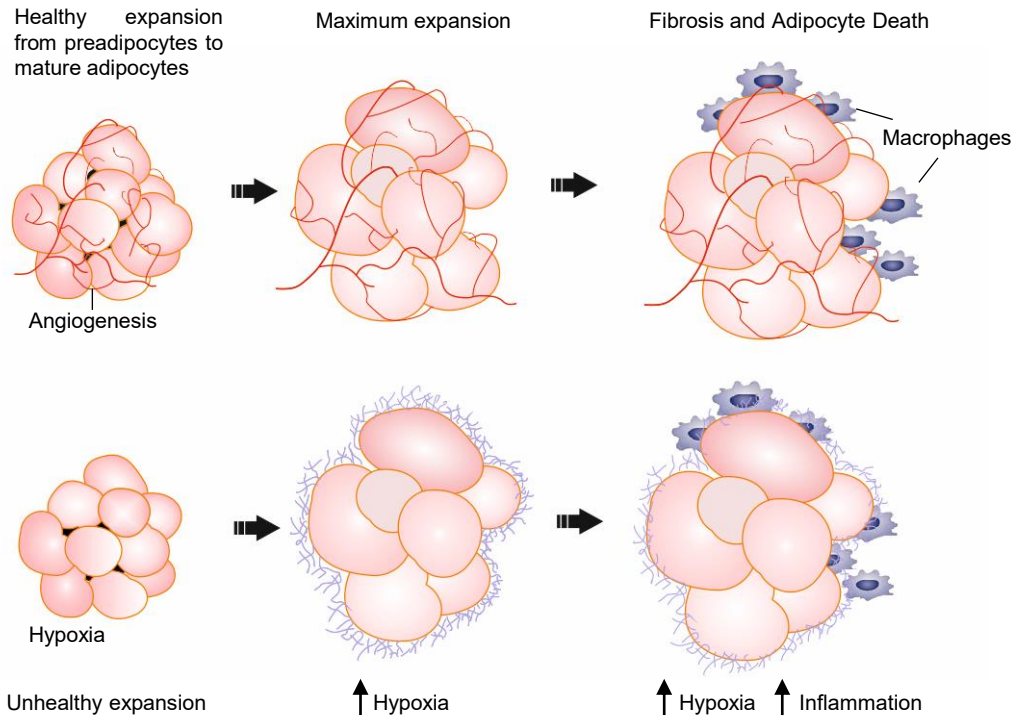

FIGURE 3 : Healthy and unhealthy adipose tissue expansion. Healthy expansion involves a controlled expansion of cells with appropriate angiogenesis further leading to apoptosis. Unhealthy tissue expansion shows reduced angiogenesis and uncontrolled enlargement of adipocytes, leading to increase in hypoxia and inflammation due to persistent macrophages. Recreated with modifications from (Sun, Kusminski et al. 2011)

It can be inferred that this limited capacity, and WAT dysfunction are key determinants for onset and progression of obesity-related pathological consequences as a result of ectopic toxic lipid deposition, for example in liver. (47, 51) One of the studies hypothesizes that T2D is likely if this personal fat threshold is exceeded. (55) However, its challenging to determine whether pathologic AT remodeling is driven by defects in adipocyte differentiation or it is a consequence of fibrosis and other metabolic changes. (50, 51).

Metabolically, accumulation of central body fat, is more harmful than the peripheral pattern fat distribution in the body. The former is related with insulin resistance. The body is more insulin resistant due to an increase in intra-abdominal or visceral fat than having increased quantities of centrally located subcutaneous fat. Where obesity is associated with insulin resistance, even lean individuals with a difference of body fat distribution tend to have a difference in insulin sensitivity. The metabolic effects of intra-abdominal fat and subcutaneous fat depots differ greatly due to variations in characteristics of AT from the both. One such example can be intra-abdominal fat which has a higher expression of genes encoding secretory proteins and proteins involved in energy production (56).

The variance in insulin resistance of a person may be expected to be determined significantly with genetics. However, there are many environmental and behavioral risk and factors as well, which have significant contribution. Studies have shown that increase in body weight causes insulin action to decline and insulin resistance is a major underpinning link between physical inactivity and metabolic syndrome (57). Obese people with less physical activity have a higher tendency of insulin resistance because of decreased burning of glucose and fatty acids. Any mismatch between meal schedules and circadian rhythms might also increase insulin resistance (58). Dietary patterns, type of food intake, sedentary lifestyle, alcohol intake, smoking, are all important factors which contribute to insulin changes in the body. However, the underlying processes with respect to causes of insulin resistance are not completely understood. More than 30 years ago, both obesity and reduced physical activity have been investigated and associated with high blood pressure, glucose intolerance and hyperinsulinemia (59). Insulin resistance has also been linked with chronic liver disease, especially nonalcoholic fatty liver disease, the occurrence of which has a strong association with obesity and other components of metabolic syndrome as well. (60).

There is recent growing evidence that changes and abnormalities in microbiota composition of the body, to be one of the factors responsible for diet-induced obesity, and in turn insulin resistance. One example of microbiota is of Gram-negative bacteria, which release an endotoxin, lipopolysaccharide, and provoke an inflammatory response in the body (61, 62). Several other factors which have been proposed to be responsible for the mechanisms of insulin resistance include mitochondrial dysfunction, endoplasmic reticulum stress, aging, oxidative stress, lipodystrophy and pregnancy (63). Specific approaches made it possible to comprehend some of the interactions between certain bacterial strains and their host, and how their metabolites may interfere with host's cell signaling, changing its metabolic profile.

1.1.7 Diabetes

Diabetes mellitus (DM) is a chronic disease characterized by prolonged hyperglycemia, caused by defects in insulin secretion and/or insulin function. Blood glucose levels increase if insulin is not produced by the β cells in enough quantities, and/or if the body does not respond to the circulating insulin. This leads to prediabetes and/or diabetes. Over the due course of time, continuous high blood glucose levels cause damage to number of tissues including nerves and blood vessels which further promotes complications such as heart disease, stroke, kidney disease, blindness etc.

Diabetes Mellitus is differentiated into following.

- **Type 1 Diabetes** contributes to nearly 5% of all diabetes worldwide. It is an autoimmune disorder where the body's immune system attacks and destroys the β cells. It is also known as juvenile diabetes as the onset occurs usually in childhood or early adulthood, i.e. <35 years.
- **Type 2 Diabetes** accounts for approximately 90% of all diagnosed diabetes worldwide. It is characterized by insufficient synthesis and secretion of insulin, coupled with impaired insulin action in target tissues such as muscle, liver and fat. The upregulation of insulin secretion from pancreatic beta cells, is unable to compensate for the increasing insulin resistance (or decreasing insulin sensitivity). This results in hyperglycemia (64, 65). T2D is normally diagnosed after 40 years of age. This can be further subdivided into two groups, diabetes with obesity and without obesity. In non-obese T2D, a deficiency in insulin production and release persists, corresponding

to insulin resistance at post receptor levels. Whereas, in obese T2D, the insulin resistance occurs due to modifications in cell insulin receptors. This associates to distribution of abdominal fat.(56)

- **Gestational Diabetes** occurs during pregnancy, where the woman, who did not have diabetes before pregnancy, develops high blood sugar levels. If not well controlled, it can lead to related complications, such as large sized baby, Cesarean section, preeclampsia, and hypoglycemia. However, it provides a unique opportunity to study the early pathogenesis of diabetes and to develop interventions to prevent the disease (66).

Apart from glucose, the onset of diabetes (both T1D and T2D) is accompanied with alterations in other metabolites, typically fructose, amino acids and lipids. Studies suggest that insulin has effects not only at the translational level but also at the transcription level of protein synthesis (67). It has been established that impaired insulin-stimulated glucose disposal and reduced mitochondrial protein synthesis is found to occur in T2D patients (67, 68). High concentrations of low-carbon lipids, and increased levels of fatty acids have been seen in individuals suffering from T2D. Although diabetes and metabolic disturbances occur jointly, it is unclear as to which of the two is causative to the other (69). Some metabolomic studies demonstrate relationships between diabetes and metabolomic profiles, which can facilitate diagnosis and prediction of occurrence of diabetes (70-72), whereas some studies focus on metabolites and their association with prediabetes measures (73, 74).

1.1.8 Comorbidities of obesity – links to adipose tissue

Human fat depots are characterized by AT morphology, i.e. size, diameter, volume and number of adipocytes, which are responsible for the total weight of the depots, thus regulating fat distribution. Any kind of modification in the adipocyte function also effects adipose distribution. The synergy between visceral obesity due to accumulation of fat mass and development of multiple metabolic syndromes like type 2 diabetes and hypertension was first investigated by Spiegelman's group (75, 76).

AT is widely distributed in different regions of the body. The distribution of fat in the central body is independent of the total adiposity. Genetic factors play a major role in heritability of this fat accumulation. One of the genome wide association studies and meta-analysis has revealed 68 loci

to be associated with body fat distribution, more closely to waist-to-hip ratio, as well as other related traits (77). Further it has been found that many genes located in the vicinity of these 68 loci were expressed in AT at mRNA level and involved in pathways responsible for regulation of fat distribution. However, nothing has been proven with respect to these loci being involved in adipocyte function.

1.1.9 Metabolically healthy obesity

The concept of Metabolically Healthy Obesity (MHO) is based on clinical observation that one third of obese individuals do not exhibit metabolic abnormalities. There is no standard definition of MHO (78-81). Apart from the general prerequisite of defining MHO by BMI being above 30 kg/m^2, clinical studies claim nearly 30 various definitions of metabolic health (Rey-López, de Rezende, Pastor-Valero, & Tess, 2014). Usually, MHO has been defined with respect to absence of any metabolic disorder and cardiovascular disease, including type 2 diabetes, dyslipidemia, hypertension, and atherosclerotic cardiovascular disease (ASCVD) in an obese person. Yet, large variations exit between different studies regarding MHO classification criteria and specific cut-off values for any considered parameter (82). This may imply that a distinct subgroup of obese individuals would not be biologically determined by MHO. Individuals with MHO (prevalence ~10–30%) are expected to have characteristic phenotypic traits like lower liver fat and visceral fat mass, higher leg fat content, greater cardiorespiratory fitness and physical activity, insulin sensitivity, normal inflammation markers, and preserved AT function compared to patients with metabolically unhealthy obesity (MUO, prevalence ~80–90%) (82). Recent studies have tried to standardize definitions of MHO for clinical research on obesity related morbidities, and differences between MHO and MUO. The risk to developing cardiometabolic diseases is lower in people with MHO compared to MUO. (82, 83)

1.1.10 Metabolic parameters as measures of obesity

Obesity is a multifaceted phenotype which cannot be reflected by a single parameter. Thus, there is a need of describing such robust phenotypes which can further explain the physiological and pathological states.

1.1.10.1 Body Mass Index

Obesity is determined by Body Mass Index (BMI) which is measured as a ratio of weight in kg to the square of height in meters. However, BMI might not be a clear indicator of how healthy a person is, or in other words, what is the likelihood to develop metabolic diseases. The numbers can be too general. Yet, broadly, according to WHO guidelines, BMI is used for a simple classification of individuals. **(Table 1)**.

BMI is not a perfect measure of health, which might indicate presence or absence of a disease. It does not measure body fat directly, but it is correlated with various metabolic and disease outcome, since these are more direct measures of body fatness (84, 85). The fat percentages do not have enough research behind them to indicate the risk of disease these individuals face. This research is improving as more and more people are using their BMI as an indicator of overall health.

As of now, BMI doesn't distinguish muscle from fat. A better predictor of weight related diseases is body-fat percentage. A study by Gallagher et al. in the past tried to correlate the body fat percentage to BMI with the existing healthy weight guidelines. It involved generating models for predicting body fat percentage using multiple regression analysis with BMI and other anthropometric variables (83).

1.1.10.2 Waist-to-hip ratio

Waist-to-hip ratio (WHR) is a quick measure of fat distribution which can help indicate a person's health. It is simply the ratio of waist circumference to the hip circumference. According to the WHO, having a WHR of over 1.0 **(Table 2)** may increase the risk of developing conditions that relate to being overweight.

A study in the past has shown that WHR was a better predictor for cardiovascular events and mortality in patients with type-2-diabetes compared to BMI and waist circumference (86). People who carry more weight around their midsection (an apple-shaped body) are at higher risk for heart disease, type 2 diabetes, and premature death than those who carry more of their weight in their hips and thighs (a pear-shaped body).

Table 1: BMI classification according to WHO

BMI	CLASSIFICATION
< 18.5 kg/m^2	Underweight
18.5 – 24.9 kg/m^2	Normal/Average
25.0 – 29.9 kg/m^2	Pre-Obesity
30.0 – 34.9 kg/m^2	Class I Obesity
35.0 – 39.9 kg/m^2	Class II Obesity
≥ 40.0 kg/m^2	Class III Obesity

Table 2: Waist-to-hip ratio in men and women as classified by WHO. A WHR above 0.90 in men and above 0.85 in women indicates obesity.

Health risk	Men	Women
Low	0.95 or lower	0.80 or lower
Moderate	0.96-1.0	0.81-0.85
High	1.0 or higher	0.86 or higher

1.2 Genetical basis of obesity

The possible links between genetics and obesity has led researchers to further investigate the molecular basis of disease. Over the course of recent two decades, multiple monogenic and polygenic genes have been discovered to be associated with obesity in some way or the other. In 2001, six genes were linked to monogenic obesity and no common variants were linked with polygenic obesity. It is important to study the overlap of these monogenic and polygenic forms of obesity in order to explore the causal genes. Previously, studies have shown obesity related genes like Melanocortin-4-Receptor *(MC4R),* Proprotein Convertase Subtilisin/Kexin Type 1 *(PCSK1),* Pro-opiomelanocortin *(POMC)* and Brain-derived neurotrophic factor *(BDNF)* to be associated with other traits as well (87). *MC4R* has been one of the first genes to be explored for mutations on the genetic level (88, 89). There have been many studies in the past, relating to specific commonly occurring small variations, termed as Single Nucleotide Polymorphisms (SNP). One such genome wide association study involving specific SNP locations, patterns and associated phenotypes, identified the variants of this gene to correlate with fat mass and weight (90). By 2008, eight monogenic and four polygenic genes were discovered with the help of genome-wide significant studies (91). These four genes were Fat mass and obesity-associated protein *(FTO), PCSK1, MC4R,* and Catenin Beta Like 1 *(CTNNBL1)* (87). *FTO* was the gene which was identified as the first locus harboring common variants with a significant impact on obesity predisposition and fat mass (92-94). More than 400 articles on different ethnic populations have been published with increasing thirst to understand the mechanisms by which this gene might be linked to obesity. All these have led to establish *FTO* as one of the 'causal genes' for obesity. On the other hand, it has also been suggested that instead of being directly correlated to weight, monozygotic twins may carry 'variability genes' which make the individuals more susceptible to environmental variation (95)

1.2.1 Genome Wide Association Studies on obesity

Genome-wide association (GWA) studies are a relatively new way for scientists to identify genes and genetic variations involved in human disease. This method searches the genome for small variations, called SNPs, that occur more frequently in people with a particular disease than in people without the disease. GWAS are robust because they investigate the entire genome (non-

candidate driven approach) instead of testing for a pre-specified set of genes (gene-specific candidate-driven studies). Each study can look at hundreds or thousands of SNPs at the same time. Researchers use data from this type of study to pinpoint genes that may contribute to a person's risk of developing a certain disease. But they cannot identify causal genes. These studies involve searching of genetic markers over genome scanning to look for the presence of disease (96).

Many GWAS have been conducted for obesity, making the disease one of the most popular phenotypes to investigate because of the relative ease of procuring the phenotypic data (96). The affect obesity has on multiple metabolic complications makes it even more intriguing to understand the disease at genetic level. GWAS look for associations with Simple nucleotide variations (SNV) and multiple genetic factors across the whole genome at once, to correlate with particular traits (96). Earlier linkage studies have identified regions on chromosomes which might harbor obesity genes, yet the known mutations have been able to explain only a small fraction of variations. Multiple meta analyses have failed to reveal novel genes. One way to study these variations is by Mendelian randomization which uses genetic variation as a randomly redistributed variable among populations to control for unobserved confounding variables in an observational setting (97). This method was implemented for establishing the causative relationship between type 2 diabetes and weight gain as a result of BMI association of *FTO* (92). GWAS provide a broader platform with exploring SNPs, however, statistically, unrealistic sample sizes are required. For example, MC4R identification needed data of 77,000 individuals; 91,000 individual samples were required for identifying just 6 loci for BMI. Although robust and novel findings exist, only small variation in BMI has been explained. The 6 loci could explain only 0.84% of variation (96).

1.2.2 Causative and Reactive genes

GWAS for obesity are not as direct as they might be expected. There are hurdles associated in finding putative causative genes for the disease.

1) Correlation of gene expression in AT with a complex phenotype like BMI does not directly imply causation. With no definite phenotype associated, finding a similar genetic pattern is a difficult task.

2) It is difficult to differentiate transcripts which may be unrelated to disease state, from those which are related to causal processes.

3) Finding causative and reactive genes involved in obesity is a challenge.

To partially overcome these problems, Pietiläinen and group have used a set of MZ twins to identify reactive and causative transcripts by means of Genome wide transcript profiling, and further carried out meta analyses of GWA data from ENGAGE consortium to explore genotyped SNPs located in coding sequence positions. Here, reactive genes are environment effected and up/down regulated only in the obese twin. Causative genes are the ones up/down regulated in both the twins as well as in unrelated obese population. (96)

1.3 Twin study cohorts in obesity

Multiple studies in the past have involved monozygotic or dizygotic twins to explore the genetics. Weight discordance is very rare in monozygotic twins. However, if found, research on such pairs has an upper hand for exploring epigenetic or environmental causes and consequences of obesity. Multiple studies have been performed on such weight discordant monozygotic twin pairs by Pietiläinen's group. One of them involved investigating the growth patterns in stature, weight and BMI where one twin is heavier than the other. The control group has weight concordant pairs. (98) Furthermore, the same set of weight discordant twins were explored to establish that increase in intra-abdominal and liver fat, and insulin resistance associate with acquired obesity, exclusive of genetic mechanisms (99). Some previous twin studies for phenotypes and genetic correlations have concluded that physiologically similar phenotypes tend to have stronger genetic correlations than other unrelated phenotypes (100, 101). Based on the same lines, Bogl and group quantified genetic and environmental factors involved in central obesity along with numerous metabolites, which are potential candidates for metabolic biomarkers (102). In support of this by means of studying BMI during childhood, few other twin studies (103-111) indicate that both genetic architecture and environmental factors have an impact on BMI variations (112). Use of relative weight as an indicative of obesity is the most important and distinctive characteristic of all these studies. Using twin dataset for investigation is unique because of the similarity of genome and pairing of the data, which is why relative values make sense. MZ twins share essentially 100% of their genes, so studying variations and the environmental factors, metabolism and other consequential effects is more significant. Pietiläinen and group refer to the arrangement of weight discordant MZ twins as 'clonal controls' of such studies as they can control genetic background whilst exploring, not only causes and effects of obesity, but also novel candidate genes for the disease. (95)

1.4 *F13A1* and its associations to obesity

F13A1 encodes the coagulation factor XIII A subunit. Coagulation Factor XIII is the last zymogen to become activated, stabilizing fibrin network as the last step in the blood coagulation cascade. Plasma factor XIII is a heterotetramer composed of 2 A subunits and 2 B subunits. The A subunits have catalytic function, and the B subunits do not have enzymatic activity and may serve as plasma carrier molecules. **(Figure 4)**

F13A1 stabilizes fibrin network as the last step of blood coagulation cascade (113, 114). In addition to being found in plasma, it is also found in tissues and is synthesized by various cells including macrophages, chondrocytes, osteoblasts and osteocytes, where it is found in the cytosol, nucleus and on the plasma membrane or cell surface, and in the ECM (114).

Recent genome-wide association study by Pietiläinen and group, in rare weight-discordant MZ twins identified 28 novel AT genes in that linked to obesity. A top gene with most significant association in this set was *F13A1*. (96). To find causative genes and to control the background, the investigators used MZ Finnish twins discordant in weight, i.e. lean-obese twin pairs. Gene expression analyses in WAT, carried out by calculating the difference values of expressions, identified significant changes in 27 genes in both lean and obese twins of which the gene with highest association to obesity was *F13A1*.

The significant association of *F13A1* with obesity was further confirmed in a large European ENGAGE consortium study of more than 21,000 unrelated individuals as well as in the GenMets cohort study which identified 7 SNPs in *F13A1* gene associated with weight. Further studies in larger human obesity cohort (GenMets) also identified seven small nucleotide polymorphisms (SNPs) in *F13A1* gene that were associated with body mass index (BMI). Also, a *F13A1* SNP (rs7766109) has also been linked to BMI and insulin resistance in polycystic ovary syndrome.

Figure 4

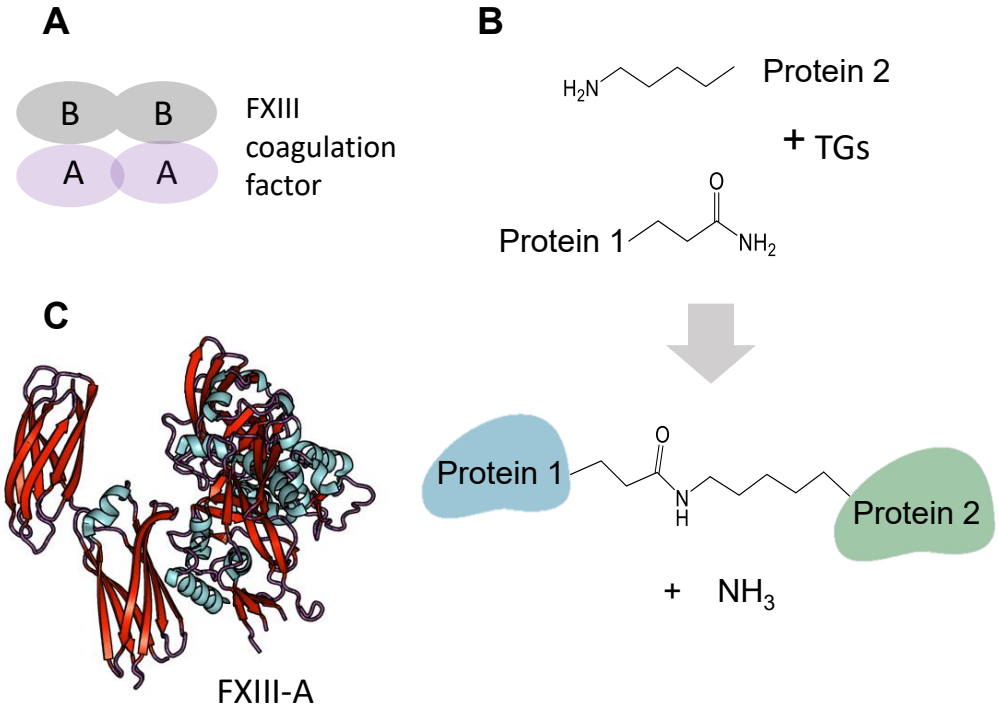

FIGURE 4: Basic structure A. FXIII coagulation factor is a heterodimer composed of FXIII-A and FXIII-B subunits. FXIII-A is a transglutaminase that is also expressed in many tissues. **B.** Isopeptide crosslinking reaction. **C.** FXIII-A protein structure (PDB id: 5mhm). Cartoon representation of the structure generated using Pymol, where the helices are colored cyan, beta-strands are colored red, and loops are colored magenta.

1.5 Whole genome screening and statistical methods of analysis

1.5.1 Normality – Graphical and statistical methods, Skewness, Kurtosis

The probability of getting an observation in a particular range of values is described by any probability distribution. The normal distribution is the familiar bell-shaped curve, with a high probability of getting an observation near the middle and lower probabilities as we get further from the middle. A normal distribution can be completely described by just two numbers, or parameters, the mean and the standard deviation. One of the assumptions for measurement variables is that the data fit the normal probability distribution, also called Gaussian distribution (after Johann Karl Gauss, 1777–1855). The tests which assume such distribution, inherently assume that the data can be described by two parameters, the mean and standard deviation, these tests are called parametric tests. Non-parametric tests on the other hand, are suitable when the data does not follow normal distribution. Various other distributions exist, where mean and standard deviation have different relationships.

Methods to test Normality:

The data to be analyzed need to be exactly normal, but the presumption that it is, shall be true, to avoid poor results. For the dataset to pass normality, there are two ways to do it.

1.5.1.1 Graphical methods

Simple graphical tools for assessing normality include Histogram, quantile-quantile plot (**QQ plot**), or even a box plot. The QQ plot is one of the most widely used graphical normality testing methods because of better accuracy to visualize the data. It plots our data against the standard normal distribution (115). The correlation measures how well the data may be modeled by the distribution. For the data to be normally distributed, the points shall lie approximately or nearly around on a 45-degree diagonal line. The plots are usually easy to interpret and helpful in identification of outliers.

Skewness and Kurtosis - A histogram with a long tail on the right side, such as the sulphate data above, is said to be skewed to the right; a histogram with a long tail on the left side is said to be skewed to the left. Another way in which data can deviate from the normal distribution is kurtosis.

A histogram that has a high peak in the middle and long tails on either side is leptokurtic; a histogram with a broad, flat middle and short tails is platykurtic.

1.5.1.2 Statistical methods

The normality tests are supplementary to the graphical assessment of normality (115). The main tests for the assessment of normality are Kolmogorov-Smirnov (K-S) test, Lilliefors corrected K-S test, Shapiro-Wilk test, Anderson-Darling test, Cramer-von Mises test, D'Agostino skewness test, Anscombe-Glynn kurtosis test, D'Agostino-Pearson omnibus test, and the Jarque-Bera test (88, 115). These compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation; the null hypothesis is that "sample distribution is normal." If the test is significant, the distribution is non-normal (115).

1.5.2 Correlations

In statistics, a correlation coefficient (r) is a quantitative assessment that measures both the direction and the strength of this tendency to vary together. Pearson's correlation takes all of the data points on the scatter plot and represents them as a single number. The correlations which involve simply the relationship of two variables (for example, X and Y) are termed as zero-order correlations.

The strength and direction of correlation are responsible to establish relationships between two variables. The value of correlation ranges from -1 to 0 or 0 to 1 representing the strength of linear relationship. Positive or negative values represent the upward or downward slope respectively.

1.5.2.1 Beyond Simple Correlations

a) The overall relationship of Y with several predictors X_i taken together can be established by multiple correlation coefficient.

b) The overall relationship of Y with some predictor X_i, after controlling several other predictors X_j, can be answered by partial correlation, also called first-order correlations when controlling for one variable.

c) The overall relationship of Y with several predictors X_i after controlling several other predictors X_j can be computed by multiple partial correlation.

1.5.3 Linear Regression

In addition to quantifying the strength of a linear relationship by computing the Pearson's correlation coefficient, it is insightful to see the relationship with a mathematical model. The most common model is a **simple linear regression**, where we estimate the Y-intercept and the slope of a line that will "best fit" the data The variable we are predicting is called the *criterion or response variable* and is referred to as Y. The variable we are basing our predictions on is called the *predictor variable* and is referred to as X. In simple linear regression, the topic of this section, the predictions of Y when plotted as a function of X form a straight line. The coefficient of determination is given by $\mathbf{R^2}$ (also called r^2 in case of two variable relationship) and it measures as to how much variation of Y is explained by X. The most common method to find the best fit model for regression is the Least Squares method.

1.5.3.1 Multiple Regression

When there are more than one predictor variables, the prediction method is called *multiple regression*. The coefficient of multiple determination ($\mathbf{R^2}$) measures how much of Y is explained by all the X's. It is the square of the multiple correlation coefficient (r).

1.5.4 Measures for Significance

1.5.4.1 P-Value

Differential analysis is carried out to identify the genes which show significant difference between groups - in our case, lean and obese individuals. There always exists some abundant difference in the groups because of randomness. However, this difference may or may not be significant. Significance can be identified by the size of this difference which in turn is considered in comparison to variance. The P-value obtained by this analysis, is the measure of likelihood of the data if there was no real difference, the reason as to why a smaller P-value is considered favorable. Usually, a probability of 0.5 is contemplated to be significant. It explains that there is a 5% chance of getting the current data if no real difference existed and therefore, we decide that the difference is abundant and significant enough.

1.5.4.2 False Positives

A true positive justifies a result which is significant. A false positive value is an indicative of a significant difference which is actually insignificant or might not exist. P-value being 0.05 states that there may be a 5% possibility of a wrong decision to be made or there is a 5% chance of getting false positives.

1.5.4.3 The Multiple testing problem and False Discovery Rate

If we have multiple hypotheses to test, one approach is to test them all separately on our data. If we do so, we increase the chances of a false positive on at least one of the tests since while a given Type I error (alpha) value may be appropriate for each individual comparison, it is not for the set of *all* comparisons. Most of the approaches to overcome the problems arising from multiple testing involve creating an adjusted P-value to reduce the chances of a false-positive. Many approaches have been described to correct for multiple comparisons including techniques like Bonferroni Correction, Sidak correction, Holm's Step-Down procedure, Hochberg's Step-Up procedure. The Bonferroni correction is a multiple-comparison correction used when several dependent or independent statistical tests are being performed simultaneously (since while a given alpha value α may be appropriate for each individual comparison, it is not for the set of *all* comparisons). In order to avoid a lot of spurious positives, the alpha value needs to be lowered to account for the number of comparisons being performed.

The traditional techniques do reduce the number of false positives; however, they reduce the total number of actual true discoveries as well. To counteract this, False Discovery Rate is used. It is a more recent technique which controls the number of false positives but only on the already significant results, or true discoveries. This is the reason as to why this approach is less conservative than others.

1.5.4.4 Multicollinearity

Multicollinearity arises when multiple independent variables in a linear regression model are associated with each other. It can lead to biased and imprecise estimates of the regression coefficients and should therefore be avoided. Collinearity between variables can be assessed

initially by reviewing the correlation matrix so that only a subset of the predictor variables that correlate highly is used for model fitting. Further, the Variation Inflation Factor (VIF) can be calculated after fitting the linear regression model to diagnose multicollinearity problems.

1.6 Gene Ontologies

Gene Ontology (GO) terms aid in the differential analysis of the desired genes, thereby signifying the biological processes and molecular functions they may be involved in. The third GO term specifically describes about the cellular component a gene may belong to.

1.6.1 Molecular Function: This GO term is suggested to be used to annotate gene products whole molecular functions are unknown. It describes essentially a process at molecular level which can be carried out by the action of a single macromolecular machine, generally via direct interactions with other entities. The function is an indicative of an activity or action performed by a complex or a gene product. This action is characterized by biochemical activity and role in a larger process/system.

1.6.2 Biological Process: A process explains a specific objective that the organism is genetically programmed to accomplish. It is defined by its outcome or the ending state, e.g., result of cell division is creation of new cells. A biological process is achieved by a sequence of molecular functions performed by specific gene products.

1.6.3 Cellular Component: It describes a location with respect to the cellular compartments and structures, occupied by a macromolecular entity while it performs a molecular function. The location of gene products may be described by gene ontology by means of relationship to cellular structures (e.g., cytoplasmic side of plasma membrane) and complexes (e.g., mitochondrion), or stable larger complexes to which they may be a part of.

1.7 Identification of *F13A1* for this study

In hunt for genes related to obesity, there have been many studies which have been able to explain only an extremely small percentage of variation responsible for phenotype of obesity. The disease is difficult to study because of the multi-variate and complex phenotypes, whereby correlation of gene expression with a phenotype like BMI does not imply causation. Moreover, the system can't be exactly replicated on animal models. Another problem arises as it gets difficult to differentiate transcripts obtained from GWAS, which are related to causal processes from the ones which are not related to the disease state in any way.

Pietiläinen's work is based on exploiting the fact that MZ twins are identical with respect to DNA sequence, due to which the differences in gene expression might correspond to regulatory and/or epigenetic changes produced in reaction to lifestyle and environment. This implies that the genes with significant differential expression in the AT between the discordant lean-obese twin pairs, can be held as reactive for obesity. On the other hand, there would be a bigger gene pool consisting both reactive and causal genes from a larger sample of unrelated individuals, where transcript levels in AT correlate significantly with BMI. Within the overlap between these two gene pools, causative genes would be the ones which correlate strongly with BMI only in unrelated sample.

Apart from identification of *F13A1* as a potential gene for obesity by Pietiläinen and group, the previous work from Kaartinen group has identified a role for FXIII-A in mouse adipocytes and investigated its mechanisms of action in adipogenesis *in vitro* (116) as well as during weight gain *in vivo* in mice (117). It was reported that in 3T3-E1 mouse adipocyte cell line and in *F13A1* deficient mouse embryonic fibroblasts undergoing adipogenesis, FXIII-A enzyme modulates insulin signaling via regulating plasma fibronectin accumulation into extracellular matrix (116). In *F13A1-/-* mice, the absence of FXIII-A in mice does not alter weight gain on a high fat diet, but decreases macrophage infiltration, collagen and fibronectin accumulation, alters adipocyte morphology and improves insulin sensitivity of obese AT (117). This suggests that absence of *F13A1* is beneficial to metabolic health in obesity. Conversely, it could be hypothesized that the presence and/or increase in FXIII-A may have opposite, negative effect on AT health.

2. Hypothesis and aim

We hypothesized that AT *F13A1/F13A1* is associated with weight gain and obesity as emphasized by the published work from Kaartinen group and Pietiläinen group.

The aim of this study was to investigate in detail if *F13A1*/FXIII-A transglutaminase was acquired excess weight in humans. For this we have used a data set from the Finnish MZ twins consisting of clinical parameters and Affymetrix transcriptome data. With data set and information of Heavy-Lean co-twins' differential values) we have examined further the clinical links of *F13A1*/Factor XIII-A in acquired weight gain and potential weight-associated molecular pathways in adipocytes. The twins offer a unique age, gender and genetically controlled setup to study changes in AT upon weight gain. **Figure 5** lays down the general idea about the approach of this study.

Figure 5

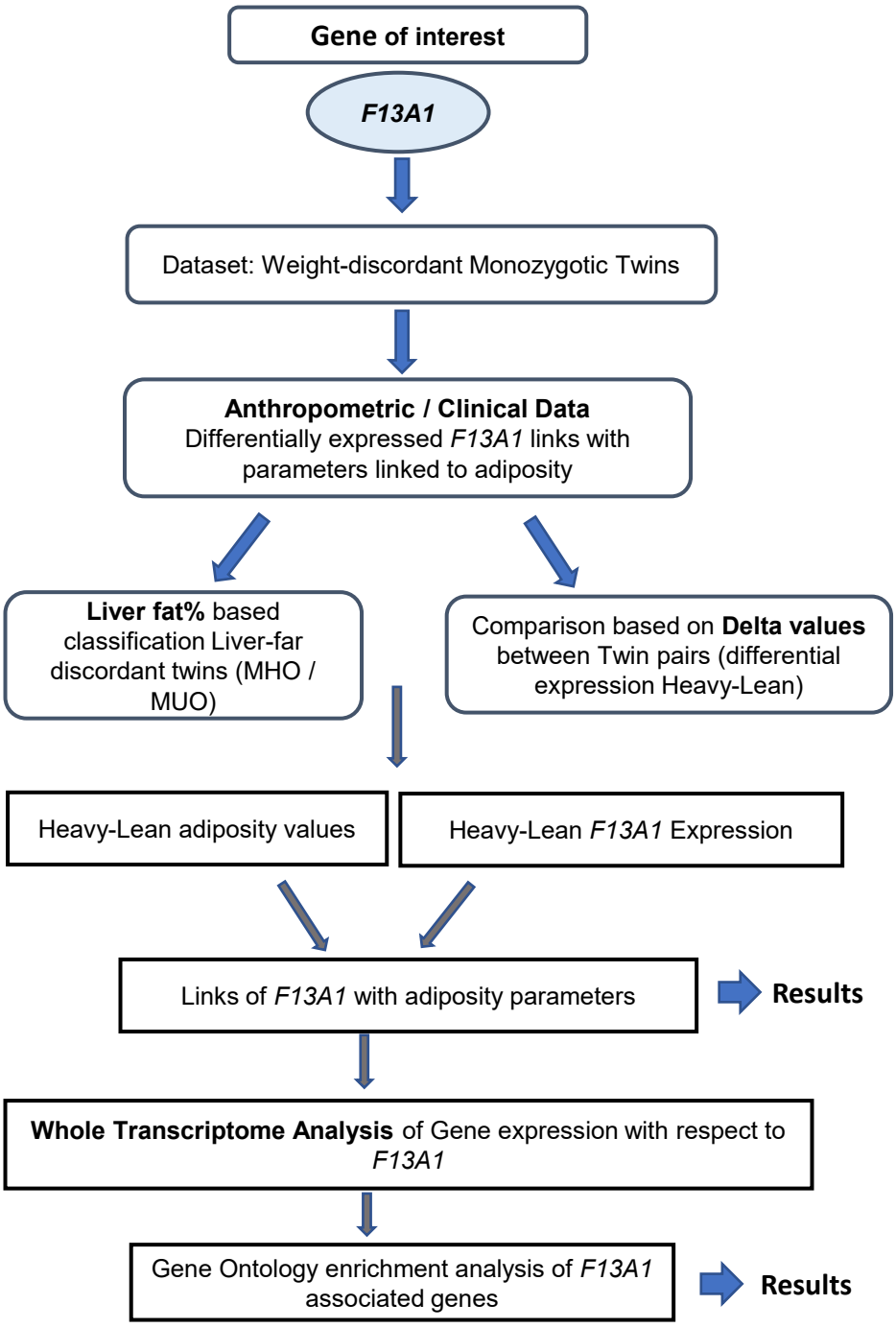

FIGURE 5: General approach of analysis pipeline. Flow chart describing organized layout of order of bioinformatics experiments design and implementation.

3. Materials and methods

3.1 Dataset Description

3.1.1 MZ twin data

The present study included 25 rare, healthy MZ pairs discordant for weight (within-pair difference in acquired excess weight, ΔBMI ≥3 kg/m^2, males n = 9, females n = 16, aged 29.8 ± 0.9 years), identified from two population-based twin cohorts, FinnTwin16 (n = 2839 pairs) and FinnTwin12 (n = 2578 pairs) (118) for whom AT transcriptomics analyses were performed. All twins were healthy, with no concomitant diseases or medications except for contraceptives. The study protocol was originally designed and performed according to the principles of the Helsinki Declaration and approved by the Ethical Committee of the Helsinki University Central Hospital. For our study to explore how genetics of genes is related with obesity, we used Affymetrix mRNA expression data from AT for 22 pairs of BMI discordant MZ twins; and the same from adipocyte data for 12 pairs among them (two men and 10 women aged 27.7 ± 1.4 years). In order to investigate AT expansion, we also utilized the anthropometric data for all these individuals.

3.1.2 Metabolic Parameters

Following were some of the important parameters has been calculated by Pietiläinen's group and focused upon in the current study.

- Body Mass Index (BMI): Body Mass Index (BMI) is a person's weight in kilograms divided by the square of height in meters. A high BMI can be an indicator of high body fatness. BMI can be used to screen for weight categories that may lead to health problems, but it is not diagnostic of the body fatness or health of an individual. In this study, height and weight were measured after an overnight fast, with the subjects barefoot and in underwear.
- Fat Percentage (fatp): This is a measure of the total body fat mass divided by the total mass of the body. It includes both, the essential body fat and storage fat. A higher fat percentage may be an indication of access fat accumulation in the AT. It was calculated using dual-energy x-ray absorptiometry (DXA) scan.
- Fatkg : The absolute weight, in kilograms, of the fat in the body, calculated by DXA.

- Body fat distribution of subcutaneous adipose tissue (SAT) and visceral adipose tissue (VAT) was calculated by magnetic resonance imaging (MRI) and liver fat by proton magnetic resonance spectroscopy.(119)
- Plasma triglyceride concentrations were determined with enzymatic methods
- Plasma leptin and adiponectin were measured by enzyme-linked immunosorbent assay (ELISA) by using DuoSet ELISA (R&D Systems Europe Ltd, Abindgon, UK)
- AT morphology: A small part of the collected AT sample was digested with collagenase and used for the measurement and calculation of adipocyte volume and number via image analysis from images taken with a microscope as previously described (120).

3.2 *F13A1* exploratory analysis

Broadly the project was divided into parts to achieve following aims:

3.2.1 Descriptive Statistics

All the descriptive statistics, mean, standard error, median, standard deviation, variance, kurtosis, skewness, range, minimum, maximum and confidence level at 95%, were calculated for both Lean and Heavy datasets separately in AT and adipocyte fraction, using R Statistical software (121, 122) and validated with MS Excel and Prism GraphPad software (version 8.0.0 for Windows, San Diego, CA, USA). Student's t-test for paired samples was used to check whether the difference in lean and obese datasets is significant. The P-value criteria selected was 0.05 at 95% confidence level. Since the two datasets had different standard deviations (both in AT and adipocyte fraction), it was preferred to use variability as a criterion of comparison instead of mean.

3.2.2 Measures of Normality

Before proceeding with the trivial assumption of normality, we verified whether the data passes the standard normality measures. Normal QQ plot was generated with Prism GraphPad software. Skewness was calculated to get an idea about how symmetrical this dataset is, and Kurtosis was calculated to check for extreme values. Both the parameters were calculated along with other descriptive statistics. Anderson-Darling test, D'Agostino & Pearson test, Shapiro-Wilk test, and

Kolmogorov-Smirnov test were calculated with a null hypothesis that this data is derived from a Normal distribution. The hypothesis cannot be rejected if the P-value is above 0.05.

3.3 Comparisons with metabolic parameters

All the delta values were estimated by the corresponding difference of Lean and Heavy twin metabolic parameters. Further, all the computations for Pearson and Spearman correlations along with respective P-values, were calculated using R statistical software and verified with GraphPad, while plotting the linear graphs. Similar pattern was followed for calculating the Coefficient of Determination and applying linear regression to the data. Correlations and linear regression were calculated for absolute values, for each metabolic parameter, with the absolute values of *F13A1*, separately for Lean and Obese groups. Further, similar computations were done for the delta values as well. Here, the results from delta values only are presented and carried forward for analysis. For nominal P-values, less than 0.05 was considered significant and multiple testing correction was performed by Benjamini-Hochberg method and reported as False discovery rate (FDR).

3.4 Liver-fat concordant and discordant twins

Pietiläinen and group conceptualized the idea of using liver-fat as a classification parameter in their previous work. Replicating the same in this study, we allocate Metabolically healthy and sick groups to each pair of twins, based on computing the values of delta liver-fat percentage. The pairs whose liver fat percentage difference of heavy and lean individual (Δliverfat) is less than 3% are classified in Group 1 (metabolically healthy) and those with Δliverfat more than 3% are classified into Group 2 (metabolically sick). *F13A1* values were compared for both these groups and further delta *F13A1* was computed as well.

3.5 Whole transcriptome wide analysis

The complete dataset of 19,623 genes was screened with Δgene values for expression in adipocyte. Transcriptome wide association (TWA)/correlation screening between intra-pair Δ*F13A1* in adipocytes to full intra-pair Δ*Transcriptome* was carried out in R statistical software and validated in MS-Excel. Significance of difference between the expression value was evaluated using Student's t-test, which was a qualification criterion for investigating differential expression of each gene. 95% confidence was chosen for the same. Those Δ*Genes* expression values showing Pearson

correlation (r) above 0.7 (r^2 above 0.5) to Δ*F13A1* and showing significantly altered expression between the co-twins (pairwise T-test, P-value below 0.05) were considered further (with this stringent criteria highest intra-pair expression difference P-value for a gene included for further analysis was 0.006). Linear regression analysis (r^2) for each significant correlation was performed with Prism GraphPad software (version 8.0 for Windows, San Diego, CA, USA).

3.6 Functional characterization of genes based on Gene Ontologies

Gene Ontology tools, PANTHER (geneontology.org) (123-125) and GOnet (tools.dice-database.org/GOnet) (126) were run on the dataset for exploring the gene ontologies, with a focus on Biological processes. GOrilla (cbl-gorilla.cs.technion.ac.il) (127, 128)was also used and further overrepresented GO terms from Biological Processes, Cellular Components and Molecular functions were selected using the criterion $P<.05$ adjusted for a false discovery rate and further analyzed for similar terms. Ontology analysis was performed on the 182 genes selected from first screening. Selected representative genes from each unique biological function were first analyzed for their linear relationships with *F13A1* through simple linear regression. Later, the relationships were also studied while controlling for best F13A1-correlating metabolic parameters, weight, body fat and adipocyte volume. Further, multiple regression was implemented using R Statistical software and SPSS (IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp), to assess how the variation in mRNA expression values of these genes was influenced by the combined effects of variations of F13A1 and aforementioned metabolic parameters. Multiple regression was corrected for multicollinearity using SPSS.

4. Results

4.1 Explore *F13A1* with respect to Anthropometric data

The levels of *F13A1* in mRNA in adipocyte and ATs, vary significantly in Lean and Heavy groups. We first examined if expression of *F13A1* mRNA is increased with excess weight within the weight discordant MZ twin pairs. The Affymetrix mRNA data from both AT and adipocyte fraction show a significant increase in *F13A1* mRNA expression in the heavier co-twins **(Figure 6)**.

4.1.1 Descriptive Statistics

Table 3 and Table 4 depict descriptive statistics calculated each for lean and heavy groups separately for AT and adipocyte *F13A1* mRNA expression.

Checking for difference in datasets

The basic underlying hypothesis was that the Lean and Obese datasets have no difference in means (Null Hypothesis).

Clearly, from the statistics, the average expression (mean) of *F13A1* in the tissue is higher for obese with a broader range and a higher minimum and maximum values. A relatively higher range in obese, explains the group's marginally higher variation, although it is insignificant.

The Student's t-test for paired samples also indicates a P-value lower than 0.0001, which makes us reject the null hypothesis at a confidence level of 95%, that both the lean and obese datasets are similar or indifferent.

4.1.2 Measures of Normality by Graphical Method

The Normal QQ plot **(Figure 7)** for our dataset shows that *F13A1* values in both Lean and Obese groups lie on the central diagonal line, indicating that the data plausibly comes from normal distribution. But graphical methods don't guarantee a distribution in this case since the dataset has a small sample size. So, it is wiser to validate the same with statistical tests as well.

Figure 6

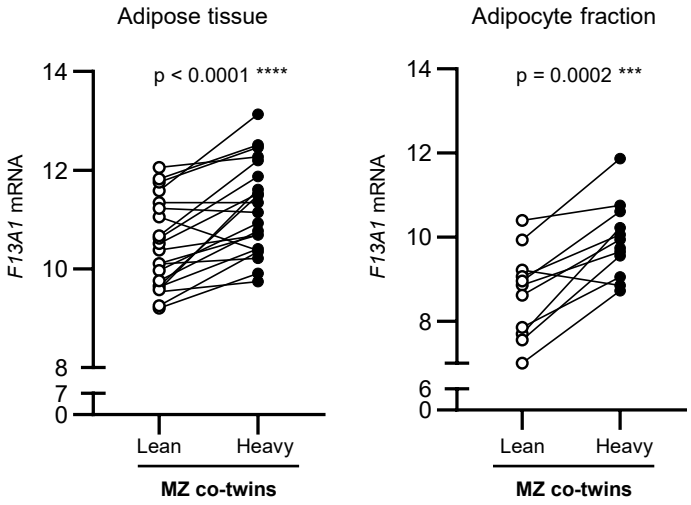

FIGURE 6. *F13A1* expression in adipose tissue and adipocyte fraction of weight-discordant monozygotic (MZ) twin pairs. (A) *F13A1* mRNA levels are significantly increased in the heavier co-twin in adipose tissue (paired T-test p<0.0001) (n=21) and in adipocyte enriched fraction (paired T-test p = 0.0002)(n=12).

Table 3 : Descriptive statistics calculated for AT F13A1 expression values separately in Lean and Heavy MZ co-twins.

Lean		Heavy	
Mean	10.518	**Mean**	11.236
Standard Error	0.193	**Standard Error**	0.195
Median	10.448	**Median**	11.242
Standard Deviation	0.906	**Standard Deviation**	0.917
Sample Variance	0.822	**Sample Variance**	0.842
Kurtosis	1.319	**Kurtosis**	0.702
Skewness	0.214	**Skewness**	0.276
Range	2.858	**Range**	3.392
Minimum	9.202	**Minimum**	9.743
Maximum	12.060	**Maximum**	13.135
Confidence Level (95%)	0.402	**Confidence Level (95%)**	0.407

Table 4 : Descriptive statistics calculated for adipocyte *F13A1* expression values separately in Lean and Heavy co-twins.

Lean		Heavy	
Mean	8.516	**Mean**	9.768
Standard Error	0.299	**Standard Error**	0.279
Median	8.760	**Median**	9.736
Standard Deviation	1.081	**Standard Deviation**	1.009
Sample Variance	1.169	**Sample Variance**	1.018
Kurtosis	-0.636	**Kurtosis**	0.589
Skewness	-0.011	**Skewness**	0.251
Range	3.609	**Range**	3.917
Minimum	6.791	**Minimum**	7.949
Maximum	10.401	**Maximum**	11.866
Confidence Level (95%)	0.653	**Confidence Level (95%)**	0.609

Figure 7

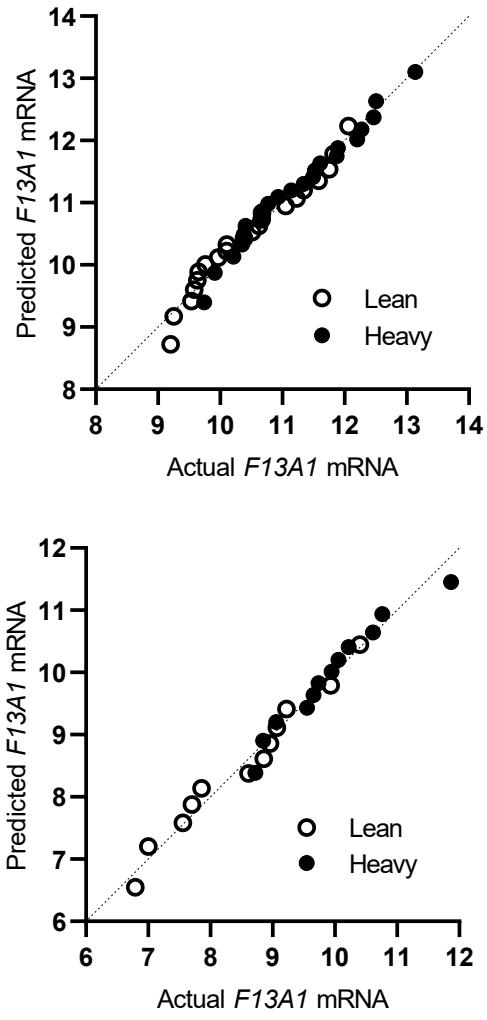

FIGURE 7. QQ plot for A) adipose tissue and B) for adipocyte *F13A1* mRNA expression data. Check for normality. All the points lie on the diagonal line confirming that the data tends towards a Gaussian/normal distribution.

4.1.3 Measures of Normality by Statistical Methods

4.1.3.1 Skewness and Kurtosis

Most of the standard statistical functions presume data to be normal. Graphical methods to judge normality of the data include histogram and normality plots. Since the data used here has a small sample size, it is wise to use the statistical measures of normality. Skewness and Kurtosis are usually considered as one of the standard measures for the same. They also indicate degree of symmetry and presence of outliers, respectively. The acceptable range of values for Skewness and Kurtosis is -1.96 to 1.96 (129) in order to prove univariate normal distribution. The value of skewness shall tend to zero for a symmetric and normal distribution. Here, both for Lean and Heavy samples, we can say that most probably, the data follows univariate normal distribution and that it is approximately symmetric (the values of skewness lying between -0.5 to 0.5).

4.1.3.2 Anderson-Darling test, D'Agostino & Pearson test, Shapiro-Wilk test, and Kolmogorov-Smirnov test

Statistical tests for normality are more precise since the actual probabilities are calculated. For all the tests carried out here, the null hypothesis stated that the sample data are not significantly different than a normal population. All the probabilities above 0.05 indicate normality, i.e., we fail to reject the null hypothesis. **(Table 5 and Table 6)**

Large samples remain indifferent to the violation of normality assumption. So, we can use parametric tests regardless of the data being normally distributed (GraphPad) . However, for smaller samples, one needs to be careful about using parametric or non-parametric tests. Ironically, normality tests have little power to reject the null hypothesis for small samples, which is why these data usually pass normality tests (115). Therefore, we tried running multiple normality tests on this data for validation. Shapiro-Wilk is a fairly powerful test even with asymmetrical data; however, it isn't robust with small samples or discrete data. (115) Also, it works well only if all the values are unique. D'Agostino on contrary is less powerful with asymmetry but robust enough for a small sample size. The Kolmogorov-Smirnov (K-S) test is an exact test which does not depend on any underlying cumulative distribution function (it is distribution free) and works well to compare two samples rather than a sample with a standard distribution. However, it works best with a large sample size. The Anderson-Darling test, a modified version of K-S utilizes a specific distribution in calculating a critical value, allowing a more sensitive testing (115, 130).

Table 5: Normality tests for AT *F13A1* mRNA

Anderson-Darling test	**Lean**	**Heavy**
P value	0.3819	0.78
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns
D'Agostino & Pearson test		
P value	0.242	0.5633
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns
Shapiro-Wilk test		
P value	0.2718	0.7615
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns
Kolmogorov-Smirnov test		
P value	>0.1000	>0.1000
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns

Table 6: Normality tests for adipocyte *F13A1* mRNA

Anderson-Darling test	**Lean**	**Heavy**
P value	0.7697	0.7507
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns
D'Agostino & Pearson test		
P value	0.8021	0.3506
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns
Shapiro-Wilk test		
P value	0.8253	0.681
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns
Kolmogorov-Smirnov test		
P value	>0.1000	>0.1000
Passed normality test (alpha=0.05)?	Yes	Yes
P value summary	ns	ns

Yet, it has its own limitations of multiple calculations for each distribution. Considering that every test is useful in its own unique ways, our data passes all these tests for normality, providing a higher confidence for a normal distribution.

4.1.4 Comparisons with Metabolic parameters

We know, there are no one-size-fits-all concept in correlations. Moreover, human statistics is harder to predict. Absolute correlations between metabolic parameters for Heavy or Lean with *F13A1* do not provide significant results for a causal variable in this case because of the small sample size and absence of a randomized, controlled experiment. Moreover, correlations here do not express the variations in dataset and are highly inconsistent with change of samples. This is because the dataset classifies all individuals into two groups, Lean and Heavy. However, we have pairs where both the twins are Lean, or they have normal BMI, or both the twins are Heavy, or, are overweight. This raw classification is based on the sheer fact that one twin is just heavier than the other. A better way of analysis is by studying the delta values, i.e. the difference between value of each metabolic parameter in Heavy and Lean (Heavy-Lean). This is a much stronger correlation in terms of confidence and a better regression in terms of indicative of variation of the dataset.

Comprehending correlations of *F13A1* expression values in AT and adipocyte fraction with each of the analyzed metabolic parameter **(Table 7)**, we can make following inferences:

1) *F13A1* is increased in AT and adipocyte enriched cellular fraction in acquired excess weight

2) Evidently, it appears that Δweight and Δfatkg is positively correlated with Δ*F13A1* with r^2 as 0.37 and 0.34. Interestingly, the most significant r^2 for Δ*F13A1* is with adipocyte size parameters: Δdiameter (r^2 0.47), Δvolume (r^2 0.48) and Δweight (r^2 0.49) with all P-values in the range of 0.01. On contrary, adipocyte Δcell number has no correlation with Δ*F13A1*. Our results support the hypothesis that these individuals respond mostly via hypertrophy, i.e., increasing adipocyte weight, volume and diameter.

3) R^2 values between Δ*F13A1 and* ΔAdipocyte diameter, volume and weight suggest moderate but significant linear positive correlation. This implies that *F13A1* increases in AT in weight gain and links to adipocyte morphology in a manner that increased *F13A1* in the tissue associate with larger adipocytes in the tissue.

Table 7: Correlation of differential F13A1 mRNA levels in adipose tissue and adipocyte fraction with differential clinical and metabolic parameters reflecting weight and adiposity between heavy and lean co-twins (Δparameter vs Δ*F13A1*).

	Δ*F13A1* | Adipose Tissue (n=21)				**Δ*F13A1* | Adipocyte (n=12)**			
ΔClinical and metabolic parameter	**Pearson (r)**	**R^2**	**p**	**FDR**	**Pearson (r)**	**R^2**	**p**	**FDR**
Weight (kg)	0.212	0.045	0.355	0.558	0.613	0.376	0.034 *	0.187
BMI (kg/m²)	0.151	0.023	0.514	0.628	0.456	0.208	0.136	0.249
Body fat (kg)	0.260	0.067	0.256	0.469	0.589	0.347	0.044 *	0.161
Subcutaneous fat (dm³)	0.376	0.142	0.093	0.512	0.461	0.213	0.131	0.288
Intra-abdominal fat (dm³)	0.270	0.073	0.236	0.519	-0.049	0.002	0.879	0.879
Liver fat (%)	0.071	0.005	0.761	0.761	-0.113	0.013	0.727	0.800
Triglycerides (mmol/l)	0.153	0.023	0.508	0.699	0.431	0.186	0.161	0.253
fP-Leptin (pg/ml)	0.291	0.085	0.201	0.553	0.500	0.250	0.098	0.270
fP-Adiponectin (ng/ml)	-0.505	0.255	0.0195 *	0.215	-0.144	0.021	0.655	0.801
Adipocyte volume (pl)	0.107	0.012	0.644	0.708	0.696	0.485	0.012 *	0.132
Adipocyte number	-0.369	0.136	0.100	0.367	-0.287	0.082	0.365	0.502

4) Significant negative correlation [Pearson (r)] between AT Δ*F13A1* and ΔAdiponectin (ng/ml) was observed. Adipocyte fraction Δ*F13A1* showed significant positive correlation to ΔWeight (kg), ΔBodyFat (kg), ΔAdipocyte volume (pl).

This implies that in AT, large difference in *F13A1* correlated with a decrease in circulating adiponectin levels **(Figure 8A)**.Conversely, a larger difference in the clinical parameter was associated with a larger difference in *F13A1* in adipocyte fraction **(Figure 8B,C,D)**.

4.2 Classification of MZ twins into liver-fat discordant groups - MHO and MUO

Although, majority of obese individuals develop comorbidities, it is estimated that 10-25% of obese humans remain healthy. This metabolically healthy obesity (MHO) has been speculated to arise from protection from development of insulin resistance and altered AT properties (such as tissue expansion modes) of individuals. Interestingly, Pietilainen group reported that the monozygotic twin pairs had clear groups with metabolic differences; MHO-Group 1: insulin sensitive, low liver fat% in obese twin, and metabolically sick obesity MSO-Group 2: insulin resistant, high liver fat% in obese twin **(Figure 9A)**. The classification is such that both the individuals of a twin pair, belong to the same group, unlike the BMI-based classification. Although one group was discordant and the other concordant for liver fat accumulation, the twin pairs were otherwise same with regards to weight and BMI discordance.

We compared the expression of *F13A1* in the two metabolically different groups. These groups reflect the metabolic health of the heavier co-twin. We see an increased *F13A1* expression in the heavier co-twin in both the liver fat concordant and the liver fat discordant group in both AT and adipocyte fraction **(Figure 9B,C)**. However, when examining whether the increase in *F13A1* levels (Δ*F13A1*) is different between the low and high liver fat groups (i.e., Heavy-Lean Δ*F13A1*), it showed no significant difference in either AT or adipocyte fraction (p=0.251). Also, regression analysis shows extremely low Coefficients of Determination and extremely high P-values both for Δ*F13A1* vs ΔBMI and Δ*F13A1* vs ΔWeight. All these facts suggest that while *F13A1* is associated with number of pathways in AT pathology in obesity it does not associate to liver fat accumulation in this data set.

Figure 8

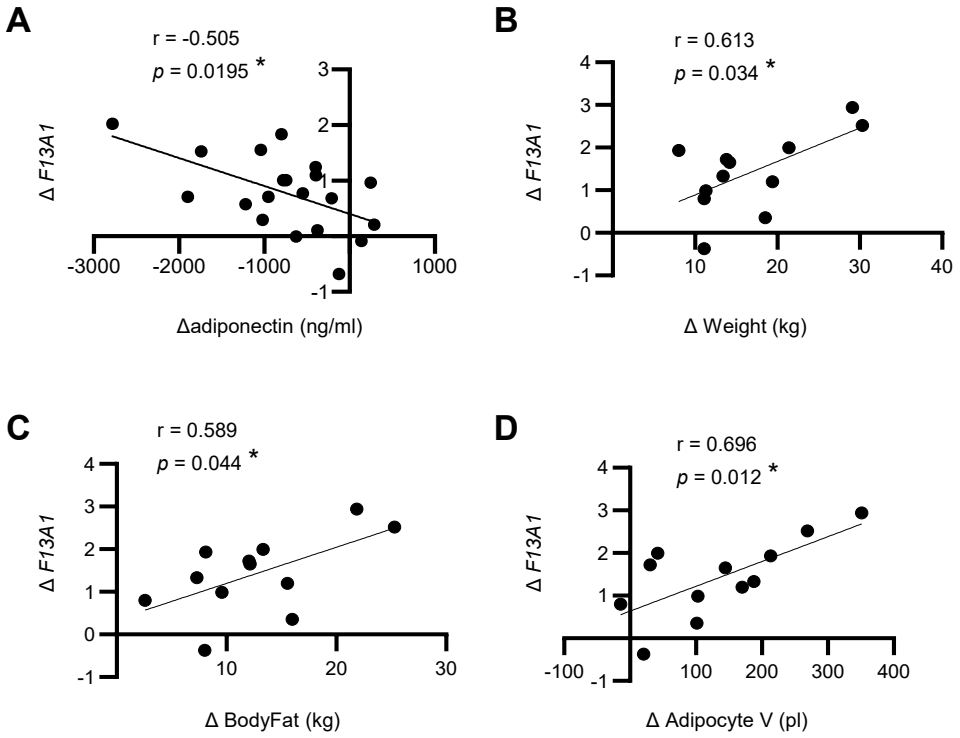

FIGURE 8. Significant correlation (r) of differential (Δ Heavy-Lean) *F13A1* expression in adipose tissue and adipocyte fraction to differential (Δ Heavy-Lean) metabolic parameters relevant to energy metabolism in weight-discordant monozygotic (MZ) twin pairs. (A) Linear, negative correlation of AT Δ*F13A1* with circulating Δadiponectin (n=21). **(B)** Linear, positive correlation of adipocyte fraction Δ*F13A1* with ΔWeight (kg) and **(C)** ΔBodyFat **(D)** Linear, positive correlation adipocyte fraction Δ*F13A1* with ΔAdipocyte volume (n=12).

Figure 9

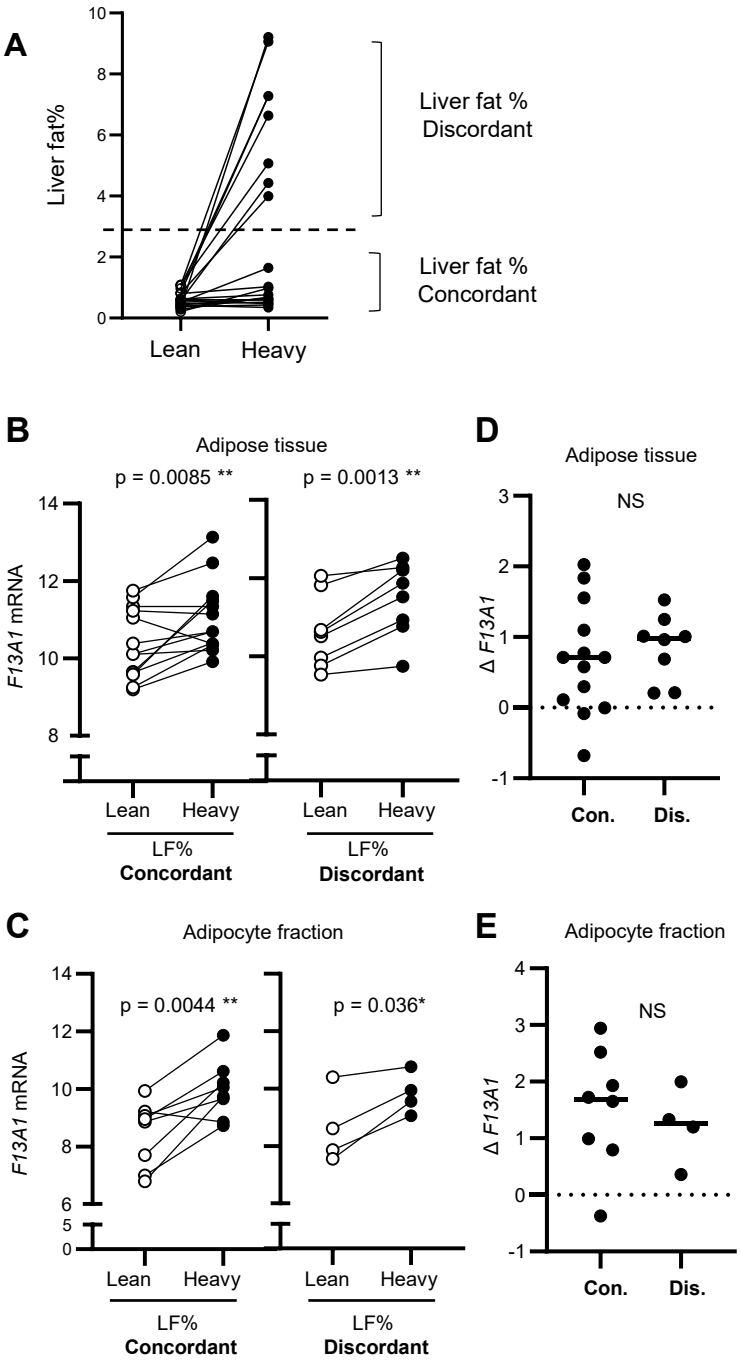

FIGURE 9. (A) Liver fat discordance in the weight-discordant monozygotic (MZ) twin pair used in this study (n=12). Data was previously presented in Naukkarinen et al. (2014) *Diabetologia* 57, 167-176[19] for 21 twin pairs. Discordance was defined by the heavier co-twin having liver fat % >3. ***F13A1* expression in weight-discordant monozygotic (MZ) twin pairs with liver fat % (LF%) discordance (B, C)** *F13A1* mRNA levels are significantly increased in both liver fat discordant and concordant twin pairs, **(D,E)** No significant difference in the differential *F13A1* expression was seen in adipose tissue or adipocyte enriched fraction. Paired T-test was used in comparing increase between the concordant (Con.) and discordant (Dis.) twin pairs. Unpaired T-test was used comparing the difference between the co-twins (Δ*F13A1*) within the liver fat groups

4.3 TWAS for correlations with *F13A1*

In order to examine what AT and weight gain pathways *F13A1* associates with, we performed a full transcriptome wide association study (TWAS) (Heavy-Lean Δ*F13A1* vs. Heavy-Lean Δ*transcriptome*). The screen identified 294 genes whose differential expression showed significant linear correlation with Δ*F13A1*, based on the criteria of Simple Linear Regression, r^2 above 0.50 and P-value within 0.05. This set was further screened to include only genes that also showed significant differential expression between the Heavy and Lean co-twins (*P*<.05, T-test for paired sample). The result was a final set of 181 genes.

4.3.1 Functional characteristics of selected genes

Gene enrichment was analysis performed for the 182 genes (including *F13A1*) using GOnet and Panther gene ontology tools showed links to 318 GOterms with FDR adjusted *P*<.05. A selection of most significant, over-represented GOterms are presented in **Figure 10** and **Table 8** demonstrating a strong link of increase in AT *F13A1* to inflammatory response, response to stress, chemotaxis as well as activation of cells/leukocytes. Also, regulation of extracellular matrix and supramolecular fiber organization, wound healing, angiogenesis and hemostasis were represented. Full list of genes in these over-represented pathways are presented in **Table 9**.

4.3.2 Regression Analysis to explore linear relationships of *F13A1* with selected genes

To illustrate the pathways that *F13A1* is involved in, linear regression graphs of the genes with coefficient of determination (r^2) for each over-represented biological function are presented in **Figure 11**.

Table 10 indicates the following along with the corresponding P-values:

- Unadjusted Pearson correlations which depict simple linear Δ*F13A1* vs ΔGenes relationship, the square of which is Coefficient of determination for Simple Linear Regression.
- Adjusted Pearson Correlations which depict Δ*F13A1* vs ΔGenes relationship, controlled for weight, fatkg, and ACY volume (also called first-order correlations), as well as adjusted combinedly for fatkg and ACY volume (also known as second-order correlations).

- Multiple correlations for Δ*F13A1* vs ΔGenes relationship which are shown with weight, fatkg, and ACY included. The square of this correlation is also called Coefficient of determination, which is the key output for Multiple Linear Regression analysis.

The selected genes that represent significant pathways are: *CYBB* (NADP oxidase, Nox2) that generates reactive oxygen species and regulates oxidative stress; *PTGS2* (cyclooxygenase 2, Cox2) ROS-reactive gene that acts as an initiator of inflammation and fever; *IL10RA* (interleukin 10 receptor subunit α); *IL8* (interleukin 8); *ITGAM* (Integrin αM / CD11b) – a leucocyte/macrophage marker; *PLAU* (urokinase, uPA) a plasmin activator and fibronolytic enzyme; *CTSS* (cathepsin S); *CD44*; and *ANGPT2* (angiopoietin 2).

Figure 10

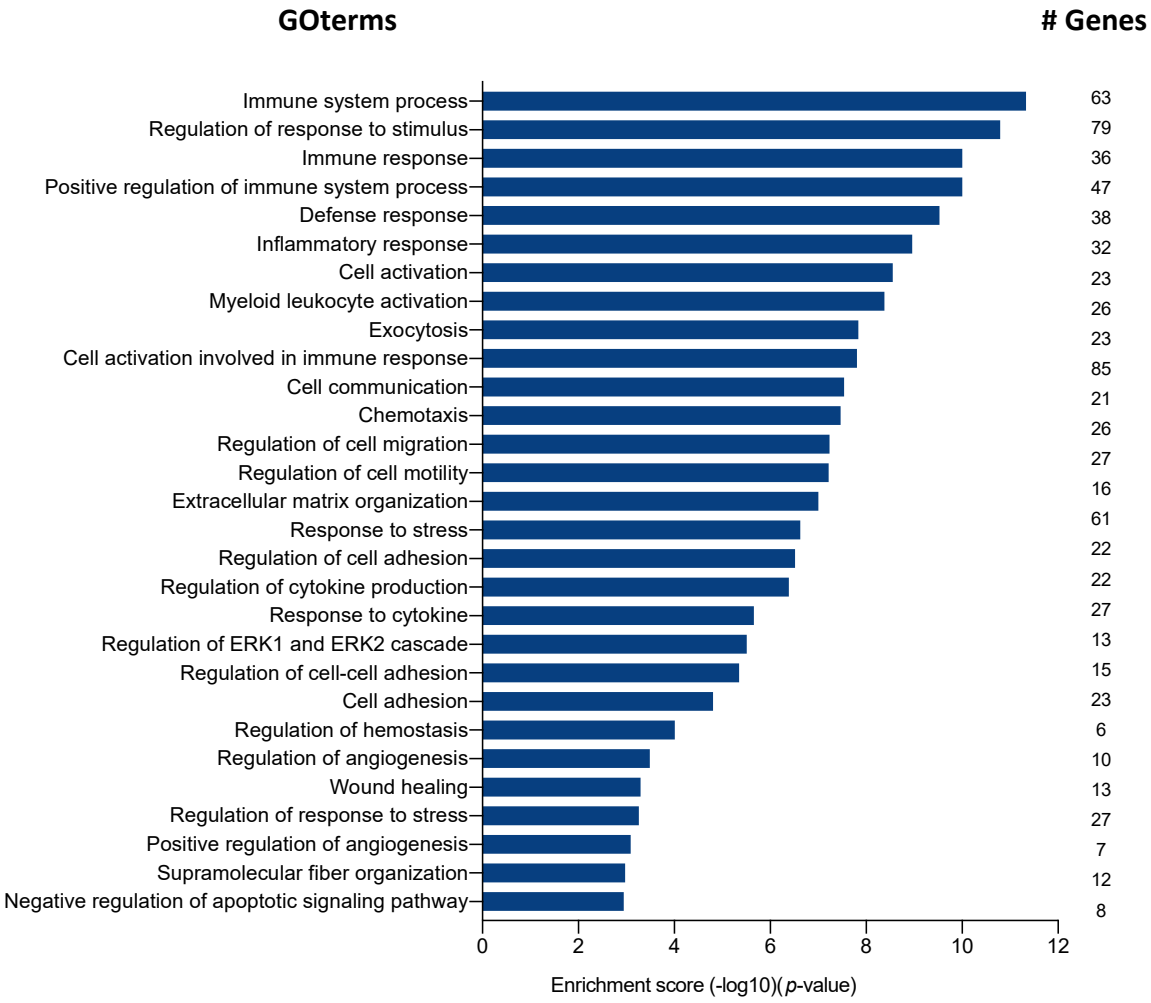

FIGURE 10. Graphical representation of over-represented GOterms with respect to Enrichment score. X-axis and the corresponding bars represent the –log10 of p-value. Number of genes which have the respective GOterm representation are also mentioned.

Table 8. Selected over-represented GOterms among the genes whose differential expression correlated with *F13A1* and is significantly altered between heavy and lean co-twins**.** Total number of genes are 182 including *F13A1.*

Gene ontology / GOterm	ID	p-value	FDR adj p	# genes/182
Immune system process	GO:0002376	4.65E-12	6.40E-09	63
Regulation of response to stimulus	GO:0048583	1.65E-11	6.72E-08	79
Positive regulation of immune system process	GO:0002684	1.00E-10	9.56E-08	36
Immune response	GO:0006955	1.00E-10	1.17E-07	47
Defense response	GO:0006952	3.00E-10	3.04E-07	38
Inflammatory response	GO:0006954	1.10E-09	9.14E-07	22
Cell activation	GO:0001775	2.80E-09	2.23E-06	32
Myeloid leukocyte activation	GO:0002274	4.20E-09	2.92E-06	23
Exocytosis	GO:0006887	1.46E-08	8.67E-06	26
Cell activation involved in immune response	GO:0002263	1.57E-08	8.67E-06	23
Cell communication	GO:0007154	2.89E-08	1.18E-05	85
Chemotaxis	GO:0006935	3.44E-08	1.35E-05	21
Regulation of cell migration	GO:0030334	5.88E-08	1.80E-05	26
Regulation of cell motility	GO:2000145	6.06E-08	1.81E-05	27
Extracellular matrix organization	GO:0030198	1.00E-07	2.61E-05	16
Response to stress	GO:0006950	2.38E-07	5.05E-05	61
Regulation of cell adhesion	GO:0030155	3.05E-07	6.22E-05	22
Regulation of cytokine production	GO:0001817	4.11E-07	7.94E-05	22
Response to cytokine	GO:0034097	2.20E-06	2.82E-04	27
Regulation of ERK1 and ERK2 cascade	GO:0070372	3.10E-06	3.80E-04	13
Regulation of cell-cell adhesion	GO:0022407	4.43E-06	5.10E-04	15
Cell adhesion	GO:0007155	1.55E-05	1.53E-03	23
Regulation of hemostasis	GO:1900046	9.86E-05	6.71E-03	6
Regulation of angiogenesis	GO:0045765	3.26E-04	1.68E-02	10
Wound healing	GO:0042060	5.08E-04	2.27E-02	13
Regulation of response to stress	GO:0080134	5.44E-04	2.40E-02	27
Positive regulation of angiogenesis	GO:0045766	8.21E-04	3.27E-02	7
Supramolecular fiber organization	GO:0097435	1.06E-03	4.01E-02	12
Negative regulation of apoptotic signalling pathway	GO:2001234	1.14E-03	4.28E-02	8

Table 9: Selected over-represented GOterms along with respective functional gene clusters

GOterm ID	GOterm	#Genes /182	Genes
GO:0002376	immune system process	63	ACIN1|ANGPT2|ANPEP|AP1S2|ARPC5|ATP1B1|C1QA|C1QB|C1QC|C1S|C5AR1|C8orf4|CCL2|CCL8|CD44|CD59|CFH|COLEC12|CRIP1|CRISPLD2|CTSS|CYBB|ELF4|FCER1G|FCGR2B|FOLR2|GPR183|HCLS1|HLA-DMB|HLA-DPA1|HLA-DPB1|HLA-DRA|IFI16|IFI30|IL1B|IL8|IQGAP2|ITGAM|LCP1|LCP2|LGMN|LPCAT1|LYZ|MAPK1|MPEG1|PDE4B|PDGFRA|PLAU|PLAUR|PLEK|PMAIP1|PYCARD|RGS1|RNASE1|SELE|SGPL1|SLFN11|SWAP70|TNF|TNFSF13B|TYROBP|VSIG4|WIPF1
GO:0048583	regulation of response to stimulus	79	ABCD1|ANGPT2|ARHGAP18|ARHGDIB|ARPC5|ASPN|BCL2A1|BOK|C1QA|C1QB|C1QC|C1S|C5AR1|C8orf4|CCL2|CCL8|CD44|CD59|CFH|CHSY1|CILP|CLIC2|CNKSR3|COLEC12|CTSS|DUSP2|EZR|FCER1G|FCGR2B|GPNMB|GPR183|HCLS1|HLA-DMB|HLA-DPA1|HLA-DPB1|HLA-DRA|IFI16|IL10RA|IL1B|IL8|ITGAM|LCP2|LGMN|LPAR1|MAPK1|MCTP1|NPTN|NR4A2|PDE4B|PDGFRA|PHLDA2|PLAU|PLAUR|PLEK|PMAIP1|PRRX1|PTGS2|PTPMT1|PTPRF|PYCARD|RASAL3|RECK|RGS1|RGS10|RGS5|RHOJ|SAMSN1|SELE|SMCR8|SOX17|SWAP70|TFPI|TIMP1|TMEM9B|TNF|TNFSF13B|TYROBP|VSIG4|WIPF1
GO:0002684	positive regulation of immune system process	36	ACIN1|ARPC5|C1QA|C1QB|C1QC|C1S|C5AR1|CCL2|CCL8|CFH|COLEC12|CTSS|FCER1G|FCGR2B|GPR183|HCLS1|HLA-DMB|HLA-DPA1|HLA-DPB1|HLA-DRA|IFI16|IL1B|IL8|ITGAM|LCP2|LGMN|MAPK1|PDE4B|PYCARD|RASAL3|SWAP70|TNF|TNFSF13B|TYROBP|VSIG4|WIPF1
GO:0006955	immune response	47	ANPEP|ARPC5|C1QA|C1QB|C1QC|C1S|C5AR1|C8orf4|CCL2|CCL8|CD44|CD59|CFH|COLEC12|CRIP1|CRISPLD2|CTSS|CYBB|ELF4|FCER1G|FCGR2B|GPR183|HLA-DMB|HLA-DPA1|HLA-DPB1|HLA-DRA|IFI16|IFI30|IL1B|IL8|IQGAP2|ITGAM|LCP1|LCP2|LPCAT1|LYZ|MAPK1|MPEG1|PLAU|PLAUR|PYCARD|RGS1|SWAP70|TNF|TNFSF13B|TYROBP|VSIG4
GO:0006952	defense response	38	C1QA|C1QB|C1QC|C1S|C5AR1|CCL2|CCL8|CD163|CD44|CFH|COLEC12|CYBB|ELF4|FCER1G|FCGR2B|FOLR2|HLA-DPA1|HLA-DPB1|HLA-DRA|IFI16|IFI30|IL1B|IL8|ITGAM|LYZ|MPEG1|NFRKB|PMAIP1|PTGS2|PYCARD|RNASE1|SELE|SLFN11|STAT5A|TIMP1|TNF|TYROBP|VSIG4
GO:0006954	inflammatory response	22	C1QA|C5AR1|CCL2|CCL8|CD163|CD44|CYBB|FCER1G|FCGR2B|FOLR2|IFI16|IL1B|IL8|ITGAM|LYZ|NFRKB|PTGS2|PYCARD|SELE|TIMP1|TNF|TYROBP
GO:0001775	cell activation	32	ANPEP|ARPC5|C1QA|C5AR1|C8orf4|CD44|CD59|CRISPLD2|CTSS|CYBB|ELF4|FCER1G|FCGR2B|GPR183|IL1B|IL8|IQGAP2|ITGAM|LCP1|LCP2|LPCAT1|LYZ|MAPK1|PDGFRA|PLAU|PLAUR|PLEK|PYCARD|SWAP70|TIMP1|TNF|TYROBP

GOterm ID	GOterm	#Genes /182	Genes
GO:0002274	myeloid leukocyte activation	23	ANPEP|ARPC5|C1QA|C5AR1|CD44|CD59|CRISPLD2|CTSS|CYBB|FCER1G|FCGR2B|IL8|IQGAP2|ITGAM|LCP2|LPCAT1|LYZ|MAPK1|PLAU|PLAUR|PYCARD|TNF|TYROBP
GO:0006887	exocytosis	26	ANPEP|ARPC5|C5AR1|CCL8|CD44|CD59|CRISPLD2|CTSS|CYBB|F13A1|FCER1G|IQGAP2|ITGAM|LPCAT1|LYZ|MAPK1|PLAU|PLAUR|PLEK|PYCARD|SCRN2|SRGN|TAGLN2|TIMP1|TUSC5|TYROBP
GO:0002263	cell activation involved in immune response	23	ANPEP|ARPC5|C5AR1|C8orf4|CD44|CD59|CRISPLD2|CTSS|CYBB|FCER1G|FCGR2B|GPR183|IQGAP2|ITGAM|LCP1|LPCAT1|LYZ|MAPK1|PLAU|PLAUR|PYCARD|SWAP70|TYROBP
GO:0007154	cell communication	85	ANGPT2|ANXA4|ARHGAP18|ARHGDIB|ARPC5|ASB8|ATP1B1|BCL2A1|BOK|C1QA|C5AR1|CCL2|CCL8|CD44|CD59|CLIC2|COL16A1|COLEC12|CRIP1|CTSS|CYBB|DCLK1|DOCK4|DPYSL2|ERG|EZR|F13A1|FCER1G|FCGR2B|FOLR2|GNB4|GPNMB|GPR183|GUCY1B3|HCLS1|HDGFRP3|HLA-DPA1|HLA-DPB1|HLA-DRA|HPGDS|HSPA13|IFI16|IFI30|IL10RA|IL1B|IL8|IQGAP2|ITGAM|LCP1|LCP2|LGMN|LPAR1|MAPK1|MCTP1|NPTN|NR4A2|NTN4|PDE4B|PDGFRA|PLAU|PLAUR|PLEK|PMAIP1|PTGS2|PTPRF|PYCARD|RASAL3|RECK|RGS1|RGS10|RGS5|RHOJ|RIN2|SELE|SGK1|SGPL1|SOX17|SRGN|STAT5A|TIMP1|TJP1|TNF|TNFSF13B|TYROBP|WIPF1
GO:0006935	chemotaxis	21	C5AR1|CCL2|CCL8|DOCK4|DPYSL2|EGR2|EZR|FCER1G|FOLR2|GPNMB|GPR183|IL1B|IL8|LPAR1|MAPK1|NPTN|NTN4|PDE4B|PDGFRA|PLAU|PLAUR
GO:0030334	regulation of cell migration	26	ANGPT2|ARHGDIB|C5AR1|CCL2|CCL8|DOCK4|GPNMB|GPR183|IL1B|IL8|LGMN|LPAR1|MAPK1|MCTP1|PDGFRA|PHLDA2|PLAU|PTGS2|PYCARD|RECK|RHOJ|SGK1|STAT5A|SWAP70|TIMP1|TNF
GO:2000145	regulation of cell motility	27	ANGPT2|ARHGAP18|ARHGDIB|C5AR1|CCL2|CCL8|DOCK4|GPNMB|GPR183|IL1B|IL8|LGMN|LPAR1|MAPK1|MCTP1|PDGFRA|PHLDA2|PLAU|PTGS2|PYCARD|RECK|RHOJ|SGK1|STAT5A|SWAP70|TIMP1|TNF
GO:0030198	extracellular matrix organization	16	ABI3BP|CD44|COL16A1|CRISPLD2|CTSS|ITGAM|LCP1|LUM|MFAP4|NTN4|PDGFRA|RECK|TGFBI|TIMP1|TNF|VCAN
GO:0006950	response to stress	61	ANGPT2|ATP1B1|BCL2A1|BOK|C1QA|C1QB|C1QC|C1S|C5AR1|C8orf4|CCL2|CCL8|CD163|CD44|CD59|CFH|COLEC12|CRIP1|CYBB|ELF4|F13A1|FAM129A|FCER1G|FCGR2B|FOLR2|GUCY1B3|HLA-DPA1|HLA-DPB1|HLA-DRA|HSPA13|IFI16|IFI30|IL1B|IL8|ITGAM|LCP1|LCP2|LYVE1|LYZ|MAPK1|MPEG1|NFRKB|NR4A2|PDGFRA|PLAU|PLAUR|PLEK|PMAIP1|PTGS2|PTPRF|PYCARD|RNASE1|SELE|SGK1|SLFN11|STAT5A|TFPI|TIMP1|TNF|TYROBP|VSIG4
GO:0030155	regulation of cell adhesion	22	ABI3BP|ANGPT2|ARHGDIB|CCL2|CD44|COL16A1|FCGR2B|GPNMB|HLA-DMB|HLA-DPA1|HLA-DPB1|IL1B|IL8|PIEZO1|PLAU|PYCARD|RASAL3|SWAP70|TGFBI|TNF|TNFSF13B|VSIG4

GOterm ID	GOterm	#Genes /182	Genes
GO:0001817	regulation of cytokine production	22	ABCD1|ANXA4|C5AR1|CYBB|EZR|FCER1G|FCGR2B|GPNMB|HLA-DPA1|HLA-DPB1|IFI16|IL1B|LRRFIP1|LUM|NPTN|PDE4B|PTGS2|PYCARD|SRGN|TNF|TYROBP|VSIG4
GO:0034097	response to cytokine	27	CCL2|CCL8|CD44|F13A1|FCER1G|HCLS1|HLA-DPA1|HLA-DPB1|HLA-DRA|IFI16|IFI30|IL10RA|IL1B|IL8|ITGAM|LAPTM5|LCP1|MAPK1|PTGS2|PYCARD|SELE|SOX17|STAT5A|TFPI|TIMP1|TNF|TNFSF13B
GO:0070372	regulation of ERK1 and ERK2 cascade	13	C5AR1|CCL2|CCL8|CD44|CNKSR3|EZR|GPNMB|GPR183|IL1B|NPTN|PDGFRA|PYCARD|TNF
GO:0022407	regulation of cell-cell adhesion	15	CCL2|CD44|FCGR2B|GPNMB|HLA-DMB|HLA-DPA1|HLA-DPB1|IL1B|PIEZO1|PYCARD|RASAL3|SWAP70|TNF|TNFSF13B|VSIG4
GO:0007155	cell adhesion	23	ATP1B1|CCL2|CD44|COL16A1|EZR|FOLR2|GPNMB|HEPACAM|IL1B|ITGAM|LYVE1|MFAP4|NPTN|NTN4|PDGFRA|PLEK|PTPRF|SELE|TGFBI|THBS2|TNF|TPBG|VCAN
GO:1900046	regulation of hemostasis	6	FCER1G|PDGFRA|PLAU|PLAUR|PLEK|TFPI
GO:0045765	regulation of angiogenesis	10	ANGPT2|C5AR1|CYBB|GPNMB|IL1B|IL8|PTGS2|RECK|RHOJ|THBS2
GO:0042060	wound healing	13	CD44|CD59|F13A1|FCER1G|LCP1|LCP2|MAPK1|PDGFRA|PLAU|PLAUR|PLEK|TFPI|TIMP1
GO:0080134	regulation of response to stress	27	ABCD1|BOK|CD44|COLEC12|CTSS|EZR|FCER1G|FCGR2B|IFI16|IL1B|ITGAM|LGMN|MAPK1|MCTP1|PDGFRA|PLAU|PLAUR|PLEK|PMAIP1|PRRX1|PTGS2|PTPRF|PYCARD|SELE|TFPI|TNF|VSIG4
GO:0045766	positive regulation of angiogenesis	7	ANGPT2|C5AR1|CYBB|IL1B|IL8|PTGS2|RHOJ
GO:0097435	supramolecular fiber organization	12	ARPC5|EZR|HCLS1|HDGFRP3|IQGAP2|LCP1|LUM|MFAP4|PDGFRA|RHOJ|SWAP70|WIPF1
GO:2001234	negative regulation of apoptotic signaling pathway	8	BCL2A1|BOK|CD44|IL1B|NR4A2|PLAUR|PTGS2|TNF

Figure 11

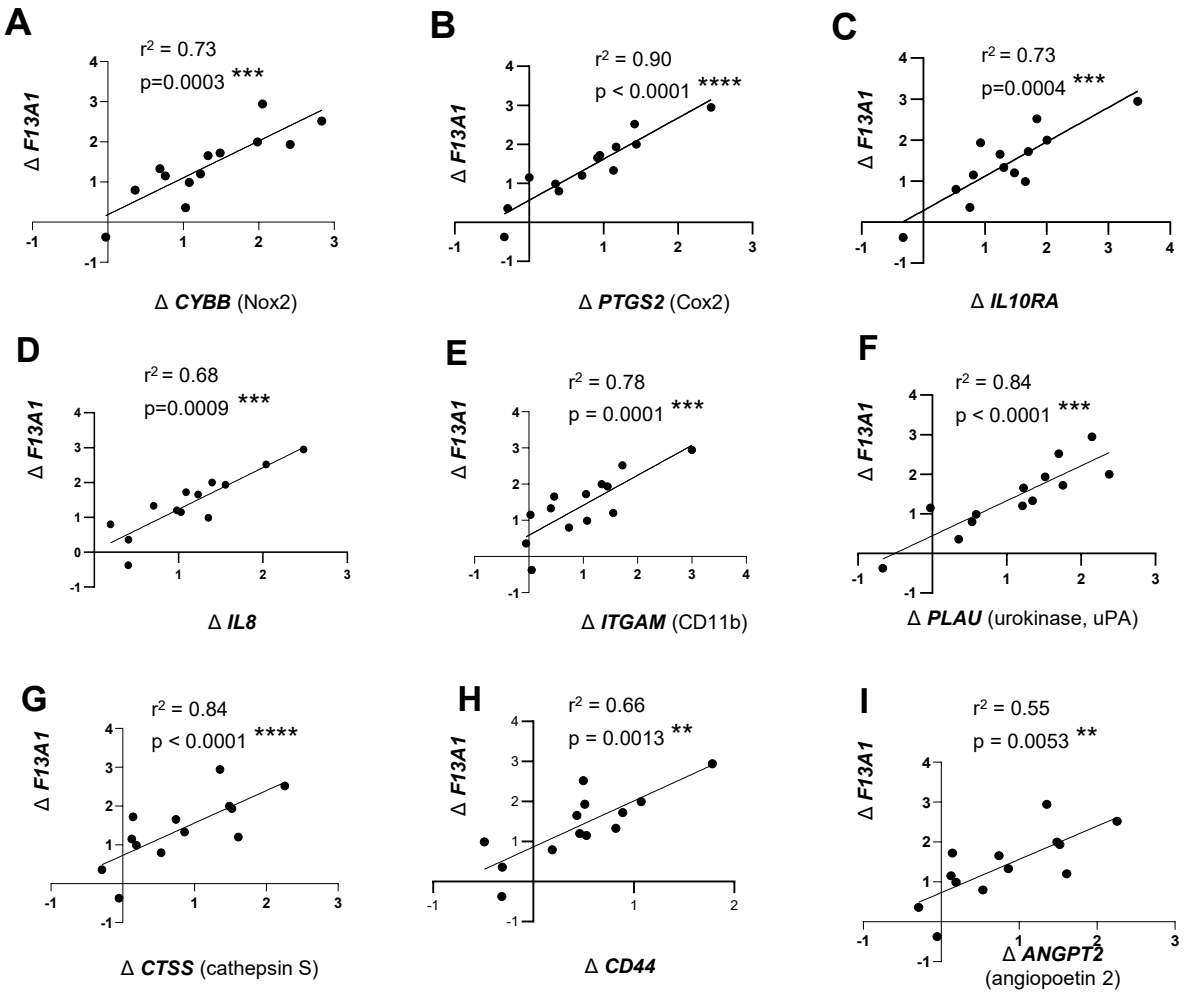

FIGURE 11. Simple linear regression to check for direct relationship of differential *F13A1* expression (Δ) in adipocyte fraction of weight-discordant monozygotic (MZ) twin pairs with selected genes (ΔHeavy-Lean twin) representing significant, over-represented GOterms listed in Table 8.
(A,B) Cellular stress, *CYBB* (Nox) and *PTGS2* (Cox2). **(C,D)** Immune response, *IL10RA* (interleukin 10 receptor a-subunit) and IL8 (interleukin 8). **(E)** leucocyte / macrophage marker *ITGAM* (CD11b). **(F)** Hemostasis/fibrinolysis, *PLAU* (urokinase uPA). **(G)** Extracellular matrix organization, *CTSS* (cathepsin S). **(H)** Wound healing/cell-cell adhesion, CD44. **(I)** Angiogenesis, ANGPT2 (angiopoetin 2) (n=12).

Table 10: Correlations of adipocyte fraction differential F13A1 mRNA levels with differential mRNA levels of representative genes from overexpressed GOterms (ΔGene vs Δ*F13A1*).

	UNADJUSTED (zero-order correlations)	**Simple Linear Regression (r^2/R^2)**	
Gene	**Correlation(r)**	**r^2**	**p-value**
CYBB	0.85	0.73	0.0003
PTGS2	0.95	0.90	<0.0001
IL10RA	0.85	0.73	0.0004
IL8	0.82	0.68	0.0009
ITGAM	0.88	0.78	0.0001
PLAU	0.92	0.84	<0.0001
CTSS	0.91	0.84	<0.0001
CD44	0.81	0.66	0.0013
ANGPT2	0.74	0.55	0.0053

	ADJUSTED Partial Correlations (r) *check for Gene-F13A1 relationship while controlling for other variables*				**Multiple Regression (R^2)** *check for combined influence of F13A1 and other variables on Gene*		
	(first-order correlation)			(second-order correlation)	Gene vs [F13A1 + fatkg + ACY volume]		
Gene	**weight**	**fatkg**	**ACY volume**	**fatkg + ACY volume**	**Multiple Correlation (R)**	**Adjusted R^2**	**p-value**
CYBB	0.78	0.77	0.75	0.74	0.88	0.70	0.0046
PTGS2	0.92	0.94	0.91	0.92	0.96	0.89	<0.0001
IL10RA	0.76	0.77	0.74	0.73	0.87	0.67	0.0068
IL8	0.72	0.73	0.67*	0.65	0.84	0.59	0.015
ITGAM	0.81	0.81	0.76	0.75	0.92	0.80	0.001
PLAU	0.89	0.90	0.93	0.92	0.94	0.86	0.0003
CTSS	0.86	0.87	0.85	0.84	0.95	0.88	0.0001
CD44	0.73	0.78	0.74	0.75	0.82	0.55	0.02
ANGPT2	0.61*	0.64*	0.55*	0.53*	0.76	0.42	0.06

*** Partial correlation significantly lower than Unadjusted correlation**

All calculations are done on delta values of each variable

5. Discussion

In this study we have examined the expression and associated, regulatory pathways of *F13A1* in AT and adipocytes in acquired excess weight. For this we have used clinical and metabolic parameters and transcriptome data of AT and adipocyte fraction from a rare set of weight discordant Finnish MZ twins. This cohort has been widely explored and provides a study model of acquired weight gain where age, sex, genetic and early environmental background is well controlled.

This thesis reports that *F13A1* mRNA expression in AT and adipocyte fraction, significantly increases in heavier twin. Its expression in AT shows a significant, negative association with circulating adiponectin and expression in adipocytes links positively with increase in weight and the total body fat and adipocyte size. Our transcriptome wide association study strongly linked *F13A1* to several relevant metabolic and tissue remodeling pathways of AT expansion that include oxidative stress, extracellular matrix organization and remodeling and angiogenesis. The strongest association of *F13A1* was to pro-inflammatory response of the adipocytes. Inflammation of AT accompanies obesity and has been associated with a greater risk of development of insulin resistance and type 2 diabetes (131-133). While the exact trigger for inflammation in obese AT is yet unknown, it is accepted that adipocyte/oxidative stress, and cell death under overnutrition causes promotion of expression of inflammatory cytokines in adipocytes resulting in macrophage infiltration to the tissue and perpetuation of the proinflammatory status (131-133).

As indicative of *F13A1* linking to AT/adipocyte stress, in our study, differential *F13A1* expression in twins shows a strong and significant linear correlation to increase in *CYBB* (Nox2)(NADPH oxidase 2) **(Figure 11A)** - a mitochondrial enzyme generating reactive oxygen species (ROS) that cause oxidative stress and has been shown to contribute to AT inflammation (134). Increase in *F13A1* also correlated with increase in *PTGS2* (Cox2)(cyclooxygenase-2) **(Figure 11B)** – an enzyme that converts arachidonic acid to prostaglandin H2 – a inflammation mediator. Cox-2 has been both associated with pro-inflammatory response of AT in obesity in rats and humans (135) but also found to be protective of fat accumulation in mice (136). Strong links of *F13A1* were also seen in interleukins and receptors (*IL10RA;* interleukin 10 receptor a subunit)(*IL8;* interleukin 8) **(Figure 11 C,D)**. *IL1B* (interleukin1β) and *TNF* (tumor necrosis factor-α) also showed significant linear association to *F13A1* (coefficient of determination for simple linear regression, r^2=0.62, and r^2=0.51 respectively, for n=13).

Figure 11

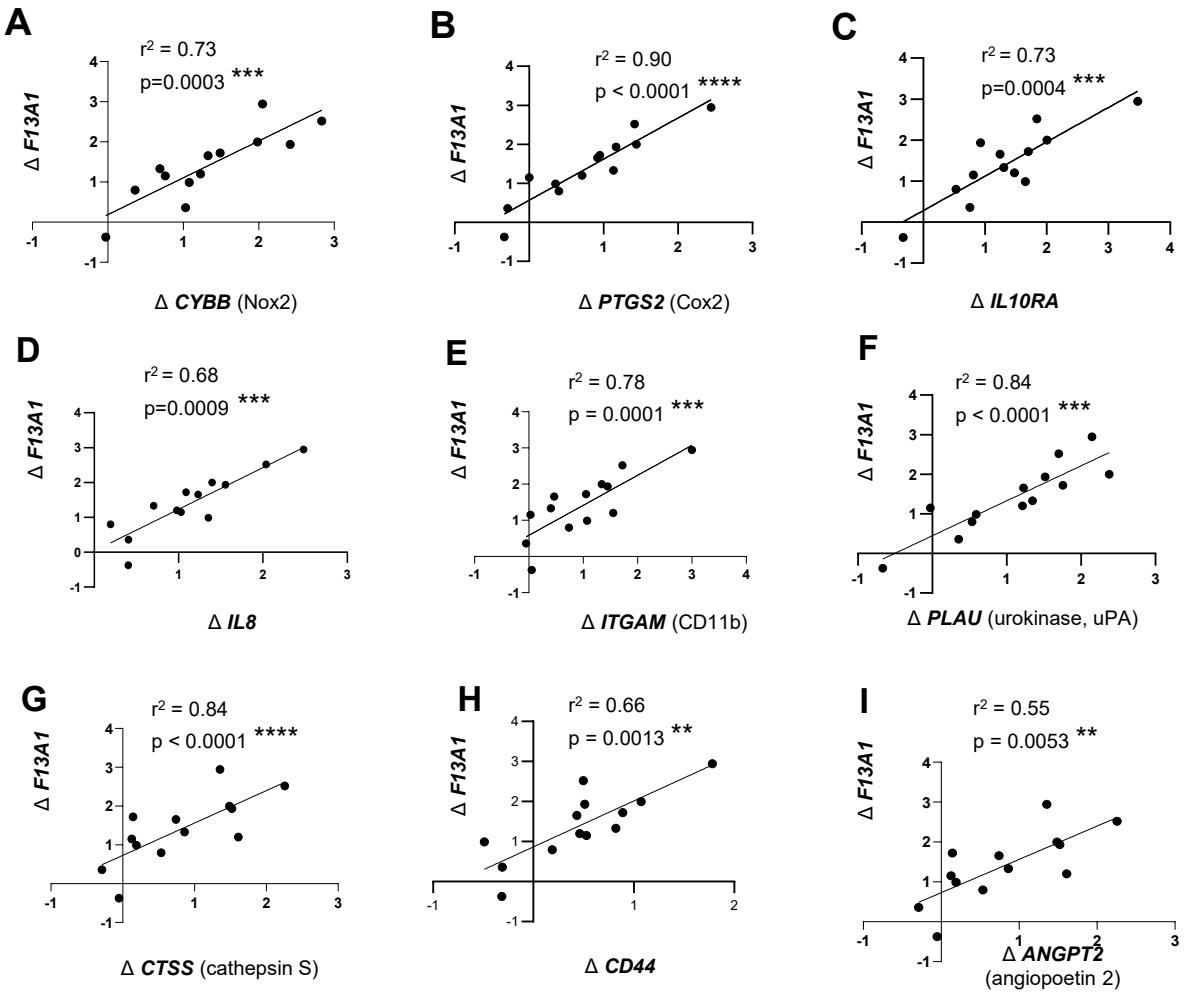

FIGURE 11. Simple linear regression to check for direct relationship of differential *F13A1* expression (Δ) in adipocyte fraction of weight-discordant monozygotic (MZ) twin pairs with selected genes (ΔHeavy-Lean twin) representing significant, over-represented GOterms listed in Table 8.
(A,B) Cellular stress, *CYBB* (Nox) and *PTGS2* (Cox2). **(C,D)** Immune response, *IL10RA* (interleukin 10 receptor a-subunit) and IL8 (interleukin 8). **(E)** leucocyte / macrophage marker *ITGAM* (CD11b). **(F)** Hemostasis/fibrinolysis, *PLAU* (urokinase uPA). **(G)** Extracellular matrix organization, *CTSS* (cathepsin S). **(H)** Wound healing/cell-cell adhesion, CD44. **(I)** Angiogenesis, ANGPT2 (angiopoetin 2) (n=12).

AT inflammation has number of downstream metabolic sequela that leads to defective insulin metabolism. In this study, we observed a negative association of *F13A1* in AT to circulating adiponectin levels. Adiponectin is an adipokine that is capable of maintaining healthy AT expansion, and it has beneficial effect on metabolic homeostasis, and insulin sensitivity resulting in protection from many comorbidities of obesity (137-139). It has been shown that number of proinflammatory cytokines and pathways can lead to decrease in adiponectin production in adipocytes (137), and thus it is possible that the association of AT *F13A1* and adiponectin mediated by the inflammatory status of the AT of the heavier twin. A recent serum proteomics study from this cohort showed elevation of number of inflammatory pathway components in circulation pointing to ongoing systemic inflammation in the heavier co-twin (140). Interestingly, association of increase of *F13A1* to decrease in circulating adiponectin was only observed at tissue level, not for *F13A1* in adipocyte fraction which may suggest that at tissue level there are additional cellular sources of *F13A1* which has negative association to metabolism. This cell may be macrophage, which is one of major cell type producing *F13A1 (114, 141).* The presence and positive correlation of *F13A1* with macrophage marker *ITGAM* (Integrin αM; CD11b) to *F13A1* **(Figure 11 E)** suggests this link, but also indicates presence of macrophages in the adipocyte fraction. The adipocyte fraction used in this study is prepared from AT biopsies via collagenase digestion after which the floating adipocyte layer was collected. This separates adipocytes from the AT stromal vascular fraction that contains many cell types such as mesenchymal stem cells, endothelial and vascular smooth muscle cells and immune cells (142). Also, macrophages are isolated from the stromal vascular fraction (143). While we cannot exclude the possibility that some different cell types are present in the adipocyte fraction, it is fair to assume that vast majority are adipocytes. Full characterization the FXIII-A/*F13A1* expressing cells in mouse AT during weight gain and investigating their specific role in metabolism is the future goal of Kaartinen lab. FXIII-A transglutaminase is best known for its function to stabilize fibrin clot in coagulation cascade, however, its cellular role has become increasingly evident. FXIII-A is expressed by fibroblasts, many mesenchymal and hematopoietic lineage cells, including monocytes and macrophages where its role follows in hemostasis and wound healing, i.e., extracellular matrix stabilization of fibrin and fibronectin and promotion of cell adhesion, proliferation, and migration as well as differentiation. It is also expressed by endothelial cells during angiogenesis. In 3T3-L1 adipocyte cultures FXIII-A activity is found on the surface of adipocytes and linked to regulation of plasma

fibronectin into extracellular matrix which in turn regulates preadipocyte proliferation, insulin signaling, adipocyte cell morphology and differentiation. FXIII-A is induced in preadipocytes immediately upon differentiating cues. It peaks early in preadipocytes and is dramatically downregulated once cells transition to mature adipocytes (116). In mice, absence of *F13A1* shows similar phenotype as predicted from *in vitro* work; insulin sensitivity is increased, macrophage infiltration to AT appears decreased, and AT morphology altered with increase in very large adipocytes as well as smaller cells (117).

In our study, FXIII-A expression also links strongly and positively to adipocyte volume but not to number, suggesting that link to hypertrophic mode of expansion. Widely collected evidence supports the concept that hypertrophic tissue expansion associates strongly with metabolic dysfunction in a manner that larger sized adipocytes are more "pathological" than small ones (38, 43, 120, 144, 145).The increased *F13A1* expression in the heavier twin also opens the interesting concept that FXIII-A produced by adipocytes might contribute to hypercoagulatory state in obesity. However, previous assessment of coagulatory and fibronolytic status of the weight-discordant MZ twin pairs showed no alterations is circulating FXIII levels (146) which strongly suggests that the increase in AT *F13A1*/FXIII-A does not affect the pool of plasma FXIII-A. However, the increased expression of FXIII-A in AT may contribute to local accumulation/stabilization of fibrin(ogen). Levels of circulating fibrin(ogen), and other procoagulant factors are increased in obesity and shown also to be significantly elevated in the heavier twin of the same Finnish MZ twin pairs. Kopec et al. recently demonstrated that fibrin(ogen) deposits accumulate to obese AT in humans and high fat diet (HFD) fed mice and contribute to weight gain, systemic as well as AT inflammation and metabolic dysfunction via mechanisms that involve it's leucocyte integrin αMβ2 binding motif which also binds to FXIII-A49, and associates in this study to *F13A1* expression **(Figure 11E)**.

In our study, within the 182 FXIII-A co-regulated genes in adipocyte fraction, six constituted the hemostasis GOterm **(Table 9).** These included *PLAU* (urokinase)**(Figure 11F)**, *PLAUR* (urokinase receptor) which are part of fibrinolytic cascade leading to plasmin activation and fibrin degradation. However, while these were also significantly increased in adipocyte fraction of the heavier co-twins in our study, it was reported earlier that circulating D-dimer, a marker of fibrin degradation, was not elevated in the heavier MZ co-twins. This again suggests that

fibrinolytic/thrombotic events likely remain local in AT and leakage of thrombotic/fibrinolytic factors from AT to circulation is likely not occurring. In fact, it is possible that the AT procoagulant/fibrinolytic factors that associate to *F13A1* are part of the angiogenic response to acquired weight gain. Increase in *THBS2* (thrombospondin 2) and *ANGPT2* (angiopoietin 2) **(Figure 11I)** between the co-twins showed strong and significant linear correlation to increase in *F13A1* expression. Angiopoietin can stimulate plasmin production and thrombospondin is among the plasmin substrates, that are part of adipocyte microenvironment and remodelled upon tissue expansion and neoangiogenesis (147, 148).

GOterms for extracellular matrix organization and supramolecular fiber organization also suggests links to tissue remodeling. Genes in this group constitutes 16 of 182 **(Table 9)** and included *bonafide* ECM proteins genes such as *COL16A*, proteoglycans, *LUM* (lumican) and *VCAN* (versican) as well as *TIMP1* (tissue inhibitor of MMPs 1). While there is no information on how COL XVI may associate to weight gain or metabolism, lumican which is a FXIII-A substrate in human plasma (149), was recently demonstrated to increase in AT of human obesity and in AT of HFD-fed mice, to positively modulate insulin sensitivity, and glucose metabolism. HFD-fed *Lum*-/- mice had reduced fat mass, and decreased insulin sensitivity and over-expressor mice showed improved insulin sensitivity and glucose clearance (150). Versican is able to modulate inflammation via multiple mechanisms which includes stimulation of secretion of inflammatory cytokines through direct binding to cellular receptors such as CD44 - which was also coregulated with *F13A1* **(Figure 11H)**. It can also directly bind and protect CCL2 protein from degradation, which similarly co-regulated gene with *F13A1*. It is interesting to note that both extracellular matrix assembly and degradation is represented in *F13A1* associate genes. In addition to the above-mentioned plasmin activators, proteolytic enzyme CTSS (cathepsin S) showed strong and significant linear association to *F13A1* **(Figure 11G)**. CTSS is upregulated in AT and circulation of obese individuals (151) and appears to be downstream to pro-inflammatory cytokines. CTSS inhibition improved glucose metabolism via suppressing hepatic glucose production (152) but did not affected liver fat accumulation. Analysis of differential expression of *F13A1* in liver fat concordant and discordant MZ twins showed no differences suggesting that also *F13A1* is not associated with fatty liver complications in obesity. This data is also supported by the observations from *F13A1*-deficient mice that had equal liver fat content compared to control mice on HFD (117).

Partial and multiple correlations were computed to check the influence of selected *F13A1*-associated metabolic parameters, on *F13A1*'s relationship with the genes discussed above. For *IL8* and *ANGPT2*, the partial correlations with *F13A1*, while controlling for adipocyte volume, were significantly lesser than the simple unadjusted correlations (**Table 10**). The result suggests chances of a potential relationship of these genes with the cell volume. A recent study supports this hypothesis for *ANGPT2*, whereby tissue-specific overexpression model of ANG-2 in double transgenic mice shows healthy AT expansion and small adipocyte size. However, more statistical evidence is required (153). Another study suggests that adipocyte size is an important determinant of adipokine secretion, such as IL-8. (154). All other representative genes from each unique GOterm correlate with *F13A1* without the influence of associated metabolic parameters.

6. Conclusions

Our findings bring further evidence to association of AT and adipocyte *F13A1* in acquired weight gain and show new evidence of the pathways that it associates with. A summary of the analytical pipeline implemented is shown in **Figure 12**. Future studies will address the triggers and mechanisms that lead to *F13A1* increase and what the downstream effects of its upregulation exactly are for AT inflammation and metabolic homeostasis. Understanding this sequela will be critical in evaluating if FXIII-A in AT could be a druggable target to achieve metabolic control in acquired weight gain.

Figure 12

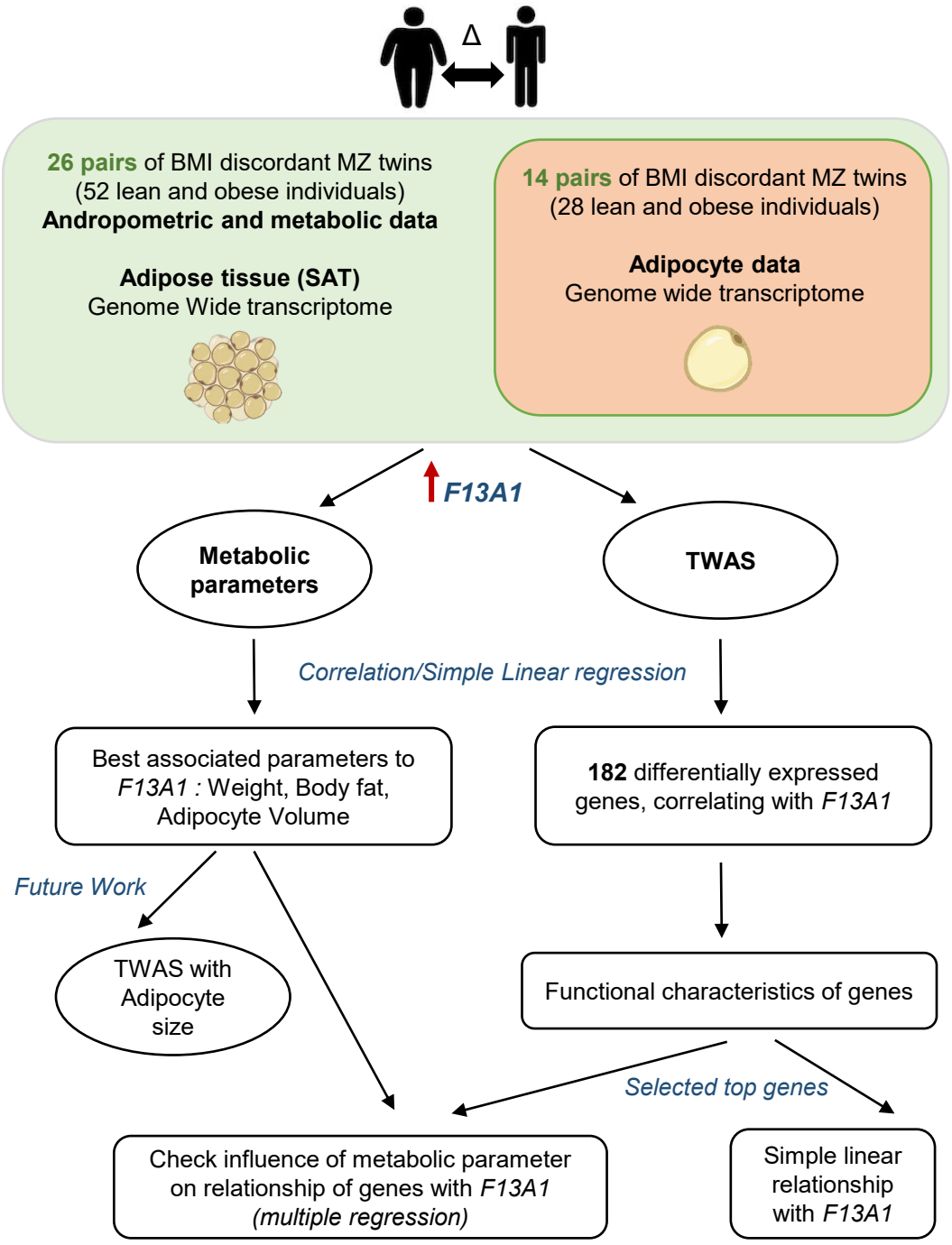

FIGURE 12: Summary. This TWAS provides association of AT and adipocyte *F13A1* in acquired weight gain and shows new evidence of the pathways that it associates with. The strongest association of *F13A1* was to pro-inflammatory response of the adipocytes. Inflammation of AT accompanies obesity and has been associated with a greater risk of development of insulin resistance and type 2 diabetes.

7. Future Directions

While our study brings robust new evidence to the association of *F13A1* to acquired weight gain in humans, the data here is limited to transcriptional level. Future studies should focus on analyzing FXIII-A enzyme at protein level and activity in AT, as well as confirming the cell type that in adipocyte fraction is mainly responsible for its production. It is also highly relevant to explore the temporal expression and activity of *F13A1*/FXIII-A during weight gain to answer if it is upregulated early or late and if it precedes the initiation of inflammation or if its expression is a result of inflammation. These are studies that can be conducted in mice on high fat diet.

Also, it is interesting to investigate further for relationship and dependencies of *F13A1* on adipocyte size. Is it a product of large adipocytes, or does it regulate the size? We conducted a second screening of the dataset with focus on transcripts linked to adipocyte size. This data was published recently (155). The study follows the same approach as here with the TWAS and GOterm analysis and has same threshold criteria as in the former screening.

8. References

1. Loos RJ, Yeo GS. The bigger picture of FTO: the first GWAS-identified obesity gene. Nature reviews Endocrinology. 2014;10(1):51-61.

2. Singla P, Bardoloi A, Parkash AA. Metabolic effects of obesity: A review. World Journal of Diabetes. 2010;1(3):76-88.

3. Dolton P, Xiao M. The intergenerational transmission of body mass index across countries. Economics and Human Biology. 2017;24:140-52.

4. Birch LL, Fisher JO. Development of eating behaviors among children and adolescents. Pediatrics. 1998;101(3 Pt 2):539-49.

5. Johnson SL, Birch LL. Parents' and children's adiposity and eating style. Pediatrics. 1994;94(5):653-61.

6. Galgani J, Ravussin E. Energy metabolism, fuel selection and body weight regulation. International Journal of Obesity (2005). 2008;32 Suppl 7:S109-19.

7. Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. Circulation. 2012;126(1):126-32.

8. Webber J. Energy balance in obesity. The Proceedings of the Nutrition Society. 2003;62(2):539-43.

9. Prentice AM. Macronutrients as sources of food energy. Public Health Nutrition. 2005;8(7a):932-9.

10. Maughan R. Carbohydrate metabolism. Surgery (Oxford). 2009;27:6 - 10.

11. Rosenbaum M, Leibel RL. Adaptive thermogenesis in humans. Int J Obes (Lond). 2010;34 Suppl 1:S47-55.

12. Sabounchi NS, Rahmandad H, Ammerman A. Best-fitting prediction equations for basal metabolic rate: informing obesity interventions in diverse populations. Int J Obes (Lond). 2013;37(10):1364-70.

13. Westerterp KR. Control of energy expenditure in humans. Eur J Clin Nutr. 2017;71(3):340-4.

14. Singh RK, Kumar P, Mahalingam K. Molecular genetics of human obesity: A comprehensive review. Comptes rendus biologies. 2017;340(2):87-108.

15. Roder PV, Wu B, Liu Y, Han W. Pancreatic regulation of glucose homeostasis. Exp Mol Med. 2016;48:e219.
16. Qaid MM, Abdelrahman MM. Role of insulin and other related hormones in energy metabolism -

A review. Cogent Food & Agriculture. 2016;2(1).
17. Yang Q, Vijayakumar A, Kahn BB. Metabolites as regulators of insulin sensitivity and metabolism. Nat Rev Mol Cell Biol. 2018;19(10):654-72.
18. Kojta I, Chacinska M, Blachnio-Zabielska A. Obesity, Bioactive Lipids, and Adipose Tissue Inflammation in Insulin Resistance. Nutrients. 2020;12(5).
19. Armani A, Mammi C, Marzolla V, Calanchini M, Antelmi A, Rosano GMC, et al. Cellular models for understanding adipogenesis, adipose dysfunction, and obesity. Journal of Cellular Biochemistry. 2010;110(3):564-72.
20. Lee M-J, Wu Y, Fried SK. Adipose tissue remodeling in pathophysiology of obesity. Current Opinion in Clinical Nutrition and Metabolic Care. 2010;13(4):371-6.
21. Schäffler A, Schölmerich J, Büchler C. Mechanisms of disease: adipocytokines and visceral adipose tissue--emerging role in intestinal and mesenteric diseases. Nature Clinical Practice Gastroenterology & Hepatology. 2005;2(2):103-11.
22. Wronska A, Kmiec Z. Structural and biochemical characteristics of various white adipose tissue depots. Acta Physiologica (Oxford, England). 2012;205(2):194-208.
23. Cinti S. Transdifferentiation properties of adipocytes in the adipose organ. American Journal of Physiology Endocrinology and Metabolism. 2009;297(5):E977-86.
24. Tordjman J. <2013_Book_PhysiologyAndPhysiopathologyOf.pdf>: Springer-Verlag Paris; 2013.
25. Divoux A, Clément K. Architecture and the extracellular matrix: the still unappreciated components of the adipose tissue. Obesity Reviews: An Official Journal of the International Association for the Study of Obesity. 2011;12(5):e494-503.
26. Mariman ECM, Wang P. Adipocyte extracellular matrix composition, dynamics and role in obesity. Cellular and molecular life sciences: CMLS. 2010;67(8):1277-92.
27. Fujimoto WY, Bergstrom RW, Boyko EJ, Chen KW, Kahn SE, Leonetti DL, et al. Preventing diabetes--applying pathophysiological and epidemiological evidence. The British Journal of Nutrition. 2000;84 Suppl 2:S173-6.

28. Montague CT, O'Rahilly S. The perils of portliness: causes and consequences of visceral adiposity. Diabetes. 2000;49(6):883-8.
29. Berry DC, Stenesen D, Zeve D, Graff JM. The developmental origins of adipose tissue. Development (Cambridge, England). 2013;140(19):3939-49.
30. Tchkonia T, Lenburg M, Thomou T, Giorgadze N, Frampton G, Pirtskhalava T, et al. Identification of depot-specific human fat cell progenitors through distinct expression profiles and developmental gene patterns. American Journal of Physiology Endocrinology and Metabolism. 2007;292(1):E298-307.
31. Yamamoto Y, Gesta S, Lee KY, Tran TT, Saadatirad P, Kahn CR. Adipose depots possess unique developmental gene signatures. Obesity (Silver Spring, Md). 2010;18(5):872-8.
32. Cannon B, Nedergaard J. Brown adipose tissue: function and physiological significance. Physiological Reviews. 2004;84(1):277-359.
33. Hassan M, Latif N, Yacoub M. Adipose tissue: friend or foe? Nature Reviews Cardiology. 2012;9(12):689-702.
34. Justesen J, Stenderup K, Ebbesen EN, Mosekilde L, Steiniche T, Kassem M. Adipocyte tissue volume in bone marrow is increased with aging and in patients with osteoporosis. Biogerontology. 2001;2(3):165-71.
35. Di Iorgi N, Mo AO, Grimm K, Wren TAL, Dorey F, Gilsanz V. Bone acquisition in healthy young females is reciprocally related to marrow adiposity. J Clin Endocrinol Metab. 2010;95(6):2977-82.
36. Muir LA, Neeley CK, Meyer KA, Baker NA, Brosius AM, Washabaugh AR, et al. Adipose tissue fibrosis, hypertrophy, and hyperplasia: Correlations with diabetes in human obesity. Obesity (Silver Spring, Md). 2016;24(3):597-605.
37. Divoux A, Tordjman J, Lacasa D, Veyrie N, Hugol D, Aissat A, et al. Fibrosis in human adipose tissue: composition, distribution, and link with lipid metabolism and fat mass loss. Diabetes. 2010;59(11):2817-25.
38. Arner E, Westermark PO, Spalding KL, Britton T, Rydén M, Frisén J, et al. Adipocyte turnover: relevance to human adipose tissue morphology. Diabetes. 2010;59(1):105-9.
39. Cotillard A, Poitou C, Torcivia A, Bouillot J-L, Dietrich A, Klöting N, et al. Adipocyte size threshold matters: link with risk of type 2 diabetes and improved insulin resistance after gastric bypass. J Clin Endocrinol Metab. 2014;99(8):E1466-70.

40. Eriksson-Hogling D, Andersson DP, Bäckdahl J, Hoffstedt J, Rössner S, Thorell A, et al. Adipose tissue morphology predicts improved insulin sensitivity following moderate or pronounced weight loss. International Journal of Obesity (2005). 2015;39(6):893-8.

41. Landgraf K, Rockstroh D, Wagner IV, Weise S, Tauscher R, Schwartze JT, et al. Evidence of early alterations in adipose tissue biology and function and its association with obesity-related inflammation and insulin resistance in children. Diabetes. 2015;64(4):1249-61.

42. Rydén M, Andersson DP, Bergström IB, Arner P. Adipose tissue and metabolic alterations: regional differences in fat cell size and number matter, but differently: a cross-sectional study. J Clin Endocrinol Metab. 2014;99(10):E1870-6.

43. Veilleux A, Caron-Jobin M, Noël S, Laberge PY, Tchernof A. Visceral adipocyte hypertrophy is associated with dyslipidemia independent of body composition and fat distribution in women. Diabetes. 2011;60(5):1504-11.

44. Andrei AM, Berbecaru-Iovan A, Din-, Anghel FRI, Stănciulescu CE, Berbecaru-Iovan S, et al. Interplay between Hypoxia, Inflammation and Adipocyte Remodeling in the Metabolic Syndrome. 2017 2017-02-01 [cited 2020-08-11 04:07:29]. In: Hypoxia and Human Diseases [Internet]. IntechOpen, [cited 2020-08-11 04:07:29]. Available from: https://api.intechopen.com/chapter/pdf-download/52447.pdf.

45. Arner P, Spalding KL. Fat cell turnover in humans. Biochem Biophys Res Commun. 2010;396(1):101-4.

46. de Ferranti S, Mozaffarian D. The perfect storm: obesity, adipocyte dysfunction, and metabolic consequences. Clin Chem. 2008;54(6):945-55.

47. Sun K, Kusminski CM, Scherer PE. Adipose tissue remodeling and obesity. The Journal of Clinical Investigation. 2011;121(6):2094-101.

48. Xu H, Barnes GT, Yang Q, Tan G, Yang D, Chou CJ, et al. Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. The Journal of Clinical Investigation. 2003;112(12):1821-30.

49. Ghaben AL, Scherer PE. Adipogenesis and metabolic health. Nat Rev Mol Cell Biol. 2019;20(4):242-58.

50. Vishvanath L, Gupta RK. Contribution of adipogenesis to healthy adipose tissue expansion in obesity. J Clin Invest. 2019;129(10):4022-31.

51. Pellegrinelli V, Carobbio S, Vidal-Puig A. Adipose tissue plasticity: how fat depots respond differently to pathophysiological cues. Diabetologia. 2016;59(6):1075-88.
52. Morley TS, Xia JY, Scherer PE. Selective enhancement of insulin sensitivity in the mature adipocyte is sufficient for systemic metabolic improvements. Nat Commun. 2015;6:7906.
53. Senol-Cosar O, Flach RJ, DiStefano M, Chawla A, Nicoloro S, Straubhaar J, et al. Tenomodulin promotes human adipocyte differentiation and beneficial visceral adipose tissue expansion. Nat Commun. 2016;7:10686.
54. Kusminski CM, Holland WL, Sun K, Park J, Spurgin SB, Lin Y, et al. MitoNEET-driven alterations in adipocyte mitochondrial activity reveal a crucial adaptive process that preserves insulin sensitivity in obesity. Nat Med. 2012;18(10):1539-49.
55. Cuthbertson DJ, Steele T, Wilding JP, Halford JC, Harrold JA, Hamer M, et al. What have human experimental overfeeding studies taught us about adipose tissue expansion and susceptibility to obesity and metabolic complications? Int J Obes (Lond). 2017;41(6):853-65.
56. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature. 2006;444(7121):840-6.
57. Roberts CK, Hevener AL, Barnard RJ. Metabolic syndrome and insulin resistance: underlying causes and modification by exercise training. Compr Physiol. 2013;3(1):1-58.
58. Stenvers DJ, Scheer F, Schrauwen P, la Fleur SE, Kalsbeek A. Circadian clocks and insulin resistance. Nature reviews Endocrinology. 2019;15(2):75-89.
59. Reaven GM. Banting lecture 1988. Role of insulin resistance in human disease. Diabetes. 1988;37(12):1595-607.
60. Moore JB. Non-alcoholic fatty liver disease: the hepatic consequence of obesity and the metabolic syndrome. Proc Nutr Soc. 2010;69(2):211-20.
61. Zhao L. The gut microbiota and obesity: from correlation to causality. Nat Rev Microbiol. 2013;11(9):639-47.
62. Saad MJ, Santos A, Prada PO. Linking Gut Microbiota and Inflammation to Obesity and Insulin Resistance. Physiology (Bethesda). 2016;31(4):283-93.
63. Ye J. Mechanisms of insulin resistance in obesity. Front Med. 2013;7(1):14-24.
64. Stumvoll M. Control of glycaemia: from molecules to men. Minkowski Lecture 2003. Diabetologia. 2004;47(5):770-81.

65. Cantley J, Ashcroft FM. Q&A: insulin secretion and type 2 diabetes: why do beta-cells fail? BMC Biol. 2015;13:33.

66. Buchanan TA, Xiang AH. Gestational diabetes mellitus. The Journal of clinical investigation. 2005;115(3):485-91.

67. Moller N, Nair KS. Diabetes and protein metabolism. Diabetes. 2008;57(1):3-4.

68. Pereira S, Marliss EB, Morais JA, Chevalier S, Gougeon R. Insulin resistance of protein metabolism in type 2 diabetes. Diabetes. 2008;57(1):56-63.

69. Arneth B, Arneth R, Shams M. Metabolomics of Type 1 and Type 2 Diabetes. Int J Mol Sci. 2019;20(10):2467.

70. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab. 2009;9(4):311-26.

71. Lucio M, Fekete A, Weigert C, WΓ¤gele B, Zhao X, Chen J, et al. Insulin sensitivity is reflected by characteristic metabolic fingerprints--a Fourier transform mass spectrometric non-targeted metabolomics approach. PLoS One. 2010;5(10):e13317-e.

72. Gall WE, Beebe K, Lawton KA, Adam K-P, Mitchell MW, Nakhle PJ, et al. alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. PLoS One. 2010;5(5):e10883-e.

73. Guasch-Ferre M, Hruby A, Toledo Ea, Clish CB, MartΓ-nez-GonzΓΎlez MA, Salas-SalvadΓi J, et al. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. Diabetes Care. 2016;39(5):833-46.

74. Ho JE, Larson MG, Vasan RS, Ghorbani A, Cheng S, Rhee EP, et al. Metabolite profiles during oral glucose challenge. Diabetes. 2013;62(8):2689-98.

75. Hotamisligil GS, Arner P, Caro JF, Atkinson RL, Spiegelman BM. Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance. The Journal of Clinical Investigation. 1995;95(5):2409-15.

76. Hotamisligil GS, Shargill NS, Spiegelman BM. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. Science (New York, NY). 1993;259(5091):87-91.

77. Rask-Andersen M, Karlsson T, Ek WE, Johansson Å. Genome-wide association study of body fat distribution identifies adiposity loci and sex-specific genetic effects. Nature Communications. 2019;10(1):339.
78. van Vliet-Ostaptchouk JV, Nuotio M-L, Slagter SN, Doiron D, Fischer K, Foco L, et al. The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. BMC endocrine disorders. 2014;14:9.
79. Magkos F. Metabolically healthy obesity: what's in a name? The American Journal of Clinical Nutrition. 2019;110(3):533-9.
80. Rey-López JP, de Rezende LF, Pastor-Valero M, Tess BH. The prevalence of metabolically healthy obesity: a systematic review and critical evaluation of the definitions used. Obesity Reviews: An Official Journal of the International Association for the Study of Obesity. 2014;15(10):781-90.
81. Eckel N, Meidtner K, Kalle-Uhlmann T, Stefan N, Schulze MB. Metabolically healthy obesity and cardiovascular events: A systematic review and meta-analysis. European Journal of Preventive Cardiology. 2016;23(9):956-66.
82. Blüher M. Metabolically Healthy Obesity. Endocrine Reviews. 2020;41(3):405-20.
83. Gallagher D, Heymsfield SB, Heo M, Jebb SA, Murgatroyd PR, Sakamoto Y. Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index. The American Journal of Clinical Nutrition. 2000;72(3):694-701.
84. Flegal KM, Graubard BI. Estimates of excess deaths associated with body mass index and other anthropometric variables. Am J Clin Nutr. 2009;89(4):1213-9.
85. Sun Q, van Dam RM, Spiegelman D, Heymsfield SB, Willett WC, Hu FB. Comparison of dual-energy x-ray absorptiometric and anthropometric measures of adiposity in relation to adiposity-related biologic factors. Am J Epidemiol. 2010;172(12):1442-54.
86. Czernichow S, Kengne AP, Huxley RR, Batty GD, de Galan B, Grobbee D, et al. Comparison of waist-to-hip ratio and other obesity indices as predictors of cardiovascular disease risk in people with type-2 diabetes: a prospective cohort study from ADVANCE. Eur J Cardiovasc Prev Rehabil. 2011;18(2):312-9.
87. Stutzmann F, Tan K, Vatin V, Dina C, Jouret B, Tichet J, et al. Prevalence of melanocortin-4 receptor deficiency in Europeans and their age-dependent penetrance in multigenerational pedigrees. Diabetes. 2008;57(9):2511-8.

88. Vaisse C, Clement K, Guy-Grand B, Froguel P. A frameshift mutation in human MC4R is associated with a dominant form of obesity. Nature Genetics. 1998;20(2):113-4.
89. Yeo GS, Farooqi IS, Aminian S, Halsall DJ, Stanhope RG, O'Rahilly S. A frameshift mutation in MC4R associated with dominantly inherited human obesity. Nature Genetics. 1998;20(2):111-2.
90. Loos RJF, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nature Genetics. 2008;40(6):768-75.
91. Choquet H, Meyre D. Genetics of Obesity: What have we Learned? Current Genomics. 2011;12(3):169-79.
92. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science (New York, NY). 2007;316(5826):889-94.
93. Dina C, Meyre D, Gallina S, Durand E, Körner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nature Genetics. 2007;39(6):724-6.
94. Scuteri A, Sanna S, Chen W-M, Uda M, Albai G, Strait J, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS genetics. 2007;3(7):e115.
95. Naukkarinen J, Rissanen A, Kaprio J, Pietiläinen KH. Causes and consequences of obesity: the contribution of recent twin studies. International Journal of Obesity (2005). 2012;36(8):1017-24.
96. Naukkarinen J, Surakka I, Pietiläinen KH, Rissanen A, Salomaa V, Ripatti S, et al. Use of genome-wide expression data to mine the "Gray Zone" of GWA studies leads to novel candidate obesity genes. PLoS genetics. 2010;6(6):e1000976.
97. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? International Journal of Epidemiology. 2003;32(1):1-22.
98. Pietiläinen KH, Rissanen A, Laamanen M, Lindholm A-K, Markkula H, Yki-Järvinen H, et al. Growth patterns in young adult monozygotic twin pairs discordant and concordant for obesity. Twin Research: The Official Journal of the International Society for Twin Studies. 2004;7(5):421-9.

99. Pietiläinen KH, Rissanen A, Kaprio J, Mäkimattila S, Häkkinen A-M, Westerbacka J, et al. Acquired obesity is associated with increased liver fat, intra-abdominal fat, and insulin resistance in young adult monozygotic twins. American Journal of Physiology Endocrinology and Metabolism. 2005;288(4):E768-74.

100. Duan H, Pang Z, Zhang D, Li S, Kruse TA, Kyvik KO, et al. Genetic and environmental dissections of sub-phenotypes of metabolic syndrome in the Chinese population: a twin-based heritability study. Obesity Facts. 2011;4(2):99-104.

101. Benyamin B, Sørensen TIA, Schousboe K, Fenger M, Visscher PM, Kyvik KO. Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? Diabetologia. 2007;50(9):1880-8.

102. Bogl LH, Kaye SM, Rämö JT, Kangas AJ, Soininen P, Hakkarainen A, et al. Abdominal obesity and circulating metabolites: A twin study approach. Metabolism: Clinical and Experimental. 2016;65(3):111-21.

103. Silventoinen K, Pietiläinen KH, Tynelius P, Sørensen TIA, Kaprio J, Rasmussen F. Genetic and environmental factors in relative weight from birth to age 18: the Swedish young male twins study. International Journal of Obesity (2005). 2007;31(4):615-21.

104. Silventoinen K, Bartels M, Posthuma D, Estourgie-van Burk GF, Willemsen G, van Beijsterveldt TCEM, et al. Genetic regulation of growth in height and weight from 3 to 12 years of age: a longitudinal study of Dutch twin children. Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies. 2007;10(2):354-63.

105. Haworth CMA, Carnell S, Meaburn EL, Davis OSP, Plomin R, Wardle J. Increasing heritability of BMI and stronger associations with the FTO gene over childhood. Obesity (Silver Spring, Md). 2008;16(12):2663-8.

106. Hur Y-M, Shin J-S. Effects of chorion type on genetic and environmental influences on height, weight, and body mass index in South Korean young twins. Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies. 2008;11(1):63-9.

107. Faith MS, Pietrobelli A, Nuñez C, Heo M, Heymsfield SB, Allison DB. Evidence for independent genetic influences on fat mass and body mass index in a pediatric twin sample. Pediatrics. 1999;104(1 Pt 1):61-7.

108. Cornes BK, Zhu G, Martin NG. Sex differences in genetic variation in weight: a longitudinal study of body mass index in adolescent twins. Behavior Genetics. 2007;37(5):648-60.
109. Hur Y-M. Sex difference in heritability of BMI in South Korean adolescent twins. Obesity (Silver Spring, Md). 2007;15(12):2908-11.
110. Pietiläinen KH, Kaprio J, Rissanen A, Winter T, Rimpelä A, Viken RJ, et al. Distribution and heritability of BMI in Finnish adolescents aged 16y and 17y: a study of 4884 twins and 2509 singletons. International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity. 1999;23(2):107-15.
111. Lajunen H-R, Kaprio J, Keski-Rahkonen A, Rose RJ, Pulkkinen L, Rissanen A, et al. Genetic and environmental effects on body mass index during adolescence: a prospective study among Finnish twins. International Journal of Obesity (2005). 2009;33(5):559-67.
112. Silventoinen K, Rokholm B, Kaprio J, Sørensen TIA. The genetic and environmental influences on childhood obesity: a systematic review of twin and adoption studies. International Journal of Obesity (2005). 2010;34(1):29-40.
113. Ariëns RA, Philippou H, Nagaswami C, Weisel JW, Lane DA, Grant PJ. The factor XIII V34L polymorphism accelerates thrombin activation of factor XIII and affects cross-linked fibrin structure. Blood. 2000;96(3):988-95.
114. Muszbek L, Bereczky Z, Bagoly Z, Komáromi I, Katona É. Factor XIII: a coagulation factor with multiple plasmatic and cellular functions. Physiological Reviews. 2011;91(3):931-72.
115. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. International Journal of Endocrinology and Metabolism. 2012;10(2):486-9.
116. Myneni VD, Hitomi K, Kaartinen MT. Factor XIII-A transglutaminase acts as a switch between preadipocyte proliferation and differentiation. Blood. 2014;124(8):1344-53.
117. Myneni VD, Mousa A, Kaartinen MT. Factor XIII-A transglutaminase deficient mice show signs of metabolically healthy obesity on high fat diet. Scientific Reports. 2016;6:35574.
118. Kaprio J. Twin Studies in Finland 2006. Twin Research and Human Genetics. 2012;9(6):772-7.
119. Granér M, Seppälä-Lindroos A, Rissanen A, Hakkarainen A, Lundbom N, Kaprio J, et al. Epicardial fat, cardiac dimensions, and low-grade inflammation in young adult monozygotic twins discordant for obesity. The American Journal of Cardiology. 2012;109(9):1295-302.

120. Heinonen S, Saarinen L, Naukkarinen J, Rodriguez A, Fruhbeck G, Hakkarainen A, et al. Adipocyte morphology and implications for metabolic derangements in acquired obesity. Int J Obes (Lond). 2014;38(11):1423-31.
121. Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2018.
122. Team RS. RStudio: Integrated Development Environment for R. RStudio, PBC; 2020.
123. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47(D1):D419-D26.
124. Gene Ontology C. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021;49(D1):D325-D34.
125. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9.
126. Pomaznoy M, Ha B, Peters B. GOnet: a tool for interactive Gene Ontology analysis. BMC Bioinformatics. 2018;19(1):470.
127. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.
128. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol. 2007;3(3):e39.
129. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod. 2013;38(1):52-4.
130. Oztuna D, Elhan AH, Tuccar E. Investigation of Four Different Normality Tests in Terms of Type 1 Error Rate and Power under Different Distributions. Turkish Journal of Medical Sciences. 2006;36(3):6.
131. Reilly SM, Saltiel AR. Adapting to obesity with adipose tissue inflammation. Nature reviews Endocrinology. 2017;13(11):633-43.
132. Stolarczyk E. Adipose tissue inflammation in obesity: a metabolic or immune response? Current Opinion in Pharmacology. 2017;37:35-40.
133. Lee B-C, Lee J. Cellular and molecular players in adipose tissue inflammation in the development of obesity-induced insulin resistance. Biochimica Et Biophysica Acta. 2014;1842(3):446-62.

134. Han CY. Roles of Reactive Oxygen Species on Insulin Resistance in Adipose Tissue. Diabetes & Metabolism Journal. 2016;40(4):272-9.
135. Chan P-C, Hsiao F-C, Chang H-M, Wabitsch M, Hsieh PS. Importance of adipocyte cyclooxygenase-2 and prostaglandin E2-prostaglandin E receptor 3 signaling in the development of obesity-induced adipose tissue inflammation and insulin resistance. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2016;30(6):2282-97.
136. Banhos Danneskiold-Samsøe N, Sonne SB, Larsen JM, Hansen AN, Fjære E, Isidor MS, et al. Overexpression of cyclooxygenase-2 in adipocytes reduces fat accumulation in inguinal white adipose tissue and hepatic steatosis in high-fat fed mice. Scientific Reports. 2019;9(1):8979.
137. Burhans MS, Hagman DK, Kuzma JN, Schmidt KA, Kratz M. Contribution of Adipose Tissue Inflammation to the Development of Type 2 Diabetes Mellitus. Comprehensive Physiology. 2018;9(1):1-58.
138. Ruan H, Dong LQ. Adiponectin signaling and function in insulin target tissues. Journal of Molecular Cell Biology. 2016;8(2):101-9.
139. Stern JH, Rutkowski JM, Scherer PE. Adiponectin, Leptin, and Fatty Acids in the Maintenance of Metabolic Homeostasis through Adipose Tissue Crosstalk. Cell Metab. 2016;23(5):770-84.
140. Sahebekhtiari N, Saraswat M, Joenväärä S, Jokinen R, Lovric A, Kaye S, et al. Plasma Proteomics Analysis Reveals Dysregulation of Complement Proteins and Inflammation in Acquired Obesity—A Study on Rare BMI-Discordant Monozygotic Twin Pairs. PROTEOMICS – Clinical Applications. 2019;13(4):1800173.
141. Mitchell JL, Mutch NJ. Let's cross-link: diverse functions of the promiscuous cellular transglutaminase factor XIII-A. J Thromb Haemost. 2019;17(1):19-30.
142. Church CD, Berry R, Rodeheffer MS. Isolation and study of adipocyte precursors. Methods in Enzymology. 2014;537:31-46.
143. Orr JS, Kennedy AJ, Hasty AH. Isolation of adipose tissue immune cells. Journal of Visualized Experiments: JoVE. 2013(75):e50707.
144. Spalding KL, Arner E, Westermark PO, Bernard S, Buchholz BA, Bergmann O, et al. Dynamics of fat cell turnover in humans. Nature. 2008;453(7196):783-7.

145. Tandon P, Wafer R, Minchin JEN. Adipose morphology and metabolic disease. J Exp Biol. 2018;221(Pt Suppl 1).
146. Kaye SM, Pietilainen KH, Kotronen A, Joutsi-Korhonen L, Kaprio J, Yki-Jarvinen H, et al. Obesity-related derangements of coagulation and fibrinolysis: a study of obesity-discordant monozygotic twin pairs. Obesity (Silver Spring). 2012;20(1):88-94.
147. Christiaens V, Scroyen I, Lijnen HR. Role of proteolysis in development of murine adipose tissue. Thromb Haemost. 2008;99(2):290-4.
148. Van Hul M, Frederix L, Lijnen HR. Role of thrombospondin-2 in murine adipose tissue angiogenesis and development. Obesity (Silver Spring). 2012;20(9):1757-62.
149. Nikolajsen CL, Dyrlund TF, Poulsen ET, Enghild JJ, Scavenius C. Coagulation factor XIIIa substrates in human plasma: identification and incorporation into the clot. J Biol Chem. 2014;289(10):6526-34.
150. Wolff G, Taranko AE, Meln I, Weinmann J, Sijmonsma T, Lerch S, et al. Diet-dependent function of the extracellular matrix proteoglycan Lumican in obesity and glucose homeostasis. Mol Metab. 2019;19:97-106.
151. Taleb S, Lacasa D, Bastard J-P, Poitou C, Cancello R, Pelloux V, et al. Cathepsin S, a novel biomarker of adiposity: relevance to atherogenesis. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2005;19(11):1540-2.
152. Lafarge JC, Pini M, Pelloux V, Orasanu G, Hartmann G, Venteclef N, et al. Cathepsin S inhibition lowers blood glucose levels in mice. Diabetologia. 2014;57(8):1674-83.
153. An YA, Sun K, Joffin N, Zhang F, Deng Y, Donze O, et al. Angiopoietin-2 in white adipose tissue improves metabolic homeostasis through enhanced angiogenesis. Elife. 2017;6.
154. Skurk T, Alberti-Huber C, Herder C, Hauner H. Relationship between adipocyte size and adipokine expression and secretion. J Clin Endocrinol Metab. 2007;92(3):1023-33.
155. Kaartinen MT, Arora M, Heinonen S, Rissanen A, Kaprio J, Pietilainen KH. Transglutaminases and Obesity in Humans: Association of F13A1 to Adipocyte Hypertrophy and Adipose Tissue Immune Response. Int J Mol Sci. 2020;21(21).