Metabolic signatures in nutrition and health : short-term diet response, sexual dimorphism and hormone chronobiology

Draper, C.F.

Chapter 4

Sexual dimorphism, age and fat mass are key phenotypic drivers of proteomic signatures

Based on

Colleen Fogarty Draper, Aoife M Curran, Marie-Pier Scott-Boyer, Armand Valsesia, Helen M Roche, Miriam F Ryan, Michael J Gibney, Martina Kutmon, Chris T Evelo, Susan L Coort, Arne Astrup, Wim H Saris, Lorraine Brennan, Jim Kaput

Sexual dimorphism, age and fat mass are key phenotypic drivers of proteomic signatures

doi: 10.1021/acs.jproteome.7b00501

ABSTRACT

Validated protein biomarkers are needed for assessing health trajectories, predicting and sub-classifying disease, and optimizing diagnostic and therapeutic clinical decision making. The sensitivity, specificity, accuracy, and precision of single or combinations of protein biomarkers may be altered by differences in physiological states limiting the ability to translate research results to clinically useful diagnostic tests. Aptamer based affinity assays were used to test whether low abundant serum proteins differed based on age, sex and fat mass in a healthy population of 94 males and 102 females from the MECHE cohort. The findings were replicated in 217 healthy male and 377 healthy female participants in the DiOGenes consortium. Of the 1129 proteins in the panel, 141, 51 and 112 proteins (adjusted $p<0.1$) were identified in the MECHE cohort and significantly replicated in DiOGenes for sexual dimorphism, age, and fat mass, respectively. Pathway analysis classified a subset of proteins from the 3 phenotypes to the complement and coagulation cascades pathways and to immune and coagulation processes. These results demonstrated that specific proteins were statistically associated with dichotomous (male v female) and continuous phenotypes (age, fat mass) which may influence the identification and use of biomarkers of clinical utility for health diagnosis and therapeutic strategies.

INTRODUCTION

The concentrations of proteins in the blood vary dynamically in the healthy state but may also change during the trajectory toward the onset of disease. Robust technologies that accurately measure protein levels have an increasingly important role in investigating and advancing health research and clinical practice [1, 2]. The full promise of protein diagnostics has yet to be realized in the clinical setting. The majority of protein diagnostic tests are based on single proteins for acute conditions (e.g., myocardial infarction) or cancers. Protein signature tests consisting of multiple markers may be needed to achieve an optimal level of sensitivity and specificity for assessing complex health and disease processes [3].

Variations in phenotype during aging, physiology (e.g., obesity or other physical conditions), or by sex dimorphism may independently affect protein levels making it difficult to optimize the utility of clinical diagnostics, especially in genetically and culturally diverse individuals. Many of the well-accepted risk factors for cardio-metabolic disease risk have defined phenotypic cut-offs. For example, HDL cholesterol levels less than 40 mg/dL (1.0 mmol/L) are used to assess increased risk of heart disease for men but that cut-off is 50 mg/dL (1.3 mmol/L) for women. Sex differences and other risk factors such as LDL cholesterol levels may independently increase risk of heart disease or other chronic medical conditions [4]. Previous work in the field of proteomics has identified 40 low-abundant proteins which differed in serum between 12 males and 12 females [5] and more than 60 plasma proteins differed by over 2 standard deviations in 29 and 30 overweight and obese women and men, respectively [6]. Age, body mass index (BMI), body fat mass, and other physiological parameters may also influence the serum proteome and therefore utility and veracity of diagnostic markers. Serum proteomic and metabolomic approaches were combined to identify circulating proteins and metabolites that differed between 5 healthy lean and 5 healthy obese men [7]. That study, albeit small, established a link between the complement system and obesity and both novel and previously reported markers of alterations in body fat mass were identified. Considering age, physiological (such as, body fat mass), dietary, and other environmental variations, additional research into sex dimorphic plasma and serum protein modulations will be needed before sex specific medical and nutritional recommendations are implemented.

One of the main challenges for analyzing the blood proteome is the large dynamic range in protein concentrations [8]. New technologies have been developed and successfully implemented to overcome this challenge [9]. Chemically modified single-stranded DNA aptamers (SOMAmers) have high specificity as affinity capture reagents for use with

undiluted and diluted plasma and serum samples to quantify low and high abundant proteins. SOMAmers are used in multiplex assays similar to DNA microarrays allowing for the simultaneous analysis of over 1000 proteins in small amounts (~65 ul) of serum. Improvements in mass spectroscopy pipelines and analysis[10] are also being made in blood proteomic analysis although these approaches require expensive equipment and expertise in the technologies.

The primary aim of the research reported here was to identify the impact of sex, age and body fat mass on the proteomic signature and replicate the findings in an independent cohort. Furthermore, the identified proteins were mapped using pathway analysis methods to provide context and a greater understanding of the biological processes that differ by phenotype. The results of this study provide a foundational understanding of the effect of these 3 phenotypic variables on protein markers.

METHODS

Study population

The research described here extended the Metabolic Challenge (MECHE) study which is part of a national research program by the Joint Irish Nutrigenomics Organisation[11]. Briefly, the MECHE study enrolled 214 participants aged 18-60 y who underwent an oral glucose tolerance test (OGTT) and/or oral lipid tolerance test (OLTT). Clinical measures, body composition, and dietary habits were assessed in the fasted state (baseline) and at multiple time points following each challenge[11]. Demographic parameters obtained at baseline were used for analysis. Height was obtained using a wall mounted Harpenden stadiometer (Holtain Limited, UK) and weight was measured using a calibrated beam balance platform scale (SECA 888, Germany). Total fat mass was determined using DXA scanning (Lunar iDXA, GE Healthcare, UK). Individuals were informed about the purpose of the study and the experimental procedures, prior to giving written consent. Good health was defined as the absence of any known chronic or infectious disease and this was verified by a number of fasting blood tests. Individuals with a BMI below 18.5 kg/m2, a low blood haemoglobin concentration (<12 g/dL), an elevated fasting plasma glucose (≥7 mM), an elevated cholesterol concentration (>7.5 mM), an elevated triglyceride concentration (>3.8 mM) and elevated enzyme indicators of liver or kidney function, any of which warranted pharmaceutical intervention, were excluded. Details of the study have been published elsewhere[11-13]. The study was registered at clinicaltrials.gov under NCT01172951. Ethical approval was obtained from the Research Ethics Committee at University College Dublin (LS-08-43-Gibney-Ryan) and the study was performed according to the Declaration of Helsinki. For the present

study, participants from the MECHE study who had proteomic data were included (n=200) (**Table 1**).

Table 1. Clinical characteristics of study participants

	MECHE (n=196)	DiOGenes (n=594)
Sex (m/f)	94/102	217/377
Age(y)	31 ± 10*	42 ± 6 **
Weight (kg)	74.42 ± 15.99	99.87 ± 17.8
BMI (kg/m^2)	24.7 ± 4.8	34.2 ± 4.8
WHR	0.84 ± 0.1	-
BP SYS (mm/Hg)	122.93 ± 12.37	-
BP DIA (mm/Hg)	73.69 ± 10.52	-
Glucose (mmol/L)	5.21 ± 0.58	-
HDL-c (mmol/L)	1.35 ± 0.35	1.2 ± 0.33
TAG (mmol/L)	1.04 ± 0.62	1.37 ± 0.65
Insulin (µIU/mL)	8.71 ± 6.93	-
HOMA-IR	1.87 ± 1.51	3.16 ± 2.56
Body fat %	25.76 ± 10.92	-
DEXA fat mass (kg)	28.28 ± 9.72	39.72 ± 11.1

Data are presented as means ± standard deviation (SD); BMI: body mass index; WHR: waist to hip ratio; HDL-c: high-density lipoprotein cholesterol; TAG: triacylglycerol; HOMA: homeostasis model assessment.*18-60 y **23-58 y.

Proteomics analysis

1,129 proteins were quantified in fasting (at least 12 hours) serum samples of 200 MECHE participants using the proteomic platform SOMAmer™ (Slow Off-rate Modified Aptamer) as previously described [9]. Dataset is available upon request. This technology has a dynamic range of more than 8 logs, allowing quantification of both low and high abundant proteins which might otherwise be missed. Pre-processing of the proteomic data included log transformation of the abundance of each protein. Principal component analysis (PCA) did not reveal any significant batch effect across the proteins analyzed. PCA identified four individuals as outliers whose data were removed. Therefore, the final proteomic dataset included 196 individuals and 1,129 proteins. Proteins measured by SOMAmers are found in the blood as secreted (431), external membrane origin (275), and intracellular proteins (423). Proteins are often shed from membranes by proteolytical cleavage and intracellular proteins may be released from cells as a part of normal or abnormal physiological cell turnover.

Replication cohort

Participants were recruited from 8 cities in 8 European counties that were healthy, overweight/obese with a BMI between 27 and 45 kg/m2 and aged <65 y. Informed consent was obtained from all participants and the study was approved by the local

Medical Ethical Committees in the respective research centers, in accordance with the Helsinki Declaration [14]. The DiOGenes intervention study was registered on Clingov (NCT00390637). The DiOGenes cohort were selected as a suitable replication cohort for this analysis due to the availability of SOMAlogic data and a large sample size. SOMAlogic proteomic data were analyzed in serum of 594 participants: 377 females and 217 males of the DiOGenes consortium, age 16-63 y, mean BMI 34 ± 4.8 kg/m2 (0 individuals with BMI<25, 122 individuals with BMI 25-30, and 472 individuals with BMI>30 [15].

Statistical methodology

Analysis for study population characteristics was carried out using IBM SPSS Statistics 20. Data are expressed as means ± standard deviation. Multivariate statistical analysis was performed using Simca-P+ software (version 14.0; Umetrics, Umea, Sweden). Prior to data analyses the MECHE dataset was scaled using UV scaling. PCA and PLS-DA was carried out to explore differences in protein levels between males and females. Orthogonal partial last-squares discriminant analysis (OPLS-DA) was performed which improves interpretation and separation between classes on a scores plot by filtering unwanted variation. A S-plot was generated from which potential proteins of interest were extracted. A value for p (corr) > 0.15 was used to select proteins that differed significantly between males and females to enhance identification of pathways.

Robust regression (R package limma) [16] was used to identify proteins that were significantly associated with either age or total body fat measured by DEXA (fat mass in kg). Robust regression was chosen over linear regression since it is less sensitive to outliers. Proteins levels were first transformed to the residuals from a linear regression on sex to correct for this covariate. The threshold of significance of Benjamini Hochberg (BH) adjusted p-values was 0.1. Baseline serum samples from DiOGenes were analyzed to test replication of the effect of sex, age and fat mass on protein levels. Robust regressions corrected for the collection center and sex, when analyzing age and fat mass.

Pathway annotation

Pathway over-representation analysis was performed with the human pathway collection from WikiPathways (curated collection with 276 pathways downloaded on 26 January 2016) using PathVisio (version 3.2.4) [17]. Permuted p-value is calculated by performing a permutation test. The data is permuted 1000 times and a rank is calculated of the actual Z score compared to the permuted Z Scores. The Z scores are calculated based on the changed proteins in a pathway out of the total proteins in the pathway that have been measured in the uploaded dataset. Pathways with a Z-score of >1.96 and a permuted p-value < 0.05 are considered important. Functional pathway enrichment

analyses were also performed with KEGG and Reactome databases with the R packages HTSAnalyzer [18] and ReactomePA, [19] respectively. The analyses were conducted with all ENTREZ proteins/genes as background and with an adjusted p-value threshold of 0.05. Pathway analyses were conducted only with proteins found to vary significantly in the same direction in both cohorts.

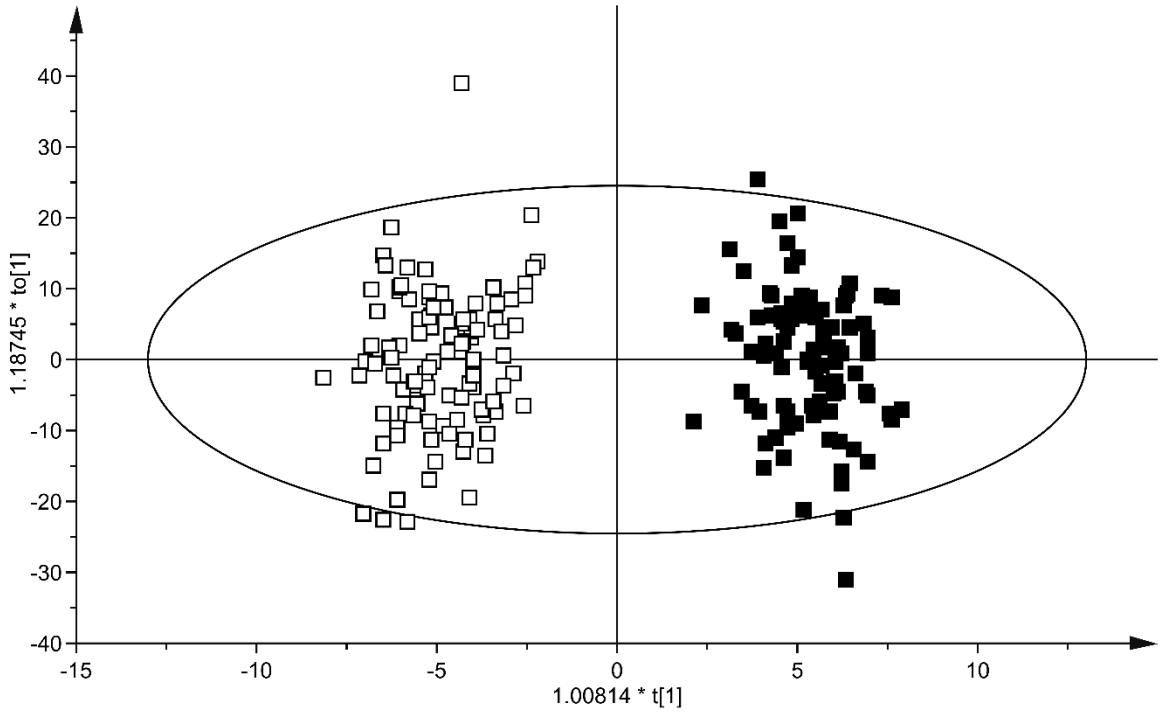

Figure 1. OPLS-DA of males vs females derived from proteomic data of MECHE participants (n = 196). (■) Males, (□) Females, R2Y = 0.945; Q2 =0.765.

RESULTS/DISCUSSION

Sexual Dimorphism

The 317 differentially expressed proteins between male and female in the MECHE cohort account for 28% of the total proteins analyzed (**Table 2** and **3, Figure 1**). From there 141 proteins were replicated in the DiOGenes cohort (**Table S5a and S5b**). The top 10 most statistically significant over-expressed proteins for each sex were compared for known physiological functions and associations with sex hormones, metabolic disease, diet, and previously characterized sex dimorphism (**Table S1** and **S2**).

The 8 most significant secreted proteins elevated in females have known associations with sex hormone metabolism (SHBG, leptin, thyroxin-binding globulin, adiponectin, angiotensinogen, fetuin B, immunoglobulin M, trefoil factor 3) (**Table 3**) [20-25]. Each of these proteins is involved in at least one diet related metabolic disease (except

immunoglobulin M) (**Table S1**). These proteins were affected by or affect glucose and insulin metabolism (SHBG, leptin, adiponectin, fetuin B, trefoil factor 3), metabolic rate (thyroxin binding globulin), and dietary carbohydrate intake (SHBG, leptin, adiponectin), and salt sensitive hypertension (angiotensinogen) [21, 22, 26-30]. Increased immunoglobulin M expression has been associated with gluten and dietary protein intakes [31, 32]. Recent studies suggest SHBG, adiponectin, angiotensinogen, and fetuin B may be involved in diabetes development [28, 33-35]. The cell-membrane located immune proteins, LAMP, and collectin placenta 1 were upregulated in females. These proteins stimulate neural growth and provide host defense with no previously described difference by sex or metabolic disease associations [36, 37]. Of note, collectin placenta 1 is the only protein within this group that was not replicated in the DiOGenes data.

Nine of the top 10 most significant proteins more abundant in males are secreted (exception is myoglobin located in the exosome - **Table 2**). Of these 10, several are associated with sex hormone metabolism (myoglobin, matrix metalloproteinase 3, serum amyloid P, tissue factor pathway inhibitor, protein S, interleukin 1 receptor like 1, LEAP) [38-46]. These top 10 proteins function in connective tissue development and growth, amyloid deposit development, blood coagulation, inflammation modulation, anti-microbial immunity and iron metabolism, as well as, immune cell migration and adhesiveness. These proteins are involved in processes contributing to metabolic diseases, specifically myocardial infarction (myoglobin), cardiovascular disease (matrix metalloproteinase), atherosclerosis (serum amyloid P, tissue factor pathway inhibitor, protein S), diabetes (ficolin-3), and iron overload (LEAP-1) [47-52]. Certain nutrients alter the abundances of some of these proteins: iron (myoglobin, LEAP), lipids (myoglobin), monounsaturated fatty acids, and n-3 PUFA (tissue factor pathway inhibitor), vitamin K (protein S), and vitamin A (LEAP) although the effect of diet was not analyzed in the MECHE cohort. The liver expressed chemokine (HCC-4) may be induced by total fat and calorie intake [53-60]. All 10 significant proteins found in MECHE were replicated in the same direction with robust regression in the DiOGenes data (**Table S5a**).

Interpreting the role of a protein by its activity, association to a disease process, or association to an individual provides information on physiological states. However, our results suggest a more holistic difference between males and females since the coagulation pathway cross-talks with and cross-regulates the immune system to maintain homeostasis [61]. Serpin Family D Member 1, a1-antitrypsin and plasminogen mapped to the complement and coagulation cascades pathway (**Figure 3**) in females.

Table 2: Proteomics results depict sexual dimorphism – proteins higher in males

Full Protein Name	UniProt ID	Protein	OPLS-DA	p (corr)	Cellular
Myoglobin	P02144	MB	0.118	0.613	Exosome
Matrix metalloproteinase 3	P08254	MMP3	0.110	0.574	Secreted
Bone morphogenetic protein-1	P13497	BMP-1	0.105	0.549	Secreted
Serum amyloid P	P02743	SAP	0.100	0.522	Secreted
Tissue factor pathway inhibitor	P10646	TFPI	0.099	0.516	Secreted
Protein S	P07225	Protein S	0.096	0.501	Secreted
Ficolin-3	O75636	Ficolin-3	0.094	0.491	Secreted
Interleukin-1 receptor-like 1	Q01638	IL-1 R4	0.093	0.483	Secreted
LEAP-1 (Hepcidin)	P81172	LEAP-1	0.090	0.470	Secreted
Liver-expressed chemokine HCC-4	O15467	HCC-4	0.088	0.459	Secreted

[a] Data presented as first 10 significant proteins out of 173 total in males using OPLS-DA with a 0.15 cut-off. [c] describes the direction of the difference in males vs. females.

Table 3: Proteomics results depict sexual dimorphism – proteins higher in females

Full Name	Uniprot ID	Protein	OPLS-DA	p (corr)	Cellular
Sex hormone-binding	P04278	SHBG	-0.102	-0.531	Secreted
Leptin (OB)	P41159	Leptin	-0.098	-0.513	Secreted
Thyroxine-binding	P05543	Thyroxine Binding	-0.097	-0.507	Secreted
Adiponectin	Q15848	Adiponectin	-0.089	-0.466	Secreted
Angiotensinogen	P01019	Angiotensinogen	-0.079	-0.414	Secreted
Fetuin B	Q9UGM5	FETUB	-0.073	-0.381	Secreted
Immunoglobulin M	P01871	IgM	-0.072	-0.378	Secreted
Trefoil factor 3 - Intestinal trefoil	Q07654	TFF3	-0.071	-0.369	Secreted
Limbic system-associated	Q13449	LSAMP	-0.069	-0.362	Membrane
Collectin Placenta 1[b]	Q5KU26	COLEC12	-0.067	-0.351	Membrane

[a] Data presented as first 10 significant proteins out of 144 total in females using OPLS-DA with a 0.15 cut-off. [b] Only protein not replicated in DiOGenes data set. [c] Describes the direction of the difference in males vs. females.

The same pathway was identified in males but through different proteins (tissue factor pathway inhibitor, thrombin activatable fibrinolysis inhibitor, plasminogen activator, serpin family A member 5, serpin family C member 1 and Protein S). Tissue factor pathway inhibitor is present in free and lipoprotein-associated forms [62] while protein S is more frequently (60% of total) bound to C4B which abolishes its anticoagulant properties [63]. Bound and free Protein S were more abundant in males compared to females [64]. This protein also is involved in phagocytosis of apoptotic cells [65]. Others studies identified Serpin Family D Member 1, Serpin Family C Member 1, Serpin Family A Member 1 and protein S among 27 significant proteins that differed in the complement and coagulation cascades between males and females [6]. Toll- like receptors, immune- and adipo-cytokine proteins were more abundant in females.

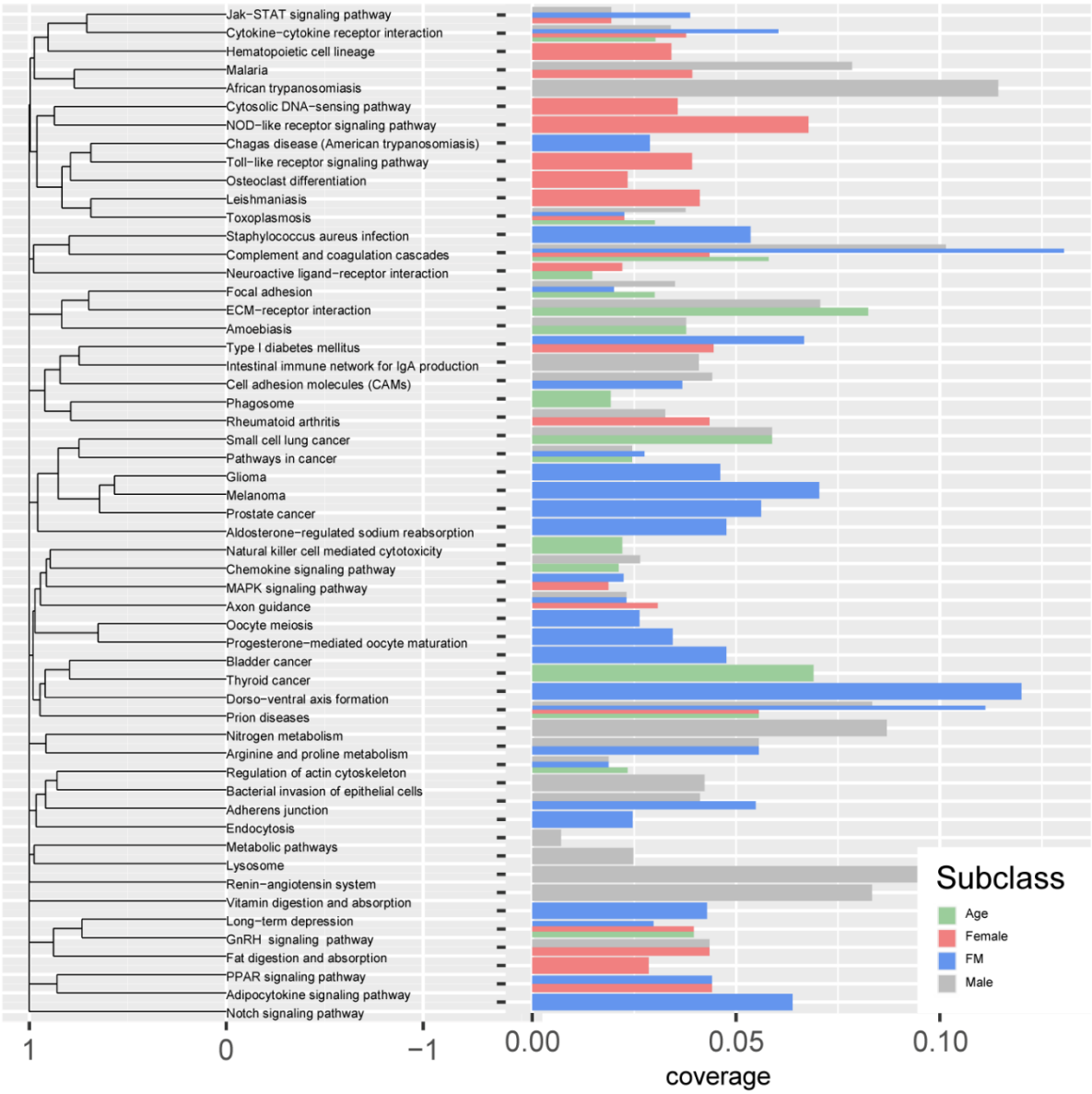

Figure 2. Overview of KEGG pathway enrichment. Bar graph displaying KEGG pathway enrichments by classes (proteins significant for male,female, age and fat mass). The size of the bar graph represent the coverage of the pathway (number significant proteins in the pathway/number of total proteins in the pathway). The dendrogram groups similar pathways (pathways that include similar genes).

These proteins mapped to pathways (**Figures 2, Figure S1, Table 6, Tables S6-S8**) previously identified as belonging to the inflammasome [1], a system of interacting networks regulating acute and chronic inflammatory conditions.

The individual proteins mapped to these pathways (e.g., members of the interleukin family) have been linked with diseases associated with chronic inflammation in (for example) obesity and T2DM [2], the pathogenesis for which differs between men and women [3, 4]. This connected complement-immune system may result from and

contribute to metabolic differences of sex dimorphism. How these related systems are regulated will require more comprehensive analysis of components of these pathways in future studies.

Age

Regression analysis revealed 167 proteins (15% of 1129 proteins) significantly associated (adjusted p-value <0.1) with age (range: 20-60 y). Fifty-one of these protein - age associations were replicated in the DiOGenes data (**Table S5c**). Coefficients of the regression for age can also be found in **Table S5c**. The top 10 proteins associated with age were IL1RL2, FSHB, ADMTS5, CHIT1 (all positively correlated with age) and AGRP, OMD, RET, CDON, IGFBP3 and IGFBP5 (negatively correlated with age) (see **Table 4**), however OMD was not significantly associated in DiOGenes. The levels of the majority of these proteins, with the exception of IL1RL2 and CDON, were previously associated with age [5-11]. The identified proteins are involved in diseases and sub-optimal states of health in relation to aging including (i) inflammation (IL1RL2, CHIT1) [10-12], (ii) arthritis (ADAMTS5, IGFBP3) [13, 14], (iii) vertebral fractures (IGFBP3) [15] (iv) bone metabolism (IGFBP5,AGRP,OMD) [16-18], (v) weight homeostasis (AGRP,CDON) [19, 20], (vi) lean body mass (IGFBP3) [21], (vii) cancer development (RET,CDON) [22, 23] and prevention (IGFBP3, IGFBP5) [24, 25], and (viii) muscle metabolism (IGFBP3, IGFBP5, CDON) [26, 27]. Levels of FSHB and AGRP were positively influenced by caloric restriction and a high-fat diet, while IGFBP3 was impacted by supplementing the diet with n-3 PUFA [28-30].

Table 4: Proteomics results significantly associated with age[a]

Full Protein Name	UniProt ID	Protein	P.Value	adj.P.Val	Cellular
Interleukin 1 receptor like 2	Q9HB29	IL1RL2	1.85E-13	2.09E-10	Membrane
alpha polypeptide - follicle	P01215	CGA	1.12E-11	6.34E-09	Secreted
Metallopeptidase with	Q9UNA0	ADAMTS5	1.41E-10	5.29E-08	Secreted
Chitinase 1	Q13231	CHIT1	2.08E-10	5.88E-08	Secreted
Agouti related neuropeptide	O00253	AGRP	4.51E-10	1.02E-07	Secreted
Osteomodulin[b]	Q99983	OMD	6.89E-10	1.30E-07	Secreted
Ret proto-oncogene 2	P07949	RET	2.85E-09	4.60E-07	Membrane
Cell adhesion associated,	Q4KMG0	CDON	1.32E-07	1.86E-05	Membrane
Insulin like growth factor	P17936	IGFBP3	1.83E-07	2.29E-05	Secreted
Insulin like growth factor	P24593	IGFBP5	2.04E-07	2.30E-05	Secreted

[a] Data presented as top 10 significant proteins out of 166 total using robust regression with correlations calculated using residuals following correction for sex. [b] Only protein not replicated in DiOGenes data set.

Annotation of the MECHE proteins statistically significant in the DiOGenes cohort identified a number of associated pathways (**Figure 2, Figure S1, Table 6, Tables S6-**

S8). Four proteins associated with aging emerged in the complement and coagulation cascade pathway, of which SERPING1 (c1 esterase inhibitor) was specific to aging. Increases in C1 esterase inhibitor are observed during inflammation [31]. All of these proteins with the exception of CCL21 were positively associated with increased age. The chemokine pattern found in this study (CCL21 lower and other CCL's higher) was consistent with other studies showing increased expression of CCL27 in senescent cells [32], levels of CCL11 and CCL7 in aged animals or humans [33], and decreased levels of CCL21 [34]. Further pathways of interest were identified through KEGG and Reactome (**Tables S6-7**). Collectively, these proteins and the pathways in which they act are processes consistent with inflammation, the interconnected processes that result from lifelong insults to the immune system resulting in chronic low-grade inflammation and immunosenescence [35].

Fat Mass

Regression analysis identified 21% of SOMAscan proteins significantly associated (adjusted p-value <0.1) with body fat mass (range: 8-58.2kg). Of these 232 proteins, 112 were replicated at an adjusted p-value <0.1 in the DiOGenes cohort, with coefficients of regression displayed in **Table S5d**. The top 10 proteins associated with body fat mass are LEP, PLAT and C1S (all positively correlated with fat mass) and IGFBP1, TFF3, SHBG, MMP2, WFIKKN2, HFE2 and TF (negatively correlated with fat mass) (**Table 5**). All 10 proteins were replicated in the DiOGenes cohort. The physiologic functions of these top proteins include inflammation, glucose metabolism, defense response, blood coagulation, regulation of cell growth, along with angiogenesis and iron homeostasis (**Table S4**). Leptin was strongly associated with fat mass [36, 37] and elevated in females, consistent with its known role in regulation of body weight and energy balance [38, 39].

Table 5: Proteomics results significantly associated with fat mass[a]

Full Protein Name	UniProt	Protein	P.Value	adj.P.Val	Cellular
Leptin	P41159	Leptin	1.57E-18	1.78E-15	Secreted
Tissue-type plasminogen	P00750	PLAT	2.92E-13	1.32E-10	Secreted
Insulin-like growth factor-binding	P08833	IGFBP1	3.51E-13	1.32E-10	Secreted
Trefoil factor 3	Q07654	TFF3	2.09E-12	5.90E-10	Secreted
Sex hormone-binding globulin	P04278	SHBG	7.32E-11	1.65E-8	Secreted
Matrix metalloproteinase 2	P08253	MMP2	3.56E-10	6.70E-8	Secreted
WAP, kazal, immunoglobulin,	Q8TEU8	WFIKKN2	4.80E-10	7.74E-8	Secreted
Complement C1s	P09871	C1S	9.65E-9	1.36E-6	Exosome
Hemojuvelin	Q6ZVN8	HFE2	2.16E-8	2.71E-6	Membrane
Transferrin	P02787	TF	2.44E-8	2.75E-6	Secreted

[a] Data presented as top 10 significant proteins out of 233 total using robust regression with correlations calculated using residuals following correction for sex. All proteins replicated with DiOGenes data set.

Increased BMI and fat mass are known risk factors for diseases such as metabolic syndrome and cardiovascular disease (CVD). Three proteins associated with fat mass in the MECHE/DiOGenes cohorts (tPA, IGFBP-1, and TFF3) have been associated with metabolic conditions. High levels of tPA antigen independently predicted cardiovascular events both in a healthy population and in individuals with prevalent coronary disease [40]. Elevated plasma tPA antigens were associated with insulin resistance, T2D, and obesity. Decreased abundance of plasma tPA (approximately 29%) was observed following a 12 week energy restricted diet in overweight women with metabolic syndrome [41]. Insulin like growth factor binding protein 1 (IGFBP-1) is negatively associated with fat mass in the MECHE/DiOGenes cohorts. Lower levels of IGFBP-1 at baseline was associated with the combination of increased percentage body fat and plasma insulin levels [42]. Trefoil factor 3 (TFF3) was negatively associated with fat mass but positively associated with female sex. Increased levels of TFF3 were observed to improve glucose tolerance in a diet-induced obesity mouse model, which supports previous reports that TFF3 plays a role in energy metabolism [43].

The 112 differentially abundant proteins associated with fat mass were mapped to KEGG, WikiPathways and Reactome pathways (**Figure 2, Figure S1, Table 6, Tables S6-S8**). Seven proteins significantly associated with fat mass mapped to the complement and coagulation cascade pathway (**Figure 3**). The complement and coagulation cascade pathway is associated with chronic disease risk [44]. In this pathway, abundances of TFPI, coagulation factor IX, tPA, Factor H and C1s were higher while anti-thrombin III and C7 were less abundant as fat mass increased. Although not directly tested in this study, enzymatic activity of thrombin would be maintained in conditions of decreased levels of anti-thrombin III with the result that coagulation would be increased. The association between increased coagulation factor IX, which is also inhibited by anti-thrombin III, and increased fat mass found in the MECHE/DiOGenes cohorts is consistent with more active coagulation processes. Evidence from cell culture demonstrated that a subset of proteins expressed in the complement pathway were altered in adipose cells from insulin resistant humans and in animal models of obesity [45]. These proteins associated with fat mass and the pathways to which they belong suggest a link between insulin resistance, T2D, and coagulation processes.

Fifteen proteins, including adiponectin, insulin, and leptin mapped to the Reactome development biology pathway (**Figure S1, Table S7**) and to KEGG cytokine - cytokine receptor interactions (**Figure 2, Table S6**).

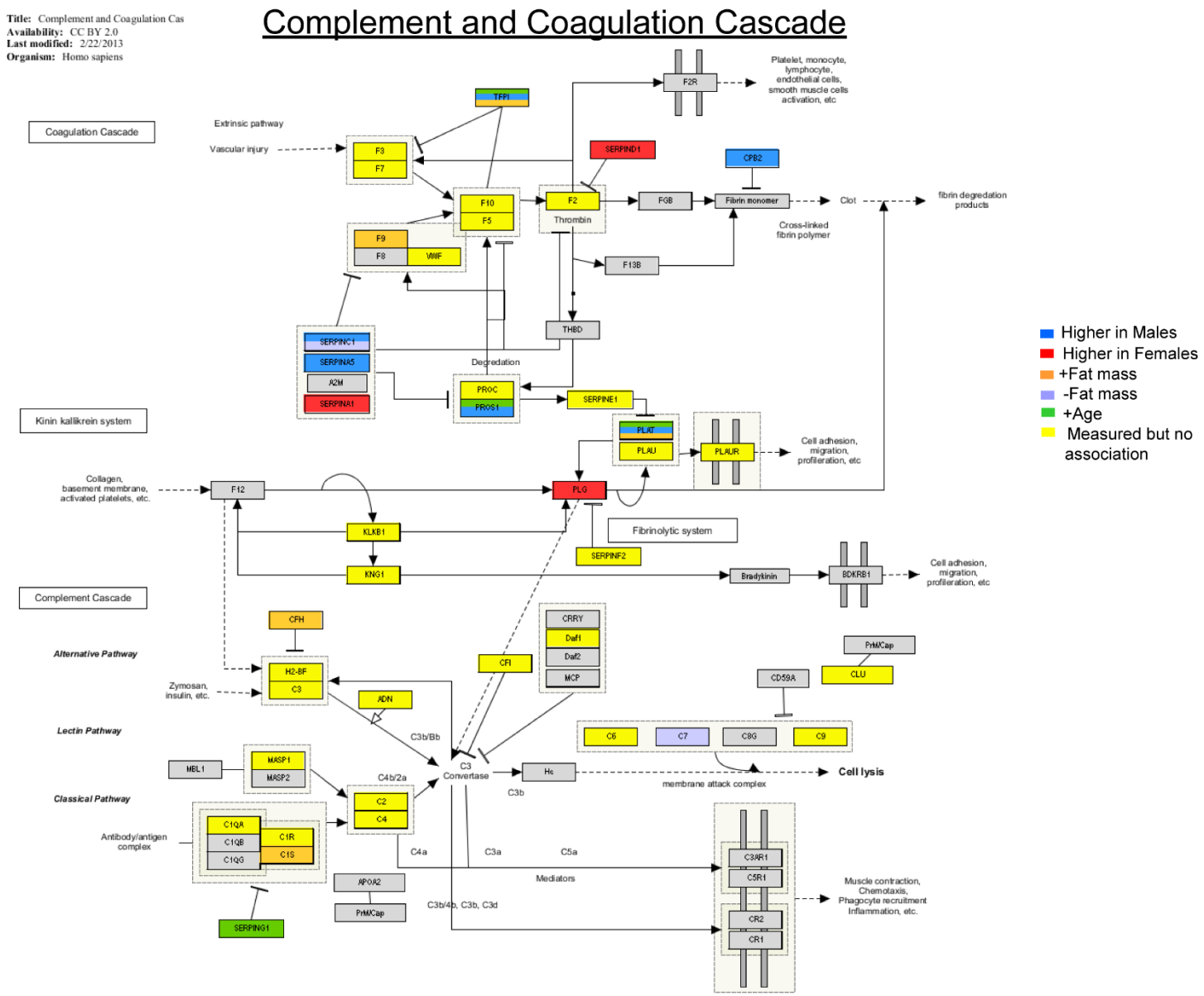

Figure 3. Complement and coagulation cascades pathway obtained from WikiPathways displaying proteins differentially expressed across sex, age, and fat mass phenotypes.

Eleven other proteins, which included ECM proteins (e.g., NCAM1, MMP2, CHL1 negative associations with FM) and growth factors (EGFR, FGFR1, negative associations with FM) were also assigned to the axon guidance pathway, a participant in Reactome's developmental biology pathway. The axon guidance pathway was identified in a transcriptomic analysis of fatty hearts in miniature pigs fed a high energy diet [1] suggesting that dysregulation of these genes may not be specific to neuronal tissues. Decreased levels of axon guidance proteins (e.g., UNC5D, RGMB CHL1) may alter neuro-adipose junctions involved in leptin regulation [2]. The mapping of individual proteins to multiple pathways also identified potential processes associated with increased fat mass. For example, NOTCH1 was found in 11 of the 33 significantly enriched Reactome pathways. NOTCH1 was inversely associated with fat mass in the present study. Decreases in levels of endothelial NOTCH1 may be a risk factor for vascular inflammation and promotion of diet-induced atherosclerosis [3]. Therefore, examination of pathways related to fat mass provides a platform for further investigation of associated biological processes.

Phenotypic variables have an impact on protein levels

Examining the impact of phenotypic variables revealed the importance of considering sex, age and fat mass in proteomics studies. Aptamer based binding assays were used to quantify low abundant serum proteins at baseline in healthy participants of the MECHE (11) and DiOGenes (15) cohorts based on sex, body fat, and age. Forty four percent of sex proteins (51% male, 35% female), 31% age proteins and 49% fat mass proteins identified as significant in the MECHE cohort were replicated in the DiOGenes cohort. The differences in replicated proteins for each phenotype group likely reflects known differences between the two cohorts. DiOGenes participants were older in age (41.6 ± 6.1 vs. 31 ± 10 y in MECHE) with a higher BMI (34.2 ± 4.8 vs. 24.7 ± 4.8 kg/m2 in MECHE) and more body fat (39.7 ± 11.1 vs. 25.76 ± 10.9 kg in MECHE) (**Table 1**). The difference in these parameters is considered a strength since the same proteins were significant in a slightly older and more obese cohort, which extends the use of these proteins in studies of individuals with wider age and BMI ranges. Histogram plots of age and sex for both cohorts are in **Figures S2a-f**. Two proteins overlapped all 3 phenotype groups, 6 proteins between age and sex, 30 proteins between sex and fat mass and 4 proteins between age and fat mass (**Figure S3**). Scatterplots of proteins related to age and fat mass respectively can be found in **Figure S6** and **Figure S7**. The present results highlight the need for including phenotypic parameters in proteomics studies and make a case for the development of phenotypic specific cut offs. Differences in sex, age and fat mass may independently induce quantitative changes in the proteins thought to be specific for a biological process or disease phenotype. That is, underlying differences in phenotype (sex, age, fat mass) may confound the identification of disease-specific

markers. The successful identification of proteins related to gender, age and fat mass in this study will allow for promising biomarkers to be tested for potential confounding factors prior to progression into a clinical setting. In addition, findings from this analysis will contribute to improved statistical modelling by including the identified proteins as confounding factors in future biomarker discovery studies.

The strengths of this study include testing whether proteins identified in the MECHE cohort were replicated in the larger DiOGenes cohort. Pathway analysis was also performed using several different pathway analysis software platforms and provided further insights into the functions of the proteins. While the present study represents an important advancement for proteomics there are a number of limitations worth noting. The proteins were identified using SOMAlogic assays which are a subset of the total protein pool. Subsequent versions of SOMAscan or mass spectroscopic methods may identify additional proteins and pathways for each of the phenotypes studied here. In addition, mapping proteins to KEGG, Reactome, and WikiPathways to create meaningful interpretation of the proteomics data is constrained by the depth and publication biases of pathway databases.

CONCLUSIONS

Phenotypic characteristics such as sex, age and body fat mass have independent associations with the levels of certain serum proteins. Mapping these proteins to pathways identified biological processes differing across phenotypic measures. Importantly, the findings were replicated in an independent cohort. Gender and sex specific health care is emerging as differences in trajectories towards disease and therapeutic responses between males and females are identified. Many of the most significant proteins identified in this study had known relationships with sex hormone metabolism indicating sex hormones play key roles in influencing metabolic health. Additionally, age and fat mass are well-established risk factors for disease. These results are relevant to the development of diagnostic and prognostic markers of health and disease trajectories. The present results will be an important consideration in the development of protein signatures for use in the clinical setting.

Table 6: Overview of pathways related to sex, age and fat mass using WikiPathways

	Pathway	Positive	Measured	Z Score	*P*	ID
Males	Complement and Coagulation Cascades	7	40	2.27	0.039	WP558
	Selenium Micronutrient Network	7	27	3.49	0.003	WP15
	Vitamin B12 Metabolism	7	24	3.88	0.002	WP1533
	Folate Metabolism	6	26	2.88	0.011	WP176
	Statin Pathway	3	4	4.95	0.001	WP430
	Urea cycle and metabolism of amino groups	3	4	4.95	0.001	WP497
Females	Adipogenesis	5	23	4.05	<0.01	WP236
	Aryl Hydrocarbon Receptor	3	19	2.41	0.032	WP2586
	FAS pathway and Stress induction of HSP regulation	3	15	2.93	0.021	WP314
	Myometrial Relaxation and Contraction Pathways	3	22	2.1	0.031	WP289
Aging	Complement and Coagulation Cascades	4	40	2.01	0.044	WP558
	Senescence and Autophagy in Cancer	4	39	2.07	0.047	WP615
Fat mass	Spinal cord injury	8	47	1.78	0.052	WP2431
	Complement and coagulation Cascades	7	40	1.74	0.053	WP558
	Adipogenesis	6	23	2.72	0.006	WP236
	Cardiac Progenitor Differentiation	5	15	3.15	0.008	WP2406
	Neural Crest Differentiation	5	18	2.65	0.017	WP2064
	Complement Activation	4	17	1.97	0.038	WP545
	AMPK signaling	3	12	1.82	0.049	WP1403
	Differentiation of white and brown adipocyte	3	3	5.33	<0.01	WP2895
	Notch Signaling Pathway	3	7	3	0.013	WP268
	NOTCH1 regulation of human endothelial cell calcification	3	10	2.2	0.039	WP3413

Table 6. Pathways obtained from pathway statistics using PathVisio software, using the curated WikiPathways directory. Pathways with a Z-Score of >1.96, a p-value of <0.05 and who have 3 or more proteins differentially expressed are considered important. P-value is permuted. Sorted by number of differentially expressed proteins in pathway.

FUNDING

This work was conducted as part of a clinical trial (NCT01172951), funded under the Food for Health Research Initiative (NDP 2007–2013; 07FHRIUCD1) by the Department of Agriculture, Fisheries and Food (http://www.agriculture.gov.ie/). This research was funded in part by Food for Health Ireland, through Enterprise Ireland Grant No. TC20130001. The DiOGenes project was supported by a contract (FP6-2005- 513946) from the European Commission Food Quality and Safety Priority of the Sixth Framework Program.

ACKNOWLEDGEMENTS

Authors would like to acknowledge participants of both the MECHE and DiOGenes studies. Special thanks to Kei Sakamoto, Sofia Moço, and Loic Dayon from Nestle Institute of Health Sciences and Jan van der Greef from Leiden University for their review of the manuscript.

REFERENCES

[1] Hanash, S., Disease proteomics. *Nature* 2003, *422*, 226-232.
[2] Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP* 2002, *1*, 845-867.
[3] Zethelius, B., Berglund, L., Sundstrom, J., Ingelsson, E., Basu, S., Larsson, A., Venge, P., Arnlov, J., Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *The New England journal of medicine* 2008, *358*, 2107-2116.
[4] American Diabetes Association, Cardiovascular disease and risk management. *Diabetes Care* 2015, *38* S49-57.
[5] Miike, K., Aoki, M., Yamashita, R., Takegawa, Y., Saya, H., Miike, T., Yamamura, K., Proteome profiling reveals gender differences in the composition of human serum. *Proteomics* 2010, *10*, 2678-2691.
[6] Al-Daghri, N. M., Al-Attas, O. S., Johnston, H. E., Singhania, A., Alokail, M. S., Alkharfy, K. M., Abd-Alrahman, S. H., Sabico, S. L., Roumeliotis, T. I., Manousopoulou-Garbis, A., Townsend, P. A., Woelk, C. H., Chrousos, G. P., Garbis, S. D., Whole serum 3D LC-nESI-FTMS quantitative proteomics reveals sexual dimorphism in the milieu interieur of overweight and obese adults. *J Proteome Res* 2014, *13*, 5094-5105.
[7] Oberbach, A., Bluher, M., Wirth, H., Till, H., Kovacs, P., Kullnick, Y., Schlichting, N., Tomm, J. M., Rolle-Kampczyk, U., Murugaiyan, J., Binder, H., Dietrich, A., von Bergen, M., Combined proteomic and metabolomic profiling of serum reveals association of the complement system with obesity and identifies novel markers of body fat mass changes. *J Proteome Res* 2011, *10*, 4769-4788.
[8] Chandramouli, K., Qian, P. Y., Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genomics Proteomics* 2009, *2009*.
[9] Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E. N., Carter, J., Dalby, A. B., Eaton, B. E., Fitzwater, T., Flather, D., Forbes, A., Foreman, T., Fowler, C., Gawande, B., Goss, M., Gunn, M., Gupta, S., Halladay, D., Heil, J., Heilig, J., Hicke, B., Husar, G., Janjic, N., Jarvis, T., Jennings, S., Katilius, E., Keeney, T. R., Kim, N., Koch, T. H., Kraemer, S., Kroiss, L., Le, N., Levine, D., Lindsey, W., Lollo, B., Mayfield, W., Mehan, M., Mehler, R., Nelson, S. K., Nelson, M., Nieuwlandt, D., Nikrad, M., Ochsner, U., Ostroff, R. M., Otis, M., Parker, T., Pietrasiewicz, S., Resnicow, D. I., Rohloff, J., Sanders, G., Sattin, S., Schneider, D., Singer, B., Stanton, M., Sterkel, A., Stewart, A., Stratford, S., Vaught, J. D., Vrkljan, M., Walker, J. J., Watrobka, M., Waugh, S., Weiss, A., Wilcox, S. K., Wolfson, A., Wolk, S. K., Zhang, C., Zichi, D., Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* 2010, *5*, e15004.
[10] Cominetti, O., Nunez Galindo, A., Corthesy, J., Oller Moreno, S., Irincheeva, I., Valsesia, A., Astrup, A., Saris, W. H., Hager, J., Kussmann, M., Dayon, L., Proteomic Biomarker Discovery in 1000 Human Plasma Samples with Mass Spectrometry. *J Proteome Res* 2016, *15*, 389-399.
[11] Morris, C., O'Grada, C., Ryan, M., Roche, H. M., Gibney, M. J., Gibney, E. R., Brennan, L., Identification of differential responses to an oral glucose tolerance test in healthy adults. *PLoS One* 2013, *8*, e72890.
[12] Wallace, M., Morris, C., O'Grada, C. M., Ryan, M., Dillon, E. T., Coleman, E., Gibney, E. R., Gibney, M. J., Roche, H. M., Brennan, L., Relationship between the

lipidome, inflammatory markers and insulin resistance. *Molecular bioSystems* 2014, *10*, 1586-1595.
[13] Ryan, M. F., O'Grada, C. M., Morris, C., Segurado, R., Walsh, M. C., Gibney, E. R., Brennan, L., Roche, H. M., Gibney, M. J., Within-person variation in the postprandial lipemic response of healthy adults. *Am J Clin Nutr* 2013, *97*, 261-267.
[14] Larsen, T. M., Dalskov, S., van Baak, M., Jebb, S., Kafatos, A., Pfeiffer, A., Martinez, J. A., Handjieva-Darlenska, T., Kunesova, M., Holst, C., Saris, W. H., Astrup, A., The Diet, Obesity and Genes (Diogenes) Dietary Study in eight European countries - a comprehensive design for long-term intervention. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2010, *11*, 76-91.
[15] Papadaki, A., Linardakis, M., Plada, M., Larsen, T. M., van Baak, M. A., Lindroos, A. K., Pfeiffer, A. F., Martinez, J. A., Handjieva-Darlenska, T., Kunesova, M., Holst, C., Saris, W. H., Astrup, A., Kafatos, A., Diet, O., Genes, P., A multicentre weight loss study using a low-calorie diet over 8 weeks: regional differences in efficacy across eight European cities. *Swiss medical weekly* 2013, *143*, w13721.
[16] Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., Smyth, G. K., limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 2015, *43*, e47.
[17] Kutmon, M., van Iersel, M. P., Bohler, A., Kelder, T., Nunes, N., Pico, A. R., Evelo, C. T., PathVisio 3: an extendable pathway analysis toolbox. *PLoS computational biology* 2015, *11*, e1004085.
[18] Wang, X., Terfve, C., Rose, J. C., Markowetz, F., HTSanalyzeR: an R/Bioconductor package for integrated network analysis of high-throughput screens. *Bioinformatics* 2011, *27*, 879-880.
[19] Yu, G., He, Q. Y., ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Molecular bioSystems* 2016, *12*, 477-479.
[20] Liu, S., Sun, Q., Sex Differences, Endogenous Sex-hormone Hormones, Sex-hormone Binding Globulin (SHBG), and Exogenous Disruptors in Diabetes and Related Metabolic Outcomes. *Journal of diabetes* 2016.
[21] Geber, S., Brandao, A. H., Sampaio, M., Effects of estradiol and FSH on leptin levels in women with suppressed pituitary. *Reproductive biology and endocrinology : RB&E* 2012, *10*, 45.
[22] Tahboub, R., Arafah, B. M., Sex steroids and the thyroid. *Best practice & research. Clinical endocrinology & metabolism* 2009, *23*, 769-780.
[23] Komukai, K., Mochizuki, S., Yoshimura, M., Gender and the renin-angiotensin-aldosterone system. *Fundamental & clinical pharmacology* 2010, *24*, 687-698.
[24] Floehr, J., Dietzel, E., Neulen, J., Rosing, B., Weissenborn, U., Jahnen-Dechent, W., Association of high fetuin-B concentrations in serum with fertilization rate in IVF: a cross-sectional pilot study. *Hum Reprod* 2016, *31*, 630-637.
[25] Strauss, J. F. B., Robert L., *Yen & Jaffe's Reproductive Endocrinology: Physiology, Pathophysiology, and Clinical Management, Seventh Edition*, Saunders, Philadelphia, Pennsylvania 2014.
[26] Friedman, J., The long road to leptin. *The Journal of clinical investigation* 2016, *126*, 4727-4734.
[27] Volp, A. P., Barbosa, K. B., Bressan, J., Nutrients can modulate the adiponectin concentrations in apparently healthy young adult. *Nutricion hospitalaria* 2016, *33*, 264.

[28] Meex, R. C., Hoy, A. J., Morris, A., Brown, R. D., Lo, J. C., Burke, M., Goode, R. J., Kingwell, B. A., Kraakman, M. J., Febbraio, M. A., Greve, J. W., Rensen, S. S., Molloy, M. P., Lancaster, G. I., Bruce, C. R., Watt, M. J., Fetuin B Is a Secreted Hepatocyte Factor Linking Steatosis to Impaired Glucose Metabolism. *Cell Metab* 2015, *22*, 1078-1089.
[29] Jiang, H., Przybyszewski, J., Mitra, D., Becker, C., Brehm-Stecher, B., Tentinger, A., MacDonald, R. S., Soy protein diet, but not Lactobacillus rhamnosus GG, decreases mucin-1, trefoil factor-3, and tumor necrosis factor-alpha in colon of dextran sodium sulfate-treated C57BL/6 mice. *The Journal of nutrition* 2011, *141*, 1239-1246.
[30] Burnier, M., Urinary angiotensinogen and salt sensitivity of blood pressure: the challenge of finding biomarkers of salt-sensitivity. *Journal of hypertension* 2015, *33*, 1368-1370.
[31] Montenegro, L., Piscitelli, D., Giorgio, F., Covelli, C., Fiore, M. G., Losurdo, G., Iannone, A., Ierardi, E., Di Leo, A., Principi, M., Reversal of IgM deficiency following a gluten-free diet in seronegative celiac disease. *World journal of gastroenterology* 2014, *20*, 17686-17689.
[32] Kim, K. O., Park, H., Kim, H. S., Effects of High-Protein Diet and/or Resveratrol Supplementation on the Immune Response of Irradiated Rats. *Preventive nutrition and food science* 2014, *19*, 156-163.
[33] Kische, H., Gross, S., Wallaschofski, H., Volzke, H., Dorr, M., Nauck, M., Haring, R., Clinical correlates of sex hormones in women: The study of health in Pomerania. *Metabolism: clinical and experimental* 2016, *65*, 1286-1296.
[34] Hara, K., Yamauchi, T., Kadowaki, T., Adiponectin: an adipokine linking adipocytes and type 2 diabetes in humans. *Curr Diab Rep* 2005, *5*, 136-140.
[35] Joyce-Tan, S. M., Zain, S. M., Abdul Sattar, M. Z., Abdullah, N. A., Renin-Angiotensin System Gene Variants and Type 2 Diabetes Mellitus: Influence of Angiotensinogen. *Journal of diabetes research* 2016, *2016*, 2161376.
[36] Pimenta, A. F., Zhukareva, V., Barbe, M. F., Reinoso, B. S., Grimley, C., Henzel, W., Fischer, I., Levitt, P., The limbic system-associated membrane protein is an Ig superfamily member that mediates selective neuronal growth and axon targeting. *Neuron* 1995, *15*, 287-297.
[37] Hansen, S. W., Ohtani, K., Roy, N., Wakamiya, N., The collectins CL-L1, CL-K1 and CL-P1, and their roles in complement and innate immunity. *Immunobiology* 2016, *221*, 1058-1067.
[38] Manttari, S., Anttila, K., Jarvilehto, M., Testosterone stimulates myoglobin expression in different muscles of the mouse. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology* 2008, *178*, 899-907.
[39] Grandas, O. H., Mountain, D. J., Kirkpatrick, S. S., Rudrapatna, V. S., Cassada, D. C., Stevens, S. L., Freeman, M. B., Goldman, M. H., Effect of hormones on matrix metalloproteinases gene regulation in human aortic smooth muscle cells. *The Journal of surgical research* 2008, *148*, 94-99.
[40] Katou, M., [Change of serum amyloid P component concentrations in women]. *Nihon Sanka Fujinka Gakkai zasshi* 1996, *48*, 481-487.
[41] Ali, H. O., Stavik, B., Dorum, E., Iversen, N., Sandset, P. M., Skretting, G., Oestrogen induced downregulation of TFPI expression is mediated by ERalpha. *Thrombosis research* 2014, *134*, 138-143.

[42] Jin, H., Lin, J., Fu, L., Mei, Y. F., Peng, G., Tan, X., Wang, D. M., Wang, W., Li, Y. G., Physiological testosterone stimulates tissue plasminogen activator and tissue factor pathway inhibitor and inhibits plasminogen activator inhibitor type 1 release in endothelial cells. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 2007, *85*, 246-251.
[43] Glueck, C. J., Friedman, J., Hafeez, A., Hassan, A., Wang, P., Testosterone therapy, thrombophilia, and hospitalization for deep venous thrombosis-pulmonary embolus, an exploratory, hypothesis-generating study. *Med Hypotheses* 2015, *84*, 341-343.
[44] Dieplinger, B., Egger, M., Poelz, W., Gabriel, C., Haltmayer, M., Mueller, T., Soluble ST2 is not independently associated with androgen and estrogen status in healthy males and females. *Clinical chemistry and laboratory medicine* 2011, *49*, 1515-1518.
[45] Guo, W., Bachman, E., Li, M., Roy, C. N., Blusztajn, J., Wong, S., Chan, S. Y., Serra, C., Jasuja, R., Travison, T. G., Muckenthaler, M. U., Nemeth, E., Bhasin, S., Testosterone administration inhibits hepcidin transcription and is associated with increased iron incorporation into red blood cells. *Aging cell* 2013, *12*, 280-291.
[46] Yang, Q., Jian, J., Katz, S., Abramson, S. B., Huang, X., 17beta-Estradiol inhibits iron hormone hepcidin through an estrogen responsive element half-site. *Endocrinology* 2012, *153*, 3170-3178.
[47] Yang, X. S., Huang, D. X., Li, Z. J., The value of radioimmunoassay of myoglobin in the diagnosis of acute myocardial infarction. *Acta cardiologica* 1982, *37*, 441-449.
[48] Yamada, Y., Izawa, H., Ichihara, S., Takatsu, F., Ishihara, H., Hirayama, H., Sone, T., Tanaka, M., Yokota, M., Prediction of the risk of myocardial infarction from polymorphisms in candidate genes. *The New England journal of medicine* 2002, *347*, 1916-1923.
[49] Song, Z., Cai, L., Guo, L., Tsukamoto, Y., Yutani, C., Li, X. A., Accumulation and expression of serum amyloid P component in human atherosclerotic lesions. *Atherosclerosis* 2010, *211*, 90-95.
[50] Chen, D., Xia, M., Hayford, C., Tham el, L., Semik, V., Hurst, S., Chen, Y., Tam, H. H., Pan, J., Wang, Y., Tan, X., Lan, H. Y., Shen, H., Kakkar, V. V., Xu, Q., McVey, J. H., Dorling, A., Expression of human tissue factor pathway inhibitor on vascular smooth muscle cells inhibits secretion of macrophage migration inhibitory factor and attenuates atherosclerosis in ApoE-/- mice. *Circulation* 2015, *131*, 1350-1360.
[51] Shang, Q., Feng, L., Yu, W., Xu, J., Liu, X., Wang, J., [Proteomics study on ficolin 3 in the human plasma of type 2 diabetics]. *Wei sheng yan jiu = Journal of hygiene research* 2016, *45*, 8-13.
[52] Powell, L. W., Seckington, R. C., Deugnier, Y., Haemochromatosis. *Lancet* 2016, *388*, 706-716.
[53] Hagler, L., Askew, E. W., Neville, J. R., Mellick, P. W., Coppes, R. I., Jr., Lowder, J. F., Jr., Influence of dietary iron deficiency on hemoglobin, myoglobin, their respective reductases, and skeletal muscle mitochondrial respiration. *Am J Clin Nutr* 1981, *34*, 2169-2177.
[54] Moreno-Navarrete, J. M., Moreno, M., Puig, J., Blasco, G., Ortega, F., Xifra, G., Ricart, W., Fernandez-Real, J. M., Hepatic iron content is independently associated with serum hepcidin levels in subjects with obesity. *Clinical nutrition* 2016.

[55] Schlater, A. E., De Miranda, M. A., Jr., Frye, M. A., Trumble, S. J., Kanatous, S. B., Changing the paradigm for myoglobin: a novel link between lipids and myoglobin. *Journal of applied physiology* 2014, *117*, 307-315.
[56] Perez-Jimenez, F., Castro, P., Lopez-Miranda, J., Paz-Rojas, E., Blanco, A., Lopez-Segura, F., Velasco, F., Marin, C., Fuentes, F., Ordovas, J. M., Circulating levels of endothelial function are modulated by dietary monounsaturated fat. *Atherosclerosis* 1999, *145*, 351-358.
[57] Berrettini, M., Parise, P., Ricotta, S., Iorio, A., Peirone, C., Nenci, G. G., Increased plasma levels of tissue factor pathway inhibitor (TFPI) after n-3 polyunsaturated fatty acids supplementation in patients with chronic atherosclerotic disease. *Thrombosis and haemostasis* 1996, *75*, 395-400.
[58] Grober, U., Reichrath, J., Holick, M. F., Kisters, K., Vitamin K: an old vitamin in a new perspective. *Dermato-endocrinology* 2014, *6*, e968490.
[59] Citelli, M., Bittencourt, L. L., da Silva, S. V., Pierucci, A. P., Pedrosa, C., Vitamin A modulates the expression of genes involved in iron bioavailability. *Biological trace element research* 2012, *149*, 64-70.
[60] Moro, T., Nakao, S., Sumiyoshi, H., Ishii, T., Miyazawa, M., Ishii, N., Sato, T., Iida, Y., Okada, Y., Tanaka, M., Hayashi, H., Ueha, S., Matsushima, K., Inagaki, Y., A Combination of Mitochondrial Oxidative Stress and Excess Fat/Calorie Intake Accelerates Steatohepatitis by Enhancing Hepatic CC Chemokine Production in Mice. *PLoS One* 2016, *11*, e0146592.
[61] Ricklin, D., Lambris, J. D., Complement in immune and inflammatory disorders: pathophysiological mechanisms. *Journal of immunology* 2013, *190*, 3831-3838.
[62] Kokawa, T., Abumiya, T., Kimura, T., Harada-Shiba, M., Koh, H., Tsushima, M., Yamamoto, A., Kato, H., Tissue factor pathway inhibitor activity in human plasma. Measurement of lipoprotein-associated and free forms in hyperlipidemia. *Arteriosclerosis, thrombosis, and vascular biology* 1995, *15*, 504-510.
[63] Griffin, J. H., Gruber, A., Fernandez, J. A., Reevaluation of total, free, and bound protein S and C4b-binding protein levels in plasma anticoagulated with citrate or hirudin. *Blood* 1992, *79*, 3203-3211.
[64] Liberti, G., Bertina, R. M., Rosendaal, F. R., Hormonal state rather than age influences cut-off values of protein S: reevaluation of the thrombotic risk associated with protein S deficiency. *Thrombosis and haemostasis* 1999, *82*, 1093-1096.
[65] Anderson, H. A., Maylock, C. A., Williams, J. A., Paweletz, C. P., Shu, H., Shacter, E., Serum-derived protein S binds to phosphatidylserine and stimulates the phagocytosis of apoptotic cells. *Nature immunology* 2003, *4*, 87-91.
[66] Triantafilou, M., Hughes, T. R., Morgan, B. P., Triantafilou, K., Complementing the inflammasome. *Immunology* 2016, *147*, 152-164.
[67] Logue, J., Walker, J. J., Colhoun, H. M., Leese, G. P., Lindsay, R. S., McKnight, J. A., Morris, A. D., Pearson, D. W., Petrie, J. R., Philip, S., Wild, S. H., Sattar, N., Scottish Diabetes Research Network Epidemiology, G., Do men develop type 2 diabetes at lower body mass indices than women? *Diabetologia* 2011, *54*, 3003-3006.
[68] Geer, E. B., Shen, W., Gender differences in insulin resistance, body composition, and energy balance. *Gend Med* 2009, *6 Suppl 1*, 60-75.
[69] Feldman, H. A., Longcope, C., Derby, C. A., Johannes, C. B., Araujo, A. B., Coviello, A. D., Bremner, W. J., McKinlay, J. B., Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the

Massachusetts male aging study. *The Journal of clinical endocrinology and metabolism* 2002, *87*, 589-598.
[70] Menni, C., Kiddle, S. J., Mangino, M., Vinuela, A., Psatha, M., Steves, C., Sattlecker, M., Buil, A., Newhouse, S., Nelson, S., Williams, S., Voyle, N., Soininen, H., Kloszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Lovestone, S., Spector, T. D., Dobson, R., Valdes, A. M., Circulating Proteomic Signatures of Chronological Age. *The journals of gerontology. Series A, Biological sciences and medical sciences* 2015, *70*, 809-816.
[71] Kurt, I., Abasli, D., Cihan, M., Serdar, M. A., Olgun, A., Saruhan, E., Erbil, M. K., Chitotriosidase levels in healthy elderly subjects. *Annals of the New York Academy of Sciences* 2007, *1100*, 185-188.
[72] Verma, P., Dalal, K., ADAMTS-4 and ADAMTS-5: key enzymes in osteoarthritis. *Journal of cellular biochemistry* 2011, *112*, 3507-3514.
[73] Zhao, C. Q., Zhang, Y. H., Jiang, S. D., Li, H., Jiang, L. S., Dai, L. Y., ADAMTS-5 and intervertebral disc degeneration: the results of tissue immunohistochemistry and in vitro cell culture. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society* 2011, *29*, 718-725.
[74] Malaguarnera, L., Ohazuruike, L. N., Tsianaka, C., Antic, T., Di Rosa, M., Malaguarnera, M., Human chitotriosidase polymorphism is associated with human longevity in Mediterranean nonagenarians and centenarians. *Journal of human genetics* 2010, *55*, 8-12.
[75] Teumer, A., Qi, Q., Nethander, M., Aschard, H., Bandinelli, S., Beekman, M., Berndt, S. I., Bidlingmaier, M., Broer, L., Group, C. L. W., Cappola, A., Ceda, G. P., Chanock, S., Chen, M. H., Chen, T. C., Chen, Y. D., Chung, J., Del Greco Miglianico, F., Eriksson, J., Ferrucci, L., Friedrich, N., Gnewuch, C., Goodarzi, M. O., Grarup, N., Guo, T., Hammer, E., Hayes, R. B., Hicks, A. A., Hofman, A., Houwing-Duistermaat, J. J., Hu, F., Hunter, D. J., Husemoen, L. L., Isaacs, A., Jacobs, K. B., Janssen, J. A., Jansson, J. O., Jehmlich, N., Johnson, S., Juul, A., Karlsson, M., Kilpelainen, T. O., Kovacs, P., Kraft, P., Li, C., Linneberg, A., Liu, Y., Loos, R. J., Body Composition Genetics, C., Lorentzon, M., Lu, Y., Maggio, M., Magi, R., Meigs, J., Mellstrom, D., Nauck, M., Newman, A. B., Pollak, M. N., Pramstaller, P. P., Prokopenko, I., Psaty, B. M., Reincke, M., Rimm, E. B., Rotter, J. I., Saint Pierre, A., Schurmann, C., Seshadri, S., Sjogren, K., Slagboom, P. E., Strickler, H. D., Stumvoll, M., Suh, Y., Sun, Q., Zhang, C., Svensson, J., Tanaka, T., Tare, A., Tonjes, A., Uh, H. W., van Duijn, C. M., van Heemst, D., Vandenput, L., Vasan, R. S., Volker, U., Willems, S. M., Ohlsson, C., Wallaschofski, H., Kaplan, R. C., Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging cell* 2016, *15*, 811-824.
[76] Johnston, A., Xing, X., Guzman, A. M., Riblett, M., Loyd, C. M., Ward, N. L., Wohn, C., Prens, E. P., Wang, F., Maier, L. E., Kang, S., Voorhees, J. J., Elder, J. T., Gudjonsson, J. E., IL-1F5, -F6, -F8, and -F9: a novel IL-1 family signaling system that is active in psoriasis and promotes keratinocyte antimicrobial peptide expression. *Journal of immunology* 2011, *186*, 2613-2622.
[77] Fosang, A. J., Little, C. B., Drug insight: aggrecanases as therapeutic targets for osteoarthritis. *Nature clinical practice. Rheumatology* 2008, *4*, 420-427.
[78] Evans, D. S., Cailotto, F., Parimi, N., Valdes, A. M., Castano-Betancourt, M. C., Liu, Y., Kaplan, R. C., Bidlingmaier, M., Vasan, R. S., Teumer, A., Tranah, G. J., Nevitt, M. C., Cummings, S. R., Orwoll, E. S., Barrett-Connor, E., Renner, J. B., Jordan, J. M.,

Doherty, M., Doherty, S. A., Uitterlinden, A. G., van Meurs, J. B., Spector, T. D., Lories, R. J., Lane, N. E., Genome-wide association and functional studies identify a role for IGFBP3 in hip osteoarthritis. *Annals of the rheumatic diseases* 2015, *74*, 1861-1867.
[79] Yamaguchi, T., Kanatani, M., Yamauchi, M., Kaji, H., Sugishita, T., Baylink, D. J., Mohan, S., Chihara, K., Sugimoto, T., Serum levels of insulin-like growth factor (IGF); IGF-binding proteins-3, -4, and -5; and their relationships to bone mineral density and the risk of vertebral fractures in postmenopausal women. *Calcified tissue international* 2006, *78*, 18-24.
[80] Rutter, M. M., Markoff, E., Clayton, L., Akeno, N., Zhao, G., Clemens, T. L., Chernausek, S. D., Osteoblast-specific expression of insulin-like growth factor-1 in bone of transgenic mice induces insulin-like growth factor binding protein-5. *Bone* 2005, *36*, 224-231.
[81] Kim, J. G., Sun, B. H., Dietrich, M. O., Koch, M., Yao, G. Q., Diano, S., Insogna, K., Horvath, T. L., AgRP Neurons Regulate Bone Mass. *Cell reports* 2015, *13*, 8-14.
[82] Wendel, M., Sommarin, Y., Heinegard, D., Bone matrix proteins: isolation and characterization of a novel cell-binding keratan sulfate proteoglycan (osteoadherin) from bovine bone. *The Journal of cell biology* 1998, *141*, 839-847.
[83] Argyropoulos, G., Rankinen, T., Neufeld, D. R., Rice, T., Province, M. A., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., Bouchard, C., A polymorphism in the human agouti-related protein is associated with late-onset obesity. *The Journal of clinical endocrinology and metabolism* 2002, *87*, 4198-4202.
[84] Corfitsen, H. T., Drago, A., Insight gained from genome-wide interaction and enrichment analysis on weight gain during citalopram treatment. *Neuroscience letters* 2017, *637*, 38-43.
[85] Waters, D. L., Yau, C. L., Montoya, G. D., Baumgartner, R. N., Serum Sex Hormones, IGF-1, and IGFBP3 Exert a Sexually Dimorphic Effect on Lean Body Mass in Aging. *The journals of gerontology. Series A, Biological sciences and medical sciences* 2003, *58*, 648-652.
[86] Kheiroddin, P., Rasihashemi, S. Z., Estiar, M. A., Mahmudian, B., Halimi, M., Mousavi, F., Nemati, M., Sakhinia, E., RET Gene Analysis in Patients with Medullary Thyroid Carcinoma. *Clinical laboratory* 2016, *62*, 871-876.
[87] Huang, S. P., Levesque, E., Guillemette, C., Yu, C. C., Huang, C. Y., Lin, V. C., Chung, I. C., Chen, L. C., Laverdiere, I., Lacombe, L., Fradet, Y., Chang, T. Y., Lee, H. Z., Juang, S. H., Bao, B. Y., Genetic variants in microRNAs and microRNA target sites predict biochemical recurrence after radical prostatectomy in localized prostate cancer. *International journal of cancer* 2014, *135*, 2661-2667.
[88] Naspi, A., Zingariello, M., Sancillo, L., Panasiti, V., Polinari, D., Martella, M., Rosa Alba, R., Londei, P., IGFBP-3 inhibits Wnt signaling in metastatic melanoma cells. *Molecular carcinogenesis* 2017, *56*, 681-693.
[89] Butt, A. J., Dickson, K. A., McDougall, F., Baxter, R. C., Insulin-like growth factor-binding protein-5 inhibits the growth of human breast cancer cells in vitro and in vivo. *The Journal of biological chemistry* 2003, *278*, 29676-29685.
[90] Spangenburg, E. E., Abraha, T., Childs, T. E., Pattison, J. S., Booth, F. W., Skeletal muscle IGF-binding protein-3 and -5 expressions are age, muscle, and load dependent. *American journal of physiology. Endocrinology and metabolism* 2003, *284*, E340-350.

[91] Kang, J. S., Mulieri, P. J., Hu, Y., Taliana, L., Krauss, R. S., BOC, an Ig superfamily member, associates with CDO to positively regulate myogenic differentiation. *The EMBO journal* 2002, *21*, 114-124.
[92] Lokiec, K., Blonska, A., Walecka-Kapica, E., Stec-Michalska, K., [Effect of treatment with diet on reducing levels of sex hormones in perimenopausal women with overweight and obesity]. *Polski merkuriusz lekarski : organ Polskiego Towarzystwa Lekarskiego* 2016, *40*, 362-368.
[93] Briggs, D. I., Lockie, S. H., Wu, Q., Lemus, M. B., Stark, R., Andrews, Z. B., Calorie-restricted weight loss reverses high-fat diet-induced ghrelin resistance, which contributes to rebound weight gain in a ghrelin-dependent manner. *Endocrinology* 2013, *154*, 709-717.
[94] Gholamhosseini, S., Nematipour, E., Djazayery, A., Javanbakht, M. H., Koohdani, F., Zareei, M., Djalali, M., omega-3 fatty acid differentially modulated serum levels of IGF1 and IGFBP3 in men with CVD: a randomized, double-blind placebo-controlled study. *Nutrition* 2015, *31*, 480-484.
[95] Caliezi, C., Wuillemin, W. A., Zeerleder, S., Redondo, M., Eisele, B., Hack, C. E., C1-Esterase inhibitor: an anti-inflammatory agent and its potential use in the treatment of diseases other than hereditary angioedema. *Pharmacol Rev* 2000, *52*, 91-112.
[96] Andriani, G. A., Almeida, V. P., Faggioli, F., Mauro, M., Tsai, W. L., Santambrogio, L., Maslov, A., Gadina, M., Campisi, J., Vijg, J., Montagna, C., Whole Chromosome Instability induces senescence and promotes SASP. *Sci Rep* 2016, *6*, 35218.
[97] Villeda, S. A., Luo, J., Mosher, K. I., Zou, B., Britschgi, M., Bieri, G., Stan, T. M., Fainberg, N., Ding, Z., Eggel, A., Lucin, K. M., Czirr, E., Park, J. S., Couillard-Despres, S., Aigner, L., Li, G., Peskind, E. R., Kaye, J. A., Quinn, J. F., Galasko, D. R., Xie, X. S., Rando, T. A., Wyss-Coray, T., The ageing systemic milieu negatively regulates neurogenesis and cognitive function. *Nature* 2011, *477*, 90-94.
[98] Lefebvre, J. S., Maue, A. C., Eaton, S. M., Lanthier, P. A., Tighe, M., Haynes, L., The aged microenvironment contributes to the age-related functional defects of CD4 T cells in mice. *Aging cell* 2012, *11*, 732-740.
[99] Monti, D., Ostan, R., Borelli, V., Castellani, G., Franceschi, C., Inflammaging and human longevity in the omics era. *Mechanisms of ageing and development* 2016.
[100] Friedman, J. M., Halaas, J. L., Leptin and the regulation of body weight in mammals. *Nature* 1998, *395*, 763-770.
[101] Baile, C. A., Della-Fera, M. A., Martin, R. J., Regulation of metabolism and body fat mass by leptin. *Annual review of nutrition* 2000, *20*, 105-127.
[102] Spiegelman, B. M., Flier, J. S., Obesity and the regulation of energy balance. *Cell* 2001, *104*, 531-543.
[103] Jequier, E., Leptin signaling, adiposity, and energy balance. *Annals of the New York Academy of Sciences* 2002, *967*, 379-388.
[104] Al-Hamodi, Z., Ismail, I. S., Saif-Ali, R., Ahmed, K. A., Muniandy, S., Association of plasminogen activator inhibitor-1 and tissue plasminogen activator with type 2 diabetes and metabolic syndrome in Malaysian subjects. *Cardiovascular diabetology* 2011, *10*, 23.
[105] Svendsen, O. L., Hassager, C., Christiansen, C., Nielsen, J. D., Winther, K., Plasminogen activator inhibitor-1, tissue-type plasminogen activator, and fibrinogen:

Effect of dieting with or without exercise in overweight postmenopausal women. *Arteriosclerosis, thrombosis, and vascular biology* 1996, *16*, 381-385.
[106] Ahmed, R. L., Thomas, W., Schmitz, K. H., Interactions between insulin, body fat, and insulin-like growth factor axis proteins. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2007, *16*, 593-597.
[107] Ge, H., Gardner, J., Wu, X., Rulifson, I., Wang, J., Xiong, Y., Ye, J., Belouski, E., Cao, P., Tang, J., Lee, K. J., Coberly, S., Gupte, J., Miao, L., Yang, L., Nguyen, N., Shan, B., Yeh, W. C., Veniant, M. M., Li, Y., Baribault, H., Trefoil Factor 3 (TFF3) Is Regulated by Food Intake, Improves Glucose Tolerance and Induces Mucinous Metaplasia. *PLoS One* 2015, *10*, e0126924.
[108] Hertle, E., Stehouwer, C. D., van Greevenbroek, M. M., The complement system in human cardiometabolic disease. *Molecular immunology* 2014, *61*, 135-148.
[109] Zhang, J., Wright, W., Bernlohr, D. A., Cushman, S. W., Chen, X., Alterations of the Classic Pathway of Complement in Adipose Tissue of Obesity and Insulin Resistance. *The FASEB Journal* 2007, *21*, A294-A294.
[110] Xia, J., Zhang, Y., Xin, L., Kong, S., Chen, Y., Yang, S., Li, K., Global Transcriptomic Profiling of Cardiac Hypertrophy and Fatty Heart Induced by Long-Term High-Energy Diet in Bama Miniature Pigs. *PLoS One* 2015, *10*, e0132420.
[111] Zeng, W., Pirzgalska, R. M., Pereira, M. M., Kubasova, N., Barateiro, A., Seixas, E., Lu, Y. H., Kozlova, A., Voss, H., Martins, G. G., Friedman, J. M., Domingos, A. I., Sympathetic neuro-adipose connections mediate leptin-driven lipolysis. *Cell* 2015, *163*, 84-94.
[112] Briot, A., Civelek, M., Seki, A., Hoi, K., Mack, J. J., Lee, S. D., Kim, J., Hong, C., Yu, J., Fishbein, G. A., Vakili, L., Fogelman, A. M., Fishbein, M. C., Lusis, A. J., Tontonoz, P., Navab, M., Berliner, J. A., Iruela-Arispe, M. L., Endothelial NOTCH1 is suppressed by circulating lipids and antagonizes inflammation during atherosclerosis. *The Journal of experimental medicine* 2015, *212*, 2147-2163.

SUPPLEMENTARY INFORMATION

The Supporting Information for **Chapter 4** is available free of charge on the ACS Publications website (https://pubs.acs.org/) at DOI: 10.1021/acs.jproteome.7b00501 (https://pubs.acs.org/doi/abs/10.1021/acs.jproteome.7b00501).

Tables S1–S8
(https://pubs.acs.org/doi/suppl/10.1021/acs.jproteome.7b00501/suppl_file/pr7b00501_si_001.pdf)

Figures S1–S7
(https://pubs.acs.org/doi/suppl/10.1021/acs.jproteome.7b00501/suppl_file/pr7b00501_si_001.pdf)

Supporting Information available
Table S1. Summary of proteins that are associated with females
Table S2. Summary of proteins that are associated with males
Table S3. Summary of proteins that are associated with age
Table S4. Summary of proteins that are associated with body fat mass
Table S5a. Proteins associated with males in MECHE and replicated in DiOGenes
Table S5b. Proteins associated with females in MECHE and replicated in DiOGenes
Table S5c. Proteins associated with age in MECHE and replicated in DiOGenes
Table S5d. Proteins associated with fat mass in MECHE and replicated in DiOGenes
Table S6. Summary of KEGG pathway analysis
Table S7. Summary of Reactome pathway analysis
Table S8. Summary of WikiPathway analysis
Figure S1. Overview of Reactome Pathway Enrichment
Figure S2.a-f. Histogram plots of age and sex per cohort
Figure S3. Venn diagram displaying number of proteins which overlap across phenotypes
Figure S4. OPLS-DA for age of <34yrs vs >35yrs derived from proteomic data of MECHE participants (n=196).
Figure S5. OPLS-DA for fat mass of <25.4kg vs >25.5kg derived from proteomic data of MECHE participants (n=196).
Figure S6. Scatterplot of proteins related to age obtained from regression analysis.
Figure S7. Scatterplot of proteins related to fat mass obtained from regression analysis.