Genome-wide DNA methylation study identifies genes associated with the cardiovascular biomarker GDF-15

Weronica E. Ek[1,*], **Åsa K. Hedman**[2], **Stefan Enroth**[1], **Andrew P. Morris**[3,4], **Cecilia M. Lindgren**[4], **Anubha Mahajan**[4], **Stefan Gustafsson**[2], **Ulf Gyllensten**[1], **Lars Lind**[5], and **Åsa Johansson**[1]

[1]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala, 75108, Uppsala University, Uppsala, Sweden

[2]Department of Medical Sciences, Molecular epidemiology and Science for Life Laboratory, 75108, Uppsala University, Uppsala, Sweden

[3]Department of Biostatistics, L69 3BX, University of Liverpool, Liverpool, UK

[4]Wellcome Trust Centre for Human Genetics, OX3 7BN, University of Oxford, Oxford, UK

[5]Department of Medical Sciences, Cardiovascular Epidemiology, 75185, Uppsala University Hospital, Uppsala, Sweden

Abstract

Growth differentiation factor 15 (GDF-15) is expressed in low to moderate levels in most healthy tissues and increases in response to inflammation. GDF-15 is associated with cardiovascular dysfunction and over-expressed in the myocardium of patients with myocardial infarction (MI). However, little is known about the function of GDF-15 in cardiovascular disease, and the underlying regulatory network of GDF-15 is not known. To investigate a possible association between GDF-15 levels and DNA methylation, we performed a genome-wide DNA methylation study of white blood cells in a population-based study (N=717). Significant loci where replicated in an independent cohort (N=963). We also performed a gene ontology (GO) enrichment analysis. We identified and replicated 16 CpG-sites (False discovery rate [FDR] < 0.05), at 11 independent loci including *MIR21. MIR21* encodes a microRNA (miR-21) that has previously been shown to be associated with the development of heart disease. Interestingly, *GDF15* mRNA contains a binding site for miR-21. Four sites were also differentially methylated in blood from participants previously diagnosed with MI and 14 enriched GO terms (FDR < 0.05, enrichment > 2) were identified, including "*cardiac muscle cell differentiation*".

This study show that GDF-15 levels are associated with differences in DNA methylation in blood cells, and a subset of the loci are also differentially methylated in participants with MI. However, there might be interactions between GDF-15 levels and methylation in other tissues not addressed in this study. These results provide novel links between GDF-15 and cardiovascular disease.

*Corresponding author: Weronica E. Ek, Address: box 815, 75108, Uppsala, Sweden, Telephone: +46703519004, Weronica.ek@igp.uu.se, Fax: +46184714931.

Conflict of interest statement
Authors declare no conflict of interest.

Introduction

Growth-differential factor-15 (GDF-15) is a member of the transforming growth factor-β (TGF-β) family (1). Under physiologic conditions, GDF-15 is expressed in low to moderate levels in most healthy tissues. GDF-15 levels may increase in response to pathological stress associated with inflammation or tissue damage and an overexpression has been found in numerous malignancies including breast, melanoma, colorectal, pancreatic and prostate cancer (2–4). Growing evidence indicate that the GDF-15 level in plasma may be a new biomarker for risk stratification and therapeutic decision making in cardiovascular disease (5, 6). GDF-15 is expressed in the myocardium of patients with acute myocardial infarction (MI)(7) and has been shown to be increased in blood from patients with acute MI (8). Elevated GDF-15 levels has also been associated with risk of mortality in patients with acute coronary syndrome (ACS) or chronic heart failure (9). GDF-15 levels are also higher in healthy individuals older than 65 years than in younger individuals (10).

Epigenetics is generally used to denote the regulation of genes that cannot be attributed to the DNA sequence. It includes a number of different mechanisms of which DNA methylation and histones modifications being the most commonly studied. DNA methylation is the heritable and reversible attachment of a methyl group to a nucleotide. In mammals, DNA methylation appears to be specific to cytosine, predominantly to CpG (cytosine-phosphate-guanine) dinucleotide. In promoter regions, CpG-sites are often found as clusters called CpG islands. Methylation of promoter CpGs can introduce stable changes in gene expression, which might lead to silencing of that gene (11, 12).

In this study, we performed a genome-wide DNA methylation study to determine the association between GDF-15 levels and DNA methylation. DNA methylation status in blood cells has been determined at more than 475,000 sites distributed throughout the genome in two independent population based study cohorts.

Results

Genome-wide DNA methylation study

DNA methylation levels of 485,512 CpG sites were measured in blood in 743 participants from NSPHS, using the Illumina 450K chip. A total of 472,477 sites (99.89%) passed the QC. Of these, 470,789 CpG -sites were autosomal and have been included in downstream analyses. 11,232 CpG-sites were located on the X chromosome and analysed in males and females separately and 416 CpG-site on the Y chromosome was analysed only in males. None of the CpG-sites located on the sex-chromosomes passed the False discovery rate (FDR) q-values < 0.05 threshold and was not further investigated. All individuals passed QC for the NSPHS population. After removing technical controls, 717 samples with DNA methylation and GDF-15 data available remained for the epigenetic analyses. In NSPHS, GDF-15 levels increased with age (p=6.98 x 10^{-41}) and with fraction of granulocytes (p=1.60 x 10^{-7}), but there was no significant difference between sexes, nor with variation in fractions of other cell types (p>0.05).

A total of 31 DNA autosomal methylation sites were significantly associated (False discovery rate (FDR) *q*-values < 0.05) with GDF-15 levels in NSPHS (Figure 1, Table 2, Supplementary Table 2). The 31 significant sites mapped to 23 genes. Of these, 16 CpG-sites, corresponding to 11 genes, were replicated in PIVUS (FDR q < 0.05) with the same direction of effect (Table 2, Supplementary Table 2). Some of the probes used for measuring DNA methylation overlaps with known SNPs that might influence the probe specificity. Including SNPs located within the probe positions (See supplementary methods for more information) did not influence the association between the CpG methylation and the GDF-15 levels for any of the replicating sites (Supplementary Table 1).

Among the replicated sites, four mapped to vacuole membrane protein 1 (*VMP1*), previously known as transmembrane protein 49 (*TMEM49*). *VMP1* has previously been associated with lipoprotein-associated phospholipase A_2 (Lp-PLA$_2$), which is a proinflammatory enzyme bound to low-density lipoprotein cholesterol and other circulating lipoproteins. Both the mass and activity of Lp-PLA$_2$ are associated with increased cardiovascular risk (13). A total of 18 of the CpG sites on the Illumina 450K methylation chip are annotated to the *VMP1* gene. Of these four were associated with GDF-15 levels. However, as many as 42 CpG sites are located in close proximity to *VMP1*. We can clearly see that among these, the four significant *VMP1* CpG sites are all located in one region towards the end of *VMP1* (Figure 2), or more precisely within or around the 11th exon (Figure 3). This region has previously been shown to be the promoter of a microRNA (miRNA) gene, *MIR21*, which is located about 3kb downstream of this promoter (14). The MirWalk database (15) identifies that miR-21 has a binding site with a match that consists of 9 bases at position 78 to 86 in the *GDF15* mRNA coding sequence (P=0.0035). The probability distribution of random matches of a miRNA 5´end sequence in the given sequence (promoter or mRNA sequence) was calculated by using a Poisson distribution. A low probability implies a significant hit (15). Interestingly, MiR-21 has been shown to play an important role in development of heart disease (16).

Meta-analysis

Meta-analysis was performed for all CpG-sites that passed QC in PIVUS and NSPHS (N = 457,898) and we identified a total of 66 GDF-15 genome-wide significant CpG-sites (p < 5×10^{-8}) (Table 2 and Supplementary Table 5).

Myocardial infarction - NSPHS

All 16 replicating CpG-sites were further investigated in relation to MI. Four sites (cg17150809, cg13033858, cg16936953 and cg18608055) showed a nominally significant decrease in DNA methylation in blood (P-values; 0.0039, 0.05, 0.05 and 0.05, respectively) in the 47 participants that have had a prior MI compared to the 271 participants (above age 50) that has not experienced a MI (Figure 4 a-d). However, none of these sites were still significant after adjusting for multiple testing.

Association between GDF15 SNPs and GDF-15 levels

A previous study identified nine SNPs, corresponding to three independent SNPs (rs3746181, rs1227731 and rs749451), on chromosome 19 (close to *GDF15*) to be

associated to GDF-15 levels (17). Of these, seven SNPs were also associated with GDF-15 levels in NSPHS (Table 3). Since these SNPs are influencing GDF-15, they might alter the precision of the GDF-15 measurement used in our analyses. Including the three independent SNPs (rs749451, rs888663 and rs1054564) as covariates did not appear to affect the association between DNA methylation at the CpG-sites and GDF-15 (Supplementary Table 3). We also performed a GWAS for GDF-15 levels in NSPHS. Three significant SNPs; rs76083784 (p = 4.80×10^{-8}), rs199517669 (p = 4.29×10^{-8}) and rs139654494 (p = 3.23×10^{-8}) were identified of which none replicated in PIVUS (Supplementary Table 4, Supplementary Figure 1-3b). Further information is found in Supplementary methods.

Causal inference

We also performed a MR test to investigate if genetically increased GDF-15 levels directly influence methylation level at our 16 replicated CpG sites, using the three previously published *GDF15* independent SNPs (rs749451, rs888663 and rs1054564). For two of the CpG -sites, the methylation level were associated (nominal P   0.05) with one or more of the *GDF-15* SNPs (Supplementary Table 3) indicating that GDF-15 might directly regulates methylation level at these sites. However, none of the tests were significant after multiple testing adjustments.

Enrichment analysis

As target list for the enrichment analyses we selected using CpG sites for which DNA methylation in blood were associated with GDF-15 levels in the meta-analysis (FDR *q*-value   0.05). This resulted in a set of 796 genes, with at leas one significant CpG site, that where entered into the enrichment analysis. All protein-coding genes in the genome that was represented by at least one CpG site on the Illumina chip (n = 16,662), were used as a background list. A total of 34 enriched biological processes were identified (FDR *q*-value < 0.05), of which 14 had an enrichment higher than 2.0. The most enriched biological processes were "*behavioral defense response*", "*behavioral fear response*" and "*cardiac muscle cell differentiation*" with fold enrichment of 6.11 (FDR *q*-value = 4.55×10^{-2}), 6.11 (FDR *q*-value = $4.39 \times 10^{-2)}$) and 5.86 (FDR *q*-value = 4.81×10^{-2}), respectively (Table 4).

Discussion

We have shown that GDF-15 levels in plasma are associated with differential DNA methylation patterns in blood cells. In total, we identified and replicated 16 CpG-sites at 11 independent genes to be associated with GDF-15 levels. By meta-analysing the results from the two cohorts, we identified as many as 66 genome-wide significant CpG-sites, which shows that the effect is similar between the two cohorts.

Among the 16 replicated associations, four CpG-sites are located within *VMP1*. This gene is an interesting candidate gene since it encodes a stress-induced protein that when overexpressed promotes formation of intracellular vacuoles followed by cell death (18). Mutations in this gene are also associated with acute pancreatitis (19). All four CpG-sites are localized within or close to the 11th exon of the gene and this region has been annotated as an active promoter in ENCODE cell lines. The fact that this region overlaps with the

promoter of miR-21(20–22) (Figure 3) and a fifth associated CpG-site is located just upstream of *MIR21* suggests that the associated CpG-sites are more likely linked to the regulation of *MIR21* rather than *VMP1*. However, it has previously been shown that *MIR21* can also be expressed together with *VMP1* through alternative polyadenylation and that *MIR21* thereby can be regulated by two promoters (23).

MiRNAs are a family of short non-coding RNAs that regulate gene expression post-transcriptionally. MiRNAs generally bind to the 3'-UTR (untranslated region) of their target mRNAs and repress protein production by destabilizing the mRNA and translational silencing (24). Results from the MirWalk database suggests that miR-21 binds to the *GDF-15* mRNA. Similarly to GDF-15, miR-21 has previously been associated with heart disease. MiR-21 expression has been shown to be increased in failing murine and human hearts (16). The same study also showed that inhibition of miRNAs in mice, using chemically modified and conjugated miRNA inhibitors (antagonists), inhibits intestinal fibrosis and improve cardiac function in a pressure overload cardiac disease mouse model (16). Taken together, our results and the similarity in association with cardiovascular diseases between miR-21 and GDF-15, suggests that miR-21 is a probable regulator of *GDF15*. However, by using SNPs that regulates the methylation level of the *MIR21* promoter as instrumental variables for miR-21 expression, we did not confirm that miR-21 levels directly influence GDF-15 levels in blood cells. DNA methylation of promoters is generally considered being associated with decreased gene expression. In our study we found that GDF-15 and DNA methylation level of the *MIR21* promoter was inversely correlated. This means that increased GDF-15 level were associated with decreased *MIR21* methylation, probably reflecting increased *MIR21* expression. This makes sense since increased GDF-15 and miR-21 levels have been associated with the same diseases. However, if miR-21 binds to the mRNA of *GDF15* we would expect an increased *MIR21* expression to result in decreased GDF-15 levels, which is the opposite of our results. It is possible that miR-21 is expressed in order to down-regulates GDF-15 and that miR-21 levels therefore reflect the GDF-15 levels. However DNA methylation was measured in blood cells, and the majority of the circulating GDF-15 might be expressed from a different cell-type, so it is possible that miR-21 regulation of GDF-15 is tissue specific and not even detectable by measuring DNA methylation for blood cells.

Another interesting gene found among our replicating CpG -sites is the *Zinc finger protein 2 (ZEB2). ZEB2* has been shown to be involved in early development of tissues in heart and digestive tract. Deletion of *ZEB2* leads to Mowat-Wilson syndrome, a rare genetic disorder characterized by a number of health defects including Hirschsprung disease, mental retardation, epilepsy and congenital heart disease (25).

GDF-15 has been associated with cardiovascular dysfunction and is known to be over-expressed in the myocardium of patients with MI. It is therefore interesting that four CpG-sites associated to GDF-15 levels showed a nominally significant decrease ($P < 0.05$) in methylation in people previously diagnosed with MI (Figure 4 a-d). One of these CpG-sites, cg16936953, is located within the promoter region of *MIR21* within the gene *VMP1*, previously shown to be involved in cardiovascular disease (13, 16). The other three CpG-sites (cg13033858, cg17150809 and cg18608055) where located within or close to genes, to

our knowledge, not previously shown to be associated to cardiovascular disease (*SSH1*, *FBXL4* and *SBNO2*). It is also interesting that genes involved in "*cardiac muscle cell differentiation*" and "*cardiocyte differentiation*" was highly enriched for association between DNA methylation in blood and GDF-15 levels. Both these GO term include almost the same genes including, Neuregulin *(NRG1)*, Calreticulin *(CALR)*, T-box 3 *(TBX3)*, Retionid X receptor *(RXRB)*, Transforming growth factor beta 2 *(TGFB2)*, Retionid X receptor alpha *(RXRA)*, Retinoic Acid receptor, alpha *(RARA)*, Chibby homolog 1 *(CBY1)* and Serum response factor *(SRF)*, with nominal p-values down to 6.48×10^{-7} (Supplementary Table 6 and 7). Most of these genes have previously been shown to be involved in cardiovascular disease. Briefly, *NRG1* is involved in cardiovascular development and disease (26), *CALR* is associated to congenital heart block disease (27), mutations in *TGFB2* has been shown to predispose to thoracic aortic disease (28) and inactivation of *SRF* in the developing heart can result in myocardial defects and embryonic lethality (29).

The aim of this study was to investigate if GDF-15 levels are associated with differential DNA methylation levels in blood cells. However, we also used a MR approach to evaluate if GDF-15 is directly influencing DNA methylation by using SNPs, previously shown to regulate GDF-15 expression, as instrumental variables. At two of the CpG -sites, the DNA methylation levels were associated (P   0.05) with one or more of the *GDF15* SNPs (Supplementary Table 3) and we can therefore not reject the hypothesis that GDF-15 directly regulates methylation level at these sites.

Limitations

To perform the epigenome-wide methylation analysis, we used whole blood. Whole blood has an inter-individual variability in the fraction of different DNA-containing blood cells. Previous studies have shown that differences in methylation can be the result of variability in cell composition when blood is used in DNA methylation studies (30). We therefore corrected for variability in cell composition between individuals. However, we do not know if the association between DNA methylation and GDF-15 levels is specific to only one cell type, or present in all of the individual cell types in blood. There might also be interactions between GDF-15 levels and methylation in other tissues that has not been addressed in this study. Another limitation is the difference between the populations in demographics, age and sex distribution. The NSPHS study includes a mixture of males and females and an age range between 14-94, while the PIVUS study used for replication consists of males and females aged 70 years. Even though the replication rate was high in our study, the overlap would probably be even higher in two more similar samples. Despite this, half of the sites replicated between these studies, lending strong support to our findings.

The MR approach indicated that GDF-15 might regulate methylation level at two sites. However, the p-values are only nominally significant. The lack of association to other CpG-sites might reflect that GDF-15 is not influencing the DNA methylation at these sites. However, since the effect of the three SNPs on GDF-15 is limited, the strength of our instrument might be somewhat limited and we cannot exclude that these results are affected by low statistical power. We did not replicate two of the SNPs from the Framingham study, which might be due to an association to non-synonymous SNPs in Framingham. Ten of our

significant GO terms include at least two genes that belong to the same gene cluster (*HOXD9*, *HOXD10* and *HOXD11*) (Supplementary Table 7). Genes (CpG-sites) that belong to the same gene cluster might pick up the same effect due to a high correlation between them, which might drive the enrichment by counting the same effect several times. We calculated the correlation between the CpG-sites included for these genes (cg22674699, cg18115040 and cg04949206) giving an $R^2 <= 0.16$ in NSPHS and an $R^2 <= 0.25$ in PIVUS, between each part of sites. These results suggest that the *HOXD*-genes might not represent independent signals.

GDF-15 levels were estimated with two different methods for the discovery and replication set (Olink Proseek Multiplex Oncology I and the immunoradiometric assay) and the concentration / scales between these methods are not directly comparable. This explains why the scales of beta estimates in Table 2 differ between the two cohorts. However, using two different methods is good because is shows that the results are not method dependent. Since GDF-15 levels were measured in different units, the estimates from PIVUS and NSPHS were standardized and z-scores included in the meta-analyses.

In summary, this study shows that level of GDF-15, which has previously been shown to be strongly correlated with cardiovascular diseases, is associated with DNA methylation levels in blood at multiple CpG-sites in the genome. GDF-15 is known to be over-expressed in the myocardium of patients with MI. It is therefore interesting that four CpG-sites associated to GDF-15 levels showed a significant decrease in methylation in people previously diagnosed with MI. Interestingly, this study also suggests that genes involved in heart tissue development are enriched for DNA methylation associations with GDF-15 levels. For some of these genes, GDF-15 appears to directly affect the DNA methylation. However, for other genes, the DNA methylation and GDF-15 might reflect coregulation probably due to underlying pathological conditions

Material and methods

Study populations

The Northern Sweden Population Health Study, NSPHS—The Northern Sweden Population Health Study (NSPHS) was initiated in 2006 to provide a health survey of the population in the parishes of Karesuando and Soppero, County of Norrbotten, and to study the medical consequences of lifestyle and genetics. This parish has about 3000 inhabitants who meet the eligibility criteria in terms of age (> 14 years) of which 1069 individuals participated in the study. The median age of these individuals was 50 years, ranging from 14 to 94 (Table 1). For each participant in the NSPHS, whole blood, plasma and serum samples were taken and immediately frozen and stored at −70°C. Genomic DNA for methylation analyses was extracted from previously frozen peripheral blood leukocytes using a phenol:chloroform protocol. The participants were asked to answer a questionnaire about their medical history, smoking habits and regular medication. The NSPHS study was approved by the local ethics committee at the University of Uppsala (Regionala Etikprövningsnämnden, Uppsala Dnr 2005:325) in compliance with the Declaration of Helsinki. All participants gave their written informed consent to the study including the examination of environmental and genetic causes of disease. In case the participant was not

full age, a legal guardian signed additionally. The procedure used to obtain informed consent and the respective informed consent form has recently been discussed in the light of present ethical guidelines (31) More information about the NSPHS has been published previously (32).

Prospective Investigation of the Vasculature in Uppsala Seniors, PIVUS—
Eligible subjects were those aged 70 years living in the community of Uppsala, Sweden. The subjects were chosen from the register of community living and were invited in a randomized order. The subjects received an invitation by letter within 2 months of their 70th birthday. Of the 2025 subjects invited, 1016 subjects participated giving a participation rate of 50.1% (33). Median age of individuals in PIVUS was 70.1 years, ranging from 69.8 to 70.7 years (Table 1). The study was approved by the Ethics Committee of the University of Uppsala and the participants gave informed consent. All subjects were investigated in the morning after an overnight fast, where collection of blood for DNA preparation and measurement of GDF-15 levels took place. No medication or smoking was allowed after midnight.

Determination of DNA Methylation Status

Genomic DNA for 743 samples from NSPHS and 1016 samples from PIVUS was bisulfite-converted using an EZ DNA methylation Kit (ZYMO research) according to the manufacturer's recommendations. The methylation status of the genomic DNA was then assessed using the Human Methylation450 BeadChip, (Illumina, San Diego, USA) according to the standard protocol. The Human Methylation450 BeadChip, (Illumina, San Diego, USA) has been validated in previous studies (34, 35).

Analysis of the raw data in NSPHS was performed using minfi. Normalization was performed using Subset-quantile Within Array Normalisation (SWAN) and a marker detection P-value $< = 0.01$ was applied, a Probe Call rate of >0.98, and an individual call rate of >0.98 was applied. A total of eight control samples were included: two positive controls, two negative controls, two duplicated samples that originated from the same blood samples and two duplicated samples that came from the same individual but blood samples that were taken with 3 years apart as has been described previously (36). Analysis of the raw data in PIVUS was performed using GenomeStudio 2011.1 from Illumina Inc. After exclusion of replicates a total of 1002 study participants had methylation data available for quality control procedures. Three samples were excluded based on poor bisulphite conversion efficiency, twelve samples due to low pass rate of CpG sites (<98.5% with a detection P-value > 0.01), six samples based on low SNP genotype match (>1 SNP mismatches) between genotypes from the methylation array and Omni/Metabochip genotyping chips and 14 samples due to abnormal leukocyte cell counts ($>10 \times 10^9$ cells/L). The signal intensities for the methylated and unmethylated state were then quantile normalised for each probe type separately.

GDF-15 measurements

In NSPHS GDF-15 levels were measured in plasma for 1004 individuals, using the Olink Proseek Multiplex Oncology I $^{96\times96}$ kit and quantified by real-time PCR using the Fluidigm

BioMark™ HD real-time PCR platform as described earlier (37, 38). In brief, two oligonucleotide-labelled antibodies probes were bound to GDF-15. If the two probes are in close proximity a PCR target sequence is formed by a proximity-dependent DNA polymerization, which is quantified using standard real-time PCR. These values are not converted into concentrations.

In PIVUS GDF-15 levels were determined by a validated immunoradiometric assay. The assay has a theoretical detection limit of 20 ng/L and allows reliable analyte quantification over a wide concentration range (~200-50 000 ng/L) (10). Serum samples were available from 1004 subjects for the present study.

SNP genotyping

DNA samples from NSPSH have previously been genotyped using Illumina Infinium HumanHap300v2 or Illumina Omni Express SNP bead microarrays, and unassayed genotypes has been imputed using the 1000 Genome Project reference panel (38). See also supplementary material for more information.

Estimation of cell fractions from the methylation data

To ensure that the results were not influenced by variation in cell fraction between samples, we estimated the fraction of CD8T-, CD4T-, NK- and B-cells, monocytes and granulocytes in the study samples. This was done using the R package minfi(39) that allows for estimating cell fractions in Illumina 450K methylation data from whole blood. This method is based on the methylation data published for flow-sorted cells (40) and algorithms derived from the study by Houseman *et al* (41).

Statistic analysis of DNA methylation data and GDF-15 levels

All statistics were performed using the stats library of R version 2.15.10 (42). The level of methylation was reported as average beta values. We used the annotation of DNA methylation sites provided by Illumina (HumanMethylation450_15017482_v.1.1.csv, accessed: 1st September, 2012), to assign CpG -sites to their corresponding genes (www.illumina.com). NSPHS; Beta values where adjusted for cell composition, year of collection and the first four principal components estimated from the DNA methylation data. Adjusted beta values for each marker were fitted to a mixed model using rank-transformed GDF-15 level, sex, age, year of collection, and smoking, as covariates. Following exclusion of individuals lacking GDF-15, DNA Methylation data or information on any of the covariates, a final set of 717 individuals was used in analysis. Since NSPHS is a population-based study including related individuals, special methods have been used to handle relatedness. All statistic analyses in NSPHS was performed using the R package GenABEL (43), which was developed to enable for statistic analyses of genetic data in related individuals as described previously (36). Population stratification was adjusted for through the kinship matrix. PIVUS; Following exclusion of individuals lacking GDF-15 phenotype data a final set of 963 individuals were used in analysis. Association between rank transformed GDF-15 and methylation beta values at each CpG-site were analyzed using linear models adjusted for sex, age, predicted white cell counts (See "Estimation of cell fractions from the methylation data" above), Bisulphite conversion plate (96-well plate) and

Bisulphite conversion efficiency mean (calculated from control probes). Meta analyses were performed using an inverse variance weighted fixed-effects model. Since GDF-15 levels were measured in different units, the estimates from PIVUS and NSPHS were standardized and z-scores included in the meta-analyses. Association between GDF-15 levels and cell fractions were estimated in NSPHS.

Myocardial infarction - NSPHS

The NSPHS is a population-based study with participants being unselected due to disease and health status. However, as many as 47 of the participants (all above 50 years of age) with methylation data available have experienced a prior MI. For the CpG-sites that were associated with GDF-15 levels, we investigated if they were also differentially methylated in the participants (above 50 years of age) with a prior MI compared to participants above age of 50 that (independent on other diseases) have not had a MI (N= 271). In the statistical test, MI was used as explanatory variable and methylation level as response variable also adjusting for precision variables (slide, array and year of sampling) and confounding variables (age and sex).

Enrichment analysis

Enrichment analysis was carried out using the gene ontology (GO) enrichment analysis and visualization web-based tool Gorilla (http://cbl-gorilla.cs.technion.ac.il/) (44). Gorilla performs enrichment analysis for two unranked lists of genes, one target and one background list using a hyper geometric model. P-values are corrected for multiple comparisons using false discovery rate (FDR). This analysis will identify GO terms (molecular function or biological processes) that are significantly enriched for genes that are differently methylated in relation to GDF-15 levels. To take into account that genes with more CpG-sites measured might be more likely to be associated with GDF-15 levels by chance we selected the most significant CpG-site within each gene and Bonferroni corrected the P-value for the number of CpG-sites that were measured within that gene.

Causal inference

To investigate if GDF-15 levels directly affect any of the associated methylation sites or vice versa, we performed a Mendelian randomization (MR) like approach. The MR approach has been described in detail previously (45) and is a commonly used method for evaluating the direct causal effect of biomarkers for various diseases (46, 47). In MR, SNPs are used as instrumental variables to divide a population into genotypic subgroups in an analogous way to the division of participants into arms in a randomized controlled trial. The basic idea of our MR approach is that individuals were randomized at birth to receive genetic variants that result in different "genetic" levels of e.g. GDF-15 levels. If GDF-15 is casually influencing the methylation level, we could also expect the known cis-regulatory variants in the *GDF15* region to influence the methylation. To avoid the risk of pleiotropy only *cis*-regulatory SNPs were considered for these analyses. In our MR analyses we used each of the three independent SNPs, previously shown to be associated with GDF-15 levels (17) (rs749451, rs888663 and rs1054564) as instrumental variables in three independent tests. The instrumental variables corresponded to the number of copies of the minor allele in each SNP assuming an additive genetic model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Abbreviations

|	
GDF-15	Growth differentiation factor 15
MI	Myocardial Infarction
FDR	False Discovery Rare
miR21	MicroRNA
GWAS	Genome wide association study
NSPHS	The Northern Sweden Population Health Study
PIVUS	Prospective Investigation of the Vasculature in Uppsala Seniors
SNP	Single nucleotide polymorphism
MR	Mendelian randomization
VMP1	vacuole membrane protein 1

References

1. Bootcov MR, Bauskin aR, Valenzuela SM, Moore aG, Bansal M, He XY, Zhang HP, Donnellan M, Mahler S, Pryor K, et al. MIC-1, a novel macrophage inhibitory cytokine, is a divergent member of the TGF-beta superfamily. Proc Natl Acad Sci U S A. 1997; 94:11514–9. [PubMed: 9326641]

2. Welsh JB, Sapinoso LM, Kern SG, Brown Da, Liu T, Bauskin AR, Ward RL, Hawkins NJ, Quinn DI, Russell PJ, et al. Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. Proc Natl Acad Sci U S A. 2003; 100:3410–5. [PubMed: 12624183]

3. Koopmann J, Rosenzweig CNW, Zhang Z, Canto MI, Brown Da, Hunter M, Yeo C, Chan DW, Breit SN, Goggins M. Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9. Clin Cancer Res. 2006; 12:442–6. [PubMed: 16428484]

4. Boyle GM, Pedley J, Martyn AC, Banducci KJ, Strutton GM, Brown Da, Breit SN, Parsons PG. Macrophage inhibitory cytokine-1 is overexpressed in malignant melanoma and is associated with tumorigenicity. J Invest Dermatol. 2009; 129:383–91. [PubMed: 18754039]

5. Wallentin L, Hijazi Z, Andersson U, Alexander J, De Caterina R, Hanna M, Horowitz J, Hylek E, Lopes R, Asberg S, et al. Growth differentiation factor 15, a marker of oxidative stress and inflammation, for risk assessment in patients with atrial fibrillation: insights from the Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE. Circulation. 2014; 130:1847–1858. [PubMed: 25294786]

6. Wallentin L, Lindholm D, Siegbahn A, Wernroth L, Becker R, Cannon C, Cornel J, Himmelmann A, Giannitsis E, Harrington R, et al. Biomarkers in relation to the effects of ticagrelor in comparison with clopidogrel in non-ST-elevation acute coronary syndrome patients managed with or without in-hospital revascularization: a substudy from the Prospective Randomized Platelet Inhibition a. Circulation. 2014; 129:293–303. [PubMed: 24170388]

7. Kempf T, Eden M, Strelau J, Naguib M, Willenbockel C, Tongers J, Heineke J, Kotlarz D, Xu J, Molkentin JD, et al. The Transforming Growth Factor-β Superfamily Member Growth-Differentiation Factor-15 Protects the Heart From Ischemia/Reperfusion Injury. Circ Res. 2006

8. Khan SQ, Ng K, Dhillon O, Kelly D, Quinn P, Squire IB, Davies JE, Ng LL. Growth differentiation factor-15 as a prognostic marker in patients with acute myocardial infarction. Eur Heart J. 2009; 30:1057–1065. [PubMed: 19168526]

9. Wollert KC, Kempf T, Peter T, Olofsson S, James S, Johnston N, Lindahl B, Horn-Wichmann R, Brabant G, Simoons ML, et al. Prognostic value of growth-differentiation factor-15 in patients with non-ST-elevation acute coronary syndrome. Circulation. 2007; 115:962–971. [PubMed: 17283261]

10. Kempf T, Horn-Wichmann R, Brabant G, Peter T, Allhoff T, Klein G, Drexler H, Johnston N, Wallentin L, Wollert KC. Circulating concentrations of growth-differentiation factor 15 in apparently healthy elderly individuals and patients with chronic heart failure as assessed by a new immunoradiometric sandwich assay. Clin Chem. 2007; 53:284–291. [PubMed: 17185363]

11. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001; 293:1068–1070. (80-.). [PubMed: 11498573]

12. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature. 2007; 447:425–432. [PubMed: 17522676]

13. Chu AY, Guilianini F, Grallert H, Dupuis J, Ballantyne CM, Barratt BJ, Nyberg F, Chasman DI, Ridker PM. Genome-wide association study evaluating lipoprotein-associated phospholipase A2 mass and activity at baseline and after rosuvastatin therapy. Circ Cardiovasc Genet. 2012; 5:676–685. [PubMed: 23118302]

14. Kumarswamy R, Volkmann I, Thum T. Regulation and function of miRNA-21 in health and disease. RNA Biol. 2011; 8:706–713. [PubMed: 21712654]

15. Dweep H, Sticht C, Pandey P, Gretz N. MiRWalk - Database: Prediction of possible miRNA binding sites by ' walking' the genes of three genomes. J Biomed Inform. 2011; 44:839–847. [PubMed: 21605702]

16. Thum T, Gross C, Fiedler J, Fischer T, Kissler S, Bussen M, Galuppo P, Just S, Rottbauer W, Frantz S, et al. MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. Nature. 2008; 456:980–984. [PubMed: 19043405]

17. Ho JE, Mahajan A, Chen MH, Larson MG, McCabe EL, Ghorbani A, Cheng S, Johnson AD, Lindgren CM, Kempf T, et al. Clinical and genetic correlates of growth differentiation factor 15 in the community. Clin Chem. 2012; 58:1582–1591. [PubMed: 22997280]

18. Qian Q, Zhou H, Chen Y, Shen C, He S, Zhao H, Wang L, Wan D, Gu W. VMP1 related autophagy and apoptosis in colorectal cancer cells: VMP1 regulates cell death. Biochem Biophys Res Commun. 2014; 443:1041–1047. [PubMed: 24365149]

19. Dusetti NJ, Jiang Y, Vaccaro MI, Tomasini R, Azizi Samir A, Calvo EL, Ropolo A, Fiedler F, Mallo GV, Dagorn J-C, et al. Cloning and expression of the rat vacuole membrane protein 1 (VMP1), a new gene activated in pancreas with acute pancreatitis, which promotes vacuole formation. Biochem Biophys Res Commun. 2002; 290:641–649. [PubMed: 11785947]

20. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA. 2004; 10:1957–1966. [PubMed: 15525708]

21. Löffler D, Brocke-Heidrich K, Pfeifer G, Stocsits C, Hackermüller J, Kretzschmar AK, Burger R, Gramatzki M, Blumert C, Bauer K, et al. Interleukin-6-dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer. Blood. 2007; 110:1330–1333. [PubMed: 17496199]

22. Fujita S, Ito T, Mizutani T, Minoguchi S, Yamamichi N, Sakurai K, Iba H. miR-21 Gene Expression Triggered by AP-1 Is Sustained through a Double-Negative Feedback Mechanism. J Mol Biol. 2008; 378:492–504. [PubMed: 18384814]

23. Ribas J, Ni X, Castanares M, Liu MM, Esopi D, Yegnasubramanian S, Rodriguez R, Mendell JT, Lupold SE. A novel source for miR-21 expression through the alternative polyadenylation of VMP1 gene transcripts. Nucleic Acids Res. 2012; 40:6821–33. [PubMed: 22505577]

24. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet. 2004; 5:522–531. [PubMed: 15211354]

25. Mowat, D., Wilson, M. Mowat-Wilson Syndrome. Management of Genetic Syndromes. Third Edition. 2010. p. 517-527.

26. Odiete O, Hill MF, Sawyer DB. Neuregulin in cardiovascular development and disease. Circ Res. 2012; 111:1376–85. [PubMed: 23104879]

27. Orth T, Dorner T, Meyer Zum Buschenfelde KH, Mayet WJ. Complete congenital heart block is associated with increased autoantibody titers against calreticulin. Eur J Clin Invest. 1996; 26:205–215. [PubMed: 8904349]

28. Boileau C, Guo D-C, Hanna N, Regalado ES, Detaint D, Gong L, Varret M, Prakash SK, Li AH, d'Indy H, et al. TGFB2 mutations cause familial thoracic aortic aneurysms and dissections associated with mild systemic features of Marfan syndrome. Nat Genet. 2012; 44:916–921. [PubMed: 22772371]

29. Parlakian A, Tuil D, Hamard G, Tavernier G, Hentzen D, Concordet J-P, Paulin D, Li Z, Daegelen D. Targeted inactivation of serum response factor in the developing heart results in myocardial defects and embryonic lethality. Mol Cell Biol. 2004; 24:5281–5289. [PubMed: 15169892]

30. Monick MM, Beach SRH, Plume J, Sears R, Gerrard M, Gene H, Philibert RA. NIH Public Access. Am J Med Genet B Neurospychiatr Genet. 2013; 159B:141–151.

31. Mascalzoni D, Janssens AC, Stewart A, Pramstaller P, Gyllensten U, Rudan I, van Duijn CM, Wilson JF, Campbell H, Quillan RM, et al. Comparison of participant information and informed consent forms of five European studies in genetic isolated populations. Eur J Hum Genet. 2010; 18:296–302. [PubMed: 19826451]

32. Johansson A, Marroni F, Hayward C, Franklin CS, Kirichenko AV, Jonasson I, Hicks AA, Vitart V, Isaacs A, Axenovich T, et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. Hum Mol Genet. 2009; 18:373–380. [PubMed: 18952825]

33. Lind L, Fors N, Hall J, Marttala K, Stenborg A. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: The Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. Arterioscler Thromb Vasc Biol. 2005; 25:2368–2375. [PubMed: 16141402]

34. Roessler J, Ammerpohl O, Gutwein J, Hasemeier B, Anwar S, Kreipe H, Lehmann U. Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. BMC Res Notes. 2012; 5:210. [PubMed: 22546179]

35. Sandoval J, Heyn HA, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics. 2011; 6:692–702. [PubMed: 21593595]

36. Besingi W, Johansson Å. Smoke-related DNA methylation changes in the etiology of human disease. Hum Mol Genet. 2014; 23:2290–2297. [PubMed: 24334605]

37. Assarsson E, Lundberg M, Holmquist G, Björkesten J, Thorsen SB, Ekman D, Eriksson A, Dickens ER, Ohlsson S, Edfeldt G, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. PLoS One. 2014; 9

38. Enroth S, Johansson Å, Enroth SB, Gyllensten U. Strong effects of genetic and lifestyle factors on biomarker variation and use of personalized cutoffs. Nat Commun. 2014; 5:4684. [PubMed: 25147954]

39. Hansen K, Aryee M. minfi: Analyze Illumina's 450k methylation arrays, Vol. R package version 1.8.7. 2013

40. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén S-E, Greco D, Söderhäll C, Scheynius A, Kere J. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One. 2012; 7:e41361. [PubMed: 22848472]

41. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012; 13:86. [PubMed: 22568884]

42. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2012. ISBN 3-900051-07-0, URL http://www.R-project.org/ *R Found. Stat. Comput. Vienna, Austria*

43. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007; 23:1294–1296. [PubMed: 17384015]

44. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009; 10:48. [PubMed: 19192299]

45. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. Int J Epidemiol. 2013; 42:1134–44. [PubMed: 24062299]

46. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Holm H, Ding EL, Johnson T, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet. 2012; 380:572–580. [PubMed: 22607825]

47. Mokry LE, Ahmad O, Forgetta V, Thanassoulis G, Richards JB. Mendelian randomisation applied to drug development in cardiovascular disease: a review. J Med Genet. 2015; 52:71–9. [PubMed: 25515070]
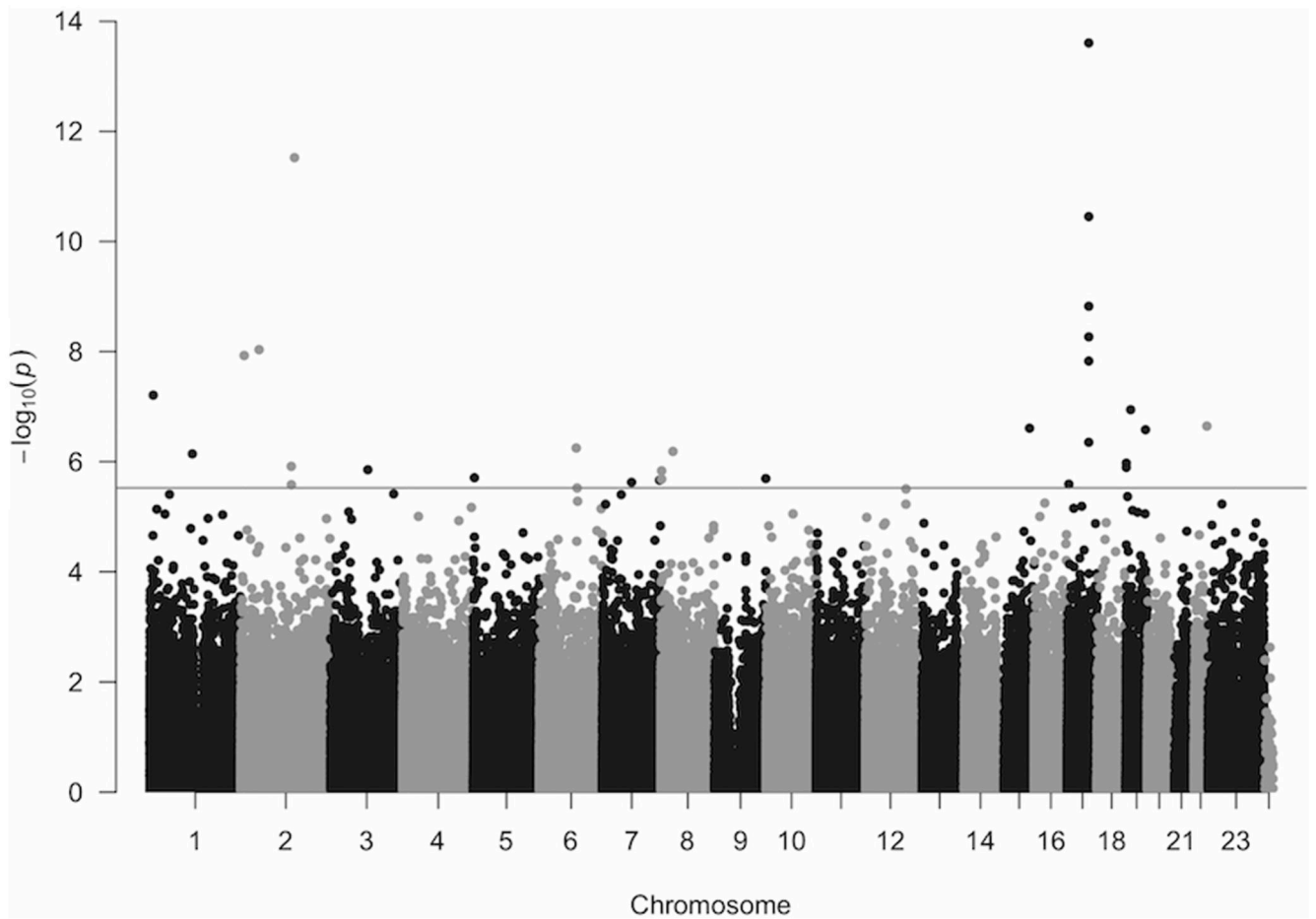
Figure 1.
Manhattan plot for the genome-wide DNA methylation study in NSPHS. Significance threshold is marked in with a grey line (FDR q-value 0.05)
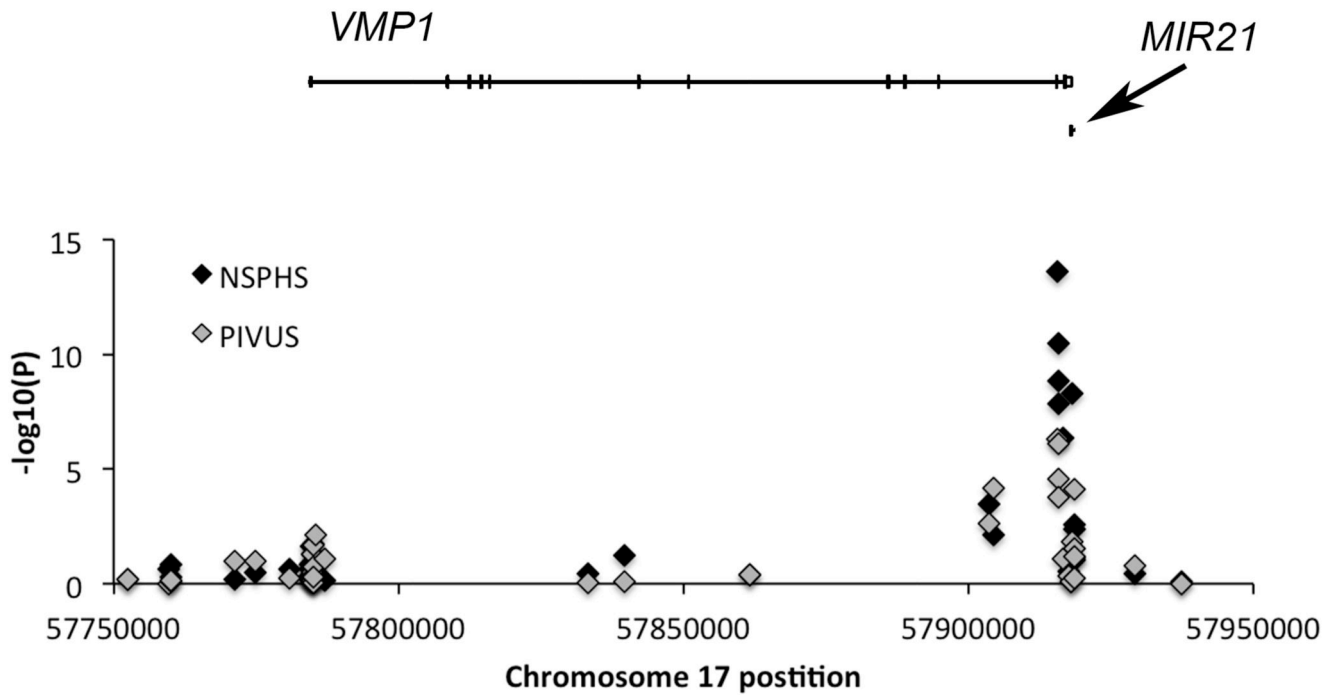
Figure 2.
Regional plot of the association between GDF-15 levels and DNA methylation in the *VNP1/MIR21* region. Both genes are coded on the positive strand with the transcription-starting site being located at the most left part of respective gene. The black marks is the –log10(*P*-values) for the association in the NSPHS and the grey marks in the PIVUS study.
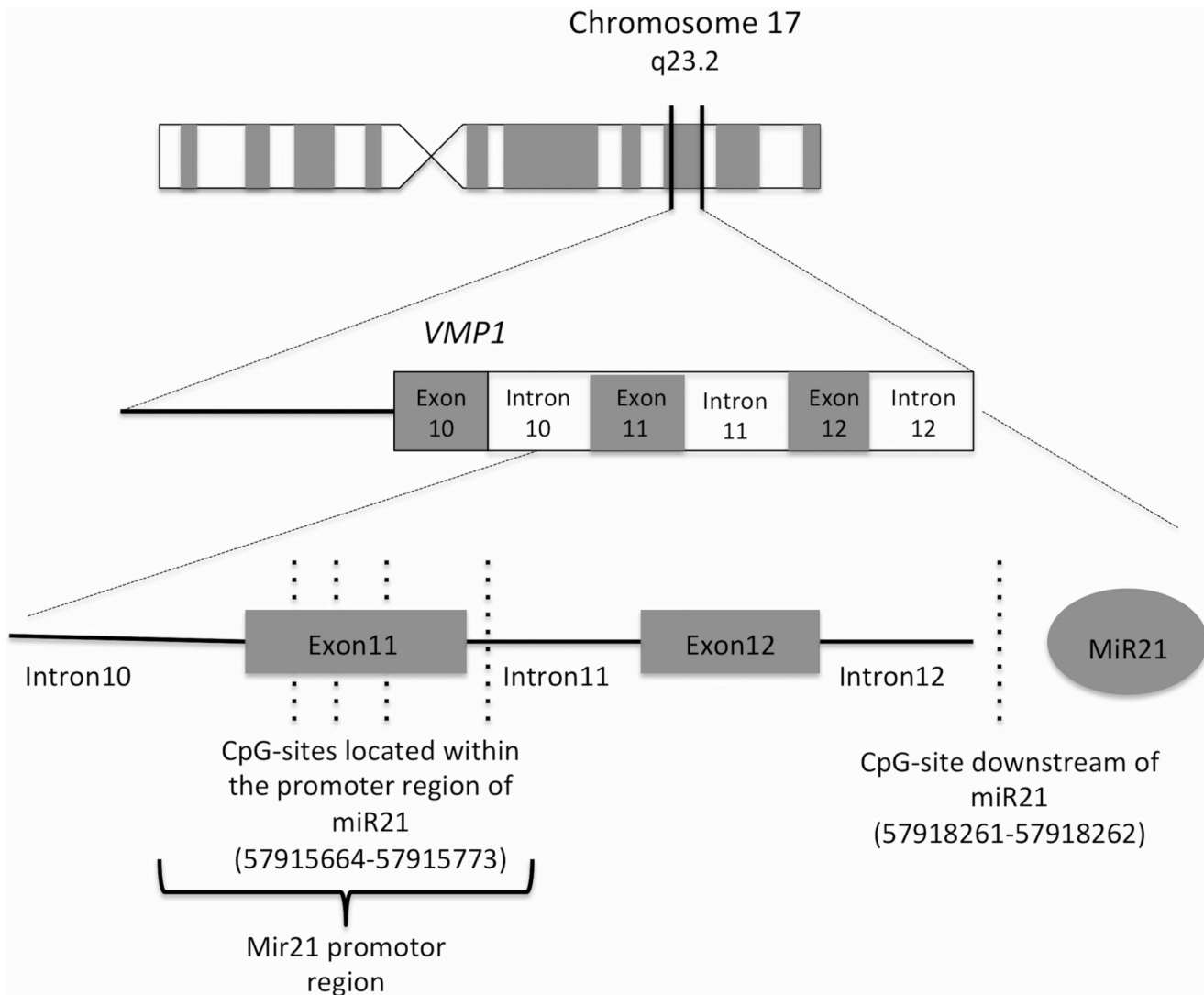
Figure 3.
Genomic location of miRNA 21 (*MIR21*), the promoter region of *MIR21*(20–22) in relation to our associated CpG -sites (dashed black lines). *MIR21* is located on chromosome 17q23.2, downstream of the protein-coding gene *VMP1* and 3kb downstream of the predicted promoter. Image is not drawn to scale.
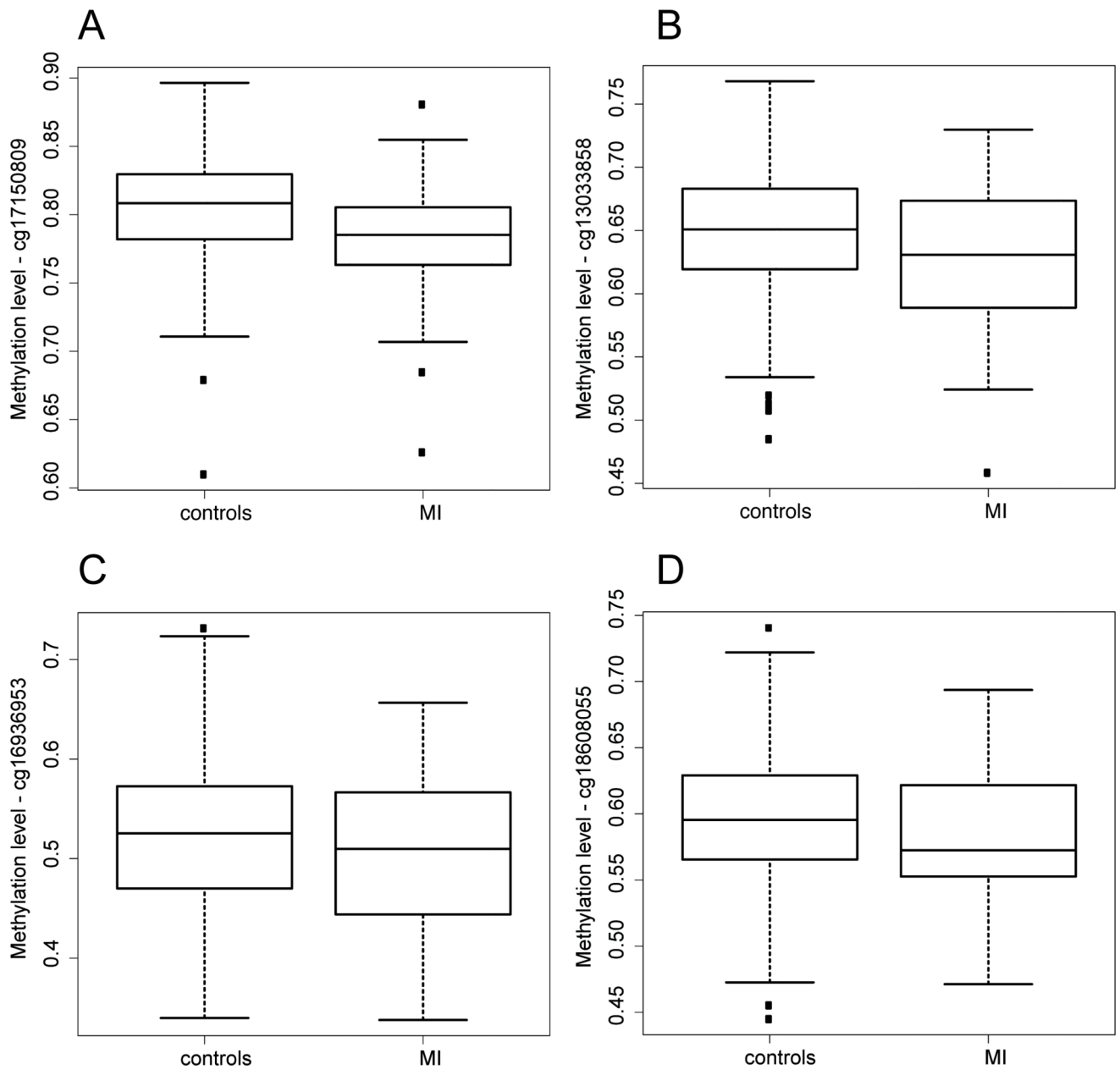
Figure 4.
(a-d). Boxplot for the methylation level at one of the CpG-sites in participants with and without a history of Myocardial Infarction (MI). Boxplots show mean and interquartile ranges. Nominal p-values for the association between MI and Methylation level (adjusted for covariates) is $P = 0.00039$, $P = 0.046$, $P = 0.049$ and $P = 0.049$, for a, b, c and d, respectively.

Table 1

Clinical characteristics of the study individuals with methylation data and GDF-15 estimates that passed QC and was included in the methylation analyses.

	Discovery (NPHSHS)	Replication (PIVUS)
N	717	963
Mean age, years	50.4	70.1
Men / Women	382 / 335	483 / 480
Men /Women (proportion)	0.53 / 0.47	0.50 / 0.50

Table 2

Genome-wide DNA methylation study summary statistics for Discovery (NSPHS), Replication cohort (PIVUS) and meta-analysis (NSPHS+PIVUS).

CHR	CpG-site	Position. bp	Gene	NSPHS		PIVUS		META-ANALYSIS	
				P	Effect	P	Effect	P	Effect
17	cg16936953	57915665	VMP1	2.48×10^{-14}	-3.52	4.80×10^{-7}	-0.88	3.34×10^{-19}	-2.93
2	cg20995564	145172035	ZEB2	3.00×10^{-12}	-4.34	0.021	-0.53	6.81×10^{-11}	-2.83
17	cg12054453	57915717	VMP1	3.53×10^{-11}	-2.85	2.62×10^{-5}	-0.71	1.48×10^{-14}	-2.39
17	cg01409343	57915740	VMP1	1.50×10^{-9}	-4.14	7.59×10^{-7}	-1.31	7.21×10^{-15}	-3.82
17	cg27023597	57918262	MIR21	5.39×10^{-9}	-3.92	0.015	-0.53	1.52×10^{-8}	-2.49
17	cg18942579	57915773	VMP1	1.48×10^{-8}	-3.20	0.0002	-0.76	3.25×10^{-11}	-2.58
19	cg19821297	12890029	Upstream of HOOK2	1.13×10^{-7}	-3.47	0.001	-0.90	4.40×10^{-10}	-2.96
22	cg09349128	50327986	Downstream of CRELD2	2.25×10^{-7}	-5.58	1.42×10^{-5}	-1.47	2.86×10^{-11}	-4.61
19	cg19369616	53447272	ZNF321	2.61×10^{-7}	3.62	0.013	0.52	2.65×10^{-7}	2.24
6	cg17150809	99461062	Upstream of FBXL4	5.62×10^{-7}	-4.69	0.003	-0.87	4.30×10^{-8}	-3.31
1	cg11393173	116369577	Downstream of NHLH2	7.18×10^{-7}	-4.76	0.001	-1.31	2.27×10^{-9}	-4.15
19	cg18608055	1130866	SBNO2	1.06×10^{-6}	-3.98	0.020	-0.69	3.93×10^{-7}	-2.89
19	cg07573872	1126342	SBNO2	1.27×10^{-6}	-3.72	3.62×10^{-7}	-1.28	2.23×10^{-12}	-3.55
3	cg12992827	101901234	Upstream of ZPLD1	1.40×10^{-6}	-3.80	2.51×10^{-6}	-1.09	2.37×10^{-11}	-3.25
17	cg13165240	3715743	C17orf85	2.56×10^{-6}	3.92	0.001	0.98	1.19×10^{-8}	3.21
12	cg13033858	109248326	SSH1	3.13×10^{-6}	-3.50	0.007	-0.72	2.32×10^{-7}	-2.68

Table 3

Replication of previously associated SNPs close to *GDF15* on chromosome 19.

SNP	A1 (MAF)	POS	P_{NSPSH}	$P_{Framingham\ meta}$	Variance Explained$_{NSPHS}$
rs3746181	G (0.085)	18477017	6.23×10^{-5}	6.22×10^{-32}	0.31%
rs1363120	C (0.083)	18482304	2.51×10^{-6}	2.38×10^{-30}	0.31%
rs888663	T (0.081)	18484922	2.93×10^{-6}	7.12×10^{-31}	0.30%
rs1227731	G (0.14)	18497903	1.38×10^{-5}	8.22×10^{-28}	0.26%
rs1054564	G (0.11)	18499815	1.38×10^{-5}	2.14×10^{-18}	0.26%
rs3195944	A (0.11)	18476711	7.08×10^{-5}	1.16×10^{-17}	0.22%
rs749451	C (0.47)	18479647	7.04×10^{-5}	1.21×10^{-29}	0.21%
rs1043063	C (0.36)	18480171	0.11	4.58×10^{-9}	0.036%
rs17725099	A (0.30)	18482358	0.211	2.59×10^{-13}	0.022%

Table 4

Biological processes that were significantly (FDR q-val 0.05) enriched (higher than 2) for genes differentially methylated for GDF-15 levels in the meta-analysis.

GO Term	Description	FDR q-value	Fold Enrichment[*]
GO:0002209	behavioral defense response	4.55×10^{-2}	6.11
GO:0001662	behavioral fear response	4.39×10^{-2}	6.11
GO:0055007	cardiac muscle cell differentiation	4.81×10^{-2}	5.86
GO:0035051	cardiocyte differentiation	4.15×10^{-2}	5.38
GO:0035136	forelimb morphogenesis	4.49×10^{-2}	4.71
GO:0008344	adult locomotory behavior	4.64×10^{-2}	3.36
GO:0030326	embryonic limb morphogenesis	3.93×10^{-2}	3.33
GO:0035113	embryonic appendage morphogenesis	3.77×10^{-2}	3.33
GO:0035107	appendage morphogenesis	3.45×10^{-2}	3.25
GO:0035108	limb morphogenesis	3.16×10^{-2}	3.25
GO:0007626	locomotory behavior	1.51×10^{-2}	2.92
GO:0044708	single-organism behavior	2.04×10^{-2}	2.18
GO:0007610	behavior	1.71×10^{-2}	2.10
GO:0007389	pattern specification process	2.75×10^{-2}	2.05

[*] Enrichment is defined as (b/n)/(B/N). N is the total number of genes; B is the total number of genes associated with a specific GO term; n is the flexible cutoff, i.e. the automatically determined number of genes in the ´target set´ and b is the number of genes in the ´target set´ that are associated with a specific GO term. For further information about these results, see Supplementary Table 6 and 7.