Title

Social regulation of inflammation related gene expression in the multi-ethnic study of atherosclerosis

Authors

Brown, Kristen M
Diez-Roux, Ana V
Smith, Jennifer A
et al.

Publication Date

2020-07-01

DOI

10.1016/j.psyneuen.2020.104654

Peer reviewed

HHS Public Access

Author manuscript

Psychoneuroendocrinology. Author manuscript; available in PMC 2021 July 01.

Published in final edited form as:

Psychoneuroendocrinology. 2020 July ; 117: 104654. doi:10.1016/j.psyneuen.2020.104654.

Social Regulation of Inflammation Related Gene Expression in the Multi-Ethnic Study of Atherosclerosis

Kristen M. Brown[1,2,*], **Ana V. Diez-Roux**[3], **Jennifer A. Smith**[1,4], **Belinda L. Needham**[1], **Bhramar Mukherjee**[5], **Erin B. Ware**[1], **Yongmei Liu**[6], **Steven W. Cole**[7], **Teresa E. Seeman**[8], **Sharon L.R. Kardia**[1]

[1]Department of Epidemiology, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI, 48109

[2]Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Rd, Atlanta, GA 30322

[3]Department of Epidemiology, Dornsife School of Public Health, Drexel University, Nesbitt Hall, 3215 Market Street, Philadelphia, PA 19104

[4]Survey Research Center, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, MI 48104

[5]Department of Biostatistics, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI, 48109

[6]Department of Epidemiology and Prevention, Wake Forest University, Medical Center Boulevard, Winston-Salem, North Carolina 27157

[7]Department of Medicine, Division of Hematology-Oncology, University of California-Los Angeles, 11-934 Factor Building, UCLA School of Medicine Campus – 167817, Los Angeles, CA 90095

[8]Department of Medicine, University of California-Los Angeles, UCLA Med-Geri, BOX 951687, 2339 PVUB, Los Angeles, CA 90095

Abstract

Background—Exposure to adverse social factors has been associated with an altered inflammatory profile, a risk factor for several acute and chronic diseases. Differential gene expression may be a biological mediator in the relationship. In this study, associations between a range of social factors and expression of inflammation-related genes were investigated.

Methods—Social factor and gene expression data were collected from 1,264 individuals in the Multi-Ethnic Study of Atherosclerosis (MESA). Inflammation-related genes were identified from the Gene Ontology database. The associations between social factors and gene expression were first assessed using the Global Analysis of Covariance (Global ANCOVA) gene set enrichment test. When the global test was significant, linear regression and elastic net penalized regression

[*]**Corresponding author**: Kristen M. Brown, PhD, MS, Rollins School of Public Health, Department of Epidemiology, University of Michigan, 1415 Washington Heights, Ann Arbor, MI, 48109, brownkri@umich.edu.

Financial disclosures:
The authors declare that they have no competing financial interests.

tests were employed to identify the individual gene transcripts within each gene set associated with the social factor.

Results—Loneliness (p=0.003), chronic burden (p=0.002), and major or lifetime discrimination (p=0.045) were significantly associated with global expression of the chronic inflammatory gene set. Of the 20 transcripts that comprise this gene set, elastic net selected 12 transcripts for loneliness, 8 for chronic burden, and 3 for major or lifetime discrimination. Major or lifetime discrimination was also associated with the inflammatory response (p=0.029), regulation of the inflammatory response (p=0.041), and immune response (p=0.025) gene sets in global analyses, and 53, 136, and 26 transcripts were selected via elastic net for these gene sets respectively. There were no significant associations in linear regression analyses after adjustment for multiple testing.

Conclusions—This study highlights gene expression as a biological mechanism through which social factors may affect inflammation.

Introduction

Epidemiologic evidence suggests that low-grade chronic inflammation is a risk factor for many common diseases. Of the 10 leading causes of death, there is evidence that inflammation is associated with eight (i.e. cardiovascular disease, cancer, chronic lower respiratory diseases, stroke, Alzheimer's disease, diabetes, kidney disease, and influenza and pneumonia) (Bertoni et al., 2010; FastStats - Deaths and Mortality, n.d.; Grivennikov, Greten, & Karin, 2010; Kaptoge et al., 2010; Monadi et al., 2016; Racanelli, Kikkers, Choi, & Cloonan, 2018; Rubio-Perez & Morillas-Ruiz, 2012). For example, prior studies have found that people with high levels of the inflammatory biomarker C-reactive protein (CRP) have 1.6 times the risk of ischemic heart disease and 1.3 times the risk of cerebrovascular disease compared to those with low levels (Zacho et al., 2008). An elevated CRP level is also associated with increased risk of mortality (Marsik et al., 2008; Ridker, 2007). Understanding the factors that affect inflammation and the processes by which inflammation leads to disease will be important for reducing overall disease burden.

Exposure to adverse social factors may alter inflammation in a manner that affects disease risk. However, the current understanding of the biological mechanisms through which social factors "get under the skin" to affect inflammatory processes has been limited. Altered gene expression patterning has been proposed as a potential mechanism. For example, Cole and colleagues report differential expression of 144 genes by loneliness status with an overrepresentation of proinflammatory genes (Steve W Cole et al., 2007). Subsequent studies have found associations between low socioeconomic status (Powell et al., 2013), caregiver stress (Miller et al., 2008), well-being (Fredrickson et al., 2013), positive vs. negative affect (Antoni et al., 2012), and grief (O'Connor, Schultze-Florey, Irwin, Arevalo, & Cole, 2014) with altered expression of inflammatory genes. Prior work in human social genomics suggests that exposure to adverse social factors has a consistent response on the genome (i.e. conserved transcriptional response to adversity). However, since most of the previous studies analyzed only one or very related social factors (e.g. complicated vs non-complicated grief) at a time and used varying methods and data sources, cross-study comparisons are currently difficult.

Although these previous studies are promising, they suffer from two important limitations that present the need for subsequent investigation. Most notably, the ability to detect differential expression is limited by small sample sizes. This heightens the concern for having adequate power to detect significant associations. Indeed, previous work suggests that human social genomics studies are prone to Type II error leading to an under detection of significant associations (Steve W. Cole, 2010a). A second limitation of many previous studies is the use of peripheral blood mononuclear cells (PBMCs) in evaluating gene expression. PBMCs are a combination of different types of blood cells with a round nucleus (e.g. T cells, B cells, natural killer cells, monocytes, and macrophages), each with different gene expression patterns (Kleiveland & Kleiveland, 2015). Since gene expression patterning is cell type specific, potential significant associations in one cell type may be muted by the presence of cell types with null associations. These limitations present the need for deeper research.

In this study, we investigate whether seven social factors are each associated with immune response genes with a focus on genes involved in the inflammatory response. Loneliness was selected as a social factor because it was previously assessed in the seminal paper of human social genomics research by Cole et. al and has been associated with inflammation in several studies (Steve W Cole et al., 2007; Hänsel, Hong, Cámara, & von Känel, 2010). Socioeconomic status (in adulthood and childhood) was selected because it is considered a fundamental cause of health and is a consistent predictor of health outcomes (Link & Phelan, 1995). Stress and chronic burden were selected as general measures that capture adverse experiences in several domains (S Cohen, Kamarck, & Mermelstein, 1983; Pilkonis, Imber, & Rubinsky, 1985). Social support was selected because it an important buffer to the effects of stress (Sheldon Cohen & Wills, 1985; Ozbay et al., 2007). Discrimination was selected because it is an understudied social factor that has been associated with altered HPA axis activity and levels on inflammatory proteins (Busse, Yim, Campos, & Marshburn, 2017; Kershaw et al., 2016). By analyzing a variety of social factors in one study while using a consistent methodology and study sample, we are able to assess whether different social factors associate with expression of the same genes.

As the field of human social genomics is relatively new, there is currently no gold standard statistical approach. We use a multi-faceted approach that accounts for the high-dimensional, correlated nature of gene expression data. We also address important limitations of previous human social genomics studies in regards to power and differing cell types. The concern for power is addressed in the present study by using a large cohort (n=1,264). We address the concern presented by using multiple cell types by using solely monocytes, an important cell of the inflammatory response. Prior work indicates that among PBMCs, the expression patterns in monocytes are among the most sensitive to social factor exposures making monocytes a prime cell type for the current analysis (Steven W Cole, Hawkley, Arevalo, & Cacioppo, 2011).

Methods and Materials

Study sample

The Multi-Ethnic Study of Atherosclerosis (MESA) was designed to investigate risk factors for the development and progression of subclinical cardiovascular disease (Bild et al., 2002). The baseline cohort was recruited between 2000 and 2002. The cohort was comprised of 6,814 adults aged 45-84 who self-identified as non-Hispanic Black, Chinese-American, non-Hispanic White, or Hispanic and were free from clinical cardiovascular disease. Six exams have been completed. Each exam has consisted of a clinic visit where questionnaires on demographic, psychosocial, and lifestyle factors were administered, and physical assessments including the blood draw needed for genetic analyses were conducted. Gene expression was assessed solely in Exam 5 on a random sample of 1,264 individuals from four of the MESA study sites (i.e. Baltimore, Maryland; New York, New York; St. Paul, Minnesota; and Forsyth County, North Carolina) (Liu et al., 2013). These individuals make up the study sample for the current study.

Social Factors

Social factor data used in the current analyses were collected throughout the first five MESA exams (Table 1). The socioeconomic variables were dichotomized in the analyses. A participant was considered to have a high adult SES if he or she had achieved a college degree or higher and a high childhood SES if either parent had achieved a high school degree or higher. The psychosocial variables were assessed using measures with Likert scale responses. Summary scores for each factor were calculated by summing across items. Some items were reverse coded such that a higher value of the score indicated higher exposure for all social factors.

Gene expression

Gene expression was assessed from purified monocytes of 1,264 participants in MESA Exam 5 using the Illumina HumanHT-12 v4 Expression BeadChip. Detailed methods have been previously described (Liu et al., 2013). The chip has probes for 47,231 transcripts (~31,000 genes) across the genome. To avoid biases due to batch, chip, or position, a stratified random sampling technique was used. Several preprocessing and quality control steps were conducted for accurate quantification of the gene expression data. Local background was corrected for using Illumina's proprietary software Genome Studio. The following preprocessing steps were conducted using Bioconductor packages in R: a bead-type summarization was produced for each transcript using the *beadarray* package (Dunning, Smith, Ritchie, & Tavaré, 2007), negative controls on the array were used to compute the detection p-value, and the *limma* package was used for background correction, quantile normalization, $\log_2$ transformation, and removal of control probes (Smyth, Michaud, & Scott, 2005). Quality control criteria for elimination of a transcript included: 'detected' expression levels in <10% of MESA samples (detection p-value cut-off=0.01), probes that contain a SNP, probes with low variance across samples (<10^{th} percentile), and probes that overlap with non-unique regions.

Gene Set Generation

Gene sets were developed based on four inflammation related biological processes derived from the Gene Ontology database: chronic inflammatory response (GO:0002544), inflammatory response (GO:0006964), regulation of inflammatory response (GO:0050727), and the immune response (GO:0006955). The lists of genes included in each investigated biological process are given in Supplementary Table 1. Selected gene sets only included genes that were represented on the Illumina HumanHT-12 v4 Expression BeadChip used in the MESA study. For cases where a gene matched to more than one transcript, we included all transcripts in the analyses. The final gene sets consisted of 20 gene transcripts for chronic inflammation, 438 for the inflammatory response, 1251 for the immune response, and 192 for regulation of the inflammatory response. There were some overlapping gene transcripts among the four categories (Supplementary Table 2).

Statistical Analyses:

To statistically test the association between the social factors and gene expression, we employed three complementary methodological approaches. Previous human genomics studies suggest that replicating individual genes may be difficult, and it may be more successful to evaluate the social sensitivity of a set of genes related to a biological process (Steve W. Cole, 2010b; Fredrickson et al., 2013). Biologically, gene expression is highly correlated and acting in concert, so a multivariate joint model is an appropriate characterization of pathway effects. First, we assessed the association between each of the seven social factors individually with each of the four gene sets using a global analysis (Global ANCOVA) for a total of 28 independent tests. The Global ANCOVA is a self-contained gene set enrichment analysis that has the null hypothesis that no gene in a set of genes is associated with a given exposure. This method is described in detail in Hummel 2008 (Hummel, Meister, & Mansmann, 2008). Starting with the Global ANCOVA test reduces the total number of tests needed and lessens the concern for multiple testing. Further, by using the permutation based approximation of the F distribution in the Global ANCOVA test, we accounted for the gene expressions being neither independent nor homoscedastic (Hummel et al., 2008).

For associations significant in Global ANCOVA analyses, two methods (i.e. linear regression and elastic net) were used to identify individual genes associated with the social factor. Using an ordinary least squares (OLS) approach, linear regression analyses provide an estimate (beta) of the expected change in gene expression by each increased unit of the social factor. Linear regression was used as it is a familiar, easy to interpret test and is similar to analytic approaches used in previous studies. However, despite being common practice, OLS estimates can have inflated variance and become unstable with a large number of correlated predictors. Single site analysis, such as this linear regression approach, has low power to identify the key gene transcripts after adjustment for multiple testing even when false discovery rate correction for correlated data is used. Elastic net, a multivariate penalized regression technique, introduces bias but reduces variance to have better prediction and mean squared error properties (Zou & Hastie, 2005). It is used to regularize the estimation by incorporating selection and shrinkage and handle correlated predictors by minimizing the residual sum of squares subject to a bound on the L1 and L2 penalties of the

coefficients. The L1 penalty lends it the properties of LASSO using variable selection and the L2 penalty handles the correlation across predictors akin to ridge regression. In the present analyses, the L1 and L2 penalties were equally weighted and the tuning parameter was optimally chosen using cross validation. These analyses were conducted using the 'glmnet' package in R. Detailed methods of the elastic net are given in Zou and Hastie 2005 (Zou & Hastie, 2005).

Results

In the sample of 1,264 MESA participants with available gene expression data, the mean age was 69.6 years and 51% of the sample was female (Table 2). Three racial/ethnic groups were represented (Hispanic, non-Hispanic Black, and non-Hispanic White). For the socioeconomic variables, 33% of the total sample had finished a level of schooling equal to a college degree or higher, and slightly over half of the sample (56%) had either a mother or father achieve at least a high school degree. There were not significant racial/ethnic differences in the psychosocial factors except for major or lifetime discrimination ($p<0.001$). Tukey's honest significance test indicated that non-Hispanic Black participants reported experiencing a higher level of discrimination than non-Hispanic Whites ($p<0.001$) or Hispanics ($p<0.001$) (Tukey, 1949). There was not a significant difference between non-Hispanic Whites and Hispanics ($p=0.16$).

Gene set enrichment analyses

For each of the seven social factors, we first conducted Global ANCOVA tests for each gene set (i.e. chronic inflammation, inflammatory response, immune response, and regulation of inflammatory response). Three social factors (loneliness, discrimination, chronic burden) were significantly associated with the chronic inflammation gene set at $p<0.05$ and an additional two factors (perceived stress and social support) were marginally significant at $p<0.10$ (Table 3). Major or lifetime discrimination was also significantly associated with the other three gene sets (i.e. inflammatory response, immune response, regulation of inflammatory response) at $p<0.05$.

Multivariable linear regression

Multivariable linear regression analyses were conducted assessing the associations between the three social factors significant in Global ANCOVA analyses (i.e. loneliness, major or lifetime discrimination, and chronic burden) with each of the 20 individual transcripts of the chronic inflammation gene set ($3\times20=60$ tests). Although none of the linear regression associations persisted after FDR multiple testing adjustment, at $p<0.05$ we found significant associations between loneliness with 5 gene transcripts and major or lifetime discrimination with 4 gene transcripts (Table 4, Supplementary Table 3). Three of these transcripts (corresponding to two unique genes: *CX3CR1* and *VNN1*) overlapped between the two social factors. Despite significant Global ANCOVA results for chronic burden with the chronic inflammation gene set, linear regression analyses of individual genes did not identify significant associations with any of the 20 transcripts in the gene set at $p<0.05$.

Since there were significant findings for major or lifetime discrimination and each of the other three gene sets in Global ANCOVA analyses, we also examined the regression relationships between major or lifetime discrimination and each of the 438 gene transcripts of the inflammatory response, 1251 gene transcripts of the immune response, and 192 gene transcripts of the regulation of the inflammatory response. For the inflammatory response, immune response, and regulation of the inflammatory response, we found 12%, 11%, and 13.5% of investigated transcripts to be significant at $p<0.05$ respectively (Supplementary Table 3). None of the associations were significant after multiple testing adjustment.

Elastic Net

Elastic net regression was also used to identify the gene transcripts associated with the social factors significant in Global ANCOVA analyses (i.e. loneliness, major or lifetime discrimination, and chronic burden). The number of transcripts selected via elastic net ranged from 3-42 depending upon the social factor and gene set (Table 4). The greatest number of transcripts was selected for discrimination with the immune response gene set, the gene set with the largest number of transcripts (1251). The greatest proportion of transcripts was selected for loneliness with the chronic inflammation gene set (12/20=0.60). The number of transcripts selected for each social factor and gene set pair is presented in Table 4, and the list of transcripts with corresponding gene names is available in Supplementary Table 4.

Overlap between linear regression and elastic net penalized regression findings

We used two complementary methods (i.e. multivariable linear regression and elastic net penalized regression) to identify the transcripts driving the significant association from the global analyses and assessed the overlap between these findings. The results indicate that transcripts that were selected via elastic net analyses were also likely to be significant at $p<0.05$ in linear regression analyses. In total, of the 53 unique gene transcripts that were selected in elastic net, 43 (81%) were also found to be significant in linear regression analyses at $p<0.05$ (Table 4).

Discussion

In this study, we take a multi-faceted statistical approach to investigate the association between a range of social factors with expression of immune system genes, with a particular focus on genes involved in the inflammatory response. By examining multiple social factors in one study, we were able to investigate whether different social factors associate with expression of the same genes. Interestingly, we found that some social factors (i.e. loneliness, chronic burden, discrimination) but not others (i.e. perceived stress, social support, adult and childhood SES) were associated with candidate gene sets in global analyses. Further, for social factors that were significant in global analyses, individual gene analyses indicate that there was substantial overlap in the implicated genes.

The finding that loneliness was significantly associated with chronic inflammation in global analysis is consistent with the literature. Loneliness is a psychosocial stressor characterized by a subjective stressful feeling due to unmet needs of social connection (Alspach, 2013). It

has been found to associate with several chronic diseases (e.g. coronary heart disease, stroke) for which chronic inflammation is a known risk factor (Hawkley & Cacioppo, 2010; Valtorta, Kanaan, Gilbody, Ronzi, & Hanratty, n.d.). While there have been several studies investigating the relationship between loneliness and inflammation at a protein level (Hackett, Hamer, Endrighi, Brydon, & Steptoe, 2012; Jaremka et al., 2013; Mezuk et al., 2016), studies on loneliness and gene expression have been limited. The small, extant literature suggests that loneliness does associate with biological functioning at the gene level, and the patterns may be altered by behavioral treatments (Steve W. Cole, 2008; Steve W Cole et al., 2007; Steven W. Cole et al., 2015; Steven W Cole et al., 2015; Creswell et al., 2012).

To identify which individual gene transcripts in the chronic inflammatory gene set were significantly associated with loneliness, we followed up the global test with linear regression and elastic net penalized regression analyses. In linear regression analyses, we found 5 of the 20 transcripts to be significantly associated with loneliness at $p<0.05$. However, none were statistically significant after false discovery rate multiple testing correction. Four of these 5 transcripts were among the 12 transcripts selected in elastic net, suggesting that a relationship may exist between loneliness and expression of these transcripts which the linear regression approach may have been ill powered to adequately detect. The 12 transcripts selected via elastic net correspond to 12 unique genes (*IDO1, VNN1, THBS1, CAMP, IL1RN, CEBPB, ADORA2B, UNC13D, IL1B, LTA, CX3CR1, CCL5*). These genes have a range of functions in the chronic inflammatory biological process. For example, *CX3CR1* plays a role in leukocyte adhesion and migration, *IL1B* produces a protein product that is involved in cell proliferation, differentiation, and apoptosis, and *THBS1* is a glycoprotein mediator of cell-to-cell and cell-to-matrix interactions (“GeneCards - Human Genes | Gene Database | Gene Search,” n.d.). The present study builds upon the existing literature by providing evidence of associations between loneliness and gene expression in chronic inflammatory response genes while addressing some of the important limitations of previous work (e.g. small sample size, differing cell types).

Chronic burden has also previously been associated with altered immune functioning including inflammation (Sheldon Cohen et al., 2012; Ranjit et al., 2007). This relationship is hypothesized to act directly through the hypothalamic-pituitary-adrenal axis (HPA) and sympathetic-adrenal-medullary (SAM) system or indirectly through risk behaviors (Sheldon Cohen, Janicki-Deverts, & Miller, 2007; Ranjit et al., 2007; Tian, Hou, Li, & Yuan, 2014). In the MESA cohort, chronic burden has been found to significantly associate with the inflammatory proteins IL-6 and CRP (Ranjit et al., 2007). In the present study, we followed up the global analyses with linear regression and elastic net to identify which genes specifically were associated with chronic burden. Although no gene transcripts were significant in linear regression analyses at $p<0.05$, 8 transcripts were selected in elastic net corresponding to 8 unique genes (*VNN1, THBS1, IL1RN, ADORA2B, LTA, CX3CR1, CCL5, IDO1*). All 8 of these genes were also selected in the elastic net analyses assessing the association between loneliness and chronic inflammatory genes.

In global analyses, major or lifetime discrimination was the only measure that yielded significant findings across all four biological processes investigated (i.e. chronic

inflammation, inflammatory response, immune response, and regulation of inflammatory processes). Discrimination, characterized by differential treatment of certain groups, has been associated with increased risk of morbidity and mortality (David R Williams & Mohammed, 2009). In the MESA cohort, there has been mixed findings on the relationship between discrimination and inflammation at the protein level. A study by Kershaw (2016) found discrimination to be significantly associated with the inflammatory biomarker IL-6, but interestingly no significant association was detected for CRP (Kershaw et al., 2016). These findings were influenced by sex and body mass index. Although the genes encoding CRP and IL-6 were not part of the chronic inflammatory gene set, we were able to provide evidence that discrimination is associated with chronic inflammatory processes at a gene level in global analyses. In linear regression analyses, four transcripts were significant at $p<0.05$, but like the relationships for loneliness and chronic burden, there were no significant associations after multiple testing adjustment. Three transcripts were selected via elastic net (*CX3CR1, VNN1, CHAMP1*) all of which were significant at $p<0.05$ in linear regression analyses. Two of these three genes (i.e. *CX3CR1, VNN1*) were also selected in elastic net analyses for both loneliness and chronic burden.

It is notable that there was substantial overlap in the elastic net findings for the three social factors that were significant in Global ANCOVA analyses (i.e. loneliness, chronic burden, discrimination). All 8 genes that were selected for chronic burden in the chronic inflammatory gene set in elastic net analyses were also selected in the loneliness analyses. Similarly, out of the 3 transcripts selected for discrimination in the chronic inflammatory gene set, 2 were selected in the chronic burden and loneliness analyses. All in all, there were differences in which social factors were significantly associated with gene expression in global analyses, but when selected, the social factors tended to be associated with the same genes.

This study has notable strengths that contribute to the human social genomics literature and provide a basis for future research. One strength of the current study is the sample size of 1,264 participants. Similar studies have been limited by very small sample sizes (Antoni et al., 2012; Steve W Cole et al., 2007; Steven W Cole et al., 2012; Fredrickson et al., 2013; Kiecolt-Glaser et al., 2005; Miller et al., 2008; O'Connor et al., 2014; Powell et al., 2013; Vedhara et al., 2015). The use of the MESA cohort makes the present study one of the largest human social genomics studies. A larger sample size enhances the power to detect significant associations. Secondly, gene expression was measured in monocytes in MESA. Since gene expression is cell type specific, it is important to have it measured in a cell type that is known to play a role in inflammation. Monocytes are phagocytic leukocytes of the innate immune response, and evidence suggests that even among the various types of white blood cells that are part of the inflammatory response, monocytes, along with dendritic cells, are most susceptible to changes in gene expression with exposure to an adverse social environment (Steven W Cole et al., 2011). Third, this study examined associations between multiple social factors and gene expression on the same study sample whereas most previous studies have investigated associations between one social factor with gene expression patterning at a time. Examining multiple social factors simultaneously helps to provide insight as to which factors associate with gene expression patterning and whether such associations affect the same genes. This provides great opportunity for future research as we

seek to better categorize how the social environment affects gene expression. Fourth, this study also helps to move the field of human social genomics forward by the employment of three complementary methodological approaches. The methods were selected to minimize the total number of tests and account for the correlated nature of gene expression data. Employing these multiple approaches using the MESA data has helped to provide important insight into understanding the associations between social factors and inflammatory gene expression and can be used as a basis for future research.

Like all studies, we acknowledge that there are some limitations to this study. Most notably, we did not have a replication sample to verify the associations between the implicated social factors and gene expression. Replication reduces the concern for Type 1 error. Prior work suggests that replication studies are needed to identify significant, consistent associations between social factors and gene expression patterning (Brown et al., 2019). Secondly, given that the Global ANCOVA test is designed to assess whether any gene transcript is associated with the exposure, the finding that loneliness and chronic burden were associated with the chronic inflammation gene set but not the other gene sets was unexpected since there was substantial overlap of gene transcripts between gene sets. It is likely that this is in part due to the common signal-to-noise ratio issue inherent to microarray studies (Subramanian et al., 2005). With only 20 transcripts and a more focused biological process with stronger cross-correlation, we would expect the chronic inflammatory gene set to be less susceptible to this issue compared to the other three investigated biological processes. More sensitive gene expression technologies (e.g. RNA-seq) may be useful in overcoming this limitation in future studies (Zhao, Fung-Leung, Bittner, Ngo, & Liu, 2014). Thirdly, the social factors, gene expression, and covariates were collected at different time points throughout the MESA exams. Because gene expression values can be transient across time, the expression measures in Exam 5 may be different from the prior exams when most of the social factors were measured. Also, the time invariance of social factors as well as the lags with which social factors may affect gene expression may vary depending on the social factor in question. Thus, future work should incorporate longitudinal time varying measures of both social factors and gene expression. Fourthly, we did not specifically investigate potential mediators (e.g. body mass index, smoking, current disease states). Future studies could consider exploring the potential mediating pathways through which social factors associate with gene expression. Lastly, it is possible that a low level of exposure to the investigated social factors and a small amount of variation in social factor scores in the MESA study sample may have inhibited the ability to detect some associations. However, even with small variation in the scores of the social factors, some significant associations were detected. Future research using samples with a greater range of social factor responses could be useful in detecting more subtle associations.

Overall, this study provides evidence that expression of inflammation-related genes may be influenced by social factors. To identify the most salient social factors, future research should replicate these tests in diverse study samples and assess how the characteristics of social factors (e.g. chronicity, perceived impact) affect association with differential gene expression patterns. Subsequent studies should also identify the downstream biological effects of gene expression differences including its association with proteomic and metabolomic profiles and overall disease risk. Better understanding of the mechanisms

through which social factors act biologically will enable the development of more precise prevention and treatment approaches.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Analysis for the present study was supported by the Center for Integrative Approaches to Health Disparities (P60 MD002249), the Center for Research on Ethnicity, Culture, and Health (R25 GM058641) and A Social Epigenomic Approach to Health Disparities in Cardiovascular Risk Factors (RO1 HL141292) . MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420, UL1-TR-001881, and DK063491. The MESA Epigenomics and Transcriptomics Studies were funded by R01HL101250, R01 DK103531-01, R01 DK103531, R01 AG054474, and R01 HL135009-01 to Wake Forest University Health Sciences, and the MESA Stress II Ancillary Study was funded by R01HL101161. Kristen Brown was also partially funded by the Multidisciplinary Research Training to Reduce Inequalities in Cardiovascular Health Fellowship (T32 HL130025).

References:

Alspach JAG (2013). Loneliness and social isolation: Risk factors long overdue for surveillance. Critical Care Nurse. 10.4037/ccn2013377

Antoni MH, Lutgendorf SK, Blomberg B, Carver CS, Lechner S, Diaz A, … Cole SW (2012). Cognitive-behavioral stress management reverses anxiety-related leukocyte transcriptional dynamics. Biological Psychiatry, 71(4), 366–372. 10.1016/j.biopsych.2011.10.007 [PubMed: 22088795]

Bertoni AG, Burke GL, Owusu JA, Carnethon MR, Vaidya D, Barr RG, … Rotter JI (2010). Inflammation and the incidence of type 2 diabetes: the Multi-Ethnic Study of Atherosclerosis (MESA). Diabetes Care, 33(4), 804–810. 10.2337/dc09-1679 [PubMed: 20097779]

Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, … Tracy RP (2002). Multi-Ethnic Study of Atherosclerosis: Objectives and design. American Journal of Epidemiology, 156(9), 871–881. 10.1093/aje/kwf113 [PubMed: 12397006]

Brown KM, Diez-Roux AV, Smith JA, Needham BL, Mukherjee B, Ware EB, … Kardia SLR (2019). Expression of socially sensitive genes: The multi-ethnic study of atherosclerosis. PLOS ONE, 14(4), e0214061 10.1371/journal.pone.0214061 [PubMed: 30973896]

Busse D, Yim IS, Campos B, & Marshburn CK (2017). Discrimination and the HPA axis: current evidence and future directions. Journal of Behavioral Medicine. 10.1007/s10865-017-9830-6

Cohen S, Kamarck T, & Mermelstein R (1983). A global measure of perceived stress. Journal of Health and Social Behavior, 24, 385–396. 10.2307/2136404 [PubMed: 6668417]

Cohen Sheldon, Janicki-Deverts D, Doyle WJ, Miller GE, Frank E, Rabin BS, & Turner RB (2012). Chronic stress, glucocorticoid receptor resistance, inflammation, and disease risk. Proceedings of the National Academy of Sciences of the United States of America, 109(16), 5995–5999. 10.1073/pnas.1118355109 [PubMed: 22474371]

Cohen Sheldon, Janicki-Deverts D, & Miller GE (2007). Psychological stress and disease. JAMA: The Journal of the American Medical Association, 298, 1685–1687. 10.1001/jama.298.14.1685 [PubMed: 17925521]

Cohen Sheldon, & Williamson GM (1988). Perceived stress in a probability sample of the United States. In The Social Psychology of Health (Vol. 13, pp. 31–67). 10.1111/j.1559-1816.1983.tb02325.x

Cohen Sheldon, & Wills TA (1985). Stress, Social Support, and the Buffering Hypothesis. Psychological Bulletin. 10.1037/0033-2909.98.2.310

Cole Steve W. (2008). Social regulation of leukocyte homeostasis: The role of glucocorticoid sensitivity. Brain, Behavior, and Immunity, 22(7), 1049–1055. 10.1016/j.bbi.2008.02.006

Cole Steve W. (2010a). Elevating the perspective on human stress genomics. Psychoneuroendocrinology, 35(7), 955–962. 10.1016/J.PSYNEUEN.2010.06.008 [PubMed: 20630660]

Cole Steve W. (2010b). Elevating the perspective on human stress genomics. Psychoneuroendocrinology. 10.1016/j.psyneuen.2010.06.008

Cole Steve W, Hawkley LC, Arevalo JM, Sung CY, Rose RM, & Cacioppo JT (2007). Social regulation of gene expression in human leukocytes. Genome Biology, 8, R189 10.1186/gb-2007-8-9-r189 [PubMed: 17854483]

Cole Steven W., Levine ME, Arevalo JMG, Ma J, Weir DR, & Crimmins EM (2015). Loneliness, eudaimonia, and the human conserved transcriptional response to adversity. Psychoneuroendocrinology, 62, 11–17. 10.1016/j.psyneuen.2015.07.001 [PubMed: 26246388]

Cole Steven W., Capitanio JP, Chun K, Arevalo JMG, Ma J, & Cacioppo JT (2015). Myeloid differentiation architecture of leukocyte transcriptome dynamics in perceived social isolation. Proceedings of the National Academy of Sciences of the United States of America, 112(49), 15142–15147. 10.1073/pnas.1514249112 [PubMed: 26598672]

Cole Steven W., Conti G, Arevalo JMG, Ruggiero AM, Heckman JJ, & Suomi SJ (2012). Transcriptional modulation of the developing immune system by early life social adversity. Proceedings of the National Academy of Sciences of the United States of America, 109(50), 20578–20583. 10.1073/pnas.1218253109 [PubMed: 23184974]

Cole Steven W., Hawkley LC, Arevalo JMG, & Cacioppo JT (2011). Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes. Proceedings of the National Academy of Sciences of the United States of America, 108(7), 3080–3085. 10.1073/pnas.1014218108 [PubMed: 21300872]

Creswell JD, Irwin MR, Burklund LJ, Lieberman MD, Arevalo JMG, Ma J, … Cole SW (2012). Mindfulness-Based Stress Reduction training reduces loneliness and pro-inflammatory gene expression in older adults: A small randomized controlled trial. Brain, Behavior, and Immunity, 26(7), 1095–1101. 10.1016/j.bbi.2012.07.006

Dunning MJ, Smith ML, Ritchie ME, & Tavaré S (2007). Beadarray: R classes and methods for Illumina bead-based data. Bioinformatics, 23(16), 2183–2184. 10.1093/bioinformatics/btm311 [PubMed: 17586828]

FastStats - Deaths and Mortality. (n.d.). Retrieved from http://www.cdc.gov/nchs/fastats/deaths.htm

Fredrickson BL, Grewen KM, Coffey KA, Algoe SB, Firestine AM, Arevalo JMG, … Cole SW (2013). A functional genomic perspective on human well-being. Proceedings of the National Academy of Sciences of the United States of America, 110(33), 13684–13689. 10.1073/pnas.1305419110 [PubMed: 23898182]

GeneCards - Human Genes | Gene Database | Gene Search. (n.d.). Retrieved May 19, 2017, from http://www.genecards.org/

Grivennikov SI, Greten FR, & Karin M (2010). Immunity, Inflammation, and Cancer. Cell. 10.1016/j.cell.2010.01.025

Hackett RA, Hamer M, Endrighi R, Brydon L, & Steptoe A (2012). Loneliness and stress-related inflammatory and neuroendocrine responses in older men and women. Psychoneuroendocrinology, 37(11), 1801–1809. 10.1016/j.psyneuen.2012.03.016 [PubMed: 22503139]

Hänsel A, Hong S, Cámara RJA, & von Känel R (2010). Inflammation as a psychophysiological biomarker in chronic psychosocial stress. Neuroscience and Biobehavioral Reviews, 35(1), 115–121. 10.1016/j.neubiorev.2009.12.012 [PubMed: 20026349]

Hawkley LC, & Cacioppo JT (2010). Loneliness Matters: A Theoretical and Empirical Review of Consequences and Mechanisms. Annals of Behavioral Medicine, 40(2), 218–227. 10.1007/s12160-010-9210-8 [PubMed: 20652462]

Hummel M, Meister R, & Mansmann U (2008). GlobalANCOVA: exploration and assessment of gene group effects. Bioinformatics (Oxford, England), 24(1), 78–85. 10.1093/bioinformatics/btm531

Jaremka LM, Fagundes CP, Peng J, Bennett JM, Glaser R, Malarkey WB, & Kiecolt-Glaser JK (2013). Loneliness promotes inflammation during acute stress. Psychological Science, 24(7), 1089–1097. 10.1177/0956797612464059 [PubMed: 23630220]

Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, Collins R, … Wood AM (2010). C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: An individual participant meta-analysis. The Lancet. 10.1016/S0140-6736(09)61717-7

Kershaw KN, Lewis TT, Roux AVD, Jenny NS, Liu K, Penedo FJ, & Carnethon MR (2016). Self-reported experiences of discrimination and inflammation among men and women: The multi-ethnic study of atherosclerosis. Health Psychology, 35(4), 343–350. 10.1037/hea0000331 [PubMed: 27018725]

Kiecolt-Glaser JK, Loving TJ, Stowell JR, Malarkey WB, Lemeshow S, Dickinson SL, & Glaser R (2005). Hostile marital interactions, proinflammatory cytokine production, and wound healing. Archives of General Psychiatry, 62, 1377–1384. 10.1001/archpsyc.62.12.1377 [PubMed: 16330726]

Kleiveland C, & Kleiveland C (2015). Peripheral blood mononuclear cells. In The Impact of Food Bioactives on Health: In Vitro and Ex Vivo Models. 10.1007/978-3-319-16104-4_15

Link BG, & Phelan J (1995). Social conditions as fundamental causes of disease. Journal of Health and Social Behavior, Spec No, 80–94. 10.2307/2626958

Liu Y, Ding J, Reynolds LM, Lohman K, Register TC, De la Fuente A, … Hoeschele I (2013). Methylomics of gene expression in human monocytes. Human Molecular Genetics, 22, 5065–5074. 10.1093/hmg/ddt356 [PubMed: 23900078]

Marsik C, Kazemi-Shirazi L, Schickbauer T, Winkler S, Joukhadar C, Wagner OF, & Endler G (2008). C-Reactive Protein and All-Cause Mortality in a Large Hospital-Based Cohort. Clinical Chemistry, 54(2). Retrieved from http://clinchem.aaccjnls.org/content/54/2/343.short

Mezuk B, Choi M, DeSantis AS, Rapp SR, Diez Roux AV, & Seeman T (2016). Loneliness, Depression, and Inflammation: Evidence from the Multi-Ethnic Study of Atherosclerosis. PLOS ONE, 11(7), e0158056 10.1371/journal.pone.0158056 [PubMed: 27367428]

Miller GE, Chen E, Sze J, Marin T, Arevalo JMG, Doll R, … Cole SW (2008). A Functional Genomic Fingerprint of Chronic Stress in Humans: Blunted Glucocorticoid and Increased NF-??B Signaling. Biological Psychiatry, 64, 266–272. 10.1016/j.biopsych.2008.03.017 [PubMed: 18440494]

Monadi M, Firouzjahi A, Hosseini A, Javadian Y, Sharbatdaran M, & Heidari B (2016). Serum C-reactive protein in asthma and its ability in predicting asthma control, A case-control study. Caspian Journal of Internal Medicine.

O'Connor M-F, Schultze-Florey CR, Irwin MR, Arevalo JMG, & Cole SW (2014). Divergent gene expression responses to complicated grief and non-complicated grief. Brain, Behavior, and Immunity, 37, 78–83. 10.1016/j.bbi.2013.12.017

Ozbay F, Johnson DC, Dimoulas E, Morgan CA, Charney D, & Southwick S (2007). Social support and resilience to stress: from neurobiology to clinical practice. Psychiatry (Edgmont (Pa.: Township)).

Pilkonis PA, Imber SD, & Rubinsky P (1985). Dimensions of life stress in psychiatric patients. Journal of Human Stress, 11, 5–10. 10.1080/0097840X.1985.9936732 [PubMed: 3843107]

Powell ND, Sloan EK, Bailey MT, Arevalo JMG, Miller GE, Chen E, … Cole SW (2013). Social stress up-regulates inflammatory gene expression in the leukocyte transcriptome via β-adrenergic induction of myelopoiesis. Proceedings of the National Academy of Sciences of the United States of America, 110(41), 16574–16579. 10.1073/pnas.1310655110 [PubMed: 24062448]

Racanelli AC, Kikkers SA, Choi AMK, & Cloonan SM (2018). Autophagy and inflammation in chronic respiratory disease. Autophagy. 10.1080/15548627.2017.1389823

Ranjit N, Diez-Roux AV, Shea S, Cushman M, Ni H, & Seeman T (2007). Socioeconomic position, race/ethnicity, and inflammation in the multi-ethnic study of atherosclerosis. Circulation, 116(21), 2383–2390. 10.1161/CIRCULATIONAHA.107.706226 [PubMed: 18025402]

Ridker PM (2007). Inflammatory Biomarkers and Risks of Myocardial Infarction, Stroke, Diabetes, and Total Mortality: Implications for Longevity. 253–259. 10.1301/nr.2007.dec.S253-S259

Rubio-Perez JM, & Morillas-Ruiz JM (2012). A review: Inflammatory process in Alzheimer's disease, role of cytokines. The Scientific World Journal. 10.1100/2012/756357

Russell D, Peplau L. a, & Cutrona CE (1980). The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. Journal of Personality and Social Psychology, 39(3), 472–480. 10.1037/0022-3514.39.3.472 [PubMed: 7431205]

Smyth GK, Michaud J, & Scott HS (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics, 21(9), 2067–2075. 10.1093/bioinformatics/bti270 [PubMed: 15657102]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette M. a, … Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America, 102, 15545–15550. 10.1073/pnas.0506580102 [PubMed: 16199517]

Tian R, Hou G, Li D, & Yuan TF (2014). A possible change process of inflammatory cytokines in the prolonged chronic stress and its ultimate implications for health. Scientific World Journal, Vol. 2014 10.1155/2014/780616

Tukey JW (1949). Comparing Individual Means in the Analysis of Variance. Source: Biometrics, 5(2), 99–114. Retrieved from http://www.jstor.org/stable/3001913

Valtorta NK, Kanaan M, Gilbody S, Ronzi S, & Hanratty B (n.d.). Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. 10.1136/heartjnl-2015-308790

Vedhara K, Gill S, Eldesouky L, Campbell BK, Arevalo JMG, Ma J, & Cole SW (2015). Personality and gene expression: Do individual differences exist in the leukocyte transcriptome? Psychoneuroendocrinology, 52, 72–82. 10.1016/j.psyneuen.2014.10.028 [PubMed: 25459894]

Williams DR, Yan Yu, Jackson JS, & Anderson NB (1997). Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination. Journal of Health Psychology, Vol. 2, pp. 335–351. 10.1177/135910539700200305 [PubMed: 22013026]

Williams David R., & Mohammed SA (2009). Discrimination and racial disparities in health: evidence and needed research. Journal of Behavioral Medicine, 32(1), 20–47. 10.1007/s10865-008-9185-0 [PubMed: 19030981]

Yang YC, Schorpp K, & Harris KM (2014). Social support, social strain and inflammation: evidence from a national longitudinal study of U.S. adults. Social Science & Medicine, 107, 124–135. 10.1016/j.socscimed.2014.02.013 [PubMed: 24607674]

Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Sillesen H, & Nordestgaard BG (2008). Genetically elevated C-reactive protein and ischemic vascular disease. The New England Journal of Medicine, 359, 1897–1908. 10.1056/NEJMoa0707402 [PubMed: 18971492]

Zhao S, Fung-Leung WP, Bittner A, Ngo K, & Liu X (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS ONE. 10.1371/journal.pone.0078644

Zou H, & Hastie T (2005). Regularization and variable selection via the elastic net. J. R. Statist. Soc. B, 67(2), 301–320.

Table 1.

Social factors assessed in the present study

Social factor	Measurement in the MESA study (# items, if applicable)	Exam Collected
Adult Socioeconomic Status	Highest level of education	1
Childhood Socioeconomic Status	Highest level of education for either parent	2
Loneliness	Scale adapted from the UCLA Loneliness Scale, 3 items (Russell, Peplau, & Cutrona, 1980)	4
Perceived Stress	Perceived Stress Scale, 4 items (S Cohen et al., 1983; Sheldon Cohen & Williamson, 1988)	5
Chronic Burden	Chronic Burden Scale, 5 items (Pilkonis et al., 1985)	3
Major or Lifetime Discrimination	Scale adapted from the Detroit Area Study, 6 items (D. R. Williams, Yan Yu, Jackson, & Anderson, 1997)	1
Social Support	Scale adapted from the MIDUS study, 4 items (Yang, Schorpp, & Harris, 2014)	4
Covariates (age, sex, and race/ethnicity)	Self-reported	1

Table 2.

Characteristics of the MESA Study Sample with Gene Expression Data

	Total sample (n=1264)	Non-Hispanic Black (n=272)	Hispanics (n=402)	Non-Hispanic White (n=590)
Demographics				
Age mean (SD)	69.6 (9.4)	69.6 (9.0)	68.4 (9.3)	70.2 (9.5)
Sex (% female)	51%	60%	50%	48%
Study Site *(n,%)*				
Forsythe County, North Carolina	49 (4%)	1 (0.3%)	0	48 (40%)
New York, New York	424 (34%)	131 (48%)	209 (52%)	84 (14%)
Baltimore, Maryland	317 (25%)	140 (51%)	0	177 (30%)
Rochester, Minnesota	474 (38%)	0	193 (48%)	281 (48%)
Total	1264 (100%)	272 (100%)	402 (100%)	590 (100%)
Socioeconomic Status				
High education- participant[a]	33%	27%	13%	49%
High education- either parent[b]	56%	55%	31%	71%
Psychosocial factors *Median, (Interquartile Range)*				
Loneliness	3 (3-5)	3 (3-4)	4 (3-5)	3 (3-5)
Major or Lifetime Discrimination	0 (0-1)	1 (0-2)	0 (0-1)	0 (0-1)
Chronic Burden	1 (0-2)	1 (0-2)	0 (0-2)	1 (0-2)
Perceived Stress	8 (5-10)	8 (5-10)	8 (5-10)	7 (6-10)
Social Support	9 (8-10)	9 (8-10)	9 (8-10)	9 (8-10)

[a]High education for participant indicates an educational level of college or greater

[b]High parental education indicates an educational level of high school or greater

Table 3.

Global ANCOVA p-values for the Association between Social Environmental Factors and the Four Gene Ontology Biological Processes

	Chronic Inflammation (k=20)[a]	**Inflammatory Response (k=438)[a]**	**Immune Response (k=1251)[a]**	**Regulation of Inflammatory Response (k=192)[a]**
Loneliness	**0.003**	0.229	0.227	0.173
Discrimination	**0.045**	**0.029**	**0.041**	**0.025**
Perceived Stress	0.092	0.123	0.179	0.181
Chronic Burden	**0.002**	0.188	0.135	0.312
Social Support	0.053	0.571	0.599	0.344
Adult SES	0.110	0.300	0.179	0.076
Childhood SES	0.403	0.548	0.506	0.467

[a] k indicates the number of transcripts in the gene set

Table 4.

Results from Linear Regression and Elastic Net Analyses

Exposure	Gene Set	# of transcripts significant at p<0.05 in linear regression analyses	# transcripts selected via elastic net	Overlap between linear regression and elastic net findings
Loneliness	Chronic Inflammation (k=20)[a]	5	12	4
Chronic Burden	Chronic Inflammation (k=20)[a]	0	8	0
Discrimination	Chronic Inflammation (k=20)[a]	4	3	3
Discrimination	Inflammatory Response(k=438)[a]	53	17	17
Discrimination	Immune Response (k=1251)[a]	136	42	36
Discrimination	Regulation of the Inflammatory Response (k=192)[a]	26	15	15
Overlap of total unique transcripts	Total: 1407	153	53	43

[a] k indicates the number of transcripts in the gene set