



(51) International Patent Classification:

C12Q 1/68 (2018.01) *C07H 21/04* (2006.01)*G01N 33/48* (2006.01) *A61K 39/395* (2006.01)*G01N 33/50* (2006.01) *C40B 40/08* (2006.01)*G01N 33/564* (2006.01)

(21) International Application Number:

PCT/US2018/028747

(22) International Filing Date:

21 April 2018 (21.04.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/488,320 21 April 2017 (21.04.2017) US

(71) Applicant: WAKE FOREST UNIVERSITY HEALTH SCIENCES [US/US]; Medical Center Boulevard, Winston Salem, North Carolina 27157 (US).

(72) Inventor: LIU, Yongmei; 391 Technology Way, Suite 199, Winston-Salem, North Carolina 27101 (US).

(74) Agent: GILES, P. Brian; Thomas | Horstemeyer LLP, 3200 Windy Hill Rd SE, Suite 1600E, Atlanta, Georgia 30339 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: ARID5B OVEREXPRESSION IN INFLAMMATORY DISEASES

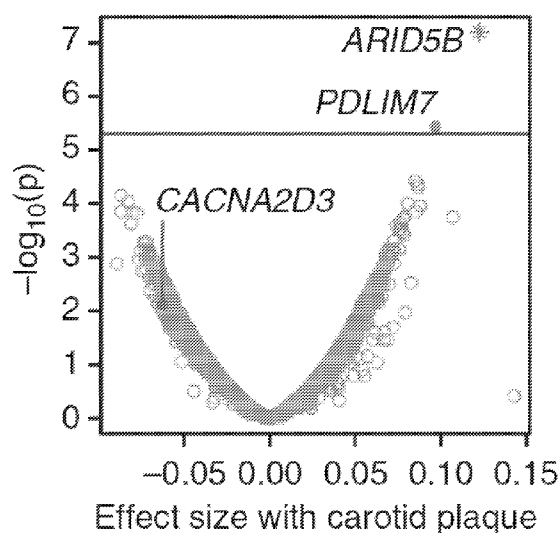


FIG. 1A

(57) Abstract: Disclosed herein are compositions and methods for diagnosing and treating inflammatory diseases based on the disclosed overexpression and hypomethylation of *ARID5B* that is associated with an increased risk of inflammatory disease. Therefore, methods are disclosed that involve assaying samples from subjects for *ARID5B* gene expression, or for methylation of an *ARID5B* CpG site that is negatively correlated with *ARID5B* gene expression. Also disclosed are methods that involves administering to the subject a composition comprising an *ARID5B* inhibitor in an amount effective to reduce endogenous expression levels of *ARID5B* in the subject.

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

ARID5B OVEREXPRESSION IN INFLAMMATORY DISEASES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims benefit of U.S. Provisional Application No. 62/488,320, filed April 21, 2017, which is hereby incorporated herein by reference in its entirety.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was made with Government Support under Grant No. HL101250 and HL135009 awarded by the National Institutes of Health. The Government has certain rights in the invention.

BACKGROUND

Despite improvements in prevention and therapy of atherosclerotic cardiovascular disease (CVD), it remains to be the leading cause of death in the developed world. Although traditional risk factors for atherosclerosis are established, there is substantial unexplained variation in the atherosclerosis phenotype. Most acute coronary syndromes and ischemic strokes are caused by rupture-prone vulnerable plaques, however, the risk factors for plaque vulnerability are largely unknown. Genome-wide association studies (GWAS) have identified a number of susceptibility variants for CVD; however, they explain only a small percent of CVD risk. The available therapies continue to primarily target hyperlipidemia and prevention of thrombosis. To develop effective therapeutic approaches for primary and secondary CVD prevention, a better understanding of the molecular mechanisms and pathophysiology of the disease is needed.

SUMMARY

Disclosed herein are compositions and methods for diagnosing and treating inflammatory diseases based on the disclosed overexpression and hypomethylation of *ARID5B* that is associated with an increased risk of inflammatory disease. Therefore, a method is disclosed that involves obtaining a sample from a subject, such as a subject having or suspected of having an inflammatory disease, and assaying the sample for *ARID5B* gene expression, or for methylation of an *ARID5B* CpG site that is negatively correlated with *ARID5B* gene expression.

In some embodiments, the CpG methylation site is one or more of the underlined CpG site found in the nucleic acid sequence:

ATAAATATTYGGTGTTATAATAGAGTAGTAATTAAATGGAAATTTTAATATA
ATTGTTAATAGTAGTAAGAATTGGAATAGYGTTAGGTATATAAAGTTGATGTTTTG
TYGTTGGTTGGTTGTTAGGAGTATAGTGA (SEQ ID NO:1). For example, the CpG
methylation site can be cg25953130.

Also disclosed is a method for diagnosing or prognosing an inflammatory
disease in a subject that involves assaying a sample from the subject for *ARID5B*
expression levels, wherein detection of elevated *ARID5B* expression levels
compared to a control is an indication that the subject has an inflammatory disease.
Since methylation of the disclosed *ARID5B* CpG sites negatively correlate with
ARID5B gene expression, in some embodiments, *ARID5B* expression levels are
assayed by assaying for methylation of the *ARID5B* CpG sites.

The inflammatory disease of the disclosed methods can be selected from the
group comprising diabetes mellitus type 2 or a cardiovascular diseases (e.g.
atherosclerosis or stroke).

Also disclosed is a method for treating an inflammatory disease in a subject
that involves obtaining a sample from a subject having or suspected of having an
inflammatory disease, assaying the sample for *ARID5B* expression levels, and
treating the subject for an inflammatory disease if the *ARID5B* expression levels are
elevated compared to a control.

For example, in some embodiments the subject presents with a high risk for
cardiovascular diseases (CVD), e.g. based on family history, obesity, diabetes,
hyperlipidemia, hypertension, or prior incidence of CVD. In these embodiments, the
subject can be treated for CVD if *ARID5B* expression levels are elevated. Treatment
for cardiovascular disease include lifestyle changes (e.g. diet and exercise), statins,
and/or blood thinners. In some embodiments, the subject is treated with an anti-
inflammatory agent if *ARID5B* expression levels are elevated.

In some cases, the choice of treatment is based on the level of *ARID5B*
expression. In particular, the disclosed methods use *ARID5B* expression levels to
stratify subjects based on disease risk. Subjects determined to be at higher risk
based on elevated *ARID5B* expression can be treated with standard of care for that
risk, including the use of aggressive treatments earlier than would otherwise be used.
For example, all subjects with CVD should be recommended lifestyle modification. As
ARID5B levels elevate, the subject can be treated with a statin. As levels elevate
further, the subject can be further treated with ezetimibe and/or PCSK9 inhibitors. As
levels elevate even further, the subject can be further treated with a PCSK9 antibody.

In some embodiments the subject presents with a high risk for diabetes or pre-diabetes, e.g. based on family history, weight, blood glucose levels, and/or prior incidence of diabetes. In these embodiments, the subject is treated for diabetes or pre-diabetes if *ARID5B* expression levels are elevated. Treatment for diabetes include lifestyle changes (diet and exercise), metformin, insulin, sulfonylurea, GLP-1 agonists, thiazolidinedione, glinide, SGLT2 inhibitors, DDP-4 inhibitors, alpha-glucosidase inhibitors, and pramlintide. In some embodiments, the subject is treated with an anti-inflammatory agent if *ARID5B* expression levels are elevated. In some cases, the choice of treatment is based on the level of *ARID5B* expression.

In some embodiments, *ARID5B* expression levels are used to screen subjects for the presence or risk of an inflammatory disease, such as CVD or diabetes, before the subject presents with any clinical symptoms. In these embodiments, detection of elevated *ARID5B* expression levels can be an indication that the subject should be screened for an inflammatory disease, such as CVD or diabetes. In some cases, the subject is directed to make lifestyle changes (e.g. diet and exercise) to reduce the risk of disease onset. In some cases, the subject is treated prophylactically with anti-inflammatory agents to prevent the onset of an inflammatory disease. In some cases, the subject is treated using standards of care for CVD and/or diabetes prevention, e.g. statins.

Also disclosed is a method of treating a subject with elevated *ARID5B* expression levels that involves administering to the subject an anti-inflammatory agent. For example, in some embodiments, the subject having elevated *ARID5B* expression levels is treated with an anti-IL6 drug, such as tocilizumab (Actemra), siltuximab (Sylvant), Sarilumab, clazakizumab, olokizumab (CDP6038), elsilimomab, BMS-945429(ALD518), sirukumab (CNTO 136), CPSI-2364, ARGX-109, FE301, and FM101.

Also disclosed is a method for treating a subject with an inflammatory disease that involves administering to the subject a composition comprising an *ARID5B* inhibitor in an amount effective to reduce endogenous expression levels of *ARID5B* in the subject.

In some embodiments, the *ARID5B* inhibitor is a gene silencing agent. As an example, the gene silencing agent can be an antisense oligonucleotide, siRNA, or shRNA. In some embodiments, the *ARID5B* inhibitor is a CRISPR-Cas system that methylates a disclosed *ARID5B* CpG site that negatively correlates with *ARID5B* gene expression. In some embodiments, prior to treatment, the method further comprises detecting elevated *ARID5B* expression levels in a sample from the

subject. Since methylation of the disclosed *ARID5B* CpG sites negatively correlate with *ARID5B* gene expression, in some embodiments, *ARID5B* expression levels are assayed by assaying for methylation of the *ARID5B* CpG sites.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF DRAWINGS

Figure 1. Transcriptomic associations with Carotid Plaque and CAC Scores in 1,208 MESA participants. The two volcano plots showing effect size (β) of $-\log_{10}(p\text{-value})$ for associations of each mRNA expression (for 10,989 unique mRNA expression) with A) carotid plaque score and B) CAC. The red line illustrates the threshold for an $FDR \leq 0.05$ (linear regression adjusting for age, sex, race, and study site). Expression of *ARID5B* and *PDLIM7* (bolded) were associated with both carotid plaque and CAC score; green star indicates mRNA expression most significantly associated with carotid plaque or CAC. The two forest plots showing the direction and effect size (β) for associations of each mRNA expression (in Panel A and B) with C) carotid plaque score and D) CAC score in the full model (including traditional CVD risk factors), overall and by the two independent studies (study 1: JHU + CO, red; study 2: UMN + WFU, blue); $\ddagger$ significantly different effect sizes observed by study (pinteraction < 0.05). E) The forest plot showing *ARID5B* mRNA expression associations with carotid plaque, stratified by age (<65 years and ≥ 65 years), sex, ethnicity/race, CVD and statin use.

Figure 2. Methylomic associations with Carotid Plaque and CAC Scores in 1,208 MESA participants. The two Manhattan plots showing chromosomal locations of $-\log_{10}(p\text{-value})$ for associations of each CpG site (for 484,817 CpG sites) with A) carotid plaque score and B) CAC score. The red line illustrates the threshold for an $FDR \leq 0.05$; linear regression adjusting for age, sex, race, and study site). One CpG at ILVBL (bolded, green dot) was associated with both carotid plaque and CAC; green star indicates CpGs most significantly associated with carotid plaque or CAC. Notably, for *ARID5B*, both methylation (cg25953130, green dot) and mRNA expression were significantly associated with carotid plaque score. The two forest plots showing the direction and effect size (β) for associations of each CpG sites (in Panel A and B, the most significant association at each unique loci is shown) with C) Carotid plaque score and D) CAC in the full model (including traditional CVD risk factors), overall and by the two independent studies (study 1: JHU + CO, red; study

2: UMN + WFU, blue); ‡ significantly different effect sizes observed by study (pinteraction<0.05). E) The forest plot showing ARID5B cg25953130 methylation associations with carotid plaque score, stratified by age (<65 years and ≥65 years), sex, ethnicity/race, CVD and statin use.

5 Figure 3. In vivo and in silico functional analysis of ARID5B CpG cg25953130
A) ARID5B mRNA expression (y-axis; normalized value) is significantly negatively correlated (Pearson r) with methylation of a CpG (cg25953130) in 1,264 CD14+ samples from MESA participants. B) A regional association plot of ARID5B CpG methylation with carotid plaque score in MESA (y-axis: -log₁₀ (p-value), x-axis:
10 position on chromosome (chr) 10) is shown in the top panel; the bottom panel shows the ARID5B expression-associated CpG (cg25953130, chr10:63,753,550, hg19, indicated by the light blue line) located in an ARID5B intron, overlaps a DNase hypersensitive site in a CD14+ sample from BLUEPRINT, histone marks indicative of a strong enhancer/promoter in a CD14+ and B cell line sample from BLUEPRINT and
15 ENCODE (see Supplementary Fig. 3 for ChromHMM color code), as well as a transcription factor binding site for EP300, detected in a neuroblastoma cell line (SK-N-SH_RA). C) A physical interaction is detected between the ARID5B promoter and the region near the ARID5B CpG cg25953130 by both HiC and ChIA-PET data. The heatmap (white, low; red, high; top panel) graphically displays Hi-C interaction counts
20 (the normalized number of contacts between a pair of loci) for the large region surrounding ARID5B in a B cell line (GM12878; reported by Rao et al. 2014); active enhancer marks (H3K27ac and H3K4me1 in GM12878 from ENCODE) are shown below; the bottom panel zooms in on ARID5B (blue highlighted region) Hi-C interaction. There were 1937 contacts between the ARID5B promoter and the
25 cg25953130 region. Below the Hi-C interaction is a depiction of detected ChIA-PET interactions (Chromatin Interaction Analysis by Paired-End Tag sequencing) for the region flanking ARID5B in the B cell line (GM12878; reported by Heidari et al. 2015), represented by blue curves.

30 Figure 4. ARID5B expression and methylation are associated with subclinical and clinical CVD risk. Odds ratio of subclinical and clinical CVD (in the full model with adjustment of traditional CVD risk factors and statin use) by tertiles of A) ARID5B mRNA expression and B) ARID5B methylation (cg25953130).

35 Figure 5. siRNA Knockdown of ARID5B alters immune/inflammatory response and lipid metabolism genes. After 3 h of LPS treatment (100 ng ml⁻¹) of THP1-monocyte ARID5B knockdown samples compared to control samples (scrambled siRNA), A) the transcriptome (N=8 per group) were enriched with the inflammatory

response genes including the listed bio-functions and canonical pathways (enrichment FDR<0.05 from IPA); proportion of genes upregulated shown as red, downregulated shown as blue. B) Relative IL1A levels (mean $\pm$ standard deviation) of mRNA expression in cells (n=3 per group, by RT-PCR) and protein expression in culture media (n=3 per group, by ELISA) decreased. C) Cell migration (mean $\pm$ standard deviation, n=10 and 4 per group in experiment 1 and 2, respectively), and phagocytosis (mean $\pm$ standard deviation, n=8, 4, and 4 per group in experiment 1, 2, and 3, respectively) were inhibited.

Figures 6A and 6B. Histograms of subclinical CVD phenotypes.

Figures 7A to 7D. Reproducibility of Microarray mRNA Expression and DNA Methylation Data Overtime. A) Correlation of the repeated measurements of transcriptomic profiles for one subject between two visits 5-month apart; B) correlation of transcriptomic profiles between two subjects at visit 1; C) Correlation of the repeated measurements of methylation profiles for one subject between two visits 5-month apart; D) correlation of methylation profiles between two subjects at visit 1.

Figure 8. Chromatin Map of Monocytes. Six chromatin states (top panel) predicted using ChromHMM and four CD14+ histone marks (from BLUEPRINT) are enriched with various genomic features (middle panel); annotation of the six chromatin states is provided (bottom panel).

Figures 9A and 9B. siRNA Knockdown of ARID5B alters immune/inflammatory response and lipid metabolism genes. A) Relative ARID5B mRNA expression levels in human THP1-monocytes after treatment with two ARID5B siRNA individually, and in combination. Levels shown are the average ($\pm$ S.E.M.). After 3 h of LPS treatment (100 ng ml⁻¹), the THP1-monocyte transcriptome in 8 ARID5B knockdown samples compared to 8 control samples (scrambled siRNA) were enriched with inflammatory response genes; B) heatmap shows mRNA expression of selected key inflammatory response genes (including cytokine, interferon signaling, and antigen processing and presentation genes) in the 8 scrambled siRNA samples compared to 8 siARID5B samples.

Figures 10A and 10B. Time Course of LPS-induced TNF and IL1A Gene Expression in Human THP1 monocytes. Gene expression of (A) TNF and (B) IL1A, in LPS (100 ng/ml) stimulated THP1 cells.

Figures 11A and 11B are bar graphs showing diabetes (Fig. 11A) and subclinical CVD (Fig. 11B) risk as a function of *ARID5B* expression with overweight/obesity or impaired glucose tolerance/diabetes status.

DETAILED DESCRIPTION

The disclosed subject matter can be understood more readily by reference to the following detailed description, the Figures, and the examples included herein.

Before the present disclosure is described in greater detail, it is to be understood that this disclosure is not limited to particular embodiments described, and as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the disclosure. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and are also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present disclosure, the preferred methods and materials are now described.

All publications and patents cited in this specification are cited to disclose and describe the methods and/or materials in connection with which the publications are cited. All such publications and patents are herein incorporated by references as if each individual publication or patent were specifically and individually indicated to be incorporated by reference. Such incorporation by reference is expressly limited to the methods and/or materials described in the cited publications and patents and does not extend to any lexicographical definitions from the cited publications and patents. Any lexicographical definition in the publications and patents cited that is not also expressly repeated in the instant application should not be treated as such and should not be read as defining any terms appearing in the accompanying claims. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present disclosure is not entitled to antedate such publication by virtue of prior disclosure. Further, the dates of publication provided could be different from the actual publication dates that may need to be

independently confirmed. As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present disclosure. Any recited method can be carried out in the order of events recited or in any other order that is logically possible.

It is understood that the disclosed methods and systems are not limited to the particular methodology, protocols, and systems described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

Definitions

Unless otherwise expressly stated, it is in no way intended that any method or aspect set forth herein be construed as requiring that its steps be performed in a specific order. Accordingly, where a method claim does not specifically state in the claims or descriptions that the steps are to be limited to a specific order, it is no way intended that an order be inferred, in any respect. This holds for any possible non-express basis for interpretation, including matters of logic with respect to arrangement of steps or operational flow, plain meaning derived from grammatical organization or punctuation, or the number or type of aspects described in the specification.

As used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

The word "or" as used herein means any one member of a particular list and can also include any combination of members of that list.

Ranges can be expressed herein as from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, a further aspect includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by use of the antecedent "about," it will be understood that the particular value forms a further aspect. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as "about" that particular value in addition to the value itself. For example, if the value "10" is disclosed, then "about 10" is also disclosed. It is also understood

that each unit between two particular units are also disclosed. For example, if 10 and 15 are disclosed, then 11, 12, 13, and 14 are also disclosed.

As used herein, the terms "optional" or "optionally" means that the subsequently described event or circumstance can or can not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

As used herein, the term "subject" refers to the target of administration, e.g., an animal. Thus, the subject of the herein disclosed methods can be a vertebrate, such as a mammal. The subject of the herein disclosed methods can be a human. The term does not denote a particular age or sex. Thus, adult and newborn subjects, as well as fetuses, whether male or female, are intended to be covered. In one aspect, the subject is a patient. A patient refers to a subject afflicted with a disease or disorder, such as, for example, cancer and/or aberrant cell growth. The term "patient" includes human and veterinary subjects. In an aspect, the subject has been diagnosed with a need for treatment for an inflammatory disease. In another aspect, the subject has not yet been diagnosed and the disclosed methods to predict a higher risk for developing an inflammatory disease.

The terms "treating", "treatment", "therapy", and "therapeutic treatment" as used herein refer to curative therapy, prophylactic therapy, or preventative therapy.

As used herein, the terms refers to the medical management of a subject or a patient with the intent to cure, ameliorate, stabilize, or prevent a disease, pathological condition, or disorder, such as, for example, cancer or a tumor. This term includes active treatment, that is, treatment directed specifically toward the improvement of a disease, pathological condition, or disorder, and also includes causal treatment, that is, treatment directed toward removal of the cause of the associated disease, pathological condition, or disorder. In addition, this term includes palliative treatment, that is, treatment designed for the relief of symptoms rather than the curing of the disease, pathological condition, or disorder; preventative treatment, that is, treatment directed to minimizing or partially or completely inhibiting the development of the associated disease, pathological condition, or disorder; and supportive treatment, that is, treatment employed to supplement another specific therapy directed toward the improvement of the associated disease, pathological condition, or disorder. In various aspects, the term covers any treatment of a subject, including a mammal (e.g., a human), and includes: (i) preventing the disease from occurring in a subject that can be predisposed to the disease but has not yet been diagnosed as having it;

(ii) inhibiting the disease, i.e., arresting its development; or (iii) relieving the disease, i.e., causing regression of the disease.

As used herein, the terms "administering" and "administration" refer to any method of providing a composition to a subject. Such methods are well known to those skilled in the art and include, but are not limited to, intracardiac administration, oral administration, transdermal administration, administration by inhalation, nasal administration, topical administration, intravaginal administration, ophthalmic administration, intraaural administration, intracerebral administration, rectal administration, sublingual administration, buccal administration, and parenteral administration, including injectable such as intravenous administration, intra-arterial administration, intramuscular administration, and subcutaneous administration. Administration can be continuous or intermittent. In various aspects, a preparation can be administered therapeutically; that is, administered to treat an existing disease or condition. In further various aspects, a preparation can be administered prophylactically; that is, administered for prevention of a disease or condition.

The term "contacting" as used herein refers to bringing a disclosed composition or peptide or pharmaceutical preparation and a cell, target receptor, or other biological entity together in such a manner that the compound can affect the activity of the target (e.g., receptor, transcription factor, cell, etc.), either directly; i.e., by interacting with the target itself, or indirectly; i.e., by interacting with another molecule, co-factor, factor, or protein on which the activity of the target is dependent.

As used herein, the terms "effective amount" and "amount effective" refer to an amount that is sufficient to achieve the desired result or to have an effect on an undesired condition. For example, in an aspect, an effective amount of the polymeric nanoparticle is an amount that kills and/or inhibits the growth of cells without causing extraneous damage to surrounding non-cancerous cells. For example, a "therapeutically effective amount" refers to an amount that is sufficient to achieve the desired therapeutic result or to have an effect on undesired symptoms, but is generally insufficient to cause adverse side effects. The specific therapeutically effective dose level for any particular patient will depend upon a variety of factors including the disorder being treated and the severity of the disorder; the specific composition employed; the age, body weight, general health, sex and diet of the patient; the time of administration; the route of administration; the rate of excretion of the specific compound employed; the duration of the treatment; drugs used in combination or coincidental with the specific compound employed and like factors well known in the medical arts.

The term "pharmaceutically acceptable" describes a material that is not biologically or otherwise undesirable, *i.e.*, without causing an unacceptable level of undesirable biological effects or interacting in a deleterious manner. As used herein, the term "pharmaceutically acceptable carrier" refers to sterile aqueous or
5 nonaqueous solutions, dispersions, suspensions or emulsions, as well as sterile powders for reconstitution into sterile injectable solutions or dispersions just prior to use. Examples of suitable aqueous and nonaqueous carriers, diluents, solvents or vehicles include water, ethanol, polyols (such as glycerol, propylene glycol, polyethylene glycol and the like), carboxymethylcellulose and suitable mixtures
10 thereof, vegetable oils (such as olive oil) and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of coating materials such as lecithin, by the maintenance of the required particle size in the case of dispersions and by the use of surfactants. These compositions can also contain adjuvants such as preservatives, wetting agents, emulsifying agents and dispersing
15 agents. Prevention of the action of microorganisms can be ensured by the inclusion of various antibacterial and antifungal agents such as paraben, chlorobutanol, phenol, sorbic acid and the like. It can also be desirable to include isotonic agents such as sugars, sodium chloride and the like. Prolonged absorption of the injectable pharmaceutical form can be brought about by the inclusion of agents, such as
20 aluminum monostearate and gelatin, which delay absorption. Injectable depot forms are made by forming microcapsule matrices of the drug in biodegradable polymers such as polylactide-polyglycolide, poly(orthoesters) and poly(anhydrides). Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Depot injectable formulations are also
25 prepared by entrapping the drug in liposomes or microemulsions which are compatible with body tissues. The injectable formulations can be sterilized, for example, by filtration through a bacterial-retaining filter or by incorporating sterilizing agents in the form of sterile solid compositions which can be dissolved or dispersed in sterile water or other sterile injectable media just prior to use. Suitable inert carriers
30 can include sugars such as lactose. Desirably, at least 95% by weight of the particles of the active ingredient have an effective particle size in the range of 0.01 to 10 micrometers.

As used herein, the term "RNA silencing agent" refers to an RNA which is capable of preventing complete processing (e.g, the full translation and/or
35 expression) of a mRNA molecule through a post-transcriptional silencing mechanism. RNA silencing agents include small (<50 b.p.), noncoding RNA molecules, for

example RNA duplexes comprising paired strands, as well as precursor RNAs from which such small non-coding RNAs can be generated. Exemplary RNA silencing agents include siRNAs, miRNAs, and siRNA-like duplexes, as well as precursors thereof.

5 As used herein, the term “small interfering RNA” (“siRNA”) (also referred to in the art as “short interfering RNAs”) refers to an RNA (or RNA analog) comprising between about 10-50 nucleotides (or nucleotide analogs) which is capable of directing or mediating RNA interference. Preferably, an siRNA comprises between about 15-30 nucleotides or nucleotide analogs, more preferably between about 16-25
10 nucleotides (or nucleotide analogs), even more preferably between about 18-23 nucleotides (or nucleotide analogs), and even more preferably between about 19-22 nucleotides (or nucleotide analogs) (e.g., 19, 20, 21 or 22 nucleotides or nucleotide analogs).

As used herein, the term “microRNA” (“miRNA”), also referred to in the art as
15 “small temporal RNAs” (“stRNAs”), refers to a small (10-50 nucleotide) RNA which is capable of directed or mediating RNA silencing. A “natural miRNA” refers to a microRNA that occurs naturally. An “miRNA disorder” shall refer to a disease or disorder characterized by a aberrant expression or activity of a natural miRNA.

As used herein, the term “RNA interference” (“RNAi”) (also referred to in the
20 art as “gene silencing” and/or “target silencing”, e.g., “target mRNA silencing”) refers to a selective intracellular degradation of RNA. RNAi occurs in cells naturally to remove foreign RNAs (e.g., viral RNAs). Natural RNAi proceeds via fragments cleaved from free dsRNA which direct the degradative mechanism to other similar RNA sequences. As used herein, the term “translational repression” refers to a
25 selective inhibition of mRNA translation. Natural translational repression proceeds via miRNAs cleaved from shRNA precursors. Both RNAi and translational repression are mediated by RISC. Both RNAi and translational repression occur naturally or can be initiated by the hand of man, for example, to silence the expression of target genes.

As used herein, the term “antisense strand” of an siRNA or RNAi agent refers
30 to a strand that is substantially complementary to a section of about 10-50 nucleotides, e.g., about 15-30, 16-25, 18-23 or 19-22 nucleotides of the mRNA of the gene targeted for silencing. The antisense strand or first strand has sequence sufficiently complementary to the desired target mRNA sequence to direct target-specific RNA interference (RNAi), e.g., complementarity sufficient to trigger the
35 destruction of the desired target mRNA by the RNAi machinery or process (RNAi interference) or complementarity sufficient to trigger translational repression of the

desired target mRNA. The term "sense strand" or "second strand" of an siRNA or RNAi agent refers to a strand that is complementary to the antisense strand or first strand. Antisense and sense strands can also be referred to as first or second strands, the first or second strand having complementarity to the target sequence and the respective second or first strand having complementarity to said first or second strand. miRNA duplex intermediates or siRNA-like duplexes include a miRNA strand having sufficient complementarity to a section of about 10-50 nucleotides of the mRNA of the gene targeted for silencing and a miRNA* strand having sufficient complementarity to form a duplex with the miRNA strand.

As used herein, the term "guide strand" refers to a strand of an RNAi agent, e.g., an antisense strand of an siRNA duplex, that enters into the RISC complex and directs cleavage of the target mRNA.

The practice of the present invention will typically employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant nucleic acid (e.g., DNA) technology, immunology, and RNA interference (RNAi) which are within the skill of the art. Non-limiting descriptions of certain of these techniques are found in the following publications: Ausubel, F., et al., (eds.), *Current Protocols in Molecular Biology*, *Current Protocols in Immunology*, *Current Protocols in Protein Science*, and *Current Protocols in Cell Biology*, all John Wiley & Sons, N. Y., edition as of December 2008; Sambrook, Russell, and Sambrook, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001; Harlow, E. and Lane, D., *Antibodies - A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1988; Freshney, R.I., "Culture of Animal Cells, A Manual of Basic Technique", 5th ed., John Wiley & Sons, Hoboken, NJ, 2005. Non-limiting information regarding therapeutic agents and human diseases is found in Goodman and Gilman's *The Pharmacological Basis of Therapeutics*, 11th Ed., McGraw Hill, 2005, Katzung, B. (ed.) *Basic and Clinical Pharmacology*, McGraw-Hill/Appleton & Lange; 10th ed. (2006) or 11th edition (July 2009).

As used herein, the terms "inflammatory disease" and "inflammatory disorder" refer to diseases, disorders, and conditions characterized by abnormal inflammation. Examples include type 2 diabetes and cardiovascular diseases, such as atherosclerosis, stroke, and ischemic heart disease.

As used herein, the term "anti-inflammatory agent" refers to substance or treatment that reduce inflammation or swelling. For example, the anti-inflammatory agent can be an anti-IL6 drug, such as tocilizumab (Actemra), siltuximab (Sylvant),

Sarilumab, clazakizumab, olokizumab (CDP6038), elsilimomab, BMS-945429(ALD518), sirukumab (CNTO 136), CPSI-2364, ARGX-109, FE301, and FM101. The anti-inflammatory agent can be an anti-IL1 drug, such as canakinumab or rilonacept. The anti-inflammatory agent can be a nonsteroidal anti-inflammatory drug (NSAID). Examples of anti-inflammatory agents include Alclofenac, Alclometasone Dipropionate, Algestone Acetonide, alpha Amylase, Amcinafal, Amcinafide, Amfenac Sodium, Amiprilose Hydrochloride, Anakinra, Anirolac, Anitrazafen, Apazone, Balsalazide Disodium, Bendazac, Benoxaprofen, Benzydamine Hydrochloride, Bromelains, Broperamole, Budesonide, Carprofen, Cicloprofen, Cintazone, Cliprofen, Clobetasol Propionate, Clobetasone Butyrate, Clopirac, Cloticasone Propionate, Cormethasone Acetate, Cortodoxone, Decanoate, Deflazacort, Delatesteryl, Depo-Testosterone, Desonide, Desoximetasone, Dexamethasone Dipropionate, Diclofenac Potassium, Diclofenac Sodium, Diflorasone Diacetate, Diflumidone Sodium, Diflunisal, Difluprednate, Diftalone, Dimethyl Sulfoxide, Drocinnonide, Endrysone, Enlimomab, Enolicam Sodium, Epirizole, Etodolac, Etofenamate, Felbinac, Fenamole, Fenbufen, Fenclofenac, Fenclorac, Fendosal, Fempipalone, Fentiazac, Flazalone, Fluazacort, Flufenamic Acid, Flumizole, Flunisolide Acetate, Flunixin, Flunixin Meglumine, Fluocortin Butyl, Fluorometholone Acetate, Fluquazone, Flurbiprofen, Fluretofen, Fluticasone Propionate, Furaprofen, Furobufen, Halcinnonide, Halobetasol Propionate, Halopredone Acetate, Ibufenac, Ibuprofen, Ibuprofen Aluminum, Ibuprofen Piconol, Ilonidap, Indomethacin, Indomethacin Sodium, Indoprofen, Indoxole, Intrazole, Isoflupredone Acetate, Isoxepac, Isoxicam, Ketoprofen, Lofemizole Hydrochloride, Lomoxicam, Loteprednol Etabonate, Meclofenamate Sodium, Meclofenamic Acid, Meclorisona Dibutyrate, Mefenamic Acid, Mesalamine, Meseclazone, Mesterolone, Methandrostenolone, Methenolone, Methenolone Acetate, Methylprednisolone Suleptanate, Momiflumate, Nabumetone, Nandrolone, Naproxen, Naproxen Sodium, Naproxol, Nimazone, Olsalazine Sodium, Orgotein, Orpanoxin, Oxandrolane, Oxaprozin, Oxyphenbutazone, Oxymetholone, Paranyline Hydrochloride, Pentosan Polysulfate Sodium, Phenbutazone Sodium Glycerate, Pirfenidone, Piroxicam, Piroxicam Cinnamate, Piroxicam Olamine, Pirprofen, Prednazate, Prifelone, Prodolic Acid, Proquazone, Proxazole, Proxazole Citrate, Rimexolone, Romazarit, Salcolex, Salnacedin, Salsalate, Sanguinarium Chloride, Seclazone, Sermetacin, Stanazolol, Sudoxicam, Sulindac, Suprofen, Talmetacin, Talniflumate, Talosalate, Tebufelone, Tenidap, Tenidap Sodium, Tenoxicam, Tesicam, Tesimide, Testosterone, Testosterone

Blends, Tetrydamine, Tiopinac, Tixocortol Pivalate, Tolmetin, Tolmetin Sodium, Triclonide, Triflumidate, Zidometacin, Zomepirac Sodium.

Methods of Diagnosis

Gene Expression Assay

Methods of "determining gene expression levels" include methods that quantify levels of gene transcripts as well as methods that determine whether a gene of interest is expressed at all. A measured expression level may be expressed as any quantitative value, for example, a fold-change in expression, up or down, relative to a control gene or relative to the same gene in another sample, or a log ratio of expression, or any visual representation thereof, such as, for example, a "heatmap" where a color intensity is representative of the amount of gene expression detected. Exemplary methods for detecting the level of expression of a gene include, but are not limited to, Northern blotting, dot or slot blots, reporter gene matrix, nuclease protection, RT-PCR, microarray profiling, differential display, 2D gel electrophoresis, SELDI-TOF, ICAT, enzyme assay, antibody assay, and MNase-based detection methods. Optionally a gene whose level of expression is to be detected may be amplified, for example by methods that may include one or more of: polymerase chain reaction (PCR), strand displacement amplification (SDA), loop-mediated isothermal amplification (LAMP), rolling circle amplification (RCA), transcription-mediated amplification (TMA), self-sustained sequence replication (3SR), nucleic acid sequence based amplification (NASBA), or reverse transcription polymerase chain reaction (RT-PCR).

A number of suitable high throughput formats exist for evaluating expression patterns and profiles of the disclosed genes. Numerous technological platforms for performing high throughput expression analysis are known. Generally, such methods involve a logical or physical array of either the subject samples, the biomarkers, or both. Common array formats include both liquid and solid phase arrays. For example, assays employing liquid phase arrays, e.g., for hybridization of nucleic acids, binding of antibodies or other receptors to ligand, etc., can be performed in multiwell or microtiter plates. Microtiter plates with 96, 384 or 1536 wells are widely available, and even higher numbers of wells, e.g., 3456 and 9600 can be used. In general, the choice of microtiter plates is determined by the methods and equipment, e.g., robotic handling and loading systems, used for sample preparation and analysis. Exemplary systems include, e.g., xMAP® technology from Luminex (Austin, TX), the SECTOR® Imager with MULTI-ARRAY® and MULTI-SPOT® technologies from Meso Scale

Discovery (Gaithersburg, MD), the ORCA™ system from Beckman-Coulter, Inc. (Fullerton, Calif.) and the ZYMATE™ systems from Zymark Corporation (Hopkinton, MA), miRCURY LNA™ microRNA Arrays (Exiqon, Woburn, MA).

Alternatively, a variety of solid phase arrays can favorably be employed to determine expression patterns in the context of the disclosed methods, assays and kits. Exemplary formats include membrane or filter arrays (e.g., nitrocellulose, nylon), pin arrays, and bead arrays (e.g., in a liquid “slurry”). Typically, probes corresponding to nucleic acid or protein reagents that specifically interact with (e.g., hybridize to or bind to) an expression product corresponding to a member of the candidate library, are immobilized, for example by direct or indirect cross-linking, to the solid support. Essentially any solid support capable of withstanding the reagents and conditions necessary for performing the particular expression assay can be utilized. For example, functionalized glass, silicon, silicon dioxide, modified silicon, any of a variety of polymers, such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, polystyrene, polycarbonate, or combinations thereof can all serve as the substrate for a solid phase array.

In one embodiment, the array is a “chip” composed, e.g., of one of the above-specified materials. Polynucleotide probes, e.g., RNA or DNA, such as cDNA, synthetic oligonucleotides, and the like, or binding proteins such as antibodies or antigen-binding fragments or derivatives thereof, that specifically interact with expression products of individual components of the candidate library are affixed to the chip in a logically ordered manner, i.e., in an array. In addition, any molecule with a specific affinity for either the sense or anti-sense sequence of the marker nucleotide sequence (depending on the design of the sample labeling), can be fixed to the array surface without loss of specific affinity for the marker and can be obtained and produced for array production, for example, proteins that specifically recognize the specific nucleic acid sequence of the marker, ribozymes, peptide nucleic acids (PNA), or other chemicals or molecules with specific affinity.

Microarray expression may be detected by scanning the microarray with a variety of laser or CCD-based scanners, and extracting features with numerous software packages, for example, IMAGEGENE™ (Biodiscovery), Feature Extraction Software (Agilent), SCANLYZE™ (Stanford Univ., Stanford, CA.), GENEPIX™ (Axon Instruments).

Methylation Assay

There are many techniques for measuring DNA methylation. For example, one can use Methylation-Specific-Quantitative PCR (MS-QPCR) or to measure DNA

Methylation. (See: Eads C. A., MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res.* 2000 Apr. 15; 28(8):E32; 2. Darst R. P., Bisulfite sequencing of DNA. *Curr Protoc Mol Biol.* 2010 July; Chapter 7: Unit 7.9.1-17, and Cottrell S. E., et al., A real-time PCR assay for DNA-methylation using methylation specific blockers, *Nucleic Acids Res.* 2004; 32(1): e10.). In some embodiments, the methylation is quantified with PyroMark™MD Pyrosequencing System (Qiagen) using PyroPyroMark® Gold Q96 Reagents (Qiagen, Cat#972804). Other approaches for methylation quantification include, for example, methylation specific QPCR or quantitative bisulfite sequencing of methylation.

ARID5B Inhibitors

Disclosed herein are ARID5B inhibitors that reduce the level of endogenous ARID5B in a subject. The term "inhibitor" refers to a compound which is capable of reducing the expression of a gene or the activity of the product of such gene to an extent sufficient to achieve a desired biological or physiological effect. Therefore, the term "inhibitor" as used herein includes one or more of an oligonucleotide inhibitor, including siRNA, shRNA, miRNA and ribozymes.

Gene Silencing

As used herein, the phrase "RNA interference" (also called "RNAi" herein) refers to its meaning as is generally accepted in the art. The term generally refers to the biological process of inhibiting, decreasing, or down-regulating gene expression in a cell, and which is mediated by short interfering nucleic acid molecules (e.g., siRNAs, miRNAs, shRNAs). Additionally, the term "RNA interference" (or "RNAi") is meant to be equivalent to other terms used to describe sequence-specific RNA interference, such as post-transcriptional gene silencing, translational inhibition, transcriptional inhibition, or epigenetics. For example, nucleic acid agents of the invention can be used to epigenetically silence *ARID5B* at either the post-transcriptional level or the pre-transcriptional level. In a non-limiting example, epigenetic modulation of gene expression by nucleic acid agents of the invention can result from modification of chromatin structure or methylation patterns to alter gene expression. In another non-limiting example, modulation of gene expression by nucleic acid agents of the invention can result from cleavage of RNA (either coding or non-coding RNA) via RISC, or via translational inhibition, as is known in the art or modulation can result from transcriptional inhibition.

The terms "inhibit," "down-regulate," "reduce" or "knockdown" as used herein refer to their meanings as are generally accepted in the art. With reference to exemplary single-stranded RNAi molecules of the invention, the terms generally refer

to the reduction in the (i) expression of a gene or target sequence and/or the level of RNA molecules encoding one or more proteins or protein subunits, and/or (ii) the activity of one or more proteins or protein subunits, below that observed in the absence of the single-stranded RNAi molecules of the invention. Down-regulation can also be associated with post-transcriptional silencing, such as RNAi-mediated cleavage, or by alteration in DNA methylation patterns or DNA chromatin structure. Inhibition, down-regulation, reduction or knockdown with an RNAi agent can be in reference to an inactive molecule, an attenuated molecule, an RNAi agent with a scrambled sequence, or an RNAi agent with mismatches. The phrase "gene silencing" refers to a partial or complete loss-of-function through targeted inhibition of an endogenous target gene in a cell. As such, the term is used interchangeably with RNAi, "knockdown," "inhibition," "down-regulation," or "reduction" of expression of a target gene.

The nucleic acid sequence for *ARID5B* is known in the art, and it is within the skill of those in the art to design and produce gene silencing oligonucleotides such as siRNA using this sequence. For example, siRNA that silence *ARID5B* are commercially available from ThermoFisher Scientific (e.g. Catalog No. S38579, S38580, S38581).

Targeted Methylation

The disclosed methods can involve the targeted methylation of CpG sites within the *ARID5B* gene. Enzymes capable of modifying methylation status are known on the art and can be targeted to *ARID5B* gene to methylate the CpG sites. In some embodiments, targeted DNA methylation of the CpG sites within the *ARID5B* gene is accomplished using a gene editing tool such as a zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), mega-nuclease, CRISPR/Cas, structure-guided endonuclease (SGN), or targetron. All of them can achieve precise genetic modifications by inducing targeted DNA double-strand breaks (DSBs). Depending on the cell cycle stage, as well as the presence or absence of a repair template with homologous terminal regions, the DSB may then be repaired by either non-homologous end joining repair system (NHEJ) or homologous recombination-based double-strand break repair pathway (HDR).

In some embodiments, targeted DNA methylation of the CpG sites within the *ARID5B* gene is accomplished using a CRISPR system. Compositions and methods for making and using CRISPR-Cas systems are described in U.S. Pat. No. 8,697,359, which is incorporated herein in its entirety. In particular, the method can involve a catalytically inactive site specific nuclease fused to an effector domain

having methylation activity; and a guide sequence or a nucleic acid that encodes a guide sequence. In some aspects, the catalytically inactive site specific nuclease is a catalytically inactive Cas protein (e.g., a Cas9 protein or a Cpf1 protein). The guide sequences may be ribonucleic acid guide sequences. In certain aspects, the guide sequence is from about 10 base pairs to about 150 base pairs in length. The one or more guide sequences may comprise two or more guide sequences.

There are various ways that a polypeptide comprising a catalytically inactive site specific nuclease fused to an effector domain having methylation or demethylation activity can be delivered to a cell or subject, e.g., by administering a nucleic acid that encodes the polypeptide, which nucleic acid may be, e.g., a viral vector or may be a translatable nucleic acid (e.g., synthetic modified mRNA). Examples of modified mRNA are described in Warren et al. (Cell Stem Cell 7(5):618-30, 2010, Mandal PK, Rossi DJ. Nat Protoc. 2013 8(3):568-82, US Pat. Pub. No. 20120046346 and/or WO2011/130624.

In some aspects, one or more guide sequences include sequences that recognize DNA in a site-specific manner. For example, guide sequences can include guide ribonucleic acid (RNA) sequences utilized by a CRISPR system or sequences within a TALEN or zinc finger system that recognize DNA in a site-specific manner. The guide sequences comprise a portion that is complementary to a portion of each of the one or more genomic sequences and comprise a binding site for the catalytically inactive site specific nuclease. In some embodiments, the RNA sequence is referred to as guide RNA (gRNA) or single guide RNA (sgRNA).

In some aspects, a single RNA sequence can be complementary to one or more (e.g., all) of the genomic sequences that are being modulated or modified. In one aspect, a single RNA is complementary to a single target genomic sequence. In a particular aspect in which two or more target genomic sequences are to be modulated or modified, multiple (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, or more) RNA sequences are introduced wherein each RNA sequence is complementary to (specific for) one target genomic sequence. In some aspects, two or more, three or more, four or more, five or more, or six or more RNA sequences are complementary to (specific for) different parts of the same target sequence. In one aspect, two or more RNA sequences bind to different sequences of the same region of DNA. In some aspects, a single RNA sequence is complementary to at least two target or more (e.g., all) of the genomic sequences. It will also be apparent to those of skill in the art that the portion of the RNA sequence that is complementary to one or more of the genomic sequences and the portion of the RNA sequence that binds to the

catalytically inactive site specific nuclease can be introduced as a single sequence or as 2 (or more) separate sequences into a cell, zygote, embryo or nonhuman animal. In some embodiments the sequence that binds to the catalytically inactive site specific nuclease comprises a stem-loop.

5 In some embodiments, the RNA sequence used to modify gene expression is a modified RNA sequence comprising a 5-methylcytidine (5mC). The RNA sequence can vary in length from about 8 base pairs (bp) to about 200 bp. In some embodiments, the RNA sequence can be about 9 to about 190 bp; about 10 to about 150 bp; about 15 to about 120 bp; about 20 to about 100 bp; about 30 to about 90
10 bp; about 40 to about 80 bp; about 50 to about 70 bp in length.

The portion of each genomic sequence to which each RNA sequence is complementary can also vary in size. In particular aspects, the portion of each genomic sequence to which the RNA is complementary can be about 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33,
15 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 100 nucleotides (contiguous nucleotides) in length. In some embodiments, each RNA sequence can be at least about 70%, 75%, 80%, 85%, 90%, 95%, 100%), etc.
20 identical or similar to the portion of each genomic sequence. In some embodiments, each RNA sequence is completely or partially identical or similar to each genomic sequence. For example, each RNA sequence can differ from perfect complementarity to the portion of the genomic sequence by about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, etc. nucleotides. In some embodiments, one or more
25 RNA sequences are perfectly complementary (100%>) across at least about 10 to about 25 (e.g., about 20) nucleotides of the genomic sequence.

As described herein, the one or more RNA sequences also comprise a (one or more) binding site for a (one or more) catalytically inactive site specific nuclease. The catalytically inactive site specific nuclease may be a catalytically inactive
30 CRISPR associated (Cas) protein. In a particular aspect, upon hybridization of the one or more RNA sequences to the one or more genomic sequences, the catalytically inactive site specific nuclease binds to the one or more RNA sequences. In some aspects, the method further comprises administering one or more catalytically inactive Cas nucleic acid or protein.

35 A variety of CRISPR associated (Cas) genes or proteins which are known in the art can be used in the methods of the invention and the choice of Cas protein will

depend upon the particular conditions of the method. Specific examples of Cas proteins include Cas1, Cas2, Cas3, Cas4, Cas5, Cas5e (or CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9, CaslO, CaslOd, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (or CasA), Cse2 (or CasB), Cse3 (or CasE), Cse4 (or CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx1O, Csx16, CsaX, Csx3, Csz1, Csx15, Csf1, Csf2, Csf3, Csf4, Cpf1, and Cui966. In a particular aspect, the Cas nucleic acid or protein used in the methods is Cas9. In some embodiments a Cas protein, e.g., a Cas9 protein, may be from any of a variety of prokaryotic species. In some embodiments a particular Cas protein, e.g., a particular Cas9 protein, may be selected to recognize a particular protospacer-adjacent motif (PAM) sequence. In certain embodiments a Cas protein, e.g., a Cas9 protein, may be obtained from a bacteria or archaea or synthesized using known methods. In certain embodiments, a Cas protein may be from a gram positive bacteria or a gram negative bacteria. In some embodiments, the Cas protein is Cpf1 protein or a functional portion thereof. In some embodiments, the Cas protein is Cpf1 from any bacterial species or functional portion thereof.

Disclosed herein is a catalytically inactive Cas9 protein tethered to all or a portion of methyltransferase to create a chimeric protein that can be guided to specific DNA sites by one or more RNA sequences (sgRNA) to methylate *ARID5B* CpG sites. DNA methylation is established by two de novo DNA methyltransferases (Dnmt3a/b), and is maintained by Dnmt1.

As will be apparent to those of skill in the art, the method can further comprise introducing other molecules or factors into the cell to facilitate methylation of the genomic sequence. An agent that enhances DNA methylation may be an inhibitor of an endogenous DNA demethylase.

Sequence-specific nucleases have been developed to increase the efficiency of gene targeting or genome editing in animal and plant systems. Among them, zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) are the two most commonly used sequence-specific chimeric proteins. Once the ZFN or TALEN constructs are introduced into and expressed in cells, the programmable DNA binding domain can specifically bind to a corresponding sequence and guide the chimeric nuclease to make a specific DNA strand cleavage. A pair of ZFNs or TALENs can be introduced to generate double strand breaks (DSBs), which activate the DNA repair systems and significantly increase the frequency of both nonhomologous end joining (NHEJ) and homologous recombination (HR).

These and other methods for DNA methylation editing are disclosed in WO2017090724, WO2018035495, and WO2018053035, which are incorporated by reference in its entirety for these teachings.

Pharmaceutical Compositions

5 The compositions disclosed can be used therapeutically in combination with a pharmaceutically acceptable carrier. By "pharmaceutically acceptable" is meant a material that is not biologically or otherwise undesirable, i.e., the material may be administered to a subject, along with the nucleic acid or vector, without causing any undesirable biological effects or interacting in a deleterious manner with any of the
10 other components of the pharmaceutical composition in which it is contained. The carrier would naturally be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject, as would be well known to one of skill in the art.

 The disclosed compositions can be formulated in buffer solutions such as
15 phosphate buffered saline solutions, liposomes, micellar structures, and capsids. Formulations of DsiRNA agent with cationic lipids can be used to facilitate transfection of the DsiRNA agent into cells. For example, cationic lipids, such as lipofectin (U.S. Pat. No. 5,705,188), cationic glycerol derivatives, and polycationic molecules, such as polylysine (published PCT International Application WO
20 97/30731), can be used. Suitable lipids include Oligofectamine, Lipofectamine (Life Technologies), NC388 (Ribozyme Pharmaceuticals, Inc., Boulder, Colo.), or FuGene 6 (Roche) all of which can be used according to the manufacturer's instructions.

 Such compositions typically include the nucleic acid molecule and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically
25 acceptable carrier" includes saline, solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions.

 A pharmaceutical composition is formulated to be compatible with its intended
30 route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils,
35 polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as

ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL.TM. (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle, which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active

ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, e.g., gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compositions are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer. Such methods include those described in U.S. Pat. No. 6,468,798.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

The compounds can also be administered by a method suitable for administration of nucleic acid agents, such as a DNA vaccine. These methods include gene guns, bio injectors, skin patches, micro-particle DNA vaccine technology, transdermal needle-free vaccination, intranasal delivery, and microencapsulation.

In one embodiment, the disclosed compositions are prepared with carriers that will protect the nucleic acids against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Such formulations can be prepared using standard techniques. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

As defined herein, a therapeutically effective amount of a nucleic acid molecule (i.e., an effective dosage) depends on the nucleic acid selected. For instance, if a plasmid encoding an siRNA agent is selected, single dose amounts in the range of approximately 1 pg to 1000 mg may be administered; in some embodiments, 10, 30, 100, or 1000 pg, or 10, 30, 100, or 1000 ng, or 10, 30, 100, or 1000 µg, or 10, 30, 100, or 1000 mg may be administered. In some embodiments, 1-5 g of the compositions can be administered. The compositions can be administered one from one or more times per day to one or more times per week; including once every other day. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

The nucleic acid molecules of the invention can be inserted into expression constructs, e.g., viral vectors, retroviral vectors, expression cassettes, or plasmid viral vectors, e.g., using methods known in the art, including but not limited to those described in Xia et al., (2002), supra. Expression constructs can be delivered to a subject by, for example, inhalation, orally, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or by stereotactic injection (see e.g., Chen et al. (1994), Proc. Natl. Acad. Sci. USA, 91, 3054-3057). The pharmaceutical preparation of the delivery vector can include the vector in an acceptable diluent, or can comprise a slow release matrix in which the delivery vehicle is imbedded. Alternatively, where the complete delivery vector can be produced intact from recombinant cells, e.g.,

retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The expression constructs may be constructs suitable for use in the appropriate expression system and include, but are not limited to retroviral vectors, linear expression cassettes, plasmids and viral or virally-derived vectors, as known in the art. Such expression constructs may include one or more inducible promoters, RNA Pol III promoter systems such as U6 snRNA promoters or H1 RNA polymerase III promoters, or other promoters known in the art. The constructs can include one or both strands of the siRNA. Expression constructs expressing both strands can also include loop structures linking both strands, or each strand can be separately transcribed from separate promoters within the same construct.

Suitably formulated pharmaceutical compositions of this invention can be administered by means known in the art such as by parenteral routes, including intravenous, intramuscular, intraperitoneal, subcutaneous, transdermal, airway (aerosol), rectal, vaginal and topical (including buccal and sublingual) administration. In some embodiments, the pharmaceutical compositions are administered by intravenous or intraparenteral infusion or injection.

In general, a suitable dosage unit of nucleic acid agent will be in the range of 0.001 to 0.25 milligrams per kilogram body weight of the recipient per day, or in the range of 0.01 to 20 micrograms per kilogram body weight per day, or in the range of 0.01 to 10 micrograms per kilogram body weight per day, or in the range of 0.10 to 5 micrograms per kilogram body weight per day, or in the range of 0.1 to 2.5 micrograms per kilogram body weight per day. Pharmaceutical composition comprising the nucleic acid agent can be administered once daily. However, the therapeutic agent may also be dosed in dosage units containing two, three, four, five, six or more sub-doses administered at appropriate intervals throughout the day. In that case, the nucleic acid agent contained in each sub-dose must be correspondingly smaller in order to achieve the total daily dosage unit. The dosage unit can also be compounded for a single dose over several days, e.g., using a conventional sustained release formulation which provides sustained and consistent release of the nucleic acid agent over a several day period. Sustained release formulations are well known in the art. In this embodiment, the dosage unit contains a corresponding multiple of the daily dose. Regardless of the formulation, the pharmaceutical composition must contain nucleic acid agent in a quantity sufficient to inhibit expression of the target gene in the animal or human being treated. The

nucleic acid agent can be compounded in such a way that the sum of the multiple units of dsRNA together contain a sufficient dose.

Data can be obtained from cell culture assays and animal studies to formulate a suitable dosage range for humans. The dosage of compositions of the invention lies within a range of circulating concentrations that include the ED₅₀ (as determined by known methods) with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For a compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range of the compound that includes the IC₅₀ (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels of dsRNA in plasma may be measured by standard methods, for example, by high performance liquid chromatography.

The pharmaceutical compositions can be included in a kit, container, pack, or dispenser together with instructions for administration.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) inflammatory disease or disorders. Subjects at risk for the disease can be identified by, for example, one or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the detection of or the manifestation of symptoms characteristic of the disease or disorder, such that the disease or disorder is prevented or, alternatively, delayed in its progression. Another aspect of the invention pertains to methods of treating subjects therapeutically, i.e., altering the onset of symptoms of the disease or disorder.

With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (e.g., a patient's "drug response phenotype", or "drug response genotype"). Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment according

to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

5 A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.

10 EXAMPLES

Example 1: Blood Monocyte Transcriptome and Epigenome Analyses Reveal Loci Associated with Human Atherosclerosis.

Epigenomics and transcriptomics can illuminate the interplay between the genome and its environment and may provide insights into the molecular basis of complex diseases, including cardiovascular disease (CVD) (Lewis, J. & Bird, A. FEBS Lett. 285:155-159 (1991); Bird, A.P. & Wolffe, A.P. Cell 99:451-454 (1999); Jones, P.A. & Takai, D. Science 293:1068-1070 (2001); Egger, G., et al. Nature 429:457-463 (2004); Petronis, A. Nature 465:721-727 (2010)). Epigenetic targeting also is an attractive treatment strategy for reordering dysregulated gene expression. To date, epigenome-wide studies of CVD traits are limited (Reinius, L.E., et al. PLoS One. 7:e41361 (2012)) and their interpretation is potentially complicated by use of data from mixed cell types, which may obscure cell-type specific functional mechanisms. Monocytes and their derived macrophages are key players in inflammation that contribute to the development of many chronic diseases, including atherosclerosis (Lessner, S.M., et al. Am. J. Pathol. 160:2145-2155 (2002); Osterud, B. & Bjorklid, E. Physiol Rev. 83:1069-1112 (2003); Hansson, G.K. N. Engl. J. Med. 352:1685-1695 (2005); Ley, K., et al. Arterioscler. Thromb. Vasc. Biol. 31:1506-1516 (2011)).

There are blood monocyte epigenome and transcriptome signatures of several CVD risk factors, including age, obesity, and cigarette smoking (Reynolds, L.M. et al. Nat. Commun. 5:5366 (2014); Reynolds, L.M. et al. BMC Genomics 16:333 (2015); Ding, J. et al. Diabetes. 64:3464-3474 (2015); Reynolds, L.M. et al. Circ Cardiovasc. Genet 8:707-716 (2015)). In this Example, transcriptome and methylome features that are associated with atherosclerosis are identified. In addition to traditional CVD risk factors, we also integrate these findings with histone modification, DNase-seq, Hi-C (genome chromosome conformation capture), and

ChIA-PET (Chromatin Interaction Analysis by Paired-End Tag) sequencing data, and *in vitro* functional data in order to characterize novel molecular mechanisms of atherosclerosis. This Example discloses a role of ARID5B in atherogenesis, and an epigenetically controlled regulatory site of *ARID5B* expression.

Methods

Participants

The present analyses are primarily based on data collected at MESA exam 5 with concurrent analyses of purified monocyte samples of 1,264 randomly selected MESA participants from four MESA sites [John Hopkins University (JHU); Columbia University (CO); University of Minnesota (UMN); Wake Forest University (WFU)]. The study protocol was approved by Institutional Review Boards of the four institutions. All participants signed informed consent.

Measurement of Atherosclerosis and Prevalent CVD

All measurements in humans were obtained at MESA exam 5 unless otherwise specified. To obtain carotid artery plaque scores, readers at the UW Atherosclerosis Imaging Research Program Laboratory adjudicated carotid plaque presence or absence, defined as a focal abnormal wall thickness (carotid IMT >1.5 mm) or a focal thickening of >50% of the surrounding IMT, as reported previously (Stein, J.H. et al. J Am Soc Echocardiogr. 21:93-111 (2008); Touboul, P.J. et al. Cerebrovasc. Dis. 23:75-80 (2007)). Presence or absence of plaque acoustic shadowing was recorded. A total plaque score (range 0-12) was calculated to describe carotid plaque burden. One point per plaque was allocated for the near and far walls of each segment (CCA, bulb, ICA) of each carotid artery that was interrogated. For carotid plaque presence, intra-reader reproducibility was excellent (per reader Kappa = 0.82-1.0, overall Kappa = 0.83, 95% confidence interval [CI] 0.70-0.96), as was inter-reader reproducibility (Kappa = 0.89; 95% CI 0.72 - 1.00) (Tattersall, M.C. et al. Stroke 45:3257-3262 (2014)). The CT Reading Center for cardiac scans in the MESA is at UCLA-Biomedical Research Institute. Coronary artery calcium (CAC) score was determined using the Agatston method, which accounts for both lesion area and calcium density using Hounsfield brightness. The re-read agreement for the CAC score (intraclass correlation coefficient = 0.99) was excellent. Prevalent CVD was defined as a past history of myocardial infarction, angina (which included definite angina and probable angina if coronary revascularization was performed at the same time or afterwards), resuscitated cardiac arrest, or stroke.

Measurement of Covariates

Weight was measured with a Detecto Platform Balance Scale to the nearest 0.5 kg. Height was measured with a stadiometer (Accu-Hite Measure Device with level bubble) to the nearest 0.1 cm. Body-mass index (BMI) was defined as weight in kilograms divided by square of height in meters. Type II diabetes mellitus (T2D) was defined as fasting glucose ≥ 7.0 mmol L⁻¹ (≥ 126 mg/dL) or use of hypoglycemic medication, and impaired fasting glucose was defined as fasting glucose 5.6-6.9 mmol L⁻¹ (100-125 mg dL⁻¹). Plasma IL-6 was measured by ultra-sensitive ELISA (Quantikine HS Human IL-6 Immunoassay; R&D Systems, Minneapolis, MN). Plasma C-reactive protein (CRP) was measured using the BNII nephelometer (High Sensitivity CRP; Dade Behring Inc., Deerfield, IL). Resting blood pressure was measured three times in the seated position using a Dinamap model Pro 100 automated oscillometric sphygmomanometer (Critikon, Tampa, FL). The average of the last two measurements was used in analysis. Hypertension was defined as systolic pressure greater than or equal to 140 mm Hg, diastolic pressure greater than or equal to 90 mm Hg or current use of anti-hypertensive medication. The single nucleotide polymorphisms (SNPs) data were derived from MESA Affymetrix 6.0 array genotype data (Fox, C.S. et al. PLoS. Genet. 8:e1002705 (2012)).

Blood Cell Count and Purification of CD14+ and CD4+ cells

Samples for complete blood count (CBC) with differential analysis were obtained by venipuncture and collected into tubes containing ethylenediaminetetraacetic acid (EDTA). Total circulating white blood cell count and cell subtype counts were performed at local LabCorp. Blood was also collected in sodium heparin-containing Vacutainer CPT™ cell separation tubes (Becton Dickinson, Rutherford, NJ) to separate peripheral blood mononuclear cells from other elements within two hours from blood draw. Subsequently, monocytes and T cells were isolated with anti-CD14 and anti-CD4 monoclonal antibody coated magnetic beads, respectively, using an autoMACS automated magnetic separation unit (Miltenyi Biotec, Bergisch Gladbach, Germany). Initially flow cytometry analysis of 18 specimens was performed, including samples from all four MESA field centers, which were found to be consistently > 90% pure.

DNA/RNA Extraction

DNA and RNA were isolated from samples simultaneously using the AllPrep DNA/RNA Mini Kit (Qiagen, Inc., Hilden, Germany). DNA and RNA QC metrics included optical density (OD) measurements, using a NanoDrop spectrophotometer and evaluation of the integrity of 18 s and 28 s ribosomal RNA using the Agilent 2100 Bioanalyzer with RNA 6000 Nano chips (Agilent Technology, Inc., Santa Clara, CA)

following manufacturer's instructions. RNA with RIN (RNA Integrity) scores > 9.0 was used for global expression microarrays. The median of RIN for our 1,264 samples was 9.9.

Global mRNA Expression Quantification

5 The Illumina HumanHT-12 v4 Expression BeadChip and Illumina Bead Array Reader were used to perform the genome-wide expression analysis, as previously described (Liu, Y. et al. Hum Mol Genet 22:5065-5074 (2013)). This data has been deposited in the NCBI Gene Expression Omnibus and is accessible through GEO Series accession number (GSE56047).

Epigenome-wide Methylation Quantification

10 Illumina HumanMethylation450 BeadChips and HiScan reader were used to perform the epigenome-wide methylation analysis, as previously described (Liu, Y. et al. Hum Mol Genet 22:5065-5074 (2013)). This methylation data has been deposited in the NCBI Gene Expression Omnibus and is accessible through GEO Series
15 accession number (GSE56046).

Quality Control and Pre-Processing of Microarray Data

Data pre-processing and quality control (QC) analyses were performed in R using *Bioconductor* packages, as previously described (Liu, Y. et al. Hum Mol Genet 22:5065-5074 (2013)). For both monocyte and T cell assays, we included 2% blind
20 duplicates. Correlations among technical replicates exceeded 0.997. Multidimensional scaling plots showed the five common control samples were highly clustered together and identified three outlier samples, which were excluded subsequently.

The Illumina HumanHT-12 v4 Expression BeadChip included >47,000 probes
25 for >30,000 genes (with unique Entrez gene IDs). Statistical analyses excluded probes with non-detectable expression in $\geq 90\%$ of MESA samples (using a detection p-value cut-off of 0.0001), probes overlapping repetitive elements or regions, probes with low variance across the samples (<10th percentile), or probes targeting putative and/or not well-characterized genes, i.e. gene names starting with KIAA, FLJ, HS,
30 MGC, or LOC, 14,619 mRNA transcripts from 10,989 unique genes were included in the analyses of the presented manuscript.

The Illumina HumanMethylation450 BeadChip included probes for 485K CpG sites. Exclusion criteria included probes with "detected" methylation levels in <90% of MESA samples using a detection p-value cut-off of 0.05 and 65 control probes
35 which assay highly-polymorphic single nucleotide polymorphisms (SNPs) rather than

DNA methylation (Pidsley, R. et al. BMC Genomics 14:293 (2013)). Methylation data for the total of 484,817 CpG sites were included in the analyses.

To estimate residual sample contamination, we generated separate enrichment scores for neutrophils, B cells, T cells, and natural killer cells as described previously (Liu, Y. et al. Hum Mol Genet 22:5065-5074 (2013)). Residual sample contamination was adjusted for with non-monocyte cell types in all the analyses. Although most of monocytes (80-90%) are expected to be CD14+CD16- (Wong, K.L. et al. Blood 118:e16-e31 (2011)), we also adjusted for expression of the *FCGR3a* gene (*CD16a*).

RNA sequencing

A subset of 374 samples was randomly selected from the 1,264 MESA monocyte samples for RNA sequencing. Total RNA samples were enriched for mRNA, by depleting rRNA using the MICROBExpress kit from Ambion and following the manufacturer's instructions. Poly(A) mRNA was enriched, and Illumina compatible, strand-specific libraries were constructed using Illumina's TruSeq Stranded mRNA HT Sample Prep Kit (Illumina, RS-122-2103). 1 µg of total RNA with RIN ≥ 8.0 was converted into a library of stranded template molecules suitable for subsequent cluster generation and sequencing by Illumina HiSeq. The libraries generated were validated using Agilent 2100 Bioanalyzer and quantitated using Quant-iT dsDNA HS Kit (Invitrogen) and qPCR. Six individually indexed cDNA libraries were pooled and sequenced on Illumina HiSeq, resulting in an average of close to 30 million reads per sample. Libraries were clustered onto flow cells using Illumina's TruSeq PE Cluster Kit v3 (PE-401-3001) and sequenced 2X100 cycles using TruSeq SBS Kit -HS (FC-401-3001) on an Illumina HiSeq™ 2500. A total of 64 lanes were run to generate approximately 30 million 2 x 101 Paired End reads per sample. The Illumina HiSeq Control Software (HCS v2.0.12) with Real Time Analysis (RTA v1.3.61) was used to provide the management and execution of the HiSeq 2500.

Illumina sequencing runs were processed to de-multiplex samples and generate FastQ files using the Illumina provided configureBclToFastq.pl script to automate running CASAVA 1.8.4 using default parameters for removal of sequencing reads failing the chastity filter (yes) and mismatches in the barcode read (0). Following generation of FastQ files, reads were trimmed to remove poor quality reads (or read tails) using Btrim (5 base sliding window average with Q > 15) (Kong, Y. Genomics 98:152-153 (2011)) and then trimmed to remove any adaptor sequence present in the reads using custom perl scripts (trim sequences containing 11 base

tag of adaptor, final length >40 bases). The Ensembl GRCh37 Homo Sapiens reference file, annotations and Bowtie2 indexes were downloaded from the igenomes website for mapping of the sequencing reads to the genome and read counting. Bowtie2 (2.1.0) and TopHat2 (2.0.8) were used to map the sequencing reads to the genome using a mate-inner-distance of 100 bp and 'firststrand' options (Langmead, B., et al. *Genome Biol.* 10, R25 (2009); Trapnell, C., et al. *Bioinformatics.* 25:1105-1111 (2009)). Following alignment, bam files were merged using the samtools (0.1.19) merge function 4, and read counts per gene were obtained using HTSeq (0.5.4p3). The 'intersection-strict' overlap resolution mode and 'stranded reverse' options were used in HTSeq.

Data pre-processing and QC analyses were performed in R using Bioconductor packages. The transcript-based raw count data files for each sample from TopHat2 were combined into a count matrix with 56,303 features (rows) and 374 MESA samples (columns). The median total count per sample was 28.8 million. Reads denoted by TopHat2 as "no_feature", "ambiguous", "too_low_aQual", "not_aligned", "alignment_not_unique" were removed. Counts were converted to Counts Per Million (CPM) using the cpm function of the edgeR package (Robinson, M.D., et al. *Bioinformatics.* 26:139-140 (2010)), and all features with CPM ≤ 0.25 in $\geq 90\%$ of the 374 MESA samples were removed. Features assigned to the mitochondrial genome were removed as well. Using the biomaRt package and querying the Ensembl BioMart database, Entrez Gene IDs, Gene Symbols, genome coordinates, gene length and percent GC content were obtained for 12,585 features which had a corresponding Entrez ID or Illumina HumanHT-12 v4 probe ID. To be able to continue to use the flexible and computationally efficient linear modeling functions in R, we transformed the raw count data to log counts per million ($y = \log\text{CPM}$):

$$y_{gs} = \log_2 \left(\frac{c_{gs} + 0.5}{T_s + 1} \cdot 10^6 \right)$$

where c_{gs} is the raw count of gene transcript g in sample s , and T_s is the normalized total count of sample s , using the Trimmed Mean of M-values (TMM) normalization method (Robinson, M.D. & Oshlack, A. *Genome Biol.* 11:R25 (2010)) as

implemented in the calcNormFactors function in the edgeR Bioconductor package (Robinson, M.D., et al. *Bioinformatics.* 26:139-140 (2010)). Either only this TMM normalization was performed, or quantile normalization (QN) was applied to the logCPM values. Because the logCPM values' variance tends to decrease with increasing count for smaller counts, we used the voom function of the limma package to estimate the mean-variance trend non-parametrically and to predict the residual

variance of each individual observation for each gene. Then the inverse residual variances were incorporated into the linear modeling (lm) as weights in a standard manner. The same low variance filter that was used for the microarray data was imposed for the logCPM data, removing another 192 features with the lowest variance and retaining 12,380 features for analysis. Weighted linear model analyses was then performed with the otherwise exact same models as for the microarray data.

Bisulfite treatment of genomic DNA and pyrosequencing

A subset of 90 samples was selected from the 1,264 MESA monocyte samples, based on carotid plaque score extremes, matched for age, sex, and race. Genomic DNA from monocytes was bisulfite modified using the EZ DNA Methylation™ Kit (Zymo Research Co., Irvine, CA) according to the manufacturer's protocol for the Infinium methylation assay. Primers (F1-Biotinated 5'-GAAAATAGGAAATGTTTAATTGTG-3' (SEQ ID NO:2), R1 5'-TCACTATACTCCTAACAACCAACC-3' (SEQ ID NO:3)) for pyrosequencing assays of ARID5B-cg25953130 were designed using PSQ assay design software version 1.0.6 (Biotage, Uppsala, Sweden). PCR was performed with the PyroMark PCR Master Mix (Qiagen Inc.). Pyrosequencing was conducted using PyroMark Gold Q98 Reagents (Qiagen Inc.). Methylation values for each CpG site were calculated using Pyro Q-CpG software 1.0.9 (Biotage).

Weighted Gene Co-Expression Network Analyses

The Weighted Gene Co-Expression Network Analysis (WGCNA) method was used to construct network modules of highly correlated transcripts, using the R package WGCNA (Langfelder, P. & Horvath, S. BMC. Bioinformatics. 9:559 (2008)). A total of 1,261 MESA samples were included in the co-expression network analysis after removing three outlier samples based on hierarchical clustering, and 13,196 mRNA transcripts were included. First, an unsigned weighted network was constructed based on the pairwise correlations among all transcripts considered, using a soft thresholding power of 5, chosen to produce approximately a scale-free topology. Then, using the topological overlap measure to estimate the network interconnectedness, the transcripts were hierarchically clustered. The default parameters of WGCNA were used, except for changing the correlation type from Pearson to biweight midcorrelation (which is more robust to outliers), the deepSplit setting from 2 to 3, the detectCutHeight value from 0.995 to 0.999, the maximum block size from 5,000 to 14,000 transcripts, and the minimum size for module detection from 20 to 10.

WGCNA produces a set of modules, each containing a unique set of mRNA transcripts. The module eigengene was obtained to represent each module, which corresponds to the first eigenvector of the within-module expression correlation matrix (or the first right-singular vector of the standardized within-module expression matrix).

Association Analyses

The overall goal of the association analysis was to characterize the associations of each measure of atherosclerosis (carotid plaque and CAC) with each of genome-wide measures of mRNA expression and DNA methylation. Analyses were performed using the linear model (*lm*) function of the *Stats* package and the *stepA/C* function of the *MASS* package in *R*. Reported correlations (*r*) represent the partial Pearson product-moment correlation coefficient. Separate linear regression models were fit for each measure of atherosclerosis, with 1) genome-wide ($\log_2$ transformed) mRNA expression profiles, 2) network module eigengenes from WGCNA, and 3) genome-wide DNA methylation profiles (M-values). The M-value is well suited for high-level analyses and can be transformed into the beta-value, an estimate of the percent methylation of an individual CpG site that ranges from 0 to 1 (M is $\logit(\text{beta-value})$). P-values were adjusted for multiple testing using the q-value FDR method (Storey, J.D. & Tibshirani, R. *Proc. Natl. Acad. Sci. U. S. A* 100:9440-9445 (2003)) and Benjamini Hochberg FDR (Benjamini, Y. & Hochberg, Y. *Journal of the Royal Statistical Society B* 57:289-300 (2006)) when applicable. All analyses accounted for effects of residual cell contamination and covariates including age, gender, ethnicity, and study site. The full model also adjusted for traditional CVD risk factors (cigarette smoking, BMI, HDL-C and LDL-C levels, hypertension, diabetes mellitus) and statin use.

Mediation analysis

To investigate the genomic features as a potential molecular link between CVD risk factors and extent of atherosclerosis, mediation analyses was performed under an assumed causal model in which a CVD risk factor leads to a change in the genomic feature, which at least partially mediates the effects of the CVD risk factor on atherosclerotic burden. The mediation analyses accounted for the biological and technical covariates and were performed by robust Structural Equation Modeling (SEM) as implemented in the *R* package *lavaan* (Rosseel, Y. *Journal of Statistical Software* 48:1-36 (2012)). SEM analysis in general, and as implemented in *lavaan*, is based on Maximum Likelihood and the normal distribution, but provides several approaches to effectively deal with non-normal data. A first approach consists of computing robust standard errors (SE) by sandwich-type covariance matrices and

scaled test statistics, in particular the Satorra-Bentler statistic whose amount of rescaling reflects the degree of kurtosis, while the second approach uses specific bootstrapping methods to obtain both SE and test statistics (Rosseel, Y. Journal of Statistical Software 48:1-36 (2012)). The reported results are based on bootstrapping which were found to be somewhat more conservative than the use of robust SE and the Satorra-Bentler statistic.

In vivo Functional Annotation Analysis

CpGs with methylation significantly associated with carotid plaque scores or CAC were investigated for association with cis-gene expression by performing a look-up in the results from our previous analysis (Liu, Y. et al. Hum Mol Genet 22:5065-5074 (2013)) of the same samples. Briefly, to identify DNA methylation associated with gene expression in *cis*, separate linear regression models were fit with the M-value for each CpG site (adjusted for methylation chip and position effects) as a predictor of transcript expression for any autosomal gene within 1 Mb of the CpG in question. Covariates were age, sex, and race/ethnicity, and study site.

mRNA expression and DNA methylation of the most significantly associated mRNA and CpGs associated with carotid plaques scores and CAC were also investigated for association with nearby genetic variants. A large window (± 1 Mb) surrounding the mRNA/CpG was investigated to avoid missing any potential effects. Covariates were age, sex, and race/ethnicity, and study site. Separate linear regression models were fit with single nucleotide polymorphisms (SNPs) located within 1 Mb as a predictor of the mRNA expression ($\log_2$ transformed) or the methylation (M-value) in the MESA samples from Caucasian participants, including SNPs with a minor allele frequency > 0.05 in the MESA Caucasian population. P-values were adjusted for multiple testing using the q-value FDR method (Storey, J.D. & Tibshirani, R. Proc. Natl. Acad. Sci. U. S. A 100:9440-9445 (2003)).

In silico Functional Annotation Analysis

For the differentially expressed co-expression modules, Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City) was used to examine the enrichment of canonical pathways and bio-functions. In silico functional prediction of chromatin states in monocytes was performed using ChromHMM (Ernst, J. & Kellis, M. Nat Methods 9:215-216 (2012)) to predict segmentation among six states, based on histone modifications in monocyte samples from the BLUEPRINTproject (Adams, D. et al. Nat Biotechnol 30:224-226 (2012); Saeed, S. et al. Science 345:1251086 (2014)) (H3K27ac, H3K4me1, H3K4me3) and ENCODE (Rosenbloom, K.R. et al. Nucleic Acids Res. 41:D56-D63 (2013)) (H3K36me3). Other functional information

utilized includes DNase hypersensitive hotspot data in a monocyte sample (C001UY46) from the BLUEPRINT project (Adams, D. et al. Nat Biotechnol 30:224-226 (2012); Saeed, S. et al. Science 345:1251086 (2014)), and transcription factor binding sites detected in any cell type from ENCODE (Rosenbloom, K.R. et al. Nucleic Acids Res. 41:D56-D63 (2013)). Data was accessed from the UCSC Genome Browser (Karolchik, D. et al. Nucleic Acids Res (2013)) and the Gene Expression Omnibus.

Hi-C data in a B-cell line (GM12878) (Rao, S.S. et al. Cell 159:1665-1680 (2014); Dixon, J.R. et al. Nature 485:376-380 (2012); Heidari, N. et al. Genome Res 24:1905-1917 (2014); Dixon, J.R. et al. Nature 518:331-336 (2015)), including the contact matrix heatmap and virtual 4C results were adapted from the YUE lab Hi-C Interactions and Virtual 4C website.

Functional Evaluation of ARID5B Using in vitro Models

To examine the functional role of ARID5B, siRNA-mediated *ARID5B* knockdown was used in LPS-stimulated human THP1-monocytes. Using the protocols as described above, transcriptomic profiling was performed using the Illumina HumanHT-12 v4 Expression BeadChip, and *ARID5B* mRNA expression measured by qPCR.

Human monocytic THP-1 cell line was purchased from the American Tissue Culture Collection (ATCC). Cells were maintained in complete RPMI 1640 medium (Invitrogen) supplemented with 100 units ml⁻¹ of penicillin, 100 ug ml⁻¹ of streptomycin, 2 mM L-glutamine, and 10% fetal bovine serum (FBS, HyClone, Logan, UT) in a humidified incubator with 5% CO₂ at 37 °C.

Two target-specific siRNAs for ARID5B exon 8&9 (life technologies, siRNA ID: s38579, designated as siARID5B1) and for exon 6 (siRNA ID: s38580, designated as siARID5B 2) were used initially to compare for the knockdown efficiency. Then, the siARID5B 1 (s38579) was chosen for the remaining *in vitro* experiments. SiRNA transfection was performed using THP1 monocytes that were split 24h prior to transfection. 10nM of siARID5B were electronically transfected into THP-1 cells for 24 h and 48h using Amaxa® Human Monocyte Nucleofector Kit and an Amaxa nucleofector II device (Lonza, Inc.). Scrambled siRNAs were transfected as negative controls. Samples were incubated for 24 h, followed by 3 h of LPS (LPS: 100 ng ml⁻¹) stimulation.

Levels of cytokines in human monocytic THP-1 cells presented in Fig. 10 were measured by quantitative real-time RT-PCR using gene-specific TaqMan probe

sets in an ABI prism 7000 sequence detection system (life technologies). GAPDH mRNA was the internal loading control.

To reduce false positive rates, two independent experiments (4 siARID5B vs. 4 scrambled siRNA samples per experiment) were performed. To detect differential expression between two groups with small sample sizes, the regularized t-test implemented in the *limma* R package was used (Smyth, G.K. Stat. Appl. Genet. Mol. Biol. 3:Article3 (2004)). For the differentially expressed genes, Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City) was used to examine the enrichment of canonical pathways and bio-functions.

To further interrogate the potential pro-inflammatory effect that ARID5B may play, the third siARID5B knockdown experiment was repeated as described above and ELISA assays of culture media performed for IL-1a, the pro-inflammatory cytokine that was most significantly reduced by ARID5B mRNA Knockdown. Supernatants collected from LPS stimulated THP1 cells were analyzed for human IL-1a production using commercial sandwich ELISA kit (R&D Systems) according to the manufacturer's instructions. The results are representative of three or more experiments performed in triplicated as means + SEM. Levels of IL-1a mRNA expression presented in Figure 5B were measured by RT-PCR as described previously.

To further examine the effects of ARID5B knockdown on cellular functions suggested by the transcriptomic profile changes, the siARID5B knockdown experiment was repeated as described above and THP1-monocyte migration and phagocytosis assays were performed. Cell migration was evaluated using Transwell inserts (6.5 mm diameter) with polycarbonate filters (5-µm pore size, Corning Costar) in 24 well plates. THP1 cells (30,000 cells per well) were added to the upper chamber of the insert. The lower chamber contained 600 µl of RPMI 1640 medium/1% FBS with chemokine, MCP-1 (40 ng/ml). The plates were incubated at 37°C in 5% CO₂ for 24 h and cells that had migrated into the lower chamber were counted using cell countess. Cell phagocytosis was evaluated using yellow-green carboxylate modified microspheres (Thermo Fisher Scientific) in 24-well plates. Cells (500,000 cells per well) were incubated with 2 µl of beads (beads:cells = 5:1) for 60 minutes. After thoroughly washing 10 times with cold PBS, % engulfed beads in cells were calculated using FACs.

Data Availability

Genome-wide gene expression and DNA methylation data has been deposited in the NCBI Gene Expression Omnibus (GEO) and is accessible through GEO Series accession number GSE56047 and GSE56046, respectively.

Results

Clinical Data and Sample Characteristic

The Multi-Ethnic Study of Atherosclerosis (MESA) is a multi-site, longitudinal study designed to investigate the prevalence, correlates, and progression of subclinical cardiovascular disease in a population cohort of 6,814 participants. Since its inception in 2000, five clinic visits have collected extensive clinical, socio-demographic, lifestyle, behavior, laboratory, nutrition, and medication data (Bild, D.E. et al. Am. J. Epidemiol. 156:871-881 (2002)). At Exam 5, carotid ultrasound and computed tomography (CT), were used to quantify carotid plaque burden (carotid plaque score, range 0-12) and coronary artery calcification (CAC Agatston score), respectively. These two measures of atherosclerosis burden independently predict future CVD events in MESA and other cohorts (Tattersall, M.C. et al. Stroke 45:3257-3262 (2014); Gepner, A.D. et al. Circ. Cardiovasc. Imaging 8 (2015); Plichart, M. et al. Atherosclerosis 219:917-924 (2011); Stein, J.H., et al. J. Am. Soc. Echocardiogr. 21:93-111 (2008)) (distribution of scores presented in Fig. 6). Table 1 shows characteristics of the MESA participants at Exam 5, overall and stratified by study site (study site 1: N=709, site 2: N=499), including demographics (age, sex, race/ethnicity, and study site), traditional CVD risk factors [cigarette smoking, body-mass index (BMI), high- and low- density lipoprotein cholesterol (HDL-C and LDL-C), hypertension, type II diabetes mellitus (T2DM)], atherosclerosis burden measures, prevalent CVD, and statin use. The transcriptome and methylome of monocytes purified at MESA Exam 5 were profiled concurrently. In a separate effort to evaluate reproducibility of single-time measures, there was a high consistency of repeated Illumina microarray data over five months (Fig. 7), suggesting both RNA expression and DNA methylation at most loci can be stable over time in an individual.

Table 1. Population Characteristics at MESA Exam 5*			
		By Site	
	Overall (N=1,208)	Study 1 (JHU+CO) (N=709)	Study 2 (UMN+WFU) (N=499)
Demographic			
Age (years)	70 ± 9	70 ± 9	69 ± 9
Sex (% female)	51.4	55.0	46.3
Race/ethnicity			

African-American (%)	21.5	36.1	0.8
Hispanic (%)	32.6	29.3	37.3
Caucasian (%)	45.9	34.6	61.9
Cigarette smoking Former (%)	51.7	49.1	55.5
Current (%)	9.1	8.1	10.5
Clinical and laboratory			
BMI, (kg m ⁻²)	30 ± 6	30 ± 6	30 ± 5
LDL-C (mg dl ⁻¹)	105 ± 32	107 ± 33	102 ± 30
HDL-C (mg dl ⁻¹)	54 ± 16	57 ± 17	50 ± 13
Impaired Glucose Tolerance (%)	45.6	44.3	47.5
Diabetes (%)	23.0	25.2	19.8
Hypertension (%)	60.9	63.8	56.6
Carotid plaque score, median (IQR)	2 (0 - 4)	2 (0 - 3)	2 (0 - 4)
Coronary artery calcium score, median (IQR)	46 (0 - 305)	32 (0 - 227)	76 (0 - 397)
Prevalent CVD (%)	5.3	4.2	6.8
Statin use (%)	38.5	35.7	42.5
*Plus-minus values are means ± SD John Hopkins University (JHU); Columbia University (CO); University of Minnesota (UMN); Wake Forest University (WFU); interquartile range (IQR)			

Blood Leukocyte Count and Atherosclerosis

Total white blood cell (WBC) count and its constituent subtypes were measured for all samples prior to monocyte purification, which provided the absolute monocyte count. Higher monocyte count, but not monocyte percentage or other leukocyte count, was marginally associated with carotid plaque score and CAC [natural log (carotid plaque score/CAC + 1), p=0.026 with 0.4% of variability explained, and p=0.079, respectively], in agreement with previous reports (Madjid, M., et al. J Am Coll Cardiol. 44:1945-1956 (2004)). Positive immunoselection (magnetic beads) (Lyons, P.A. et al. Genomics 8:64 (2007)) were used to produce samples of monocytes with greater than 90% purity. Residual contamination with neutrophils, B cells, T cells, and natural killer cells was estimated as previously reported (Liu, Y. et al. Hum Mol Genet 22:5065-5074 (2013)). Likewise, the generally small percentage of CD14⁺ monocytes that are also CD16⁺ (~10%) (Wong, K.L. et al. Blood 118:e16-e31 (2011)) was estimated based on expression of *FCGR3a* (*CD16a*). Neither the surrogates of residual cell contamination nor the % of CD14⁺CD16⁺ cells was associated with the measures of atherosclerosis.

Transcriptome Signature of Atherosclerosis

Transcriptomic studies of atherosclerosis in mouse models and in humans (Berisha, S.Z., et al. PLoS. One. 8:e65003 (2013); Maiwald, S., et al. Curr. Opin. Clin. Nutr. Metab Care 16:411-417 (2013); Sivapalaratnam, S. et al. PLoS. One. 7:e32166 (2012)) have been reported. However, no consensus atherosclerosis biomarkers or pathways have been identified. Of the 10,989 unique genes with RNA expression detectable in monocytes, genes were identified with expression associated with carotid plaque score (n=2) and CAC (n=13) at a q-value-based false discovery rate (FDR) (Storey, J.D. & Tibshirani, R. Proc. Natl. Acad. Sci. U. S. A 100:9440-9445 (2003)) of 0.05 after adjusting for demographics (Figure 1A and 1B). Using the FDR level of 0.05 for the genome-wide search, signals with small effect size would inevitably be missed due to the limited statistical power for the sample size we have, thus additional signals for FDR of 0.20 were shown in Table 2 (21 gene transcripts with carotid plaque, and 104 gene transcripts with CAC). Expression of two genes, *ARID5B* and *PDLIM7* (*PDZ and LIM domain protein 7*), were positively associated with both measures of atherosclerosis (FDR≤0.05); *ARID5B* was most significantly associated with carotid plaque score ($p=6.30 \times 10^{-8}$, FDR= 1.08×10^{-3} ; with CAC: $p=2.47 \times 10^{-5}$, FDR=0.03). *ARID5B* plays a key metabolic role in adipose, liver and smooth muscle, and was previously implicated in lipid metabolism and adipogenesis in mice (Whitson, R.H., et al. Biochem. Biophys. Res Commun. 312:997-1004 (2003); Yamakawa, T., et al. Mol Endocrinol. 22:441-453 (2008)). *PDLIM7* is an actin and protein kinase adaptor that promotes mineralization, which might be relevant to plaques with calcium. The most significant signal associated with CAC was lower expression of *CACNA2D3* (Calcium Channel, Voltage-Dependent, Alpha 2/Delta Subunit 3; $p=2.08 \times 10^{-8}$, FDR= 3.19×10^{-4} ; with carotid plaque: $p=6.71 \times 10^{-3}$, FDR=0.40) a tumor suppressor gene that can induce mitochondrial-mediated apoptosis (Wong, A.M. et al. Int J Cancer. 133:2284-2295 (2013)). Additionally adjusting for other traditional CVD risk factors and statin use (in a model designated as the full model) had minimal impact on the significant associations (Figures 1C and 1D). All of these associations had a consistent direction of effect across the independent study sites, particularly for three genes (*ARID5B*, *TSPYL1*, and *ADM*) which were cross-validated ($p<0.05$, Figures 1C and 1D). Race/ethnicity- and sex-specific analyses also were consistent across various strata (Table 3). For the top signal, *ARID5B*, the associations with carotid plaque were consistent in direction and nominally significant across subgroups of age (< or ≥ 65 years), sex, and statin use (Figure 1E). Results utilizing RNA-sequencing derived expression levels from a subset (N=354) of the monocyte samples validated significant

associations between *ARID5B* and CAC ($\beta \pm SE = 0.45 \pm 0.20$, $p = 0.03$) and carotid plaque ($\beta \pm SE = 0.13 \pm 0.06$, $p = 0.04$).

To identify network modules of highly correlated transcripts in an unbiased fashion, the weighted gene co-expression network analysis was applied (Langfelder, P. & Horvath, S. BMC. Bioinformatics. 9:559 (2008)), identifying 40 co-expressed gene network modules. Three modules significantly ($FDR \leq 0.05$) associated with CAC (Table 4), including a cholesterol metabolism transcriptional network (CMTN) with 12 functionally coupled genes that was also associated with carotid plaque. The CMTN eigengene (the first principal component) of CMTN is associated with a transcriptional profile expected to increase intracellular cholesterol, including up-regulation of cholesterol uptake (*LDLR* and *MYLIP*) and cholesterol and fatty acid synthesis genes (*HMGCS1*, *FDFT1*, *SQLE*, *CYP51A1*, *SC4MOL*, *SC5DL*, *SCD* and *FADS1*), as well as down-regulation of cholesterol efflux genes (*ABCG1* and *ABCA1*). The CMTN is a signature feature of obesity, which was associated with CAC ($p=3.34 \times 10^{-4}$) (Ding, J. et al. Diabetes. 64:3464-3474 (2015)). Here, it is further showed that the CMTN was positively associated with carotid plaque ($p=2.31 \times 10^{-4}$; $FDR=0.009$). Among the CMTN members, *LDLR* expression was most significantly associated with CAC ($p=1.34 \times 10^{-4}$, $FDR=0.10$) and *ABCG1* expression was most significantly associated with carotid plaque ($p=1.48 \times 10^{-4}$, $FDR=0.18$, Table 2). The other two network modules significantly ($FDR \leq 0.05$) associated with CAC were enriched with genes involved in phagosome formation (module 39; $FDR=1.32 \times 10^{-2}$) and migration of cells (module 23; $FDR=3.96 \times 10^{-2}$; Table 4).

Methylome Signature of Atherosclerosis

Of the 484,817 CpG sites, 31 and 7 had methylation significantly associated with carotid plaque and CAC, respectively, including one CpG (cg23661483 in exon of *ILVBL*) associated with both carotid plaque and CAC ($FDR \leq 0.05$, adjusting for demographics, Figure 2A and 2B, and Table 5 showing additional signals for $FDR \leq 0.10$). The most significant CpG (cg06126421) associated with carotid plaque score ($p=2.00 \times 10^{-10}$, $FDR=8.91 \times 10^{-5}$) is an intergenic CpG, which recently was reported as one of the most significant smoking-associated methylation sites genome-wide (Shenker, N.S. et al. Hum Mol Genet 22:843-851 (2013)). Notably, the carotid plaque-associated methylation sites include one CpG in *ARID5B* (cg25953130, intron 2, $p=4.31 \times 10^{-7}$, $FDR=0.01$), which tends to be hypomethylated in the individuals with higher carotid plaque scores. Similar results were found between cg25953130 methylation and CAC ($p=6.80 \times 10^{-5}$, $FDR=0.32$).

The 37 differentially methylated sites associated with carotid plaque or CAC presented higher variable CpG methylation levels (across MESA population) with the interdecile range (IDR) of the percentage of methylation (measured by beta-value (Du, P. et al. BMC Bioinformatics. 11:587 (2010)) ranging from 4 to 24% (median: 10%), compared to the whole methylome (median: 4%). The majority of the methylation sites associated with carotid plaque or CAC had inverse associations with the atherosclerosis measures (Table 5). The 37 atherosclerosis-associated CpGs are distributed among 30 unique genomic loci.

Additional adjustments in the full model had minimal impact on the significant associations, which were also consistent across the two study sites for the majority (61%) of differentially methylated sites ($p < 0.05$, as shown for each unique loci in Figures 2C and 2D). Race/ethnicity- and sex-specific analyses also showed high consistency across the various strata (Table 6), suggesting shared effects across ancestries. At the genome-wide level, race/ethnicity-specific analyses did not uncover additional significant associations. The associations of *ARID5B* methylation (cg25953130, intron) with carotid plaque were consistent in direction and nominally significant across subgroups of age ($<$ or ≥ 65 years), sex, CVD status, and statin use (Figure 2E). Pyrosequencing derived methylation of cg25953130 from a subset ($n=90$) of the monocyte samples significantly correlated with microarray-based methylation levels ($r = 0.92$, $p = 5.2 \times 10^{-37}$), and validated significant associations between cg25953130 methylation and carotid plaque ($\beta \pm se = -0.07 \pm 0.03$, $p = 6.31 \times 10^{-3}$).

In Vivo and In Silico Functional Validation

DNA methylation has been viewed as an important potential regulator of gene expression (Jones, P.A. Nat Rev Genet 13:484-492 (2012)). To prioritize the list of differentially methylated CpGs, expression-associated methylation sites (eMS) reported in the same monocyte samples from a previous study were assessed (Wong, K.L. et al. Blood 118:e16-e31 (2011)). Five atherosclerosis-associated CpGs whose methylation was significantly related to mRNA expression of at least one nearby gene (Table 7) was identified. Out of the five-atherosclerosis associated eMS, four had methylation correlated with mRNA expression profiles that were nominally ($p < 0.05$) associated with atherosclerosis, including expression of *SC4MOL*. The CpG most significantly associated with CAC (cg05119988, located in the 5' UTR region of *SC4MOL*) significantly correlated ($r = -0.17$, $p = 1.2 \times 10^{-10}$, $FDR = 7.3 \times 10^{-8}$) with mRNA expression of *SC4MOL*, a member of the identified CMNT, which was nominally associated with both carotid plaque and CAC. The *SC4MOL* eMS resides in a

predicted weak promoter region of a B cell line; however, this region was identified as heterochromatin in monocytes (using ChromHMM (Ernst, J. & Kellis, M. Nat Methods 9:215-216 (2012)) for histone modification data from BLUEPRINT (Adams, D. et al. Nat Biotechnol 30:224-226 (2012); Saeed, S. et al. Science 345:1251086 (2014)),
 5 Fig. 8, Table 5 and Table 7).

The *ARID5B* CpG (cg25953130) was the only CpG site significantly associated with expression of *ARID5B* ($p=3.84 \times 10^{-14}$, $FDR=5.70 \times 10^{-6}$; Figure 3A). As shown in Figure 3B, it overlaps a DNase hypersensitive hotspot (BLUEPRINT monocyte data (Adams, D. et al. Nat Biotechnol 30:224-226 (2012); Saeed, S. et al.
 10 Science 345:1251086 (2014))), a predicted strong enhancer (using both monocyte and B cell line histone mark data from the BLUEPRINT (Adams, D. et al. Nat Biotechnol 30:224-226 (2012); Saeed, S. et al. Science 345:1251086 (2014)) and ENCODE (Rosenbloom, K.R. et al. Nucleic Acids Res. 41:D56-D63 (2013)) projects, respectively), and a transcription factor binding site occupied by EP300 (in a
 15 neuroblastoma cell line). More importantly, chromatin-capture sequencing technologies (both Hi-C and ChIA-PET) confirmed direct interactions between regions in the *ARID5B* cg25953130 locus and the *ARID5B* promoter region in B cell line (Rao, S.S. et al. Cell 159:1665-1680 (2014); Dixon, J.R. et al. Nature 485:376-380 (2012); Heidari, N. et al. Genome Res 24, 1905-1917 (2014); Dixon, J.R. et al.
 20 Nature 518:331-336 (2015)) (Figure 3C). These data, together with the publically available functional data, strongly support the presence of an *ARID5B* regulatory region in the *ARID5B* gene body flanking *ARID5B* cg25953130.

To test whether the assumed methylation effects on atherosclerosis burden were mediated through its associated mRNA expression, Structural Equation
 25 Modeling was used with bootstrapping (SEM, *R* package *lavaan* (Rosseel, Y. Journal of Statistical Software 48:1-36 (2012))) to perform mediation analyses. *ARID5B* mRNA expression significantly mediated 15% and 14% of the total effect of this *ARID5B* CpG on carotid plaque score (indirect effect, $p=2.1 \times 10^{-4}$) and CAC ($p=2.1 \times 10^{-3}$), respectively.

The inverse association between methylation of this *ARID5B* CpG and carotid plaque score after adjusting for *ARID5B* expression (direct effect) was also significant. Jointly, the associations of *ARID5B* gene expression and methylation levels with atherosclerosis explain an additional 2.3% of the variability in carotid plaque score beyond well-known CVD risk factors and statin use. The effect sizes of
 35 *ARID5B* gene expression and methylation levels associating with carotid plaque score are higher than the effect sizes of T2DM (1% of variability) or hypertension

(0.9%) in the same model. These data suggest that different types of related genomic features (mRNA expression and DNA methylation) may offer additive values in prediction of CVD susceptibility.

To further demonstrate the clinical relevance of *ARID5B*, the tertiles of *ARID5B* mRNA expression and *ARID5B* cg25953130 methylation were associated with presence of carotid plaque (defined as carotid plaque score greater than zero, N=816 cases), presence of CAC (defined as CAC greater than zero, N=844 cases), and prevalent CVD (history of a coronary heart event or stroke, N=64 cases).

ARID5B mRNA expression was positively associated with presence of carotid plaque (3rd tertile odds ratio=2.10, 95% CI: 1.42–3.09, $p=1.87 \times 10^{-4}$), presence of CAC (3rd tertile odds ratio=2.10, 95% CI: 1.42–3.09, $p=1.87 \times 10^{-4}$), and prevalent CVD (3rd tertile odds ratio=2.33, 95% CI: 1.08–5.02, $p=3.12 \times 10^{-2}$, Figure 4A); while *ARID5B* cg25953130 methylation was inversely associated with presence of carotid plaque (3rd tertile odds ratio=0.64, 95% CI: 0.45–0.92, $p=1.59 \times 10^{-2}$) and presence of CAC (3rd tertile odds ratio=0.64, 95% CI: 0.45–0.92, $p=1.59 \times 10^{-2}$, Figure 4B) in the full model. To examine the dose response relationship between *ARID5B* and extent of atherosclerosis, which may indicate their potential contribution to the progression of plaques, we performed linear regression analysis while excluding those with zero value of carotid plaque score. The associations with carotid plaque score remain significant for the *ARID5B* mRNA expression (unique variability explained: 1.2%, $p=1.1 \times 10^{-3}$) and cg25953130 methylation (unique variability explained: 0.73%, $p=1.1 \times 10^{-2}$).

Identified Genomic Features and Known CVD Risk Factors

Among the atherosclerosis-associated genomic features, eight mRNA (Table 8) and 33 CpGs (Table 9) were also associated with one or more traditional CVD risk factors, particularly demographics (age, sex, race/ethnicity), cigarette smoking, and obesity. For three mRNA and 29 CpGs (bolded in Table 6 and Table 9), the predicted effects of the majority of CVD risk factors on atherosclerosis that were mediated through the genomic features have the similar direction as the observed associations between the CVD risk factors and the measure of atherosclerosis.

ARID5B expression was positively associated with many CVD risk factors (FDR<0.05 with adjustment for demographic variables), such as age, ($p=2.06 \times 10^{-13}$), BMI ($p=1.67 \times 10^{-6}$), T2DM ($p=3.43 \times 10^{-5}$), and inflammatory stress (measured by plasma interleukin-6 levels (IL6; $p=1.36 \times 10^{-10}$), and inversely associated with HDL-C levels ($p=6.86 \times 10^{-6}$). *ARID5B* methylation (cg25953130) was associated inversely with age ($p=3.33 \times 10^{-11}$) and plasma IL-6 levels ($p=0.004$), and the *ARID5B*

CpG tended to be hypomethylated in current smokers ($p=7.19 \times 10^{-7}$). If *ARID5B* is in fact causal, mediation analyses showed the expression of *ARID5B* significantly mediated 10% and 25% of the total effect of age and IL-6 on carotid plaque score ($p=7.46 \times 10^{-6}$, 4.33×10^{-5} , respectively), and the *ARID5B* methylation significantly mediated 7% and 10% of the total effect of age and IL6 on carotid plaque score ($p=2.57 \times 10^{-4}$, 8.28×10^{-3} , respectively).

ARID5B RNA Expression and Methylation in CD4+ T cells

To examine the *ARID5B* expression and methylation across cell-types, similar analyses were performed in a subset of 517 MESA CD4+ T cell samples.

Correlations for *ARID5B* expression and methylation between monocyte and T cells were weak ($r=0.12$, 0.30 , respectively). Within T cell samples, *ARID5B* expression was inversely correlated with cg25953130 methylation ($r = -0.45$, $p = 1.27 \times 10^{-31}$), as seen in monocytes. The *ARID5B* mRNA and cg25953130 methylation associations with carotid plaque and CAC were not significant in T cell samples ($p>0.05$), but remained statistically significant (p ranges from 3.8×10^{-3} to 7.2×10^{-3}) when analyzed in the monocyte samples from the same subset. In the same data, however, there was an association of *AHRR* cg05575921 with carotid plaque score and CAC score in T cell samples ($p=3.18 \times 10^{-5}$, 5.89×10^{-5} , respectively), with a strong correlation between the two cell types for the *AHRR* methylation site ($r=0.97$). *AHRR* hypomethylation is a well-known, robust biomarker of smoking that is linked to carotid plaque score in MESA (Reynolds, L.M. et al. Circ Cardiovasc. Genet 8:707-716 (2015)). These results demonstrate examples of both potential cell-type specific and shared genomic features in relation to burden of atherosclerosis.

Functional Evaluation of ARID5B Using In Vitro Models

Although little is known about the majority of identified genomic features, pleiotropic effects of *ARID5B* in adipogenesis (Whitson, R.H., et al. Biochem. Biophys. Res Commun. 312:997-1004 (2003)), chondrogenesis (Hata, K. et al. Nat Commun. 4:2850. (2013)), autoimmune diseases (Yang, W. et al. Am J Hum Genet. 92:41-51 (2013); Okada, Y. et al. Nat Genet. 44:511-516 (2012)), lipid metabolism (Yamakawa, T., et al. Mol Endocrinol. 22:441-453 (2008)), and smooth muscle cell differentiation (Watanabe, M. et al. Circ Res. 91:382-389 (2002)) have been previously reported. As a transcription coactivator that is part of the H3K9me2 demethylase complex with PHD finger protein 2 (PHDF2), *ARID5B* is expected to activate its target genes by removing the repressive H3K9Me2 histone mark (demethylation of H3K9me2) from the promoter region of target genes (Hata, K. et al. Nat Commun. 4:2850. (2013); Baba, A. et al. Nat Cell Biol. 13:668-675 (2011)). To

evaluate the functional role of ARID5B in monocytes/macrophages, the effects of its knockdown on transcriptomic profiles we was in LPS stimulated human THP1-monocytes in two independent experiments (LPS: 100 ng ml⁻¹, n=4 per group at each experiment).

5 *ARID5B* mRNA knockdown efficiency was 85% and 76% in the two experiments (Fig. 9); both consistently showed that *ARID5B* knockdown decreased expression of 1,320 genes and increased expression of 1,162 genes (FDR<0.005 in the discovery set and FDR<0.05 in the replication set). The 2,482 ARID5B-modified genes displayed a significant overrepresentation of related pathways including
10 inflammatory/immune response, chemotaxis, migration, extravasation signaling, and phagocytosis, and also the lipid synthesis functional pathway, compared to the background list of genes detectable on the array (enrichment FDR<0.05, Figure 5A, full list of genes provided in Table 10). Enriched bio-functions and canonical pathways in the inflammatory response pathway consisted of mainly down-regulated
15 genes, including key proinflammatory cytokines (e.g. TNF, IL1a), activator and effector cytokines from the type I interferon signaling pathway (e.g. IRF3, IFNB1, STAT1 and STAT2), and antigen processing and presentation genes (e.g. MHC class II cell surface receptors: HLA-DRA, and HLA-DRBs, Fig. 9, bolded in Table 10), collectively indicating decreased pathway activation. Decreased expression of genes
20 in two of the enriched pathways, phagocytosis and lipid synthesis (Table 10), overlapped with the members of two atherosclerosis-associated networks we identified in MESA (CMTN and module 39 - enriched for the phagosome formation pathway, Table 4), which suggests ARID5B may contribute to the gene network associations with atherosclerosis.

25 To interrogate the potential pro-inflammatory role that ARID5B may play, the third *ARID5B* knockdown experiment was repeated and ELISA assays of culture media performed for IL1A, the pro-inflammatory cytokine that was most significantly reduced by *ARID5B* mRNA knockdown (FDR: 3.79x10⁻⁸, 4.48x10⁻⁴ in the first two experiments). *ARID5B* knockdown decreased levels of both *IL1A* mRNA (p=0.02) in
30 the THP1-monocytes and IL1A protein expression (p=4.28x10⁻⁴) in their culture media, as shown in Figure 5B.

To further examine the effects of *ARID5B* knockdown on cellular functions suggested by the transcriptomic profile changes, THP1-monocyte migration and phagocytosis assays were performed. *ARID5B* knockdown suppressed monocyte
35 migration (p=3x10⁻⁷, 0.004 in experiment 1 and 2, respectively, Figure 5C). ARID5B

knockdown also moderately inhibited monocyte phagocytosis ($p=0.008$) as shown in Figure 5C.

Table 2. Monocyte Genes with Expression Associated with Atherosclerosis (FDR < 0.20)

Gene			Carotid Plaque Score (n=1,181)				Coronary Artery Calcium (n=1,208)			
Gene name	illumina probe id	Chr	Beta	SE	P-value	FDR	Beta	SE	P-value	FDR
CACNA2D3	ILMN_1754076	3	-0.06	0.02	6.71E-03		-0.44	0.08	2.08E-08	3.19E-04
ARID5B	ILMN_1721626	10	0.12	0.02	6.30E-08	1.08E-03	0.32	0.08	2.47E-05	3.45E-02
PDLIM7	ILMN_1690125	5	0.10	0.02	3.56E-06	3.05E-02	0.31	0.07	8.49E-06	1.95E-02
LILRB4	ILMN_2355953	19	0.07	0.02	1.41E-03		0.32	0.07	6.44E-06	1.95E-02
TSPYL1	ILMN_1779014	6	-0.07	0.02	2.10E-03		-0.34	0.07	3.57E-06	1.95E-02
P2RX4	ILMN_1813846	12	0.04	0.02	4.82E-02		0.31	0.07	8.06E-06	1.95E-02
RSU1	ILMN_2383344	10	-0.02	0.02	0.29		-0.33	0.07	5.29E-06	1.95E-02
FAM168B	ILMN_1677122	2	-0.02	0.02	0.40		-0.32	0.07	1.01E-05	1.95E-02
CANX	ILMN_2401057	5	0.00	0.02	0.86		-0.33	0.07	9.58E-06	1.95E-02
IARS	ILMN_1733956	9	-0.04	0.02	0.07		-0.31	0.07	2.21E-05	3.39E-02
ADM	ILMN_1708934	11	0.02	0.02	0.33		0.32	0.07	2.07E-05	3.39E-02
FAM10A4	ILMN_1726308	13	-0.04	0.02	4.72E-02		-0.29	0.07	3.39E-05	4.31E-02
EIF5A2	ILMN_1735151	3	0.03	0.02	0.16		0.29	0.07	3.64E-05	4.31E-02
EBAG9	ILMN_1791896	8	0.00	0.02	0.99		-0.30	0.07	5.17E-05	0.06
PTPRN2	ILMN_1728844	7	-0.05	0.02	1.34E-02		-0.27	0.07	7.93E-05	0.08
ZNF395	ILMN_1772876	8	-0.02	0.02	0.25		-0.28	0.07	9.50E-05	0.09
IMPDH2	ILMN_1705737	3	-0.01	0.02	0.70		-0.28	0.07	1.00E-04	0.09
TRNT1	ILMN_1764098	3	-0.05	0.02	2.72E-02		-0.26	0.07	1.26E-04	0.10
LDLR	ILMN_2053415	19	0.05	0.02	3.14E-02		0.27	0.07	1.34E-04	0.10
F11R	ILMN_2342033	1	-0.06	0.02	2.81E-03		-0.26	0.07	1.67E-04	0.11
MEF2C	ILMN_1742544	5	-0.06	0.02	3.49E-03		-0.27	0.07	1.86E-04	0.11
GAS6	ILMN_1779558	13	0.06	0.02	6.38E-03		0.26	0.07	2.11E-04	0.11
SHKBP1	ILMN_1765493	19	0.06	0.02	7.86E-03		0.27	0.07	1.69E-04	0.11
FER1L3	ILMN_2370976	10	0.06	0.02	8.68E-03		0.27	0.07	1.91E-04	0.11
CCNB1IP1	ILMN_2347349	14	-0.04	0.02	0.05		-0.27	0.07	2.19E-04	0.11
OLR1	ILMN_1723035	12	0.04	0.02	0.05		0.26	0.07	2.14E-04	0.11
MXD1	ILMN_2214678	2	0.04	0.02	0.05		0.28	0.07	1.60E-04	0.11
USP14	ILMN_1806804	18	-0.03	0.02	0.12		-0.26	0.07	2.07E-04	0.11
FAM160B2	ILMN_1802971	8	-0.01	0.02	0.75		0.26	0.07	2.06E-04	0.11
CDC42EP4	ILMN_1745223	17	0.00	0.02	0.93		0.29	0.08	2.15E-04	0.11
POLR1E	ILMN_2110167	9	-0.03	0.02	0.15		-0.26	0.07	2.48E-04	0.11

FGFRL1	ILMN_1795865	4	0.04	0.02	0.06		0.29	0.08	2.82E-04	0.12
SERBP1	ILMN_1773968	1	-0.03	0.02	0.13		-0.25	0.07	3.02E-04	0.12
DAB2	ILMN_1764228	5	0.02	0.02	0.37		0.25	0.07	3.03E-04	0.12
C19ORF39	ILMN_1707943	19	-0.02	0.02	0.26		-0.25	0.07	3.33E-04	0.13
RPN2	ILMN_1693421	20	-0.02	0.02	0.28		-0.25	0.07	3.43E-04	0.13
ALDH1A1	ILMN_2096372	9	-0.06	0.02	1.44E-03		-0.24	0.07	3.66E-04	0.13
CIAO1	ILMN_2048793	2	-0.05	0.02	1.33E-02		-0.26	0.07	4.01E-04	0.13
TPST2	ILMN_2329679	22	0.05	0.02	2.03E-02		0.26	0.07	3.94E-04	0.13
C6ORF225	ILMN_3248941	6	-0.03	0.02	0.12		-0.26	0.07	4.00E-04	0.13
COX17	ILMN_2187718	3	0.03	0.02	0.20		0.26	0.07	3.89E-04	0.13
SLC39A8	ILMN_1695316	4	0.09	0.02	3.67E-05	0.14	0.19	0.07	6.33E-03	
RBM38	ILMN_2404049	20	0.09	0.02	4.28E-05	0.14	0.12	0.07	0.09	
ALDOA	ILMN_1741148	16	0.09	0.02	4.84E-05	0.14	0.19	0.07	8.76E-03	
BLCAP	ILMN_1675612	20	-0.04	0.02	0.11		-0.26	0.07	4.44E-04	0.15
PAQR7	ILMN_1746618	1	-0.04	0.02	4.35E-02		-0.24	0.07	4.54E-04	0.15
FIGNL2	ILMN_3249343	12	-0.04	0.02	0.08		-0.24	0.07	4.86E-04	0.15
GGCT	ILMN_1745005	7	-0.05	0.02	2.96E-02		-0.25	0.07	5.06E-04	0.16
FDFT1	ILMN_2144088	8	0.07	0.02	7.31E-04		0.23	0.07	6.02E-04	0.16
CREBL2	ILMN_1660223	12	-0.04	0.02	4.87E-02		-0.24	0.07	6.22E-04	0.16
TCIRG1	ILMN_1711994	11	0.04	0.02	0.07		0.24	0.07	5.74E-04	0.16
CENTB2	ILMN_2088825	3	-0.04	0.02	0.07		-0.24	0.07	5.75E-04	0.16
RRAS	ILMN_1780825	19	0.03	0.02	0.24		0.26	0.08	6.12E-04	0.16
SNHG11	ILMN_1791884	20	0.02	0.02	0.39		0.24	0.07	5.78E-04	0.16
ENPP2	ILMN_2373791	8	0.01	0.02	0.53		0.24	0.07	5.71E-04	0.16
AKR1B1	ILMN_1701731	7	-0.01	0.02	0.55		-0.25	0.07	6.22E-04	0.16
CCNB1IP1	ILMN_1752394	14	-0.01	0.02	0.68		-0.25	0.07	5.57E-04	0.16
SLC25A14	ILMN_2347044	X	-0.01	0.02	0.80		0.24	0.07	6.11E-04	0.16
CR1	ILMN_1767193	1	0.05	0.02	1.86E-02		0.24	0.07	6.97E-04	0.16
ELMO2	ILMN_2317730	20	0.05	0.02	2.53E-02		0.25	0.07	7.00E-04	0.16
ACAA1	ILMN_1738921	3	0.04	0.02	3.77E-02		0.24	0.07	7.17E-04	0.16
CHRA1	ILMN_1722390	8	-0.03	0.02	0.13		-0.24	0.07	6.58E-04	0.16
RNF144B	ILMN_1752526	6	0.03	0.02	0.15		0.24	0.07	7.09E-04	0.16
PDIA5	ILMN_1695763	3	0.01	0.02	0.52		-0.24	0.07	6.62E-04	0.16
STX10	ILMN_1753712	19	0.01	0.02	0.76		0.24	0.07	6.83E-04	0.16

CDKN1A	ILMN_1784602	6	0.00	0.02	1.00		0.26	0.08	6.91E-04	0.16
RAB37	ILMN_2255579	17	-0.02	0.02	0.32		-0.25	0.07	7.37E-04	0.16
RAB37	ILMN_2243912	17	-0.04	0.02	4.56E-02		-0.25	0.07	7.72E-04	0.16
PRPF4	ILMN_1697440	9	-0.01	0.02	0.63		-0.23	0.07	7.78E-04	0.16
TPST2	ILMN_1695432	22	0.07	0.02	2.62E-03		0.24	0.07	8.73E-04	0.17
PIM1	ILMN_1815023	6	0.05	0.02	2.36E-02		0.25	0.07	8.11E-04	0.17
PFAS	ILMN_1755862	17	-0.03	0.02	0.10		-0.24	0.07	8.67E-04	0.17
ALKBH6	ILMN_1808792	19	-0.01	0.02	0.61		0.23	0.07	8.52E-04	0.17
ALOX15B	ILMN_2393149	17	-0.01	0.02	0.61		-0.24	0.07	8.39E-04	0.17
SET	ILMN_1742238	9	0.00	0.02	0.99		-0.24	0.07	8.52E-04	0.17
ATP2B1	ILMN_1813399	12	-0.09	0.02	6.91E-05	0.17	-0.09	0.07	0.24	
MYL6	ILMN_2326071	12	0.02	0.02	0.29		0.24	0.07	9.22E-04	0.18
ABCG1	ILMN_2329927	21	-0.06	0.02	2.39E-03		-0.23	0.07	9.35E-04	0.18
INPP1	ILMN_1667239	2	-0.08	0.02	8.82E-05	0.18	-0.16	0.07	1.96E-02	
MT1H	ILMN_2124802	16	0.08	0.02	9.82E-05	0.18	0.22	0.07	1.62E-03	
PGD	ILMN_1794165	1	0.09	0.02	1.09E-04	0.18	0.20	0.08	1.16E-02	
GALK2	ILMN_1723124	15	-0.09	0.02	1.37E-04	0.18	-0.15	0.08	4.85E-02	
PSMD1	ILMN_1702837	2	0.09	0.02	1.37E-04	0.18	0.06	0.07	0.40	
C10ORF58	ILMN_2118472	10	-0.08	0.02	1.41E-04	0.18	-0.03	0.07	0.64	
ABCG1	ILMN_1794782	21	-0.08	0.02	1.48E-04	0.18	-0.21	0.07	2.82E-03	
BCL2	ILMN_2246956	18	0.11	0.03	1.74E-04	0.18	0.16	0.10	0.09	
RPS19	ILMN_1784717	19	0.08	0.02	1.75E-04	0.18	0.13	0.07	0.06	
NICN1	ILMN_1807211	3	0.08	0.02	1.86E-04	0.18	0.18	0.07	1.14E-02	
MAPK13	ILMN_1749327	6	0.08	0.02	1.88E-04	0.18	0.20	0.07	5.43E-03	
PFDN2	ILMN_1678754	1	0.08	0.02	2.19E-04	0.18	0.16	0.07	2.28E-02	
SAMSN1	ILMN_1684887	21	0.08	0.02	2.27E-04	0.18	0.20	0.07	5.08E-03	
MID1IP1	ILMN_1668960	X	-0.08	0.02	2.30E-04	0.18	-0.14	0.07	0.06	
C10ORF24	ILMN_1667966	1	0.07	0.02	1.43E-03		0.23	0.07	9.70E-04	0.18
UBE2J1	ILMN_1713759	6	0.05	0.02	3.61E-02		0.25	0.07	9.84E-04	0.18
HIST3H2A	ILMN_1779648	1	0.04	0.02	4.78E-02		0.23	0.07	9.94E-04	0.18
ANAPC1	ILMN_1804812	2	-0.06	0.02	6.11E-03		-0.23	0.07	1.03E-03	0.18
CRTAP	ILMN_1720484	3	-0.04	0.02	0.06		-0.24	0.07	1.02E-03	0.18
RASSF2	ILMN_2352303	20	0.00	0.02	0.92		-0.24	0.07	1.03E-03	0.18
PDCD4	ILMN_1768004	10	-0.05	0.02	2.34E-02		-0.23	0.07	1.06E-03	0.18

FDFT1	ILMN_1741096	8	0.07	0.02	2.44E-04	0.18	0.04	0.07	0.55	
LYZ	ILMN_2162972	12	0.06	0.02	1.36E-02		0.25	0.08	1.25E-03	0.19
PIP5K1B	ILMN_1736654	9	-0.04	0.02	0.05		-0.23	0.07	1.18E-03	0.19
RAB27A	ILMN_2329773	15	-0.03	0.02	0.12		-0.24	0.08	1.23E-03	0.19
NCF1C	ILMN_2112988	7	0.03	0.02	0.13		0.23	0.07	1.24E-03	0.19
TMEM19	ILMN_3251207	12	-0.03	0.02	0.14		-0.24	0.07	1.15E-03	0.19
DLEU7	ILMN_3307752	13	-0.02	0.02	0.30		-0.24	0.08	1.26E-03	0.19
ZBTB17	ILMN_1711048	1	0.02	0.02	0.45		0.23	0.07	1.24E-03	0.19
EIF4EBP3	ILMN_2197225	5	-0.01	0.02	0.66		-0.23	0.07	1.27E-03	0.19
METAP2	ILMN_1670420	12	0.00	0.02	0.98		-0.23	0.07	1.26E-03	0.19
ARAP2	ILMN_3244434	4	0.05	0.02	2.85E-02		0.23	0.07	1.31E-03	0.19
NR3C1	ILMN_2389347	5	-0.04	0.02	4.84E-02		-0.23	0.07	1.31E-03	0.19
PTPRO	ILMN_1720113	12	0.03	0.02	0.12		0.24	0.07	1.30E-03	0.19
C10ORF32	ILMN_1772706	10	-0.03	0.02	0.10		-0.23	0.07	1.33E-03	0.19
STX12	ILMN_1773901	1	-0.04	0.02	0.06		-0.23	0.07	1.35E-03	0.19
ALPK1	ILMN_2078697	4	0.06	0.02	5.23E-03		0.23	0.07	1.41E-03	0.20
FAM63A	ILMN_1689552	1	-0.05	0.02	4.11E-02		-0.25	0.08	1.40E-03	0.20
ZNF770	ILMN_1734254	15	0.00	0.02	0.88		-0.23	0.07	1.41E-03	0.20
CALML4	ILMN_1717757	15	0.04	0.02	0.08		0.22	0.07	1.44E-03	0.20
SLC26A6	ILMN_1785252	3	0.00	0.02	0.83		0.23	0.07	1.45E-03	0.20
LAIR1	ILMN_1768598	19	0.06	0.02	7.62E-03		0.22	0.07	1.49E-03	0.20
ATXN10	ILMN_2169676	22	-0.05	0.02	1.79E-02		-0.22	0.07	1.47E-03	0.20
FBL	ILMN_1719205	19	0.03	0.02	0.15		-0.24	0.08	1.50E-03	0.20
CRY1	ILMN_2202637	12	-0.04	0.02	0.05		-0.22	0.07	1.52E-03	0.20
Sorted by most significant association (FDR) with carotid plaque burden or CAC; FDR ≥ 0.20 are left blank.										

Table 3. Race/ethnicity and sex-stratified analysis of genes with mRNA expression associated with Atherosclerosis														
Carotid Plaque Score associated gene expression (FDR≤0.05):														
Gene	African American (N=260)			Caucasian (N=532)			Hispanic (N=372)			Women (N=603)			Men (N=561)	
	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	P-val
ARID5B	0.13	0.06	0.019	0.15	0.03	2.3E-06	0.06	0.04	0.17	0.08	0.03	0.014	0.14	1.0E-05
PDLIM7	0.05	0.05	0.27	0.09	0.03	2.6E-03	0.08	0.04	0.034	0.10	0.03	8.6E-04	0.05	0.13
Coronary artery calcium associated gene expression (FDR≤0.05):														
Gene	African American (N=259)			Caucasian (N=544)			Hispanic (N=387)			Women (N=613)			Men (N=577)	
	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	P-val
CACNA2D3	-0.64	0.16	7.4E-05	-0.25	0.12	0.029	-0.26	0.14	0.078	-0.12	0.11	0.28	-0.56	1.6E-06
ARID5B	0.36	0.19	0.057	0.14	0.10	0.16	0.20	0.14	0.15	0.06	0.11	0.56	0.33	1.9E-03
PDLIM7	0.07	0.15	0.63	0.21	0.10	0.042	0.27	0.13	0.033	0.11	0.10	0.28	0.26	0.010
TSPYL1	-0.28	0.14	0.054	-0.28	0.09	2.4E-03	-0.19	0.14	0.15	-0.28	0.08	8.9E-04	-0.29	9.5E-03
RSU1	-0.32	0.16	0.044	-0.19	0.10	0.056	-0.24	0.12	0.047	-0.27	0.09	3.6E-03	-0.19	0.067
LILRB4	0.39	0.14	6.1E-03	0.23	0.09	0.014	0.29	0.14	0.032	0.23	0.09	0.011	0.33	9.7E-04
P2RX4	0.20	0.16	0.22	0.36	0.10	2.3E-04	0.10	0.12	0.37	0.22	0.09	0.016	0.26	9.4E-03
CANX	-0.19	0.16	0.22	-0.32	0.10	2.1E-03	-0.21	0.13	0.12	-0.31	0.09	8.6E-04	-0.23	0.034
FAM168B	-0.27	0.16	0.090	-0.24	0.10	0.013	-0.26	0.12	0.038	-0.24	0.09	7.4E-03	-0.29	5.6E-03
ADM	0.29	0.14	0.045	0.26	0.11	0.016	0.27	0.14	0.049	0.17	0.10	0.084	0.39	2.3E-04
IARS	-0.30	0.17	0.078	-0.21	0.10	0.032	-0.23	0.12	0.061	-0.18	0.09	0.055	-0.30	3.6E-03
FAM10A4	-0.36	0.15	0.016	-0.18	0.10	0.065	-0.19	0.13	0.16	-0.30	0.09	1.1E-03	-0.14	0.17
EIF5A2	0.32	0.14	0.025	0.30	0.10	2.5E-03	0.04	0.13	0.76	0.24	0.09	6.8E-03	0.18	0.098

Table 4. Gene co-expression networks associated with subclinical atherosclerosis

Module	Gene co-expression networks associated with subclinical atherosclerosis		tCarotid Plaque Score			tCoronary Artery Calcium		
			Beta	SE	P-value	Beta	SE	P-value
Module 35	LDLR, SC5DL, FADS1, CYP51A1, FDFT1, SCD, MYLIP, SC4MOL, HMGCS1, SQLE, ABCG1, ABCA1	*Enriched Bio-functions	0.08	0.02	2.31E-04	0.25	0.07	3.34E-04
Module 39	PIM1, GAS6, MTHFD2, FCGR1B, FCGR1A, FCGR1C, CASP5	Steroid metabolism (FDR = 4.23E-11; LDLR, CYP51A1, SQLE, HMGCS1, ABCA1, ABCG1, SC5DL, FDFT1, SC4MOL) Phagosome formation (FDR = 1.32E-02; FCGR1C, FCGR1A, FCGR1B)	0.06	0.02	4.62E-03	0.24	0.07	5.26E-04
Module 23	OSBPL8, CCDC109A, CX3CR1, CBARA1, RRAGD, RAPH1, FAM129A, UBE2J1, HDAC9, EPS8, ITGB5, C1ORF24, LPCAT3, CDK6, ELMO2, TPST2, SORT1, KCNMB4, PLA2G15, P2RX4, SNTB1, SLITRK4, PDGFC, MYD88, MPHOSPH6, TCF4	Migration of cells (FDR = 3.96E-02; ITGB5, EPS8, CX3CR1, MYD88, CDK6, OSBPL8, ELMO2, TCF4, SORT1)	0.05	0.02	0.02	0.22	0.07	1.63E-03
*Enrichment from IPA Ingenuity, module genes compared to a background of all genes with detectable expression. tModule eigengene associations FDR _{BH} ≤ 0.05 shown in bold.								

Table 5. Monocyte CpGs with Methylation Associated with Atherosclerosis (FDR ≤ 0.10)

CpG methylation characteristics										Methylation ~ Carotid Plaque (N=1,176)				Methylation ~ CAC (N=1,203)			
CpG ID	Chr	Location	CpG island	*TFBS	±DMR	±DNase HS	±Predicted state in B cell line (GM12878)	Beta	SE	P-value	FDR	Beta	SE	P-value	FDR		
cg06126421	6	Intergenic		1		1	13 Heterochrom/lo	-0.14	0.02	2.00E-10	8.91E-05	-0.19	0.07	8.33E-03			
cg05575921	5	AHRR (Body)	Shore	1			4 Strong Enhancer	-0.13	0.02	4.05E-10	9.01E-05	-0.25	0.07	3.40E-04			
cg21566642	2	Intergenic	Island		DMR		12 Repressed	-0.12	0.02	3.99E-09	5.92E-04	-0.27	0.07	1.44E-04			
cg05951221	2	Intergenic	Island			1	12 Repressed	-0.12	0.02	2.15E-08	2.40E-03	-0.26	0.07	1.34E-04			
cg06644428	2	Intergenic	Island		DMR		12 Repressed	-0.11	0.02	6.44E-08	5.73E-03	-0.26	0.07	2.04E-04			
cg05119988	4	SC4MOL (5'UTR)	Shelf	1			2 Weak Promoter	-0.08	0.02	1.63E-04		-0.36	0.07	9.83E-08	0.03		
cg19572487	17	RARA (5'UTR)	Shore	1		1	1 Active Promoter	-0.11	0.02	9.89E-08	7.33E-03	-0.20	0.07	4.89E-03			
cg23661483	19	ILVBL (Body)	Shelf				11 Weak Txn	0.11	0.02	6.09E-07	0.02	0.38	0.07	1.28E-07	0.03		
cg14753356	6	Intergenic		1		1	13 Heterochrom/lo	-0.11	0.02	1.43E-07	9.09E-03	-0.15	0.07	0.04			
cg09935388	1	GFI1 (Body)	Island		CDMR	1	11 Weak Txn	-0.11	0.02	1.77E-07	9.84E-03	-0.17	0.07	0.01			
cg05546038	16	NOL3 (TSS1500)	Shore	1			11 Weak Txn	0.11	0.02	2.05E-07	0.01	0.16	0.07	0.03			
cg10549018	10	TLL2 (Body)					7 Weak Enhancer	0.06	0.02	5.08E-03		0.36	0.07	2.28E-07	0.03		
cg21161138	5	AHRR (Body)		1			5 Strong Enhancer	-0.11	0.02	2.78E-07	0.01	-0.21	0.07	2.86E-03			
cg15342087	6	Intergenic		1		1	13 Heterochrom/lo	-0.11	0.02	2.79E-07	0.01	-0.06	0.07	0.36			
cg03295554	11	ETS1 (Body)	Shelf	1		1	5 Strong Enhancer	-0.10	0.02	3.10E-07	0.01	-0.20	0.07	3.22E-03			
cg14026106	10	MARCH8 (Body)		1		1	13 Heterochrom/lo	-0.06	0.02	3.19E-03		-0.36	0.07	3.26E-07	0.03		
cg24859433	6	Intergenic		1		1	13 Heterochrom/lo	-0.11	0.02	3.57E-07	0.01	-0.10	0.07	0.18			
cg00889709	8	FAM110B (5'UTR)				1	13 Heterochrom/lo	0.03	0.02	0.10		0.36	0.07	3.83E-07	0.03		
cg25963130	10	ARID5B (Body)		1		1	5 Strong Enhancer	-0.11	0.02	4.31E-07	0.01	-0.28	0.07	6.80E-05			
cg00021275	2	Intergenic	Shelf			1	9 Txn Transition	0.11	0.02	4.39E-07	0.01	0.23	0.08	2.03E-03			
cg24247537	11	PTDSS2 (Body)	Shore				4 Strong Enhancer	0.06	0.02	0.01		0.37	0.07	4.84E-07	0.03		

cg18857467	1	HDAC1 (Body)	1		1		4 Strong Enhancer	-0.07	0.02	2.01E-03		-0.35	0.07	4.89E-07	0.03
cg16107001	4	C4orf48 (TSS1500)		Island			2 Weak Promoter	0.10	0.02	5.00E-07	0.01	0.10	0.07	0.13	
cg21271420	6	Intergenic					13 Heterochrom/fo	-0.11	0.02	6.49E-07	0.02	-0.05	0.07	0.50	
cg08867471	5	Intergenic	1				13 Heterochrom/fo	-0.12	0.02	9.09E-07	0.02	-0.08	0.08	0.31	
cg03636183	19	F2RL3 (Body)	1	Shore			6 Weak Enhancer	-0.10	0.02	9.17E-07	0.02	-0.16	0.07	0.02	
cg01940273	2	Intergenic	1	Island			12 Repressed	-0.10	0.02	9.21E-07	0.02	-0.22	0.07	1.56E-03	
cg16663980	1	MNDA (5'UTR)					11 Weak Txn	0.07	0.02	1.33E-03		0.34	0.07	1.04E-06	0.06
cg04761231	9	RPL35 (Body)	1	Island			1 Active Promoter	-0.11	0.02	1.23E-06	0.02	-0.27	0.08	3.09E-04	
cg19286329	1	Intergenic	1				4 Strong Enhancer	-0.10	0.02	1.34E-06	0.03	-0.18	0.07	0.01	
cg15501219	3	Intergenic		Shelf			13 Heterochrom/fo	0.13	0.03	1.67E-06	0.03	0.42	0.09	5.15E-06	
cg24574691	11	Intergenic					13 Heterochrom/fo	0.10	0.02	1.83E-06	0.03	0.17	0.07	0.01	
cg00834988	7	AUTS2 (Body)		Shore			13 Heterochrom/fo	0.10	0.02	1.84E-06	0.03	0.29	0.07	4.65E-05	
cg02160023	16	CLCN7 (3'UTR)		Shore			13 Heterochrom/fo	0.02	0.02	0.42		0.33	0.07	2.01E-06	0.10
cg24420089	11	PTDSS2 (Body)		Shore			4 Strong Enhancer	0.05	0.02	0.03		0.34	0.07	2.27E-06	0.10
cg13200854	4	MMRN1 (Body)				1	13 Heterochrom/fo	-0.10	0.02	2.27E-06	0.04	-0.24	0.07	7.39E-04	
cg07033253	7	SRM3 (Body)		Shore			13 Heterochrom/fo	-0.09	0.02	1.05E-05	0.09	-0.33	0.07	2.44E-06	0.10
cg05246522	17	KSR1 (TSS200)	1			1	4 Strong Enhancer	-0.07	0.02	6.58E-04		-0.35	0.07	2.70E-06	0.10
cg09076584	17	FLJ25006 (TSS200)				1	10 Txn Elongation	0.08	0.02	1.24E-04		0.33	0.07	2.74E-06	0.10
cg08958747	16	RAB28 (Body)	1	Shelf			4 Strong Enhancer	-0.10	0.02	2.83E-06	0.04	-0.17	0.07	0.02	
cg23848152	1	NOTCH2NL (5'UTR)		Shore			3 Poised Promoter	-0.10	0.02	2.93E-06	0.04	-0.15	0.07	0.03	
cg12798040	14	XRCC3 (Body)	1			1	10 Txn Elongation	0.07	0.02	7.47E-04		0.32	0.07	3.29E-06	0.10
cg13913475	X	Intergenic	1	Shore		1	8 Insulator	-0.05	0.02	0.01		-0.32	0.07	3.31E-06	0.10
cg13699363	3	EDEM1 (TSS1500)	1	Island		1	1 Active Promoter	-0.04	0.02	0.04		-0.32	0.07	3.62E-06	0.10
cg24736734	5	Intergenic	1			1	7 Weak Enhancer	-0.10	0.02	3.66E-06	0.05	-0.21	0.07	4.02E-03	
cg01763719	2	Intergenic					13 Heterochrom/fo	0.11	0.02	3.70E-06	0.05	0.22	0.08	9.20E-03	

cg03355101	19	Intergenic		1	DMR			2 Weak Promoter	-0.09	0.02	3.99E-06	0.06	0.04	0.07	0.59	
cg10512376	4	EMCN (TSS1500)						13 Heterochrom/lo	-0.10	0.02	5.71E-06	0.08	-0.24	0.07	9.98E-04	
cg22773522	13	DAOA (TSS1500)						13 Heterochrom/lo	-0.09	0.02	6.19E-06	0.08	-0.08	0.07	0.25	
cg18446336	7	GNA12 (Body)		1				11 Weak Txn	-0.09	0.02	6.45E-06	0.08	-0.07	0.07	0.31	
cg10508317	17	SOC3 (Body)	Island	1	DMR			3 Poised Promoter	-0.10	0.02	6.98E-06	0.09	-0.20	0.07	7.12E-03	
cg00294684	11	OR52B6 (1stExon)						13 Heterochrom/lo	-0.09	0.02	7.48E-06	0.09	-0.16	0.07	0.02	
cg02475408	7	Intergenic	Shore					13 Heterochrom/lo	-0.09	0.02	7.74E-06	0.09	-0.18	0.07	7.67E-03	
cg05888181	8	KBTD11 (3'UTR)	Shore				1	13 Heterochrom/lo	0.11	0.02	7.90E-06	0.09	0.14	0.08	0.09	
cg20507228	15	MAN2A2 (Body)						10 Txn Elongation	0.09	0.02	7.91E-06	0.09	0.29	0.07	3.84E-05	
cg03738331	4	HTT (TSS200)	Island	1				1 Active Promoter	-0.09	0.02	8.24E-06	0.09	-0.12	0.07	0.07	
cg01692968	9	Intergenic	Shore	1				6 Weak Enhancer	-0.10	0.02	8.46E-06	0.09	-0.11	0.08	0.15	
cg23840027	1	EPB41 (5'UTR)						11 Weak Txn	0.09	0.02	8.49E-06	0.09	0.15	0.07	0.04	
cg09768249	X	DDX3X (TSS200)	Island	1				1 Active Promoter	0.09	0.02	8.50E-06	0.09	0.003	0.069	0.97	
cg06656152	5	Intergenic						13 Heterochrom/lo	-0.09	0.02	8.69E-06	0.09	-0.08	0.07	0.23	
cg16661609	19	LILRB4 (TSS1500)						5 Strong Enhancer	-0.09	0.02	9.47E-06	0.09	-0.13	0.07	0.08	
cg03072035	10	Intergenic	Shelf				1	13 Heterochrom/lo	0.10	0.02	9.55E-06	0.09	0.17	0.08	0.02	
cg01110839	11	SLC39A13 (3'UTR)		1				11 Weak Txn	0.10	0.02	1.04E-05	0.09	0.23	0.07	1.63E-03	
cg08170227	14	ACTN1 (Body)		1				11 Weak Txn	-0.09	0.02	1.05E-05	0.09	-0.21	0.07	2.14E-03	
cg00913954	1	STK40 (TSS1500)	Shore	1	RDMR			4 Strong Enhancer	0.11	0.02	1.05E-05	0.09	0.14	0.08	0.08	
cg19979108	5	SH3PXD2B (Body)		1				12 Repressed	-0.10	0.02	1.11E-05	0.09	-0.18	0.07	0.02	
cg05249393	19	LDLR (TSS200)	Shore	1			1	4 Strong Enhancer	-0.09	0.02	1.12E-05	0.09	-0.16	0.07	0.02	
cg15344028	2	ICOS (5'UTR)						13 Heterochrom/lo	-0.09	0.02	1.13E-05	0.09	-0.13	0.07	0.06	
cg12912426	12	ATF7IP (5'UTR)						5 Strong Enhancer	-0.09	0.02	1.15E-05	0.09	-0.12	0.07	0.08	
cg05850338	17	Intergenic	Shelf					13 Heterochrom/lo	0.09	0.02	1.16E-05	0.09	0.11	0.07	0.13	

cg185561976	2	ICOS (5'UTR)				1	13 Heterochrom/lo	-0.09	0.02	1.18E-05	0.09	-0.17	0.07	0.01
cg03059174	20	SFRS6 (TSS1500)	Shore	1		1	1 Active Promoter	-0.09	0.02	1.18E-05	0.09	-0.10	0.07	0.14
cg12547807	1	Intergenic		1		1	4 Strong Enhancer	-0.09	0.02	1.19E-05	0.09	-0.20	0.07	3.79E-03
cg17038626	12	CACNA1C (Body)					12 Repressed	-0.09	0.02	1.20E-05	0.09	-0.06	0.07	0.36
cg26712743	10	PYROXD2 (Body)		1		1	4 Strong Enhancer	-0.10	0.02	1.22E-05	0.09	-0.24	0.07	1.25E-03
cg18188448	1	LRRCS2 (1stExon)					12 Repressed	-0.09	0.02	1.27E-05	0.09	-0.17	0.07	0.01
cg09646173	5	PDE6A (Body)		1			13 Heterochrom/lo	-0.10	0.02	1.32E-05	0.09	-0.24	0.07	1.39E-03
cg16006965	3	GCEI2 (TSS1500)		1		1	4 Strong Enhancer	-0.09	0.02	1.34E-05	0.09	-0.07	0.07	0.29
cg08882503	1	NPHP4 (Body)				1	11 Weak Txn	-0.09	0.02	1.38E-05	0.09	-0.12	0.07	0.08
cg16016063	1	Intergenic					13 Heterochrom/lo	0.09	0.02	1.40E-05	0.09	0.15	0.07	0.03
cg03928367	20	FER1L4 (Body)	Shore		CDMR	1	12 Repressed	0.09	0.02	1.45E-05	0.10	0.11	0.07	0.11
cg18050903	4	RASGEF1B (Body)		1		1	13 Heterochrom/lo	-0.09	0.02	1.56E-05	0.10	-0.18	0.07	0.01

Sorted by most significant association (p-value) with carotid plaque or CAC; FDR > 0.10 are left blank
 *TFBS in any cell type from UCSC browser; † DMR: diff methylated region from Illumina (RDMR: reprogramming-specific, CDMR: cancer-specific); ‡DNase HS in any cell type and predicted state from ENCODE.

Table 6. Race/ethnicity and sex-stratified analysis of CpGs with methylation associated with Atherosclerosis

Carotid Plaque Score associated CpG methylation (FDR≤0.05):																
CpG Methylation		African American (N=259)			Caucasian (N=530)			Hispanic (N=370)			Women (N=601)			Men (N=558)		
CpG ID	Location	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val
cg16107001	C4orf48 (TSS1500)	0.10	0.05	0.026	0.14	0.03	2.3E-06	0.05	0.03	0.15	0.06	0.03	0.020	0.14	0.03	9.7E-07
cg21271420	Intergenic	-0.07	0.04	0.10	-0.12	0.03	9.4E-05	-0.09	0.04	0.029	-0.09	0.03	2.2E-03	-0.12	0.03	1.6E-04
cg23848152	NOTCH2NL (5'UTR)	-0.13	0.05	5.4E-03	-0.09	0.03	1.2E-03	-0.07	0.03	0.061	-0.05	0.03	0.065	-0.15	0.03	1.6E-06
cg08867471	Intergenic	-0.04	0.05	0.37	-0.17	0.04	2.2E-06	-0.07	0.04	0.11	-0.13	0.03	7.6E-05	-0.10	0.03	5.3E-03
cg05575921	AHRR (Body)	-0.05	0.06	0.38	-0.17	0.04	1.0E-05	-0.09	0.05	0.077	-0.05	0.04	0.20	-0.16	0.04	7.1E-06
cg25953130	ARID5B (Body)	-0.08	0.05	0.073	-0.13	0.03	6.3E-05	-0.04	0.04	0.31	-0.10	0.03	7.3E-04	-0.09	0.03	5.9E-03
cg06126421	Intergenic	-0.09	0.06	0.10	-0.14	0.03	4.4E-05	-0.05	0.04	0.19	-0.06	0.03	0.080	-0.16	0.03	5.8E-06
cg05546038	NOL3 (TSS1500)	0.10	0.04	0.022	0.11	0.03	1.2E-03	0.09	0.04	0.014	0.08	0.03	7.7E-03	0.11	0.03	4.0E-04
cg24574691	Intergenic	0.07	0.03	0.035	0.09	0.05	0.072	0.09	0.03	9.8E-04	0.08	0.02	1.8E-04	0.08	0.04	0.036
cg09935388	GFI1 (Body)	-0.11	0.05	0.035	-0.11	0.03	1.6E-04	-0.03	0.04	0.50	-0.08	0.03	9.5E-03	-0.10	0.03	1.6E-03
cg01763719	Intergenic	0.13	0.05	0.019	0.10	0.04	7.9E-03	0.10	0.04	0.017	0.15	0.03	2.1E-05	0.07	0.03	0.038
cg00021275	Intergenic	0.12	0.06	0.032	0.09	0.03	7.7E-03	0.10	0.04	8.7E-03	0.11	0.03	2.4E-04	0.07	0.03	0.025
cg23661483	ILVBL (Body)	0.07	0.05	0.13	0.06	0.03	0.058	0.12	0.04	7.7E-04	0.11	0.03	1.7E-04	0.07	0.03	0.022
cg08958747	RAB26 (Body)	-0.10	0.05	0.049	-0.09	0.03	8.0E-03	-0.08	0.04	0.039	-0.08	0.03	9.9E-03	-0.10	0.03	1.5E-03
cg13200854	MMRN1 (Body)	-0.08	0.05	0.11	-0.11	0.03	1.4E-04	-0.06	0.04	0.083	-0.04	0.03	0.14	-0.13	0.03	1.2E-05
cg24736734	Intergenic	-0.14	0.05	3.1E-03	-0.07	0.03	0.023	-0.11	0.04	7.6E-03	-0.11	0.03	9.3E-05	-0.07	0.03	0.024
cg19572487	RARA (5'UTR)	-0.04	0.05	0.46	-0.12	0.03	1.9E-04	-0.06	0.04	0.14	-0.05	0.03	0.11	-0.12	0.03	7.6E-05
cg03295554	ETS1 (Body)	-0.08	0.04	0.061	-0.09	0.03	2.6E-03	-0.08	0.04	0.028	-0.10	0.03	3.1E-04	-0.06	0.03	0.045
cg19266329	Intergenic	-0.14	0.05	3.2E-03	-0.10	0.03	1.4E-03	-0.02	0.04	0.68	-0.07	0.03	0.015	-0.10	0.03	1.3E-03
cg04761231	RPL35 (Body)	-0.06	0.05	0.22	-0.12	0.03	2.2E-04	-0.05	0.04	0.27	-0.07	0.03	0.016	-0.11	0.03	1.4E-03
cg21566642	Intergenic	-0.07	0.06	0.20	-0.11	0.03	2.1E-03	-0.06	0.04	0.16	-0.06	0.03	0.067	-0.12	0.03	5.0E-04
cg15501219	Intergenic	0.13	0.06	0.037	0.06	0.04	0.13	0.13	0.04	4.0E-03	0.11	0.04	7.0E-03	0.10	0.04	4.6E-03
cg15342087	Intergenic	-0.09	0.04	0.039	-0.09	0.03	9.9E-03	-0.03	0.04	0.42	-0.07	0.03	0.035	-0.09	0.03	3.7E-03
cg00834988	AUTS2 (Body)	0.05	0.04	0.24	0.06	0.03	0.043	0.10	0.03	2.8E-03	0.08	0.03	6.4E-03	0.07	0.03	0.024
cg24859433	Intergenic	-0.08	0.05	0.077	-0.09	0.03	8.0E-03	-0.02	0.04	0.62	-0.07	0.03	0.029	-0.09	0.03	7.5E-03
cg14753356	Intergenic	-0.06	0.05	0.28	-0.11	0.03	1.5E-03	-0.04	0.04	0.32	-0.03	0.03	0.42	-0.13	0.03	5.9E-05
cg21161138	AHRR (Body)	-0.13	0.05	8.8E-03	-0.09	0.03	9.9E-03	-0.03	0.04	0.54	-0.05	0.03	0.13	-0.11	0.03	1.4E-03
cg05951221	Intergenic	-0.04	0.06	0.48	-0.09	0.03	5.9E-03	-0.05	0.04	0.23	-0.06	0.03	0.077	-0.10	0.03	5.9E-03

cg06644428	Intergenic	-0.07	0.05	0.15	-0.06	0.03	0.063	-0.07	0.04	0.086	-0.04	0.03	0.25	-0.10	0.03	1.4E-03
cg03636183	F2RL3 (Body)	0.00	0.04	0.95	-0.13	0.04	4.1E-04	-0.03	0.04	0.40	-0.03	0.03	0.35	-0.10	0.03	1.2E-03
cg01940273	Intergenic	-0.08	0.05	0.16	-0.09	0.03	9.5E-03	-0.02	0.04	0.59	-0.04	0.03	0.27	-0.10	0.03	2.8E-03
Coronary artery calcium associated CpG methylation (FDR≤0.05):																
CpG ID	Location	African American (N=258)			Caucasian (N=542)			Hispanic (N=385)			Women (N=611)			Men (N=574)		
		Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val	Beta	SE	P-val
cg18857467	HDAC1 (Body)	-0.25	0.17	0.15	-0.31	0.09	4.6E-04	-0.18	0.13	0.16	-0.26	0.09	4.0E-03	-0.23	0.10	0.022
cg05119988	SC4MOL (5'UTR)	-0.40	0.14	5.9E-03	-0.23	0.10	0.016	-0.07	0.13	0.56	-0.17	0.09	0.054	-0.21	0.10	0.037
cg00889709	FAM110B (5'UTR)	0.19	0.13	0.15	0.21	0.11	0.062	0.35	0.12	3.5E-03	0.17	0.10	0.080	0.30	0.10	2.8E-03
cg14026106	42071 (Body)	-0.22	0.12	0.076	-0.25	0.11	0.021	-0.40	0.14	4.4E-03	-0.20	0.09	0.035	-0.30	0.10	4.1E-03
cg10549018	TLL2 (Body)	0.20	0.15	0.20	0.26	0.10	0.011	0.42	0.12	3.7E-04	0.35	0.09	1.5E-04	0.26	0.10	0.011
cg24247537	PTDSS2 (Body)	0.47	0.14	1.3E-03	0.22	0.11	0.039	0.34	0.13	8.3E-03	0.44	0.10	6.7E-06	0.18	0.11	0.084
cg23661483	ILVBL (Body)	0.36	0.16	0.029	0.23	0.10	0.021	0.36	0.13	4.9E-03	0.33	0.10	5.5E-04	0.27	0.10	8.7E-03
Sorted by most significant association (p-value) with carotid plaque burden or CAC																

CpG methylation characteristics				CpG methylation ~ Cis-gene expression (N=1,264)				mRNA expression ~ Carotid Plaque Score (N=1,181)				mRNA expression ~ CAC (N=1,208)					
CpG Illumina ID	chromosome	Location	CD14+ in silico predicted state*	CD14+ DNase HS †	~ Carotid plaque (CP) or CAC FDR<0.05	Gene	corr†	P-value	FDR	Beta	SE	P-value	FDR	Beta	SE	P-value	FDR
cg25953130	10	ARID5B (Body)	5	1	CP	ARID5B	-0.21	1.3E-08	5.7E-06	0.12	0.02	6.3E-08	1.1E-03	0.32	0.08	2.5E-05	0.034
cg05119988	4	SC4MOL (5'UTR)	2		CAC	SC4MOL	-0.17	1.2E-10	7.3E-08	0.07	0.02	1.4E-03	0.31	0.17	0.07	0.014	0.31
cg05546038	16	NOL3 (TSS1500)	6	1	CP	DPEP2	-0.21	1.5E-11	1.0E-08	-0.06	0.02	7.6E-03	0.41	-0.01	0.07	0.91	0.82
cg05546038	16	NOL3 (TSS1500)	6	1	CP	ZDHHC1	-0.19	1.5E-10	9.1E-08	-0.05	0.02	0.030	0.50	-0.03	0.07	0.67	0.77
cg05546038	16	NOL3 (TSS1500)	6	1	CP	ATP6V0D1	0.08	2.4E-06	6.5E-04	0.03	0.02	0.12	0.64	0.04	0.07	0.63	0.76
cg05546038	16	NOL3 (TSS1500)	6	1	CP	NOL3	-0.18	2.2E-09	1.1E-06	-0.01	0.02	0.53	0.86	0.05	0.07	0.49	0.71
cg05546038	16	NOL3 (TSS1500)	6	1	CP	DYNC1LI2	-0.12	1.6E-06	4.4E-04	-0.01	0.02	0.75	0.90	0.01	0.07	0.91	0.82
cg24736734	5		2		CP	PAM	-0.13	1.2E-08	5.1E-06	0.04	0.02	0.066	0.56	0.09	0.07	0.19	0.57
cg00889709	8	FAM110B (5'UTR)	3	1	CAC	FAM110B	-0.14	3.8E-08	1.5E-05	-0.01	0.02	0.66	0.89	-0.02	0.07	0.81	0.80

Table includes CpGs with significant (FDR<0.05) associations between CD14+ CpG methylation and atherosclerosis traits carotid plaque or coronary artery calcium (CAC), and cis (± 1 MB) gene expression profiles (in CD14+ samples from MESA, FDR<0.001)

*States predicted using ChromHMM with CD14+ histone data from BLUEPRINT; state annotations are 1: transcribed, 2: heterochromatin/low abundance, 3: weak/poised enhancer, 4: strong enhancer, 5: active promoter, 6: weak promoter

† DNase hypersensitive sites in a CD14+ sample from BLUEPRINT

‡ correlation between methylation and cis (± 1MB) mRNA expression profiles

Table 8. Atherosclerosis-associated gene expression in monocytes associations with CVD risk factors*											
Gene	~ Carotid plaque (CP) or CAC FDR<0.05	Beta direction (+/-) or P-val	Age	Sex (male)	Race (Black)	Race (Hispanic)	Current smoker	BMI	HDL cholesterol	Hypertension	Diabetes
CACNA2D3	CAC	+ - P-val	- 1.81E-10		- 7.75E-25	- 3.32E-15		- 8.47E-12	+ 5.33E-04	- 4.96E-05	- 5.55E-05
ARID5B	CP, CAC	+ - P-val	+ 2.06E-13	+ 4.39E-02	+ 2.89E-06	+ 8.96E-04	+ 4.18E-02	+ 1.67E-08	- 6.86E-06	+ 7.20E-03	+ 3.43E-05
PDLIM7	CP, CAC	+ - P-val		- 6.85E-04		+ 1.19E-02	+ 1.40E-03	+ 9.14E-14	- 1.46E-08	+ 5.02E-03	+ 2.94E-04
LILRB4	CAC	+ - P-val		- 4.32E-07		+ 1.13E-03			- 4.84E-02	+ 6.38E-03	
P2RX4	CAC	+ - P-val		- 8.79E-03	+ 5.50E-10	+ 2.36E-06		+ 1.14E-03	- 3.05E-02	+ 2.37E-02	+ 8.01E-04
FAM168B	CAC	+ - P-val		+ 1.94E-09				- 2.94E-02			
ADM	CAC	+ - P-val		- 1.04E-13	+ 4.97E-06	+ 1.11E-03		+ 1.12E-08	- 9.21E-03	+ 3.11E-02	
IARS	CAC	+ - P-val	- 2.60E-04	+ 1.47E-11		- 3.52E-03		- 7.61E-06	+ 2.78E-02	- 4.45E-02	- 1.01E-02
* Adjusting for demographic variables; sorted by most significant association with carotid plaque score or CAC; non-significant (p ≥ 0.05) associations are left blank; three genes bolded had genome-wide FDR<0.05 for a direction of association with the majority of CVD risk factors similar to those expected based on the mRNA expression association with atherosclerosis and atherosclerosis associations with risk factors: LDL cholesterol was not associated with any atherosclerosis-associated gene expression.											

* Adjusting for demographic variables; sorted by most significant association with carotid plaque score or CAC; non-significant ($p \geq 0.05$) associations are left blank; three genes bolded had genome-wide FDR<0.05 for a direction of association with the majority of CVD risk factors similar to those expected based on the mRNA expression association with atherosclerosis and atherosclerosis associations with risk factors; LDL cholesterol was not associated with any atherosclerosis-associated gene expression.

Table 9. Atherosclerosis-associated CpG methylation in monocytes associations with CVD risk factors*											
CpG Illumina ID and nearby gene location	~ Carotid Plaque (CP) or CAC (FDR≤0.05)	Beta direction (+/-) or P-value	Age	Sex (male)	Race (Black)	Race (Hispanic)	Current smoker	BMI	HDL cholesterol	Hypertension	Diabetic
cg06126421 Intergenic	CP	+ P-val	- 8.51E-12	- 2.67E-14		+ 4.83E-08	- 1.47E-42				
cg05575921 AHRR (Body)	CP	+ P-val		- 1.18E-11	- 2.26E-02	+ 4.50E-02	- 1.59E-110	+ 6.23E-03			
cg21566642 Intergenic	CP	+ P-val	- 2.29E-02	- 1.61E-07		+ 1.24E-03	- 1.74E-74				
cg05951221 Intergenic	CP	+ P-val		- 2.47E-04			- 5.99E-71				
cg06644428 Intergenic	CP	+ P-val	+ 2.76E-02				- 4.07E-30				
cg05119988 SC4MOL (5'UTR)	CAC	+ P-val						- 1.09E-04	+ 1.09E-06	- 4.99E-03	- 1.27E-09
cg19572487 RARA (5'UTR)	CP	+ P-val	- 7.24E-06	- 7.28E-06	+ 5.59E-04	+ 2.01E-05	- 1.20E-25				
cg23661483 ILVBL (Body)	CP, CAC	+ P-val	+ 2.78E-09		- 1.30E-03	- 2.49E-07	+ 3.60E-04				+ 2.76E-03
cg14753356 Intergenic	CP	+ P-val	- 5.08E-12	- 3.86E-06		+ 1.09E-04	- 8.09E-32				
cg09935388 GFI1 (Body)	CP	+ P-val				+ 2.71E-03	- 1.84E-40				
cg05546038 NOL3 (TSS1500)	CP	+ P-val		+ 9.11E-08	- 4.96E-04		+ 1.59E-08				
cg10549018 TLL2 (Body)	CAC	+ P-val	+ 2.09E-10	+ 8.23E-06	- 2.92E-04		+ 2.03E-03				
cg21161136 AHRR (Body)	CP	+ P-val		- 4.20E-04		+ 8.06E-04	- 3.01E-59	+ 9.39E-03			
cg15342087 Intergenic	CP	+ P-val	- 2.35E-06	- 1.74E-06	- 3.70E-05		- 3.21E-40				
cg03295554 ETS1 (Body)	CP	+ P-val				+ 3.50E-02	- 1.54E-09	- 4.72E-03		- 5.11E-03	- 1.99E-03
cg14026106 MARCB1 (Body)	CAC	+ P-val	- 1.66E-30							- 6.84E-03	
cg24859433 Intergenic	CP	+ P-val	- 1.26E-06	- 1.32E-10	- 1.63E-05		- 5.88E-34				- 3.68E-02
cg00889709	CAC	+ P-val	+ 2.78E-09	+ 9.11E-08	- 4.96E-04		+ 1.59E-08	+ 9.39E-03	- 5.11E-03	+ 2.76E-03	+ 1.27E-09

FAM110B (5'UTR)		P-val	1.36E-03	6.08E-04	4.11E-15	3.69E-02	7.34E-07	2.27E-03	2.20E-02	4.10E-07
cg25953130	CP	+ -	-	+	-	-		+		
ARID5B (Body)		P-val	3.33E-11	1.14E-06	1.10E-04	2.25E-02		3.91E-03		
cg00021275	CP	+ -	+	+		+	+	-	+	+
Intergenic		P-val	2.96E-04	3.03E-13		2.50E-02	8.31E-09	1.36E-03	1.14E-03	1.28E-06
cg24247537	CAC	+ -	+	+	+	+				+
PTDS2 (Body)		P-val	1.89E-08	4.80E-05	9.08E-16	2.05E-05				9.70E-04
cg21271420	CP	+ -	-	-		+				
Intergenic		P-val	9.67E-27	7.86E-07		3.41E-04				
cg0867471	CP	+ -		-		-				
Intergenic		P-val		5.76E-06		6.96E-07				
cg03636183	CP	+ -		-	-	-	+			-
F2RL3 (Body)		P-val		2.42E-03	3.23E-09	8.48E-45	4.73E-02			1.39E-02
cg01940273	CP	+ -	-	-		-				
Intergenic		P-val	4.84E-06	2.82E-10		1.09E-62				
cg04761231	CP	+ -	-	+	+	-				
RPL35 (Body)		P-val	2.50E-04	1.46E-48	6.44E-07	1.18E-04				
cg19266329	CP	+ -	-	-	-	-				-
Intergenic		P-val	1.10E-19	1.72E-08	1.60E-03	1.45E-06				1.80E-05
cg15501219	CP	+ -	+	-		+	+			+
Intergenic		P-val	9.82E-05	1.17E-11		3.69E-04	7.31E-05			5.80E-06
cg24574691	CP	+ -		+	-	-			+	
Intergenic		P-val		1.21E-02	3.91E-06	2.11E-02			5.72E-03	
cg00834968	CP	+ -		+		+			+	+
AUTS2 (Body)		P-val		3.87E-02		2.38E-10			2.05E-02	2.68E-03
cg13200854	CP	+ -	-	-		-			-	
MMFN1 (Body)		P-val	4.70E-14	6.14E-03		8.55E-03			1.54E-02	
cg08958747	CP	+ -	-	+	+	-				
RAB26 (Body)		P-val	7.46E-29	3.70E-07	2.56E-05	1.15E-03				
cg01763719	CP	+ -				+	+			
Intergenic		P-val				1.26E-06	1.24E-02			

* Adjusting for demographic variables; sorted by most significant association with carotid plaque score or CAC; non-significant ($p \geq 0.05$) associations are left blank; 29 CpGs bolded had genome-wide FDR < 0.05 for a direction of association with the majority of CVD risk factors similar to those expected based on the methylation association with atherosclerosis and associations with risk factors; LDL cholesterol was not associated with any atherosclerosis-associated CpG methylation.

Table 10. siRNA Knockdown of ARID5B Alters Immune/Inflammatory Response and Lipid Metabolism Genes.		
Terms		
Bio-Functions	FDR	
Inflammatory response	1.10E-06	Down-regulated genes ABCD1, ABHD12, ABR, ADA, ADORA2B, ADORA3, ADRBK1, AGT, AKT1, ALOX5, APOE, ATF3, ATM, AXL, BCL6, C3AR1, C5AR1, CASP1, CASP4, CCL14, CCL3, CCL3L1, CCL4, CCRL2, CD14, CD1D, CD36, CD4, CD74, CD9, CDC42, CIRBP, CKLF, CLCF1, CLCN7, CORO1A, CSF1R, CST3, CTSS, CTSG, CXCL2, CYBB, CYSLTR1, DUSP10, EDN1, EGR1, EIF2AK1, ELANE, ELMO1, ENG, EPAS1, ERN1, F11R, FABP4, FCGR1A, FCGR2A, FCGRT, FGR, FOXF1, FPR1, GNA13, GNAI2, GRN, HAVCR2, HDAC4, HEBP1, HLA-DMA, HLA-DRA, HLA-DRB1, HLA-DRB5, HLA-DRB6, HMOX1, HPSE, IFNB1, IL1A, INPP5D, IRF3, ITGB1, KLF4, LAT, LGALS9, LILRB3, LIPA, LTB4R, LY96, MCL1, MDK, MPP1, MR1, MYO1F, MYO9B, NCKAP1L, NFIL3, NFKB1Z, NLRP3, NR4A2, OSTM1, PARP4, PDE4B, PECAM1, PIK3CB, PLA2G7, PLAU, PLCG2, PON2, PPARG, PPT1, PRDM1, PRKCA, PRNP, PROCR, PRTN3, PSEN2, PSMB8, PSMB9, PTGER2, PTGER4, PTGS1, PTGS2, PTPN11, PTPRO, PYCARD, RAC2, RGS1, RHOB, RIPK3, RNASE2, RPL13A, SERPINF1, SIGIRR, SLC11A1, SLC1A5, SMAD7, SPHK1, STAT1, STAT2, STAT6, STK39, SYK, TBXAS1, TGM2, TLR4, TNF, TNFRSF14, TNFSF14, TREM2, TYROBP, ULBP1, WAS, ZFP36 Up-regulated genes ABCF1, AHR, ALOX5AP, ANGPTL4, AQP9, BID, BRD2, CCND1, CD40, CD44, CEBPD, CXCL6, E2F2, FEM1A, FYN, GJA1, HDAC9, HLA-H, HNRPA0, ICAM1, IDO1, IFNAR1, IKBKB, IL10, IL10RA, IL18RAP, IL23A, JMJD6, LILRB3, LRIG2, LYN, LYST, MAPK7, MGLL, MIF, NCL, NFE2L1, NFKB1, NFX1, NOD2, NOTCH1, OLR1, ORM1, P2RY6, PGLYRP2, PIK3CD, PPM1D, PRKCD, PTX3, S100A12, S100A8, S100A9, SBNO2, SFRP1, SIRPA, SLC33A1, SLC7A2, SMAD3, SRC, STAT4, STAT5A, TICAM2, TNFAIP8L2, TNFRSF1A, TNFRSF1B, TNFRSF4, TNFRSF9, TNIP1, TRADD, TRAF3IP2, TRNT1, TSC1, VAV3
Chemotaxis of leukocytes	1.38E-02	Down-regulated genes ADORA3, ALOX5, APOE, C3AR1, C5AR1, CCL14, CCL3, CCL3L1, CCL4, CCRL2, CD4, CD74, CD9, CDC42, CKLF, CORO1A, CSF1R, CST3, CTSG, CXCL2, CYBB, CYSLTR1, EDN1, ELANE, ELMO1, FCGR2A, FGR, FPR1, GNA13, GNAI2, GRN, HEBP1, HPSE, IFNB1, IL1A, INPP5D, LAT, LGALS9, LILRB3, LTB4R, MDK, MPP1, MYO1F, MYO9B, NCKAP1L, PDE4B, PIK3CB, PLA2G7, PLAU, PLCG2, PON2, PPARG, PRKCA, PRNP, PRTN3, PTPN11, PTPRO, RAC2, RGS1, RHOB, RNASE2, RPL13A, SPHK1, STAT6, SYK, TLR4, TNF, TNFRSF14, TNFSF14, WAS Up-regulated genes AQP9, CCND1, CD40, CXCL6, FYN, ICAM1, IKBKB, IL10, IL23A, JMJD6, LYN, LYST, MIF, NOD2, PIK3CD, S100A12, S100A8, S100A9, SMAD3, TNFRSF1A, TRAF3IP2, TSC1, VAV3
Migration of phagocytes	1.69E-03	Down-regulated genes ADAM15, AGT, APOE, BMPR2, CCL3, CCL4, CD209, CD74, CD86, CDC42, CSF1R, CTSG, CTSS, CXCL2, CYBB, DOCK8, ENG, EPAS1, F11R, FGR, FLNA, FPR1, HCLS1, IFNB1, INPP5D, ITGAX, ITGB1, MDK, MYO9B, NAB2, NARS, NLRP3, PDE4B, PECAM1, PLAU, PLCB2, PPARG, PRNP, PTGDS, PTGS2, RAC2, RHOB, RTN4, SEMA4D, SERPINF1, TIMP2, TIMP3, TLN1, TLR4, TNF, TREM2, TYROBP, VCAN, WAS Up-regulated genes ABL1, CCND1, CD40, CD44, CD58, CD80, CD99, CXCL6, ICAM1, ICAM2, IL10, IL10RA, MIF, P2RY6, PIK3CD, PPM1D, PVR, QPCT, S100A8, S100A9, SH2B3, SIRPA, TNFAIP8L2, TNFRSF1A, TNFRSF9, VAV3

Phagocytosis of blood cells	3.00E-03	BIN1, BTK, CD14, CD36, CD93, CDC42, CERK, CLCN3, CORO1A, CORO1C, CYP2S1, DEF6, EIF2AK1, FCGR1A, FCGR1B, FCGR1C, FCGR2A, FGR, FPR1, GAB2, GAS6, GLRX, HMOX1, INPP5D, ITGAX, MERTK, PIK3CB, PLAU, RAC2, STIM1, SYK, TGM2, TLR4, TNF, TREM2, UCP2, WAS	CD44, FYN, IFNAR1, IRF8, JMD6, LYN, MIF, NOD2, PTX3, S100A9, SIRPA, SRC
Apoptosis of leukocytes	6.79E-04	ABCG1, ADA, ADORA3, AKT1, ANPEP, APOE, ASAH1, ATF3, ATM, BTK, C3AR1, CASP1, CASP4, CAT, CCL3, CCL4, CD14, CD4, CD9, CDC42, CORO1A, CXCL2, CYBB, DOCK8, EGR1, EGR3, ETS2, FCGR1A, FGL2, FOSL1, GADD45B, GPR132, HAVCR2, IFNB1, IL1A, INPP5D, IRF3, IRF4, IRS2, ITGB1, ITPR1, KLF4, LGALS9, MAP2K5, MAP4K1, MCL1, MERTK, MYC, NFATC1, NR4A1, NR4A3, PECAM1, PKN1, PLCG2, PPARG, PRKCA, PRTN3, PTGER4, RAC2, RCAN1, RIPK3, SERPINF1, SLC39A10, SMAD7, SPHK1, ST14, STAT1, STAT6, TLR4, TMEM109, TNF, TNFRSF21, TNFSF14, TOP2A, WAS	ABL1, AHR, BCL11A, BCL2, BCL2L1, BCL3, BID, BMF, CASP2, CCDC86, CD40, CD44, CD70, CD80, CDKN1B, CREB1, DLL4, DNAJA3, E2F2, FBXW7, FYN, ICAM1, IDO1, IKBKB, IKBKE, IL10, IL7R, IRF8, LILRB1, LYN, MIF, MYB, NFKB1B, NOTCH1, OGT, PERP, PIK3CD, PIM1, POLR3E, PPM1D, PRELID1, PRKCD, RCAN2, S100A8, SERPINE9, SMAD3, SOD2, ST6GAL1, STAT5A, TICAM2, TNFAIP8L2, TNFRSF1A, TNFRSF1B, TNFRSF4, TNFRSF9, TNFSF10, TNFSF13B, TRAF3IP2, TSC1, TSC22D3, VAV3
Synthesis of lipid	2.10E-04	ABCG1, ACSL3, ACSS2, ADA, ADORA2B, ADRBK1, AGPAT1, AGPAT2, AGT, AKT1, ALDH1A2, ALOX5, ANPEP, APOC1, APOC2, APOE, ARV1, ASAH1, ATF3, B3GNT5, BMP2, BMPR2, BRE, C5AR1, CASP1, CD14, CD209, CD36, CD4, CD74, CD9, CDC42, CDIPT, CERK, CHKB, CLN3, CYP1B1, DAB2, DGAT1, DHRS9, DOLK, EDN1, EEF1A1, EGR1, ERLIN2, FABP5, FAR1, FCGR1A, FCGR2A, FGR, FOSL1, FOXA1, FOXO4, FPR1, GBA, GBGT1, GDF15, GNPAT, HMGCS1, HMOX1, HOXA10, IL1A, INPP5D, INPPL1, INSIG1, ITGB1, LTA4H, LTB4R, LTC4S, MAPK3, ME2, MGST2, MITF, MKNK1, MYC, NPC2, NR4A1, NR4A2, NR4A3, NSMAF, PDHB, PDSS1, PDSS2, PIGB, PIGP, PIGY, PIK3C2A, PIK3CB, PISD, PITPNM1, PLAU, PLCG2, PMVK, PNPLA2, PNPLA6, PON2, POR, PPARG, PRKAA1, PRKAG2, PRKCA, PRNP, PTGDS, PTGER2, PTGS1, PTGS2, PTPN11, RBP1, RGS2, RXRA, SC5DL, SCD, SH3KBP1, SIGIRR, SLC9A3R1, SPHK1, SPTLC1, ST3GAL5, SYK, TBXAS1, TLR4, TNF, TRERF1, TRIB3	ACAT2, ACSL5, AHR, AKR1B1, ALOX5AP, ATP5J, BCL2, BCL2L1, BMP6, BNIP3, BRD2, C1QTNF1, CD40, CD82, CES1, CREB1, CYP7B1, DEGS1, FAF2, FYN, IGFBP4, IL10, IL18RAP, INSIG2, LYN, MAP3K1, MAP3K8, MCAT, ME1, MIF, NR5A2, OLR1, OSBP, PEX11B, PGS1, PIAS1, PIGV, PIK3C2B, PIK3CD, PPAT, PRDX2, PRKCD, PTPMT1, PTX3, RDH10, SLC1A3, SLC35A2, SMAD3, SMAD4, STAT5A, TNFRSF1A, TNFRSF1B, TNFSF10, UGCG
Canonical pathways	FDR		

Leukocyte Extravasation Signaling	8.31E-03	ACTB, ACTG1, ACTN1, ACTN4, ARHGAP4, ARHGAP5, ARHGAP9, ATM, BTK, CDC42, CLDN1, CLDN23, CTNNA1, CYBB, F11R, GNAI2, ITGB1, MAP3K4, MAPK13, MMP10, NCF2, NCF4, PECAM1, PIK3C2A, PIK3CB, PIK3R2, PLCG2, PRKCA, PRKCH, PTPN11, RAC2, SIPA1, TIMP2, TIMP3, VCL, WAS	ABL1, CD44, CD99, CRKL, ICAM1, PIK3C2B, PIK3CD, PRKCD, PTK2, RASGRP1, RDX, SRC, VAV3
Enrichment from IPA ingenuity, including genes significantly (FDR<0.05) altered after siARID5B in in LPS stimulated human THP1-monocytes, compared to a background of all genes with detectable expression			

Example 2: Risk Stratification Role of ARID5B Expression in Type II Diabetes and Atherosclerosis

Using the data reported in the Example 1, Figure 11A gives the odds ratio of type II diabetes (T2DM) according to ARID5B levels after prior stratification of persons according to their overweight and obesity status, the strongest risk factor for T2DM. Tertiles of ARID5B were used in these analyses to ensure an adequate representation of ARID5B levels for all strata under investigation. The ARID5B tertiles showed a strong effect in overweight and obesity subgroups, subjects with diabetes more frequently had ARID5B values in the medium and highest tertile (Odds ratio for obese subjects with the highest tertile of ARID5B was 11, compared to those with normal weight and the lowest tertile of ARID5B, p values: $1.1E-6$)

Figure 11B showed the mean carotid plaque scores (log transformed) according to ARID5B levels after prior stratification of persons according to their impaired glucose tolerance and diabetes status, one of the strongest risk factors for atherosclerotic CVD. The ARID5B tertiles showed a significant ($p < 0.001$) effect at normal, impaired glucose tolerance, and diabetes subgroups).

In conclusion, the type II diabetes and atherosclerosis risk increases with the higher levels of ARID5B in all the low, medium, and high risk groups. ARID5B significantly improves the risk stratifications for both T2DM and atherosclerosis, above and beyond their known risk factors.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of skill in the art to which the disclosed invention belongs. Publications cited herein and the materials for which they are cited are specifically incorporated by reference.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

WHAT IS CLAIMED IS:

1. A method, comprising obtaining a sample from a subject having or suspected of having an inflammatory disease, and assaying the sample for methylation of an *ARID5B* CpG site affecting *ARID5B* gene expression.
2. The method of claim 1, wherein the CpG methylation site comprises a CpG site found in the nucleic acid sequence (SEQ ID NO:1).
3. The method of claim 2, wherein the CpG methylation site comprises cg25953130.
4. A method for diagnosing or prognosing an inflammatory disease in a subject, comprising assaying a sample from the subject for *ARID5B* expression levels, wherein detection of elevated *ARID5B* expression levels compared to a control is an indication that the subject has an inflammatory disease.
5. The method of claim 4, wherein the *ARID5B* expression levels are assayed by assaying for methylation of an *ARID5B* CpG site that negatively correlates with *ARID5B* gene expression.
6. The method of claim 5, wherein the CpG methylation site is found in the nucleic acid sequence SEQ ID NO:1.
7. The method of claim 6, wherein the CpG methylation site comprises cg25953130.
8. The method of any one of claims 4 to 6, wherein the inflammatory disease is diabetes mellitus type 2 or a cardiovascular disease.
9. The method of claim 8, wherein the cardiovascular disease comprises atherosclerosis.
10. A method for treating an inflammatory disease in a subject, comprising
obtaining a sample from a subject having or suspected of having an inflammatory disease,
assaying the sample for *ARID5B* expression levels, and
treating the subject for an inflammatory disease if the *ARID5B* expression levels are elevated compared to a control.
11. The method of claim 10, wherein the *ARID5B* expression levels are assayed by assaying for methylation of an *ARID5B* CpG site that negatively correlates with *ARID5B* gene expression.
12. The method of claim 11, wherein the CpG methylation site is found in the nucleic acid sequence SEQ ID NO:1.
13. The method of claim 12, wherein the CpG methylation site comprises cg25953130.

14. The method of any one of claims 10 to 13, wherein the subject is at high risk for cardiovascular disease, wherein the subject is treated for cardiovascular disease if *ARID5B* expression levels are elevated.
15. The method of any one of claims 10 to 13, wherein the subject is at high risk for diabetes mellitus type 2, wherein the subject is treated for diabetes if *ARID5B* expression levels are elevated.
16. The method of claim 14 or 15, wherein the subject is treated with an anti-inflammatory agent if *ARID5B* expression levels are elevated.
17. A method for treating a subject with an inflammatory disease, comprising administering to the subject a composition comprising an *ARID5B* inhibitor in an amount effective to reduce endogenous expression levels of *ARID5B* in the subject.
18. The method of claim 17, wherein the *ARID5B* inhibitor is a gene silencing agent.
19. The method of claim 18, wherein the gene silencing agent comprises an antisense oligonucleotide, siRNA, or shRNA.
20. The method of claim 17, wherein the *ARID5B* inhibitor is a CRISPR-Cas system that methylates an *ARID5B* CpG site that negatively correlates with *ARID5B* gene expression.
21. The method of claim 20, wherein the CpG methylation site is found in the nucleic acid sequence SEQ ID NO:1.
22. The method of claim 21, wherein the CpG methylation site comprises cg25953130.
23. The method of any one of claims 17 to 22, wherein the inflammatory disease is diabetes mellitus type 2 or a cardiovascular disease.
24. The method of claim 23, wherein the cardiovascular disease comprises atherosclerosis.
25. The method of any one of claims 17 to 24, wherein prior to treatment the method further comprises detecting elevated *ARID5B* expression levels in a sample from the subject.
26. The method of claim 25, wherein the *ARID5B* expression levels are detected by assaying for methylation of an *ARID5B* CpG site that negatively correlates with *ARID5B* gene expression.

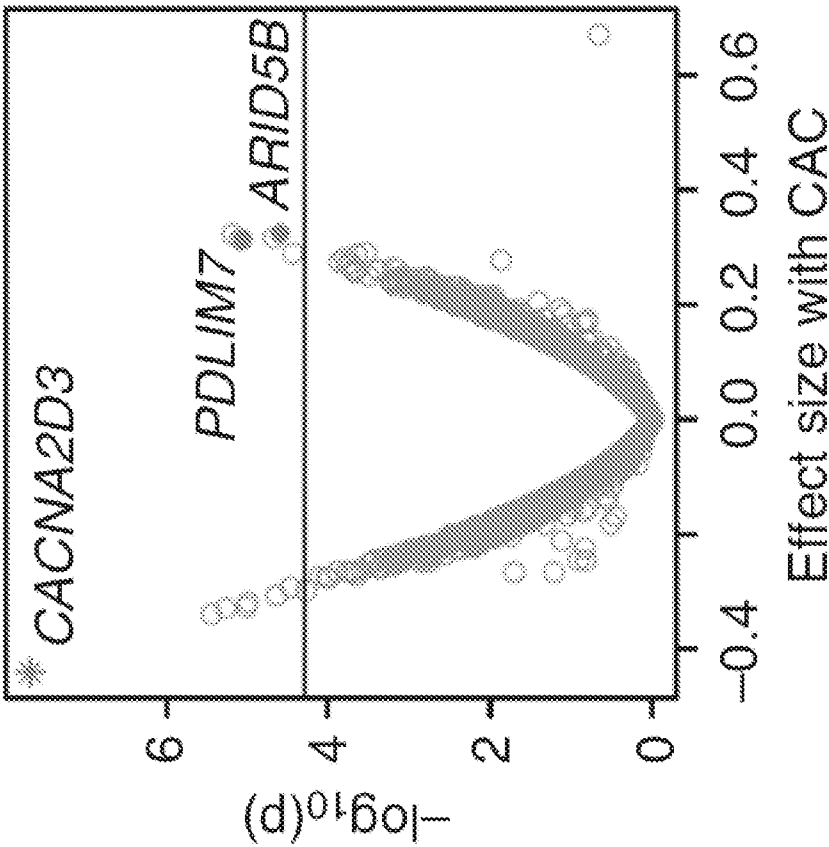


FIG. 1B

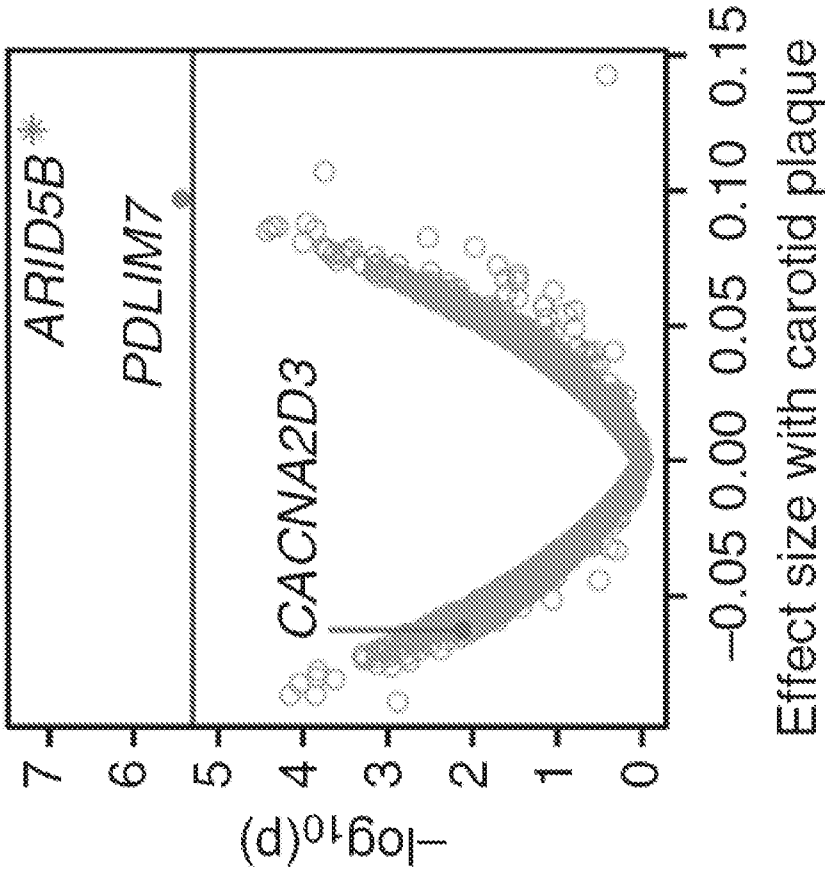


FIG. 1A

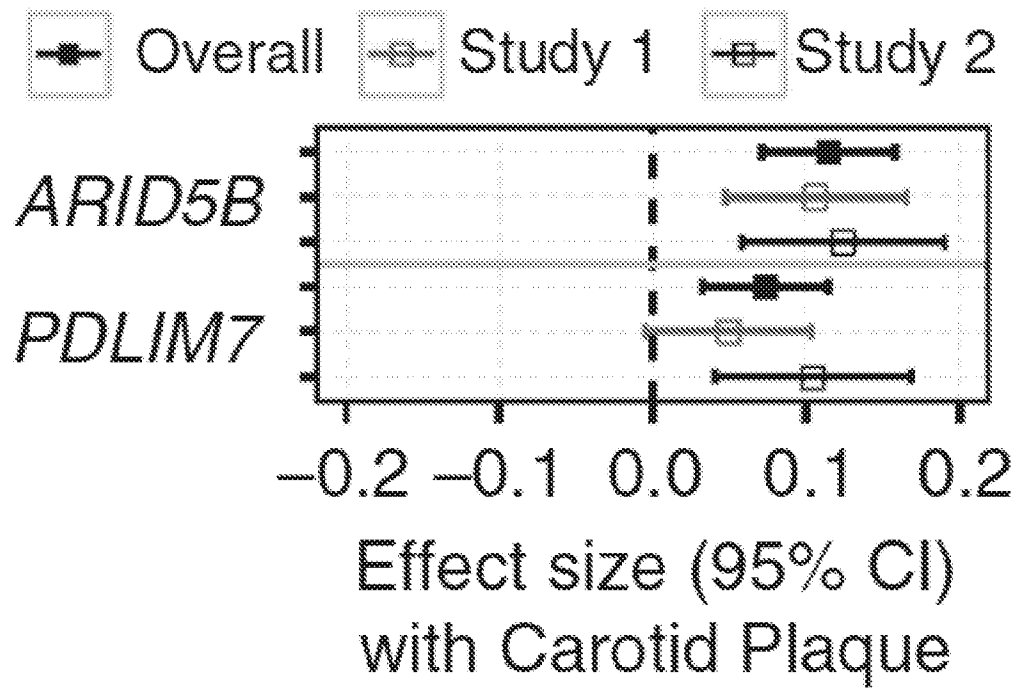


FIG. 1C

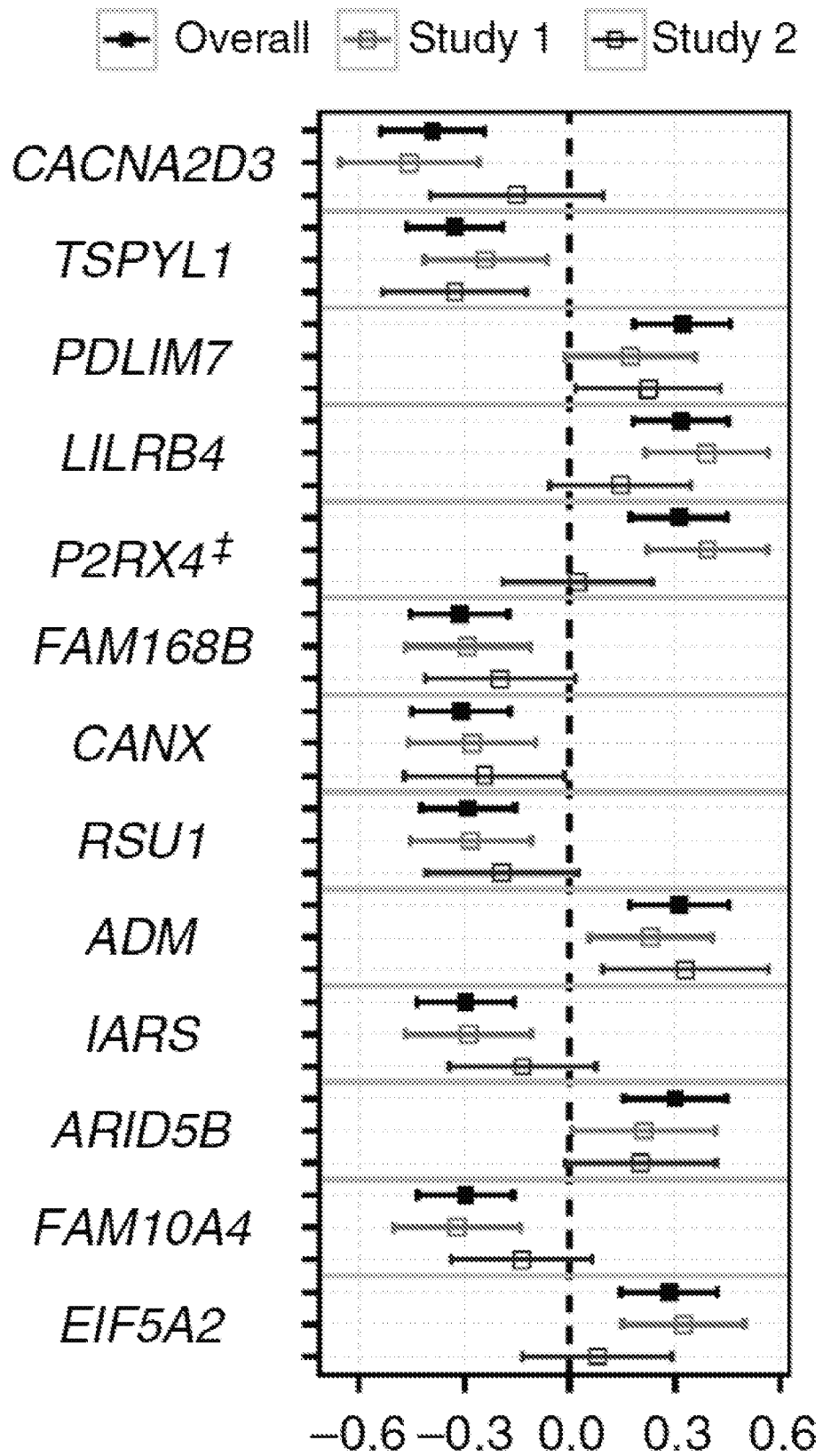


FIG. 1D

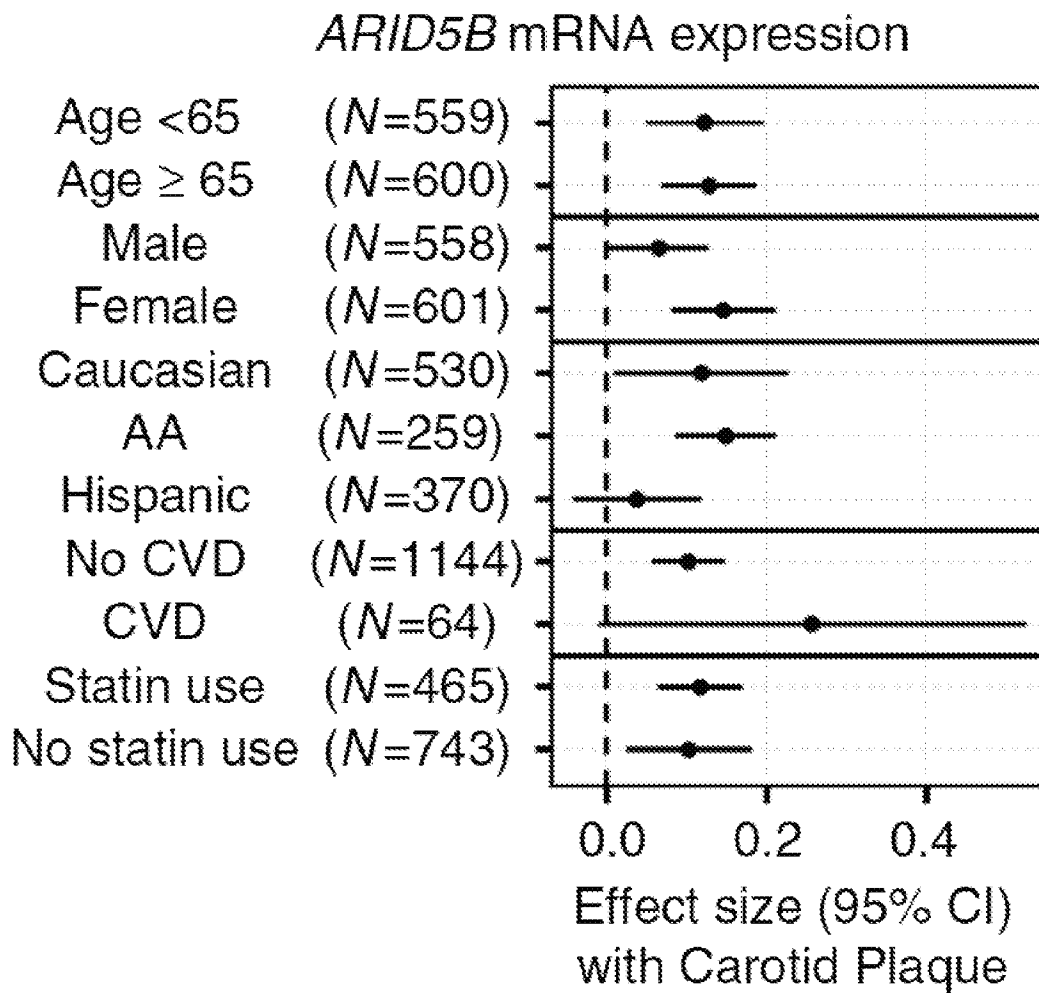


FIG. 1E

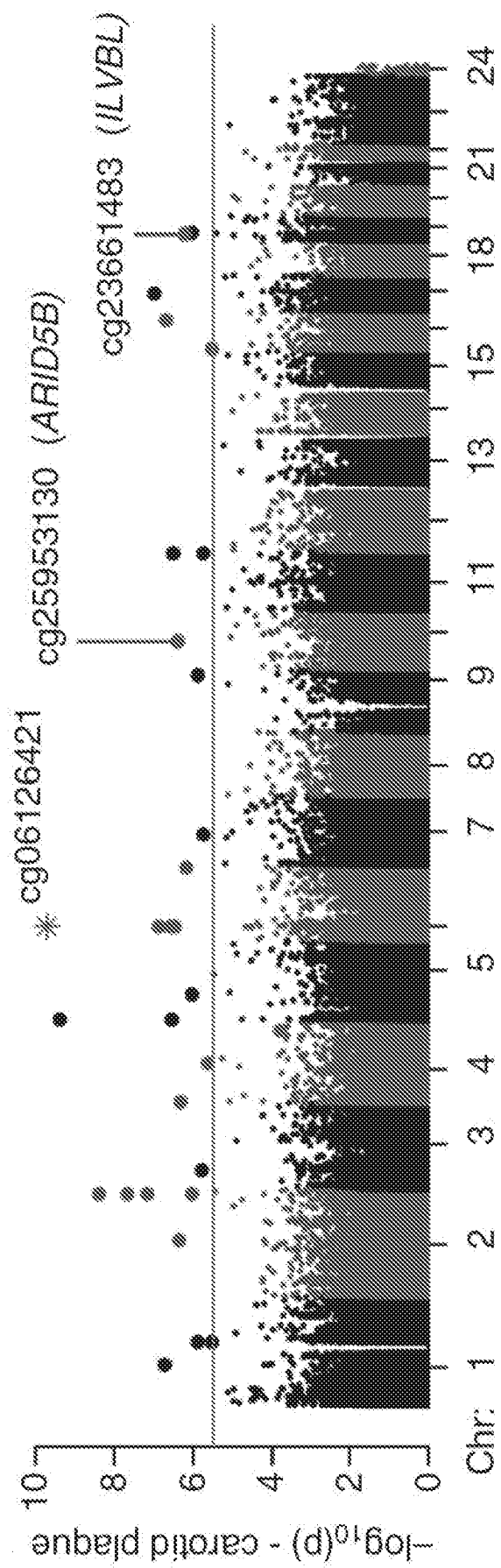


FIG. 2A

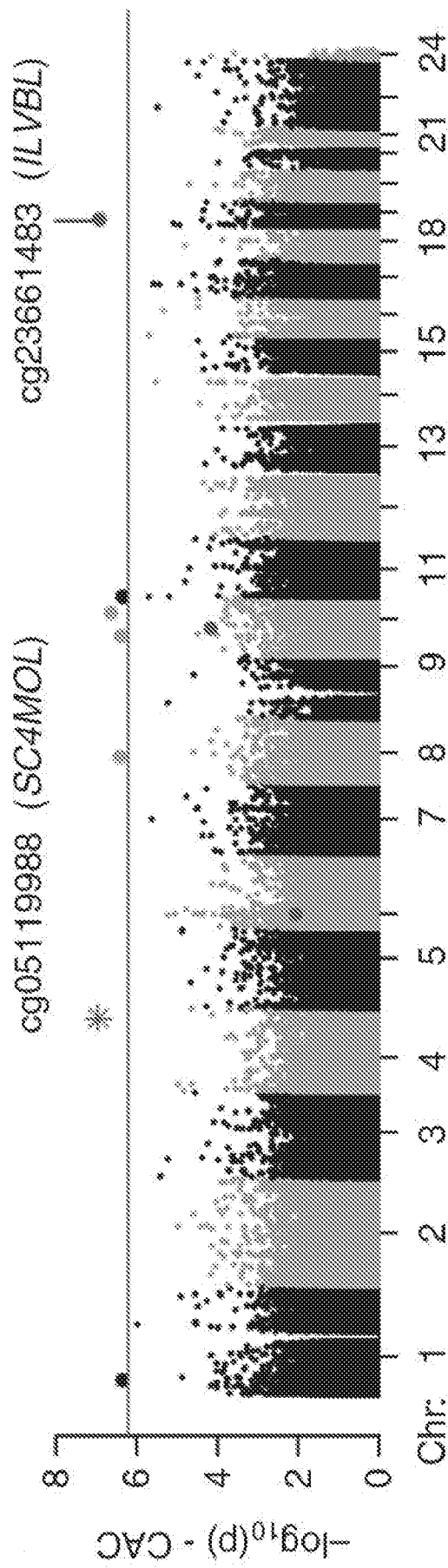


FIG. 2B

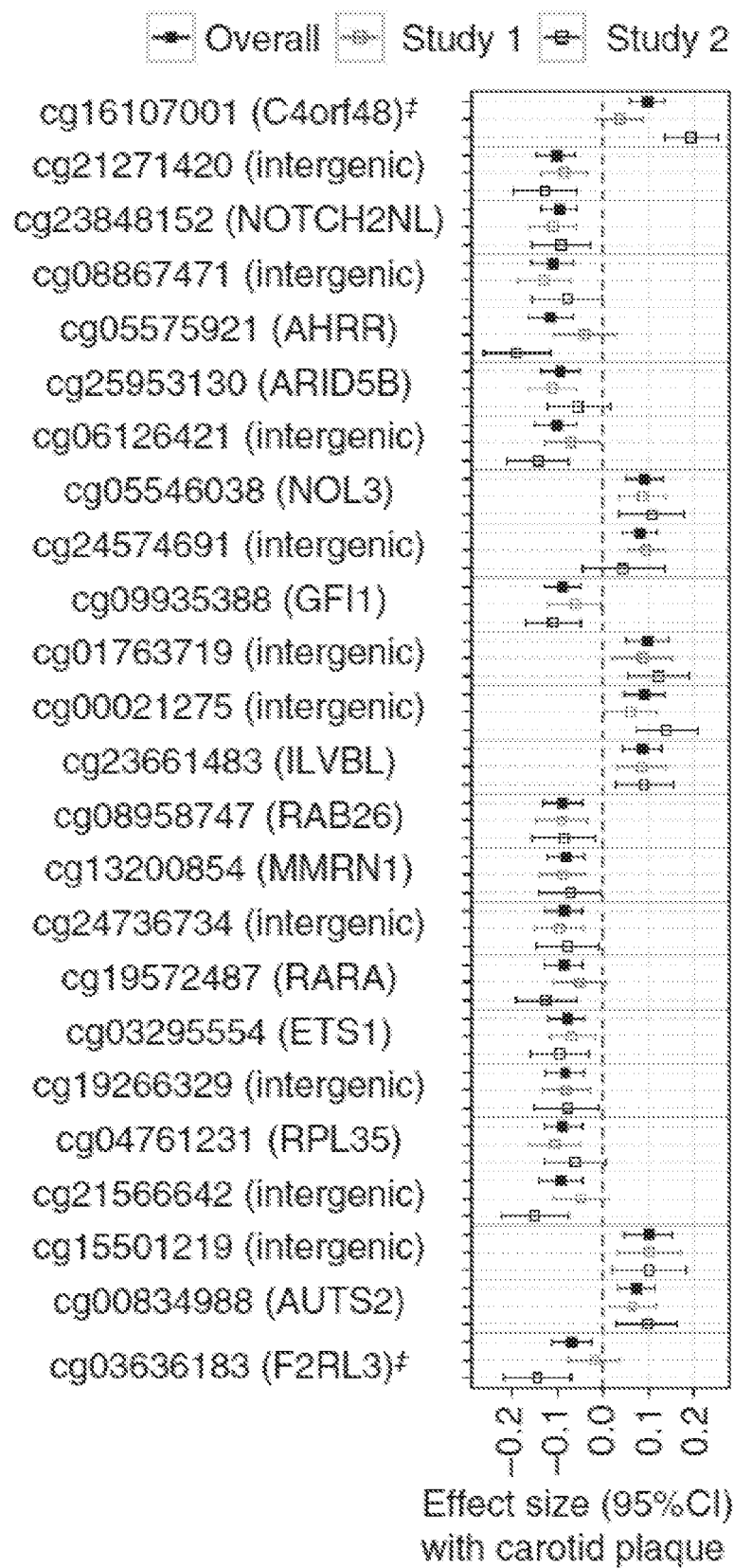


FIG. 2C

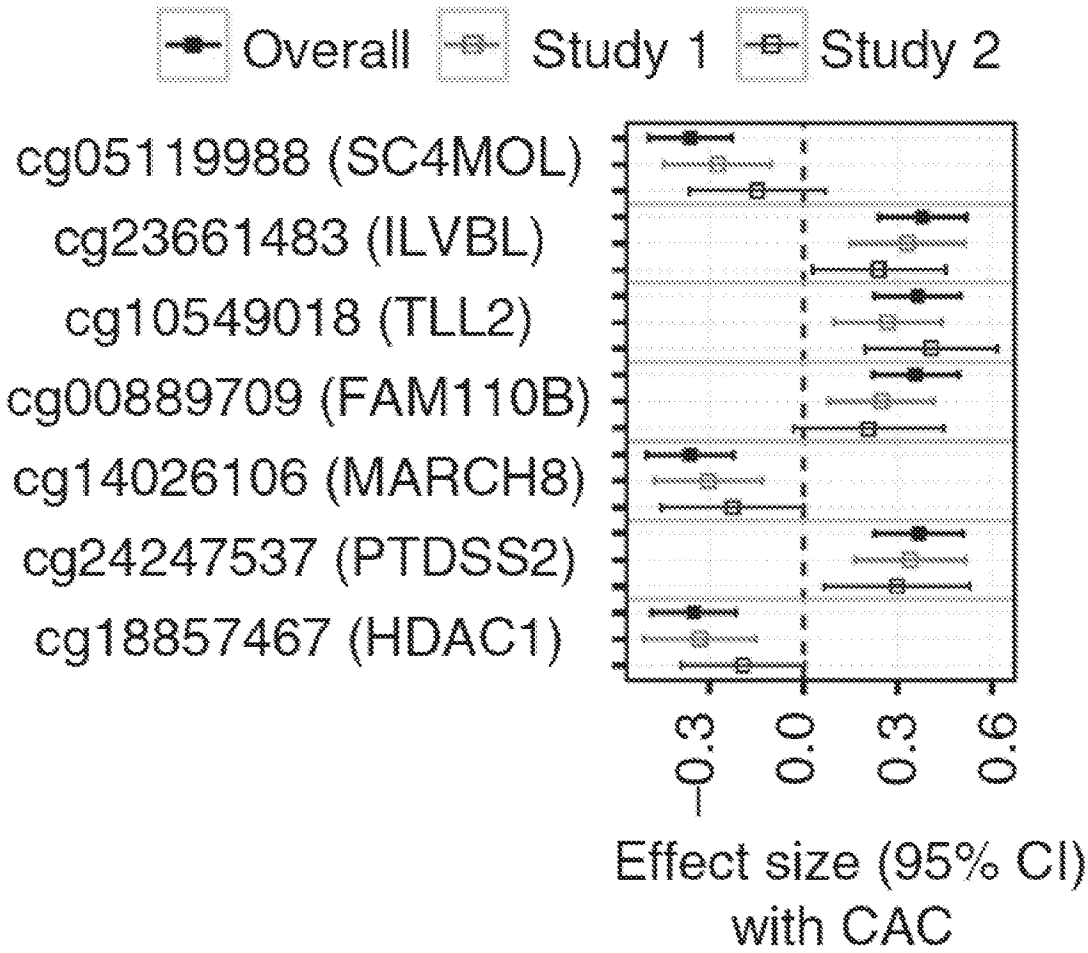


FIG. 2D

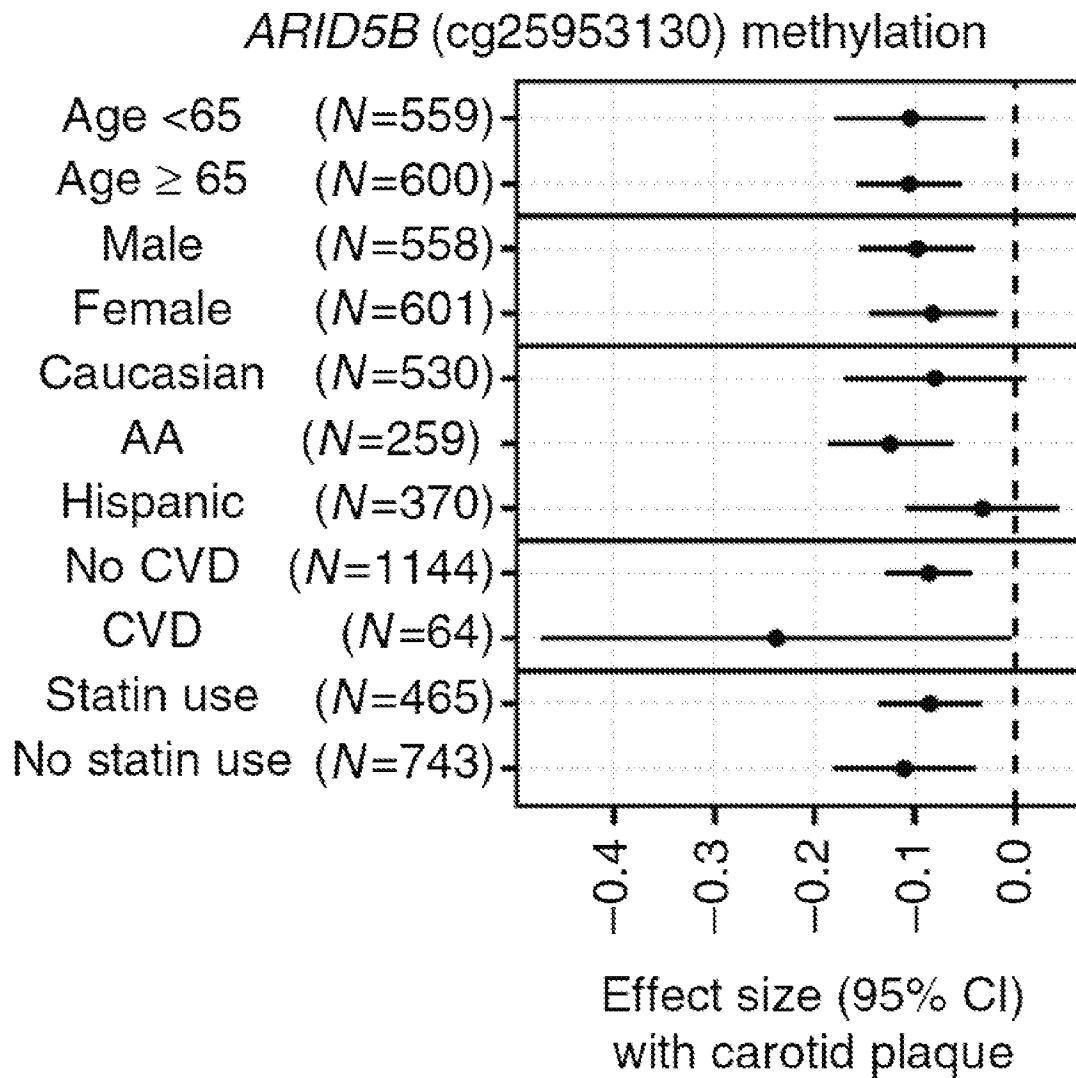


FIG. 2E

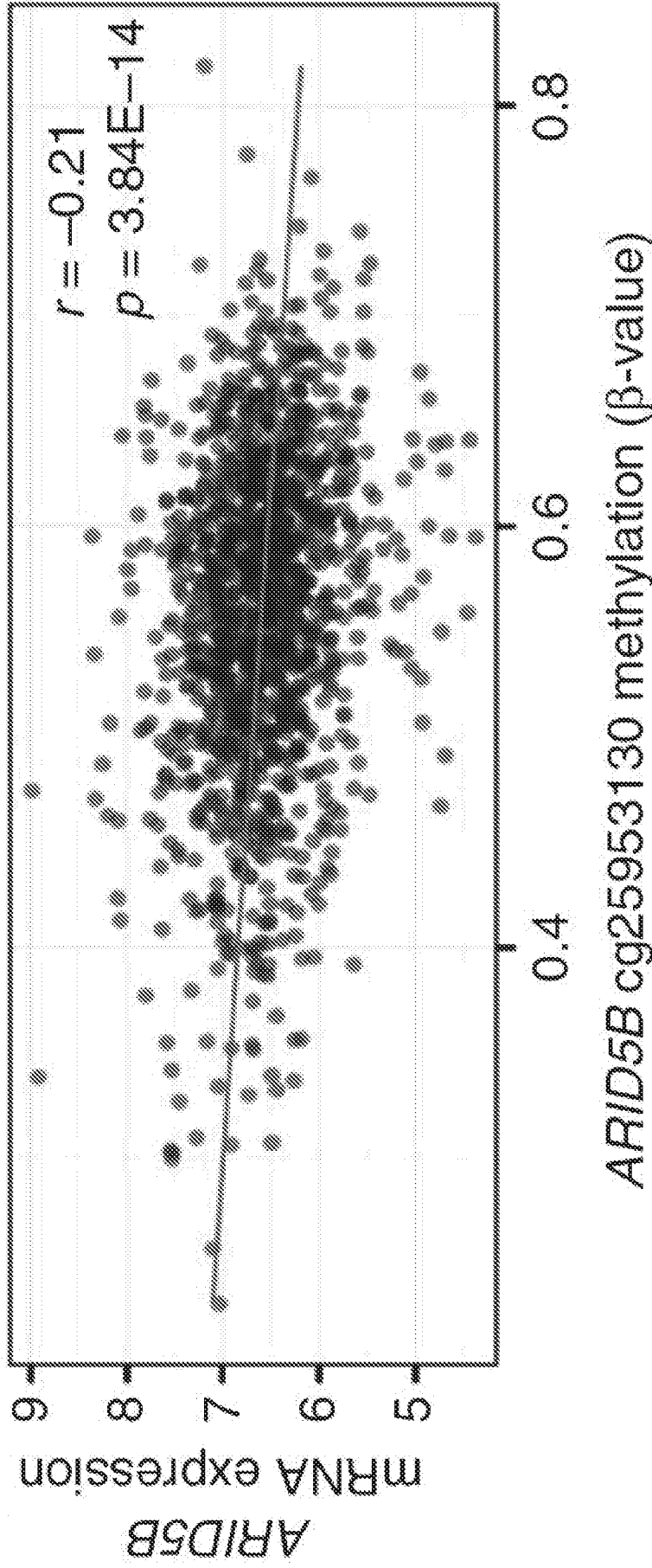


FIG. 3A

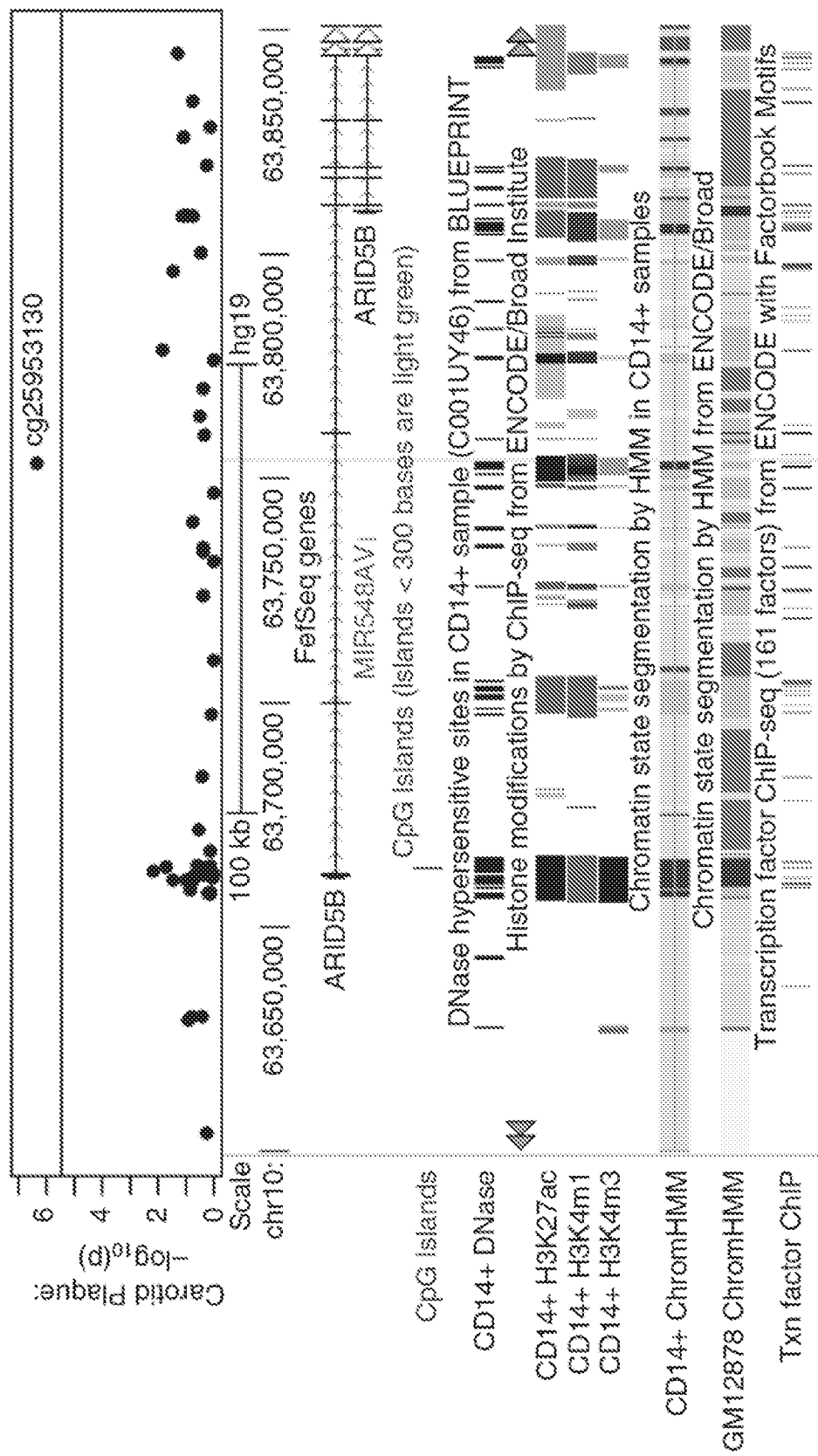


FIG. 3B

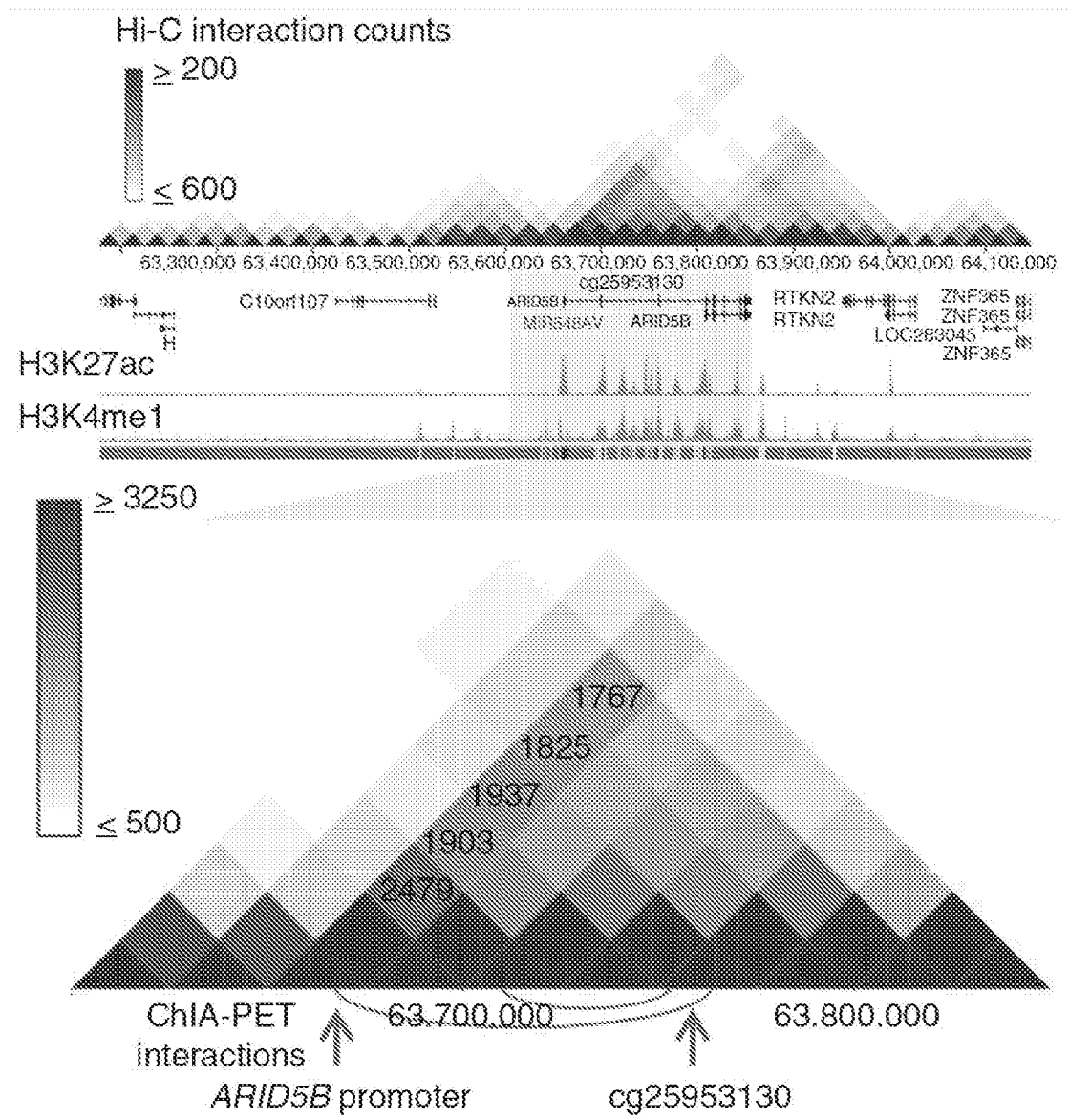


FIG. 3B

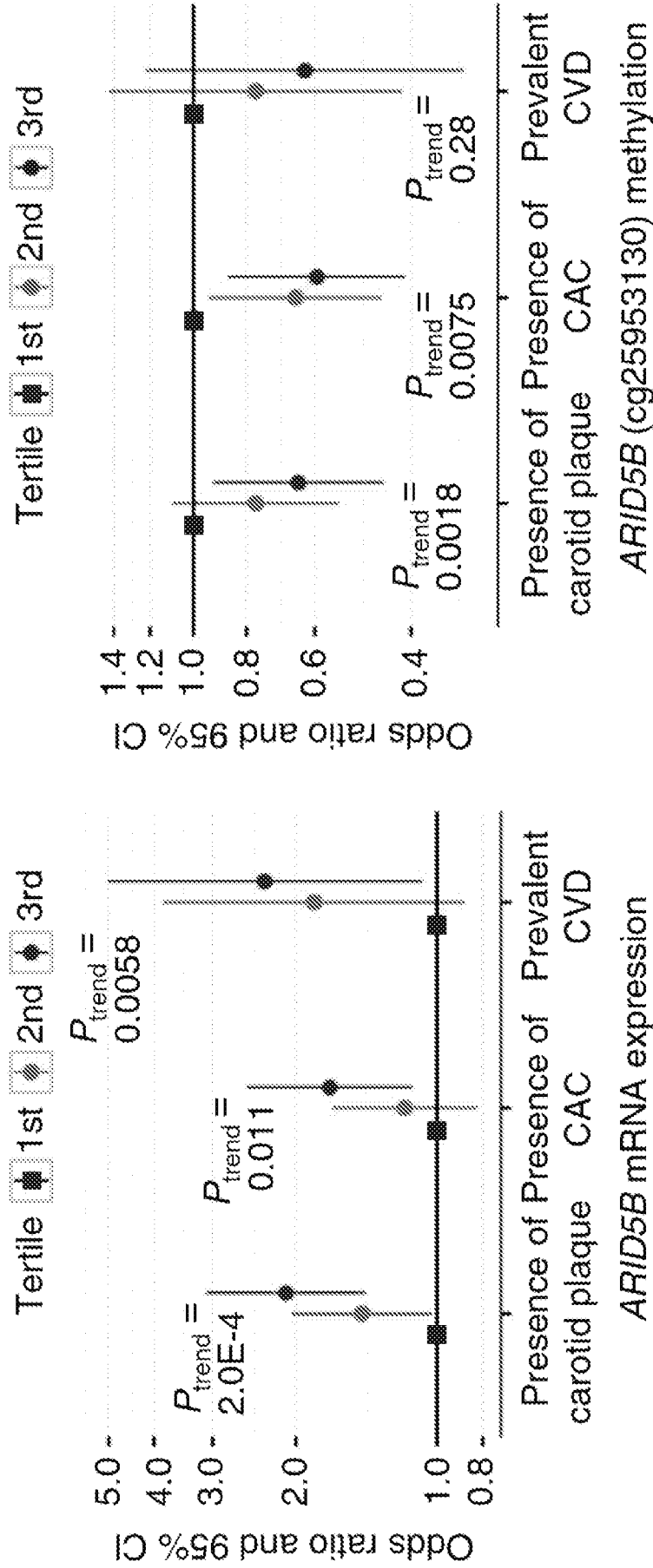


FIG. 4A

FIG. 4B

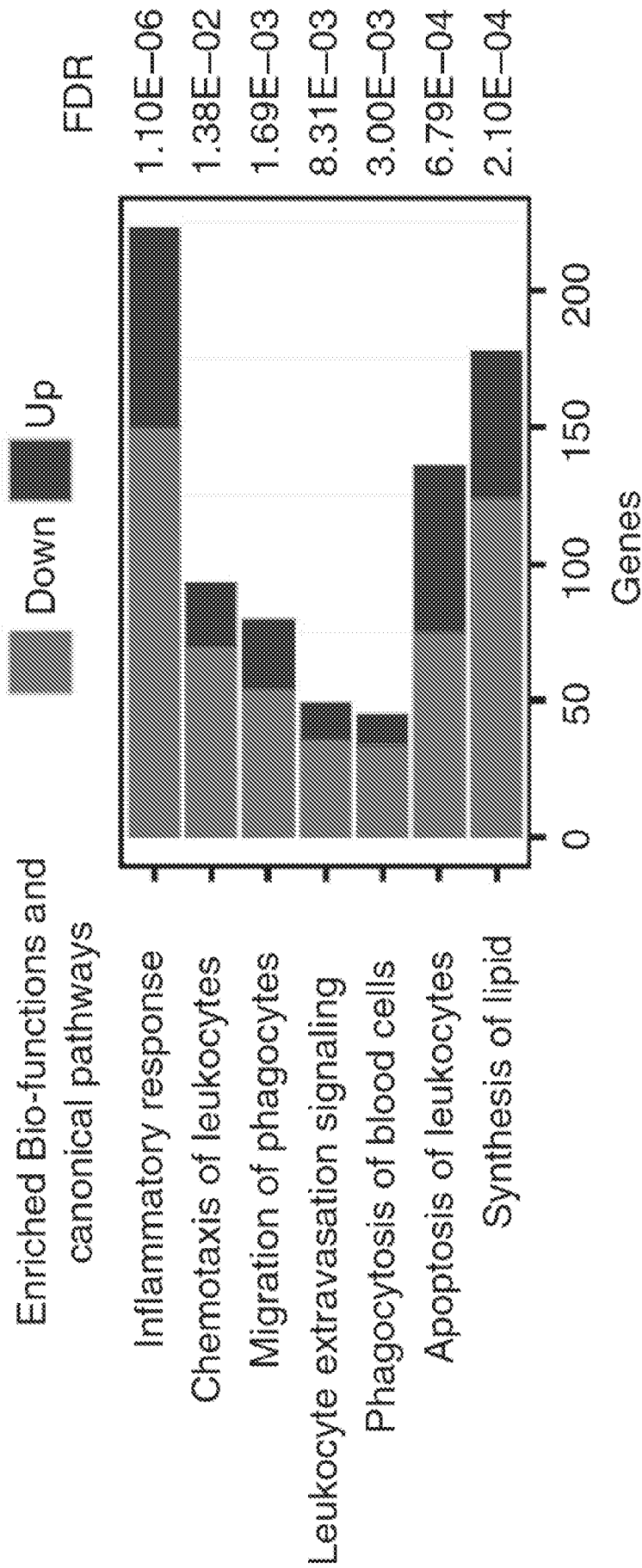


FIG. 5A

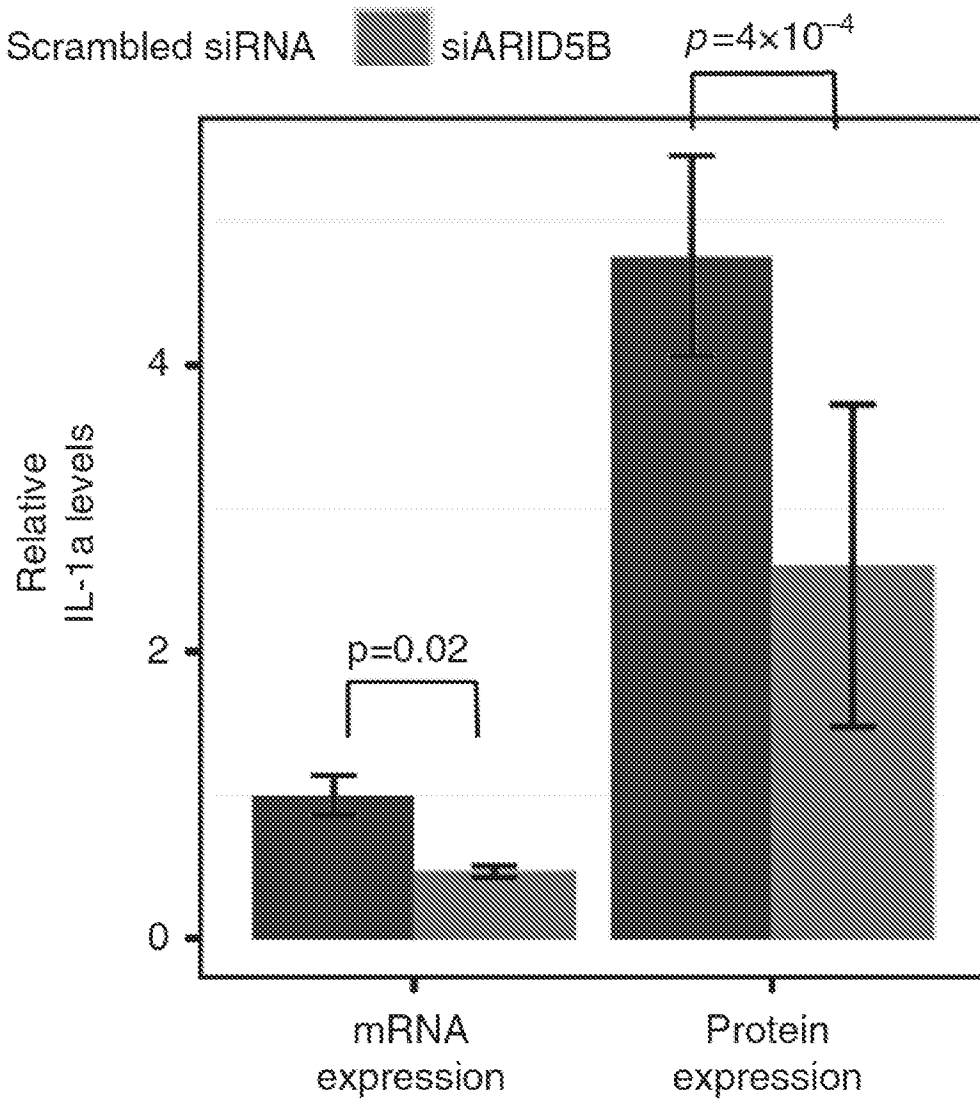


FIG. 5B

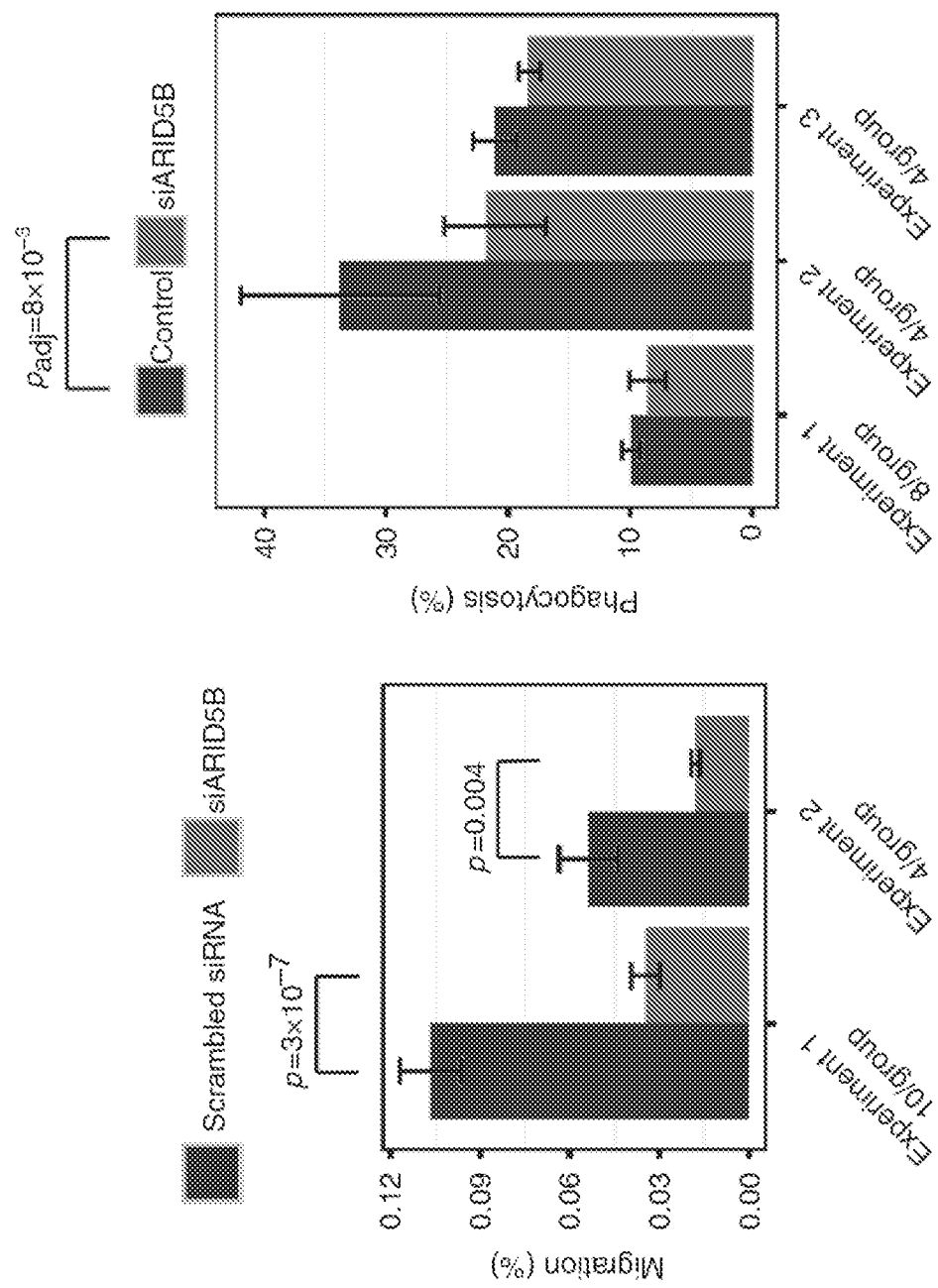


FIG. 5C

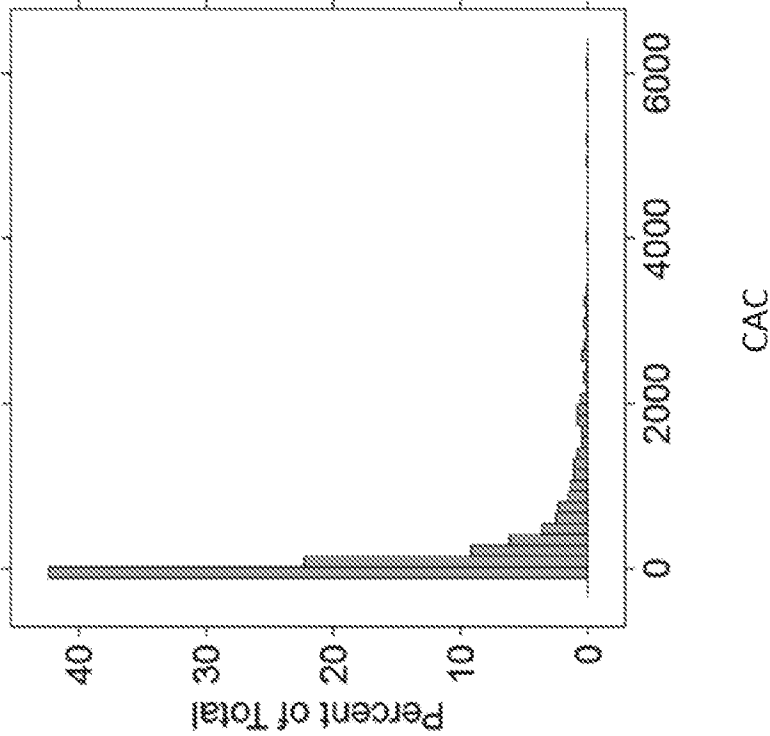


FIG. 6A

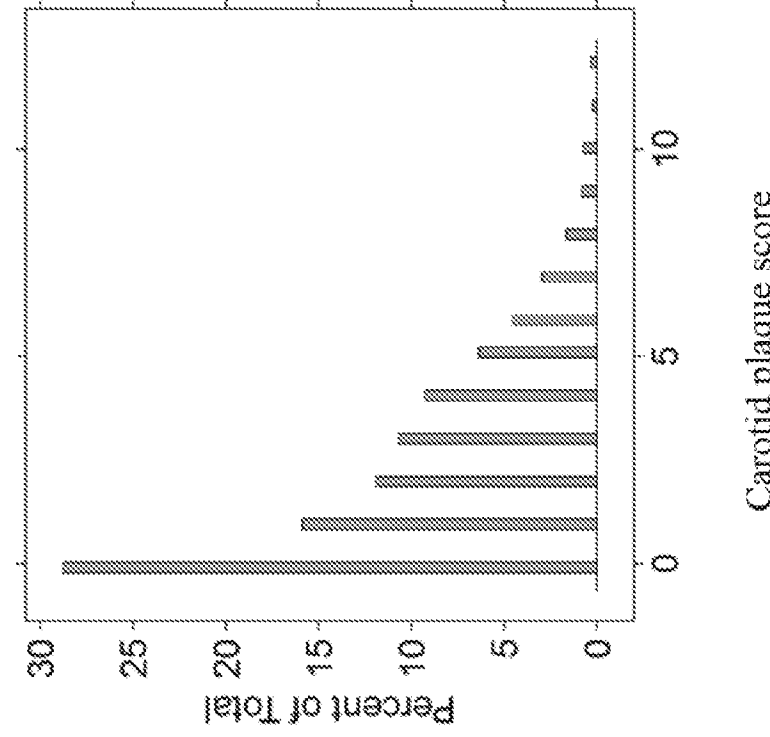


FIG. 6B

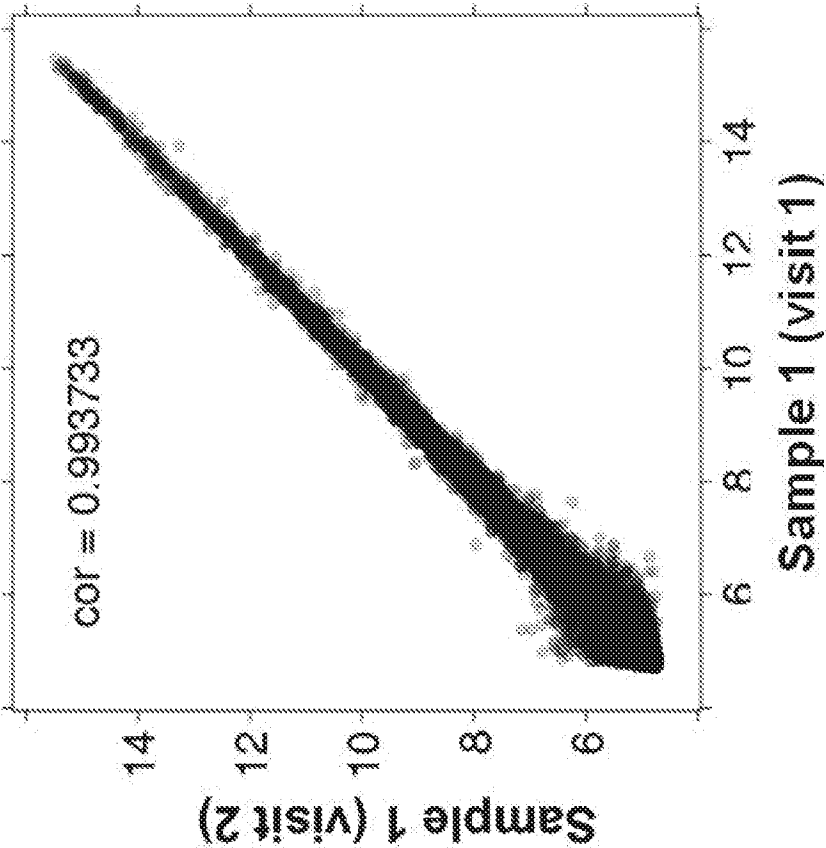


FIG. 7A

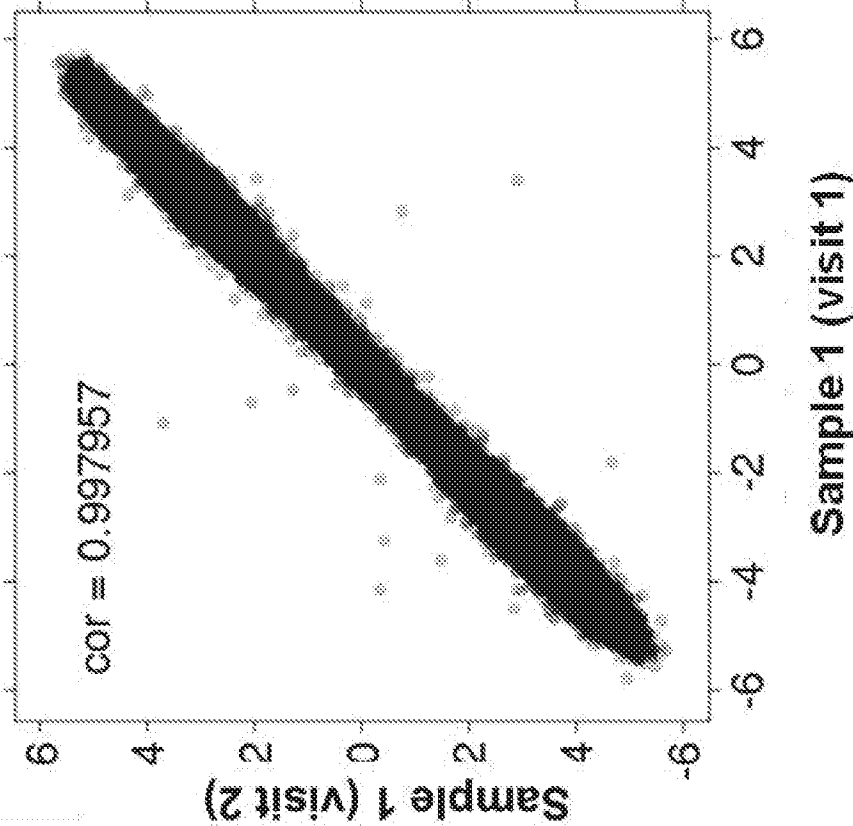


FIG. 7B

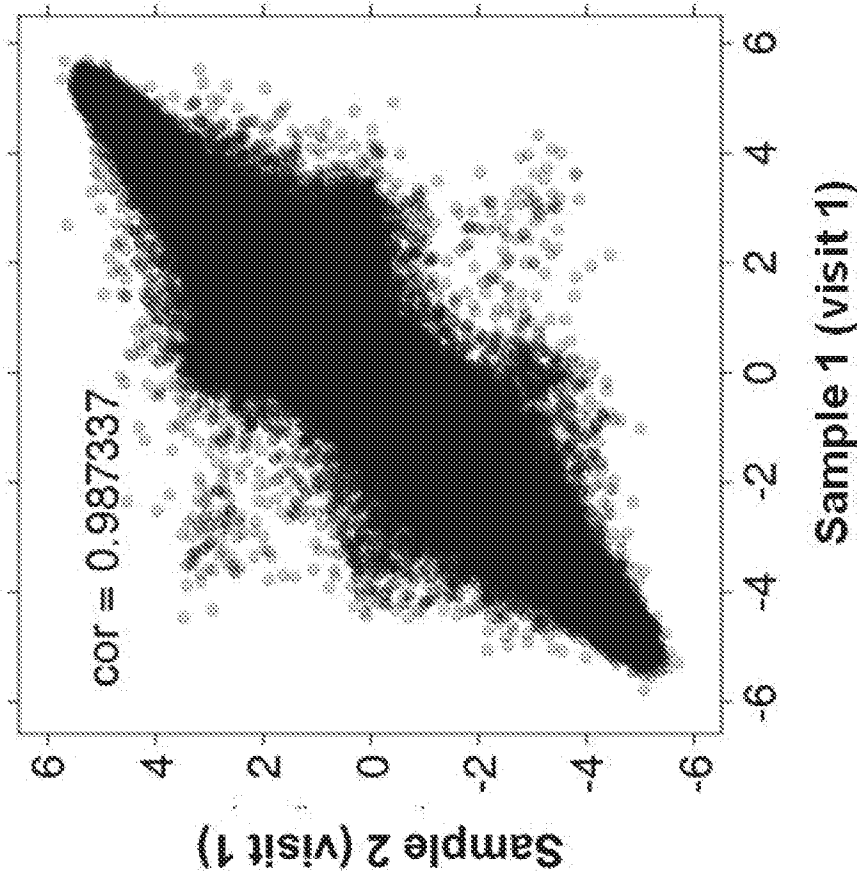


FIG. 7D

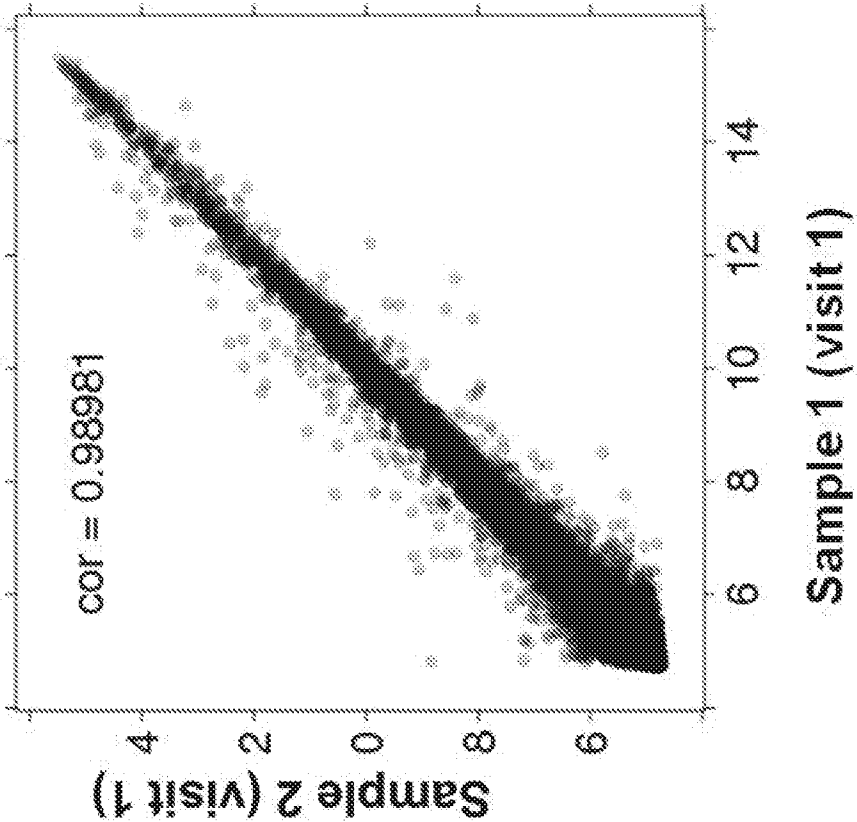


FIG. 7C

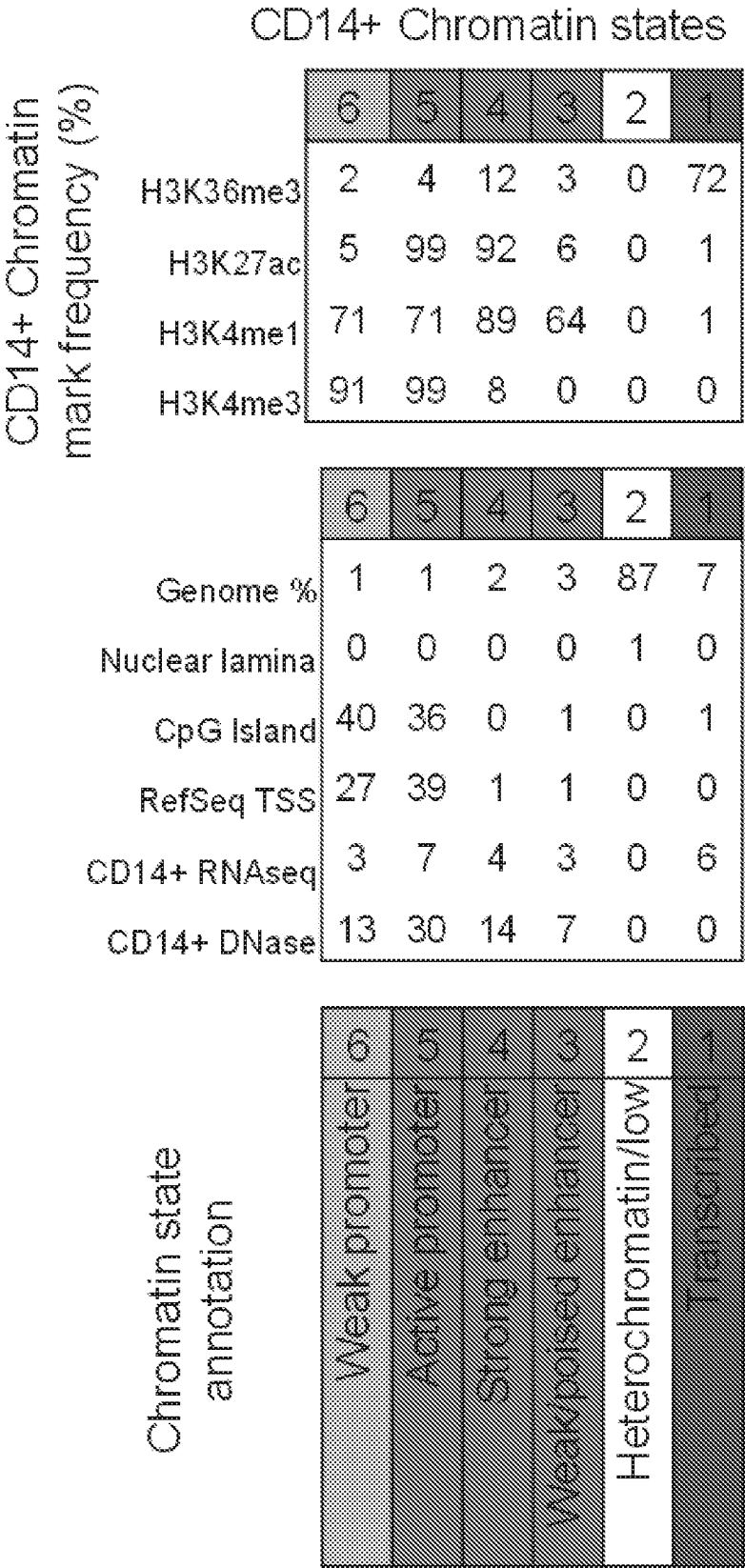


FIG. 8

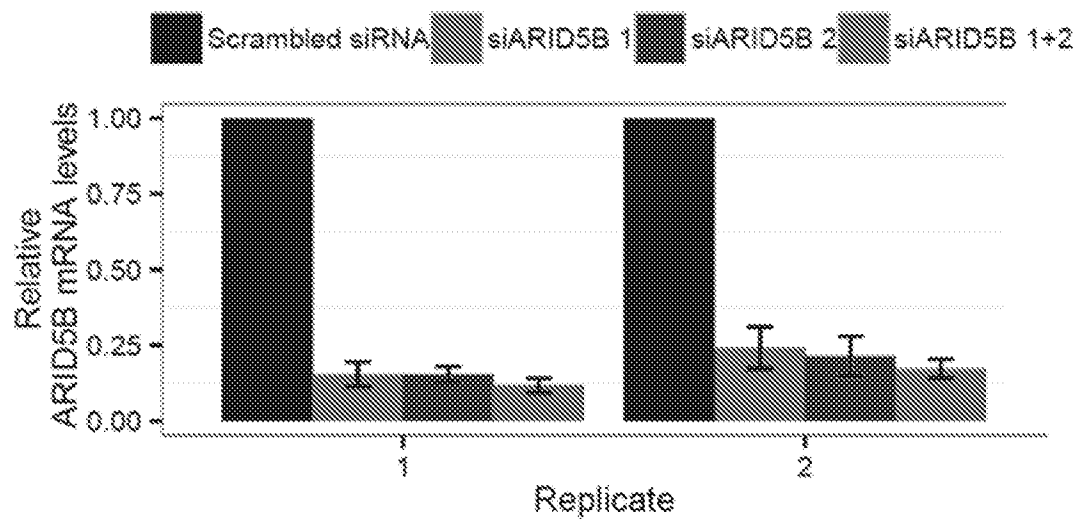


FIG. 9A

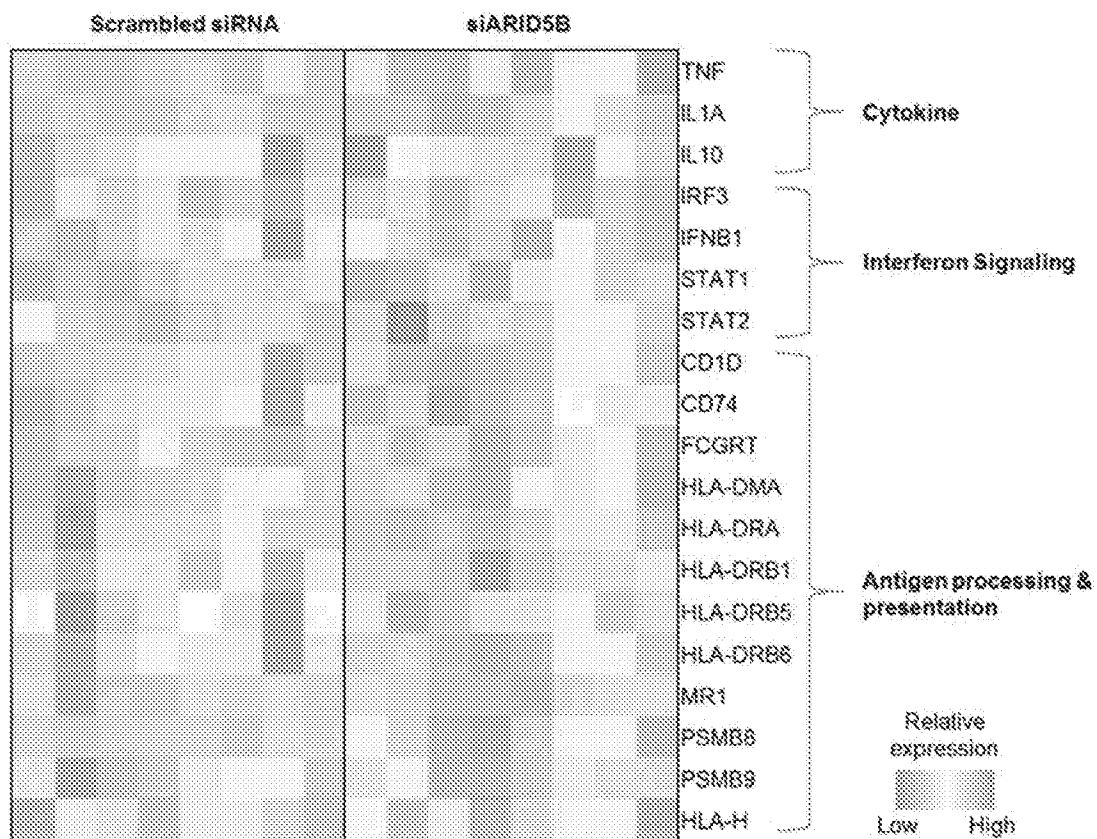


FIG. 9B

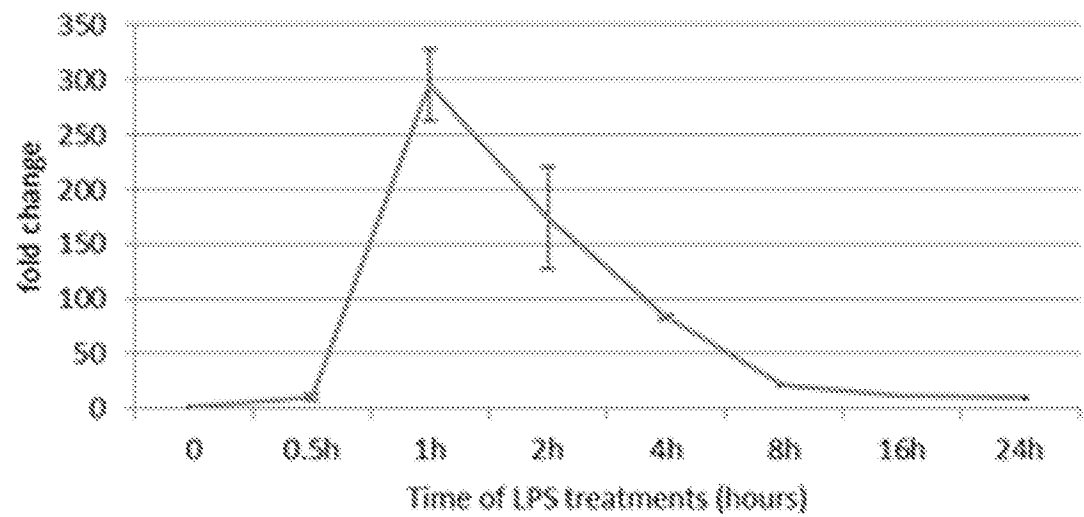


FIG. 10A

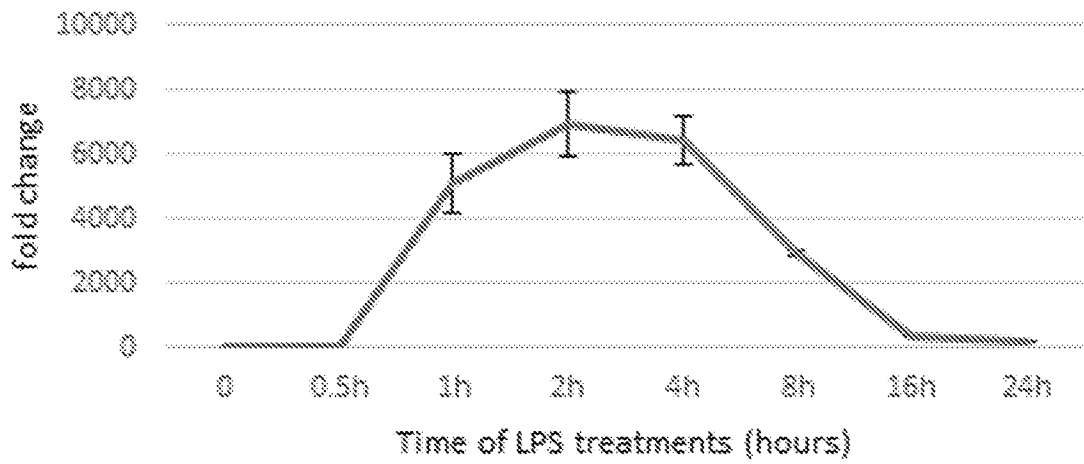


FIG. 10B

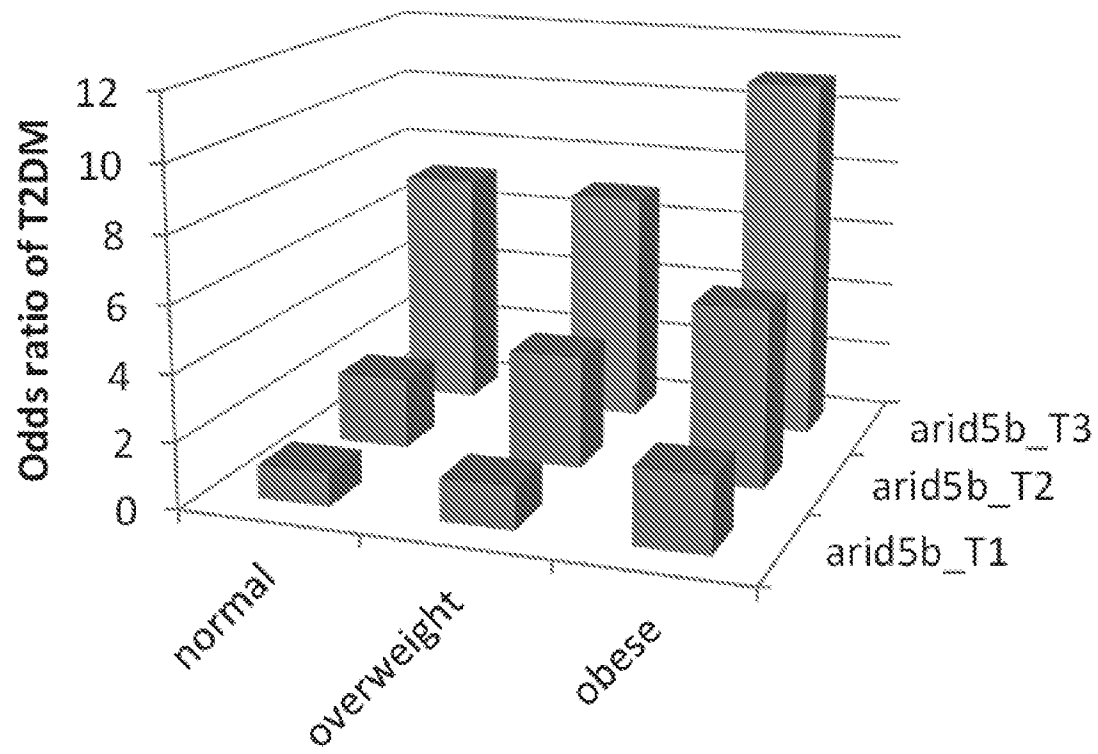


FIG. 11A

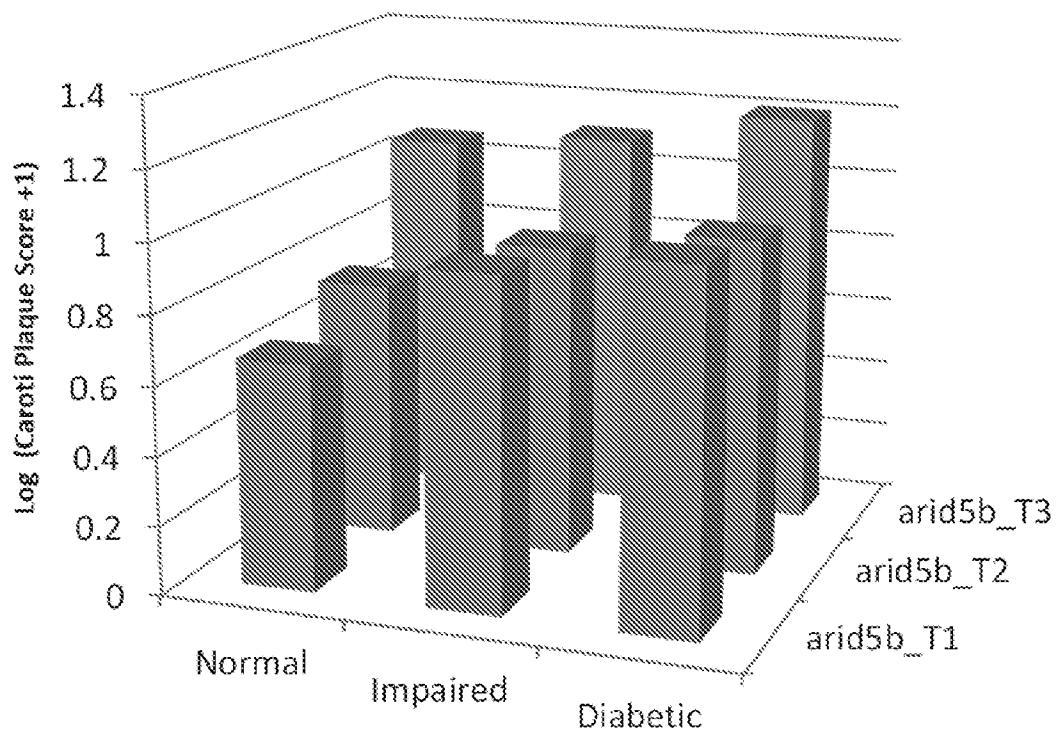


FIG. 11B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/28747

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C12Q 1/68; G01N 33/48, 33/50, 33/564; C07H 21/04; A61K 39/395; C40B 40/08 (2018.01)

CPC - C12Q 1/68; G01N 33/48, 33/50, 33/564; C07H 21/04; A61K 39/395; C40B 40/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	(LIU, Y et al.) Transcriptomics and Methylomics of Atherosclerosis in Circulating Monocytes - the Multi-Ethnic Study of Atherosclerosis. Abstract 53. Circulation. 10 March 2015, Vol. 131, Suppl No. 1; page 1, 1st paragraph; page 2, 2nd paragraph; page 2, 3rd paragraph	1, 4-5, 8/4-5, 9/8/4-5 ----- 11, 14/11, 15/11, 20, 23/20, 24/23/17-20
X -- Y	WO 2016/105535 A2 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY et al.) 30 June 2016; paragraphs [0007], [0011], [0073], [0079], [0081], [0086], [00112], [0114], [0122]	10, 14/10, 15/10, 17-19, 23/17-19 ----- 11, 14/11, 15/11, 20, 23/20, 24/23/17-20
Y	WO 2016/049024 A2 (THE BROAD INSTITUTE INC. et al.) 31 March 2016; paragraphs [0008], [0012]	20, 23/20, 24/23/20
A	US 2002/0182586 A1 (MORRIS, DW et al.) 05 December 2002; Table 1	2-3, 6-7, 8/6, 9/8/6, 12-13, 14/12-13, 15/12-13, 21-22, 23/21-22, 24/23/21-22
A	US 2009/0130096 A1 (SEIGNEURIC, RG et al.) 21 May 2009; Table 2D	2-3, 6-7, 8/6, 9/8/6, 12-13, 14/12-13, 15/12-13, 21-22, 23/21-2



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

29 June 2018 (29.06.2018)

Date of mailing of the international search report

16 JUL 2018

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/28747

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2015/124921 A1 (THE UNIVERSITY COURT OF THE UNIVERSITY OF EDINBURGH) 27 August 2015; entire document	1-7, 8/4-6, 9/8/4-6, 10-13, 14/10-13, 15/10-13, 17-22, 23/17-22, 24/23/17-22
A	WO 2014/036314 A2 (IGNYTA, INC.) 06 March 2014; entire document	1-7, 8/4-6, 9/8/4-6, 10-13, 14/10-13, 15/10-13, 17-22, 23/17-22, 24/23/17-22
A	(REYNOLDS, LM et al.) Age-related variations in the methylome associated with gene expression in human monocytes and T cells. Nature Communications. 18 November 2014, Vol. 5; pages 1-8; DOI: 10.1038/ncomms6366	1-7, 8/4-6, 9/8/4-6, 10-13, 14/10-13, 15/10-13, 17-22, 23/17-22, 24/23/17-22
P, X	(LIU, Y et al.) Blood monocyte transcriptome and epigenome analyses reveal loci associated with human Atherosclerosis. Nature Communications. 30 August 2017, Vol. 8, No. 1; pages 1-12; DOI: 10.1038/s41467-017-00517-4	1-7, 8/4-6, 9/8/4-6, 10-13, 14/10-13, 15/10-13, 17-22, 23/17-22, 24/23/17-22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/28747

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 16, 25-26
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.