Transcriptome profiling from adipose tissue during a low-calorie diet reveals predictors of weight and glycemic outcomes in obese, nondiabetic subjects

Claudia Armenise,[1] Gregory Lefebvre,[2] Jérôme Carayol,[2] Sophie Bonnel,[3] Jennifer Bolton,[3] Alessandro Di Cara,[1] Nele Gheldof,[2] Patrick Descombes,[2] Dominique Langin,[3,4] Wim HM Saris,[5] Arne Astrup,[6] Jörg Hager,[2] Nathalie Viguerie,[3] and Armand Valsesia[2]

[1]QuartzBio SA, Geneva, Switzerland; [2]Nestlé Institute of Health Sciences, Lausanne, Switzerland; [3]Institute of Metabolic and Cardiovascular Diseases, French National Institute of Health and Medical Research, Paul Sabatier University, UMR1048, Obesity Research Laboratory, University of Toulouse, Toulouse, France; [4]Department of Clinical Biochemistry, Toulouse University Hospitals, Toulouse, France; [5]Department of Human Biology, School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre, Maastricht, Netherlands; and [6]University of Copenhagen, Department of Nutrition, Exercise and Sports, Faculty of Science, Copenhagen, Denmark

ABSTRACT

Background: A low-calorie diet (LCD) reduces fat mass excess, improves insulin sensitivity, and alters adipose tissue (AT) gene expression, yet the relation with clinical outcomes remains unclear.

Objective: We evaluated AT transcriptome alterations during an LCD and the association with weight and glycemic outcomes both at LCD termination and 6 mo after the LCD.

Design: Using RNA sequencing (RNAseq), we analyzed transcriptome changes in AT from 191 obese, nondiabetic patients within a multicenter, controlled dietary intervention. Expression changes were associated with outcomes after an 8-wk LCD (800–1000 kcal/d) and 6 mo after the LCD. Results were validated by using quantitative reverse transcriptase-polymerase chain reaction in 350 subjects from the same cohort. Statistical models were constructed to classify weight maintainers or glycemic improvers.

Results: With RNAseq analyses, we identified 1173 genes that were differentially expressed after the LCD, of which 350 and 33 were associated with changes in body mass index (BMI; in kg/m^2) and Matsuda index values, respectively, whereas 29 genes were associated with both endpoints. Pathway analyses highlighted enrichment in lipid and glucose metabolism. Classification models were constructed to identify weight maintainers. A model based on clinical baseline variables could not achieve any classification (validation AUC: 0.50; 95% CI: 0.36, 0.64). However, clinical changes during the LCD yielded better performance of the model (AUC: 0.73; 95% CI: 0.60, 0.87]). Adding baseline expression to this model improved the performance significantly (AUC: 0.87; 95% CI: 0.77, 0.96; Delong's $P = 0.012$). Similar analyses were performed to classify subjects with good glycemic improvements. Baseline- and LCD-based clinical models yielded similar performance (best AUC: 0.73; 95% CI: 0.60, 0.86). The addition of expression changes during the LCD improved the performance substantially (AUC: 0.80; 95% CI: 0.69, 0.92; $P = 0.058$).

Conclusions: This study investigated AT transcriptome alterations after an LCD in a large cohort of obese, nondiabetic patients. Gene expression combined with clinical variables enabled us to distinguish weight and glycemic responders from nonresponders. These potential biomarkers may help clinicians understand intersubject variability and better predict the success of dietary interventions. This trial was registered at clinicaltrials.gov as NCT00390637.

Am J Clin Nutr 2017;106:736–46.

Keywords: obesity, insulin resistance, low-calorie diet, transcriptome analysis, adipose tissue

Supported by the European Commission, Food Quality and Safety Priority of the Sixth Framework Program (FP6-2005-513946 and ESGI 262055 DIOGenATS), and Nestlé Institute of Health Sciences.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplemental Figures 1 and 2 and Supplemental Tables 1–11 are available from the "Online Supporting Material" link in the online posting of the article and from the same link in the online table of contents at http://ajcn.nutrition.org.

CA and AV contributed equally to this work.

Address correspondence to AV (e-mail: armand.valsesia@rd.nestle.com).

Abbreviations used: AT, adipose tissue; *CAV2*, caveolin 2; *C7*, complement component 7; *ELOVL5*, ELOVL fatty acid elongase 5; *ENSG*, Ensembl gene identifier; eQTL, expression quantitative trait loci; FDR, false discovery rate; *GSDMB*, gasdermin B; LCD, low-calorie diet; *LEP*, leptin; *LOX*, lysyl oxidase; *ME1*, malic enzyme 1; *MTCH2*, mitochondrial carrier 2; *NPY1R*, neuropeptide Y receptor Y1; *PCK2*, phosphoenolpyruvate carboxykinase 2; *PPAP2A*, phosphatidic acid phosphatase type 2A; RNAseq, RNA sequencing; ROC, receiver operating characteristic; RSD, relative SD; RT-qPCR, reverse transcription quantitative polymerase chain reaction; SNP, single nucleotide polymorphism; *SPARC*, secreted protein acidic and cysteine rich; SVF, stromal vascular fraction; T2D, type 2 diabetes.

Received March 3, 2017. Accepted for publication June 21, 2017.

First published online August 9, 2017; doi: https://doi.org/10.3945/ajcn.117.156216.

INTRODUCTION

Obesity is characterized by an excess of fat mass that affects adipocyte metabolism and associates with comorbidities such as cardiovascular disease, insulin resistance, type 2 diabetes (T2D), and cancer (1–3). Dietary interventions aim to reduce fat mass, restore normal adipose tissue (AT) function, and improve dysfunctions linked with the metabolic syndrome (4, 5). Yet, high variability is observed in the capacity to lose and maintain weight (6). Weight loss generally leads to an improvement of glycemic control (4, 5), but this corresponds to an average improvement and may not be true for each subject. We demonstrated that among subjects with strong weight loss (>8% of initial body mass), only half of these subjects had sustainable glycemic control improvements (7). The plasticity of AT has been studied in response to weight loss through expression studies either for candidate genes with a large sample size or at the transcriptome-level with a small sample size (8–14). However, to date our understanding remains limited regarding the link between AT gene expression changes, and weight and glycemic outcomes of a low-calorie diet (LCD). We investigated transcriptome-wide differences between 20 weight maintainers (0–10% weight regained) and 20 weight regainers (>50% weight regained) and highlighted differences in fatty acid metabolism, citric acid cycle, oxidative phosphorylation, and apoptosis pathways (14). By design, the scope was restricted to differences between extreme responders. Therefore, our knowledge of subjects with intermediate responses remains insufficient for stratification purposes and underscores the need for biomarkers (7, 15). In addition, although the role of the β cell dysfunction and peripheral insulin resistance are well established in T2D development (16), the contribution of AT to T2D remains unclear. Recently, a measure of AT insulin resistance was shown to be a prognostic marker of insulin resistance and T2D progression (17), indicating that the progressive decline in β cell function correlates with a progressive increase in AT insulin resistance and dyslipidemia. This finding further emphasizes the importance of AT plasticity in response to LCD interventions. In this report, we studied the transcriptome of overweight and obese, nondiabetic subjects, from the DiOGenes (Diet, Obesity and Genes) study. The subjects followed an 8-wk LCD and a 6-mo weight-maintenance diet intervention (18). We analyzed a large RNA sequencing (RNAseq) dataset (191 subjects × 2 time points) from AT biopsies and identified genes that were differentially expressed and associated with clinical outcomes after an LCD. The expression changes of these genes were then validated by using RT-qPCR in a larger sample size (350 subjects × 2 time-points). Finally, we investigated whether these gene expression changes could distinguish responders from nonresponders 6 mo after LCD.

METHODS

Ethics

The study was performed according to the latest version of the Declaration of Helsinki. Local ethics committees approved of all procedures that involved human participants, and written, informed consent was obtained from all participants.

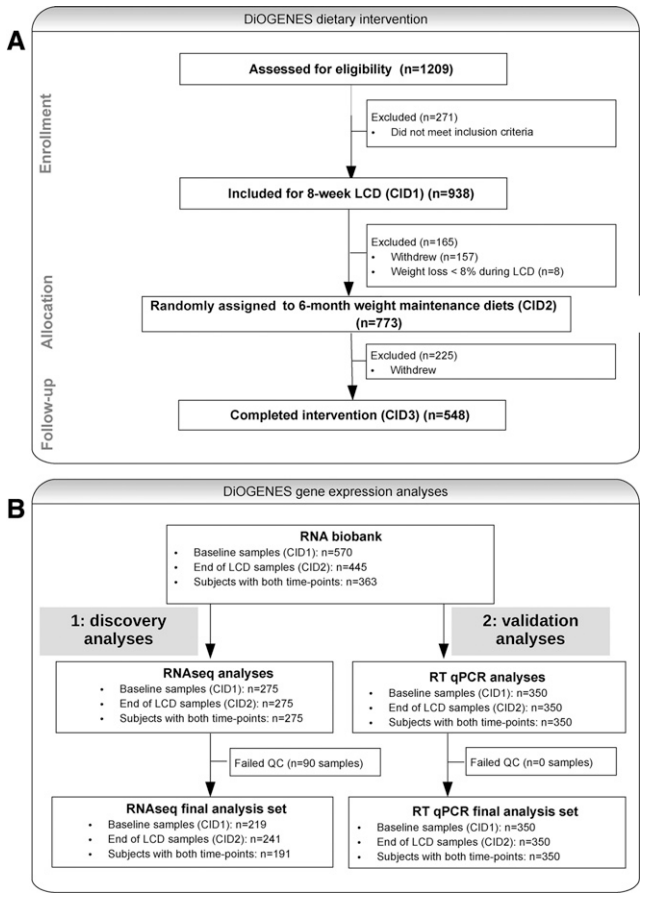

FIGURE 1 Flowcharts of DiOGENES intervention and expression studies. (A) Flowchart of the DiOGENES dietary intervention. (B) Flowchart of the DiOGENES gene expression analyses (path 1: discovery analyses with the use of RNAseq; path 2: validation analyses with the use of RT-qPCR). CID, clinical investigation day; DiOGENES, Diet, Obesity and Genes; LCD, low-calorie diet; QC, quality control; RT-qPCR, reverse transcription quantitative polymerase chain reaction; RNAseq, RNA sequencing.

Study design

The DiOGenes study is a pan-European, multicenter, randomized controlled dietary intervention program (NCT00390637). The study was conducted in 8 European countries: Netherlands, Denmark, United Kingdom, Greece, Bulgaria, Germany, Spain, and Czech Republic. The study has been described in detail previously (18, 19), and a summary of the study is shown in **Figure 1**A. Briefly, overweight or obese nondiabetic subjects followed an LCD for 8 wk. The LCD provided 800 kcal/d with the use of a meal-replacement product (Modifast, Nutrition et Santé France). Participants could also eat ≤400 g vegetables/d (corresponding to a maximal addition of 200 kcal/d). Subjects who lost ≥8% of weight were then included in a 6-mo weight-maintenance diet phase and followed one of 5 diets (a 2×2 factorial combination diet of low and high protein and glycemic index or a control diet).

In total, 938 participants were included for the LCD intervention, and 773 (82%) achieved the 8% weight loss and were included in the weight-maintenance diet. Among those 773 participants, 548 (71%) completed this 6-mo weight-maintenance intervention. The dropout rates (18% for the LCD and 29% for weight maintenance) have been discussed in detail previously (18) and are in line with reports from other dietary studies (20–22).

Gene expression analyses by using RNAseq

Total RNA was extracted from abdominal subcutaneous AT needle biopsies as previously described (23). Gene expression was then examined by using 100-nucleotide long paired-end RNA sequencing with an Illumina HiSeq 2000 of libraries prepared by using the Illumina TruSeq kit following the manufacturer's standard protocols. Sequencing was performed for a total of 550 samples corresponding to 275 subjects having both the baseline and after-weight-loss time points (Figure 1B). Demultiplexing was carried out with Casava (24); the resulting FASTQ files were then mapped onto the human genome (GRCh37 assembly) with RNA-STAR (25) with the use of default parameters. Sequencing quality was evaluated by using FastQC (26). Mapping quality was assessed by using Rsamtools (27). Multimapper reads, reads with >5 mismatches, as well as PCR duplicates, were subsequently discarded.

For each sample, the number of reads mapping onto genes was retrieved by using GenomicAlignments (28). Annotation was performed by using 64,102 genes from the GRCh37.75 assembly generated with the use of the AnnotationDbi R package (29). Only reads with both ends mapping onto one single gene were considered.

Samples were discarded if they had unusually low sequencing depths (<1 million reads), unusually low mapping percentages (<60%), unusually low final mapping percentages (<40%), highly variable sequencing depths across replicates (SD >5 million reads), and highly variable sequencing quality across replicates (SD >5). On feature annotation, samples were discarded if they had unusually low percentages of reads annotated to genes (<75%) and samples with unusually few detected genes (<25,000). Principal component analyses on expression levels were also performed to exclude outlier samples and did not reveal sex or country effects. After all QC steps, 191 subjects remained with data for both time points. Expression data are available from the Gene Expression Omnibus under accession GSE95640.

Validation by using RT-qPCR assays

Complementary DNA was prepared from 500 ng total RNA and processed by using the Biomark HD system with 96.96 Dynamic Array IFC (BioMark) and TaqMan assays (Applied Biosystems) following the protocol described by Viguerie et al. (23). Raw data were obtained by using the Fluidigm software (Real-time PCR Analysis V4.1.1; BioMark) and were checked by using the graphical representation of plate layout. Duplicate gene raw Ct values were averaged and then normalized to a reference gene as described by Viguerie et al. (23). Relative gene expression (30) was calculated as $2^{-\Delta CT}$. Of the 29 genes identified in RNAseq analyses to be associated with BMI (in kg/m^2) and Matsuda outcomes, 4 genes [Ensembl gene identifier (*ENSG*) *00000230058*, *ENSG00000236404*, *ENSG00000264968*, *ENSG00000266968*] did not have any available Taqman probes and could not be tested. During quality-control checks, 7 genes [aldo-keto reductase family 1 member C2 (*AKR1C2*), keratin 5 (*KRT5*), Rieske Fe-S domain containing (*RFESD*), fibulin 7 (*FBLN7*), myosin heavy polypeptide 1 (*MYH1*), protein tyrosine phosphatase, receptor type T (*PTPRTI*) , and spermatogenesis associated 9 (*SPATA9*)] were found with very low expression amounts. Their respective Ct values were close to detection limits, indicating the measurements were not reliable. Those 7 genes were discarded from further analyses. Principal component analysis of the expression fold-change on the quality-controlled dataset did not identify any outliers and did not reveal sex or country effects.

Expression amounts in AT fractions

The same 18 genes that were successfully quantified in the RT-qPCR validation phase were also investigated in different AT fractions [isolated adipocytes, stromal vascular fraction (SVF), and total abdominal subcutaneous AT] from 7 women [age: 27–50 y, BMI (mean ± SEM): 25.3 ± 4.5] undergoing plastic surgery at the University Hospital of Toulouse. Gene expression was measured by using the same TaqMan assays mentioned above. The study was approved by the ethical committee of the University Hospital of Toulouse and conformed to the Declaration of Helsinki.

Statistical analyses

Differential expression analyses

Analysis of differentially expressed genes during the LCD was performed by using *DESeq2* (31). Analyses were adjusted for sex, age, and center. Adjustment for multiple testing was performed by using the Benjamini-Hochberg (32) false discovery rate (FDR) method. Genes were considered significantly differentially expressed if their FDR-adjusted *P* value <10% and their absolute $\log_2$ fold-change was ≥0.2.

Functional analyses

Pathway and enrichment analyses were performed by using the EnrichR tool (33). Statistical significance was set at an FDR-adjusted *P* value <5%. Network-based gene set enrichment analyses were also performed by using the EnrichNet tool (34).

Association with clinical outcomes

Genes differentially expressed were tested for association with clinical endpoints by using linear mixed effect models adjusting for sex and age as fixed effects and center as a random effect. For RNAseq data (discovery analyses), normalized gene counts were obtained from the *DESeq2* package (31) after running the differential expression analysis. For qRT-PCR data (validation analyses), the difference in relative gene expression value was used ($2^{-\Delta\Delta Ct}$). For discovery analyses, FDR thresholds were arbitrarily set at 5% for BMI outcomes and 10% for Matsuda outcomes. For validation analyses, statistical significance was set at an FDR-adjusted *P* value <5%.

Expression analysis of AT fractions

The difference in expression pattern between isolated adipocytes and SVFs was assessed by using a Kruskal-Wallis paired test. Adjustment for multiple testing was performed by using the Benjamini-Hochberg (32) procedure.

Expression quantitative trait loci analyses

Genotyping was performed by using Illumina 660 single nucleotide polymorphism (SNP) arrays, followed with imputation

based on the 1000 Genome European reference panel (35) and by using Minimach3 (36). The principal component analyses on genotype data did not reveal any strong population stratification. Linear mixed models were used to test quantitative trait loci associations between SNPs and $\log_2$-transformed values of fragments per kilobase of transcript per million mapped reads. Expression was adjusted for age, sex, BMI, and center. Analyses were performed for baseline expression amounts (adjusted for baseline BMI) and change in expression amounts during the LCD (adjusted for BMI changes). The GCTA software (37) was used for linear mixed model computation with the "loco" option. *cis*–Expression quantitative trait loci (eQTL) effects were tested by investigating SNPs within 1 Mb of each gene. Adjustment for multiple testing was performed by using the simpleM method (38). Significance levels were set at an adjusted α of 5%.

Predictive analyses

These analyses aim to construct models to distinguish (classify) 2 classes of patients: those with BMI improvement better than the median population's improvement compared with the remaining patients having less improvement. This improvement was defined as the percentage of changes after the weight-maintenance diet phase (6-mo after the LCD) compared with baseline. Similar analyses were also performed for Matsuda index improvement, with the 2 classes defined by patients having a Matsuda index improvement higher or lower than the median population's improvement.

In these analyses, several types of models were tested:

- A clinical model based on baseline BMI, Matsuda index, sex, and age;
- A model including variables from the clinical model together with baseline gene expression;
- A model including variables from the best clinical model, with gene expression changes during LCD;
- A model including variables from the best clinical model, with baseline gene expression; and
- A model based on baseline clinical variables and baseline gene expression.

The RT-qPCR dataset (n = 350 subjects) was randomly separated into a training set (for model fitting) and a validation set (to assess the model performance). The training dataset included 75% of the whole RT-qPCR dataset. Data were standardized (mean-centered and scaled to unit variance), and training was performed by using 10-fold cross-validation repeated 5 times. In the training phase, the best classification model was identified based on AUC from the receiving operating characteristic (ROC) curve. Models were fitted by using a generalized linear model with the binomial family distribution. Model training and evaluation were performed by using the caret framework (39). Results from the model evaluation on the validation set are presented in the form of ROC curves. Superiority of a model compared with another was tested with a one-sided Delong test (40), a nonparametric method used to compare ROC AUCs. Simplification of the models was attempted by using the recursive feature elimination. Values are expressed as means ± SEMs.

RESULTS

Baseline characteristics and overall clinical outcome

Baseline characteristics of the participants are described in **Table 1**. Participants were on average 41 y old, had a baseline BMI of 34.62 ± 0.22, and a baseline Matsuda index of 4.61 ± 0.22.

After the LCD intervention, participants lost on average 10.82% of their initial body weight (95% CI: 10.33, 11.38) and had glycemic control improvements (mean Matsuda index change: +2.29; 95% CI: +1.85, +2.73). A modest, yet significant, correlation was found between weight loss and glycemic improvements (Pearson's $r = -0.16$; 95% CI: −0.31, −0.01; $P = 0.036$). In addition, moderate variability was observed in weight loss [relative SD (RSD) = 25% for BMI changes], but high variability was found in glycemic improvements (RSD = 120% for Matsuda changes). **Figure 2**A illustrates the variability in clinical outcomes and shows that ∼15% of the subjects had no improvements in glycemic control (changes in Matsuda index < 0), despite a weight loss >10% of initial body weight.

Differential expression analyses during the LCD and association with clinical outcomes

We assessed gene expression changes in AT biopsies during the LCD intervention. We identified 1173 genes differentially expressed, of which 350 were significantly associated with changes in BMI (**Supplemental Table 1**) and 33 were associated with changes in Matsuda index (**Supplemental Table 2**). Interestingly, despite the modest correlation between BMI and Matsuda index changes, nearly all (29 of 33) of the genes

TABLE 1
Subject characteristics[1]

	All (n = 191)	Women (n = 121)	Men (n = 70)	P[2]
Baseline				
Age, y	41.63 ± 0.44	41.24 ± 0.56	42.31 ± 0.69	0.2271
BMI, kg/m^2	34.62 ± 0.33	34.89 ± 0.43	34.14 ± 0.53	0.2687
Weight, kg	99.49 ± 1.22	95.07 ± 1.34	107.13 ± 2.10	<0.001
Matsuda index	4.61 ± 0.22	5.01 ± 0.29	3.97 ± 0.33	0.0207
Weight loss on LCD, kg	−10.86 ± 0.26	−9.91 ± 0.27	−12.51 ± 0.50	<0.001
Weight loss on LCD, %	−10.82 ± 0.19	−10.36 ± 0.21	−11.61 ± 0.36	0.0036
Change in Matsuda index on LCD	2.29 ± 0.22	2.07 ± 0.25	2.66 ± 0.43	0.2283

[1] Values are means ± SEs. LCD, low-calorie diet.
[2] t Test comparison of the differences between men and women.

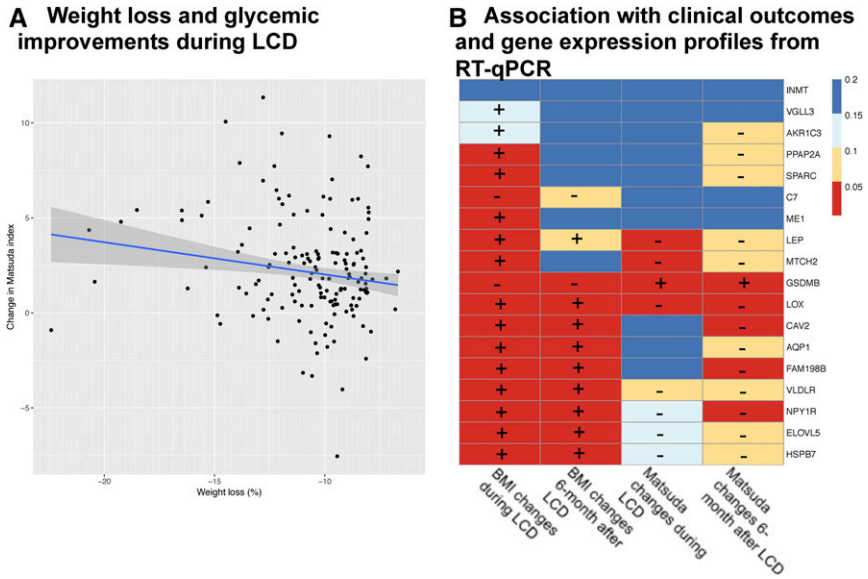

FIGURE 2 Results summary from the RT-qPCR analyses. (A) Correlation between weight loss and glycemic improvement changes during the LCD (Pearson's $r = -0.16$; 95% CI: -0.31, -0.01; $P = 0.036$). About 15% of the subjects had a change in the Matsuda index of <0. The fitted line corresponds to a robust linear regression, with the gray area representing its SE. (B) Processed summary of the association between gene expression changes during the LCD and clinical outcomes during the LCD or during the 6-mo weight-maintenance diet phase. Association was tested by using linear mixed effect models. Color indicates statistical significance, with FDRs below a given threshold. For example, red shows associations with FDR <0.05. For associations having FDR <15%, the sign of correlation between expression changes and clinical outcomes is indicated. AKR1C3, aldo-keto reductase family 1 member C3; AQP1, aquaporin 1; CAV2, caveolin 2; C7, complement component 7; ELOVL5, ELOVL fatty acid elongase 5; FAM198B, family with sequence similarity 198 member B; FDR, false discovery rate; GSDMB, gasdermin B; HSPB7, heat shock protein family B (small) member 7; INMT, indolethylamine N-methyltransferase; LCD, low-calorie diet; LEP, leptin; LOX, lysyl oxidase; ME1, malic enzyme 1; MTCH2, mitochondrial carrier 2; NPY1R, neuropeptide Y receptor Y1; PPAP2A, phosphatidic acid phosphatase type 2A; RT-qPCR, reverse transcription quantitative polymerase chain reaction; SPARC, secreted protein acidic and cysteine rich; VGLL3, vestigial like family member 3; VLDLR, very low density lipoprotein receptor.

associated with changes in Matsuda index were also associated with changes in BMI. **Table 2** provides a list of these 29 genes, including their $\log_2$ fold-change in expression. Most of these genes were downregulated during the LCD, except 5 genes that were upregulated [*FBLN7*, gasdermin B (*GSDMB*), complement component 7 (*C7*), *ENSG00000264968*, and *PTPRT*].

Pathway and enrichment analyses

A pathway analysis identified 4 KEGG pathways significantly enriched in BMI hits (**Supplemental Table 3**) pertaining to fatty acid metabolism [including peroxisome proliferator-activated receptor (*PPAR*) signaling] and to *AMPK* signaling. Enrichment in the fatty acid metabolism pathway was defined by 13 genes [acetyl-CoA carboxylase alpha (*ACACA*), acyl-CoA dehydrogenase, long chain (*ACADL*), acyl-CoA oxidase 1 (*ACOX1*), acyl-CoA synthease long-chain familiy member 1 (*ACSL1*), diazepam binding inhibitor (*DBI*), ELOVL fatty acid elongase 5 (*ELOVL5*), fatty acid desaturase 2 (*FADS2*), malic enzyme 1 (*ME1*), mitochondrial trans-2-enoyl-CoA reductase (*MECR*), phosphoenolpyruvate carboxykinase 2 (*PCK2*), protein tyrosine phosphatase-like B (*PTPLB*), stearoyl-CoA desaturase (*SCD*), and solute carrier family 27 member 2 (*SLC27A2*)]. All these genes, except *PCK2*, were downregulated during the LCD. The *AMPK* pathway was mostly defined by *ACACA*, cyclin D1 (*CCDN1*), fructose-bisphosphatase 1 (*FBP1*), glycogen synthase (*GYS*) 1, *GYS2*, leptin (*LEP*), *PCK2*, protein phosphatase 2 scaffold subunit abeta (*PPP2R1B*), and *SCD* genes, with all genes except *PCK2* and *FBP1* being downregulated. To address known limitations from pathway analyses (34, 41, 42), we repeated these analyses using different annotation databases (Reactome, Gene Ontology). This confirmed the enrichment in pathways linked with lipid and glucose metabolism (Supplemental Table 3). Enrichment with the use of the MGI Mammalian Phenotype ontology (43) identified significant enrichments in lipid homeostasis (MP0002118) and liver physiology (MP0000609) (see constituent genes in Supplemental Table 3) and marginal enrichment in glucose homeostasis (FDR-adjusted $P = 0.096$). Network-based gene set enrichment analyses (34) also confirmed the significant enrichment in fatty acid metabolism (*PPAR* signaling and biosynthesis of unsaturated fatty acids) and marginal enrichment in several pathways related to glucose homeostasis.

Results validation by using RT-qPCR

From the identified 29 genes in the RNAseq analyses, 18 were quantified with RT-qPCR in all available DiOGenes RNA samples ($n = 350$ subjects) and passed the quality-control procedure (see Methods). With these data, we confirmed that all 18 genes were significantly differentially expressed during the LCD (**Supplemental Table 4**). The direction of change was also consistent with results from RNAseq, with *GSDMB* and *C7* being upregulated and all remaining genes being downregulated.

To complete the analysis, we reassessed the association between changes in expression and LCD clinical outcomes. Full results are shown in Supplemental Table 4, and a simplified overview is shown in Figure 2B. Fifteen of eighteen genes were significantly associated with BMI changes. Two additional genes [aldo-keto reductase family 1 member C3 (*AKR1C3*) and

TABLE 2
List of genes differentially expressed during the LCD and associated with both weight and glycemic LCD outcomes[1]

Ensembl gene ID	Gene symbol	Median expression at baseline (FPKM)	Expression $\log_2$ FC	Differential expression FDR	Association with BMI changes		Association with Matsuda index changes	
					P	Adjusted *P*	*P*	Adjusted *P*
ENSG00000065833	*ME1*	64.18	−0.3247	1.16×10^{-59}	1.36×10^{-05}	1.16×10^{-04}	2.94×10^{-03}	8.01×10^{-02}
ENSG00000113140	*SPARC*	1225.38	−0.2722	1.61×10^{-52}	2.64×10^{-12}	1.71×10^{-09}	4.96×10^{-04}	3.13×10^{-02}
ENSG00000109919	*MTCH2*	45.56	−0.229	8.14×10^{-51}	1.62×10^{-05}	1.32×10^{-04}	4.84×10^{-03}	9.47×10^{-02}
ENSG00000113083	*LOX*	35.49	−0.295	3.08×10^{-42}	2.53×10^{-07}	6.01×10^{-06}	3.66×10^{-04}	3.13×10^{-02}
ENSG00000067113	*PPAP2A*	78.59	−0.214	1.00×10^{-45}	2.11×10^{-05}	1.64×10^{-04}	4.69×10^{-03}	9.47×10^{-02}
ENSG00000012660	*ELOVL5*	202.07	−0.2428	8.05×10^{-41}	2.69×10^{-04}	1.20×10^{-03}	4.56×10^{-04}	3.13×10^{-02}
ENSG00000164125	*FAM198B*	51.46	−0.2336	2.68×10^{-38}	9.25×10^{-05}	5.23×10^{-04}	4.59×10^{-03}	9.47×10^{-02}
ENSG00000144152	*FBLN7*	0.85	0.3216	2.41×10^{-37}	1.82×10^{-04}	8.56×10^{-04}	4.20×10^{-03}	9.47×10^{-02}
ENSG00000105971	*CAV2*	129.47	−0.2226	3.24×10^{-38}	4.70×10^{-06}	4.90×10^{-05}	1.27×10^{-03}	5.83×10^{-02}
ENSG00000174697	*LEP*	503.49	−0.2991	1.50×10^{-38}	6.88×10^{-09}	3.41×10^{-07}	6.13×10^{-06}	3.48×10^{-03}
ENSG00000196139	*AKR1C3*	74.88	−0.2473	1.91×10^{-31}	8.85×10^{-06}	8.39×10^{-05}	3.80×10^{-03}	9.37×10^{-02}
ENSG00000230058	*ENSG00000230058*	7.56	−0.3348	2.11×10^{-28}	5.22×10^{-04}	1.97×10^{-03}	9.96×10^{-04}	5.55×10^{-02}
ENSG00000145757	*SPATA9*	3.76	−0.3257	4.01×10^{-30}	2.37×10^{-06}	2.94×10^{-05}	5.54×10^{-03}	9.76×10^{-02}
ENSG00000186081	*KRT5*	1.65	−0.4236	4.00×10^{-24}	4.65×10^{-06}	4.90×10^{-05}	5.68×10^{-03}	9.76×10^{-02}
ENSG00000164128	*NPY1R*	87.12	−0.2762	8.48×10^{-30}	5.42×10^{-07}	8.96×10^{-06}	4.66×10^{-04}	3.13×10^{-02}
ENSG00000147852	*VLDLR*	38.69	−0.2795	8.87×10^{-27}	1.46×10^{-09}	1.35×10^{-07}	5.67×10^{-03}	9.76×10^{-02}
ENSG00000151632	*AKR1C2*	137.58	−0.2147	7.63×10^{-24}	2.74×10^{-05}	2.01×10^{-04}	4.58×10^{-03}	9.47×10^{-02}
ENSG00000240583	*AQP1*	132.57	−0.2171	9.44×10^{-22}	1.07×10^{-04}	5.72×10^{-04}	2.02×10^{-03}	8.01×10^{-02}
ENSG00000175449	*RFESD*	0.76	−0.2223	1.41×10^{-19}	6.22×10^{-04}	2.27×10^{-03}	1.34×10^{-03}	5.83×10^{-02}
ENSG00000206538	*VGLL3*	17.64	−0.2103	3.81×10^{-20}	6.83×10^{-04}	2.42×10^{-03}	3.56×10^{-04}	3.13×10^{-02}
ENSG00000241644	*INMT*	5.37	−0.2718	6.29×10^{-20}	6.70×10^{-04}	2.39×10^{-03}	3.97×10^{-03}	9.37×10^{-02}
ENSG00000073605	*GSDMB*	6.78	0.2488	3.17×10^{-16}	5.19×10^{-05}	3.22×10^{-04}	2.46×10^{-04}	3.13×10^{-02}
ENSG00000236404	*ENSG00000236404*	4.83	−0.275	2.17×10^{-17}	3.44×10^{-07}	7.10×10^{-06}	5.61×10^{-03}	9.76×10^{-02}
ENSG00000112936	*C7*	18.01	0.2618	2.13×10^{-13}	3.90×10^{-05}	2.59×10^{-04}	2.97×10^{-03}	8.01×10^{-02}
ENSG00000173641	*HSPB7*	81.79	−0.2279	4.52×10^{-16}	4.92×10^{-07}	8.37×10^{-06}	8.76×10^{-05}	2.48×10^{-02}
ENSG00000266968	*ENSG00000266968*	4.12	−0.2858	1.97×10^{-13}	2.95×10^{-09}	2.11×10^{-07}	4.22×10^{-04}	3.13×10^{-02}
ENSG00000264968	*ENSG00000264968*	0.83	0.3305	1.38×10^{-10}	4.21×10^{-05}	2.72×10^{-04}	2.64×10^{-03}	8.01×10^{-02}
ENSG00000196090	*PTPRT*	0.29	0.2379	2.94×10^{-10}	1.11×10^{-04}	5.88×10^{-04}	2.23×10^{-03}	8.01×10^{-02}
ENSG00000109061	*MYH1*	2.60	−0.2547	8.96×10^{-07}	2.17×10^{-08}	8.73×10^{-07}	2.72×10^{-03}	8.01×10^{-02}

[1] The differential expression was tested by using the DESeq2 framework, and the association between the change in gene expression and clinical outcomes was assessed by using linear mixed effect models. *AKR1C2*, aldo-keto reductase family 1 member C2; *AKR1C3*, aldo-keto reductase family 1 member C3; *AQP1*, aquaporin 1; *CAV2*, caveolin2; *C7*, complement component 7; *ELOVL5*, ELOVL fatty acid elongase 5; *FAM198B*, family with sequence similarity 198 member B; *FBLN7*, fibulin 7; FDR, false discovery rate; FPKM, fragments per kilobase of transcript per million mapped reads; FC, fold-change; *GSDMB*, gasdermin B; *HSPB7*, heat shock protein family B (small) member 7; ID, identifier; *INMT*, indolethylamine N-methyltransferase; *KRT5*, keratin 5; LCD, low-calorie diet; *LEP*, leptin; *LOX*, lysyl oxidase; *ME1*, malic enzyme 1; *MTCH2*, mitochondrial carrier 2; *MYH1*, myosin heavy polypeptide 1; *NPY1R*, neuropeptide Y receptor Y1; *PPAP2A*, phosphatidic acid phosphatase type 2A; *PTPRT*, protein tyrosine phosphatase, receptor type T; *RFESD*, Rieske Fe-S domain containing; *SPARC*, secreted protein acidic and cysteine rich; *SPATA9*, spermatogenesis associated 9; *VGLL3*, vestigial like family member 3; *VLDLR*, very low density lipoprotein receptor.

vestigial like family member 3 (*VGLL3*)] had marginal associations with BMI changes (FDR-adjusted $P < 0.15$). Four genes were significantly associated with changes in the Matsuda index [lysyl oxidase (*LOX*), *LEP*, mitochondrial carrier 2 (*MTCH2*), and *GSDMB*)], and 4 additional genes [*ELOVL5*, heat shock protein family B (small) member 7 (*HSPB7*), neuropeptide Y receptor Y1 (*NPY1R*), and very low density lipoprotein receptor (*VLDLR*)] showed a trend with corrected *P* value <15%.

The association between changes in expression during the LCD and weight outcomes, 6 mo after LCD, was also tested. Nine genes were found significantly associated with, BMI and 5 genes were associated with the Matsuda index 6 mo after LCD (Figure 2B). Using a less conservative threshold (FDR-adjusted *P* value <0.1), we identified 2 additional genes (*LEP* and *C7*) associated with BMI and 9 additional genes associated with the Matsuda index.

The directionality of correlation between expression changes and clinical outcomes is indicated in Figure 2B. All downregulated genes had a positive correlation with BMI changes, indicating that downregulation of gene expression correlated with weight loss and maintenance. Conversely, these genes had a negative correlation with Matsuda outcomes, reflecting the correlation between gene downregulation and glycemic improvements (improvement of insulin sensitivity). The 2 upregulated genes (*C7* and *GSDMB*) had a negative correlation with BMI outcomes and a positive correlation with Matsuda outcomes.

Predictive models

To further challenge the potential of the 18 identified genes as biomarkers, we constructed models to predict subjects with distinct clinical outcomes 6 mo after the LCD (see Methods).

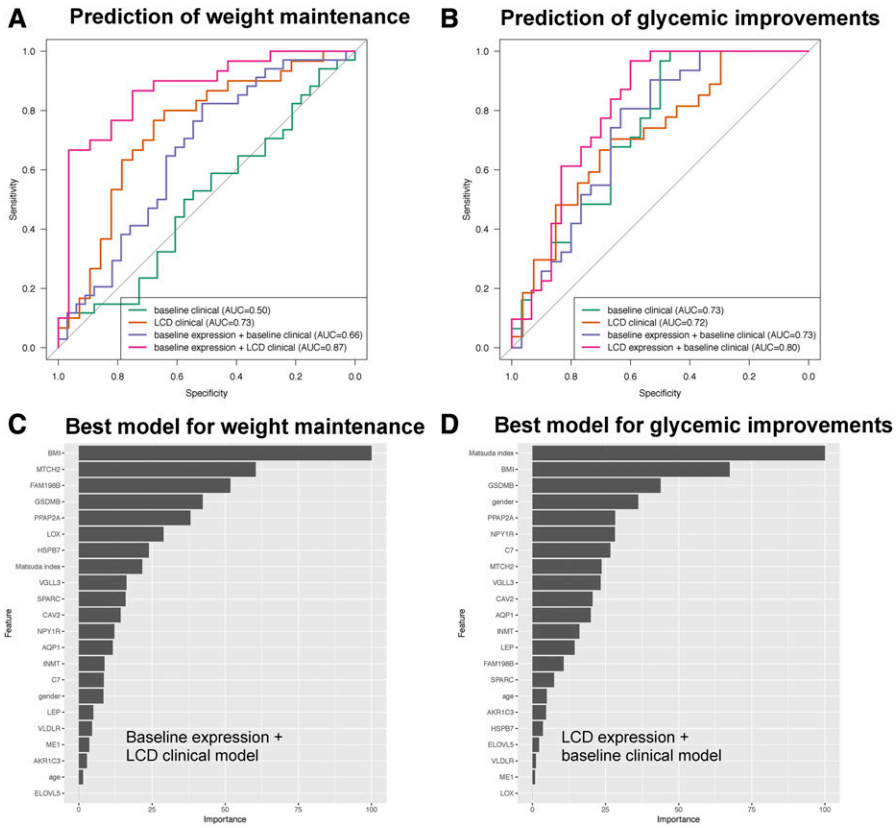

FIGURE 3 Results from predictive models. Panels A and B show receiver operating characteristic curves showing classification performance as obtained on a validation dataset for classification of weight maintainers (A) and classification of subjects with glycemic improvements after the 6-mo follow-up (B). Two clinical models were fitted, one with the use of baseline variables (BMI, Matsuda index, sex, and age) and the second model with the use of the same variables with BMI and Matsuda index changes during the LCD. The best clinical model was then combined with the expression data of 18 genes. Two models were derived: one with the use of baseline expression, the second with the use of expression changes during the LCD; only the best combined model is shown in the plot. For comparison, a combined model based only on baseline variables (clinical and expression) is also displayed. Classification models were constructed with the use of GLMs with a binomial function. Panels C and D show the relative importance of the predictors in the best identified models with panel C as the best model for the prediction of weight maintainers and panel D as the best model for the prediction of subjects with glycemic improvements. Importance is defined according to the *z* scores from the GLM model and is relative to the top predictor (the one with the highest *z* score). AKR1C3, aldo-keto reductase family 1 member C3; AQP1, aquaporin 1; CAV2, caveolin2; C7, complement component 7; ELOVL5, ELOVL fatty acid elongase 5; FAM198B, family with sequence similarity 198 member B; GLM, generalized linear model; GSDMB, gasdermin B; HSPB7, heat shock protein family B (small) member 7; INMT, indolethylamine N-methyltransferase; LCD, low-calorie diet; LEP, leptin; LOX, lysyl oxidase; ME1, malic enzyme 1; MTCH2, mitochondrial carrier 2; NPY1R, neuropeptide Y receptor Y1; PPAP2A, phosphatidic acid phosphatase type 2A; SPARC, secreted protein acidic and cysteine rich; VGLL3, vestigial like family member 3; VLDLR, very low density lipoprotein receptor.

For BMI analyses, 2 classes were defined: *1*) subjects with BMI changes greater than the median population's change and *2*) the remaining subjects having less weight maintenance. Next, we defined a clinical model based only on baseline clinical variables (BMI, Matsuda index, sex, and age). This model had very poor classification performance (ROC AUC: 0.50; 95% CI: 0.36, 0.64; **Figure 3**A), equivalent to a random classification. We next assessed a clinical model using clinical changes during the LCD (changes in weight and the Matsuda index, with baseline age and sex). This model, referred to as the "LCD clinical model," had significantly better performance with (AUC: 0.73; 95% CI: 0.60, 0.87) and was kept for subsequent analyses. Using AT transcriptomics (RT-qPCR) data, we explored whether the classification performance could be further improved. To this aim, 2 additional models were evaluated:

- A model combining changes in clinical variables during the LCD, together with changes in AT gene expression during LCD. This model was referred to as the "LCD expression + LCD clinical model."
- A model combining changes in clinical variables during LCD, together with baseline gene expression. This model was referred to as the "baseline expression + LCD clinical model."

The first model yielded slightly higher performance (AUC: 0.79; 95% CI: 0.67, 0.91). The second model provided even better performance (AUC: 0.87; 95% CI: 0.77, 0.96) and significantly outperformed the LCD clinical model (Delong $P = 0.012$) and LCD expression + LCD clinical model ($P = 0.029$).

We also investigated a model combining baseline clinical variables with baseline expression amounts ("baseline expression + baseline clinical"), potentially the easiest to use in a clinical setting. This model achieved better performance than the baseline clinical model (AUC: 0.66; 95% CI: 0.52, 0.79). However, such performance remained lower than all models based on LCD clinical changes (Figure 3A).

Figure 3C shows the features ranked by relative importance for the best identified model ("baseline expression levels with changes in clinical variables during LCD"). Change in BMI during the LCD was the strongest predictor of weight maintenance, with baseline *MTCH2*, family with sequence similarity 198 member B (*FAM198B*), *GSDMB*, and phosphatidic acid phosphatase type 2A (*PPAP2A*) gene expression having very strong contributions to the model. Validation AUCs from all possible pairwise combinations of clinical and transcriptomics models are available in **Supplemental Table 5,** and their corresponding ROC curves are displayed in **Supplemental Figure 1**. Additional details about the 2 BMI groups can be found in Supplemental Tables (see **Supplemental Table 6** for clinical characteristics, **Supplemental Table 7** for a summary of gene expression, and **Supplemental Table 8** for the full details of the best classification model, as obtained during the training phase).

For changes in Matsuda index, we performed similar analyses and defined 2 groups of patients according to changes higher or lower than the median population's change. Both baseline and LCD clinical models yielded good performance (Figure 3B). The baseline clinical model had a AUC of 0.73 (95% CI: 0.60, 0.86), and the LCD clinical model had an AUC of 0.72 (95% CI: 0.58, 0.85). Deriving a baseline expression + baseline clinical model provided similar performance (AUC: 0.73; 95% CI: 0.60, 0.86). Thus, between these models, the simpler model (based only on clinical variables) would be preferred.

However, the model combining baseline clinical variables and changes in gene expression during the LCD had the highest performance (AUC: 0.80; 95% CI: 0.69, 0.92; Figure 3B). This model had marginally higher performance than the baseline clinical model ($P = 0.058$) and had significantly higher performance than the baseline expression + baseline clinical ($P = 0.035$). Repeating these analyses with the use of the LCD clinical model reached the same conclusions. A summary from all tested models is presented in Supplemental Table 5, and all corresponding ROC curves are displayed in **Supplemental Figure 2**. Figure 3D shows the variable importance of the best model. As expected, clinical variables have a strong contribution (the Matsuda index and BMI are ranked as the top 2 variables), but clearly the model benefits from several gene expression profiles. In particular, the *GSDMB*, *PPAP2A*, *NPY1R*, *C7*, and *MTCH2* are the top 5 genes that contribute to the model.

As with the BMI classes, complete information on the characteristics of the Matsuda groups and details about the best classification model can be found in **Supplemental Tables 9–11**.

Finally, we checked whether adding the weight-maintenance diet label would improve further the performance of the models. For classification of both BMI and Matsuda classes, adding the weight-maintenance diet label in the models did not improve significantly the performance compared with the existing models (Delong's $P > 0.30$). We also attempted to simplify the various models by applying a recursive feature elimination. Yet, the simplified models did not yield better performance than the full models. On the validation dataset, the AUCs from the simplified models were slightly smaller than the AUCs from the full models; however, these differences were not significant ($P > 0.05$). Thus, the combination of both strong and weak predictors (e.g., genes marginally associated with the outcomes) was useful for the overall classification's performance.

Expression profiles in isolated adipocytes and SVF

The 18 genes were also quantified in different AT fractions, including total AT, isolated adipocytes, and SVF from 7 non-DiOGenes subjects (see Methods). These analyses aimed to distinguish genes expressed mostly in adipocytes from genes mostly expressed in SVF. We compared expression between paired isolated adipocytes and SVF and found that 14 of 18 genes had significantly higher expression in adipocytes than in SVF (corrected $P < 0.05$). Three genes presented higher expression in SVF. The *LOX* gene was expressed significantly lower in adipocytes than in SVF (corrected $P = 0.019$), and the *C7* and aquaporin 1 (*AQP1*) genes had marginally lower expression in adipocytes than in SVF (corrected $P = 0.067$). No significant differences were observed for indolethylamine N-methyltransferase (*INMT*) (corrected $P = 0.23$).

eQTL analyses

We tested whether expression of the 18 genes was under some genetic control (eQTL effect). Using 57,854 SNPs located within 1 Mb of each gene, we assessed whether there was any *cis*-eQTL effect on baseline expression or on expression fold-change during the LCD. For baseline gene expression, only 2 SNPs (rs73206101 and rs73208115) were significantly associated with caveolin 2 (*CAV2*) gene expression (adjusted $P < 0.05$). These 2 SNPs were in moderate linkage disequilibrium ($r^2 = 0.68$) with each other. No significant *cis*-eQTL effect was detected for expression changes during the LCD. By repeating this analysis with expression fold-changes adjusted for weight loss, the same negative results were produced. Using different type of multiple-testing corrections (FDR, familywise error rate) also led to the same conclusions.

DISCUSSION

In this study, we documented transcriptome-wide AT changes in response to an LCD and the link with clinical outcomes after the LCD. We identified 1173 genes differentially expressed during the LCD, with most genes being downregulated. This is consistent with previous studies (23) and in agreement with fat-mass reduction and adaptation of the AT metabolism (44). Pathway analyses of the 350 genes associated with BMI changes pointed toward enrichment in lipid and glucose metabolism pathways. This is in line with recent analyses of the DiOGenes lipidomic data (7) showing the importance of dyslipidemia during LCD. This is also in agreement with gene expression findings distinguishing weight maintainers from weight regainers (14). Enrichment analyses with the use of the 33 genes associated with Matsuda outcomes were not performed, because this type of analysis usually requires a larger gene set. However, the individual gene functions pointed to similar biological processes.

Nearly all genes (29 of 33) associated with glycemic improvements were also associated with weight loss, marking them as potentially clinically relevant biomarkers. We thus focused on the 29 genes associated with both endpoints, with the aim to better understand the link between expression changes and glycemic improvements. Remarkably, within this list, 7 of the 10 most differentially expressed genes [*ME1*, secreted protein acidic and cysteine rich (*SPARC*), *MTCH2*, *LOX*, *ELOVL5*, *CAV2*, and *LEP*] are well established in the context of obesity, T2D onset,

or adipocyte regulation. *ME1* is a known causal obesity gene (45), *SPARC* is involved with adipocyte differentiation and strongly associated with both obesity and diabetes (46), *MTCH2* is a known obesity risk factor gene (47), *LOX* levels were found altered on bariatric surgery (48), *ELOVL5* plays a key role in regulating triglycerides and MUFA/PUFA synthesis (49), genetic polymorphisms in *CAV2*, in interaction with fat intake, are associated with T2D (50), and *LEP* is a major hallmark gene in weight loss and obesity (51–53). Such enrichment in obesity biomarkers lends further credibility to our results.

RT-qPCR analyses in a larger sample set (n = 350 subjects instead of n = 191) showed that all tested genes (18) displayed differential expression, with a consistent direction of expression changes compared with RNAseq results. Overall, we observed a very strong replication of BMI hits and less replication for Matsuda hits. The latter can be explained by numerous factors, such as a slightly higher rate of false positives in the discovery phase, the higher variability of Matsuda outcomes (RSD = 120% compared with 25% for BMI outcomes), increased heterogeneity in the replication cohort, or simply technological differences between RNAseq and RT-qPCR leading to differences in the dynamic range. In subsequent analyses, we kept all genes because their known function clearly marked them as plausible biological candidates. To assess their potential as biomarkers, we defined 2 classes of patients: good and lesser improvers both for BMI and Matsuda outcomes and attempted to construct classification models. Classification of good weight maintainers could not be achieved with a model based only on baseline clinical variables: the performance was comparable to a random classification. This illustrates the difficulty in predicting weight maintenance based on initial body weight. A model based on clinical changes during the LCD provided much better performance. Performance was further increased, up to an AUC of 0.87, by integrating baseline gene expression. In this model, BMI had some predictive value, and the model benefited from several key gene expression amounts such as *MTCH2*, *FAM198B*, *GSDMB*, and *PPAP2A*. Our choice to focus on genes associated with both BMI and the Matsuda index potentially excluded other relevant BMI biomarkers. Yet, with the best model already having an AUC close to 0.87, the room for significant improvement is limited. For classification of the Matsuda groups, the clinical models already yielded good performance (AUC: 0.73). Here, a model integrating baseline gene expression did not improve the classification. By contrast, the model using changes in expression during the LCD improved the AUC by +7%, corresponding to an increase in both sensitivity and specificity. In this model, the clinical variables, the Matsuda index, BMI, and sex, were among the strongest predictors (ranked as 1, 2, and 4, respectively) as expected, and the model strongly benefited from inclusion of the following genes: *GSDMB*, *PPAP2A*, *NPY1R*, *C7*, and *MTCH2* (ranked as 3, then 5–8, respectively).

Expression analyses of 18 genes in AT fractions showed that 14 were predominantly expressed in adipocytes compared with the SVF, whereas 3 were more specific to the SVF. Among the 18 genes, 5 encoded for secreted proteins (*LEP*, *LOX*, *SPARC*, *C7*, and *VLDLR*) and may thus have potential as circulating biomarkers.

eQTL analyses revealed *cis*-effects on *CAV2* baseline expression, and no significant eQTL effect was detected for expression changes during the LCD in any of the 18 genes. This is consistent with previous observations (23), where only a marginal *cis*-eQTL effect was found on baseline expression, and no significant effect was detected for expression fold-changes on LCD.

Our results are strengthened by the fact that the DiOGenes study is the largest study of its kind, combining data both from an LCD and a weight-maintenance diet intervention. Our initial analyses were based on a discovery RNAseq analysis, and findings were replicated in a larger sample from the same cohort by using RT-qPCR. Such a large sample size enabled us to perform classification analyses, splitting the data into training and testing sets, and using internal cross-validation during the training phase. We could thereby maximize the use of each observation and prevent model overfitting. Further follow-up in isolated adipocytes demonstrated the cell specificity of our findings, and potential confounding factors linked with genotypes could be excluded by eQTL analyses. Additional preclinical and clinical studies are beyond the scope of the current study but remain needed to fully validate the identified biomarkers and understand the underlying molecular differences between responders and nonresponders.

In conclusion, our study demonstrates the importance of molecular phenotyping to identify potential biomarkers and studying the possible mechanisms involved in the dynamics of weight loss and maintenance. To our knowledge, this study is the first to link LCD-induced transcriptome-wide alterations with weight and glycemic outcomes in a large population of overweight and obese patients. We propose plausible biomarkers that outperform clinical models for classification of weight maintainers and that help in improving clinical models for prediction of glycemic improvers. Our methodologies and findings might benefit clinicians and scientists by providing a better understanding of the large intersubject variability in response to dietary interventions.

We thank Esther Danenberg for very precious help with the RNAseq data, Hélène Ruffieux and Margherita Springer for helpful comments on our manuscript.

The authors' responsibilities were as follows—WHMS and AA: designed the DiOGenes clinical study; AV, NV, WHMS, JH, and DL: designed the transcriptomics studies; AV: designed and led the present study; CA, GL, JB, ADC, JC, and AV: implemented the statistical analyses; SB: generated the RT-qPCR data; PD, DL, and NV: supervised the expression data generation; JC and AV: performed the eQTL analyses; AV, CA, and NV: interpreted the results and wrote the manuscript with input from all authors; AV: had primary responsibility for the final content; and all authors: read and approved the final manuscript. GL, JC, NG, PD, JH, and AV are full-time employees at Nestlé Institute of Health Sciences SA. CA and ADC are full-time employees at QuartzBio SA. WHMS reports having received research support from several food companies such as Nestlé, DSM, Unilever, Nutrition et Sante, and Danone as well as Pharmaceutical companies such as GSK, Novartis, and Novo Nordisk; he is an unpaid scientific advisor for the International Life Science Institute, ILSI Europe. AA reports grants and personal fees from Global Dairy Platform, personal fees from McCain Foods, personal fees from McDonald's, personal fees from Arena Pharmaceuticals Inc., personal fees from Basic Research, personal fees from Dutch Beer Knowledge Institute, Netherlands, personal fees from Gelesis, personal fees from Novo Nordisk, Denmark, personal fees from Orexigen Therapeutics Inc., personal fees from S-Biotek, Denmark, personal fees from Twinlab, personal fees from Vivus Inc., grants from Arla Foods, Denmark, grants from Danish Dairy Research Council, grants from Nordea Foundation, Denmark, outside the submitted work, and royalties received for the book first published in Danish as "Verdens Bedste Kur" (Politiken; Copenhagen, Denmark), and subsequently published in Dutch as "Het beste dieet ter wereld" (Kosmos Uitgevers; Utrecht/Antwerpen,

Netherlands), in Spanish as "Plan DIOGENES para el control del peso. La dieta personalizada inteligente" (Editorial Evergráficas; Léon, Spain), and in English as "World's Best Diet" (Penguin, Australia).

REFERENCES

1. Haslam DW, James WP. Obesity. Lancet 2005;366:1197–209.
2. Dixon JB. The effect of obesity on health outcomes. Mol Cell Endocrinol 2010;316:104–8.
3. Lean ME. Pathophysiology of obesity. Proc Nutr Soc 2000;59:331–6.
4. Soare A, Weiss EP, Pozzilli P. Benefits of caloric restriction for cardiometabolic health, including type 2 diabetes mellitus risk. Diabetes Metab Res Rev 2014;30 Suppl 1:41–7.
5. Alves NE, Enes BN, Martino HS, Alfenas Rde C, Ribeiro SM. Meal replacement based on Human Ration modulates metabolic risk factors during body weight loss: a randomized controlled trial. Eur J Nutr 2014;53:939–50.
6. Neiberg RH, Wing RR, Bray GA, Reboussin DM, Rickman AD, Johnson KC, Kitabchi AE, Faulconbridge LF, Kitzman DW, Espeland MA, et al. Patterns of weight change associated with long-term weight change and cardiovascular disease risk factors in the Look AHEAD Study. Obesity (Silver Spring) 2012;20:2048–56.
7. Valsesia A, Saris WH, Astrup A, Hager J, Masoodi M. Distinct lipid profiles predict improved glycemic control in obese, nondiabetic patients after a low-caloric diet intervention: the Diet, Obesity and Genes randomized trial. Am J Clin Nutr 2016;104:566–75.
8. Capel F, Viguerie N, Vega N, Dejean S, Arner P, Klimcakova E, Martinez JA, Saris WH, Holst C, Taylor M, et al. Contribution of energy restriction and macronutrient composition to changes in adipose tissue gene expression during dietary weight-loss programs in obese women. J Clin Endocrinol Metab 2008;93:4315–22.
9. Dahlman I, Linder K, Arvidsson Nordström E, Andersson I, Lidén J, Verdich C, Sørensen TIA, Arner P. Changes in adipose tissue gene expression with energy-restricted diets in obese women. Am J Clin Nutr 2005;81:1275–85.
10. Franck N, Gummesson A, Jernås M, Glad C, Svensson P-A, Guillot G, Rudemo M, Nyström FH, Carlsson LM, Olsson B. Identification of adipocyte genes regulated by caloric intake. J Clin Endocrinol Metab 2011;96:E413–8.
11. Viguerie N, Vidal H, Arner P, Holst C, Verdich C, Avizou S, Astrup A, Saris WH, Macdonald IA, Klimcakova E, et al. Adipose tissue gene expression in obese subjects during low-fat and high-fat hypocaloric diets. Diabetologia 2005;48:123–31.
12. 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) 2017;41:309–16.
13. Márquez-Quiñones A, Mutch DM, Debard C, Wang P, Combes M, Roussel B, Holst C, Martinez JA, Handjieva-Darlenska T, Kalouskova P, et al. Adipose tissue transcriptome reflects variations between subjects with continued weight loss and subjects regaining weight 6 mo after caloric restriction independent of energy intake. Am J Clin Nutr 2010;92:975–84.
14. Mutch DM, Pers TH, Temanni MR, Pelloux V, Marquez-Quiñones A, Holst C, Martinez JA, Babalis D, van Baak MA, Handjieva-Darlenska T, et al. A distinct adipose tissue gene expression response to caloric restriction predicts 6-mo weight maintenance in obese subjects. Am J Clin Nutr 2011;94:1399–409.
15. Polsky S, Ogden LG, MacLean PS, Giles ED, Brill C, Wyatt HR. Biomarker profile does not predict weight loss success in successful and unsuccessful diet-reduced obese individuals: a prospective study. ISRN Obes 2013;2013:804129.
16. Abdul-Ghani MA, Tripathy D, DeFronzo RA. Contributions of beta-cell dysfunction and insulin resistance to the pathogenesis of impaired glucose tolerance and impaired fasting glucose. Diabetes Care 2006;29:1130–9.
17. Gastaldelli A, Gaggini M, DeFronzo RA. Role of adipose tissue insulin resistance in the natural history of T2DM: results from the San Antonio Metabolism Study. Diabetes 2017;66:815–22.
18. Larsen TM, Dalskov S-M, van Baak M, Jebb SA, Papadaki A, Pfeiffer AF, Martinez JA, Handjieva-Darlenska T, Kunešová M, Pihlsgård M, et al. Diets with high or low protein content and glycemic index for weight-loss maintenance. N Engl J Med 2010;363:2102–13.
19. Larsen TM, Dalskov S, van Baak M, Jebb S, Kafatos A, Pfeiffer A, Martinez JA, Handjieva-Darlenska T, Kunesová M, Holst C, et al. The Diet, Obesity and Genes (Diogenes) dietary study in eight European countries - a comprehensive design for long-term intervention. Obes Rev 2010;11:76–91.
20. Miller M, Beach V, Sorkin JD, Mangano C, Dobmeier C, Novacic D, Rhyne J, Vogel RA. Comparative effects of three popular diets on lipids, endothelial function, and C-reactive protein during weight maintenance. J Am Diet Assoc 2009;109:713–7.
21. Hession M, Rolland C, Kulkarni U, Wise A, Broom J. Systematic review of randomized controlled trials of low-carbohydrate vs. low-fat/low-calorie diets in the management of obesity and its comorbidities. Obes Rev 2009;10:36–50.
22. Hemmingsson E, Johansson K, Eriksson J, Sundström J, Neovius M, Marcus C. Weight loss and dropout during a commercial weight-loss program including a very-low-calorie diet, a low-calorie diet, or restricted normal food: observational cohort study. Am J Clin Nutr 2012;96:953–61.
23. Viguerie N, Montastier E, Maoret J-J, Roussel B, Combes M, Valle C, Villa-Vialaneix N, Iacovoni JS, Martinez JA, Holst C, et al. Determinants of human adipose tissue gene expression: impact of diet, sex, metabolic status, and cis genetic regulation. PLoS Genet 2012;8:e1002959.
24. Illumina. bcl2fastq conversion software [Internet]. [cited 2015 Aug 27]. Available from: http://support.illumina.com/sequencing/sequencing_software/casava.html.
25. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
26. Andrews, S. FastQC: a quality control tool for high throughput sequence data. [Internet]. [cited 2015 Aug 27]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
27. Morgan M, Pagès H, Obenchain V, Hayden N. Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import [Internet]. [cited 2015 Aug 27]. Available from: http://bioconductor.org/packages/release/bioc/html/Rsamtools.html.
28. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. Software for computing and annotating genomic ranges. PLOS Comput Biol 2013;9:e1003118.
29. Pagès H, Carlson M, Falcon S, Li N. AnnotationDbi: annotation database interface [Internet]. [cited 2015 Aug 27]. Available from: https://bioconductor.org/packages/release/bioc/html/AnnotationDbi.html.
30. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 2001;25:402–8.
31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995;57:289–300.
33. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90–7.
34. Glaab E, Baudot A, Krasnogor N, Schneider R, Valencia A. EnrichNet: network-based gene set enrichment analysis. Bioinformatics 2012;28:i451–7.
35. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491:56–65.
36. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–7.
37. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82.
38. Gao X. Multiple testing corrections for imputed SNPs. Genet Epidemiol 2011;35:154–8.
39. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28:1–26.
40. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.
41. Jin L, Zuo X-Y, Su W-Y, Zhao X-L, Yuan M-Q, Han L-Z, Zhao X, Chen Y-D, Rao S-Q. Pathway-based analysis tools for complex diseases: a review. Genomics Proteomics Bioinformatics 2014;12:210–20.

42. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol 2012;8: e1002375.
43. Smith CL, Eppig JT. The mammalian phenotype ontology: enabling robust annotation and comparative analysis. Wiley Interdiscip Rev Syst Biol Med 2009;1:390–9.
44. Larson-Meyer DE, Heilbronn LK, Redman LM, Newcomer BR, Frisard MI, Anton S, Smith SR, Maplstat AA, Ravussin E. Effect of calorie restriction with or without exercise on insulin sensitivity, β-cell function, fat cell size, and ectopic lipid in overweight subjects. Diabetes Care 2006;29:1337–44.
45. Yang X, Deignan JL, Qi H, Zhu J, Qian S, Zhong J, Torosyan G, Majid S, Falkard B, Kleinhanz RR, et al. Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. Nat Genet 2009;41:415–23.
46. Challa TD, Straub LG, Balaz M, Kiehlmann E, Donze O, Rudofsky G, Ukropec J, Ukropcova B, Wolfrum C. Regulation of de novo adipocyte differentiation through cross talk between adipocytes and preadipocytes. Diabetes 2015;64:4075–87.
47. Graff M, Gordon-Larsen P, Lim U, Fowke JH, Love S-A, Fesinmeyer M, Wilkens LR, Vertilus S, Ritchie MD, Prentice RL, et al. The influence of obesity-related single nucleotide polymorphisms on BMI across the life course: the PAGE study. Diabetes 2013;62:1763–7.
48. Liu Y, Aron-Wisnewsky J, Marcelin G, Genser L, Le Naour G, Torcivia A, Bauvois B, Bouchet S, Pelloux V, Sasso M, et al. Accumulation and changes in composition of collagens in subcutaneous adipose tissue after bariatric surgery. J Clin Endocrinol Metab 2016; 101:293–304.
49. Tripathy S, Lytle KA, Stevens RD, Bain JR, Newgard CB, Greenberg AS, Huang L-S, Jump DB. Fatty acid elongase-5 (Elovl5) regulates hepatic triglyceride catabolism in obese C57BL/6J mice. J Lipid Res 2014;55:1448–64.
50. Fisher E, Schreiber S, Joost H-G, Boeing H, Döring F. A two-step association study identifies CAV2 rs2270188 single nucleotide polymorphism interaction with fat intake in type 2 diabetes risk. J Nutr 2011;141:177–81.
51. Kilpeläinen TO, Carli JFM, Skowronski AA, Sun Q, Kriebel J, Feitosa MF, Hedman ÅK, Drong AW, Hayes JE, Zhao J, et al. Genome-wide meta-analysis uncovers novel loci influencing circulating leptin levels. Nat Commun 2016;7:10494.
52. Ahima RS, Prabakaran D, Mantzoros C, Qu D, Lowell B, Maratos-Flier E, Flier JS. Role of leptin in the neuroendocrine response to fasting. Nature 1996;382:250–2.
53. Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, Sewter CP, Digby JE, Mohammed SN, Hurst JA, et al. Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature 1997;387:903–8.