Poverty shapes the transcriptome of immune cells

Nicole S. Arnold[1,2], Justyna Resztak[2], Adnan Alazizi[2], Nicole Noren Hooten[4], Michele K. Evans[4], Valerie Odero-Marah[1], Douglas F. Dluzen[3]*, Roger Pique-Regi[2,*], Francesca Luca[2,5,*]

Morgan State University[1], Wayne State University[2], Johns Hopkins University[3], National Institute on Aging, National Institutes of Health[4], University of Rome Tor Vergata[5]

*To whom correspondence should be addressed

Abstract

Social determinants of health influence health outcomes and life expectancy. Specifically, individuals living in poverty often have adverse health outcomes related to chronic inflammation that affect the cardiovascular, renal, and pulmonary systems. However, the underlying mechanisms by which poverty increases the risk of disease and health disparities are still not fully understood. To bridge the gap in our understanding of the link between living in poverty and adverse health outcomes, we performed RNA sequencing of blood immune cells from 204 participants of the Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study in Baltimore City, Maryland. This study cohort included men and women self-identified as African American and White. We identified 581 genes differentially expressed in association with poverty. A larger number of differentially expressed genes were detected in women, compared to men, and 64 genes had distinct sex-by-poverty interaction effects. Genes differentially expressed in women living in poverty were enriched in wound healing and coagulation processes, while in men were mostly related to immunoglobulin production and humoral immune response. Of the genes differentially expressed in individuals living in poverty, 275 are also associated with complex diseases in transcriptome-wide association studies. Our results suggest that living in poverty influences inflammation and the risk for chronic disease through gene expression changes in immune cells, and that some of the effects of living in poverty are different in women and men.

Introduction

Socioeconomic inequalities in health and mortality are very well documented[1–3]. These inequalities have been observed at similar levels since the early nineteenth century. More recently health inequalities are also reflected in current major causes of death including cancer and cardiovascular disease[4]. Individuals living in poverty are more frequently exposed to risk factors for these conditions including poor diet, inadequate exercise, and smoking. Additionally, psychosocial factors that may influence health status are also common when living in poverty, such as housing insecurity, discrimination, structural racism, and lack of social support [5]. Living below the federal poverty line increases cardiovascular risk[6] and low socioeconomic status (SES) is associated with increase of cardiovascular disease biomarkers in the circulation, including red-blood-cell width[7], high-density lipoprotein cholesterol (HDL-C)[8], intimal medial thickness of the arteries and pulse-wave velocity[9]. These risks are also further amplified in vulnerable groups such as African American adults[9]. Low SES is associated with a diminished ability to suppress CMV infection and maintain quiescence.[10], It has consistently been associated with a greater risk of infectious and cardiovascular diseases, more rapid progression, and decreased survival[4]. The biological mechanisms underlying these inequalities still need to be better understood. Elevated innate immunity responsiveness is observed in low SES children when compared to children that live in a higher SES[11]. In a broad cohort across the United States, low childhood SES was linked to poorer metabolic outcomes in midlife[12]. For example, participants raised in families where neither parent finished high school were 1.4 times more likely to develop metabolic syndrome than counterparts raised in university-educated households[12].
Adverse health is consistently linked to lower SES, but little is known about the molecular mechanisms behind this disparity. Subjects from low-income families have higher cortisol output in daily life, higher expression of transcripts containing NFkappaB response components, and higher production of the proinflammatory cytokine interleukin 6[13]. Children from low-income homes have increased cortisol levels in adulthood, regardless of social standing[13,14], which could potentially lead to resistance to glucocorticoid signaling and allow for increased adrenocortical and inflammatory responses[13]. Additionally, metabolic syndrome components such as central obesity, insulin resistance, and blood pressure can worsen with chronic cortisol exposure[15].

Social genomics investigates the various dimensions of a person's social and psychological environment that influence gene expression[16–20]. Numerous studies have linked changes in gene expression in whole blood to socioeconomic status and psychosocial experiences[21–2728]. Along with the single-gene analysis, earlier studies in this field[16,17,19] identified a pattern of differentially expressed genes known as the conserved transcriptional response to adversity (CTRA). The CTRA comprises genes in stress response pathways mediated by the sympathetic nervous system and beta-adrenergic signaling to induce the immune system's transcriptional activation of inflammatory genes[29,30]. However, socioeconomic status and psychosocial experiences have a broader effect on immune cell gene expression and impact additional pathways beyond those included in the CTRA[28].
A systematic genome-wide characterization of gene expression in immune cells for individuals living in poverty and stratified by age, self-identified race, and sex has not been done before. Importantly, there are few studies that examine the influence of chronic low SES within the immune cells of individuals living in poverty. Yet, such a characterization is essential to understand the biological mechanisms underlying poor health outcomes and health disparities in immune and cardiovascular diseases among individuals with low SES. Here we hypothesized that differential gene expression is associated with poverty status and that understanding the molecular pathways affected by poverty may shed light on health disparities in disease risk and outcome. We analyzed a sub-cohort of participants from the National Institute on Aging's (NIA)

Healthy Aging in Neighborhoods of Diversity Across the Life Span (HANDLS) study in Baltimore, MD. To determine the association between poverty status and gene expression, we studied 204 participants from the HANDLS cohort and performed RNA-sequencing on peripheral blood mononuclear cells (PBMCs) collected at two different time points (Figure 1).

Results

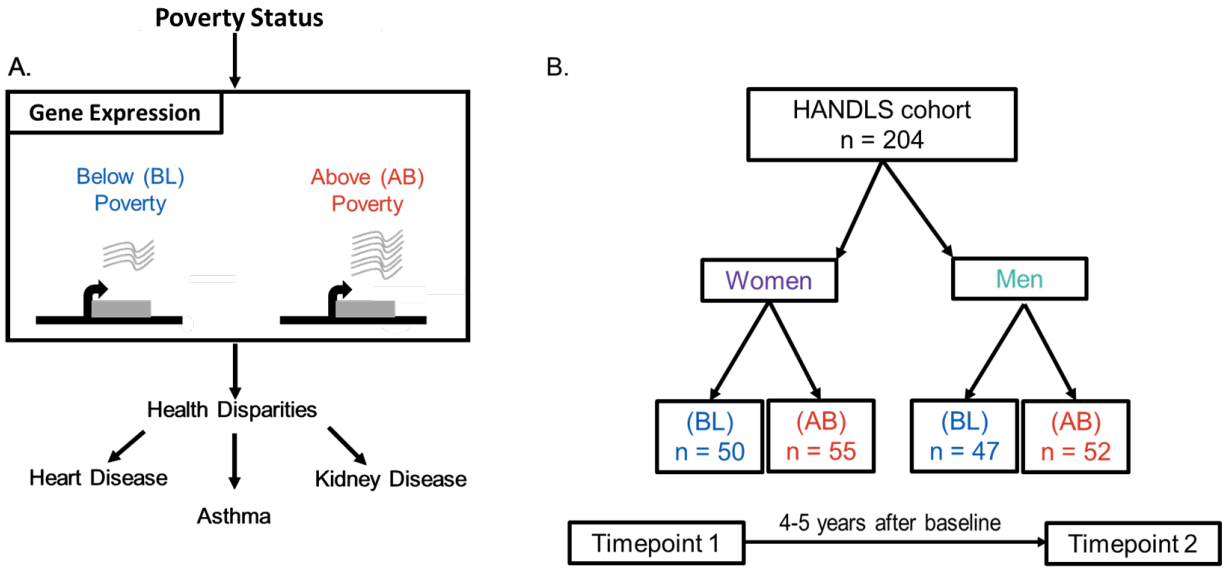


Figure 1: *Study Design*. A. Conceptual framework showing the workflow of the study to examine the relationship between poverty status and gene expression patterns leading to health disparities. In this study poverty status was determined as 125% above or below the federal poverty line set for 2004. Below the poverty line *BL* is shown in blue, and above the poverty line *AB* is shown in red. B. Study cohort. The sample cohort consists of 204 participants. Blood samples were collected at two different timepoints ~ 4-5 years apart for a total sample size of 408. Women (purple) n=105, men (green) n=99, below poverty line n=97, above poverty line n=107.

Poverty status influences gene expression in the immune system

To identify genes differentially expressed between individuals living in poverty and those living above poverty, we analyzed differential gene expression using DESeq2 within each time point. We controlled for technical variables, sex, age, and the first three genotype PCs. Note that samples cluster into two groups along the first PC (which explains 34% of the variance) and the groups correspond to self-identified race, while the other PCs likely capture other ancestry components (Supplementary Figure 1). We observed 396 differentially expressed genes in the first timepoint (FDR<10%); however, at the same FDR threshold, we found only 2 genes in the second time point. However, when we assessed reproducibility we found that 128 of the 396 genes are differentially expressed at the second time point at a nominal $p<0.05$. To investigate whether we have limited power to detect associations between poverty and gene expression, we

asked whether we could find consistent effects between the two time points, even if they did not meet the genome-wide significance threshold. This provides a method to assess reproducibility. When comparing the two timepoints, we found a strong correlation in the gene expression log-fold change (r^2=0.5, p<0.001) (Figure 2A, Supplementary Table 1). This result confirms that the gene expression changes associated with poverty are largely consistent across time points. To leverage the information from both time points and increase our statistical power, we analyzed the entire dataset and considered the two timepoints as replicate measurements. We identified 581 genes differentially expressed (FDR=10%) in individuals living in poverty compared to those above poverty(Supplementary Table 2).

We found that approximately 67% of the differentially expressed genes had lower expression in individuals living in poverty (Figure 2B). We performed an overrepresentation analysis to identify the biological processes impacted by poverty. We found an enrichment for biological processes pertaining to wound healing, blood coagulation, platelet activation, focal adhesion, Rap 1 signaling pathway, and serotonergic synapse (Figure 2C). The two genes with the most significant differences in gene expression in individuals living in poverty are *CABP5* (calcium binding protein 5) and *MGAM2 (*Maltase-glucoamylase 2) (Figure 2D). The product of the *CABP5* gene belongs to a subfamily of calcium-binding proteins, which share similarities to calmodulin[38,39]. Calcium-binding proteins are an essential component of calcium-mediated cellular signal transduction. *MGAM2* is a protein coding gene involved in glycogen metabolism and degradation for carbohydrate digestion and absorption[40]. Individuals living in poverty had a lower expression of both these genes when compared to those living above poverty.

Effects of poverty on gene expression in men and women

Sex differences in physiology and pathology, along with sex steroid receptor signaling have an influence on health outcomes and contribute to life expectancy differences between men and women[41–43]. Sex hormones modulate the inflammatory system, and sexual dimorphism drives inflammatory activity[44,45]. Sex bias occurs in tissue-specific gene expression and may affect specific biological functions in immune cells[46]. To investigate whether men and women experience different transcriptional effects due to poverty, we split our study cohort by sex and investigated gene expression patterns associated with poverty in men and women separately. We found 804 differentially expressed genes in women, 324 upregulated and 480 downregulated in women living in poverty (Supplementary Figure 3A). However, we found only 74 differentially expressed genes in men (69% displayed higher expression in poverty) (Figure 3A, Supplementary Figure 3B). Overall, we found that the effect of poverty on gene expression is correlated between sexes (Spearman's correlation ρ=0.18, p = 4.9 x 10^{-134}, Supplementary Figure 3C), with only 4 genes being differentially expressed both in men and women at FDR<10%. Genes differentially expressed in women living in poverty were enriched for the biological processes of wound healing, coagulation, platelet activation, and clot formation; while in men we identified enrichment for a different set of biological processes, including immunoglobulin production, receptor-mediated endocytosis, and complement activation. These results indicate that living in poverty may affect different molecular mechanisms and biological functions in men and women with potential sex-specific health disparity outcomes.

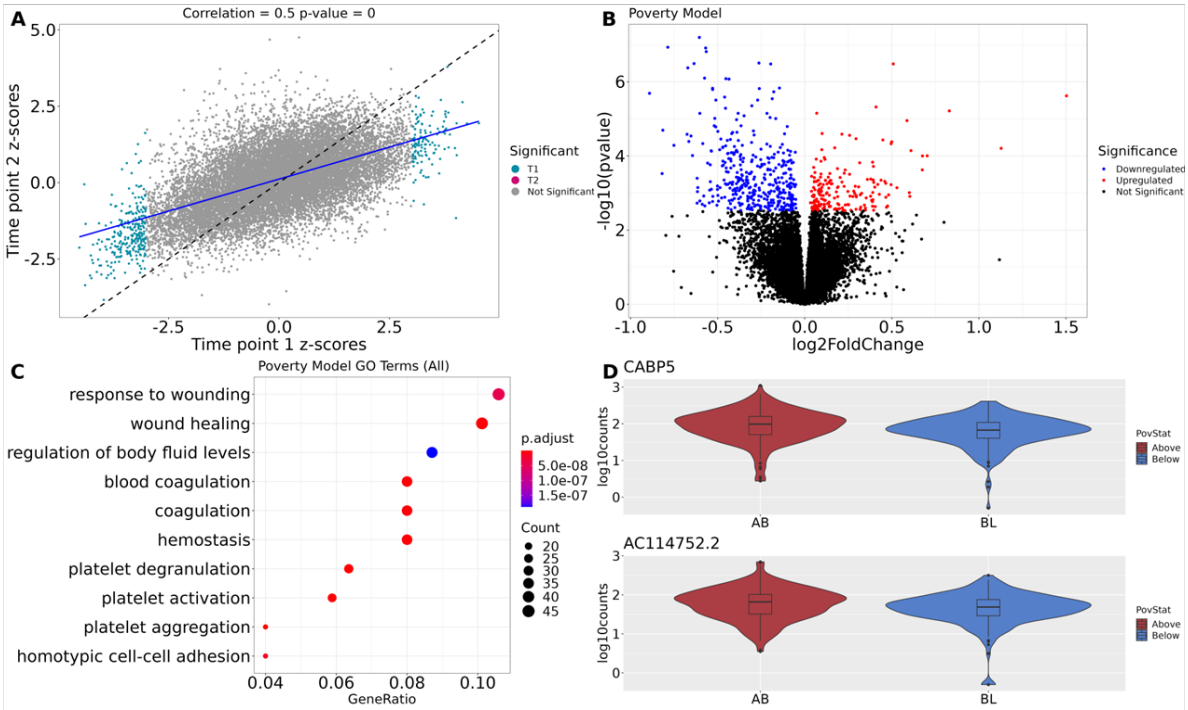


Figure 2: *Poverty status influences gene expression in immune cells*. A. Scatter plot of the correlation between gene expression log fold changes estimated for each time point separately. Significantly differentially expressed genes (FDR<10%) are indicated with colors: Time point 1 (blue), and time point 2 (pink). B. Volcano plot of the differentially expressed genes in the entire cohort, while considering the time points as replicate measurements. The 391 downregulated genes are indicated in blue and the 190 upregulated genes are in red. C. Results of the enrichment analysis for the gene ontology terms in the differentially expressed genes. The X-axis displays the proportion of gene (gene ratio), y-axis gene ontology terms (biological processes), legend indicates the number of associated genes (n=10-25), and the p.adjust legend. D. Examples of differentially expressed genes. The x-axis indicates the poverty status (as in Figure 1), the y-axis indicates the normalized log 10 expression counts.

Sex-Specific Effects of Poverty on Gene Expression

The observed differences between men and women in gene expression associated with poverty status suggest sex-specific effects. We tested this hypothesis by directly testing for an interaction between sex and poverty status on immune cell gene expression (model 2). We identified 64 differentially expressed genes, 23 upregulated in women living in poverty and 41 downregulated in women living in poverty. Two examples of the most significant genes in this model are *DCLK2* (Doublecortin Like Kinase 2) and *ETV7* (ETS variant transcription factor 7)(Figure 3 E, F). We observed a higher expression of *DCLK2* in women living in poverty when compared to men living in poverty. *DCLK2* is a gene encoding for protein kinase which plays an important role in binding and regulating microtubule polymerization. *DCLK2,* a sex-specific gene, is associated with cardiovascular and metabolic health involved with lipid-related pathways[47]. In addition, high expression of *DCLK2* is associated with an increased risk of chronic lymphocytic leukemia[48]. *ETV7* is upregulated in men living in poverty, compared to women. *ETV7* encodes for transcription

factors functioning in development and differentiation. *ETV7* ETS family of transcription factors produces an evolutionarily conserved protein. Members of this family control a broad range of cellular processes throughout development, differentiation, and oncogenesis. This gene encodes multiple isoforms and has several alternatively spliced transcript variants. The biological processes that the gene *ETV7* affects include animal organ morphogenesis, RNA polymerase II regulation, transcription, and negative regulation of transcription[49–51].

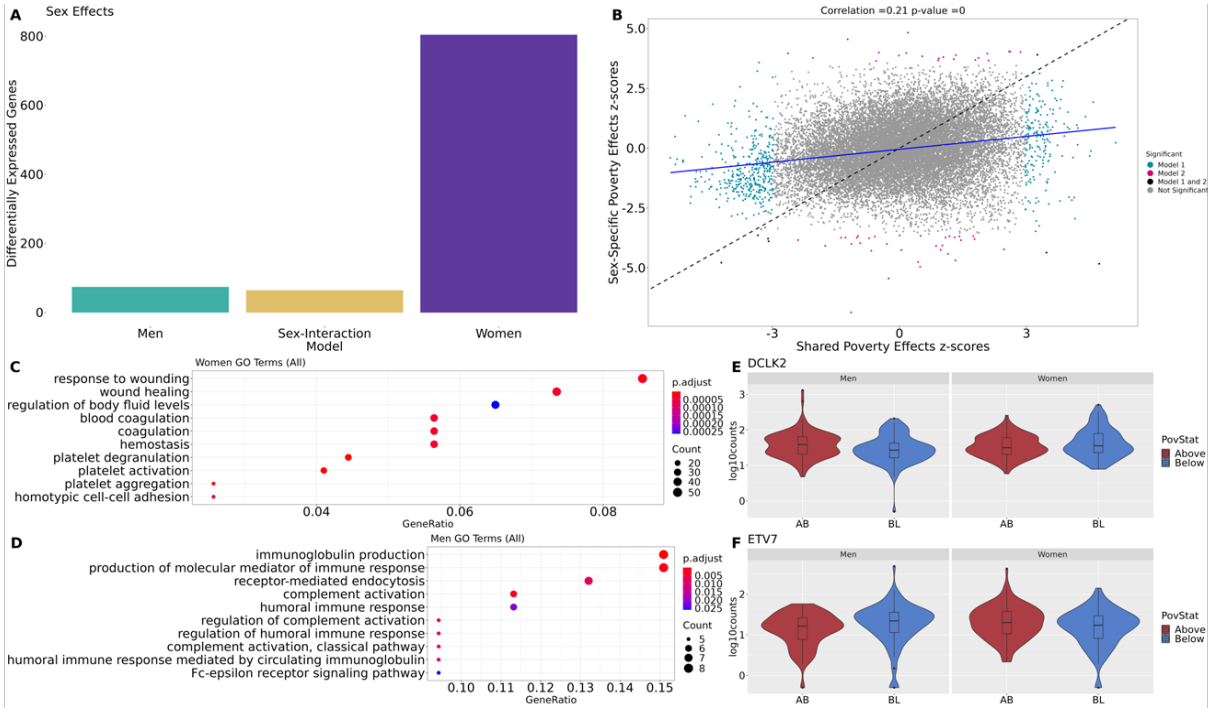


Figure 3: ***Association between poverty and gene expression in men and women.*** A. Bar plot of differentially expressed genes in men (green), and women (purple) living in poverty, and from the sex-by-poverty interaction model (yellow) (model 2). B. Scatter plot of the z-scores for the differentially expressed genes with shared poverty effects between sexes (x-axis, model 1) and with sex-specific poverty effects (y-axis, model 2). Significant DEGs found in model 1(green), model 2 (pink), significant in both model 1 and 2 (black) and not significant (gray). The z-scores were weakly correlated (r=0.18 p<0.001). C-D. Gene ontology terms enriched in the genes differentially expressed in women and men living in poverty. The X-axis displays the proportion of gene (gene ratio), y-axis gene ontology terms (biological processes), legend indicates the number of associated genes (n=10-25), and the p.adjust legend (Bonferroni correction). E-F. Examples of genes with significant sex-by-poverty effects in gene expression. The x-axis indicates the poverty status (as in Figure 1), and the y-axis indicates the normalized log 10 gene expression counts. In both plots, the first panel shows gene expression in men, and the second panel in women.

Understanding diseases and conditions influenced by poverty through transcriptional changes

To examine the adverse effects of poverty on disease risk that may be mediated by changes in gene expression in immune cells (Figure 1A), we considered genes associated with disease risk

through a probabilistic transcriptome-wide association study (PTWAS[37]). We focused on traits with an inflammatory or immunological component (Table 1) and identified 70 genes associated with disease risk that are also differentially expressed in individuals living in poverty compared to those living above the poverty line. Positive PTWAS *z*-scores indicate that increased gene expression is associated with increased risk for the disease considered. The direction of the DESeq2 z-scores was compared to the PTWAS z-scores to examine for concordant and discordant genetic effects. We define as concordant[52] the instances where the PTWAS effect and the poverty effect on gene expression are in the same direction, i.e., poverty is associated with high expression, and increased gene expression is associated with increased risk. We define as discordant the instances where the PTWAS effect and the poverty effect on gene expression are in opposite directions, i.e., poverty is associated with high expression, and increased gene expression is associated with decreased risk. Note that the direction of the PTWAS effect also reflects the original context of the eQTL study and does not capture epistasis, therefore caution should be exercised in over-interpreting the sign of the z-score.

We found more than 10 genes associated with asthma (n=12), eczema (n=11) and hypertension (n=27) (Table 1, Supplementary Table 2). For example, lower expression of *GGTA1P*, *AGPAT1, CFAP161, UNC13D, and WBP2* were associated with increased asthma risk (Figure 4), and we also find lower expression of these genes in individuals living in poverty compared to those living above poverty line, thus suggesting that poverty may impact the molecular pathways that lead to asthma. We also find concordant transcriptional effects of poverty and hypertension for the *VCL* and *VIL1* genes. We observed lower expression of these genes in individuals living in poverty, and low expression is also associated with higher risk for hypertension[53]. For other genes, instead, the transcriptional effect of poverty is discordant with disease risk.

Out of the 64 genes with a significant interaction between sex and poverty status, we found 9 genes associated with disease risk[37] (Figure 4), including in asthma (n=2), eczema (n=3), hypertension (n=3), and multiple sclerosis (n=1) (Supplementary Table 3).

Traits	Concordance	Discordance
Alzheimer's Disease	1	0
Asthma	15	22
Coronary Artery Disease	6	2
Crohn's Disease	6	2
Eczema	15	12
Hypertension	24	25
Lupus	4	2
Multiple Sclerosis	6	1
Psoriasis	1	4
Rheumatoid Arthritis	0	1

Table 1: Clinical significance of the genes differentially expressed with poverty. Number of genes (column 2) and differentially expressed genes with poverty (column 3) associated with selected complex traits (column 1) in PTWAS[37]. Columns 4 and 5 indicate whether the association of gene expression with the complex trait is in the same direction (concordant) or opposite direction (discordant) compared to the association between gene expression and poverty status.

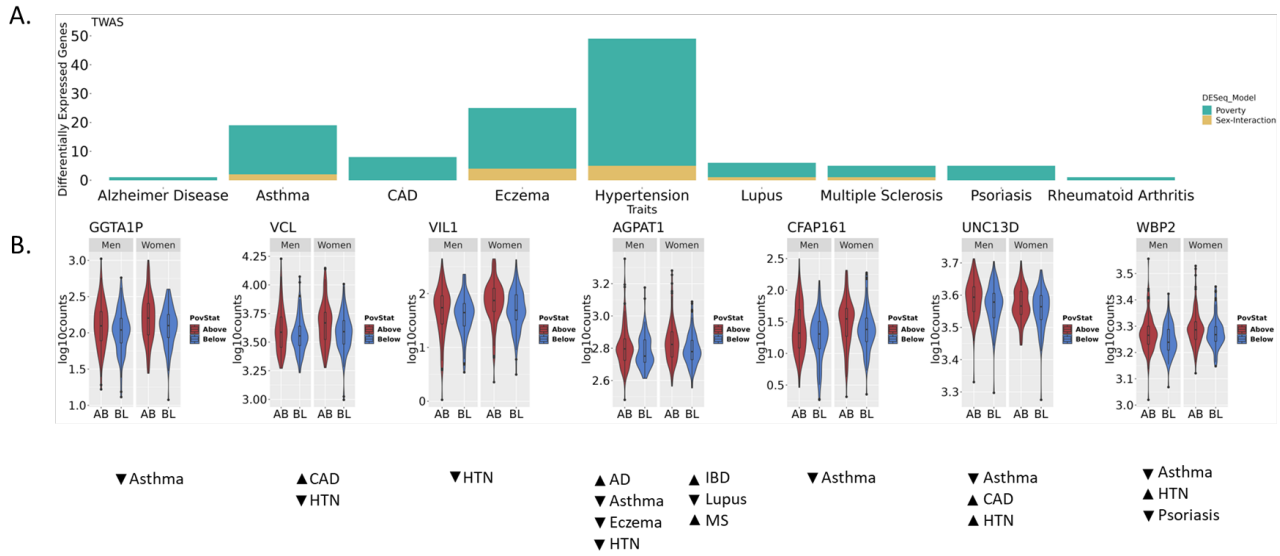


Figure 4: ***Genes associated with poverty status are also associated with disease risk.*** A. Bar plot of the inflammatory traits that are also associated with disease risk in PTWAS. The x-axis lists the traits and the y-axis lists the number of differentially expressed genes in the poverty (green) and poverty by sex-interaction (gold) models. B. Violin plots of the differentially expressed genes significantly associated with poverty status (as in Figure 1). The poverty status is displayed on the x-axis, and normalized log 10 counts on the y-axis.

Discussion

Our study advances the knowledge of how poverty may influence health, using a well-balanced, inner-city cohort. Using RNA-sequencing in a large cohort of 204 individuals, we found 581 genes differentially expressed between individuals living above or below 125% of the federal poverty line. We discovered that in women living in poverty, genes differentially expressed were enriched for coagulation and wound healing processes. In contrast, many of the genes linked to poverty in men are related to immunoglobulin production and humoral immune response. We validated the sex-specific impacts of poverty on gene expression by directly testing for an interaction between sex and poverty, and discovered 64 genes with differential expression. Poverty may contribute to health disparities in asthma, coronary artery disease, eczema, hypertension, inflammatory bowel disease, lupus, multiple sclerosis, psoriasis, and rheumatoid arthritis through gene expression changes in immune cells. For example, differential gene expression of *AGAPT1* was associated with increased risk for Alzheimer's Disease, eczema, hypertension, inflammatory bowel disease, lupus, multiple sclerosis, and asthma (Figure 4). Our findings further support that living in poverty has adverse consequences that may influence immunological health and chronic inflammatory disorders and identifies several important pathways through changes in gene expression.

Our previously published work conducted on a smaller subset of 52 individuals[22] found suggestive indication of an association between poverty and immune cell gene expression. Here, using a larger cohort study design, we found many genes differentially expressed. Our results show that it is possible to identify transcriptional signatures associated with poverty in well powered population studies. Furthermore, these effects are detected consistently across two time points.

We identified wound healing among the significant biological processes dysregulated in individuals that live in poverty. The role of immune cells in wound healing is critical for integrity restoration of damaged tissues[54]. Chronic inflammation of immune cells delays tissue wound healing repair[54,55]. Monocytes are one of the cell types in PBMCs and single cell studies have shown that their transcriptional status influences other PBMC components (e.g. T cells)[56,57]. Lack of plasticity of monocytes/macrophages in the inflammatory to anti-inflammatory signaling promotes ineffective wound healing, which is a main complication of type 2 Diabetes[58]. In addition, chronic pro-inflammatory states of monocytes have been linked to cardiovascular diseases like coronary artery disease and myocardial infarction[59,60]. From the list of immune/inflammatory complex traits that we selected, we identified genes differentially expressed in poverty that are also associated with risk for coronary artery disease, lupus, rheumatoid arthritis, psoriasis, asthma, multiple sclerosis, and Alzheimer's disease. Previous studies suggested that lack of plasticity between monocytes and macrophages contributes to the development of adverse health outcomes for these traits[61–65]. We analyzed gene expression within a mixed population of peripheral blood mononuclear cells and did not discern whether transcriptional profiles were associated with unique cell types. Follow-up analyses are warranted to experimentally validate these initial findings and identify if transcriptional profiles in specific cell types serve as better biological biomarkers for the effect of poverty status on gene expression.

Men and women differ in their risk for many complex diseases (e.g. [66,67]). These sex differences are attributed to hormones, sex chromosomes, genotype-sex effects, behavioral differences, and differential environmental exposures; however, the mechanisms and underlying biology need clarification. Oliva et al[46], studied differences in gene expression between men and women in the Genotype-Tissue Expression (GTEx) dataset and identified 13,294 genes differentially expressed in at least one tissue. In blood they found 1710 differentially expressed genes. The sex-specific differences in gene expression observed between individuals living in poverty or above poverty may reflect differences between sexes in average life expectancy of at-risk populations[68]. Compared with individuals who are not living in poverty, men living in poverty had fewer genes differentially expressed than women (Figure 2, Supplementary Table 3). Gene expression differences between men and women could reflect that living in poverty has a less profound effect on gene expression levels in PBMCs in men.

Poverty is consistently linked to social and environmental conditions that act as stressors and are associated with risk of chronic illness[69]. Earlier studies in social genomics identified a pattern of differentially expressed genes known as the conserved transcriptional response to adversity (CTRA)[16,17,19]. Stress response pathways are mediated by the sympathetic nervous system and beta-adrenergic signaling to induce the immune system's transcriptional activation of inflammatory genes[29,30]. Socioeconomic status and psychosocial experiences have a broader effect on immune cell gene expression and impact additional pathways beyond those included in the CTRA[28]. We identified 581 differentially expressed genes, of which only 1 was also included in the CTRA. Our study thus identifies relevant biological processes and pathways associated with poverty that were previously unknown. Potential stress response to poverty includes response to wound healing, platelet activation and aggregation, and vascular smooth muscle contraction.

Sex hormone levels are associated with health status and disease risk. Sex steroid hormone signaling offers protection in both men and women; however, the depletion of estrogen and testosterone following menopause elevates the risk for cardiovascular disease[41,70,71], highlighting the significance of sex steroid hormone signaling. In addition, women generally have a lower risk of cardiovascular disease than men of comparable age[72], but this protection is lost with menopause. In our study, biological processes that may be directly related with cardiovascular events (e.g. coagulation) are enriched for genes differentially expressed in women living in poverty. Future directions include examining the effects of sex (estrogen, testosterone, and progesterone) and stress hormones (cortisol, glucagon, and prolactin) to evaluate the potential effects of poverty as environmental stressors on immune cell gene expression patterns.

Here we focus on poverty strictly defined as household income, however several factors of living in poverty may be directly associated with gene expression differences, including for example food insecurity or nutritional factors, psychosocial stress or perceived violence, or any other complex environmental influences resulting from living in an impoverished urban area. Future studies aiming to dissect the relative effect of each of these factors will be important to guide interventions aimed at improving health outcomes for inner-city residents.

In conclusion, our study provides evidence that poverty influences gene expression patterns in a large diverse cohort. Impoverished environments have long been associated with the prevalence of poor health outcomes. This work helps elucidate the molecular pathways that potentially transduce the effects of poverty into poor health outcomes, thus providing an understanding of how social determinants of health alter gene expression and can provide additional information about pathways that contribute to disparities in health outcomes in disadvantaged populations.

Materials and Methods

Study Participants and sample collection

The Healthy Aging in Neighborhoods of Diversity Across the Life Span (HANDLS) study [31] is a longitudinal, epidemiological study of a diverse cohort of individuals living in Baltimore City and was designed and implemented by the National Institute on Aging Intramural Research Program (NIA IRP). Poverty status was assigned based on self-reported household income, by grouping participants as either 125% above or below the 2004 federal poverty line. Samples were collected at two different time points ~4-5 years apart for each participant (time point 1 and time point 2). Participant blood samples were collected in the morning. The isolation of the PBMCs was conducted within 3 hours of blood draw, aliquoted, and stored at -80˚C. The Institutional Review Boards of the National Institute of Environmental Health Sciences, National Institutes of Health and Morgan State University approved this study and a written informed consent was signed by all participants.

In the present study, HANDLS participants were selected across 4 variables: sex, self-identified race, poverty status, and age group (10-yr groups between 35-65). The cohort was divided in 24 batches balanced for all variables, for a total of 239 individuals across two time points, for a total of 478 samples. Following RNA-sequencing QC (see section below), the final sample size consisted of 204 participants (Figure 1B).

RNA Extraction, Library preparation and Sequencing

Cells were thawed and transferred into RPMI1640 (Gibco Cat# 11835030) medium containing 10% FBS (Gibco Cat# 10082147). Cells were then collected by centrifugation at 1200 RPM for 7 min and washed twice using ice-cold PBS (Gibco Cat# 10010002). Collected pellets were lysed using Lysis/Binding Buffer (Invitrogen Cat# 61012), and frozen at −80°C. Polyadenylated mRNAs were

subsequently isolated from thawed lysates using the Dynabeads mRNA Direct Kit (Invitrogen), following the manufacturer's instructions. RNA-seq libraries were prepared using a protocol modified from the NEBNext Ultradirectional (Illumina® - NEB Cat# E7760L) library preparation protocol. RNA was fragmented at 94°C for 5 minutes to obtain 200–1500 bp fragments in size. SPRISelect beads (Beckman Coulter) were used in all purification steps and size selection. Barcodes from PerkinElmer were added by ligation. The individual libraries were quantified using the KAPA real-time PCR system (ROCHE DIAGNOSTICS Cat # 07960441001), following the manufacturer's instructions and using a custom-made series of standards obtained from serial dilutions of the PhiX DNA (Illumina). Next, libraries were pooled and sequenced using the Novaseq instrument. Median number of reads/sample was 33,015,582 Supplementary File 1.

Data pre-processing

We aligned the RNA-seq reads with HISAT232 to the human genome version 'grch38_snp_tran', which considers splicing and common genetic variants. Aligned and cleaned (deduplicated) reads were counted using HTSeq and Homo_sapiens.GRCh38.103 transcriptome assembly genes. We started with an initial sample size of 478 libraries corresponding to 239 individuals. Due to poor RNA quality 26 libraries failed to sequence, which resulted in 452 libraries. For all gene expression analyses, genes on sex chromosomes and genes with expression below six reads or 0.1 counts per million in at least 20% of samples were dropped wich resulted in 21,251 genes.

To validate the identity of each sample from the same individual across the two time point measurements, we used the genotypes obtained from the RNA-seq reads. We obtained genotype information from the RNA-seq data by using samtools pileup and bcftools on known genotypes from the 1000 genomes project[33]. The resulting vcf was filtered for biallelic SNPs and the samples were further analyzed with plink2[34]. We then obtained the first 10 principal components of the genotypes. The difference in the 10 principal components for both time point 1 and 2 was calculated and if the resulting variance deviated more than 0.001 units, the sample pairs were removed. This resulted in a final set of 408 samples that were used in all subsequent analyses (Supplementary Figure 2). The first 3 genotype principal components were added in the gene expression model to control for genetic ancestry.

Additionally, the following covariates were included in the model to correct for confounders: library preparation batch (confounded with library sequencing batch), age, sex, percent reads mapping to exons, percent non-duplicate reads, and the first three principal components on the genotype matrix.

Differential gene expression analysis

We used DESeq2 v1.22.1[35] to test for differential gene expression across 204 individuals in two time points.

Model 1 (Poverty Model): library preparation batch + age + sex + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status

Model 2 (Poverty x Sex Interaction Model): library preparation batch + age + sex + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status + poverty status:sex

Model 3 (Women Poverty Model): library preparation batch + age + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status

Model 4 (Men Poverty Model): library preparation batch + age + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3+ poverty status

We also examined sex separately with models 3 and 4. To control for FDR, we used the Benjamin-Hochberg independent filtering step and multiple test correction implemented in DESeq2. Supplementary File 2 contains the full results of the analysis.
For each of these models, we also processed each time point separately, in which case we dropped the time variable. We then compared the study variable (e.g., poverty status in model 1) between time 1 and time 2 (e.g., Model 1-T1 and Model 1-T2) and correlated their effects to assess reproducibility (Supplementary Table 1).

Gene Ontology and pathway enrichment analyses

To identify relevant pathways related to the significantly differentially expressed genes in individuals living in poverty, we performed an enrichment analysis (hypergeometric test) for GO, KEGG and REACTOME terms with clusterProfiler in R[36]. We tested each list of differentially expressed genes against a background of all tested genes in our dataset. Significant biological processes contained a minimum of 10 genes and a maximum of 500 genes with a p adjusted < 0.1.

TWAS analyses

To directly investigate whether discovered effects on gene expression may contribute to differential health outcomes we used the results from PTWAS (Probabilistic Transcriptome-Wide Association Study)[37] (5% FDR) for Alzheimer's Disease (AD), Asthma (AS), Coronary Artery Disease (CAD), Crohn's Disease (CD), Eczema, Hypertension (HTN), Inflammatory bowel disease (IBD), Lupus, Multiple Sclerosis (MS), Psoriasis, and Rheumatoid Arthritis (RA) risk. We overlapped the significantly differentially expressed genes from models 1 and 2 with the PTWAS results from Additional file 5, Table S4 in[37].

Acknowledgments

We thank Samuele Zilioli and members of the Luca/Pique-Regi group for helpful discussions and comments on an early version of this manuscript. We would like to thank the HANDLS participants and clinical staff, and Nicolle Mode for helping with cohort design. This work was supported by the NIH grant #U54MD013376 and by the National Institute on Aging Intramural Research Program, National Institutes of Health project #AG000513.

Supplementary Information

Methods

DESEQ models used to identify differentially expressed genes at each time point separately:

Model 1-T1: Timepoint 1 Poverty Model: library preparation batch + age + sex + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status

Model 1- T2: Timepoint 2 Poverty Model: library preparation batch + age + sex + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status

Model 2-T1: Timepoint 1 Sex-Interaction Model: library preparation batch + age + sex + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 ++ poverty status + poverty status:sex

Model 2-T2: Timepoint 2 Sex-Interaction Model: library preparation batch + age + sex + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + + poverty status + poverty status:sex

Supplementary Table 1: Correlations between T1 and T2 for the different models. Models tested to find significantly differential expressed genes while controlling for specific confounding variables at a 10% false discovery rate.

Gene Expression Models

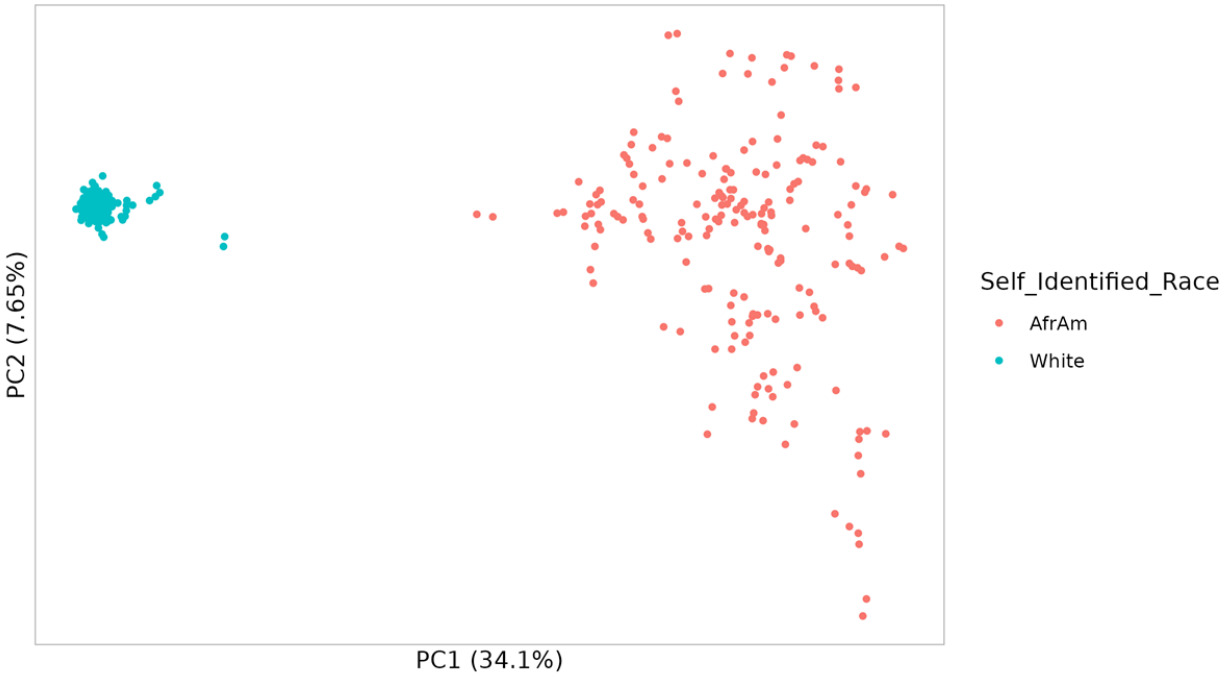
Model 1 (Poverty Model): library preparation batch + age + sex + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status

Model 2 (Poverty x Sex Interaction Model): library preparation batch + age + sex + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + + poverty status + poverty status:sex

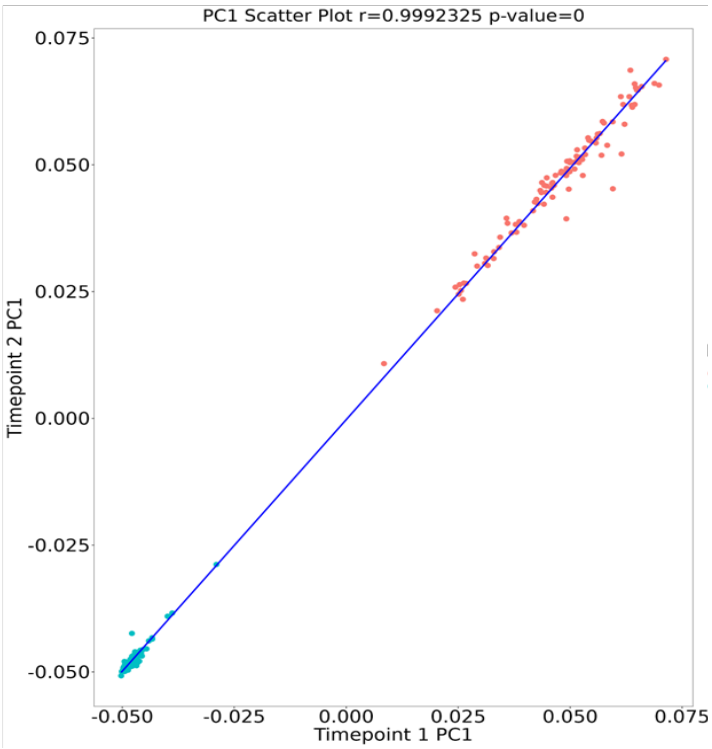
Model 3 (Women Poverty Model): library preparation batch + age + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3 + poverty status

Model 4 (Men Poverty Model): library preparation batch + age + time + percent reads mapping to exons + percent non-duplicate reads + PC1 + PC2 + PC3+ poverty status

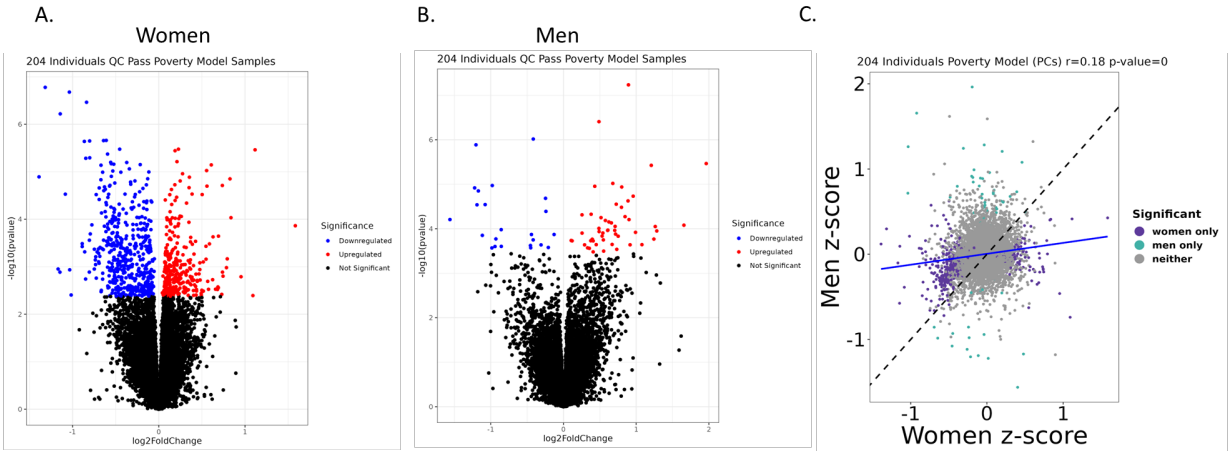
Model	Tested Variable	Confounding Variable(s)	Individuals	T1T2 DEGs (10% FDR)	Correlation	P-value
1	Poverty	Sex, PC1, PC2, PC3	204	581	0.4968831	0
2	Sex-Interaction	Sex, PC1, PC2, PC3	204	64	0.46599	0
3	Poverty	PC1, PC2, PC3	105	804	0.5567745	0
4	Poverty	PC1, PC2, PC3	99	74	0.3137857	0



Supplementary Figure 1: Principal Component Analysis of individual genotypes. Scatter plot of principal component 1 (x axis) and principal component 2 (y axis) calculated on the genotypes inferred from the RNAseq reads. Each dot represents an individual and their self-identified race (African American in orange and White in green).



Supplementary Figure 2: QC based on Principal Component Analysis of individual genotypes. Scatter plot of principal component 1 calculated on the genotypes inferred from the RNAseq reads at time point 1 (x-axis) and at time point 2 (y-axis). Each dot represents an individual and their self-reported census-based race category (African American in orange and White in green).



Supplementary Figure 3: Effects of poverty on gene expression in men and women. A. Volcano plot of the differentially expressed genes in the entire women cohort, while considering the time points as replicate measurement. The 480 downregulated genes are indicated in blue and the 324 upregulated genes are in red. B. Volcano plot of the differentially expressed genes in the entire men cohort, while considering the time points as replicate measurement. The 23 downregulated genes are indicated in blue and the 51 upregulated genes are in red. C. Scatter plot of the correlation between gene expression log fold changes estimated for men and women separately. Significantly differentially expressed genes (FDR<10%) are indicated with colors: women (purple), and men (green).

References

1. Antonovsky, A. Social class, life expectancy and overall mortality. *Milbank Mem. Fund Q.* **45**, 31–73 (1967).
2. Sorlie, P. D., Backlund, E. & Keller, J. B. US mortality by economic, demographic, and social characteristics: the National Longitudinal Mortality Study. *Am. J. Public Health* **85**, 949–956 (1995).
3. Kunst, A. E., Groenhof, F., Mackenbach, J. P. & Health, E. W. Occupational class and cause specific mortality in middle aged men in 11 European countries: comparison of population based studies. EU Working Group on Socioeconomic Inequalities in Health. *BMJ* **316**, 1636–1642 (1998).
4. Phelan, J. C., Link, B. G. & Tehranifar, P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J. Health Soc. Behav.* **51 Suppl**, S28–40 (2010).
5. Shahu, A. *et al.* Income disparity and utilization of cardiovascular preventive care services among U.S. adults. *Am J Prev Cardiol* **8**, 100286 (2021).
6. Odutayo, A. *et al.* Income Disparities in Absolute Cardiovascular Risk and Cardiovascular Risk Factors in the United States, 1999-2014. *JAMA Cardiol* **2**, 782–790 (2017).
7. Tajuddin, S. M., Nalls, M. A., Zonderman, A. B. & Evans, M. K. Association of red cell distribution width with all-cause and cardiovascular-specific mortality in African American and white adults: a prospective cohort study. *J. Transl. Med.* **15**, 208 (2017).
8. Waldstein, S. R. *et al.* Cross-sectional relations of race and poverty status to cardiovascular risk factors in the Healthy Aging in Neighborhoods of Diversity across the Lifespan (HANDLS) study. *BMC Public Health* **16**, 258 (2016).
9. Wendell, C. R., Waldstein, S. R., Evans, M. K. & Zonderman, A. B. Distributions of Subclinical Cardiovascular Disease in a Socioeconomically and Racially Diverse Sample.

Stroke **48**, 850–856 (2017).

10. Dowd, J. B. & Aiello, A. E. Socioeconomic differentials in immune response. *Epidemiology* **20**, 902–908 (2009).
11. Azad, M. B. *et al.* Influence of socioeconomic status trajectories on innate immune responsiveness in children. *PLoS One* **7**, e38669 (2012).
12. Miller, G. E. *et al.* Pathways to resilience: maternal nurturance as a buffer against the effects of childhood poverty on metabolic syndrome at midlife. *Psychol. Sci.* **22**, 1591–1599 (2011).
13. Miller, G. E. *et al.* Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14716–14721 (2009).
14. Li, L., Power, C., Kelly, S., Kirschbaum, C. & Hertzman, C. Life-time socio-economic position and cortisol patterns in mid-life. *Psychoneuroendocrinology* **32**, 824–833 (2007).
15. Björntorp, P. & Rosmond, R. Hypothalamic origin of the metabolic syndrome X. *Ann. N. Y. Acad. Sci.* **892**, 297–307 (1999).
16. Slavich, G. M. & Cole, S. W. The Emerging Field of Human Social Genomics. *Clin. Psychol. Sci.* **1**, 331–348 (2013).
17. Cole, S. W. Social regulation of human gene expression. *Curr. Dir. Psychol. Sci.* **18**, 132–137 (2009).
18. Toyokawa, S., Uddin, M., Koenen, K. C. & Galea, S. How does the social environment ‘get into the mind’? Epigenetics at the intersection of social and psychiatric epidemiology. *Soc. Sci. Med.* **74**, 67–74 (2012).
19. Cole, S. W. Human social genomics. *PLoS Genet.* **10**, e1004601 (2014).
20. Galea, S., Uddin, M. & Koenen, K. The urban environment and mental disorders: Epigenetic links. *Epigenetics* **6**, 400–404 (2011).
21. Chen, E. *et al.* Genome-wide transcriptional profiling linked to social class in asthma.

Thorax **64**, 38–43 (2009).

22. Arnold, N. S. *et al.* The association between poverty and gene expression within peripheral blood mononuclear cells in a diverse Baltimore City cohort. *PLoS One* **15**, e0239654 (2020).
23. Robles, T. F. *et al.* Family environments and leukocyte transcriptome indicators of a proinflammatory phenotype in children and parents. *Dev. Psychopathol.* **30**, 235–253 (2018).
24. Stanton, S. C. E. *et al.* Mothers' Attachment is Linked to their Children's Anti-Inflammatory Gene Expression via Maternal Warmth. *Soc. Psychol. Personal. Sci.* **8**, 796–805 (2017).
25. Powell, N. D. *et al.* Social stress up-regulates inflammatory gene expression in the leukocyte transcriptome via β-adrenergic induction of myelopoiesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16574–16579 (2013).
26. Farrell, A. K. *et al.* Socioeconomic status, family negative emotional climate, and anti-inflammatory gene expression among youth with asthma. *Psychoneuroendocrinology* **91**, 62–67 (2018).
27. Segman, R. H. *et al.* Blood mononuclear cell gene expression signature of postpartum depression. *Mol. Psychiatry* **15**, 93–100, 2 (2010).
28. Resztak, J. A. *et al.* Psychosocial experiences modulate asthma-associated genes through gene-environment interactions. *Elife* **10**, (2021).
29. Thames, A. D., Irwin, M. R., Breen, E. C. & Cole, S. W. Experienced discrimination and racial differences in leukocyte gene expression. *Psychoneuroendocrinology* **106**, 277–283 (2019).
30. Cole, S. W. The Conserved Transcriptional Response to Adversity. *Curr Opin Behav Sci* **28**, 31–37 (2019).
31. Evans, M. K. *et al.* Healthy aging in neighborhoods of diversity across the life span (HANDLS): overcoming barriers to implementing a longitudinal, epidemiologic, urban study

of health, race, and socioeconomic status. *Ethn. Dis.* **20**, 267–275 (2010).

32. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
33. 1000 Genomes Project Consortium *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
34. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
35. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
36. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
37. Zhang, Y. *et al.* PTWAS: investigating tissue-relevant causal molecular mechanisms of complex traits using probabilistic TWAS analysis. *Genome Biol.* **21**, 232 (2020).
38. Haeseleer, F. *et al.* Five members of a novel Ca(2+)-binding protein (CABP) subfamily with similarity to calmodulin. *J. Biol. Chem.* **275**, 1247–1260 (2000).
39. McCue, H. V., Burgoyne, R. D. & Haynes, L. P. Membrane targeting of the EF-hand containing calcium-sensing proteins CaBP7 and CaBP8. *Biochem. Biophys. Res. Commun.* **380**, 825–831 (2009).
40. Ren, L., Cao, X., Geng, P., Bai, F. & Bai, G. Study of the inhibition of two human maltase-glucoamylases catalytic domains by different α-glucosidase inhibitors. *Carbohydr. Res.* **346**, 2688–2692 (2011).
41. Boese, A. C., Kim, S. C., Yin, K.-J., Lee, J.-P. & Hamblin, M. H. Sex differences in vascular physiology and pathophysiology: estrogen and androgen signaling in health and disease. *Am. J. Physiol. Heart Circ. Physiol.* **313**, H524–H545 (2017).
42. Ahmed, S. & Spence, J. D. Sex differences in the intestinal microbiome: interactions with risk factors for atherosclerosis and cardiovascular disease. *Biol. Sex Differ.* **12**, 35 (2021).

43. Spence, J. D. & Pilote, L. Importance of sex and gender in atherosclerosis and cardiovascular disease. *Atherosclerosis* **241**, 208–210 (2015).

44. Rainville, J. R., Tsyglakova, M. & Hodes, G. E. Deciphering sex differences in the immune system and depression. *Front. Neuroendocrinol.* **50**, 67–90 (2018).

45. Dolsen, E. A., Crosswell, A. D. & Prather, A. A. Links Between Stress, Sleep, and Inflammation: Are there Sex Differences? *Curr. Psychiatry Rep.* **21**, 8 (2019).

46. Oliva, M. *et al.* The impact of sex on gene expression across human tissues. *Science* **369**, (2020).

47. Feitosa, M. F. *et al.* Gene discovery for high-density lipoprotein cholesterol level change over time in prospective family studies. *Atherosclerosis* **297**, 102–110 (2020).

48. Barrow, T. M. *et al.* Analysis of retrotransposon subfamily DNA methylation reveals novel early epigenetic changes in chronic lymphocytic leukemia. *Haematologica* **106**, 98–110 (2021).

49. Gu, X. *et al.* Tel-2 is a novel transcriptional repressor related to the Ets factor Tel/ETV-6. *J. Biol. Chem.* **276**, 9421–9436 (2001).

50. Gaudet, P., Livstone, M. S., Lewis, S. E. & Thomas, P. D. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief. Bioinform.* **12**, 449–462 (2011).

51. Potter, M. D., Buijs, A., Kreider, B., van Rompaey, L. & Grosveld, G. C. Identification and characterization of a new human ETS-family transcription factor, TEL2, that is expressed in hematopoietic tissues and can associate with TEL1/ETV6. *Blood* **95**, 3341–3348 (2000).

52. Findley, A. S. *et al.* Interpreting Coronary Artery Disease Risk Through Gene-Environment Interactions in Gene Regulation. *Genetics* **213**, 651–663 (2019).

53. Kirschbaum, T. K. *et al.* The Association of Socioeconomic Status With Hypertension in 76 Low- and Middle-Income Countries. *J. Am. Coll. Cardiol.* **80**, 804–817 (2022).

54. Boniakowski, A. E., Kimball, A. S., Jacobs, B. N., Kunkel, S. L. & Gallagher, K. A.

Macrophage-Mediated Inflammation in Normal and Diabetic Wound Healing. *J. Immunol.* **199**, 17–24 (2017).

55. Kimball, A. *et al.* Ly6CHi Blood Monocyte/Macrophage Drive Chronic Inflammation and Impair Wound Healing in Diabetes Mellitus. *Arterioscler. Thromb. Vasc. Biol.* **38**, 1102–1114 (2018).
56. Resztak, J. A. *et al.* Genetic control of the dynamic transcriptional response to immune stimuli and glucocorticoids at single cell resolution. *bioRxiv* (2022) doi:10.1101/2021.09.30.462672.
57. Randolph, H. E. *et al.* Genetic ancestry effects on the response to viral infection are pervasive but cell type specific. *Science* **374**, 1127–1133 (2021).
58. Jordan, A. R., Racine, R. R., Hennig, M. J. P. & Lokeshwar, V. B. The Role of CD44 in Disease Pathophysiology and Targeted Treatment. *Front. Immunol.* **6**, 182 (2015).
59. van der Laan, A. M. *et al.* A proinflammatory monocyte response is associated with myocardial injury and impaired functional outcome in patients with ST-segment elevation myocardial infarction: monocytes and myocardial infarction. *Am. Heart J.* **163**, 57–65.e2 (2012).
60. Peet, C., Ivetic, A., Bromage, D. I. & Shah, A. M. Cardiac monocytes and macrophages after myocardial infarction. *Cardiovasc. Res.* **116**, 1101–1112 (2020).
61. Kapellos, T. S. *et al.* Human Monocyte Subsets and Phenotypes in Major Chronic Inflammatory Diseases. *Front. Immunol.* **10**, 2035 (2019).
62. Nally, F. K., De Santi, C. & McCoy, C. E. Nanomodulation of Macrophages in Multiple Sclerosis. *Cells* **8**, (2019).
63. Cheng, J. B. *et al.* Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution. *Cell Rep.* **25**, 871–883 (2018).
64. Ross, E. A., Devitt, A. & Johnson, J. R. Macrophages: The Good, the Bad, and the Gluttony. *Front. Immunol.* **12**, 708186 (2021).

65. Steinbach, F. *et al.* Monocytes from systemic lupus erythematous patients are severely altered in phenotype and lineage flexibility. *Ann. Rheum. Dis.* **59**, 283–288 (2000).

66. Arslan, M. *et al.* Coronary CT angiography for suspected acute coronary syndrome: sex-associated differences. *Neth. Heart J.* **29**, 518–524 (2021).

67. Kim, J.-W., Kim, H.-A., Suh, C.-H. & Jung, J.-Y. Sex hormones affect the pathogenesis and clinical characteristics of systemic lupus erythematosus. *Front. Med.* **9**, 906475 (2022).

68. Chetty, R. *et al.* The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA* **315**, 1750–1766 (2016).

69. Gaye, A., Gibbons, G. H., Barry, C., Quarells, R. & Davis, S. K. Influence of socioeconomic status on the whole blood transcriptome in African Americans. *PLoS One* **12**, e0187290 (2017).

70. Pérez-López, F. R., Chedraui, P., Gilbert, J. J. & Pérez-Roncero, G. Cardiovascular risk in menopausal women and prevalent related co-morbid conditions: facing the post-Women's Health Initiative era. *Fertil. Steril.* **92**, 1171–1186 (2009).

71. Nilsson, P. M., Viigimaa, M., Giwercman, A. & Cifkova, R. Hypertension and Reproduction. *Curr. Hypertens. Rep.* **22**, 29 (2020).

72. Writing Group Members *et al.* Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. *Circulation* **133**, e38–360 (2016).