Population Architecture using Genomics and Epidemiology (PAGE)

Ver. 02/25/21

PAGE Manuscript Proposal Template

Submit proposals by email to the PAGE Coordinating Center to Natalie Nudelman (nmakow@hginj.rutgers.edu)

Before submitting this form to Natalie, please fill out the yellow portion of the **PAGE II & III Ongoing Papers Google Sheet**, see, https://docs.google.com/spreadsheets/d/17E7E5G4u_nYWZEkzqgeRigwRbGxYzVNsuSUiVjXCytY/edit#gid=1451131466

Do not exceed 3 pages in length (not including references).

ARIC PAGE Ms. Number: 3893 **Submission Date: 7/13/21** **Approval Date: _____**

Title of Proposed MS: Methylation patterns associated with inflammation traits in multi-ethnic populations

Abbreviated Title of Proposed MS: DNA methylation and inflammation

I. INVESTIGATOR INFORMATION:

Name of Lead Author: Jessica Lundin | **Junior Investigator? Y**

Email Address: jlundin2@fredhutch.org

Authorship model*:

Name of Corresponding Author (**if different**):
Email Address:

Names, affiliations and email address of PAGE Investigators proposed as co-authors:

Name	Affiliation in PAGE	Email
Charles Kooperberg	WHI	clk@fredhutch.org
Riki Peters	WHI	upeters@fredhutch.org
Chris Carlson	WHI	ccarlson@fredhutch.org
Yao Hu	WHI	Yhu23@fredhutch.org
Chris Haiman	MEC	haiman@usc.edu
Kari North	SOL, ARIC	Kari_north@unc.edu
Tara Matise	Coordinating center	matise@rutgers.edu
Chris Gignoux	Coordinating center	Chris.gignoux@cuanschutz.edu
Eimear Kenny	Mt Sinai Biobank	Eimear.kenny@mssm.edu

Ruth Loos	Mt Sinai Biobank	Ruth.loos@mssm.edu
Laura Raffield	UNC	laura_raffield@unc.edu
Christy Avery	UNC	christy_avery@unc.edu
Alexander Reiner	UW	apreiner@uw.edu
Others TBD		

Partner studies collaborating with PAGE in this ms. proposal:

Name	Study	Email

II. SCIENTIFIC RATIONALE

Inflammation is a complex immune response including not only a rapid response to injury and pathogens, but also chronic low-grade inflammation that contributes to the pathophysiology of chronic diseases such as cardiovascular disease and diabetes.[1-3] Biomarkers of inflammation, such as C-reactive protein (CRP), can be extremely high after injury or in response to a pathogen (> 10 mg/L), however moderate-levels (> 3 mg/L) measured in peripheral blood samples have been used as indicators of disease.[4] Even with the high sensitivity of CRP measurements, and use as an indicator of increased disease risk, it is non-specific and the precise role of CRP in the development and progression of disease is not known. Genome-wide association studies in European Ancestry populations have identified candidate loci of genetic variants associated with CRP clustered by immune pathways and liver metabolic pathways, with the strongest association for a variant at the *CRP* locus.[5; 6] In all, the lead variants at all distinct loci explained ~6% of the CRP variance.[5] As such, the proportion of phenotypic variance explained remains small, likely due to trait complexity in the epigenome, the influence of environmental factors such as smoking, among other plausible explanations.

DNA methylation may contribute to the variation in disease phenotype biomarkers and mediate the effects of genetic and environmental factors. DNA methylation is an epigenetic modification characterized by the addition of methyl groups predominantly to cytosines at CpG sites and plays a pivotal role in gene expression through promoter silencing.[7] Gene-specific DNA methylation has been linked to inflammatory markers in previous studies (Table 1). In the largest epigenome-wide association study to date,[8] DNA methylation and CRP levels used a European population (8,863) as the discovery set and African American population (4,111) as a transethnic replication. This study identified and validated 58 CpG sites located in 45

unique loci in whole blood. The top signal was near the AIM2 gene, and was inversely associated with the CRP levels, with other significant associations with CRP for both ethnic groups including the PHOSPHO1 gene (identified in an EWAS of incident type 2 diabetes), and SOC3 gene (associated with atherosclerosis and increased risk of CHS). These findings were replicated in a separate study of inflammatory markers and DNA methylation that was focused on repeated samples from a European ancestry cohort (10 years between sample collection).[9] A significant association with CRP levels at the SOCS3 locus was also found in a EWAS study incorporating four European cohorts.[10] A study of CRP-associated DNA methylation in 966 participants of African Ancestry (representing 492 sibships) reported 257 CRP-associated DNA methylation sites (Infinium27k); the most significant CpG sites were near the KLK10, LMO2, and TM4SF4 genes, and genes associated with the immune response, with a single CpG-site, cg10636246 near gene AIM2 overlapping with the other EWAS studies described above (all Infinium450k).[11] In addition to the above findings, higher levels of IL-6 has been associated with (hyper and hypo) methylation of USP2, TMEM49, among others.[12-14] A review of the epigenome wide analysis and candidate gene studies summarize the genes associated with inflammation markers by pathways such as atherosclerosis, IL-6, IL-9, IL-8, growth hormones, and JAK/STAT signaling pathways.[15]

All but two of the analyses described above[8; 11] were performed in European-ancestry populations, with moderate sample sizes, highlighting the importance of examining inflammation-associated methylation sites in ancestrally diverse populations. The two multiethnic studies only extended to African Ancestry and their reporting findings did not overlap. It remains unknown whether the identified CpG-lipid associations could be generalized to Hispanic/Latinos and other racial/ethnic groups or replicated in African Americans. Another research need moving forward is to improve the understanding of the role of CRP in disease risk. As it is well established that CRP levels increase during infections and inflammatory disease, the development of therapeutic strategies directed toward cellular processes associated with increased CRP levels requires a deeper understanding of the molecular mechanisms across multi-ethnic populations. Evaluating the role of the epigenome in disease risk and along the causal pathway of these complex diseases, could be valuable in understanding the specific cellular processes and pathogenic mechanisms, as well as racial/ethnic variation.

III. OBJECTIVES AND PLAN

a. **Study Questions/Hypotheses**.

- To identify novel CpG sites associated with markers of inflammation in ancestrally diverse populations.
- To explore the generalization and potential heterogeneity of the previously reported and newly discovered CpG sites by examining their effect sizes and association directions across ethnic groups.
- To explore the modification effects of environmental/lifestyle covariates on CpG-inflammation associations.
- To infer the causality between differential methylation and the change of profiles of inflammation markers.

b. **Outcomes, Covariates, OMICs Data Considered.**

Outcomes (phenotype): Inflammation markers: CRP, IL-6, fibrinogen, and, possibly, D-dimer.

Covariates, as appropriate (main model): Age, sex, race/ethnicity, ancestry, study, center, family structure, white blood cell species, smoking, BMI, technical variables.

Covariates (potential covariates in extended model): prevalent CHD, diabetes, hypertension, and lipid levels (TG, tChol, HDL, LDL), fasting glucose, fasting insulin, inflammation related disease (colitis, diverticulitis, pancreatitis; n~550), medication use (related to any of the above), and other identified covariates on CpG-inflammation associations.

OMICS Data: Methylation data. DNA methylomic data is available in 11,093 individuals from whole blood assays and is drawn from Illumina arrays (the majority using the 450k array).

c. **Analytical Approach**

Before analysis, all CRP levels will be natural log transformed within each study, other inflammation biomarkers will be adjusted or transformed as appropriate. Individuals with > 4 standard deviations from the cohort mean (transformed) value will be excluded from the analysis. Beta values, which estimate the methylation level using a ratio of intensities between methylated and unmethylated alleles, will be used. These will be calculated and normalized using a Beta-MIxture Quantile dilation (BMIQ) approach. Any value with a detection p-value above 0.01 was set to missing (with detection *p*-values representing the likelihood of detection relative to background), and samples with more than 1.5% missing data were removed. Additionally, CpGs with greater than 1% missing data was excluded from the analysis. Based on prior work, we will eliminate any CpGs where the probe sequence maps either to a location that does not match the annotation file or to more than one locus.

In the methylome-wide association analyses, differentially methylated CpG sites will be evaluated for association with inflammation markers (i.e., phenotypes) using linear mixed effects regression with inflammation marker (e.g. CRP, transformed for normal distribution) as the dependent variable, and methylation B-values as the independent variable. The multiethnic models will be adjusted for age, sex (as appropriate), ancestry (PC1-10), center/study site (if applicable), family structure (if applicable), proportion of white blood cell species (using the Houseman method estimates or direct measurements), and technical covariates (chip ID, chip position, batch, etc as random effect). Smoking status and BMI will be included in the baseline model because these variables alter baseline CRP levels.[4] Other potential confounders will be evaluated for newly identified and previously reported inflammation marker-associated CpG sites in an extension of the baseline model. These may include (as available): inflammation related disease, prevalent diabetes, hypertension, CHD, and lipid levels (HDL, LDL, tChol, TG), and related medication use. The summary statistics from different studies will be combined through inverse variance-weighted fixed-effect meta-analyses. Bonferroni corrections will be applied to these models where $\alpha=0.05/(\text{number of CpG sites tested})$ in order to define significant CpG sites. The effect size will be represented as the change in phenotype value (e.g. lnCRP) per 1-unit change in the beta estimate (model coefficient). The genome coordinates provided by Illumina (GRCh37/hg19) will be used to annotate the CpG sites to loci.

We will then perform race/ethnic-specific methylome-wide association analyses to explore the generalizability and heterogeneity of inflammation marker-associated CpG sites. An examination of the association directions and effect estimates will be evaluated across ethnic groups.

To identify newly identified and previously reported inflammation marker-associated CpG sites that interplay with environmental/lifestyle covariates, we will perform 2-degree-of-freedom tests[16] that jointly evaluate main effects (CpG sites) and interaction (CpG sites by environmental/lifestyle covariate) with the same adjustment applied in the main model. We will explore sex as an interactive covariate, among other identified covariates as appropriate. A previous analysis demonstrated significant heterogeneity in effect estimates between men and women at four genetic variants,[5] although the effect estimates were all in the same direction. Tests of effect modification in methylation models by sex have been previously evaluated using sex-mDNA interaction terms in the multivariate models, but none have reported significant findings.[10; 11] The joint meta-analysis of the main and interactive effects has higher power than an interaction regression model and, thus, ability to detect a significant interaction if one exists.

In the Mendelian randomization analyses, we plan to use genetic variants as instrumental variables to determine whether differential methylation is consequential to the change of inflammation marker profiles or vice versa. We will conduct two-sample MR in relation to inflammation markers. For each analysis, estimates will be calculated primarily using inverse variance weighted regression methods which involve taking the ratio of the SNP-outcome effect estimate to the SNP-exposure effect estimate. Sensitivity analyses using MR-Egger[17] and weighted median[18] models will also be performed to examine evidence for directional pleiotropy.

d. **Anticipated date of draft manuscript to P&P: __TBD_____**

e. What manuscript proposals listed on *www.pagestudy.org/index.php/manuscripts/* **are most related to the work proposed here? Approved PAGE ms. numbers: _____**

- **If any: Have the lead authors of these proposals been contacted for comments and/or collaboration? Yes__ No__**

IV. SOURCE OF DATA TO BE USED (Check all that apply)

Study Name	Included
ARIC	**x**
HCHS/SOL	
CCHC	
MEC	**x**
BioME	
WHI	**x**
CARDIA	
MESA	**x**

I, _(initials)_, affirm that this proposal has been reviewed and approved by all listed investigators.

* We suggest inclusion of PAGE coauthors from all participating studies.

V. REFERENCES

1. Donath, M.Y., and Shoelson, S.E. (2011). Type 2 diabetes as an inflammatory disease. Nature reviews immunology 11, 98-107.
2. Peikert, A., Kaier, K., Merz, J., Manhart, L., Schäfer, I., Hilgendorf, I., Hehn, P., Wolf, D., Willecke, F., and Sheng, X. (2020). Residual inflammatory risk in coronary heart disease: incidence of elevated high-sensitive CRP in a real-world cohort. Clinical Research in Cardiology 109, 315-323.
3. Hage, F.G., and Szalai, A.J. (2007). C-reactive protein gene polymorphisms, C-reactive protein blood levels, and cardiovascular disease risk. Journal of the American College of Cardiology 50, 1115-1122.
4. Sproston, N.R., and Ashworth, J.J. (2018). Role of C-reactive protein at sites of inflammation and infection. Frontiers in immunology 9, 754.
5. Ligthart, S., Vaez, A., Võsa, U., Stathopoulou, M.G., De Vries, P.S., Prins, B.P., Van der Most, P.J., Tanaka, T., Naderi, E., and Rose, L.M. (2018). Genome analyses of> 200,000 individuals identify 58 loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. The American Journal of Human Genetics 103, 691-706.
6. Dehghan, A., Dupuis, J., Barbalic, M., Bis, J.C., Eiriksdottir, G., Lu, C., Pellikka, N., Wallaschofski, H., Kettunen, J., Henneman, P., et al. (2011). Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. Circulation 123, 731-738.
7. van der Harst, P., de Windt, L.J., and Chambers, J.C. (2017). Translational perspective on epigenetics in cardiovascular disease. Journal of the American College of Cardiology 70, 590-606.
8. Ligthart, S., Marzi, C., Aslibekyan, S., Mendelson, M.M., Conneely, K.N., Tanaka, T., Colicino, E., Waite, L.L., Joehanes, R., and Guan, W. (2016). DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. Genome biology 17, 1-15.
9. Myte, R., Sundkvist, A., Van Guelpen, B., and Harlid, S. (2019). Circulating levels of inflammatory markers and DNA methylation, an analysis of repeated samples from a population based cohort. Epigenetics 14, 649-659.
10. Marzi, C., Holdt, L.M., Fiorito, G., Tsai, P.-C., Kretschmer, A., Wahl, S., Guarrera, S., Teupser, D., Spector, T.D., and Iacoviello, L. (2016). Epigenetic signatures at AQP3 and SOCS3 engage in low-grade inflammation across different tissues. PLoS One 11, e0166015.
11. Sun, Y.V., Lazarus, A., Smith, J.A., Chuang, Y.-H., Zhao, W., Turner, S.T., and Kardia, S.L. (2013). Gene-specific DNA methylation association with serum levels of C-reactive protein in African Americans. PloS one 8, e73480.
12. Smith, A.K., Conneely, K.N., Pace, T.W., Mister, D., Felger, J.C., Kilaru, V., Akel, M.J., Vertino, P.M., Miller, A.H., and Torres, M.A. (2014). Epigenetic changes associated with inflammation in breast cancer patients treated with chemotherapy. Brain, behavior, and immunity 38, 227-236.
13. Lai, N.-S., Chou, J.-L., Chen, G.C., Liu, S.-Q., Lu, M.-C., and Chan, M.W. (2014). Association between cytokines and methylation of SOCS-1 in serum of patients with ankylosing spondylitis. Molecular biology reports 41, 3773-3780.

14. Uddin, M., Koenen, K., Aiello, A., Wildman, D., de Los Santos, R., and Galea, S. (2011). Epigenetic and inflammatory marker profiles associated with depression in a community-based epidemiologic sample. Psychological medicine 41, 997.
15. Gonzalez-Jaramillo, V., Portilla-Fernandez, E., Glisic, M., Voortman, T., Ghanbari, M., Bramer, W., Chowdhury, R., Nijsten, T., Dehghan, A., and Franco, O.H. (2019). Epigenetics and inflammatory markers: a systematic review of the current evidence. International journal of inflammation 2019.
16. Manning, A.K., LaValley, M., Liu, C.T., Rice, K., An, P., Liu, Y., Miljkovic, I., Rasmussen-Torvik, L., Harris, T.B., and Province, M.A. (2011). Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP× environment regression coefficients. Genetic epidemiology 35, 11-18.
17. Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International Journal of Epidemiology 44, 512-525.
18. Bowden, J., Davey Smith, G., Haycock, P.C., and Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 40, 304-314.
19. Miller, M., Maniates, H., Wolf, E., Logue, M., Schichman, S., Stone, A., Milberg, W., and McGlinchey, R. (2018). CRP polymorphisms and DNA methylation of the AIM2 gene influence associations between trauma exposure, PTSD, and C-reactive protein. Brain, behavior, and immunity 67, 194-202.
20. Arpon, A., Riezu-Boj, J.I., Milagro, F., Marti, A., Razquin, C., Martínez-González, M., Corella, D., Estruch, R., Casas, R., and Fitó, M. (2016). Adherence to Mediterranean diet is associated with methylation changes in inflammation-related genes in peripheral blood cells. Journal of physiology and biochemistry 73, 445-455.
21. Su, S., Zhu, H., Xu, X., Wang, X., Dong, Y., Kapuku, G., Treiber, F., Gutin, B., Harshfield, G., and Snieder, H. (2014). DNA methylation of the LY86 gene is associated with obesity, insulin resistance, and inflammation. Twin Research and Human Genetics 17, 183-191.
22. Jhun, M.A., Smith, J.A., Ware, E.B., Kardia, S.L., Mosley Jr, T.H., Turner, S.T., Peyser, P.A., and Park, S.K. (2017). Modeling the causal role of DNA methylation in the association between cigarette smoking and inflammation in African Americans: a 2-step epigenetic Mendelian randomization study. American journal of epidemiology 186, 1149-1158.

Table 1. Summary of previously identified CpG sites associated with inflammation markers

Inflammatory marker	CpG site	Chr	Position	Gene	Reference
mDNA and CRP	cg12992827	3	101901234	.	9
mDNA and CRP	cg16936953	17	57915665	*TMEM49*	9
mDNA and CRP	cg18181703	17	76354621	*SOCS3*	9, 11
mDNA and CRP	cg12054453	17	57915717	*TMEM49*	9
mDNA and CRP	cg06192883	15	52554171	*MYO5C*	9
mDNA and CRP	cg18942579	17	57915773	*TMEM49*	9
mDNA and CRP	cg18608055	19	1130866	*SBNO2*	9
mDNA and CRP	cg20995564	2	145172035	*ZEB2*	9
mDNA and CRP	cg01409343	17	57915740	*TMEM49*	9
mDNA and CRP	cg17980786	3	32933637	*TRIM71*	9
mDNA and CRP	cg07573872	19	1126342	*SBNO2*	9
mDNA and CRP	cg10636246	1	159046973	*AIM2*	9, 12, 16
mDNA and CRP	cg06690548	4	139162808	*SLC7A11*	9
mDNA and CRP	cg05316065	8	130799007	*GSDMC*	9
mDNA and CRP	cg15551881	9	123688715	*TRAF1*	9
mDNA and CRP	cg26470501	19	45252955	*BCL3*	9, 11
mDNA and CRP	cg26804423	7	8201134	*ICA1*	9
mDNA and CRP	cg27184903	15	29285727	*APBA2*	9
mDNA and CRP	cg27469606	19	1154485	*SBNO2*	9
mDNA and CRP	cg02481950	16	21665002	*METTL9*	9
mDNA and CRP	cg26610247	8	142297175	.	9
mDNA and CRP	cg07094298	4	2748026	*TNIP2*	9
mDNA and CRP	cg27023597	17	57918262	*MIR21*	9
mDNA and CRP	cg19821297	19	12890029	.	9, 11
mDNA and CRP	cg02650017	17	47301614	*PHOSPHO1*	9
mDNA and CRP	cg15721584	3	181326755	*SOX2OT*	9
mDNA and CRP	cg03957124	6	37016869	.	9
mDNA and CRP	cg00851028	1	234905772	.	9
mDNA and CRP	cg21429551	7	30635762	*GARS*	9
mDNA and CRP	cg01059398	3	172235808	*TNFSF10*	9
mDNA and CRP	cg02341197	21	34185927	*C21orf62*	9
mDNA and CRP	cg15310871	8	20077936	*ATP6V1B2*	9
mDNA and CRP	cg26846781	17	61620942	*KCNH6*	9
mDNA and CRP	cg27637521	17	76355202	*SOCS3*	9
mDNA and CRP	cg02003183	14	103415882	*CDC42BPB*	9
mDNA and CRP	cg04523589	3	48265146	*CAMP*	9
mDNA and CRP	cg25325512	6	37142220	*PIM1*	9
mDNA and CRP	cg22749855	17	76353952	*SOCS3*	9
mDNA and CRP	cg02734358	4	90227074	*GPRIN3*	9

mDNA and CRP	cg19769147	14	105860954	*PACS2*	9
mDNA and CRP	cg23761815	10	73083123	*SLC29A3*	9
mDNA and CRP	cg13585930	10	72027357	*NPFFR1*	9
mDNA and CRP	cg27050612	17	46133198	*NFE2L1*	9
mDNA and CRP	cg17501210	6	166970252	*RPS6KA2*	9
mDNA and CRP	cg15020801	17	46022809	*PNPO*	9
mDNA and CRP	cg12053291	12	125282342	*SCARB1*	9
mDNA and CRP	cg08548559	22	31686097	*PIK3IP1*	9
mDNA and CRP	cg18663307	21	46341389	*ITGB2*	9
mDNA and CRP	cg05575921	5	373378	*AHRR*	9
mDNA and CRP	cg09182678	22	50328711	.	9
mDNA and CRP	cg00812761	4	53799391	*SCFD2*	9
mDNA and CRP	cg00159243	12	109023799	*SELPLG*	9
mDNA and CRP	cg04987734	14	103415873	*CDC42BPB*	9
mDNA and CRP	cg06126421	6	30720080	.	9
mDNA and CRP	cg12269535	6	43142014	*SRF*	9
mDNA and CRP	cg24174557	17	57903544	*TMEM49*	9
mDNA and CRP	cg03128029	2	203143288	*NOP58*	9
mDNA and CRP	cg25392060	8	142297121	.	9
mDNA and CRP	cg02716826	9	33447032	*SUGT1P1; AQP3*	9 (discovery only), 11
mDNA and CRP	cg24204847	not reported	not reported	*EEF2*	17
mDNA and CRP	not reported	NA	NA	*LY86*	18
mDNA and CRP	not reported	NA	NA	*SOCS-1*	19
mDNA and CRP	not reported	NA	NA	*IL-6*	20
mDNA and CRP	cg07073964	19	649371	KLK10	12
mDNA and CRP	cg09358725	11	33870664	LMO2	12
mDNA and CRP	cg04121771	3	150674314	TM4SF4	12
mDNA and CRP	cg08458487	10	81699171	SFTPD	12
mDNA and CRP	cg09305224	9	139047066	FUT7	12
mDNA and CRP	cg00645579	11	607140	IRF7	12
mDNA and CRP	cg05556717	7	75257240	CCL26	12
mDNA and CRP	cg17496921	19	11267993	TSPAN16	12
mDNA and CRP	cg03801286	21	34806378	KCNE1	12
mDNA and CRP	cg21969640	12	53043844	GPR84	12
mDNA and CRP	cg05501357	11	33264845	HIPK3	12
mDNA and CRP	cg03600318	10	81698971	SFTPD	12
mDNA and CRP	cg18084554	19	880046	ARID3A	12
mDNA and CRP	cg06625767	5	176769301	F12	12
mDNA and CRP	cg15248035	9	36159949	CCIN	12
mDNA and CRP	cg05546038	16	65764534	NOL3	12
mDNA and CRP	cg09303642	12	52977085	NFE2	12
mDNA and CRP	cg03330678	17	72827828	9-Sep	12
mDNA and CRP	cg17753124	19	13120872	IER2	12

mDNA and CRP	cg22242539	17	1611970	SERPINF1	12
mDNA and CRP	cg17166812	1	159436198	NDUFS2	12
mDNA and CRP	cg22266967	4	6746599	S100P	12
mDNA and CRP	cg12380764	1	205037818	IL19	12
mDNA and CRP	cg10275770	17	59437937	ICAM2	12
mDNA and CRP	cg21492378	9	122890100	CEP1	12
mDNA and CRP	cg22381196	16	70598877	DHODH	12
mDNA and CRP	cg23140706	12	52975545	NFE2	12
mDNA and CRP	cg20283107	8	124858150	FAM91A1	12
mDNA and CRP	cg27606341	5	39255389	FYB	12
mDNA and CRP	cg26861460	22	42906788	PARVG	12
mDNA and IL-6	cg26077811	not reported	not reported	*USP2*	21
mDNA and IL-6	cg18942579	not reported	not reported	*TMEM49*	21
mDNA and IL-6	cg12054453	not reported	not reported	*TMEM49*	21
mDNA and IL-6	cg16936953	not reported	not reported	*TMEM49*	21
mDNA and IL-6	cg05438378	not reported	not reported	*SMAD3*	21
mDNA and IL-6	cg25446789	not reported	not reported	*DTNB*	21
mDNA and IL-6	cg01409343	not reported	not reported	*TMEM49*	21
mDNA and IL-6	cg13518625	not reported	not reported	.	21
mDNA and IL-6	None	NA	NA	*NA*	22
mDNA and IL-6	not reported	NA	NA	*SOCS-1*	19
mDNA and IL-6	not reported	NA	NA	*IL-6*	20
mDNA and fibrinogen	not reported	NA	NA	*LY86*	18

Table 1 references with brief description of study

[8]Ligthart et al 2016, EWAS n=8,863 (Discovery, European ancestry [EA]) and 4,111 (Replication, African ancestry [AA]), Infinium450, 58 of 218 discovery CpG sites replicated in replication mega analysis

[10]Marzi et al. 2016, EWAS n=1,741 (Discovery) and 503 (Replication), Infinium450, EA only

[11]Sun et al. 2013, EWAS n=966, Infinium27; AA only (only top 30 of 257 significant CpG-sites listed; non-listed sites were cross-checked with findings from other studies, one site was replicated and is noted above (near *AIM2* gene)

[19]Miller et al. 2017, candidate gene n=286, Infinium450, mostly EA men, PTSD Veterans

[20]Arpon et al. 2017, candidate gene n=36, Infinium450, EA, Mediterranean diet

[21]Su et al. 2014, candidate gene n>1400, Infinium450, EA and AA, obesity, insulin resistance, inflammation

[13]Lai et al. 2014, candidate gene n=46, Bisulfite method, ancestry not stated: maybe Taiwan (recruited from Taiwan hospital, all authors have Taiwan affiliations), ankylosing spondylitis

[14]Uddin et al. 2011, candidate gene n=100, abstract only, depression

[12]Smith et al. 2014, EWAS (baseline) then candidate gene (follow-up), n=61 then 39 women, Infinium450, EA and AA, breast cancer chemotherapy

[22]Jhun et al. 2017, candidate gene n=822, Infinium27, AA, mDNA, smoking, and inflammation (MR study)