Molecular biomarkers for weight control in obese individuals subjected to a multi-phase dietary intervention

Jennifer Bolton[1,2], Emilie Montastier[1,2,3], Jérôme Carayol[4], Sophie Bonnel[1,2], Lucile Mir[1,2], Marie-Adeline Marques[1,2], Jason Iacovoni[1,2], Nathalie Villa-Vialaneix[5], Armand Valsesia[4], Dominique Langin[1,2,3], and Nathalie Viguerie[1,2]

[1] Institut National de la Santé et de la Recherche Médicale (INSERM), UMR1048, Obesity Research Laboratory, Institute of Metabolic and Cardiovascular Diseases (I2MC), Toulouse, France;

[2] University of Toulouse, UMR1048, Paul Sabatier University, Toulouse, France;

[3] Toulouse University Hospitals, Departments of Endocrinology, Metabolism and Nutrition, Toulouse, France;

[4] Nestlé Institute of Health Sciences SA, EPFL Innovation Park, Bâtiment H, 1015 Lausanne, Switzerland;

[5] Unité de Mathématiques et Informatique Appliquées de Toulouse (MIAT), Université de Toulouse, INRA, Castanet Tolosan, France

***Abbreviated Title*:** Biomarkers of weight control

***Key terms*:** gene expression; weight loss; calorie restriction; adipose tissue

***Word count*:** 3579 (excluding references)

***Number of figures and tables*:** 7

Corresponding author and person to whom reprint requests should be addressed:

Nathalie Viguerie

INSERM, UMR1048, Institut des Maladies Métaboliques et Cardiovasculaires (I2MC)

1 avenue Jean Poulhès

31432 Toulouse, France

Phone: +33561325631

Fax: +33561325600

email: nathalie.viguerie@inserm.fr

Grants

This work was supported by the Innovative Medicines Initiative Joint Undertaking (grant agreement n° 115372), Inserm, Paul Sabatier University and the Commission of the European Communities (FP6-513946 DiOGenes).

Disclosure Statement:

DL is a member of Institut Universitaire de France, JC and AV are employed by Nestlé, the other authors have nothing to disclose.

Clinical Trial Identifier:

NCT00390637

ABSTRACT

Context. While calorie restriction has proven beneficial for weight loss, long-term weight control is variable between individuals.

Objective. We aimed to identify biomarkers of weight control during a dietary intervention (DI), which included 8-weeks of calorie-restriction and 6-months of follow-up.

Design. Adipose tissue (AT) transcriptomes were compared between 21 obese individuals that had either good (maintained weight loss) or poor (regained weight) weight control during the DI. Selected genes were validated on 310 individuals from the same study using RT-qPCR, and protein levels of potential circulating biomarkers were measured by ELISA.

Results. We evaluated 4 genes that had altered expression during the DI, encode secreted proteins, and have not previously been implicated in weight control (*EGFL6*, *FSTL3*, *CRYAB*, *IGFBP3*); as well as 2 genes for which baseline expression was different between those with good and poor weight control (*ASPN*, *USP53*). Changes in plasma concentration of EGFL6, FSTL3, and CRYAB mirrored AT mRNA expression, all decreased during DI in individuals with good weight control. *ASPN* and *USP53* had higher baseline expression in individuals that went on to have good weight control, and eQTL analysis found polymorphisms associated with expression levels of *USP53* in AT. A regulatory network was identified in which TGFβ1 was responsible for down-regulation of certain genes during DI in good-controllers. Interestingly, ASPN is a TGFβ1 inhibitor.

Conclusions. This study found circulating biomarkers associated with weight control, which could serve to adjust weight management strategies, and genes that may be prognostic for successful weight control.

INTRODUCTION

Although obesity may appear to be a simple issue of increased body fat due to excess energy intake, effective guidance for weight control is lacking. Caloric restriction is generally first prescribed to lose weight, however maintenance of weight loss often remains an obstacle. While the extent to which a hypocaloric diet induces weight loss is heterogeneous, subsequent weight control shows even greater inter-individual variation. So far most attempts to predict weight loss during, or weight control after, caloric restriction have failed to provide useful predictive biomarkers (1).

Adipose tissue (AT) plays a pivotal role in obesity-related complications. In addition to storing and releasing excess energy loads, AT also secretes numerous bioactive factors, thus making it a potential source of biomarkers. Nutritional genomics can be used to determine how dietary interventions impact AT, and to identify genes that may cause or contribute to the development of obesity-related disease (2). Gene expression profiling studies in humans to identify transcriptional responses to diet and their molecular targets have shown that weight changes are a major contributor to altered AT gene expression (3-8).

In the present study, we used subcutaneous AT from obese individuals that followed a 8-month dietary intervention (DI), consisting of a 8-week calorie-restriction (CR) diet followed by a 6-month *ad libitum* follow-up. Discovery analyses used transcriptomics to identify genes that were differentially expressed between individuals that successfully maintained weight loss (good-controllers), and those that returned to their baseline weight during follow-up (poor-controllers). Validation analyses using RT-qPCR on a larger cohort confirmed identification of genes that show altered expression in response to weight changes during DI, as well as genes that have different expression levels at baseline, representing markers that are potentially indicative of an individual's ability to successfully maintain weight loss. Finally, for validated genes that encode secreted proteins, plasma protein levels were measured.

MATERIALS & METHODS

Subjects and clinical evaluation

These analyses used samples obtained from the DiOGenes Study (9), all participants signed an informed consent document after verbal and written instructions. As shown in Figure 1a, overweight individuals followed a low-calorie (800-1000 kcal/day) diet for 8 weeks; those that lost at least 8% of their baseline weight were randomized to one of four *ad libitum* follow-up diets or a control diet for 6 months (71% completed). Transcriptome analyses used individuals from the extremes of percentage weight change during DI (exclusion criteria are detailed in Supplementary Figure 1) that were group-matched for baseline age, weight, BMI, waist circumference, blood pressure, and insulin resistance measured by HOMA-IR (homeostatic model assessment index). This resulted in selection of 22 individuals: 11 good-controllers (either maintained weight loss or continued to lose during follow-up) and 11 poor-controllers (regained during follow-up, returning to their baseline weight).

Adipose tissue fractionation and ex vivo cell culture

Abdominal subcutaneous AT was obtained from 7 women (BMI 25.3±4.5 kg/m^2, age 27–50 years) undergoing plastic surgery. The study was approved by the University Hospital of Toulouse ethical committee, and conforms to the Declaration of Helsinki. From each AT sample, 1 gram was flash frozen and stored at -80°C, and 10 grams were digested using collagenase (10), adipocytes were separated from the SVF by washing and centrifugation. For use in gene expression analyses, adipocytes and SVF cells were homogenized in lysis buffer (miRNeasy kit, Qiagen) and stored at −80°C until RNA extraction. For use in secretion analyses, isolated packed adipocytes and SVF cells were maintained *ex vivo* at 37°C in endothelial culture basal medium with 0.1% fatty acid free bovine serum albumin at 2ml (500,000) adipocytes in 10ml medium or 300,000 SVF cells per 1ml medium, respectively. These conditioned media were collected after 24 hours, centrifuged, and stored at -80°C.

Enzyme-linked immunosorbent assays (ELISA)

Protein levels of EGFL6 (csb-el007475hu, Cusabio, Clinisciences, Nanterre, France), FSTL3 (CEK1166, Cohesion Biosciences, Clinisciences), CRYAB (csb-el006008hu, Cusabio, Clinisciences) and IGFBP3 (CEK1195, Cohesion Biosciences, Clinisciences) were measured in duplicate, following manufacturer's instructions.

Gene expression studies

Transcriptome microarray assays

Total RNA was extracted from AT (11) and transcriptomes measured using Agilent Whole Human Genome Microarray 4x44K v2 according to the manufacturer's recommendations (Agilent Technologies, design ID 026652) (6). Arrays were scanned using an InnoScan®710 scanner (Innopsys, Carbonne, France), and images were quantified using MAPIX® v6.5.0 software (Innopsys). Microarray processing included background subtraction, loess intra-array normalization, and Gquantile inter-array normalization in limma (12).

RT-qPCR assays

cDNA was prepared from 500 ng of total RNA and processed using the Biomark™ HD system with 96.96 Dynamic Array IFC (BioMark) and TaqMan assays (Applied Biosystems) as described in (11). Raw data from the default global threshold setting (BioMark Real-time PCR Analysis V4.1.1, Fluidigm) were checked using the graphical representation of plate layout. Duplicate raw Ct values for the same gene were averaged, then relative gene expression was calculated as $2^{-\Delta Ct}$.

Statistical analyses

Clinical characteristics

All analyses used R version 3.2.2. Differences between groups used nonparametric Mann-Whitney U test. Robust mixed ANOVA with bootstrapping (13) applying multiple trimmed group means (default level of 20%) was used to compare within subject changes between groups, reporting the interaction effect.

Adipose tissue transcriptome: differential expression

Principal component analysis (PCA) of transcriptome data identified one individual as an outlier and as such was excluded, analyses included 11 good-controllers and 10 poor-controllers. Differential expression (DE) consisted of 3 separate analyses, comparing $\log_2$ transformed measures between groups at: i) baseline, ii) end of the DI, and iii) intra-individual $\log_2$ fold-changes during DI. Analyses used limma package (13) and all 35,274 spots, because limma relies on the spread of variances pre-filtering is not recommended. Duplicate probes were removed after modeling, keeping that with the smallest p-value (27,385 unique probes), q-values were calculated using the qvalue package (14) on gene-level data (18,568 Entrez Genes). Comparisons of results always represent the same probe.

RT-qPCR validation

RT-qPCR expression was measured for 310 individuals from the DiOGenes study. We applied linear regression models using age, sex, and change in BMI during DI as predictors, and either: i) $\log_2$ transformed baseline expression, or ii) intra-individual $\log_2$ fold-changes during DI, as dependent. For the latter, additional models added baseline BMI or an interaction with baseline BMI . PCA of RT-qPCR expression at end of DI showed that diet had no effect on global gene expression; addition of diet as a random effect to the above models did not alter results. Addition of centre as a random effect did not affect results when using intra-individual changes in expression, nor our top result when using baseline expression.

eQTL analysis

Genome-wide associations were performed using the Illumina 660 chip imputed with European 1000 Genomes (GRCh37) using Minimac3 (15), and $\log_2$ transformed baseline mRNA expression measured by RT-qPCR in 346 individuals. QTL associations between single nucleotide polymorphisms (SNPs) and gene expression used linear mixed models (LMM). Transformed gene expression residuals from regression on age, sex, BMI and center were used as dependent, and individual SNPs as independent variables. GCTA software (16) was used for LMM computation with the 'loco' option to avoid deflation of the test statistics. The Genotype-Tissue Expression (GTEx) Portal (version 4.1, build 201), a database of human genome expression and regulation (17), was used to confirm results, reporting single-tissue eQTL p-values. LocusZoom (18) was used to display regional information of SNPs identified by eQTL analyses. eQTL were considered *cis* if lead SNPs were within 1Mb of the gene, and $p<5E-08$ was considered genome-wide significant.

Pathway analyses

Ingenuity Pathway Analysis (IPA, Qiagen, USA) was used to identify pathways and/or networks, using transcriptome expression ratios obtained from: i) DE results at end of DI for 86 genes that diverged ($q<0.20$ and absolute $\log_2$ expression ratios >0.6), or ii) baseline DE results for 209 genes that classified poor- and good-responders ($p<0.05$) (Figure 1b). The 18,568 genes were used as reference dataset (direct and indirect relationships were permitted), and genes reported as located in "Extracellular Space" were considered to encode secreted factors.

RESULTS

Clinical characteristics of good- and poor-controllers

There was no difference in baseline BMI between good- and poor-controllers ($p=0.504$), nor were there any differences in clinical measures at baseline, although fasting fructosamine was marginally higher in good-controllers ($p=0.078$, Supplementary Table l). Good-controllers lost more weight during CR than poor-controllers (13% *vs.* 9%, $p<0.001$). As expected, at the end of the DI good-controllers showed improved health status, with significant differences in changes of fat mass, waist circumference, SBP, LDL-cholesterol, and C-reactive protein between groups (Table 1).

Altered gene expression in response to the dietary intervention

Discovery analysis using microarrays

We considered genes to have diverged expression if they were differentially expressed at the end of DI, and had differences in $\log_2$ fold-changes between poor- and good-controllers during DI. There were 202 diverged genes with $q<0.20$, of these 27 had $q<0.05$: 22 were down-regulated and 5 were up-regulated in good-controllers, but remained unchanged in poor-controllers (Table 2 & Figure 2a). Of these 27 genes, 6 encoded secreted proteins (*LOXL2*, *IGFBP3*, *HTRA1*, *LEP*, *EGFL6*, *SPARC*).

Pathway analysis

IPA analysis using 86 genes with diverged expression ($q<0.20$ & expression ratio at end of DI>0.6) found a regulatory network centered on *TGFB1*, due to the observed higher expression levels of *LOX*, *LOXL2*, *LAMB3*, *SPARC*, *CCND*1, and *INHBB* in poor-controllers at the end of the DI (Supplementary Figure 2). These genes were up-regulated in poor-controllers and down-regulated in good-controllers during DI (Supplementary Table 3).

Validation analyses using RT-qPCR

We validated 22 out of 24 genes selected from the 202 diverged genes ($q<0.20$), using RT-qPCR . These 24 genes (9 with $q<0.05$, 15 with $0.05<q<0.20$) were selected based on either the largest absolute DE, or potential to encode secreted factors. All 9 genes with $q<0.05$ in DE analyses were replicated by RT-qPCR,

EGFL6, TNMD, CES1, HSPB7, LEP, SPARC, VLDLR, LOXL2 were down-regulated and *IGFBP3* was up-regulated with increased weight loss during DI (Table 2). These genes were also associated with percent weight lost during CR ($p<0.029$), and percent weight change during follow-up ($p<0.005$). Of the 15 genes with $0.05<q<0.20$ in DE analyses, 13 were replicated by RT-qPCR, all were down-regulated with increased weight loss during DI (Table 3). *AES, CCND1, CRYAB, FAM198B, FSTL3, INHBB,* and *LOX* were associated with percent weight lost during CR ($p<0.012$), and all except *FSTL3* were associated with percent weight change during follow-up ($p<0.010$) (Supplementary Table 7). Of these 22 genes, 19 (excluding *TNMD, NOMO1,* and *TPST2*) had significant associations with changes in fat mass, showing consistent directional effects with changes in BMI (Supplementary Table 7).

Figure 2b shows trajectories of expression during DI for the 22 validated genes, plotted by groups of decrease in BMI during the DI (Supplementary Table 2 shows clinical characteristics of the groups). *IGFBP3* had an inverted profile, being up-regulated during CR, then during follow-up had stabilized (higher) expression in individuals that had the greatest decrease in BMI, and was down-regulated in individuals that regained weight. All other genes were down-regulated during CR. During follow-up *LEP, SPARC, HSPB7,CES1, VLDLR, AES,* and *LOX* had stabilized (lower) expression in individuals that had the greatest decrease in BMI, while *EGFL6, TNMD, CRYAB, AKR1C3, FSTL3, FAM198B,* and *MTCH2* had continued down-regulation in individuals that had the greatest decrease in BMI.

Potential secreted biomarkers

To characterize potential circulating biomarkers we measured expression of selected genes in adipocytes and SVF isolated from AT, checked for secretion, and measured secreted factors in plasma from individuals that decreased BMI by >10 or <0, representing the top and bottom 5th percentile of change in BMI (Supplementary Table 2). *EGFL6, TNMD, SPARC, FSTL3,*and *CRYAB* were predominantly or exclusively expressed in adipocytes, whereas *IGFBP3* was predominantly expressed in the SVF (Figure 2c). Regarding secretion, EGFL6, FSTL3 and CRYAB were detected in media from adipocytes but not SVF, IGFBP3 was detected in both media (Table 4). Figure 2d shows changes in plasma concentrations during DI, there were significant decreases in EGFL6 (57%, $p=0.03$), FSTL3 (26%, $p=0.01$) and CRYAB (23%, $p=0.07$) in the group that decreased BMI by >10, and no significant changes in the group with a change in BMI <0. There were no significant changes in plasma IGFBP3 in either group ($p=0.31$). We also found a positive correlation between BMI and FSTL3 levels in media from adipocytes ($r=0.79$, $p<0.05$; Supplementary Figure 3), and plas ma ($r=0.52$, $p<0.05$, not shown).

Differential expression independent of the dietary intervention

Discovery analysis using microarray

We identified 209 genes that were differentially expressed at both baseline and at the end of DI, which we considered to have classified poor- and good-controllers independent of DI, with about half more highly

expressed in either group (Supplementary Table 4).

Validation analyses using RT-qPCR

We selected 17 of the 209 classifier genes for validation by RT-qPCR, based on largest expression ratios at baseline and enriched for genes encoding secreted proteins. Of these, only *ASPN* and *USP53* showed associations with changes in BMI during DI. Higher baseline expression of *ASPN* was associated with a greater decrease in BMI during DI ($p<0.001$), as well as higher baseline BMI ($p<0.001$) and continued weight loss during follow-up ($p=0.001$). When an interaction with baseline BMI was included, baseline expression of *USP53* was found to be associated with changes in BMI during DI ($p=0.012$, interaction $p=0.008$), having higher baseline expression in individuals that lost more weight during DI for individuals with higher baseline BMI. This association was attenuated when adjusted for centre, a potential confounder. There was also a positive association between baseline expression of *USP53* and baseline BMI ($p=0.002$). Figure 3a shows trajectories of expression during DI for *ASPN* and *USP53*, for which baseline expression appears to be associated with weight control after CR.

Properties of the validated classifiers

ASPN was predominantly expressed in the SVF, whereas *USP53* was expressed in both adipocytes and SVF (Figure 3b). Evaluation of secretion of ASPN in AT fractions and plasma failed, as most samples were below the detection limit (8.59 pg/ml).

eQTL analysis

There was a genome-wide significant *cis*-eQTL and an almost significant *trans* eQTL between SNPs and RT-qPCR expression levels of *USP53*. The *cis*-eQTL included SNPs downstream of *USP53* (lead SNP: rs2168987; p=3.1E-08; minor allele frequency=0.44), shown in Figure 3c. The minor T allele of rs2168987 has been previously shown to be associated with higher USP53 expression in AT (FDR<5%) (19). The *trans*-eQTL was on another chromosome, within *ZAK* (MAP3K MLT; lead SNP: rs3769187; p=2.3E-07; minor allele frequency=0.21).

DISCUSSION

We aimed to identify biomarkers of weight control using individuals from the DiOGenes Study, a 2-phase DI including a 8-week CR phase and a 6-month *ad libitum* follow-up. To this end, we used AT transcriptomics to identify genes affected by weight change during the DI, as well as genes that were indicative of successful weight control after CR. We validated our results using RT-qPCR on a larger cohort form the same study, and focused on genes encoding secreted proteins. Discovery analysis compared groups of extreme responders: good-controllers (maintained weight loss during follow-up), and poor-controllers (regained weight during follow-up). Our discovery analyses made use of a small sample size, thus we

applied a relaxed selection criteria and assessed the robustness of the identified genes using a larger replication cohort. We did not adjust for energy intake as data was missing for almost half of the individuals used in these analyses. Diverged genes were altered in response to weight changes during DI; among the 22 validated diverged genes we confirmed 3 as potential circulating biomarkers. Two genes for which baseline expression was indicative of weight control after CR were validated, their expression was not altered in response to the DI.

We have previously shown that AT signatures reflect the capacity to maintain body weight after CR (6), and that genes are generally down-regulated during CR (7,11). Here we found an exception, I*GFBP3* (Insulin-like growth factor binding protein 3) was up-regulated during CR, and subsequently down-regulated with weight gain during follow-up. At the end of DI individuals that decreased BMI by >10 had 56% higher AT expression of *IGFBP3* than those that returned to baseline weight. *IGFBP3* encodes the main insulin-like growth factor transport protein in blood and is known to inhibit adipogenesis (20) and to repress the transforming growth factor β1 (TGFβ1, a secreted cytokine in the TGFβ superfamily) signaling pathway (21). Pathway analysis found a regulatory network controlled by *TGFB1*, although *TGFB1* was not identified as differentially expressed in these analyses. We found that genes identified as diverged during DI were predominantly expressed in adipocytes rather than the SVF. Amongst the best candidate genes was *TNMD*, encoding tenomodulin, a type II transmembrane glycoprotein, whose expression was down-regulated in response to weight loss, and has been positively correlated with BMI (22). TNMD is known to be required for adipocyte differentiation and has been suggested as a protective factor against insulin resistance by promoting hyperplasia and beneficial lipid storage in visceral AT (23). We focused further work on other top candidates (*EGFL6*, *FSTL3* and *CRYAB)* encoding secreted proteins and with little known in the context of weight control.

To evaluate whether these could serve as circulating biomarkers, we compared plasma protein levels from individuals that decreased BMI >10 points and those that returned to baseline BMI at the end of DI. We found significant intra-individual decreases in circulating EGFL6, FSTL3 and CRYAB during DI in individuals with good weight control, but no change in individuals with poor weight control. This was consistent with changes in AT expression of *EGFL6, FSTL3* and *CRYAB* and suggests that secretion from AT likely contributes to plasma levels of these proteins. *EGFL6* encodes epidermal growth factor-like domain multiple-6, a member of the epidermal growth factor repeat superfamily. It has been suggested that this paracrine/autocrine growth factor of AT is an extracellular matrix protein (24). *EGFL6* has previously been shown to have higher AT expression and secretion in obese versus lean individuals, to be down-regulated in obese patients after surgery-induced weight loss, and is potentially involved in the process of AT expansion and the development of obesity (24,25). Here, we showed long-term down-regulation of EGFL6 after CR induced weight loss. *CRYAB*, encoding an α-crystallin B chain, has been previously shown to have a positive association between BMI and AT expression, and increased levels during adipogenesis (26). While EGFL6 and CRYAB are known adipokines, this is the first report of FSTL3 as an adipokine. *FSTL3*, encodes follistatin-like 3, a member of the follistatin (FST)-related protein family (27). FST is known as an adipokine

with reduced expression and secretion in obese versus lean women (28). Both FST and FSTL3 are antagonists of activin and myostatin (29). FSTL3is released by muscles (30) and adipose tissue (27). Here, we found a positive relationship between changes in BMI and *FSTL3* expression in AT, secretion by human adipocytes, and FSTL3 in plasma. Studies on FSTL3 null mice have shown a differential role of FST and FSTL3 on glucose homeostasis and body composition (31). Our observation of decreased *FSTL3* expression and plasma FSTL3 levels with greater decreases in BMI reveals a discrepancy between FST and FSTL3 regarding body weight control.

A remarkable outcome of this study was the identification of genes for which baseline expression was associated with changes in BMI during DI. *ASPN* and *USP53* had higher baseline expression in individuals that exhibited better weight control after CR. *ASPN* (Asporin) had 2-fold higher baseline expression in individuals that went on to decrease BMI by >10 points versus those that returned to baseline weight during follow-up. High *ASPN* expression in AT appears to be a hallmark of individuals that successfully maintained weight loss, as *ASPN* expression was not regulated during CR or follow-up. We also found that *ASPN* was predominantly expressed in the SVF of AT, rather than adipocytes. ASPN belongs to a family of leucine-rich repeat proteins associated with the extracellular matrix and has been found to be expressed in many tissues (32). It has been suggested that extracellular matrix may constrain AT expandability (33), and here higher expression of *ASPN* in AT was relevant for the prevention of weight (re)gain. ASPN is a tumor suppressor and a TGFβ1 inhibitor (34). This corresponds with our gene expression data showing higher baseline expression of *ASPN* associated with a greater decrease in BMI; suggesting that increased inhibition of the TGFβ1 pathway by *ASPN* resulted in increased weight control after CR. *USP53* encodes ubiquitin specific peptidase 53, a tight junction-associated protein (35). The association between weight control and *USP53* was dependent on baseline BMI; here we found that in more obese individuals, higher expression was associated with increased weight loss. *USP53* expression was found to be genetically controlled as well, with both *cis*- and *trans*-eQTL. Our results indicate that AT mRNA levels of *ASPN* and *USP53* might be of interest as prognostic indicators of long-term response to weight reducing diets.

An interesting observation is the implication of TGFβ1, a multifunctional growth factor with pro-fibrotic properties (32), both as a regulator of expression of certain genes that were down-regulated during DI, and as a target for ASPN that had higher expression during DI, in individuals with good weight control after CR. It has been suggested that excess fibrosis in AT may alter tissue remodeling and restrain loss of fat mass (36). The consistency of these observations emphasizes the potential role of AT fibrosis in long-term weight control.

In the present study, we identified a novel adipokine (FSTL3), as well as circulating biomarkers of weight control after CR that are secreted from adipocytes (EGFL6, CRYAB, and FSTL3). We also identified genes for which higher expression was associated with increased weight control after weight loss (*ASPN* and *USP53*). For use as biomarkers, these genes and circulating factors now need to be evaluated in other cohorts.

REFERENCES

1. Stubbs J, Whybrow S, Teixeira P, Blundell J, Lawton C, Westenhoefer J, Engel D, Shepherd R, McConnon A, Gilbert P, Raats M. Problems in identifying predictors and correlates of weight loss and maintenance: implications for weight control therapies based on behaviour change. Obesity reviews : an official journal of the International Association for the Study of Obesity 2011; 12:688-708
2. Bordoni A, Capozzi F. Foodomics for healthy nutrition. Current opinion in clinical nutrition and metabolic care 2014; 17:418-424
3. Capel F, Viguerie N, Vega N, Dejean S, Arner P, Klimcakova E, Martinez JA, Saris WH, Holst C, Taylor M, Oppert JM, Sorensen TI, Clement K, Vidal H, Langin D. Contribution of energy restriction and macronutrient composition to changes in adipose tissue gene expression during dietary weight-loss programs in obese women. The Journal of clinical endocrinology and metabolism 2008; 93:4315-4322
4. Dahlman I, Linder K, Arvidsson Nordstrom E, Andersson I, Liden J, Verdich C, Sorensen TI, Arner P. Changes in adipose tissue gene expression with energy-restricted diets in obese women. The American journal of clinical nutrition 2005; 81:1275-1285
5. Franck N, Gummesson A, Jernas M, Glad C, Svensson PA, Guillot G, Rudemo M, Nystrom FH, Carlsson LM, Olsson B. Identification of adipocyte genes regulated by caloric intake. The Journal of clinical endocrinology and metabolism 2011; 96:E413-418
6. Marquez-Quinones A, Mutch DM, Debard C, Wang P, Combes M, Roussel B, Holst C, Martinez JA, Handjieva-Darlenska T, Kalouskova P, Jebb S, Babalis D, Pfeiffer AF, Larsen TM, Astrup A, Saris WH, Mariman E, Clement K, Vidal H, Langin D, Viguerie N. Adipose tissue transcriptome reflects variations between subjects with continued weight loss and subjects regaining weight 6 mo after caloric restriction independent of energy intake. The American journal of clinical nutrition 2010; 92:975-984
7. Viguerie N, Vidal H, Arner P, Holst C, Verdich C, Avizou S, Astrup A, Saris WH, Macdonald IA, Klimcakova E, Clement K, Martinez A, Hoffstedt J, Sorensen TI, Langin D. Adipose tissue gene expression in obese subjects during low-fat and high-fat hypocaloric diets. Diabetologia 2005; 48:123-131
8. Vink RG, Roumans NJ, Fazelzadeh P, Tareen SH, Boekschoten MV, van Baak MA, Mariman EC. Adipose tissue gene expression is differentially regulated with different rates of weight loss in overweight and obese humans. Int J Obes (Lond) 2016;
9. Larsen TM, Dalskov SM, van Baak M, Jebb SA, Papadaki A, Pfeiffer AF, Martinez JA, Handjieva-Darlenska T, Kunesova M, Pihlsgard M, Stender S, Holst C, Saris WH, Astrup A. Diets with high or low protein content and glycemic index for weight-loss maintenance. The New England journal of medicine 2010; 363:2102-2113
10. Curat CA, Wegner V, Sengenes C, Miranville A, Tonus C, Busse R, Bouloumie A. Macrophages in human visceral adipose tissue: increased accumulation in obesity and a source of resistin and visfatin. Diabetologia 2006; 49:744-747
11. Viguerie N, Montastier E, Maoret JJ, Roussel B, Combes M, Valle C, Villa-Vialaneix N, Iacovoni JS, Martinez JA, Holst C, Astrup A, Vidal H, Clement K, Hager J, Saris WH, Langin D. Determinants of human adipose tissue gene expression: impact of diet, sex, metabolic status, and cis genetic regulation. PLoS genetics 2012; 8:e1002959
12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research 2015; 43:e47
13. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology 2004; 3:Article3
14. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proceedings of the

National Academy of Sciences of the United States of America 2003; 100:9440-9445
15. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, Loh PR, Iacono WG, Swaroop A, Scott LJ, Cucca F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C. Next-generation genotype imputation service and methods. Nature genetics 2016; 48:1284-1287
16. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. American journal of human genetics 2011; 88:76-82
17. The Genotype-Tissue Expression (GTEx) project. Nature genetics 2013; 45:580-585
18. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010; 26:2336-2337
19. Lucas AO. Surveillance of communicable diseases in tropical Africa. International journal of epidemiology 1976; 5:39-43
20. Chan SS, Schedlich LJ, Twigg SM, Baxter RC. Inhibition of adipocyte differentiation by insulin-like growth factor-binding protein-3. American journal of physiology Endocrinology and metabolism 2009; 296:E654-663
21. de Silva HC, Firth SM, Twigg SM, Baxter RC. Interaction between IGF binding protein-3 and TGFbeta in the regulation of adipocyte differentiation. Endocrinology 2012; 153:4799-4807
22. Saiki A, Olsson M, Jernas M, Gummesson A, McTernan PG, Andersson J, Jacobson P, Sjoholm K, Olsson B, Yamamura S, Walley A, Froguel P, Carlsson B, Sjostrom L, Svensson PA, Carlsson LM. Tenomodulin is highly expressed in adipose tissue, increased in obesity, and down-regulated during diet-induced weight loss. The Journal of clinical endocrinology and metabolism 2009; 94:3987-3994
23. Senol-Cosar O, Flach RJ, DiStefano M, Chawla A, Nicoloro S, Straubhaar J, Hardy OT, Noh HL, Kim JK, Wabitsch M, Scherer PE, Czech MP. Tenomodulin promotes human adipocyte differentiation and beneficial visceral adipose tissue expansion. Nature communications 2016; 7:10686
24. Oberauer R, Rist W, Lenter MC, Hamilton BS, Neubauer H. EGFL6 is increasingly expressed in human obesity and promotes proliferation of adipose tissue-derived stromal vascular cells. Molecular and cellular biochemistry 2010; 343:257-269
25. Gerhard GS, Styer AM, Strodel WE, Roesch SL, Yavorek A, Carey DJ, Wood GC, Petrick AT, Gabrielsen J, Ibele A, Benotti P, Rolston DD, Still CD, Argyropoulos G. Gene expression profiling in subcutaneous, visceral and epigastric adipose tissues of patients with extreme obesity. Int J Obes (Lond) 2014; 38:371-378
26. Lehr S, Hartwig S, Lamers D, Famulla S, Muller S, Hanisch FG, Cuvelier C, Ruige J, Eckardt K, Ouwens DM, Sell H, Eckel J. Identification and validation of novel adipokines released from primary human adipocytes. Molecular & cellular proteomics : MCP 2012; 11:M111 010504
27. Allen DL, Cleary AS, Speaker KJ, Lindsay SF, Uyenishi J, Reed JM, Madden MC, Mehan RS. Myostatin, activin receptor IIb, and follistatin-like-3 gene expression are altered in adipose tissue and skeletal muscle of obese mice. American journal of physiology Endocrinology and metabolism 2008; 294:E918-927
28. Flanagan JN, Linder K, Mejhert N, Dungner E, Wahlen K, Decaunes P, Ryden M, Bjorklund P, Arver S, Bhasin S, Bouloumie A, Arner P, Dahlman I. Role of follistatin in promoting adipogenesis in women. The Journal of clinical endocrinology and metabolism 2009; 94:3003-3009
29. Mukherjee A, Sidis Y, Mahan A, Raher MJ, Xia Y, Rosen ED, Bloch KD, Thomas MK, Schneyer AL. FSTL3 deletion reveals roles for TGF-beta family ligands in glucose and fat homeostasis in adults. Proceedings of the National Academy of Sciences of the United States of America 2007; 104:1348-1353
30. Henningsen J, Rigbolt KT, Blagoev B, Pedersen BK, Kratchmarova I. Dynamics of the

skeletal muscle secretome during myoblast differentiation. Molecular & cellular proteomics : MCP 2010; 9:2482-2496

31. Brown ML, Bonomi L, Ungerleider N, Zina J, Kimura F, Mukherjee A, Sidis Y, Schneyer A. Follistatin and follistatin like-3 differentially regulate adiposity and glucose homeostasis. Obesity (Silver Spring) 2011; 19:1940-1949

32. Luo T, Nocon A, Fry J, Sherban A, Rui X, Jiang B, Xu XJ, Han J, Yan Y, Yang Q, Li Q, Zang M. AMPK Activation by Metformin Suppresses Abnormal Extracellular Matrix Remodeling in Adipose Tissue and Ameliorates Insulin Resistance in Obesity. Diabetes 2016; 65:2295-2310

33. Rodriguez A, Ezquerro S, Mendez-Gimenez L, Becerril S, Fruhbeck G. Revisiting the adipocyte: a model for integration of cytokine signaling in the regulation of energy metabolism. American journal of physiology Endocrinology and metabolism 2015; 309:E691-714

34. Maris P, Blomme A, Palacios AP, Costanza B, Bellahcene A, Bianchi E, Gofflot S, Drion P, Trombino GE, Di Valentin E, Cusumano PG, Maweja S, Jerusalem G, Delvenne P, Lifrange E, Castronovo V, Turtoi A. Asporin Is a Fibroblast-Derived TGF-beta1 Inhibitor and a Tumor Suppressor Associated with Good Prognosis in Breast Cancer. PLoS medicine 2015; 12:e1001871

35. Kazmierczak M, Harris SL, Kazmierczak P, Shah P, Starovoytov V, Ohlemiller KK, Schwander M. Progressive Hearing Loss in Mice Carrying a Mutation in Usp53. The Journal of neuroscience : the official journal of the Society for Neuroscience 2015; 35:15582-15598

36. Divoux A, Tordjman J, Lacasa D, Veyrie N, Hugol D, Aissat A, Basdevant A, Guerre-Millo M, Poitou C, Zucker JD, Bedossa P, Clement K. Fibrosis in human adipose tissue: composition, distribution, and link with lipid metabolism and fat mass loss. Diabetes 2010; 59:2817-2825

TABLES

Table 1.

Clinical characteristics of study participants with good- or poor- weight control during DI.

		Good-controllers (n=11)				Poor-controllers (n=10)			
Sex (M/F)		4 / 7				2 / 8			
Diet (n)		1(3), 2(1), 3(2), 5(5)				1(0), 2(4), 3(1), 5(5)			
	n	Mean (sd)			n	Mean (sd)			P*
Age (yr)	11	44.5 (7.6)			10	46.7 (5.2)			0.777
Percentage Weight Lost (%)	11	21.1 (2.0)			10	1.1 (0.7)			<0.001
LCD Weight Lost (kg)	11	12.7 (4.6)			10	8.2 (1.5)			0.002
Percent LCD Weight Lost (%)	11	13.2 (3.5)			10	9.0 (1.1)			<0.001
		Baseline		After DI		Baseline		After DI	P**
BMI (kg/m2)	11	33.0 (2.7)	11	26.0 (2.2)	10	32.2 (3.6)	10	31.9 (3.5)	<0.001
Total Energy (kJ/day)	11	9412 (3804)	5	7056 (2725)	9	8707 (2671)	8	8823 (2517)	0.062
Fat Mass (%)	9	37.5 (6.4)	11	29.3 (5.9)	8	40.1 (8.5)	10	37.9 (10.0)	<0.001
Waist Circumference (cm)	10	103.3 (13.4)	11	86.5 (9.6)	10	102.1 (5.5)	10	99.7 (4.9)	0.028
Waist to Hip Ratio	10	0.90 (0.14)	11	0.87 (0.08)	10	0.92 (0.09)	10	0.90 (0.08)	0.648
SBP (mmHg)	9	124 (11)	11	119 (10)	10	117 (13)	10	123 (12)	<0.001
DBP (mmHg)	9	73 (8)	11	71 (11)	10	69 (10)	10	74(10)	0.072
Fasting Cholesterol (mmol/L)	11	4.97 (0.33)	11	4.73 (0.30)	10	5.07 (0.98)	10	5.50 (0.76)	0.058
Fasting HDL (mmol/L)	11	1.45 (0.3)	11	1.59 (0.31)	10	1.33 (0.42)	10	1.51 (0.31)	0.563
Fasting LDL (mmol/L)	11	3.02 (0.41)	11	2.68 (0.47)	10	3.17 (0.64)	10	3.43 (0.67)	0.031
Fasting TG (mmol/L)	11	1.12 (0.23)	11	1.02 (0.35)	10	1.27 (0.55)	10	1.25 (0.35)	0.058
Fasting Fructosamine (µmol/L)	11	216.3 (16.2)	11	224.6 (14.7)	10	204.8 (15.9)	10	215.6 (15.3)	0.802
Fasting Adiponectin (µg/mL)	11	10.49 (5.34)	11	13.64 (3.73)	10	11.05 (5.18)	10	11.49 (5.51)	0.292
Fasting CRP (mg/L)	10	2.69 (1.24)	10	1.71 (2.34)	10	3.71 (2.86)	9	2.96 (1.38)	0.015

Fasting Glucose (mmol/L)	10	4.99 (0.64)	11	4.52 (0.27)	10	5.15 (0.43)	10	4.98 (0.43)	0.458
Fasting Insulin (µIU/mL)	11	8.11 (4.37)	7	5.98 (4.58)	9	9.23 (2.62)	10	9.85 (5.26)	0.208
HOMA-IR	10	1.96 (1.07)	7	1.19 (0.82)	10	2.27 (0.82)	10	2.12 (1.12)	0.565

* P from Mann-Whitney U test ** P from bootstrapped mixed robust ANOVA (interaction term) testing whether the intra-individual changes in measures between baseline and after the dietary intervention were different between the groups (Mann-Whitney U test comparing the groups at baseline and end of follow-up are available in Supplementary Table 1).

Data are presented as (mean + sd). Groups represent good- and poor- controllers used in microarray analyses.

DI – dietary intervention; BMI – Body mass index; SBP – Systolic blood pressure; DBP – Diastolic blood pressure; HDL – High density lipoprotein; LDL – low density lipoprotein; TG – triglycerides; CRP – C-reactive protein; HOMA-IR- homeostatic model assessment index for insulin resistance

486
487

Table 2.

Differential expression and validation results for 27 genes (q<0.05) with diverged expression associated with changes in BMI during the dietary intervention

	Discovery analyses (n=21)			*Validation analyses (n=310)*	
	ratio after DI *	q-value	mean log_2FC ‡ (poor / good)	*estimate* §	*p-value*
Genes more highly expressed in poor-controllers					
EGFL6 [s]	2.70	0.037	0.16 / -2.54	*0.28*	*<0.001*
TNMD	1.41	0.042	0.41 / -1.00	*0.13*	*<0.001*
CES1	1.31	0.003	0.22 / -1.10	*0.82*	*<0.001*
UCHL1	1.26	0.035	0.04 / -1.22	-	-
HSPB7	1.16	0.011	0.21 / -0.95	*0.81*	*<0.001*
LEP [s]	1.16	0.014	0.17 / -0.99	*0.99*	*<0.001*
TNFRSF25	1.03	0.031	0.42 / -0.61	-	-
SPARC [s]	0.96	0.042	0.16 / -0.80	*5.89*	*<0.001*
ABCC6	0.90	0.015	0.14 / -0.77	-	-
NANOS1	0.90	0.044	0.49 / -0.41	-	-
VLDLR	0.88	0.042	0.28 / -0.59	*0.24*	*<0.001*
SYNPO	0.87	0.050	0.33 / -0.54	-	-
LOXL2 [s]	0.86	0.006	0.32 / -0.54	*0.03*	*0.003*
ASAH1	0.82	0.006	0.10 / -0.73	-	-
MRAS	0.76	0.015	0.29 / -0.48	-	-
VKORC1L1	0.74	0.026	0.25 / -0.49	-	-
GLIPR2	0.73	0.050	0.12 / -0.61	-	-
GPX1	0.72	0.042	0.03 / -0.69	-	-
MSTO1	0.70	0.026	0.24 / -0.46	-	-
CALU	0.64	0.045	0.27 / -0.37	-	-
HTRA1[s]	0.63	0.024	0.12 / -0.51	-	-
LOC729013	0.53	0.031	0.09 / -0.44	-	-
Genes more highly expressed in good-controllers					
EIF4B	-0.47	0.044	-0.09 / 0.38	-	-
BTF3P11	-0.56	0.035	-0.13 / 0.43	-	-
AASS	-0.56	0.035	-0.01 / 0.55	-	-
ADH1B	-0.85	0.015	-0.09 / 0.76	-	-
IGFBP3 [s]	-0.89	0.003	-0.16 / 0.73	*-0.33*	*<0.001*

[S] *genes that encode secreted proteins.*

* ratio of expression at end of dietary intervention = poor-controllers/good-controllers

‡ positive log_2FC means expression increased during dietary intervention (DI)

§ estimate for age and sex adjusted association between change in expression and change in BMI during dietary intervention (change calculated as end of DI - baseline)

Table 3.

Differential expression and validation results for 13 genes (0.05<q<0.20) with validated diverged expression associated with changes in BMI during the dietary intervention

	Discovery analyses (n=21)			*Validation analyses (n=310)*	
	ratio after DI *	q-value	mean log_2FC ‡ (poor / good)	*estimate* §	*p-value*
AKR1C3	0.97	0.077	0.16 / -0.80	*1.11*	*<0.001*
VGLL3[S]	0.86	0.139	0.22 / -0.64	*0.004*	*0.036*
FSTL3[S]	0.85	0.080	0.26 / -0.59	*0.01*	*0.021*
MTCH2	0.81	0.085	0.28 / -0.52	*0.08*	*0.008*
FAM198B	0.72	0.176	0.28 / -0.44	*0.22*	*<0.001*
LOX	0.69	0.189	0.29 / -0.40	*0.09*	*<0.001*
CRYAB[S]	0.63	0.156	0.14 / -0.49	*6.92*	*<0.001*
CCND1	0.62	0.081	0.19 / -0.43	*0.21*	*<0.001*
MECR	0.57	0.080	0.20 / -0.38	*0.02*	*0.001*
TPST2	0.52	0.149	0.08 / -0.45	*0.03*	*0.016*
NOMO1	0.44	0.124	0.14 / -0.30	*0.11*	*<0.001*
INHBB	0.43	0.154	0.25 / -0.19	*0.04*	*0.002*
AES	0.36	0.189	0.18 / -0.18	*0.19*	*0.001*

[S] *genes that encode secreted proteins.*

* ratio of expression at end of dietary intervention = poor-controllers/good-controllers

‡ positive log_2FC means expression increased during dietary intervention (DI)

§ estimate for age and sex adjusted association between change in expression and change in BMI during dietary intervention (change calculated as end of DI - baseline)

Table 4.

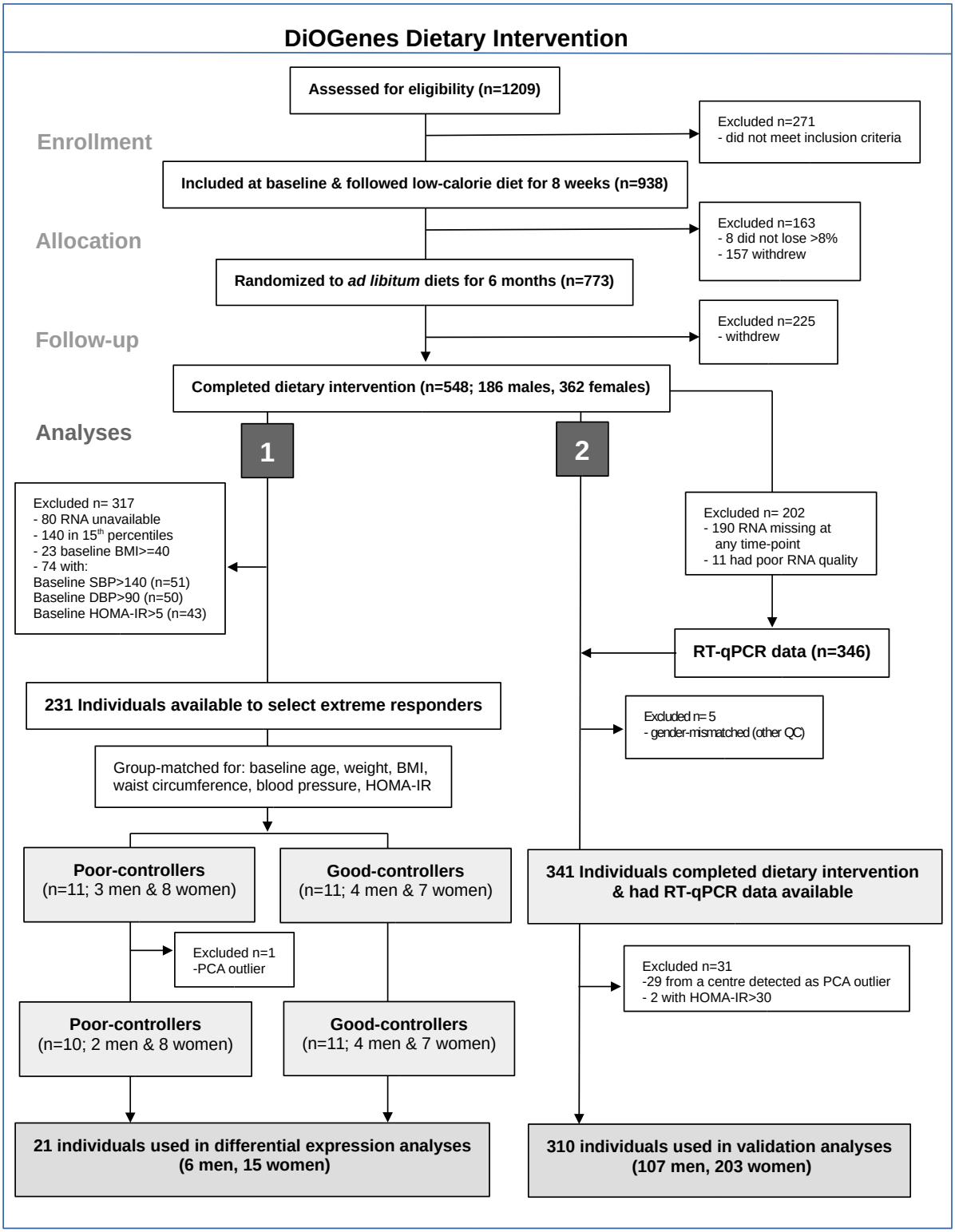
Localization of secretion of protein in adipose tissue fractions

Adipokine	Adipocytes (pg/ml)	SVF (pg/ml)
EGFL6	26.8 ± 1.9	<19.5
FSTL3	25.9 ± 1.9	<10
CRYAB	107.3 ± 31.0	<3.2
IGFBP3	1432 ± 357	451 ± 335

Adipokines concentration was measured in media from isolated adipocytes and stroma-vascular cells (SVF) from human subcutaneous abdominal adipose tissue cultured for 24h (n=7). Data are presented as mean ± sem.

FIGURES

Figure 1. Flowcharts of DiOGenes study design and analyses results.

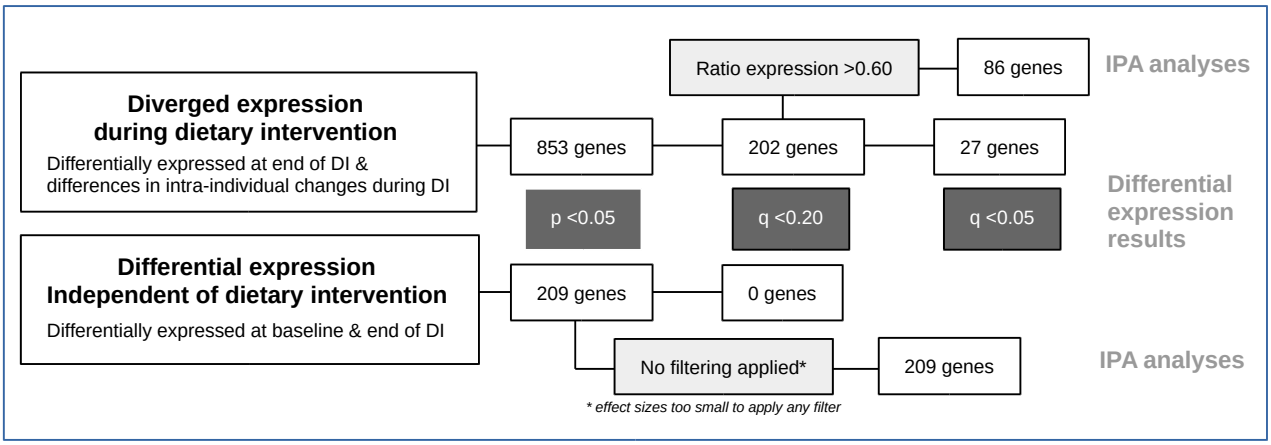


a. Flowchart of the DiOGenes study population used in differential expression (1) and validation (2) analyses

Path 1: Selection of individuals to include in differential expression (transcriptome) analyses of extreme responders.

Path 2: Selection of individuals for use in validation (RT-qPCR) analyses.

RT-qPCR = Reverse transcription quantitative polymerase chain reaction



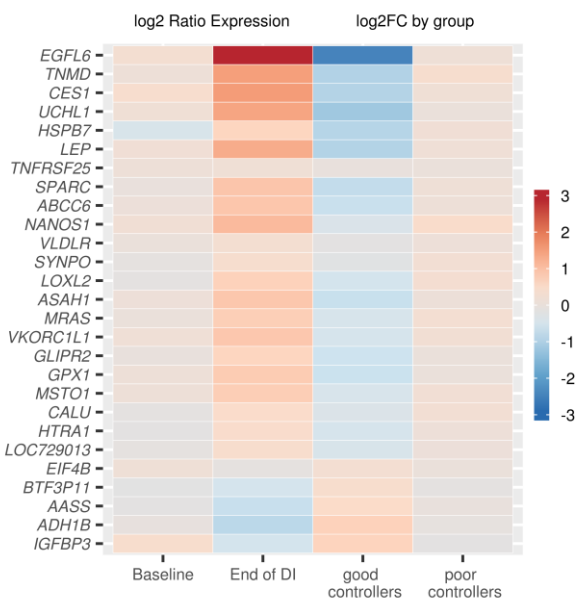
b. Flowchart of number of genes identified based on level of significance applied for each: diverged expression during dietary intervention, and differential expression independent of dietary intervention.

Genes that showed diverged expression or were independent of dietary intervention required the stated level of significance at both time-points.

DI = dietary intervention

IPA = Ingenuity pathway analyses

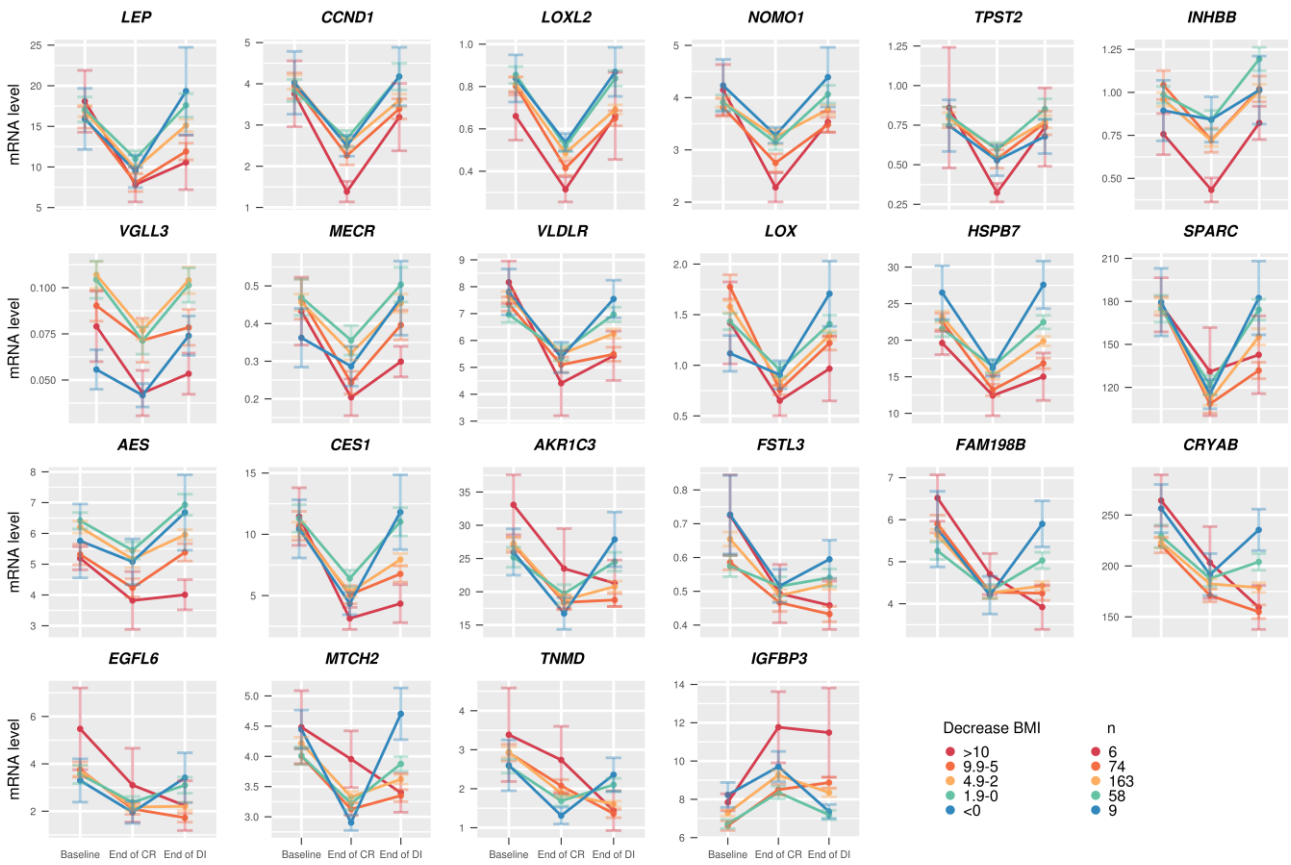
Figure 2. Evaluation of genes identified as having altered expression during the dietary intervention.



a. Heatmap of expression ratios and fold-changes for the 27 genes identified as significantly diverged (q<0.05) during the dietary intervention.

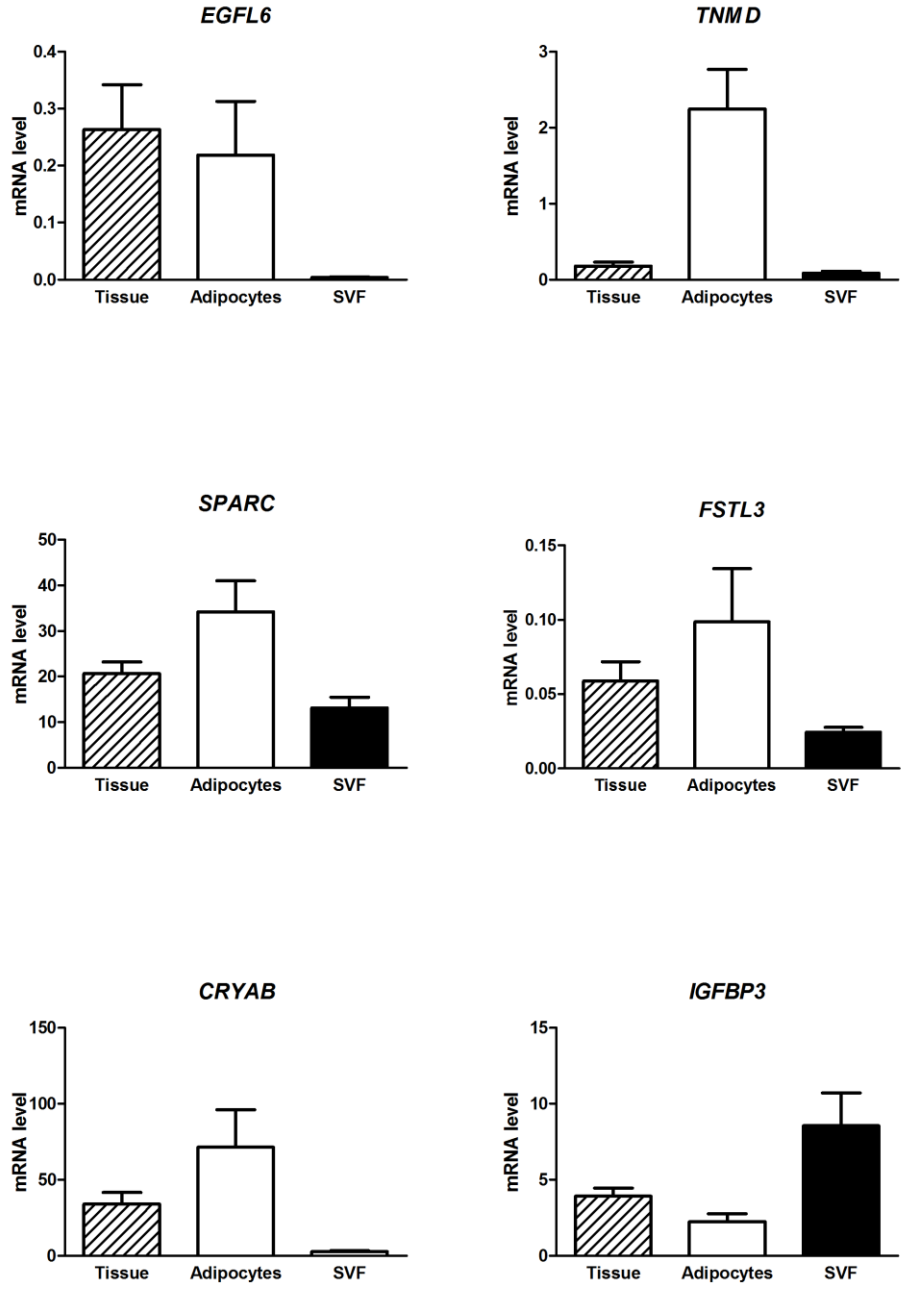
Summary of the top results obtained from differential expression analyses of microarray expression data (n=21). The two columns on the left represent $\log_2$ ratios of expression (poor-controllers / good-controllers) at baseline and after DI. The two columns on the right represent $\log_2$ FC during the dietary intervention for good-controllers and poor-controllers. Legend shows $\log_2$ values.

$\log_2$ FC = $\log_2$ Fold-change.



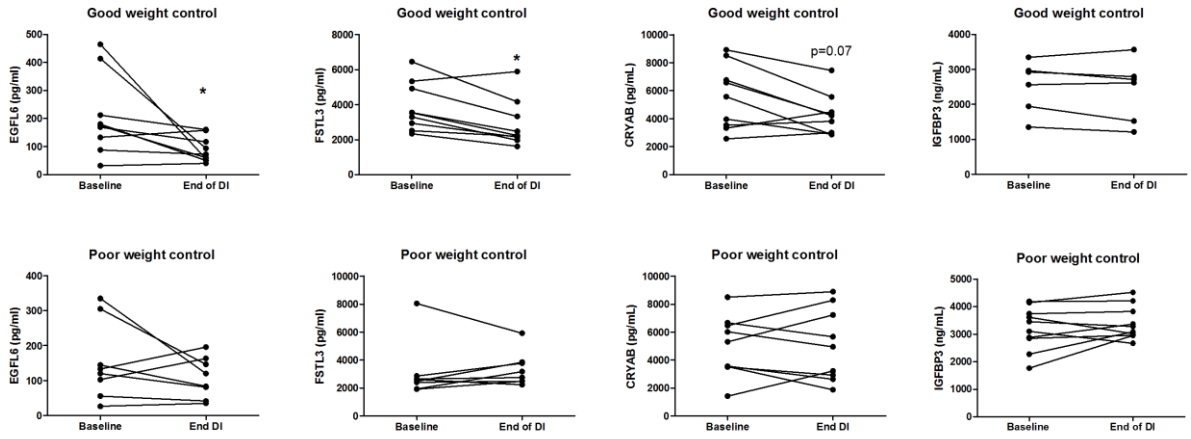
b. Evolution of expression for 22 genes validated by RT-qPCR.

Evolution of relative expression measured by RT-qPCR (n=310) at baseline, end of calorie-restriction (CR), and end of the dietary intervention (DI), grouped by decrease in BMI during the DI, and ordered by patterns of changes in expression. Points represent mean relative expression for each group and bars represent mean +/- sem.



c. Localization of expression in adipose tissue for 6 genes encoding secreted proteins.

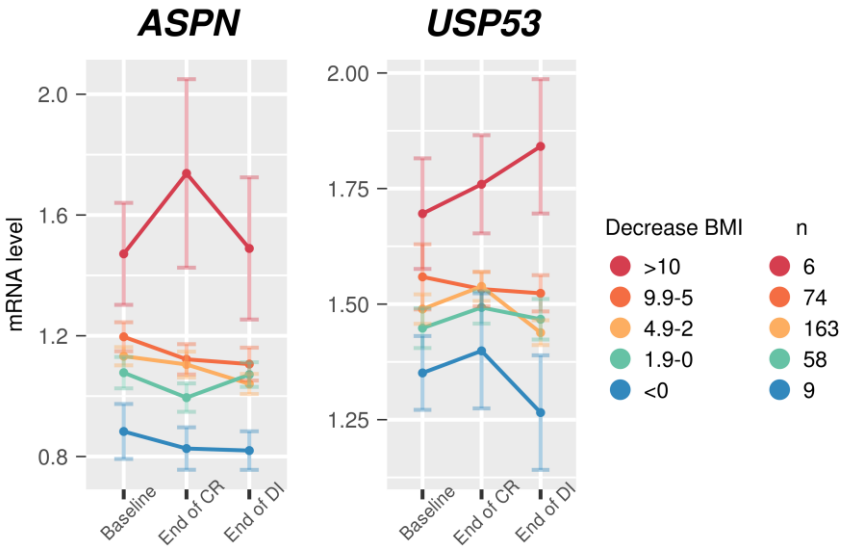
Expression levels of EGFL6, TNMD, SPARC, FSTL3, CRYAB, and IGFBP3 in adipose tissue, adipocytes, and stroma-vascular fraction. mRNA levels were measured in paired samples of freshly isolated adipocytes (n=7) and stroma-vascular fraction (SVF, n=6) from human subcutaneous abdominal adipose tissue (n=5). Data are presented as mean ± sem.



d. Changes in plasma levels for 4 genes encoding secreted proteins.

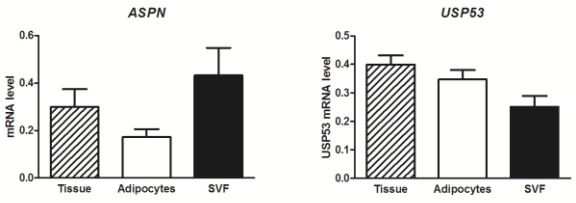
Intra-individual changes in plasma levels of EGFL6, CRYAB, FSTL3 and IGFBP3. Protein levels were measures in plasma samples obtained before and at the end of dietary intervention, for individuals from the top 5 percentiles of changes in BMI during the dietary intervention (poor weight control, n =7- 9; good weight control, n=8-9).

Figure 3. Evaluation of genes identified as classifiers of weight control independent of the dietary intervention.



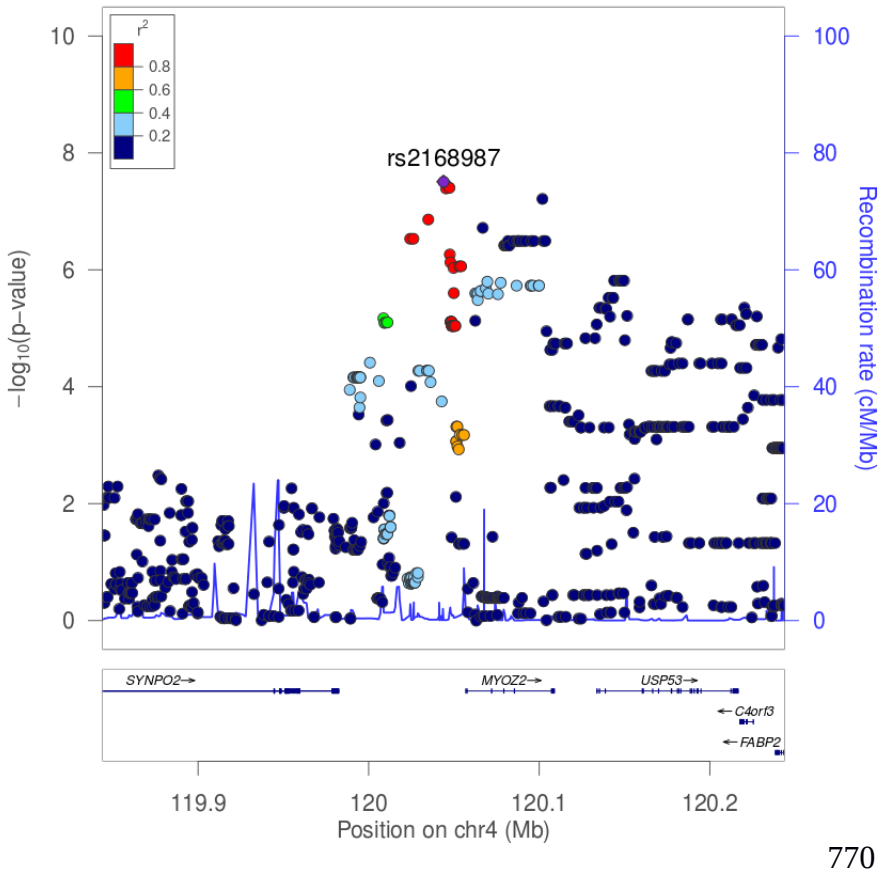
a. Evolution of expression for ASPN and USP53 identified as classifiers at validated by RT-qPCR.

Evolution of relative expression measured by RT-qPCR (n=310) at baseline, end of calorie-restriction (CR), and end of the dietary intervention (DI), grouped by decrease in BMI during the DI. Points represent mean relative expression for each group and bars represent mean +/- sem.



b. Localization of expression in adipose tissue cells for ASPN and USP53.

Expression levels of ASPN and USP53 in adipose tissue, adipocytes, and stroma-vascular fraction (SVF). mRNA level was determined in paired samples of freshly isolated adipocytes (n=7) and SVF (n=6) from human subcutaneous abdominal adipose tissue (n=5). Data are presented as mean ± sem.



c. LocusZoom plot of rs2168987 identified by eQTL analyses to be associated with expression of USP53

Plot showing the lead SNP identified by eQTL analysis as a purple diamond. The y-axis represents -log10 p-values obtained from eQTL analyses, and the points are coloured to represent correlation with the lead SNP. Points in red are interchangeable with the lead SNP, whereas points in blue are independent.