

Full title: Alterations of a cellular cholesterol metabolism network is a molecular feature of obesity-related type 2 diabetes and cardiovascular disease

Short title: Cellular cholesterol in obesity-related disease

Authors: Jingzhong Ding MD, PhD¹, Lindsay M Reynolds PhD², Tanja Zeller PhD³, Christian Müller PhD³, Kurt Lohman Mstat², Barbara J. Nicklas, PhD¹, Stephen B. Kritchevsky, PhD¹, Zhiqing Huang PhD⁴, Alberto de la Fuente PhD⁵, Nicola Soranzo PhD⁶, Robert E. Settlage PhD⁷, Chia-Chi Chuang PhD⁸, Timothy Howard PhD¹, Ning Xu MD, PhD⁹, Mark O. Goodarzi MD, PhD⁹, Y.-D. Ida Chen PhD¹⁰, Jerome I. Rotter MD¹⁰, David S. Siscovick MD, MPH¹¹, John S. Parks PhD⁸, Susan Murphy PhD⁴, David R. Jacobs Jr. PhD¹², Wendy Post MD, MS¹³, Russell P. Tracy PhD¹⁴, Philipp S. Wild PhD¹⁵, Stefan Blankenberg PhD³, Ina Hoeschele PhD⁷, David Herrington MD, MHS¹, Charles E. McCall MD¹, Yongmei Liu MD, PhD²

Affiliations:

¹ Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157, USA

² Department of Epidemiology & Prevention, Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157, USA

³ University Heart Center Hamburg, Clinic for General and Interventional Cardiology, and DZHK (German Centre for Cardiovascular Research), partner site Hamburg/Kiel/Lübeck, Hamburg, Germany

⁴ Duke University, Durham, North Carolina 27708, USA

⁵ Department of Biomathematics and Bioinformatics, Institute of Genetics and Biometry, Leibniz-Institute for Farm Animal Biology (FBN), Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany

⁶ CRS4 Bioinformatica, Polaris Edificio 3, Loc. Piscina Manna, 09010 Pula (CA), Italy

⁷ Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia 24061, USA

⁸ Department of Internal Medicine-Section on Molecular Medicine, Wake Forest School of
Medicine, Winston-Salem, North Carolina 27157, USA

⁹ Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Cedars-Sinai
Medical Center, Los Angeles, California 90048, USA

¹⁰ Institute for Translational Genomics and Population Sciences, Los Angeles BioMedical
Research Institute at Harbor-UCLA Medical Center, Torrance, California 90502, USA

¹¹ The New York Academy of Medicine, New York, NY 10029, USA

¹² School of Public Health, University of Minnesota, Minneapolis, Minnesota 55454, USA

¹³ Johns Hopkins University, Baltimore, Maryland 21224, USA

¹⁴ University of Vermont, Colchester, Vermont 05446, USA

¹⁵ Center for Thrombosis and Hemostasis and Department of Medicine 2, University Medical
Center of the Johannes, Gutenberg-University Mainz, and DZHK (German Centre for
Cardiovascular Research), partner site Rhine-Main, Mainz, Germany

Corresponding Author:

Yongmei Liu MD, PhD

Department of Epidemiology and Prevention

Division of Public Health Sciences

Medical Center Blvd.

Wake Forest School of Medicine

Winston-Salem, NC 27157

Fax # 336-716-6427

Tele #: 336-716-9503

yoliu@wakehealth.edu

Word count: 4,452

Tables: 1

Figures: 4

Abstract

Obesity is linked to type 2 diabetes (T2D) and cardiovascular diseases; however, the underlying molecular mechanisms remain unclear. We aimed to identify obesity-associated molecular features that may contribute to obesity-related diseases. Using circulating monocytes from 1,264 Multi-Ethnic Study of Atherosclerosis participants, we quantified the transcriptome and epigenome. We discovered that alterations in a network of co-expressed cholesterol metabolism genes are a signature feature of obesity and inflammatory stress. This network included 11 body mass index (BMI)-associated genes related to sterol uptake (\uparrow *LDLR*, \downarrow *MYLIP*), synthesis (\uparrow *SCD*, *FADS1*, *HMGCS1*, *FDFT1*, *SQLE*, *CYP51A1*, *SC4MOL*) and efflux (\downarrow *ABCA1*, *ABCG1*) – producing a molecular profile expected to increase intracellular cholesterol. Importantly, these alterations were associated with T2D and coronary artery calcium (CAC), independent from cardiometabolic factors including serum lipid profiles. This network mediated the associations between obesity and T2D/CAC. Several genes in the network harbored CpG dinucleotides (e.g. *ABCG1*/cg06500161) which overlapped ENCODE-annotated regulatory regions, and had methylation profiles that mediated the associations between BMI/inflammation and expression of their cognate genes. Taken together with several lines of previous experimental evidence, these data suggest that alterations of the cholesterol metabolism gene network represent a molecular link between obesity/inflammation and T2D/CAC.

Text

Obesity is a major risk factor for type 2 diabetes (T2D) and cardiovascular disease. Traditionally, obesity-related insulin resistance and atherosclerosis are viewed as lipid accumulation disorders, with fatty acid accumulation in insulin target tissues including liver, muscles, and adipose tissue,

and cholesterol accumulation in atheromatous plaques(1). In the past decade, experimental studies have demonstrated an essential role for immune cells, especially monocytes/macrophages, and chronic inflammation in the progression of insulin resistance and atherosclerosis (2-6). However, the precise molecular mechanisms accounting for these relationships remain uncertain. Genome-wide analyses of gene expression modifications associating with obesity could help define the cellular features of obesity, and suggest potential molecular mechanisms that contribute to obesity-associated diseases (7). Our goal is to better understand the biological mechanisms underlying obesity-related risk for common diseases by identifying T2D and/or subclinical cardiovascular disease-associated molecular characteristics that are common in blood monocytes of obese individuals. Accordingly, using purified peripheral monocytes from 1,264 Multi-Ethnic Study of Atherosclerosis (MESA) participants, we identified transcriptomic features associated with body mass index (BMI) and determined which of these features were also associated with T2D and coronary artery calcium (CAC).

Methods:

Study population

The Multi-Ethnic Study of Atherosclerosis (MESA) was designed to investigate the progression of subclinical cardiovascular disease in a community-based cohort of 6,814 participants. Since its inception in 2000, five clinic visits (exams) collected extensive clinical, socio-demographic, lifestyle and behavior, and laboratory data (8). The present analysis is primarily based on analyses of purified monocyte samples from the April 2010-February 2012 examination (exam 5) of 1,264 randomly selected MESA participants from four MESA field centers (Baltimore,

MD; Forsyth County, NC; New York, NY; and St. Paul, MN). At exam 5, these participants also underwent assessment of BMI, T2D and CAC. The study protocol was approved by the Institutional Review Board at each site. All participants signed informed consent.

Purification of monocytes

Monocytes were isolated with anti-CD14 monoclonal antibody coated magnetic beads, respectively, using AutoMACs automated magnetic separation unit (Miltenyi Biotec, Bergisch Gladbach, Germany; see **Supporting document** for details).

DNA/RNA extraction

DNA and RNA were isolated from samples simultaneously using the AllPrep DNA/RNA Mini Kit (Qiagen, Inc., Hilden, Germany; see **Supporting document** for details).

Global expression quantification

The Illumina HumanHT-12 v4 Expression BeadChip and Illumina Bead Array Reader were used to perform the genome-wide expression analysis, following the Illumina expression protocol (see **Supporting document** for details). These data have been deposited in the NCBI Gene Expression Omnibus and is accessible through GEO Series accession number (GSE56045).

Epigenome-wide methylation quantification

The Illumina HumanMethylation450 BeadChip and HiScan reader were used to perform the epigenome-wide methylation analysis (see **Supporting document** for details). These methylation data have been deposited in the NCBI Gene Expression Omnibus and is accessible through GEO Series accession number (GSE56046).

Quality control and pre-processing of microarray data

Data pre-processing and QC analyses were performed in *R* (<http://www.r-project.org/>) using *Bioconductor* (<http://www.bioconductor.org/>) packages (see **Supporting document** for details).

The Illumina HumanHT-12 v4 Expression BeadChip included 48K transcripts. Statistical analyses were limited to probes retained after applying the following QC elimination: non-detectable expression in $\geq 90\%$ of MESA samples using a detection p-value cut-off of 0.0001, overlap with a repetitive element or region, low variance across the samples ($< 10^{\text{th}}$ percentile), or putative and/or not well-characterized genes, i.e. gene names starting with KIAA, FLJ, HS, Cxorf, MGC, or LOC. We included 14,619 gene transcripts (10,898 unique genes) for analysis.

The Illumina HumanMethylation450 BeadChip included probes for 485K CpG sites. Of these 485K CpG sites, 448,588 passed the QC elimination criteria including: “detected” methylation levels $< 90\%$ of MESA samples using a detection p-value cut-off of 0.05, existence of any SNPs within 10 base pairs of the targeted CpG site, or overlap with a repetitive element or region.

These pre-processing pipelines effectively removed large effects of sample position on chips. However, probe-specific position effects existed for some CpG sites. Therefore, sample position on chips was adjusted in the analysis of methylation data.

Measurement of BMI, inflammatory stress, T2D and CAC

Weight was measured with a Detecto Platform Balance Scale to the nearest 0.5 kg. Height was measured with a stadiometer (Accu-Hite Measure Device with level bubble) to the nearest 0.1 cm. BMI was defined as weight in kilograms divided by square of height in meters. Plasma IL-6 was measured by ultra-sensitive ELISA (Quantikine HS Human IL-6 Immunoassay; R&D

Systems, Minneapolis, MN). Plasma CRP was measured using the BNII nephelometer (N High Sensitivity CRP; Dade Behring Inc., Deerfield, IL).

T2D was defined as fasting glucose ≥ 7.0 mmol/L (≥ 126 mg/dL) or use of hypoglycemic medication, and impaired fasting glucose was defined as fasting glucose 5.6-6.9 mmol/L (100-125 mg/dL). Fasting serum glucose at each examination was measured by rate reflectance spectrophotometry using thin-film adaptation of the glucose oxidase method on the Vitros analyzer (Johnson & Johnson Clinical Diagnostics, Rochester, NY). To achieve consistency of the serum glucose assay over examinations, 200 samples from each examination were reanalyzed and then recalibrated the original observations. The CT Reading Center for cardiac scans in the MESA is at UCLA-Biomedical Research Institute. CAC is scored using the Agatston method, which accounts for both lesion area and calcium density using Hounsfield brightness. The re-read agreement for the Agatston score (intraclass correlation coefficient, 0.99) was excellent.

Association analysis

Association analyses were performed using the linear model (*lm*) function and the *stepAIC* function of the *MASS* package in *R*. To identify gene transcripts or methylation sites associated with BMI, we fit separate linear regression models with BMI as a predictor of each transcript expression or the M-value for each CpG site. Covariates were age, gender, race/ethnicity, study site, technical covariates (expression/methylation chip, sample position on chips for methylation analyses), and residual sample contamination with non-targeted cells (e.g. non-monocytes, see below). To identify methylation sites associated with *cis*-gene expression, we fit separate linear regression models with the M-value for each CpG site (adjusted for methylation chip and position effects) as a predictor of transcript expression for any autosomal gene within 1 Mb of

the CpG site in question. Covariates were age, gender, and race/ethnicity, study site, expression chip, and residual sample contamination with non-targeted cells. Gender- and ethnicity-stratified analyses were performed as an internal validation and check of generalizability. P-values were adjusted for multiple testing using the q-value false discovery rate (FDR) method (9). To minimize false-positive results, we used a commonly used FDR threshold of 0.05. All FDR control was performed at the genome- and epigenome-wide scales for the tested gene transcripts and methylation probes, respectively.

Stability analysis enhanced weighted gene co-expression network identification

We applied the weighted correlation network analysis as implemented in the R package WGCNA (10) to construct network modules of highly correlated transcripts. We have used the less stringent threshold of $FDR \leq 0.10$ to pre-select a subset of 2,807 BMI-associated genes in order to satisfy the scale-free topology criterion. . We first obtained a weighted network based on the pairwise correlations among all transcripts considered, with an adjustment to produce a scale-free topology. Then, using the topological overlap measure to estimate the network interconnectedness, the transcripts were hierarchically clustered. With respect to the default parameters of WGCNA, we changed only the correlation type from Pearson to biweight midcorrelation (which is more robust to outliers) and set the minimum size for module detection to 10 (see **Supporting document** for detailed assessment of the stability of the identified networks).

Using this novel approach of stability analysis enhanced WGCNA, we then computed, for each consensus module, the eigengene defined as the first right-singular vector in the singular value decomposition of the standardized $p_m \times N$ expression matrix (or the first eigenvector of the $N \times N$

correlation matrix, $N = 1,264$) based on the p_m genes within module m . Association analyses of the eigengenes of individual sub-networks (modules) with T2D and CAC were performed using logistic and linear regression, respectively. CAC was log transformed ($\log(\text{Agatston score} + 1)$). We fit separate models with the first eigenvector for each network as a predictor of T2D or CAC. Covariates were age, gender, race/ethnicity, study site, expression/methylation chip, methylation position (for CpG methylation analyses only), and residual sample contamination with non-monocytes.

Functional annotation analysis

DAVID Bioinformatics Resources 6.7 was used to examine enrichment of GO pathway genes among BMI-associated genes ($\text{FDR} < 0.05$) (11). The Encyclopedia of DNA Elements (ENCODE)(12) project data was accessed through the UCSC Genome Browser (13) at <http://genome.ucsc.edu/> to evaluate potential functional regions surrounding BMI-associated methylation sites. Downloaded data sets included histone modifications detected in a CD14+ sample (ChIP-seq: H3K4m1/H3K4me3) from the Broad Institute/Bernstein lab, as well as DNaseI Hypersensitivity data in a CD14+ sample by Digital DNaseI from the University of Washington ENCODE group. These data were released by the ENCODE Consortium/Data Coordination Center (DCC) to be freely downloaded, analyzed, and published.

Analyses of risk gene dosage

Within the identified cholesterol (cyan) module of the co-expression network, we converted the continuous gene expression values to a dichotomized score coded 0 vs. 1 (-1) for an expression profile with positive (negative) association with BMI, with 1 (-1) representing expression values greater than the median value of this profile in the study population. We then sum the score for

the 11 genes as "risk score", defined as the number of risk genes for each participant. Association analyses for risk gene dosage with T2D and CAC were performed using the logistic and linear model, respectively. CAC was log transformed ($\log(\text{Agatston score} + 1)$). We fit separate models with risk gene dosage as a predictor of T2D or CAC. Covariates were age, gender, race/ethnicity, and study site. BMI, HDL cholesterol, triglycerides, systolic blood pressure, cigarette smoking, and physical activity were further adjusted in the models.

Mediation analyses

Methylation of DNA, which occurs predominately at CpG dinucleotides, has been viewed as an important potential regulator of gene expression (14). To investigate DNA methylation as a potential molecular link between BMI and gene expression, we performed mediation analyses under an assumed causal model in which BMI leads to a change in the methylation level of a CpG site, which at least partially mediates the effects of BMI on the expression of a nearby gene. The mediation analyses accounted for the biological and technical covariates and were performed by robust Structural Equation Modeling (SEM) as implemented in the R package lavaan (15). SEM analysis in general, and as implemented in lavaan, is based on Maximum Likelihood and the normal distribution, but provides several approaches to effectively deal with non-normal data. A first approach consists of computing robust standard errors (SE) by sandwich-type covariance matrices and scaled test statistics, in particular the Satorra-Bentler statistic whose amount of rescaling reflects the degree of kurtosis, while the second approach uses specific bootstrapping methods to obtain both SE and test statistics (15). The results we report are based on bootstrapping which we found to be somewhat more conservative than the use of robust SE and the Satorra-Bentler statistic. Lastly, mediation analyses for the eigengene of the cholesterol metabolism network (cyan) module were conducted in the same way and under

the assumed causal model that BMI leads to a change in T2D or CAC which is at least partially mediated by a change in gene expression in the cyan module.

RNA sequencing

A subset of 374 samples was randomly selected from the 1,264 MESA monocyte samples for RNA sequencing of mRNA (see **Supporting document** for details).

Replication study

Gene expression profiles from monocytes enriched from the whole blood of 1,285 German men and women were used to replicate main findings (see **Supporting document** for details).

Bisulfite treatment of genomic DNA and pyrosequencing

A subset of 176 samples was randomly selected from the 1,264 MESA monocyte samples for bisulfite treatment of genomic DNA and pyrosequencing (see **Supporting document** for details).

Weight loss intervention study

We quantified the transcriptome in pre- and post- intervention monocytes of 16 sedentary obese men and women, age 65-80 years, by leveraging an ongoing 5-month weight loss intervention trial. The weight loss intervention involved provision of a hypocaloric (-250-400 kcal/day) diet, behavioral counseling with registered dietitian, and 4 days/week of supervised moderate-intensity aerobic exercise training (16). An estimate of insulin sensitivity by the homeostasis model assessment (HOMA) score (17) was calculated with the formula: fasting plasma insulin ($\mu\text{U/ml}$) \times glucose (mmol/l)/22.5. We used the same protocol for monocyte purification and mRNA quantification as in MESA. Paired t-test was used to compare the cyan module gene

expression before and after the intervention. Pearson correlation coefficient was used to assess the relationships between changes in the cyan module gene expression and changes in HOMA measures.

Results

Transcriptome-wide analysis of BMI

The study population consisted of 1,264 MESA participants, 47% Caucasian, 32% Hispanic, 21% African American, 51% female, aged 55 – 94 years (**Table S1**). To determine if an obesity-associated transcriptional signature exists in circulating monocytes, we examined the relationships between BMI and individual gene expression level of 10,898 genes detected in monocytes. Using FDR threshold of 0.05, we identified 1,741 BMI-associated genes (**Figures 1 & S1**), among which the top ten most significantly BMI-associated gene transcripts were listed in **Table S2a**. Similar associations were observed for waist circumference (**Table S2b**). The top BMI-associated gene was guanine nucleotide binding protein gamma 10 (*GNG10*, $p=3.44 \times 10^{-23}$, **Table S2a**). *GNG10* encodes a G-protein gamma subunit which facilitates G-protein activation of Phospholipase A2 (PLA2) and the subsequent hydrolysis of phospholipids to form fatty acid and lysophospholipid products. Gene set enrichment analysis using DAVID (11) revealed that, overall, lipid biosynthesis genes were significantly over-represented among all the BMI-associated transcripts. Other biological processes significantly enriched among the BMI-associated genes included mitochondrial transport, apoptosis, inflammation, and glucose metabolism (**Table S3**).

Co-expression Network analysis of BMI: Identification of the Cholesterol Metabolism Gene Network

Next, we performed a weighted gene co-expression network analysis using the R package WGCNA (10) and a network stabilization procedure to identify robust networks of genes with coordinated gene expression patterns associated with BMI. The identified consensus modules are labeled by different colors (with one color, grey, denoting genes not assigned to modules), hence we will refer to specific modules below by their color (e.g., cyan). We identified 15 consensus modules associated with BMI (p-value range equal to 1.60×10^{-18} – 7.39×10^{-6} ; **Figure 2** and **Tables S4a and S4b**). Gene Ontology (GO) or Kyoto Encyclopedia of Genes and Genomes (KEGG) terms were significantly enriched (FDR<0.05) in five of the 15 modules identified, including sterol metabolism (cyan), antigen processing and presentation (magenta), non-coding RNA metabolism (brown), and ribosome machinery/translation (black and lightcyan) (**Figure 2** and **Table S5**).

The co-expressed network module most significantly associated with BMI was the cyan module ($p = 1.6 \times 10^{-18}$), which contained 11 functionally coupled genes related to sterol metabolism (**Figure 3** and **Table S5**), including eight genes which were up-regulated with increasing BMI (*LDLR*, *HMGCS1*, *FDFT1*, *SQLE*, *CYP51A1*, *SC4MOL*, *SCD*, and *FADS1*), and three which were down-regulated with increasing BMI (*MYLIP*, *ABCG1*, and *ABCA1*). These genes are known to be key contributors to cholesterol uptake (\uparrow *LDLR*, \downarrow *MYLIP*) (18;19), fatty acid and cholesterol synthesis (\uparrow *SCD*, *FADS1*, *HMGCS1*, *FDFT1*, *SQLE*, *CYP51A1*, *SC4MOL*) (20-22), and cholesterol efflux (\downarrow *ABCA1*, *ABCG1*) (23). Collectively, the expression profiles of the cyan module genes suggest that obesity is associated with cholesterol and lipid accumulation in circulating monocytes.

To determine if the relationship between cyan module gene expression and BMI was driven by the serum lipid profiles, we examined the association between the eigengene for the cyan module

and BMI after adjustment for plasma high-density lipoprotein (HDL), low-density lipoprotein (LDL), and triglycerides measured at the same time as the blood draw for monocyte separation. In this analysis, the association with BMI was attenuated, but remained significant ($p = 8.7 \times 10^{-8}$). We observed that LDL, HDL, and triglycerides levels were negatively associated with cholesterol synthesis/uptake gene expression ($p < 2.2 \times 10^{-16}$), and positively associated with cholesterol efflux gene expression ($p < 2 \times 10^{-16}$). To further evaluate the internal validity of the cyan module, we examined the association of individual cyan module genes with BMI across sex and race subgroups. In general, the associations were qualitatively consistent across subgroups and all p-values for interactions between sex or race and gene expression levels were non-significant ($p > 0.05$, **Table S6**). In addition, we validated the gene expression signals using RNA sequencing data in a subset of samples (N=374; **Table S7**).

We then performed similar analyses of the cyan module genes in 1,285 German men and women from the Gutenberg Heart Study (GHS) using monocytes enriched from whole blood with the RosetteSep monocyte enrichment cocktail (24). Seven of the eleven cyan module genes (*ABCG1*, *MYLIP*, *SQLE*, *ABCA1*, *SC4MOL*, *CYP51A1*, and *FDFT1*) were replicated in the GHS cohort (genome-wide FDR<0.05; **Table S8**). However, the cyan module was not detected in the GHS, which could be due to the expected low purity of monocytes in the GHS given the different methods we had for monocyte collection (see the methods in the **Supporting document**).

Weight Loss and the Cholesterol Metabolism Gene Network

We further investigated the effects of a 5-month weight loss intervention on the cyan module gene expression in 16 obese older adults. Mean weight loss was $-6.7 \pm 1.1\%$, and the HOMA measure of insulin resistance decreased an average of $33 \pm 8\%$. The intervention significantly

down-regulated *SQLE* (p: 0.04) while its association with other members of the cyan modules did not reach statistical significance (p: 0.06-0.83, **Table S9**), although the effect directions were all consistent with the observed BMI associations in MESA. Furthermore, changes in the HOMA measure of insulin resistance were inversely associated with changes in *ABCA1* and *MYLIP* (p: 0.05) while the association of HOMA changes with other members of the cyan modules did not reach statistical significance (p: 0.08-0.95, **Table S9**), although the effect directions were all consistent with the observed T2D associations in MESA.

Associations of Cholesterol Metabolism Gene Network with T2D and CAC

To determine if the cyan (or any of the other BMI-associated networks) were related to T2D or cardiovascular disease, we examined the association between the eigengene of each BMI-associated network and prevalent T2D (excluding individuals with impaired fasting glucose) or CAC in the MESA participants (using a Bonferroni adjusted significance threshold of $p < 0.003$ (0.05/15 BMI-associated networks)). In this analysis, the cyan module was the most significantly associated with both T2D and CAC (**Figure 2** and **Tables S4a and S4b**). These associations did not differ by the status of statin use (**Table S10**), and were attenuated but significant after adjustment for statin use (T2D: odds ratio per SD increment: 1.47, p-value: 2.6×10^{-5} ; CAC: fold change per SD increment: 1.13; p-value: 0.046). To alleviate the concern that the detection of cyan module may be influenced by statin use or pre-selection of BMI-associated genes, we performed an unsupervised WGCNA using the full set of genes after adjusting the expression values for the statin effect and confirmed that the cyan module remained present.

We also created a cyan module risk score, defined as the number of risk genes for each participant. There was an increase in the odds of having T2D (odds ratio: 5.4; 95% CI: 3.0 – 9.4;

p-value: 5.9×10^{-9} ; excluding individuals with impaired fasting glucose) associated with a risk score of 9 – 11 vs. a score of 0 – 2 genes, and the test for a linear trend across the full range of the risk score (0 – 11) was also highly significant ($p_{trend}: 5.07 \times 10^{-10}$, **Figure 4**). Likewise, the amount of CAC was significantly increased in subjects with a score of 9 – 11 vs. 0 – 2 (fold change: 2.1; 95% CI: 1.4 – 3.2; p-value: 0.001), and the corresponding p_{trend} was 2.02×10^{-3} . These associations were attenuated but remained significant after additional adjustment for smoking, BMI, HDL cholesterol, triglycerides, systolic blood pressure, and physical activity. BMI were positively associated with both T2D and CAC ($p: 3.43 \times 10^{-21}$, $p: 6.05 \times 10^{-4}$, respectively). Mediation analyses using robust Structural Equation Modeling (SEM) showed significant indirect effects of BMI on T2D and CAC that were mediated through the eigengene of the cyan network ($p: 9.27 \times 10^{-4}$, $p: 0.015$, respectively). However, this type of mediation analysis is not sufficient to derive a causal relationship since we cannot rule out other possible competing causal models.

Associations of Cholesterol Metabolism Gene Network with Inflammatory Mediators

To investigate systemic inflammation in relation to the cholesterol metabolism gene network variation, we also tested systemic inflammatory markers, plasma interleukin 6 (IL-6) and C-reactive protein (CRP) measured at MESA Exam 1 as correlates for the gene expression network profiles measured 9 years later in monocytes (MESA Exam 5). In these analyses, the cyan module eigengene significantly associated with logIL-6 (beta: 0.007; $p: 2.0 \times 10^{-7}$) and logCRP measurements (beta: 0.004; $p: 1.7 \times 10^{-7}$, **Table S4b**).

Methylation Associations with the Cholesterol Metabolism Gene Network

To investigate epigenetic modification of DNA as a potential regulator of the cholesterol metabolism gene expression, we integrated genome-wide DNA methylation data from our methylomics of gene expression study (25), which was measured in the same 1,264 monocyte samples, and tested for associations between expression of the cholesterol metabolism network genes and DNA methylation of CpG sites located near the network genes.

Using an FDR threshold of 0.001 (a more stringent FDR threshold used to retain less false positives), we identified 30 CpG sites (of 6,253 CpG sites located within 1 MB of any cyan network gene) that were *cis*-gene expression associated methylation sites (eMS), whose degree of methylation was linked to expression of nine of the cyan network genes (*ABCA1*, *ABCG1*, *CYP51A1*, *FADS1*, *FDFT1*, *LDLR*, *SC4MOL*, *SCD*, and *SQLE*, **Table S11**). Among these 30 eMS, only four CpG sites had methylation profiles also associated with BMI (cg06500161, cg05323251, cg10192877, and cg05119988, FDR<0.001; **Table 1**). Mediation analyses, using SEM and accounting for population and technical covariates, provided statistical evidence that the relationship between BMI and expression of three cyan module genes (*ABCG1*, *SC4MOL*, and *LDLR*) was at least partially mediated through methylation (See “Indirect” effects in **Table 1**). *ABCG1* harbored two eMS, including the eMS most strongly correlated with BMI (cg06500161, $p=2.8 \times 10^{-7}$, $\text{FDR}_{\text{epigenome-wide}}=7.4 \times 10^{-4}$; **Figure S2** and **Table 1**). Methylation profiles of these two *ABCG1*-eMS were independently associated with decreasing expression of *ABCG1* (from multiple regression analysis, P-value<0.05). *ABCG1*-eMS are located within the *ABCG1* gene body, and occupy ENCODE-annotated regulatory sites (12;13), characterized by features indicative of active *cis*-regulatory sequences (DNaseI hypersensitive hotspots and enhancer histone methylation marks, H3K4m1/H3K4m3). In a subset of samples (N=176),

pyrosequencing-based validation of the *ABCG1* eMS corroborated the association between BMI and methylation of cg06500161 ($p: 3.35 \times 10^{-4}$ for pyrosequencing methylation vs. BMI).

Discussion

In summary, we discovered a large number of genes whose expression is strongly associated with BMI. Notable among these genes is a well-defined network of co-expressed cholesterol metabolism genes (the cyan network) whose coordinated action would be expected to increase intracellular cholesterol content. Furthermore, we discovered that expression of the co-expressed cholesterol metabolism genes at Exam 5 was also correlated with IL-6 and CRP measured at Exam 1 although our data cannot infer their temporal relationships. These findings are intriguing since cellular cholesterol homeostasis is expected to be tightly regulated by a cholesterol-mediated feedback system that depends on two transcription factors, sterol response element-binding proteins (SREBP) and liver X receptor ($LXR\alpha$). When plasma LDL levels are high, more LDL uptake by cells through LDLR potentially leads to a compensatory down-regulation of cellular cholesterol synthesis and uptake genes, which may explain the observed negative relationship between the plasma LDL levels and the expression of cholesterol synthesis and uptake genes. However, this negative feedback regulation of cholesterol metabolism can be overridden by inflammatory stress. *In vitro* and *in vivo* studies confirm that inflammatory stress results in altered expression of the network genes and can increase intracellular cholesterol content in several cell types including hepatocytes, vascular smooth muscle cells, and macrophages (26-30). In addition, human experimental data from us (monocytes) and others (adipose tissue) (31-34) show that weight loss rebalanced expression of members of the cholesterol metabolism gene network. These data suggest that obesity and/or inflammatory stress may predispose intracellular cholesterol accumulation.

Importantly, we discovered that the same lipid metabolic co-expression network is also strongly associated with both CAC and T2D. Cholesterol-loading of macrophages (foam cell formation) is typically viewed as a critical cellular component of the atherosclerotic lesion (35), while the pathophysiology of obesity-related T2D involves fatty acids and triglycerides, rather than cholesterol as in atherosclerosis. However, recently several lines of experimental evidence (36-43) have shown that disruption of cellular cholesterol homeostasis in several cell types (e.g. macrophage, β -cell, adipocyte) can lead to pathological processes preceding T2D. Our data for the first time made the link between potential disruption of cellular cholesterol homeostasis and CAC or T2D using human monocytes in the population setting. The longitudinal relationships between changes in members of the cyan module gene expression and changes in the HOMA measure observed in the small weight loss intervention study indirectly support the cross-sectional findings in MESA. Taken together with the aggregation of previously reported evidence, our data suggest an emerging hypothesis that the cholesterol metabolism gene network represents a molecular link between obesity/inflammation and its most important complications - T2D and cardiovascular disease. If confirmed, it will provide a rationale for developing novel therapeutics targeting intracellular cholesterol by systematically modulating this cholesterol metabolism gene network, rather than individual genes, for optimizing the prevention and treatment of obesity-related diseases.

We also report a convergence of DNA methylation and gene expression data. Our findings of BMI-associated increases in methylation that are linked to expression alterations of several members of the cholesterol metabolism gene network, coupled with the mediation analyses, suggest a hypothesis that obesity leads to changes in the molecular architecture of monocytes, in part through alteration of DNA methylation. ABCG1 plays a critical role in mediating

cholesterol efflux and preventing cellular lipid accumulation (44); however, effects of genetic polymorphisms on *ABCG1* expression have not been reported from expression quantitative trait loci (eQTLs) studies or observed in our study. The predicted regulatory regions of *ABCG1* which harbor expression-associated methylation sites (such as the genomic loci surrounding cg06500161) merit special attention as potential targets for novel therapies designed to attenuate the adverse metabolic and clinical effects of obesity. Additional functional evaluation of the relationship between methylation of this genomic loci and gene expression are warranted.

As a cross-sectional study, the associations observed in MESA do not support causation (in either direction). While our weight loss intervention data support that obesity precedes the cholesterol metabolism gene changes in monocytes, inferences should be made with caution due to the small sample size and lack of a non-weight loss and non-exercise control group. We also cannot discern whether obesity itself or other unmeasured obesity-related factors are directly related to expression of the cholesterol metabolism genes. Furthermore, the weight loss intervention trial cannot reveal the temporal relationships between the cholesterol metabolism gene network and cardiometabolic traits. More research is needed to further clarify the causes as well as the functional and clinical consequences of obesity-related alterations in expression of this complex network of genes. Although BMI is a commonly used indicator of general obesity, it is suboptimal for estimation of visceral adiposity, which may be the best predictor for obesity-related metabolic and cardiovascular risks (45). However, our data showed that the BMI-associations were similar to the associations with waist circumference, a surrogate of visceral adiposity (46).

A distinction of our study is the use of purified circulating monocytes, as opposed to mixed human cell types (e.g. whole blood or adipose tissue) (47-49) which may produce type I and type

IL errors (50;51). We choose to study monocytes because they are key cells of innate immunity and major contributors to the pathogenesis of inflammatory diseases including T2D and cardiovascular disease (52). The involvement of circulating monocytes has been further indicated by other lines of evidence. For example, high-fat meals can promote lipid-loading in circulating monocytes even before their migration into tissues and differentiation to macrophages (53). It is also plausible to hypothesize that the associations with the cholesterol metabolism network genes that we have observed in monocyte samples may also reflect changes in other tissues or cells which are not accessible in the population setting, e.g. hepatic cells, β -cells. Supporting this notion, inflammatory stimuli produced gene expression changes in several cell types, including hepatocytes and vascular smooth muscle cells (27;30), which are consistent with the correlations we observed between IL-6 and the network genes in human monocytes samples. We and others have also highlighted the importance of β -cell cholesterol homeostasis in T2D (43;54). Collectively, our data suggest that circulating monocytes can serve as a useful *in vivo* model to better understand complex molecular mechanisms underlying obesity-related diseases.

Acknowledgements

Dr. Yongmei Liu is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions, and thank Carole Grenier for provision of technical support. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>. The MESA Epigenomics & Transcriptomics Study was funded by NHLBI grant

R01HL101250 to Wake Forest University Health Sciences. The weight loss intervention study was supported by NIH grant R01 HL093713 and the WFU Claude D. Pepper Older Americans Independence Center (P30-AG21332). This research was supported by contracts N01-HC-from the National Heart, Lung and Blood Institute, by grants UL1-RR-024156 and UL1-RR-025005 from the National Center for Research Resources. For MESA investigators - Y.L. oversaw the study, carried out experiments, researched the data, and wrote the manuscript. J.D. carried out experiments, researched the data, and wrote the manuscript. I.H., C.E.M, D.H., R.P.T., and L.M.R. carried out experiments, researched the data, and helped write the manuscript. K.L., B.J.N., S.B.K., Z.H., A.d.I.F., N.S., R.E.S., C.-C.C., T.H., J.S.P., S.M., D.J., and W.P. carried out experiments and researched the data. N.X., M.O.G., Y.-D.I.C., J.I.R., D.S.S. researched the data. For GHS investigators - Acquisition of data: T.Z. and P.S.W.; analysis and interpretation of data: C.M. and T.Z.; reviewing and editing of the manuscript: C.M., T.Z., and S.B.; obtained funding: S.B. and P.S.W.; study supervision: S.B. and T.Z. No conflicts exist.

Reference List

1. Rocha,VZ, Libby,P: Obesity, inflammation, and atherosclerosis. *Nat Rev Cardiol* 6:399-409, 2009
2. Weisberg,SP, McCann,D, Desai,M, Rosenbaum,M, Leibel,RL, Ferrante,AW, Jr.: Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest* 112:1796-1808, 2003
3. Xu,H, Barnes,GT, Yang,Q, Tan,G, Yang,D, Chou,CJ, Sole,J, Nichols,A, Ross,JS, Tartaglia,LA, Chen,H: Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. *J Clin Invest* 112:1821-1830, 2003
4. Arkan,MC, Hevener,AL, Greten,FR, Maeda,S, Li,ZW, Long,JM, Wynshaw-Boris,A, Poli,G, Olefsky,J, Karin,M: IKK-beta links inflammation to obesity-induced insulin resistance. *Nat Med* 11:191-198, 2005
5. Hotamisligil,GS: Inflammation and metabolic disorders. *Nature* 444:860-867, 2006
6. Moore,KJ, Tabas,I: Macrophages in the pathogenesis of atherosclerosis. *Cell* 145:341-355, 2011
7. Petronis,A: Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 465:721-727, 2010

8. Bild,DE, Bluemke,DA, Burke,GL, Detrano,R, ez Roux,AV, Folsom,AR, Greenland,P, Jacob,DR, Jr., Kronmal,R, Liu,K, Nelson,JC, O'Leary,D, Saad,MF, Shea,S, Szklo,M, Tracy,RP: Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 156:871-881, 2002
9. Storey,JD, Tibshirani,R: Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100:9440-9445, 2003
10. Langfelder,P, Horvath,S: WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559, 2008
11. Dennis,G, Jr., Sherman,BT, Hosack,DA, Yang,J, Gao,W, Lane,HC, Lempicki,RA: DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4:3, 2003
12. Dunham,I, Kundaje,A, Aldred,SF, Collins,PJ, Davis,CA, Doyle,F, Epstein,CB, Frietze,S, Harrow,J, Kaul,R, Khatun,J, Lajoie,BR, Landt,SG, Lee,BK, Pauli,F, Rosenbloom,KR, Sabo,P, Safi,A, Sanyal,A, Shores,N, Simon,JM, Song,L, Trinklein,ND, Altshuler,RC, Birney,E, Brown,JB, Cheng,C, Djebali,S, Dong,X, Dunham,I, Ernst,J, Furey,TS, Gerstein,M, Giardine,B, Greven,M, Hardison,RC, Harris,RS, Herrero,J, Hoffman,MM, Iyer,S, Kellis,M, Khatun,J, Kheradpour,P, Kundaje,A, Lassman,T, Li,Q, Lin,X, Marinov,GK, Merkel,A, Mortazavi,A, Parker,SC, Reddy,TE, Rozowsky,J, Schlesinger,F, Thurman,RE, Wang,J, Ward,LD, Whitfield,TW, Wilder,SP, Wu,W, Xi,HS, Yip,KY, Zhuang,J, Bernstein,BE, Birney,E, Dunham,I, Green,ED, Gunter,C, Snyder,M, Pazin,MJ,

Lowdon,RF, Dillon,LA, Adams,LB, Kelly,CJ, Zhang,J, Wexler,JR, Green,ED, Good,PJ, Feingold,EA, Bernstein,BE, Birney,E, Crawford,GE, Dekker,J, Elinitski,L, Farnham,PJ, Gerstein,M, Giddings,MC, Gingeras,TR, Green,ED, Guigo,R, Hardison,RC, Hubbard,TJ, Kellis,M, Kent,WJ, Lieb,JD, Margulies,EH, Myers,RM, Snyder,M, Starnatoyannopoulos,JA, Tennebaum,SA, Weng,Z, White,KP, Wold,B, Khatun,J, Yu,Y, Wrobel,J, Risk,BA, Gunawardena,HP, Kuiper,HC, Maier,CW, Xie,L, Chen,X, Giddings,MC, Bernstein,BE, Epstein,CB, Shores,N, Ernst,J, Kheradpour,P, Mikkelsen,TS, Gillespie,S, Goren,A, Ram,O, Zhang,X, Wang,L, Issner,R, Coyne,MJ, Durham,T, Ku,M, Truong,T, Ward,LD, Altshuler,RC, Eaton,ML, Kellis,M, Djebali,S, Davis,CA, Merkel,A, Dobin,A, Lassmann,T, Mortazavi,A, Tanzer,A, Lagarde,J, Lin,W, Schlesinger,F, Xue,C, Marinov,GK, Khatun,J, Williams,BA, Zaleski,C, Rozowsky,J, Roder,M, Kokocinski,F, Abdelhamid,RF, Alioto,T, Antoshechkin,I, Baer,MT, Batut,P, Bell,I, Bell,K, Chakraborty,S, Chen,X, Chrast,J, Curado,J, Derrien,T, Drenkow,J, Dumais,E, Dumais,J, Duttagupta,R, Fastuca,M, Fejes-Toth,K, Ferreira,P, Foissac,S, Fullwood,MJ, Gao,H, Gonzalez,D, Gordon,A, Gunawardena,HP, Howald,C, Jha,S, Johnson,R, Kapranov,P, King,B, Kingswood,C, Li,G, Luo,OJ, Park,E, Preall,JB, Presaud,K, Ribeca,P, Risk,BA, Robyr,D, Ruan,X, Sammeth,M, Sandu,KS, Schaeffer,L, See,LH, Shahab,A, Skancke,J, Suzuki,AM, Takahashi,H, Tilgner,H, Trout,D, Walters,N, Wang,H, Wrobel,J, Yu,Y, Hayashizaki,Y, Harrow,J, Gerstein,M, Hubbard,TJ, Reymond,A, Antonarakis,SE, Hannon,GJ, Giddings,MC, Ruan,Y, Wold,B, Carninci,P, Guigo,R, Gingeras,TR, Rosenbloom,KR, Sloan,CA, Learned,K, Malladi,VS, Wong,MC, Barber,GP, Cline,MS, Dreszer,TR, Heitner,SG, Karolchik,D, Kent,WJ, Kirkup,VM, Meyer,LR, Long,JC, Maddren,M, Raney,BJ, Furey,TS, Song,L, Grasfeder,LL, Giresi,PG, Lee,BK,

- Battenhouse,A, Sheffield,NC, Simon,JM, Showers,KA, Safi,A, London,D, Bhinge,AA, Shestak,C, Schaner,MR, Kim,SK, Zhang,ZZ, Mieczkowski,PA, Mieczkowska,JO, Liu,Z, McDaniell,RM, Ni,Y: An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57-74, 2012
13. Rosenbloom,KR, Sloan,CA, Malladi,VS, Dreszer,TR, Learned,K, Kirkup,VM, Wong,MC, Maddren,M, Fang,R, Heitner,SG, Lee,BT, Barber,GP, Harte,RA, Diekhans,M, Long,JC, Wilder,SP, Zweig,AS, Karolchik,D, Kuhn,RM, Haussler,D, Kent,WJ: ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res* 41:D56-D63, 2013
14. Jones,PA: Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13:484-492, 2012
15. Rosseel,Y: lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* 48:1-36, 2012
16. Nicklas,BJ, Gaukstern,JE, Legault,C, Leng,I, Rejeski,WJ: Intervening on spontaneous physical activity to prevent weight regain in older adults: design of a randomized, clinical trial. *Contemp Clin Trials* 33:450-455, 2012
17. Matthews,DR, Hosker,JP, Rudenski,AS, Naylor,BA, Treacher,DF, Turner,RC: Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412-419, 1985

18. Brown,MS, Goldstein,JL: Receptor-mediated endocytosis: insights from the lipoprotein receptor system. *Proc Natl Acad Sci U S A* 76:3330-3337, 1979
19. Zelcer,N, Hong,C, Boyadjian,R, Tontonoz,P: LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. *Science* 325:100-104, 2009
20. Flowers,MT, Ntambi,JM: Role of stearoyl-coenzyme A desaturase in regulating lipid metabolism. *Curr Opin Lipidol* 19:248-256, 2008
21. Glaser,C, Heinrich,J, Koletzko,B: Role of FADS1 and FADS2 polymorphisms in polyunsaturated fatty acid metabolism. *Metabolism* 59:993-999, 2010
22. Seth,G, McIvor,RS, Hu,WS: 17Beta-hydroxysteroid dehydrogenase type 7 (Hsd17b7) reverts cholesterol auxotrophy in NS0 cells. *J Biotechnol* 121:241-252, 2006
23. Wang,X, Collins,HL, Ranalletta,M, Fuki,IV, Billheimer,JT, Rothblat,GH, Tall,AR, Rader,DJ: Macrophage ABCA1 and ABCG1, but not SR-BI, promote macrophage reverse cholesterol transport in vivo. *J Clin Invest* 117:2216-2224, 2007
24. Zeller,T, Wild,P, Szymczak,S, Rotival,M, Schillert,A, Castagne,R, Maouche,S, Germain,M, Lackner,K, Rossmann,H, Eleftheriadis,M, Sinning,CR, Schnabel,RB, Lubos,E, Mennerich,D, Rust,W, Perret,C, Proust,C, Nicaud,V, Loscalzo,J, Hubner,N, Tregouet,D, Munzel,T, Ziegler,A, Tiret,L, Blankenberg,S, Cambien,F: Genetics and

- beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One* 5:e10693, 2010
25. Liu,Y, Ding,J, Reynolds,LM, Lohman,K, Register,TC, de la Fuente,A, Howard,TD, Hawkins,GA, Cui,W, Morris,J, Smith,SG, Barr,RG, Kaufman,JD, Burke,GL, Post,W, Shea,S, McCall,CE, Siscovick,D, Jacobs,DR, Jr., Tracy,RP, Herrington,DM, Hoeschele,I: Methyloomics of gene expression in human monocytes. *Hum Mol Genet* 22:5065-5074, 2013
 26. Zhao,L, Chen,Y, Tang,R, Chen,Y, Li,Q, Gong,J, Huang,A, Varghese,Z, Moorhead,JF, Ruan,XZ: Inflammatory stress exacerbates hepatic cholesterol accumulation via increasing cholesterol uptake and de novo synthesis. *J Gastroenterol Hepatol* 26:875-883, 2011
 27. Wang,Y, Moser,AH, Shigenaga,JK, Grunfeld,C, Feingold,KR: Downregulation of liver X receptor-alpha in mouse kidney and HK-2 proximal tubular cells by LPS and cytokines. *J Lipid Res* 46:2377-2387, 2005
 28. Khovidhunkit,W, Kim,MS, Memon,RA, Shigenaga,JK, Moser,AH, Feingold,KR, Grunfeld,C: Effects of infection and inflammation on lipid and lipoprotein metabolism: mechanisms and consequences to the host. *J Lipid Res* 45:1169-1196, 2004
 29. Ruan,XZ, Moorhead,JF, Fernando,R, Wheeler,DC, Powis,SH, Varghese,Z: Regulation of lipoprotein trafficking in the kidney: role of inflammatory mediators and transcription factors. *Biochem Soc Trans* 32:88-91, 2004

30. Chen,Y, Ku,H, Zhao,L, Wheeler,DC, Li,LC, Li,Q, Varghese,Z, Moorhead,JF, Powis,SH, Huang,A, Ruan,XZ: Inflammatory stress induces statin resistance by disrupting 3-hydroxy-3-methylglutaryl-CoA reductase feedback regulation. *Arterioscler Thromb Vasc Biol* 34:365-376, 2014
31. Mutch,DM, Pers,TH, Temanni,MR, Pelloux,V, Marquez-Quinones,A, Holst,C, Martinez,JA, Babalis,D, van Baak,MA, Handjieva-Darlenska,T, Walker,CG, Astrup,A, Saris,WH, Langin,D, Viguerie,N, Zucker,JD, Clement,K: A distinct adipose tissue gene expression response to caloric restriction predicts 6-mo weight maintenance in obese subjects. *Am J Clin Nutr* 94:1399-1409, 2011
32. Dahlman,I, Linder,K, Arvidsson,NE, Andersson,I, Liden,J, Verdich,C, Sorensen,TI, Arner,P: Changes in adipose tissue gene expression with energy-restricted diets in obese women. *Am J Clin Nutr* 81:1275-1285, 2005
33. Johansson,LE, Danielsson,AP, Parikh,H, Klintonberg,M, Norstrom,F, Groop,L, Ridderstrale,M: Differential gene expression in adipose tissue from obese human subjects during weight loss and weight maintenance. *Am J Clin Nutr* 96:196-207, 2012
34. Shea,J, French,CR, Bishop,J, Martin,G, Roebathan,B, Pace,D, Fitzpatrick,D, Sun,G: Changes in the transcriptome of abdominal subcutaneous adipose tissue in response to short-term overfeeding in lean and obese men. *Am J Clin Nutr* 89:407-415, 2009

35. Tabas,I: Consequences of cellular cholesterol accumulation: basic concepts and physiological implications. *J Clin Invest* 110:905-911, 2002
36. Bie,J, Zhao,B, Song,J, Ghosh,S: Improved insulin sensitivity in high fat- and high cholesterol-fed Ldlr-/- mice with macrophage-specific transgenic expression of cholesteryl ester hydrolase: role of macrophage inflammation and infiltration into adipose tissue. *J Biol Chem* 285:13630-13637, 2010
37. Bie,J, Zhao,B, Marqueen,KE, Wang,J, Szomju,B, Ghosh,S: Macrophage-specific transgenic expression of cholesteryl ester hydrolase attenuates hepatic lipid accumulation and also improves glucose tolerance in ob/ob mice. *Am J Physiol Endocrinol Metab* 302:E1283-E1291, 2012
38. Le,LS, Ferre,P, Dugail,I: Adipocyte cholesterol balance in obesity. *Biochem Soc Trans* 32:103-106, 2004
39. Yu,BL, Zhao,SP, Hu,JR: Cholesterol imbalance in adipocytes: a possible mechanism of adipocytes dysfunction in obesity. *Obes Rev* 11:560-567, 2010
40. de,HW, Bhattacharjee,A, Ruddle,P, Kang,MH, Hayden,MR: ABCA1 in adipocytes regulates adipose tissue lipid content, glucose tolerance, and insulin sensitivity. *J Lipid Res* 55:516-523, 2014

41. Fryirs,M, Barter,PJ, Rye,KA: Cholesterol metabolism and pancreatic beta-cell function. *Curr Opin Lipidol* 20:159-164, 2009
42. Maxfield,FR, Tabas,I: Role of cholesterol and lipid organization in disease. *Nature* 438:612-621, 2005
43. Brunham,LR, Kruit,JK, Pape,TD, Timmins,JM, Reuwer,AQ, VasANJI,Z, Marsh,BJ, Rodrigues,B, Johnson,JD, Parks,JS, Verchere,CB, Hayden,MR: Beta-cell ABCA1 influences insulin secretion, glucose homeostasis and response to thiazolidinedione treatment. *Nat Med* 13:340-347, 2007
44. Kennedy,MA, Barrera,GC, Nakamura,K, Baldan,A, Tarr,P, Fishbein,MC, Frank,J, Francone,OL, Edwards,PA: ABCG1 has a critical role in mediating cholesterol efflux to HDL and preventing cellular lipid accumulation. *Cell Metab* 1:121-131, 2005
45. Britton,KA, Massaro,JM, Murabito,JM, Kreger,BE, Hoffmann,U, Fox,CS: Body fat distribution, incident cardiovascular disease, cancer, and all-cause mortality. *J Am Coll Cardiol* 62:921-925, 2013
46. Despres,JP, Lemieux,I: Abdominal obesity and metabolic syndrome. *Nature* 444:881-887, 2006
47. Emilsson,V, Thorleifsson,G, Zhang,B, Leonardson,AS, Zink,F, Zhu,J, Carlson,S, Helgason,A, Walters,GB, Gunnarsdottir,S, Mouy,M, Steinthorsdottir,V, Eiriksdottir,GH,

- Bjornsdottir,G, Reynisdottir,I, Gudbjartsson,D, Helgadottir,A, Jonasdottir,A, Jonasdottir,A, Styrkarsdottir,U, Gretarsdottir,S, Magnusson,KP, Stefansson,H, Fossdal,R, Kristjansson,K, Gislason,HG, Stefansson,T, Leifsson,BG, Thorsteinsdottir,U, Lamb,JR, Gulcher,JR, Reitman,ML, Kong,A, Schadt,EE, Stefansson,K: Genetics of gene expression and its effect on disease. *Nature* 452:423-428, 2008
48. Hulsmans,M, Geeraert,B, De,KD, Mertens,A, Lannoo,M, Vanaudenaerde,B, Hoylaerts,M, Benhabiles,N, Tsatsanis,C, Mathieu,C, Holvoet,P: Interleukin-1 receptor-associated kinase-3 is a key inhibitor of inflammation in obesity and metabolic syndrome. *PLoS One* 7:e30414, 2012
49. Feinberg,AP, Irizarry,RA, Fradin,D, Aryee,MJ, Murakami,P, Aspelund,T, Eiriksdottir,G, Harris,TB, Launer,L, Gudnason,V, Fallin,MD: Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med* 2:49ra67, 2010
50. Zhong,H, Yang,X, Kaplan,LM, Molony,C, Schadt,EE: Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am J Hum Genet* 86:581-591, 2010
51. Dimas,AS, Deutsch,S, Stranger,BE, Montgomery,SB, Borel,C, Attar-Cohen,H, Ingle,C, Beazley,C, Gutierrez,AM, Sekowska,M, Gagnebin,M, Nisbett,J, Deloukas,P, Dermitzakis,ET, Antonarakis,SE: Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 325:1246-1250, 2009

52. Shi,C, Pamer,EG: Monocyte recruitment during infection and inflammation. *Nat Rev Immunol* 11:762-774, 2011

53. Varela,LM, Ortega,A, Bermudez,B, Lopez,S, Pacheco,YM, Villar,J, Abia,R, Muriana,FJ: A high-fat meal promotes lipid-load and apolipoprotein B-48 receptor transcriptional activity in circulating monocytes. *Am J Clin Nutr* 93:918-925, 2011

54. Fryirs,M, Barter,PJ, Rye,KA: Cholesterol metabolism and pancreatic beta-cell function. *Curr Opin Lipidol* 20:159-164, 2009

Tables

Table 1 | Expression-associated methylation sites (eMS) that potentially mediate the relationship between BMI and cholesterol network gene expression

eMS		BMI-methylation				BMI-gene expression				
CpG ID	CpG location	Beta	SE	P-value	FDR	Gene	Indirect		Direct	
							Beta	P-value	Beta	P-value
cg06500161	Body (ABCG1)	0.005	0.001	2.8E-07	5.4E-04	<i>ABCG1</i>	-0.005	2.6E-06	-0.025	3.7E-12
cg05323251	Body (ICAM3)	0.003	0.001	3.0E-06	2.9E-03	<i>LDLR</i>	0.001	2.4E-03	0.012	1.5E-07
cg10192877	Body (ABCG1)	0.004	0.001	5.3E-05	0.02	<i>ABCG1</i>	-0.002	1.1E-03	-0.025	3.6E-14
cg05119988	5'UTR (SC4MOL)	-0.006	0.001	1.0E-04	0.03	<i>SC4MOL</i>	0.001	1.2E-03	0.009	1.2E-06

Effects of body mass index (BMI) on methylation of cyan module gene expression-associated methylation sites (Beta, standard error – SE, P-value, and genome-wide FDR), and the indirect (mediated) and direct effects (beta, P-value) of BMI on gene expression; indirect effects estimated using Structural Equation Modeling (SEM) for methylation mediating the effect of BMI on gene expression (\log_2 fold change of expression per 1 unit increase in BMI (kg/m^2)); some eMS were located within gene bodies (gene containing CpG in parentheses) or in the 5' untranslated regions (UTR) of genes. Analyses included 1,263 CD14+ monocyte samples and were adjusted for age, sex, race, study site, and residual sample contamination with non-monocytes.

Figure Legends

Figure 1 | Schematic summary of main findings in MESA.

Figure 2 | Gene co-expression network modules identified associating with BMI.

Significance ($-\log_{10}$ P-value on the y-axis) resulting from the association of each consensus network module of co-expressed genes (eigengene) with BMI (black), T2D (dotted), and CAC (grey) in 1,263 monocyte samples; dashed line represents a Bonferroni correction for fifteen modules ($\alpha = 0.05$). The number of genes and the biological pathways significantly enriched within each module (from DAVID, $FDR < 0.05$) are also shown below, corresponding to each co-expression network module (x-axis). See also **Table S4 & S5**.

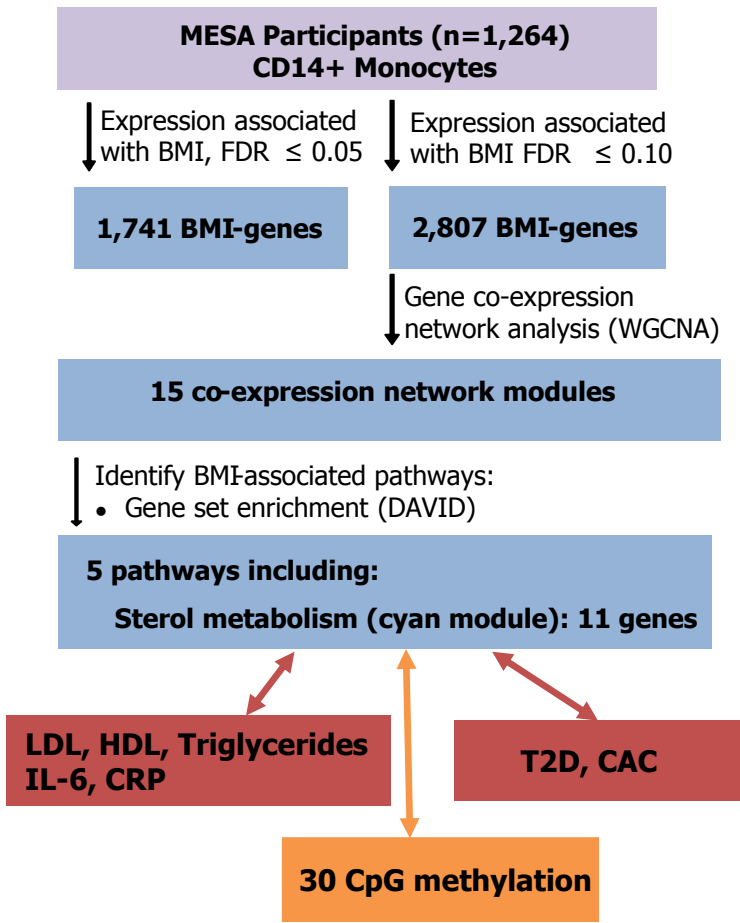
Figure 3 | BMI-associated expression patterns of the 11 inter-correlated genes in the cyan module reveal altered cholesterol regulation.

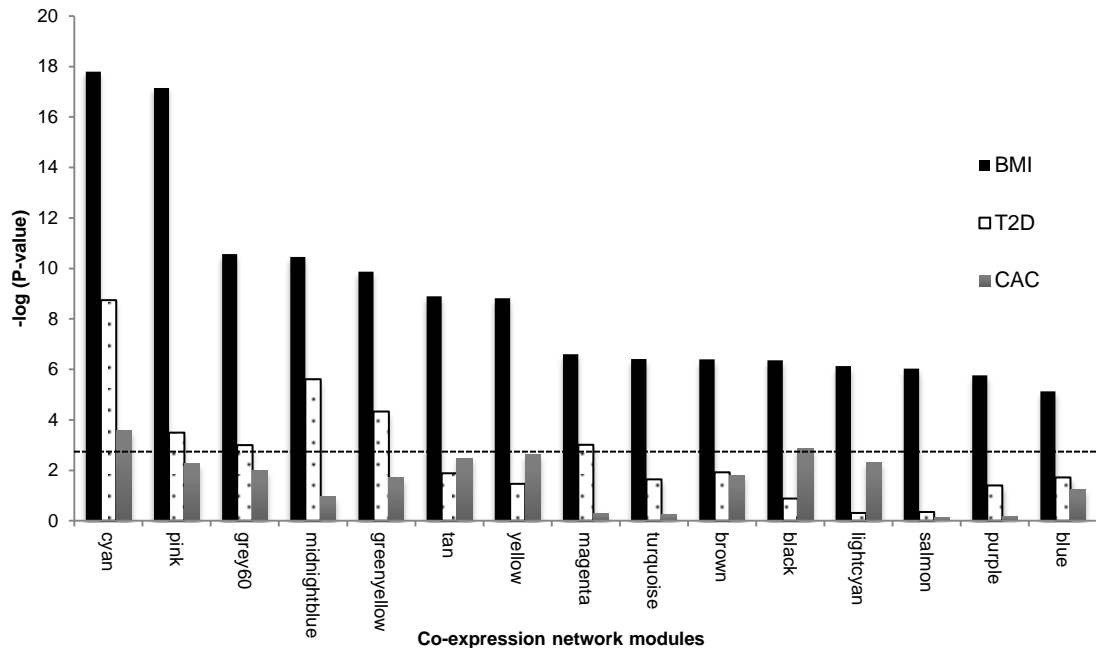
Absolute correlation coefficients between expression of cyan module genes in 1,263 monocyte samples range from 0.3 to 0.7, indicated by line thickness. Positive correlations between genes are indicated by red lines (negative = blue); positive associations between gene expression and BMI are indicated by red nodes (negative = blue); strength of correlation between genes and the cyan module eigengene indicated by node size. The anticipated effects of BMI-associated gene expression on intracellular cholesterol levels and fatty acid synthesis are indicated by brackets/arrows. Cyan module genes include: 3-hydroxy-3-methylglutaryl-CoA synthase 1 (*HMGCS1*), farnesyl-diphosphate farnesyltransferase 1 (*FDFT1*), squalene epoxidase (*SQLE*), methylsterol monooxygenase 1 (*SC4MOL*), cytochrome P450, family 51, subfamily A, polypeptide 1 (*CYP51A1*), stearoyl-CoA desaturase (*SCD*), fatty acid desaturase 1 (*FADS1*), low density lipoprotein receptor (*LDLR*), myosin regulatory light

chain interacting protein (*MYLIP*), ATP-binding cassette, sub-family G (*ABCG1*), and ATP-binding cassette, sub-family A, member 1 (*ABCA1*).

Figure 4 | Odds of T2D and amount of CAC increase as intracellular cholesterol accumulation risk score increases. (a) Odds ratio of T2D (individuals with impaired fasting glucose were excluded) and **(b)** fold change of CAC (with 95% confidence intervals - CI), by risk gene dosage group. Dosage groups defined by the number of ‘risk genes’ (expression levels predicted to promote intra-cellular cholesterol accumulation) from the cyan network: reference group (ref.) = 0 – 2, Low = 3 – 5, Medium = 6 – 8, High = 9 – 11. Model 1 was adjusted for age, gender, race, and study sites; Model 2 was adjusted for all variables in Model 1 plus smoking, HDL cholesterol, triglycerides, systolic blood pressure, and physical activity; and Model 3 was adjusted for all variables in Model 2 plus BMI.

Figure 1





Co-expression network modules

Gene count	11	25	12	14	18	16	42	20	576	61	29	13	16	19	63
Enriched pathways	Cyan: Sterol/ cholesterol Metabolic Process						Magenta: Antigen processing and presentation		Brown: ncRNA metabolic process		Lightcyan: Translation		Black: Ribosome		

Figure 3

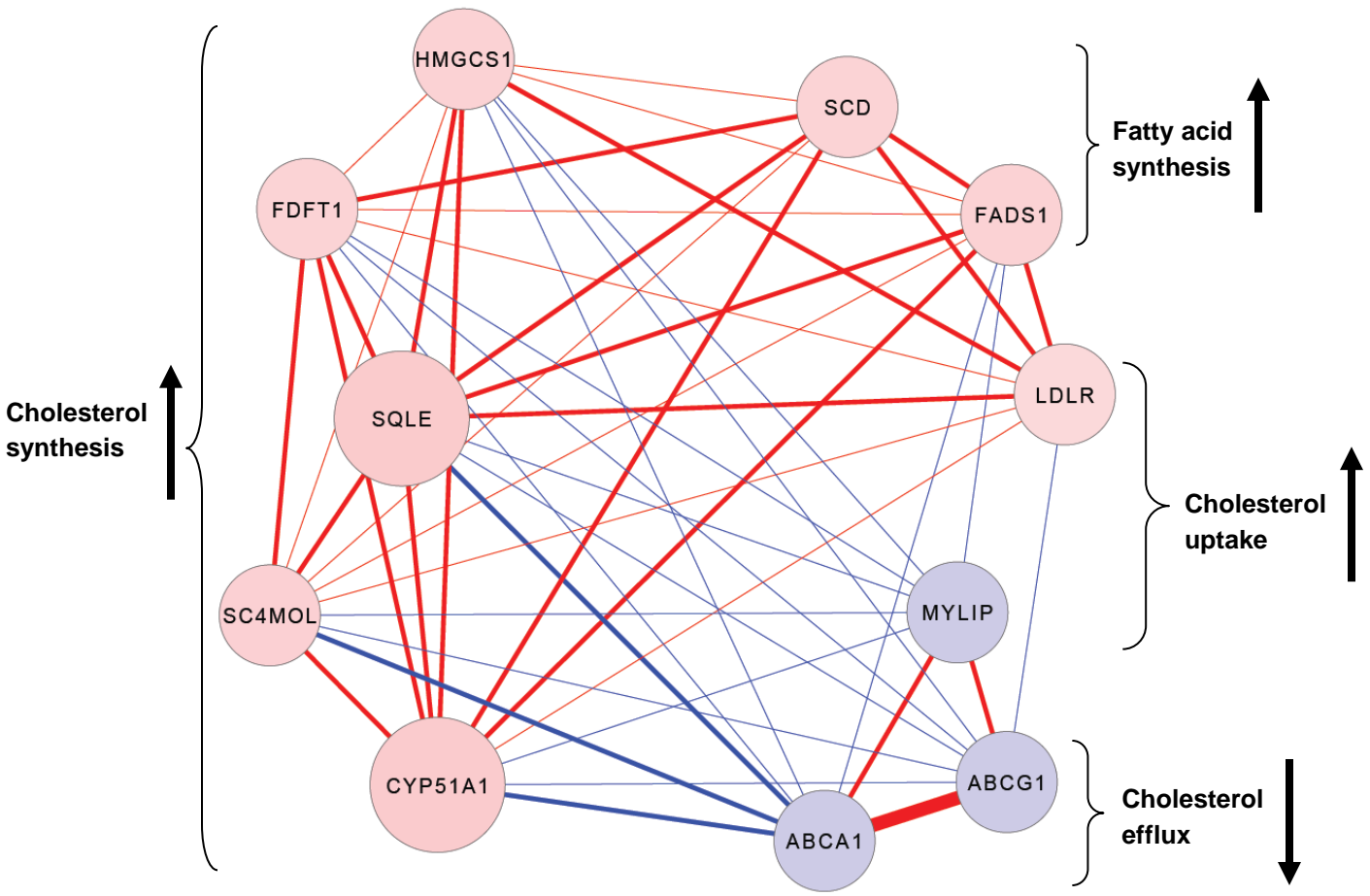


Figure 4

