Genetic regulation of differentially methylated genes in visceral adipose tissue of severely obese men discordant for the metabolic syndrome

Frédéric Guénard[a,b], André Tchernof[b,c], Yves Deshaies[c,d], Simon Biron[e], Odette Lescelleur[e], Laurent Biertho[e], Simon Marceau[e], Louis Pérusse[a,f] and Marie-Claude Vohl[a,b]

[a]Institute of Nutrition and Functional Foods (INAF), Québec, Canada; [b]School of Nutrition, Laval University, Québec, Canada; [c]Québec Heart and Lung Institute, Québec, Canada; [d]Department of Medicine, Laval University, Québec, Canada; [e]Department of Surgery, Laval University, Québec, Canada; [f]Department of Kinesiology, Laval University, Québec, Canada.

Email addresses:
Frédéric Guénard: frederic.guenard@fsaa.ulaval.ca
André Tchernof: andre.tchernof@criucpq.ulaval.ca
Yves Deshaies: Yves.Deshaies@fmed.ulaval.ca
Simon Biron: simon.biron@fmed.ulaval.ca
Odette Lescelleur: Odette.Lescelleur@criucpq.ulaval.ca
Laurent Biertho: laurentbiertho@gmail.com
Simon Marceau: simon.marceau@ssss.gouv.qc.ca
Louis Pérusse: Louis.Perusse@kin.ulaval.ca

Corresponding author:
Marie-Claude Vohl, Ph.D.
Institute of Nutrition and Functional Foods (INAF)
Université Laval
2440 Hochelaga Blvd
Quebec, Quebec
Canada, G1V 0A6
Email: marie-claude.vohl@fsaa.ulaval.ca
Phone number: (418) 656-2131 ext. 4676

Running title: Methylation-associated SNPs in adipose tissue

List of abbreviations:
BMI, body mass index; CpG-SNPs, methylation-associated SNPs; DBP, diastolic blood pressure; GEO, Gene Expression Omnibus; GWAS, genome-wide association studies; HDL-C, HDL-cholesterol; HWE, Hardy-Weinberg equilibrium; LD, linkage disequilibrium; LDL-C, LDL-cholesterol; MAF, minor allele frequency; meQTLs, methylation quantitative trait loci; MetS, metabolic syndrome; NCEP-ATPIII, National Cholesterol Education Program Adult Treatment Panel III; SBP, systolic blood pressure; TF, transcription factor; TG, triglycerides; total-C, total-cholesterol; UTR, untranslated region; VAT, visceral adipose tissue; VEP, Variant Effect Predictor.

Abstract

A genetic influence on methylation levels has been reported and methylation quantitative trait loci (meQTLs) have been identified in various tissues. The contribution of genetic and epigenetic factors in the development of the metabolic syndrome (MetS) has also been noted. In order to pinpoint candidate genes for testing association of SNPs with MetS and its components, we aimed to evaluate the contribution of genetic variations to differentially methylated CpG sites in severely obese men discordant for MetS. Genome-wide differential methylation analysis was conducted in visceral adipose tissue (VAT) of 31 severely obese men discordant for MetS (16 with and 15 without MetS) and identified ~17,800 variable CpG sites. Genome-wide association study conducted to identify SNPs (meQTL) associated with methylation levels at variable CpG sites revealed 2292 significant associations ($P<2.22x10^{-11}$) involving 2182 unique meQTL regulating methylation levels of 174 variable CpG sites. Two meQTLs disrupting CpG sites located within the collagen encoding *COL11A2* gene were tested for associations with MetS and its components in a cohort of 3021 obese individuals. Rare allele of these meQTLs showed association with plasma fasting glucose levels. Further analysis conducted on these meQTL suggested a biological impact mediated through disruption of transcription factor (TF) binding sites based on prediction of TF binding affinities. The current study identified meQTL in VAT of severely obese men and revealed associations of two *COL11A2* meQTL with fasting glucose levels.

Introduction

Epigenetic mechanisms are involved in the acquisition and maintenance of organized tissues[1] and represent a potential link through which genetics and environment may cause phenotypic variations.[2] Specifically, several pieces of evidence point to a contribution of genetic[3-5] and environmental factors in the establishment of DNA methylation levels at individual genomic loci or at the genome-wide level.[6,7] Alterations in DNA methylation are known to affect gene transcription, phenotypes and methylation patterns associated with complex traits.[8,9] Changes in methylation levels at specific CpG sites have been demonstrated over time[6,10] and have been associated with multiple environmental factors.[7,11] Support for a genetic influence on methylation levels has emerged from estimates of methylation heritability in family as well as in twin studies,[12,13] and the identification of methylation quantitative traits loci (meQTL) in multiple samples and tissues.[3,4,14] Current evidence in the literature report a dependence of DNA methylation on local sequence content with associations observed close to the methylation site.[3,4,12,14] A striking impact of SNPs at methylation sites (hereafter termed CpG-SNPs) disrupting methylation potential and resulting in regions exhibiting allele-specific DNA methylation has been reported.[15]

Over the years, genome-wide association studies (GWAS) and candidate gene-based association studies identified several loci associated with complex traits including the metabolic syndrome (MetS),[16] a clustering of metabolic abnormalities defined by abdominal obesity, impaired glucose tolerance, dyslipidemia and hypertension. Despite the fact that heritability of MetS has been established,[17,18] genetic factors explain a small proportion of MetS variability. In line with the involvement of metabolically active tissues in the pathogenesis of complex diseases[19] and with the presence of tissue-specific methylation patterns,[20,21] our group previously established

methylation profiles in visceral adipose tissue (VAT) of men with MetS (MetS+, N=7) vs. men without MetS (MetS-, N=7) and revealed a potential association of specific metabolic pathways with the presence of MetS.[22]

The current study aimed to decipher the contribution of DNA sequence variations on methylation levels of severely obese men with versus without MetS. We extend our previous analysis of differentially methylated genes in VAT of individuals discordant for MetS by conducting an epigenome-wide association study for most variable sites. Focusing on genetic variations influencing methylation status in VAT of men discordant for MetS, we present an approach integrating genotype and methylation data in order to pinpoint candidate genes for testing associations of CpG-SNPs with MetS and its components. Through the analysis of CpG-SNPs identified in biologically relevant and metabolically active VAT, we provide a proof of concept for the identification of MetS-associated loci, thus integrating mechanistic insights to classic genetic association studies.

Patients and Methods

Subjects

Blood samples and VAT were obtained as previously described[23] from a subset of 31 severely obese men (BMI >40 kg/m^2) undergoing biliopancreatic diversion with duodenal switch (BPD-DS)[24] at the Institut universitaire de cardiologie et de pneumologie de Québec - Université Laval (IUCPQ-UL) (Quebec City, Quebec, Canada). This subset of individuals, including 13 individuals previously analyzed for differences in methylation levels according to MetS group,[22] and not taking any medication to treat MetS features, was selected on the basis of the presence (N=16) or absence (N=15) of MetS from a cohort of 3021 Caucasian individuals (941 men, 2080 women). Blood buffy coat has been collected for all individuals of the whole cohort. All tissue specimens were obtained from the biobank of IUCPQ according to institutionally approved management modalities. Waist girth, resting systolic (SBP) and diastolic (DBP) blood pressure, plasma lipids [total-C, LDL-C, HDL-C and TG] as well as fasting glucose concentrations were measured preoperatively using standardized procedures.[25] BMI was calculated as weight in kilograms divided by height in meters squared. Presence of MetS was determined using the National Cholesterol Education Program Adult Treatment Panel III (NCEP-ATPIII) guidelines[16] when an individual fulfilled three or more criteria. The study was approved by the Université Laval and IUCPQ-UL ethics committees and and was performed in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent before their inclusion.

DNA extraction from VAT and genome-wide methylation analysis

Genomic DNA was extracted from VAT using the DNeasy Blood & Tissue kit (QIAGEN, Mississauga, Ontario, Canada) for the subset of 31 men. DNA was first quantified using both

NanoDrop Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and PicoGreen DNA methods, and 1 μg was bisulfite converted. Quantitative genome-wide methylation analysis was conducted using the Infinium HumanMethylation450 BeadChip according to the manufacturer's instructions (Illumina, San Diego, CA). Arrays were processed at the McGill University and Génome Québec Innovation Centre (Montréal, Canada). The Infinium HumanMethylation450 BeadChip array assesses more than 485,000 methylation sites at single-nucleotide resolution and was designed for genome-wide methylation analysis with coverage targeted mostly across gene regions (i.e. promoter region, 5'UTR, first exon, gene body and 3'UTR).

Methylation data were visualized and analyzed using the GenomeStudio® software version 2011.1 (Illumina Inc.) and the Methylation Module. Background subtraction and normalization of methylation data were conducted using internal control probe pairs. Two different probe designs are found on the Infinium® HumanMethylation450 BeadChip array (Infinium I and Infinium II assays). Assay-specific normalization was therefore conducted using internal control probes. All samples fulfilled quality controls (bisulfite conversion, extension, staining, hybridization, target removal, negative and non-polymorphic control probes). Probes with detection *P* value >0.05 in either of the two groups were removed from analysis. Methylation levels (beta values; β) were estimated as the ratio of signal intensity of the methylated alleles to the sum of methylated and unmethylated intensity signals of the alleles and varied from 0 (no methylation) to 1 (100% methylation). Differences in methylation levels (mean β values) were tested between MetS+ and MetS- groups using Student *t*-test in GenomeStudio software. Aiming to identify CpG-SNPs related to the presence of MetS, a relaxed *P* value cutoff of 0.05 was used to identify differentially methylated CpG sites (variable sites) for subsequent SNPs association analyses.

Genotyping

DNA was isolated from the blood buffy coat using the GenElute™ Blood Genomic DNA kit (Sigma, St Louis, MO, USA). Quantification and verification of DNA quality were conducted via both NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and PicoGreen DNA methods. Genome-wide genotyping was performed for 4,301,331 markers using Illumina HumanOmni-5-Quad BeadChip® (Illumina Inc., San Diego, CA) in the subset of 31 individuals selected for genome-wide methylation analysis. Samples were loaded on six different arrays to minimize any batch effect while it can accommodate 12 samples simultaneously. Arrays were processed at McGill University and Genome Quebec Innovation Centre (Montreal, Canada), according to the manufacturer's instructions. Genotyping of selected CpG-SNPs in the whole cohort of 3021 individuals was carried out using pre-designed primers and TaqMan probes (Applied Biosystems) following isolation, quantification and quality assessment as described above. Genotypes were determined using Applied Biosystems® 7500 Fast Real-Time PCR System and analyzed using ABI Prism SDS software version 1.2.3 (Applied Biosystems).

SNPs and sample quality control for genome-wide genotyping

Samples were tested for call rate (>95%), ethnicity (Caucasian; HapMap) and gender mismatch based on genotyping data. No subject was excluded; all 31 samples were used in subsequent analysis. Calculations of allele frequencies and tests of SNP data for Hardy-Weinberg equilibrium (HWE) were performed using PLINK[26] (version 1.07). SNPs with a call rate <90%, genotype distribution deviating from HWE (*P* values less than 10^{-5}), monomorphic or with a minor allele frequency (MAF) < 0.01, located on mitochondrial DNA or unmapped were removed from analyses. A total of 1,818,106 SNPs were excluded leaving 2,483,225 SNPs for statistical analyses.

Gene expression analysis

Gene expression levels in VAT were retrieved from previously published data[27] for 13 of the 31 men used in CpG site methylation analysis. Briefly, gene expression profiling was performed using Affymetrix HG-U133 plus 2 arrays (Affymetrix, Santa Clara, CA) in severely obese men (BMI >40 kg/m^2) not taking any medication to treat MetS features.

Functional analyses

Potential impacts of meQTLs and selected CpG-SNPs were analyzed using VEP. Analysis of the impact of SNPs on TF binding affinities was conducted using TRAP[28] which allowed the prediction of TF binding affinities based on DNA sequences and identified potential TF binding sites affected by SNPs in specific sequences. Sequences overlapping SNPs of interest (30 bp upstream and downstream) were submitted for analysis as input sequences. This software was also used to identify overrepresentation of TF among meQTL-associated differentially methylated CpG sites. The Transfac vertebrates 2010.1 database was used as TF matrix file and human promoter sequences were introduced as background model. Haplotype estimation from selected CpG-SNPs was conducted using PHASE v2.1.[29]

Statistical Analysis

Clinical data were expressed as mean ± SD according to MetS status. Differences in clinical data between MetS+ and MetS- men were tested using Student's *t*-test for continuous variables and with Chi-square test for categorical variables. Associations between SNPs and methylation levels of CpG sites showing differential methylation between MetS+ and MetS- individuals were tested using PLINK. Chromosome-specific associations were tested under an additive model to identify SNPs associated with methylation levels of CpG sites located on the same chromosome, thus

allowing the identification of regulatory SNPs and increasing the probability of functional links between SNPs and CpG sites methylation levels. Associations were first tested using linear regression. Methylation levels of specific CpG sites being reported to be associated with age[6] and smoking status,[11] associations were also tested after adjustment for the effect of age or smoking (current smoker or not). Bonferroni correction was applied to account for multiple testing thus leading to a *P* value cutoff of $P<2.22\text{x}10^{-11}$ ($0.05/\sim2.25\text{x}10^{9}$ tests) to claim statistical significance. Stratification of CpG sites, SNPs and the number of association tested is provided in Supplementary Table S1. Associations between selected CpG-SNPs, MetS and its components were tested under two different models. Associations were first tested under dichotomous model treating MetS and its components as high or low according to NCEP-ATPIII criteria.[16] Second, differences between genotype groups were tested using analysis of variance (general linear model, type III sum of squares) with adjustments for the effects of age, sex and BMI for each individual component. Transformations were applied to non-normally distributed variables in order to meet the criteria for normality [logarithmic transformation (log 10) for fasting glucose levels; negative inverse transformation (1/(-1*(1 + X))) for HDL-C]. LD (r^2) between SNPs demonstrating significant associations as well as between selected CpG-SNPs was calculated from HapMap CEU population and 1000 Genomes Project phase 1v3 data using Haploview,[30] SNAP[31] or LD calculator. Statistical analyses for differences of clinical data between MetS+ and MetS- groups in the subset of 31 individuals and for association testing in the whole cohort were conducted using the SAS software version 9.3 (SAS Institute Inc), whereas PLINK was used to test genome-wide associations.

Results

Subjects' Description

The current study included 31 middle-aged severely obese men (mean body mass index [BMI] 55.7±11.5 kg/m^2) candidates for biliopancreatic diversion with duodenal switch. All subjects met MetS criterion for abdominal obesity and 16 of them had three or more MetS features (Table 1). MetS+ and MetS- individuals showed similar BMI and waist girth while higher total-cholesterol (total-C), triglyceride (TG), fasting glucose levels and blood pressure (P<0.05 for all) were observed in MetS+ individuals. MetS+ group also had lower mean HDL-cholesterol levels (HDL-C; P<0.0001) than MetS- group.

Differential Methylation Analysis

Methylation levels were obtained for all 31 individuals with a mean call rate of 99.96%. Classification of methylation levels for all sites according to fraction of probes with low (<25%), medium (≥25% and ≤75%), or high (>75%) methylation revealed a difference in the distribution of CpG site methylation levels between MetS+ and MetS- groups with a higher number of highly (> 75%) methylated sites in MetS+ individuals (P<0.0001). Overall, high correlations in CpG site methylation levels were observed between groups (r=0.999, P<2.2x10^{-16}). Nonetheless, differential methylation analysis conducted between MetS+ and MetS- individuals revealed 17,801 CpG sites with significant differences (P<0.05; Table 2) in mean methylation levels, among which 17,610 CpG sites showed differences in methylation levels lower than 0.1. High proportions of differentially methylated CpG sites were located in intergenic and in gene body regions (Figure 1) with 12,970 being located in gene regions (promoter, 5'untranslated region [UTR], 1st exon, gene body or 3'UTR regions) and assigned to 6901 unique genes. These 17,801 differentially methylated CpG sites between groups, thereafter called variable CpG sites, were

used as the set of CpG sites of interest for the identification of SNPs (meQTL) associated to CpG site methylation levels.

Association between SNPs and methylation levels for differentially methylated CpG sites

Associations were tested between 2,483,225 SNPs and methylation levels for the set of 17,801 variable CpG sites in a chromosome-specific manner (Supplementary Table S1). A total of 2292 significant associations ($P<2.22x10^{-11}$) were identified involving 2182 SNPs (meQTL) associated with methylation levels of 174 variable CpG sites (Supplementary Data S1). CpG sites under genetic control were overrepresented among variable CpG sites which showed differences in methylation levels greater than 0.1 between MetS+ and MetS- groups ($P<0.0001$). Specifically, 191 of the 17,801 differentially methylated sites showed group differences greater than 0.1. Concomitantly, 114 of the 174 meQTL-associated variable CpG sites showed such group difference. Chromosome-based linkage disequilibrium (LD) analysis conducted in our cohort for the 2182 meQTL revealed moderate pairwise LD (mean $r^2=0.70$) between meQTLs while a similar analysis conducted in the CEU population (Utah Residents with Northern and Western Ancestry) from the 1000 Genomes Project revealed very low LD ($r^2 = 0.02$) between these meQTL. It suggests that LD observed in our population resulted from a limited sample size and genetic variability, and that most of the meQTL identified were independent. Associated CpG sites were mostly located in gene regions (70.1%), including the promoter, the gene body, the 5'UTR, the 1st exon and the 3'UTR regions and were assigned to 105 genes (Figure 2A) among which some had been associated with BMI (*PDZRN4*, *FAM101A*),[32] glucose homeostasis (*DOCK5*, *MGAT5*),[33] HDL-C (*FAM101A*),[34,35] TG (*FAM101A*, *DLGAP2*)[34-36] and diabetes (*SND1*, *RBMS1*).[37,38] Analysis of transcription factors (TF) found in surrounding regions (100 bp) of meQTL-associated CpG sites showed overrepresentation of TF involved in tissue development

(HNF4, NKX32, PAX8, PITX2), lipid metabolism (FXR), initiation of transcription (TBP), cell cycle progression (FOXM1) or known as steroid receptors (GR, AR, PR) (FDR-adjusted *P* value <0.05; Table 3). The list of meQTLs was examined using the Ensembl Variant Effect Predictor (VEP),[39] with each meQTL assigned on the basis of the most severe predicted consequence. Most SNPs involved in associations were intronic (48.8%) or intergenic (28.0%) with 98 overlapping regulatory regions (Figure 2B). Notably, SNPs previously associated with BMI[32] (rs285575) and diabetes[40] (rs3764021) were found among meQTL. Chromosomal distance between meQTL and associated-CpG sites was calculated for all pairs and revealed that the effect of meQTLs on CpG site methylation levels is a function of distance, a higher number of associated CpG sites being found at shorter distances (Supplementary Figure S1A) with a median distance of 90.1 kb between SNP and CpG site. Specifically, 64 of the 2292 associations identified involved meQTLs located within a CpG site, thereafter termed CpG-SNPs, thus introducing or removing a CpG site. Beyond direct effects on methylation levels, these 64 CpG-SNPs contributed to 74 of the 2292 significant associations, thereby implying that some CpG-SNPs are also associated with methylation level of nearby CpG sites and demonstrating a limited contribution of CpG-SNPs to associations identified through patterns of correlation in methylation levels and LD. Globally, a mean reduction of 0.25 in methylation level was associated with each minor allele of SNPs, which suggests an important impact of meQTL on methylation levels (Supplementary Figure S1B) and a strong contribution of CpG-SNPs. A directional effect of meQTLs on change in methylation levels according to CpG site localization was also observed, CpG sites localized in the 1st exon demonstrating an increase in methylation levels with increasing number of meQTL rare alleles (Supplementary Figure S1C).

Since changes in methylation levels were reported with age[6] and smoking,[11] associations were further tested after adjustment for age or smoking and revealed a minor effect of these

confounders on the list of significant associations. Adjustment for the effect of age had small impact on the list of meQTL-associated CpG sites, 168 of the 174 CpG sites initially identified being found among significant associations ($P<2.22x10^{-11}$) following adjustment for age. Similarly, testing association between SNPs and methylation levels for variable CpG sites with adjustment for the effect of smoking led to the identification of significant association for 166 of the 174 CpG sites initially identified. Among the CpG sites no longer significant following adjustment for smoking, four (cg09654329, *RGS12*; cg12556325, *C7orf29*; cg20438820, intergenic; cg25106783, *PRR4*) were common to CpG sites no longer significant following adjustment for age, thus suggesting that methylation levels at these sites are environmentally-driven. However, these results must be interpreted with caution since only 5 individuals were smokers (see Table 1).

Selection of CpG-SNPs for association with cardiometabolic risk factors

In search for CpG-SNPs associated with MetS and its components, we looked for a gene with multiple CpG-SNPs showing associations for further analyses. The collagen type XI alpha 2 (*COL11A2*) gene, known to encode a minor fibrillar collagen, showed two CpG-SNPs among the 39 significant associations observed. Overall, meQTL located within *COL11A2* demonstrated medium LD (mean LD of 0.518) in our study sample. Specifically, the presence of each rare allele of CpG-SNP rs2855430 was associated with a reduction of 0.198 in methylation level of CpG site cg13224161 (chr6:33141279) while CpG site cg16507569 (chr6:33158020) showed a reduction of 0.293 in methylation levels in presence of a minor allele of rs114894582 (Supplementary Figure S2). The presence of further association between CpG-SNPs, MetS and its components was tested in a cohort of 3021 severely obese individuals (Table 4) to demonstrate the usefulness of an approach using CpG-SNPs for the identification of MetS-related

SNPs. To overcome failure in assay design for rs114894582, which could not be provided by manufacturer due to the presence of SNPs in surrounding region, genotyping of rs12526336 was conducted in the cohort of severely obese individuals, the latter showing perfect LD with rs114894582 (r^2=1.0) in our study sample of 31 severely obese men and in the CEU population of the 1000 Genomes Project. A difference in genotype group frequencies was observed between individuals with high vs. low fasting glucose levels (χ^2=8.65, *P*=0.01), specifically a higher frequency of homozygotes for the rare allele was found in the high glucose level group. Similarly, following adjustments for age, sex and BMI, rs12526336 was associated with fasting glucose levels (*P*=0.01). Homozygotes for the rare allele (A) for this SNP demonstrated higher fasting glucose levels (7.78 mmol/l) than common homozygotes (6.61 mmol/l, *P*=0.003) and heterozygotes (6.61 mmol/l, *P*=0.003). Similarly, rs2855430 was associated with fasting glucose levels (*P*=0.04), homozygotes for the rare allele of rs2855430 showed higher fasting glucose levels (7.26 mmol/l) than common homozygotes and heterozygotes (6.63 mmol/l, *P*=0.02 and 6.53 mmol/l, 0.01, respectively).

Functional impact of MetS component-associated CpG-SNPs

To test the functional impact of MetS component-associated CpG-SNPs, we used a combination of *in silico* analyses and tested association of CpG-SNPs with gene expression levels. Identification of differences in TF binding affinity using TRAP revealed a potential disruption of a binding site for leukemia/lymphoma-related factor and the creation of a BCL6 binding site in the presence of minor allele of CpG-SNP rs2855430. The presence of minor allele of CpG-SNP rs114894582 was predicted to disrupt an estrogen binding site while the presence of the surrogate SNP rs12526336 potentially creates MYB and AP3 binding sites, thus revealing a potential

impact of CpG-SNP on gene expression through an impact on the binding of TF. Associations between CpG-SNPs and *COL11A2* gene expression were then tested in 13 severely obese men using expression data retrieved from published array data[27] and were non-significant. Based on patterns of chromatin state maps and histone binding in adipocytes from the NIH Roadmap Epigenomics project for the *COL11A2* gene and surrounding region (1 Mb), the potential impact of MetS component-associated CpG-SNPs was further analyzed for genes upstream *COL11A2* (300 kb) and revealed an association of the rs2855430 CpG-SNP with *PFDN6* and *TAPBP* gene expression following adjustment for age (P=0.04 for both). A trend towards an association with *RGL2* gene expression levels was also identified.

Discussion

This study describes how genetic and epigenetic mechanisms interact in VAT of severely obese men discordant for the MetS. It also provides an analytical scheme to unveil phenotype-SNP associations through analysis of potential underlying molecular mechanisms. Focusing here on CpG-SNPs potentially affecting gene function through disruption of gene methylation, we present a mechanistic approach for the identification of candidate genes for MetS and its components. The prioritization method used here first demonstrated a high correlation of methylation levels between MetS+ and MetS- groups but also identified a set of variable (differentially methylated) CpG sites potentially reflecting MetS specific methylation differences. Although intergenic CpG sites represent a minority of the sites interrogated by the array used here, a high proportion (27.1%) of variable CpG sites identified were located in intergenic regions. Accordingly, environment-dependent variability and tissue-specific differential methylation were previously reported to occur preferentially at gene-poor regions,[41,42] consistent with previous results from our group and others arguing for a contribution of genetically- and environmentally-driven epigenetic events leading to metabolic dysfunction and MetS.[22,43]

Our study distinguishes itself from others that have assessed the genetic component of DNA methylation[3,4,12,14,44,45] by its focus on the genetic contribution to variable CpG sites between individuals with and without the MetS in VAT. We explored the hypothesis that genetic factors underlying differentially methylated CpG sites may reveal MetS-associated SNPs. Through testing of the associations between SNPs and the set of variable CpG sites in a chromosome specific manner, we revealed meQTL potentially modulating methylation levels of 174 CpG sites in VAT. These results reveal a genetic influence on CpG methylation levels through proximal and distal effects, similar to a previous study focusing on subcutaneous adipose tissue samples.

Further adjustment for age or smoking identified small effects of the confounding factors on the number of significant associations in contrast to previous reports.[6,46] The narrow age range and the low frequency of smokers in MetS+ and MetS- groups from the present study sample may have limited our capacity to reveal these confounding effects. Yet, our study may provide potential mechanisms for associations previously reported in GWAS through an impact on CpG site methylation levels and potential modification of TF binding. Accordingly, meQTL identified here were localized in genes previously associated with MetS features (HDL-C: *FAM101A*; glucose homeostasis: *DOCK5*, *MGAT5*; and diabetes: *SND1*, *RBSM1*).[33-35,37,38] Furthermore, specific meQTLs identified were previously associated with BMI (rs2855775) [32] and diabetes (rs3764021).[40] TF analysis conducted from meQTL-associated CpG sites identified overrepresentation of transcription factors or receptors acting as ligand-activated transcription factors involved in hormone response (GR, AR, PR, PAX8)[47,48] as well as in liver development, glucose and lipid metabolism (HNF4, FXR),[49,50] thus arguing for a mechanistic effect of meQTL through impacts at the systemic level. Accordingly, meQTL common to various tissues were previously reported.[3,51]

The proportion of CpG sites (0.98%) reported as being under genetic control among the 17,801 variable sites suggests a strong contribution of non-genetic factors in VAT methylation differences between MetS+ and MetS- groups. Most of these differences in methylation levels may potentially reflect environmentally-driven methylation changes or pathophysiological modifications occurring through the development of MetS. However, a high proportion of meQTL-associated CpG sites among those showing large (>10%) differences in methylation levels between MetS+ and MetS- groups suggest an important genetic control for most variable CpG sites, especially through disruption of the CpG sites by the presence of CpG-SNPs.

Regarding the genetic component of variable CpG sites between VAT of severely obese men discordant for the MetS, comparison of the number of meQTLs identified here with numbers reported in other studies is limited as 17,801 variable CpG sites were assessed here.[44,51,52] The current results are concordant with significant heritability of methylation levels previously reported[12,52,53] through an impact of genetic variations on methylation levels, and with environmental influences on methylation levels reported in twin comparisons.[41] The observed median distance of 90.1 kb between meQTLs and associated CpG sites was higher than the median distance of ~80 kb[44] or ~30 kb[45] previously reported in adipose tissue. These inconsistencies might reflect the larger range of the associations tested. Our study conducted a chromosome-restricted analysis on the basis that SNPs impact may act on larger genomic regions, whereas other groups restricted their analyses on smaller 500 kb regions.[44,45] A larger median distance between meQTLs and associated-CpGs may also be an artifact of the relative genetic homogeneity of our population, higher LD being identified between meQTL in our sample than in the CEU population from the 1000 Genomes Project. Based on such higher LD observed, meQTLs identified may alternatively reflect causal SNPs which have not been interrogated by the array.

Our study revealed a limited number of meQTLs being classified as CpG-SNPs adding to the evidence supporting an impact of meQTLs on methylation levels without disruption of CpG sites,[3,12,44] in contrast with results from the GOLDN study[54] demonstrating a major contribution of CpG-SNPs in meQTL signals. Although limited in number, the strong effect on CpG site methylation levels and their potential physiological impacts drew our attention on two CpG-SNPs located within the *COL11A2* gene. As a proof of principle for the identification of disease-

associated SNPs, associations of *COL11A2* CpG-SNPs with MetS and its components were tested. Although associations between *COL11A2* CpG-SNPs and fasting glucose levels were identified here, associations were not observed for *COL11A2* gene expression levels. Similarly, partial overlap between meQTLs and expression levels of other closely localized genes in adipose tissue has been reported.[6,12] Alternatively, impact of meQTL may be exerted in specific cell type and masked by cellular heterogeneity. Otherwise, it may be exacerbated in specific tissues, tissue depots or conditions. Consequently, up-regulation of extracellular matrix constituents including members of the collagen family was reported in subcutaneous white adipose tissue.[55] Testing CpG-SNPs for association with expression levels of surrounding genes based on patterns of chromatin state maps and histone binding suggested a distal impact of these CpG-SNPs in VAT potentially through regulation of surrounding genes. Such distal effect was recently reported for a SNP located near the *FTO* gene.[56] Alternatively, CpG-SNPs analyzed here may exert an action in other metabolically active tissues, *COL11A2* gene being highly expressed in tissues other than adipose, including muscle (EMBL-EBI Expression Atlas). The fasting glucose levels-associated CpG-SNPs identified here may thus exert their impact in muscle tissue, an important player of whole-body glucose homeostasis.[57]

The present study on the genetic components underlying methylation levels of differentially methylated CpG sites in obese men discordant for MetS provided an analytical scheme to identify MetS-associated SNP based on potential molecular impacts of CpG-SNPs. Providing a proof of concept, the current study design and associations reported merit further replication in independent studies. Relying on potential disruption of molecular mechanisms for the selection and testing of SNPs for association with MetS and its components may nonetheless fill the gap of knowledge between associations reported and disease physiopathology.

Acknowledgments

The authors thank participants in the study for their excellent collaboration and the members of the department of bariatric surgery (Frédéric-Simon Hould, Stéfane Lebel and Picard Marceau) for their valuable contribution in clinical care and patient recruitment. We acknowledge the contribution of the biobank staff of the IUCPQ-UL for sample management and the work of the McGill University and Génome Quebec Innovation Centre for genotyping and gene methylation array processing. M.C.V. is a Canada Research Chair in Genomics Applied to Nutrition and Health.

Conflicts of Interest: The authors have read the journal's policy on disclosure of potential conflicts of interest and declare that there is no conflict of interest associated with this manuscript.

Funding: This study was supported by a grant-in-aid from the Heart and Stroke Foundation of Canada (G-14-0005824). This funding source had no involvement in study design, collection, analysis and interpretation of data, nor in the writing and the submission of this article.

Reference List

1. Cantone I, Fisher AG. Epigenetic programming and reprogramming during development. Nat.Struct.Mol.Biol. 2013;20:282-9.

2. Feinberg AP. Epigenetics at the epicenter of modern medicine. JAMA 2008;299:1345-50.

3. Bell JT, Pai AA, Pickrell JK, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 2011;12:R10.

4. Gibbs JR, van der Brug MP, Hernandez DG, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS.Genet. 2010;6:e1000952.

5. Schalkwyk LC, Meaburn EL, Smith R, et al. Allelic skewing of DNA methylation is widespread across the genome. Am.J.Hum.Genet. 2010;86:196-212.

6. Bell JT, Tsai PC, Yang TP, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoS.Genet. 2012;8:e1002629.

7. Christensen BC, Houseman EA, Marsit CJ, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS.Genet. 2009;5:e1000602.

8. Fackler MJ, Umbricht CB, Williams D, et al. Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. Cancer Res. 2011;71:6195-207.

9. Heyn H. A symbiotic liaison between the genetic and epigenetic code. Front Genet. 2014;5:113.

10. Talens RP, Boomsma DI, Tobi EW, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J. 2010;24:3135-44.

11. Breitling LP, Yang R, Korn B, et al. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. Am.J.Hum.Genet. 2011;88:450-7.

12. Grundberg E, Meduri E, Sandling JK, et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. Am.J.Hum.Genet. 2013;93:876-90.

13. McRae AF, Powell JE, Henders AK, et al. Contribution of genetic variation to transgenerational inheritance of DNA methylation. Genome Biol. 2014;15:R73.

14. Zhang D, Cheng L, Badner JA, et al. Genetic control of individual differences in gene-specific methylation in human brain. Am.J.Hum.Genet. 2010;86:411-9.

15. Shoemaker R, Deng J, Wang W, et al. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. Genome Res. 2010;20:883-9.

16. Expert Panel on Detection EaToHBCiA. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA 2001;285:2486-97.

17. Henneman P, Aulchenko YS, Frants RR, et al. Prevalence and heritability of the metabolic syndrome and its individual components in a Dutch isolate: the Erasmus Rucphen Family study. J.Med.Genet. 2008;45:572-7.

18. Teran-Garcia M, Bouchard C. Genetics of the metabolic syndrome. Appl.Physiol Nutr.Metab 2007;32:89-114.

19. Tchernof A, Despres JP. Pathophysiology of human visceral obesity: an update. Physiol Rev. 2013;93:359-404.

20. Guenard F, Deshaies Y, Hould F, et al. Use of blood as a surrogate model for the assessment of visceral adipose tissue methylation profiles associated with the metabolic syndrome in men. J Mol Genet Med 2016;10.

21. Ribel-Madsen R, Fraga MF, Jacobsen S, et al. Genome-wide analysis of DNA methylation differences in muscle and fat from monozygotic twins discordant for type 2 diabetes. PLoS.One. 2012;7:e51302.

22. Guenard F, Tchernof A, Deshaies Y, et al. Differential methylation in visceral adipose tissue of obese men discordant for metabolic disturbances. Physiol Genomics 2014;46:216-22.

23. Vohl MC, Sladek R, Robitaille J, et al. A survey of genes differentially expressed in subcutaneous and visceral adipose tissue in men. Obes.Res. 2004;12:1217-22.

24. Marceau P, Hould FS, Simard S, et al. Biliopancreatic diversion with duodenal switch. World J.Surg. 1998;22:947-54.

25. Robitaille J, Despres JP, Perusse L, et al. The PPAR-gamma P12A polymorphism modulates the relationship between dietary fat intake and components of the metabolic syndrome: results from the Quebec Family Study. Clin.Genet. 2003;63:109-16.

26. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am.J.Hum.Genet. 2007;81:559-75.

27. Bouchard L, Tchernof A, Deshaies Y, et al. ZFP36: a promising candidate gene for obesity-related metabolic complications identified by converging genomics. Obes.Surg. 2007;17:372-82.

28. Thomas-Chollier M, Hufton A, Heinig M, et al. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. Nat.Protoc. 2011;6:1860-9.

29. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am.J.Hum.Genet. 2006;78:629-44.

30. Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21:263-5.

31. Johnson AD, Handsaker RE, Pulit SL, et al. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008;24:2938-9.

32. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197-206.

33. Palmer ND, Goodarzi MO, Langefeld CD, et al. Genetic Variants Associated With Quantitative Glucose Homeostasis Traits Translate to Type 2 Diabetes in Mexican Americans: The GUARDIAN (Genetics Underlying Diabetes in Hispanics) Consortium. Diabetes 2015;64:1853-66.

34. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707-13.

35. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. Nat.Genet. 2013;45:1274-83.

36. Ko A, Cantor RM, Weissglas-Volkov D, et al. Amerindian-specific regions under positive selection harbour new lipid variants in Latinos. Nat.Commun. 2014;5:3983.

37. Ma RC, Hu C, Tam CH, et al. Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. Diabetologia 2013;56:1291-305.

38. Qi L, Cornelis MC, Kraft P, et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. Hum.Mol.Genet. 2010;19:2706-15.

39. McLaren W, Pritchard B, Rios D, et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. 2010;26:2069-70.

40. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.

41. Fraga MF, Esteller M. Epigenetics and aging: the targets and the marks. Trends Genet. 2007;23:413-8.

42. Slieker RC, Bos SD, Goeman JJ, et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. Epigenetics.Chromatin. 2013;6:26.

43. Smith CJ, Ryckman KK. Epigenetic and developmental influences on the risk of obesity, diabetes, and metabolic syndrome. Diabetes Metab Syndr.Obes. 2015;8:295-302.

44. Drong AW, Nicholson G, Hedman AK, et al. The presence of methylation quantitative trait loci indicates a direct genetic influence on the level of DNA methylation in adipose tissue. PLoS.One. 2013;8:e55923.

45. Volkov P, Olsson AH, Gillberg L, et al. A Genome-Wide mQTL Analysis in Human Adipose Tissue Identifies Genetic Variants Associated with DNA Methylation, Gene Expression and Metabolic Traits. PLoS.One. 2016;11:e0157776.

46. Rakyan VK, Down TA, Maslau S, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res. 2010;20:434-9.

47. Marotta P, Amendola E, Scarfo M, et al. The paired box transcription factor Pax8 is essential for function and survival of adult thyroid cells. Mol.Cell Endocrinol. 2014;396:26-36.

48. Tchernof A, Mansour MF, Pelletier M, et al. Updated survey of the steroid-converting enzymes in human adipose tissues. J.Steroid Biochem.Mol.Biol. 2015;147:56-69.

49. Battle MA, Konopka G, Parviz F, et al. Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. Proc.Natl.Acad.Sci.U.S.A 2006;103:8419-24.

50. Wagner M, Zollner G, Trauner M. Nuclear receptors in liver disease. Hepatology 2011;53:1023-34.

51. Shi J, Marconett CN, Duan J, et al. Characterizing the genetic basis of methylome diversity in histologically normal human lung tissue. Nat.Commun. 2014;5:3365.

52. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. Genome Biol. 2016;17:61.

53. Tremblay BL, Guenard F, Lamarche B, et al. Familial Resemblances in Blood Leukocyte DNA Methylation Levels. Epigenetics. 2016;0.

54. Zhi D, Aslibekyan S, Irvin MR, et al. SNPs located at CpG sites modulate genome-epigenome interaction. Epigenetics. 2013;8:802-6.

55. Henegar C, Tordjman J, Achard V, et al. Adipose tissue transcriptomic signature highlights the pathological relevance of extracellular matrix in human obesity. Genome Biol. 2008;9:R14.

56. Claussnitzer M, Dankel SN, Kim KH, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. N.Engl.J.Med. 2015;373:895-907.

57. DeFronzo RA, Tripathy D. Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. Diabetes Care 2009;32 Suppl 2:S157-S163.

Figure Legends

Figure 1. Localization of differentially methylated CpG sites between MetS+ and MetS- severely obese men (N=17,801). Probes were localized according to Infinium HumanMethylation450 BeadChip annotation. Abbreviations: MetS+, with the metabolic syndrome; MetS-, without the metabolic syndrome.

Figure 2. Localization of CpG sites and SNPs involved in associations identified. a. Localization of SNP-associated CpG sites (N=174). Probes were localized according to Infinium HumanMethylation450 BeadChip annotation. **b.** Localization of methylation-associated SNPs according to Illumina HumanOmni-5-Quad BeadChip annotation (N=2182).

Supplementary Figure Legends

Supplementary Figure S1. Descriptive statistics of associations identified. A. Distance between meQTLs and associated CpG sites for significant associations identified (N=2292). Details for meQTLs and associated CpG sites located less than 100 kb apart are shown in upper right part (N=1169). **B.** Distribution of changes in methylation levels per copy of the rare allele for meQTL-associated CpG sites (N=2292). **C.** Change in methylation levels per meQTL rare allele according to CpG site localization (N=2292). CpG site localization according to Infinium HumanMethylation450 BeadChip annotation.

Supplementary Figure S2. Methylation levels of disrupted CpG sites according to CpG-SNPs genotype groups in methylation cohort (N=31). A. Methylation levels of cg13224161 (chr6:33141279) according to rs2855430 and **B.** cg16507569 (chr6:33158020) according to rs114894582. Positions relative Genome build 37.