


















A multi-omics approach to investigate the inflammatory response to life course socioeconomic position

Raphaële Castagné^{*,1} , Michelle Kelly-Irving¹ , Vittorio Krogh² , Domenico Palli³ , Salvatore Panico⁴, Carlotta Sacerdote⁵ , Rosario Tumino⁶ , Dennie GAJ Hebels⁷ , Jos CS Kleinjans⁸, Theo MCM de Kok⁸ , Panagiotis Georgiadis⁹ , Soterios A Kyrtopoulos⁹ , Roel Vermeulen¹⁰ , Silvia Stringhini¹¹ , Paolo Vineis^{12,13} , Marc Chadeau-Hyam^{‡,12}  & Cyrille Delpierre^{‡,1} 

¹LEASP, UMR 1027, Inserm-Université Toulouse III Paul Sabatier, Toulouse, France

²Epidemiology & Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan 20133, Italy

³Cancer Risk Factors & Lifestyle Epidemiology Unit, Institute for Cancer Research Prevention & Clinical Network-ISPRO, Florence 50141, Italy

⁴Department of Clinical Medicine & Surgery, University of Naples Federico II, Naples 80131, Italy

⁵Unit of Cancer Epidemiology, Città della Salute e della Scienza University-Hospital & Center for Cancer Prevention (CPO), Turin 10133, Italy

⁶Cancer Registry & Department of Histopathology, Provincial Health Authority (ASP) Ragusa 97100, Italy

⁷MERLN Institute, Maastricht University, Universiteitssingel 40, 6229ER Maastricht, The Netherlands

⁸Department of Toxicogenomics, GROW Institute & Developmental Biology, Maastricht University, Maastricht 6211LK, The Netherlands

⁹National Hellenic Research Foundation, Institute of Chemical Biology, Vas. Constantinou 48, 11635 Athens, Greece

¹⁰Division of Environmental Epidemiology, Institute for Risk Assessment Sciences (IRAS), Utrecht University, PO Box 80178, 3508 TD, Utrecht, The Netherlands

¹¹Institute of Social & Preventive Medicine, Lausanne University Hospital, Lausanne 1010, Switzerland

¹²MRC-PHE Centre for Environment & Health, School of Public Health, Department of Epidemiology & Biostatistics, Imperial College London, SW7 2BU, London, UK

¹³Molecular & Genetic Epidemiology Unit, Italian Institute for Genomic Medicine (IIGM), Torino 10126, Italy

*Author for correspondence: raphaele.castagne@inserm.fr

‡Authors contributed equally

Aim: Inflammation represents a potential pathway through which socioeconomic position (SEP) is biologically embedded. **Materials & methods:** We analyzed inflammatory biomarkers in response to life course SEP by integrating multi-omics DNA-methylation, gene expression and protein level in 178 European Prospective Investigation into Cancer and Nutrition-Italy participants. **Results & conclusion:** We identified 61 potential *cis* acting CpG loci whose methylation levels were associated with gene expression at a Bonferroni correction. We examined the relationships between life course SEP and these 61 *cis*-acting regulatory methylation sites individually and jointly using several scores. Less-advantaged SEP participants exhibit, later in life, a lower inflammatory methylome score, suggesting an overall increased expression of the corresponding inflammatory genes or proteins, supporting the hypothesis that SEP impacts adult physiology through inflammation.

First draft submitted: 9 September 2019; Accepted for publication: 29 June 2020; Published online: 2 September 2020

Keywords: DNA methylation • gene expression • inflammation • life course epidemiology • protein • socioeconomic position

Socioeconomic position (SEP) is one of the most consistent determinants of health [1]. Epidemiological studies worldwide have demonstrated strong, graded and persistent associations between indicators of SEP across the life course and a large range of health outcomes including cardiovascular and respiratory diseases, hypertension, diabetes and cancer [2]. Although it is now established that disadvantaged SEP is associated with poorer health [3],

the relationship between SEP and health is not fully explained by established behavioral and lifestyle risk factors for chronic diseases [4,5]. Persistent socioeconomic disadvantage and psychosocial adversities experienced over the life course can induce a chronic overactivation and/or dysregulation of the physiological stress response [6]. The latter has been shown to alter many biological mechanisms, in particular inflammatory and immune responses [7]. In addition, there is increasing evidence that the inflammatory response is socially patterned, through a wide range of possible mediators at different molecular levels. A better understanding of the pathways and mechanisms through which social disadvantage affects health is key to reduce the harmful effect of SEP and ultimately address socioeconomic inequalities in health.

The association between disadvantaged life course SEP and elevated inflammatory proteins in adulthood – mainly including CRP and IL-6 – is supported by several recent meta-analyses [8–10]. Some studies indicate that different types of adverse social conditions such as low SEP [11–14], social isolation [15,16] and early-life social deprivation [17] are related to the upregulation of the expression of several genes involved in the inflammatory pathways [18], namely the conserved response to adversity [19]. Epigenetics, as a potential regulator of gene expression without altering the genome, has been postulated to link adverse social exposures to health through the development of a proinflammatory state [20]. Previous studies have demonstrated a link between global DNA methylation [21–23], at individual genome-wide CpG site [24–27] or at site-specific DNA methylation patterns related to inflammation [28–30] and socially patterned stressors including disadvantaged childhood SEP [24,25,28–31], adulthood SEP [28,29], perceived stress [25] and neighborhood crime [32]. These studies support the hypothesis that disadvantaged SEP may influence biological functioning, in particular the inflammation response, at different omics levels. However, studies considering the different omics levels in an integrative analysis are still sparse.

In previous studies in 268 participants from the Italian component of the European Prospective Investigation into Cancer and Nutrition (EPIC-Italy), we examined how SEP over the life course is likely to be embedded through the inflammatory system using: circulating levels of 28 inflammation-related proteins [33] and 845 inflammatory genes [34] considered separately, or combined into an transcriptome inflammatory score. Our results suggest that early-life SEP is associated with elevated levels of both inflammatory-related proteins and transcripts in adulthood. In the present study, we adopted a hypothesis-driven investigation focusing on inflammation measured at three different omic levels: proteins, transcripts and DNA methylation. We hypothesized that participants with a disadvantaged SEP will have lower inflammatory DNA methylation level compared with their advantaged counterparts that in turn will upregulate expression and translation of these genes. Our approach first relied on identifying CpG sites associated with inflammatory-related transcripts and proteins, and subsequently investigate their relationships with three SEP indicators at different life stages. We examined these associations for each CpG site separately and through a composite inflammatory methylome score. Finally, we investigated life course effects of early-life SEP experiences by sequentially controlling for time-ordered SEP. We further explored the relationship between each molecular stratum at the score level by life course SEP assuming that a low inflammatory methylome score will result in a higher level of inflammatory genes and proteins.

Materials & methods

Study population

The EnviroGenoMarkers (EGM) project was initially designed as two nested case–control studies with the objective to identify novel biomarkers of non-Hodgkin's lymphoma and breast cancer risk from multiple '-omics' profiles [35]. EGM includes participants from the Italian component of the EPIC-Italy [36] and the Northern Sweden Health and Disease Study [37]. In this study, 268 EGM participants from EPIC-Italy healthy at enrollment were included. For these participants, anthropometric, lifestyle, dietary and socioeconomic factors as well as blood samples were collected at recruitment. All participants provided informed consent, and the EPIC study protocol was approved by the review board of the International Agency for Research on Cancer and by all local institutes recruiting participants. The study was conducted in accordance with the approved ethical guidelines. Biosamples underwent genome wide expression and methylation as well as targeted proteomic inflammatory profiling in two distinct batches.

Life course SEP

As previously described, we used dichotomized indicators of each life course SEP factors from the EPIC questionnaire to preserve both power and interpretability [33,34].

Briefly 'childhood SEP' was approached by father's occupation and recorded in the two following categories: 'Manual' grouping unskilled workers, skilled workers and farmers; and 'Non-manual' grouping retailers, employees and self-employed. Young 'adulthood SEP' was measured through participant's own education, which was dichotomized as 'High' (above the minimum legal education level, 15 years of age) and 'Low' (below the minimum legal education level). For 'Adulthood SEP', we used the highest occupational position in the household defined by either the participant's own occupation or his/her partner's and classified as 'Manual' and 'Non-manual,' using the same categorization described above for father's occupation.

Laboratory analyses

Genome-wide methylation profiling experiments were conducted as previously described [38], using the Illumina Infinium HumanMethylation450 platform according to manufacturers' protocols. Briefly, the QIAsymphony DNA Midi Kit (Qiagen, Crawley, UK) was used to extract DNA samples from buffy coats. The DNA (500 ng) of each sample was bisulfite-converted using the EZ-96 DNA Methylation-Gold™ Kit (Zymo Research, CA, USA) and hybridized on the Infinium HumanMethylation 450 BeadChip. The beadchips were then scanned using the Illumina HiScan SQ scanner and intensities of the images were extracted using the GenomeStudio (v.2011.1) Methylation module (1.9.0) software. Data preprocessing was carried out using in-house software written for the R statistical computing environment resulting in a subset of 485,512 probes targeting autosomal CpG loci. Probes with SNPs in their sequence and/or detected in less than 10% of the samples were subsequently excluded from the analyses, leaving us with 367,648 probes in 188 samples. For each CpG, the methylation level at each locus was expressed as a β -value according to the fluorescent intensity ratio representing any value between 0 (unmethylated) and 1 (completely methylated).

Genome-wide expression profiling experiments were carried out using a previously published protocol [35]. Briefly, the Agilent 4×44 K human whole genome microarray platform was used to obtain gene expression profiles. After assessing both technical performance and quality of the microarrays according to a protocol described [35,39], a total of 29,662 transcripts were successfully analyzed in 246 samples.

Proteomic assay

We used the MILLIPLEX HCYTOMAG-60K and HSCYTMAG-60SK kits (Millipore, MA, USA) following the manufacturers' protocols to measure a panel of inflammatory-related proteins ($N = 32$) including 10 chemokines, 12 cytokines and 6 growth factors. Due to a high rate of nondetects ($>75\%$), we excluded four analytes (IL-12, IL1-RA, sIL2-RA and Flt3ligand) from further statistical analyses leaving 28 proteins in 268 samples.

Inflammatory methylome definition

Transcriptomic-based definition of the inflammatory methylome

The 1027 inflammation-related genes list assembled by Loza *et al.* was used to define our inflammatory transcriptome and of these, 845 genes were assayed in our population [34]. Using the manufacturer supplied annotation data, we looked up for the 845 inflammatory genes and mapped 11,502 CpG sites corresponding to 824 genes. The methylation-expression analyses were conducted in the subset of 173 participants with both DNA methylation and gene. Out of the 11,502 selected CpG sites, 4389 had on average seven missing values and were consequently imputed by the mean. We then 'de-noised' DNA methylation data, by subtracting potential technical confounding effects from the methylation levels. As previously described [39], we ran a linear mixed model (including chip ID and position on the chip – as random intercepts) and subtracting the random effect estimates from the observed methylation levels. Expression levels of each of the 824 inflammatory genes were then regressed against methylation levels in the subset of CpG sites present in the given gene. In these analyses, linear mixed models were used between mRNA expression (dependent variable) and DNA methylation (independent variable) controlling for age, sex and phase (fixed effect) and technical covariates for transcriptomic profiling (i.e., isolation, hybridization and labeling dates) [39]. We declared CpG–transcript pairs as significant based on Bonferroni 5% significance level (per-test significance level $\alpha' = 0.05/[11,502 \text{ tests}]$), and the unique list of CpG sites involved in these associations defines the *cis*-acting CpG site. We identified 61 *cis*-acting regulatory CpG sites constitutive of the inflammatory methylome detailed in Supplementary Table 1.

Inflammatory methylome score

The inflammatory methylome score is defined as the average methylation level from the ‘de-noised’ DNA methylation data, across all the 61 *cis*-acting CpG sites (*cis*-methylation sites associated with gene expression). The higher the methylome score, the higher the level of methylation.

Alternative inflammatory methylome scores

In order to make the DNA methylation level distribution comparable between samples for each CpG sites, normalized *z*-scores were computed for each of the 61 *cis*-acting CpG sites. The *z*-scores from all CpG sites were subsequently summed up to create a combined inflammatory methylome *z*-score. As a hypothesis-free alternative, we used the first principal component (PC) from a principal component analyses (PCA) based on denoised DNA methylation data of all the *cis*-acting CpG sites. Each alternative of the inflammatory methylome score was defined from the from the ‘de-noised’ DNA methylation.

Proteins-based definition of the inflammatory methylome

Using the same approach, we sought for CpG sites whose methylation levels were linked the circulating levels of any of the 28 assayed inflammatory proteins. Using the manufacturer supplied annotation data, we looked up for the 28 proteins, and mapped 205 CpG sites corresponding to 27 proteins (seven CpG sites per protein on average). The methylation–protein analyses were conducted in the subset of 173 participants with both DNA methylation and protein concentration available from the same blood samples. Protein expression of each of the 27 inflammatory proteins were then regressed against methylation levels in the subset of the corresponding 205 CpG sites. In these analyses, linear mixed models were used between protein concentration (dependent variable) and DNA methylation (independent variable) controlling for age, sex and phase (fixed effect) and technical covariates for proteomic profiling including the identifier of the plate (random effect). As an exploratory approach among the 205 CpG–protein pairs tested, we declared a CpG–protein pair as significant based on 5% significance level. The five potential CpG sites associated with protein concentration are described in [Supplementary Table 2](#). The ‘de-noised’ DNA methylation data for these five CpGs sites were used to derive the same three scores detailed above (mean derived score, *z*-score and principal components 1).

Statistical analyses

Single CpG site analyses

We modeled the relationship between each *cis*-acting CpG sites (dependent variable) and the binary SEP indicator (the independent variable of interest) using linear mixed models. We included chip ID and position on the chip as random effects to account for technical confounding. Age, gender, center (recorded in three categories: North, Central and South Italy), phase (corresponding to the two experimental batches) were included as fixed effects. We also included two binary variables indicating whether a participant is a prospective breast cancer or lymphoma case to account for the case–control design of EGM (Model 1). The 61 *cis*-acting CpG sites exhibit a strong pairwise correlation ([Supplementary Figure 1](#)). Results from a PCA of these 61 inflammatory markers revealed that 33 PCs explained more than 95% of the total variation seen in the dataset ([Supplementary Figure 2](#)). To account for the correlation in the data, the multiple testing corrected significance level was defined as $p = 0.05/33$ ($p = 0.0015$).

Inflammatory methylome score analyses

Because the inflammatory methylome scores were based on the ‘de-noised’ DNA methylation data, these were analyzed using linear models. Starting from the fixed effects included in Model 1 and detailed above, we considered smoking status (categorical: current, former, never); alcohol consumption (continuous, g/day); physical activity (categorical: inactive, moderately inactive, moderately active, active) and BMI (continuous variable, kg/m²) as potential mediators. Based on Model 1, we controlled for each behavioral factor and BMI separately. Model 2 included the behavioral factors listed above and BMI.

Life course analyses of the inflammatory methylome scores

We developed four time-sequenced linear models adjusted for the chronologically ordered of life course SEP resulting in:

Model A: Age, gender, case–control status, phase, center and father’s occupation;

Model B-1: Model A + education;

Model B-2: Model A + highest household occupational position;

Model C: Model B-1 + highest household occupational position;

Model D was subsequently built upon Model C to control for potential mediators including the three behavioral factors described above and BMI.

Sensitivity analyses

Model 1, 2 and life course models were also run for each alternative definition of the score (z-score and PC1) and the two definitions of the inflammatory methylome (transcriptomic or proteomic based).

Cell type estimation

For each type of analyse, we further adjusted the regression models for blood cell composition, estimated according to the methods proposed by Houseman using the R package EpiDISH. In brief, the EpiDISH package uses blood cell-specific DNA methylation reference profiles of known cell-type-specific DNase hypersensitive sites to provide cell-count estimates by robust partial correlations. We obtained relative proportions of B-cells, TCD4⁺ cells, TCD8⁺ cells, natural killer (NK) cells, monocytes, neutrophils and eosinophils subsequently combined to derive the proportion of lymphocytes, granulocytes and monocytes. We found that the EPIC samples included in this study (N = 173) contained an average of 29.7% lymphocytes, 62.2% granulocytes and 8.1% of monocytes.

Targeted integration of corresponding gene expression & protein levels data

For both transcriptomic- and proteins-based definition of the inflammatory methylome, we defined the corresponding score at the transcriptomic and proteomic level. Briefly, for each of the 61 genes, we defined a dichotomized indicator: 'high gene expression level' = 1, and 'low gene expression level' = 0 based on the highest quartile of the gene expression level, and summed these across the 61 genes. For each of the five proteins, the score was calculated following the same procedure based on the highest quartile of the log-transformed concentrations. Spearman correlation was then used to assess the relationship between the inflammatory methylome and transcriptome scores and the inflammatory methylome and proteome scores.

Statistical analyses were performed using R v3.5.0.

Results

Study population

Characteristics of the 173 participants with both gene expression and DNA methylation are summarized in Table 1. For each SEP indicator, participants in the disadvantaged group tend to be older, to have a higher BMI and to be nonsmokers compared with their advantaged counterparts.

Description of the inflammatory methylome

We identified 61 out of 11,502 potential *cis*-acting regulatory methylation sites (expression Methylation Site, eMS) as CpG sites for which methylation was associated with expression of the CpG site's closest gene at a Bonferroni threshold ($p < 4.3 \times 10^{-6}$; Supplementary Table 1). Most of CpG sites were from the leukocyte (23%) and cytokine (14.8%) signaling subpathways (Figure 1A). The vast majority (78.7%) of the CpG sites were inversely associated with transcripts levels in *cis* (Figure 1B). Overall, the majority of probes (>50%) were situated in gene bodies, followed by approximately 15% of probes situated within 1500 bp upstream of TSS (Figure 1C). We identified the majority of *cis*-eMS (62.3%) were in open sea, followed by shelves (16.4%), shores (14.8%) and islands (6.6%) (Figure 1D). A similar pattern was observed between CpG sites whose methylation was negatively associated with expression (Supplementary Figure 3), while CpG sites whose methylation was positively associated with expression were mostly located in gene body (Supplementary Figure 4). Few CpG sites were no longer associated with gene expression after controlling for cell type composition (N = 9; $p \geq 0.05$; Supplementary Table 1).

Transcriptomic-based definition of the inflammatory methylome profiles at a single-base resolution

We first examined associations between the 61 CpG sites individually and each SEP along the life course separately using as reference the most advantaged socioeconomic group. Therefore, a negative association indicates a hypomethylation in less advantaged socioeconomic groups. Although no association remains significant after correction for multiple testing with our three SEP indicators, DNA methylation level exhibited general negative

Table 1. Summary characteristics of the study population by life course socioeconomic position.

Characteristics	Father's occupational position (N = 158)			Participant's education (N = 172)			Household's highest occupational position (N = 160)			Participants with complete data (N = 155)
	Nonmanual	Manual	p-value [†]	High	Low	p-value [†]	Nonmanual	Manual	p-value [†]	
N	54 (34.2)	104 (65.8)		77 (44.8)	95 (55.2)		102 (63.7)	58 (36.2)		
Age, years, mean (SD)	51.9 (8.1)	53.7 (7.8)	2.04 × 10 ⁻¹	50.5 (7.3)	54.8 (7.8)	2.34 × 10 ⁻⁴	52.2 (7.9)	54.9 (7.8)	3.93 × 10 ⁻²	53.2 (7.9)
Gender										
– Women	45 (83.3)	75 (72.1)	1.71 × 10 ⁻¹	59 (76.6)	75 (78.9)	8.57 × 10 ⁻¹	76 (74.5)	45 (22.4)	8.07 × 10 ⁻¹	118 (76.1)
– Men	9 (16.7)	29 (27.9)		18 (23.4)	20 (21.1)		26 (25.5)	13 (22.4)		37 (23.9)
Cancer case	26 (48.1)	39 (37.5)	2.63 × 10 ⁻¹	28 (36.4)	41 (43.2)	4.55 × 10 ⁻¹	48 (47.1)	17 (29.3)	4.24 × 10 ⁻²	63 (40.6)
Center [‡]										
– South	1 (1.9)	11 (10.6)	3.21 × 10 ⁻²	12 (15.6)	9 (9.5)	2.73 × 10 ⁻²	7 (6.9)	5 (8.6)	0.3571444	11 (7.1)
– Central	34 (63.0)	46 (44.2)		43 (55.8)	40 (42.1)		56 (54.9)	25 (43.1)		79 (51.0)
– North	19 (35.2)	47 (45.2)		22 (28.6)	46 (48.4)		39 (38.2)	28 (48.3)		65 (41.9)
BMI	25.3 (3.6)	25.8 (3.5)	4.33 × 10 ⁻¹	25.1 (3.3)	26.2 (3.6)	3.36 × 10 ⁻²	25.4 (3.4)	26.2 (3.7)	1.61 × 10 ⁻¹	25.7 (3.5)
– Missing	–	2		–	1		–	2		–
Smoking status										
– Never	23 (42.6)	56 (53.8)	3.42 × 10 ⁻¹	28 (36.4)	56 (58.9)	1.20 × 10 ⁻²	45 (44.1)	34 (59.6)	1.70 × 10 ⁻¹	78 (50.3)
– Former	15 (27.8)	25 (24.0)		26 (33.8)	19 (20.0)		29 (28.4)	12 (21.1)		40 (25.8)
– Current	16 (29.6)	22 (21.2)		23 (29.9)	20 (21.1)		28 (27.5)	11 (19.3)		37 (23.9)
– Missing	0	1 (1)		–	–		–	1		–
Grams alcohol/day	13.7 (19.3)	9.2 (14)	1.30 × 10 ⁻¹	12 (16.4)	9 (15.1)	2.20 × 10 ⁻¹	10.8 (15.3)	11.5 (17.7)	8.12 × 10 ⁻¹	10.9 (16.2)
– Missing	–	1		–	–		–	1		–

Counts and percentages are reported for categorical variable, and means and SDs for continuous variables.

[†] p-value for difference was calculated using the Chi-squared test for categorical variables and the student's t-test for continuous variables.

[‡] North: Turin & Varese; Central: Florence; South: Naples & Ragusa.

SD: Standard deviation.

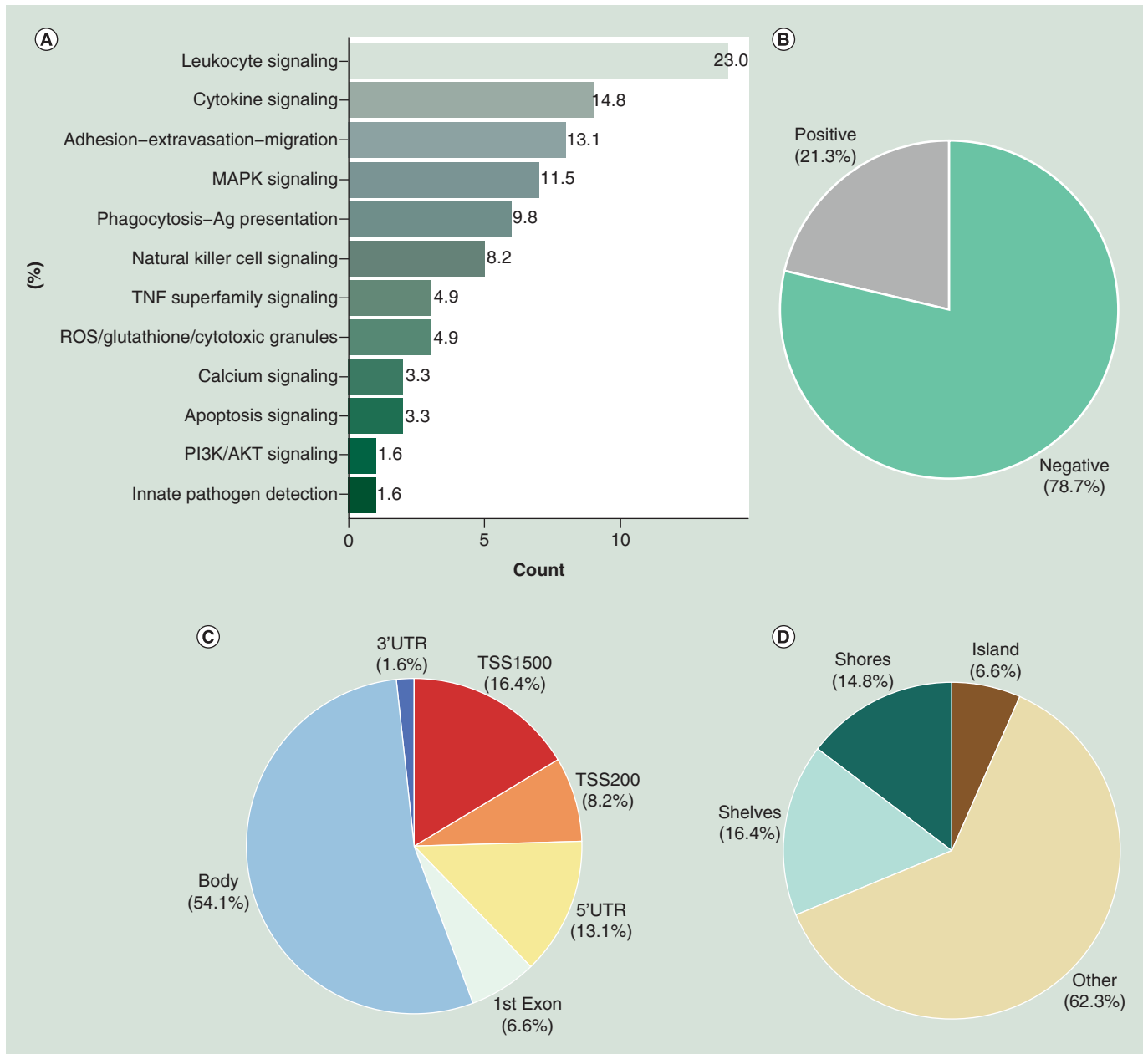


Figure 1. Description of the inflammatory methylome. Frequency of the 61 *cis*-eMS defining the inflammatory methylome by (A) inflammatory subpathway; (B) correlation between DNA methylation and gene expression; (C) physical distribution and (D) CpG island and neighborhood context. eMS: Expression methylation site.

associations with father’s occupational position, educational level and the highest household occupational position, respectively, 57, 82 and 80% of the regression coefficient were negative (Supplementary Table 3 & Figures 5–7).

Transcriptomic-based definition of the inflammatory methylome at a score resolution

As detailed in the methods section, DNA methylation score was calculated for the inflammatory methylome by taking the arithmetic mean of the methylation levels from the 61 *cis*-eMS. Both educational level and the highest household occupational position were found significantly associated with the inflammatory methylome score: participants from disadvantaged socioeconomic groups had a lower inflammatory methylome score ($\beta = -0.0075$; $p = 0.0067$; $\beta = -0.0076$; $p = 0.0073$ for educational level and the highest household occupational position,

Table 2. Linear regression results for the inflammatory methylome and each of the three life course socioeconomic position factors in the European Prospective Investigation into Cancer and Nutrition-Italy participants from EnviroGenoMarkers (N = 155).

Inflammatory methylome scores	Father's occupational position		Participant's education		Highest household occupational position	
	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
Inflammatory methylome score	-0.0023 (0.0028)	4.18×10^{-1}	-0.0075 (0.0028)	6.72×10^{-3}	-0.0076 (0.0028)	7.37×10^{-3}
Inflammatory methylome z-score	-2.325 (4.0788)	5.70×10^{-1}	-9.7219 (3.9778)	1.57×10^{-2}	-11.4191 (4.0258)	5.20×10^{-3}
Principal component 1	-0.713 (0.9193)	4.39×10^{-1}	-1.6264 (0.9055)	7.45×10^{-2}	-2.0508 (0.9172)	2.69×10^{-2}

Results are given for Model 1 controlling for age, gender, lymphoma and breast cancer case-control status, phase and center.
SE: Standard error.

respectively). The inflammatory methylome score was lower in participants whose father had a low occupational position but the association was not significant (Table 2). Adjusting for either behavioral factors, BMI or all covariates together only marginally affected our results: effect size estimates showed consistent signs, and associations remained statistically significant for both participant's education and the highest household occupational position ($p < 0.05$ across all models investigated, Supplementary Table 4).

The association between education level and lower inflammatory methylome score was not affected by adjusting for father's occupational position (Table 3A, Model B-1, $\beta = -0.008$; $p = 0.008$), and weakened while controlling for the highest household occupational position (Table 3A, Model C, $\beta = -0.006$; $p = 0.08$). Adjustments for potential (behavioral) mediators further weakened this association, which lost statistical significance (Table 3A, Model D, $\beta = -0.005$; $p = 0.139$). Further adjustment for the estimated cell type proportions slightly strengthened the association with participant's education (Table 3A, fully adjusted model, $\beta = -0.006$; $p = 0.030$). The association between the inflammatory methylome score and the highest household occupational position was weakened by adjusting for father's occupational position (Table 3A, Model B-1, $\beta = -0.008$; $p = 0.009$), and no longer significant when controlling participant's education (Table 3A, Model B-2, $\beta = -0.008$; $p = 0.011$ and Model C, $\beta = -0.005$; $p = 0.103$) and for potential (behavioral) mediators (Table 3A, Model D, $\beta = -0.005$; $p = 0.106$) as well as in the fully adjusted model ($\beta = -0.003$; $p = 0.175$).

As a sensitivity analyses, the inflammatory methylome score was also investigated by considering two alternative definitions: we calculated a z-score and ran a principal component analyses on the 61 *cis*-eMS (PC1, explaining 44.2% of the variance). Our results were slightly weakened but remained consistent; significant associations between educational level and highest household occupational position and the inflammatory methylome scores were observed (Table 2). We identified consistent associations for both alternative scores (z-score and PC1) after adjusting for either behavioral factors, BMI or all covariates together (Supplementary Table 4). Our results and conclusions from the sequentially adjusted model for early-life, young adulthood and adulthood SEP indicators remained markedly stable albeit weakened irrespective of the score considered: the z-score (Table 3B) and the first principal component from the 61 DNA methylation level (Table 3C).

Testing the functional relevance of the inflammatory methylome at the protein level

To evaluate whether DNA methylation levels affect inflammatory protein concentrations, we looked for CpG sites associated with 27 inflammatory proteins in a subset of 173 EPIC-Italy participants (see Methods section). We identified five potential regulatory methylation sites (pMS) as CpG sites whose methylation was associated with protein level of the CpG site's closest gene ($p < 0.05$, Supplementary Table 2). Adjustment for the estimated cell type proportions (see Methods section, lymphocytes and monocytes) slightly weakened the association between DNA methylation level and protein concentrations for two CpG sites. As above, the DNA methylation score was calculated for the inflammatory methylome by taking the arithmetic mean of the methylation levels from the five pMS. We observed a significant association between participant's education and the highest household occupational position and the inflammatory methylome score where disadvantaged participants had a lower inflammatory methylome score compared with their advantaged counterparts (Supplementary Table 5, Model 1, $\beta = -0.0044$; $p = 0.025$; $\beta = -0.0056$; $p = 0.005$ for educational level and the highest household occupational position, respectively). Effect size estimates showed consistent signs, and associations reached statistical significance for both participant's education and the highest household occupational position after controlling for either behavioral factors and BMI, or all of them together (Supplementary Table 5, Model 2, $\beta = -0.0043$; $p = 0.033$; $\beta = -0.0051$; $p = 0.0115$ for educational

Table 3. Life course multiple regression analyses and the inflammatory methylome. Results are presented for the inflammatory methylome score (A), the inflammatory methylome z-score (B) and the first PC (C). Estimates are based on 155 participants with complete socioeconomic position, lifestyle information and DNA methylation data.

Inflammatory methylome scores		Model A		Model B-1		Model B-2		Model C		Model D		Fully adjusted model†	
Variables	Levels	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
(A) Inflammatory methylome score													
Father's occupational position	Manual	-0.002 (0.003)	4.18 × 10 ⁻¹	0 (0.003)	6.22 × 10 ⁻¹	0 (0.003)	9.96 × 10 ⁻¹	0.002 (0.003)	5.07 × 10 ⁻¹	0.002 (0.003)	5.12 × 10 ⁻¹	0.002 (0.002)	3.87 × 10 ⁻¹
Participant's education	Low	-0.008 (0.003)	8.67 × 10 ⁻³					-0.006 (0.003)	8.05 × 10 ⁻²	-0.005 (0.003)	1.39 × 10 ⁻¹	-0.006 (0.003)	2.98 × 10 ⁻²
Highest household occupational position	Manual			-0.008 (0.003)	1.09 × 10 ⁻²			-0.005 (0.003)	1.03 × 10 ⁻¹	-0.005 (0.003)	1.07 × 10 ⁻¹	-0.003 (0.002)	1.76 × 10 ⁻¹
BMI								0 (0)	3.02 × 10 ⁻¹	0 (0)	3.02 × 10 ⁻¹	0 (0)	8.94 × 10 ⁻¹
Smoking status	Former							0.002 (0.004)	5.14 × 10 ⁻¹	0.001 (0.003)	8.04 × 10 ⁻¹	0.001 (0.003)	8.04 × 10 ⁻¹
	Current							-0.001 (0.004)	7.13 × 10 ⁻¹	0.001 (0.003)	6.58 × 10 ⁻¹	0.001 (0.003)	6.58 × 10 ⁻¹
Alcohol								0 (0)	7.18 × 10 ⁻¹	0 (0)	7.18 × 10 ⁻¹	0 (0)	7.91 × 10 ⁻¹
(B) Inflammatory methylome z-score													
Variables	Levels	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
Father's occupational position	Manual	-2.325 (4.079)	5.70 × 10 ⁻¹	2.791 (4.523)	5.38 × 10 ⁻¹	1.259 (4.192)	7.64 × 10 ⁻¹	3.706 (4.506)	4.12 × 10 ⁻¹	3.969 (4.613)	3.91 × 10 ⁻¹	3.864 (3.222)	2.32 × 10 ⁻¹
Participant's education	Low	-11.003 (4.494)	1.55 × 10 ⁻²					-7.073 (4.889)	1.50 × 10 ⁻¹	-5.481 (4.971)	2.72 × 10 ⁻¹	-6.129 (3.474)	7.99 × 10 ⁻²
Highest household occupational position	Manual			-11.811 (4.243)	6.09 × 10 ⁻³			-9.038 (4.642)	5.35 × 10 ⁻²	-9.186 (4.706)	5.29 × 10 ⁻²	-6.105 (3.311)	6.73 × 10 ⁻²
BMI								-0.783 (0.568)	1.70 × 10 ⁻¹	-0.224 (0.399)	5.75 × 10 ⁻¹	-0.224 (0.399)	5.75 × 10 ⁻¹
Smoking status	Former							5.024 (5.357)	3.50 × 10 ⁻¹	2.356 (3.752)	5.31 × 10 ⁻¹	2.356 (3.752)	5.31 × 10 ⁻¹
	Current							-0.946 (5.101)	8.53 × 10 ⁻¹	2.772 (3.585)	4.41 × 10 ⁻¹	2.772 (3.585)	4.41 × 10 ⁻¹
Alcohol								0.086 (0.136)	5.30 × 10 ⁻¹	0.006 (0.096)	9.54 × 10 ⁻¹	0.006 (0.096)	9.54 × 10 ⁻¹

†Model adjusted for cell blood composition (see Methods section). SE: Standard error.

Table 3. Life course multiple regression analyses and the inflammatory methylome. Results are presented for the inflammatory methylome score (A), the inflammatory methylome z-score (B) and the first PC (C). Estimates are based on 155 participants with complete socioeconomic position, lifestyle information and DNA methylation data (cont.).

(C) Principal component 1										
Variables	Levels	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value	p-value
Father's occupational position	Manual	-0.713 (0.919)	4.39×10^{-1}	0.055 (1.031)	9.58×10^{-1}	-0.1 (0.955)	9.17×10^{-1}	0.222 (1.032)	8.30×10^{-1}	8.14×10^{-1}
Participant's education	Low	-1.652 (1.024)	1.09×10^{-1}	-	-	-0.952 (1.119)	4.06×10^{-1}	-0.625 (1.143)	5.85×10^{-1}	5.85×10^{-1}
Highest household occupational position	Manual	-	-	-2.02 (0.967)	3.85×10^{-2}	-1.654 (1.063)	1.22×10^{-1}	-1.663 (1.081)	1.26×10^{-1}	1.26×10^{-1}
BMI								-0.155 (0.13)	2.37×10^{-1}	2.37×10^{-1}
Smoking status	Former							1.027 (1.231)	4.06×10^{-1}	4.06×10^{-1}
	Current							-0.101 (1.172)	9.31×10^{-1}	9.31×10^{-1}
Alcohol								0.013 (0.031)	6.81×10^{-1}	6.81×10^{-1}

†Model adjusted for cell blood composition (see Methods section).
SE: Standard error.

level and the highest household occupational position, respectively). By sequentially controlling for time-ordered SEP, we observed that the association between educational attainment and inflammatory methylome score was slightly weakened by adjusting for father's occupational position, and no longer significant when controlling for the highest household occupational position (Supplementary Table 6A, Model B-1 and C, respectively, $\beta = -0.005$; $p = 0.032$; $\beta = -0.003$; $p = 0.24$). The association between the highest household occupational position and the inflammatory methylome was not affected upon adjustment for early-life SEP (Model B-2, $\beta = -0.006$; $p = 0.0072$) and weakened when controlling for young adulthood SEP (Model C, $\beta = -0.005$; $p = 0.047$) and upon adjustment for potential mediators (Model D, $\beta = -0.004$; $p = 0.0821$; Supplementary Table 6A). Associations were consistent using either alternative scores (z-score and the principal component 1 explaining 35.6% of the variance, Supplementary Table 6B & C, respectively) and controlling for cell composition marginally affected the results (Supplementary Table 6, fully adjusted model).

Integration of the inflammatory methylome with the inflammatory transcriptome & proteome

To investigate the potential increased inflammatory responses to a lower inflammatory methylome score at the gene expression level and proteins concentration, we further explored the relationship between the inflammatory methylome and the corresponding transcriptome by life course SEP. Briefly, the 61 genes measured in the same 155 samples with detailed information on SEP and behavioral factors and BMI were combined into a score by taking the mean. We first observed a negative correlation between the inflammatory methylome and transcriptome scores (Spearman correlation = -0.35). Results by life course SEP are given Figure 2, a modest increase of the corresponding inflammatory transcriptome was observed only in adulthood (Figure 2C), while no differences were observed for early-life and young adulthood SEP (Figure 2A & B). Additional analyses using the proteins-based definition revealed a modest increase of the corresponding inflammatory proteome only in participant's whose father had a manual occupation (Supplementary Figure 8A) and no differences were observed later in life (Supplementary Figure 8B & C).

Discussion

In the present study, we employed a gene candidate approach to specifically look for inflammatory related CpG sites associated with their closest gene at two molecular levels (gene expression and proteins) in order to define a functionally relevant inflammatory methylome. We further investigated the association between SEP at different time points in life and the inflammatory methylome as defined by the methylation levels at the 61 *cis*-acting CpG sites. While we did not identify any significant association after correction for multiple testing when examining each of the 61 CpGs associated with gene expression in relation to the three SEP indicators separately (which could be due to the limited statistical power our population size yields), the use of an inflammatory methylome score combining methylation levels across all 61 *cis*-acting CpG sites revealed significant associations with SEP. In particular, we found that while no difference was observed between participants reporting a father with a 'manual' occupation compared with those whose father had a 'non-manual' occupation, participants with a less advantaged SEP in young adulthood or in adulthood, later in life, had a lower inflammatory methylome score compared with participants with a more advantaged SEP. These associations, independent of major lifestyle-related risk factors, were robust to the definition of the inflammatory methylome score. Analysis using a life course approach indicated a stronger effect of young adulthood SEP on the inflammatory methylome albeit weakened when controlling for SEP in adulthood. This is suggestive of a shared pathway between SEP at each life stage as well as an indirect pathway between educational attainment and the inflammatory methylome. The exploratory analyses testing the functional relevance of CpGs associated with protein concentration showed consistent, though statistically weakened, associations between life course SEP and the protein-based inflammatory methylome.

Some limitations of this study should be considered. We used data from a cancer case-control study nested within the EPIC-Italy cohort meaning that half of the participants developed either a non-Hodgkin's lymphoma or a breast cancer during the follow-up affecting both representativeness and generalizability of our results. However, blood samples were collected at enrollment when the participants were healthy. We further selected participants with complete information on life course SEP, behavioral variables, BMI and measure of inflammation at multiple molecular levels that are subjected to measurement errors and leading to a small-sample size limiting the statistical power and restricting our methodological choices. All three life course SEP were recorded into binary indicators leading to some degree of misclassification. Associations between SEP and the inflammatory methylome score were independent of major behaviors and lifestyle-related risk factors. We did not investigate the mediating roles of

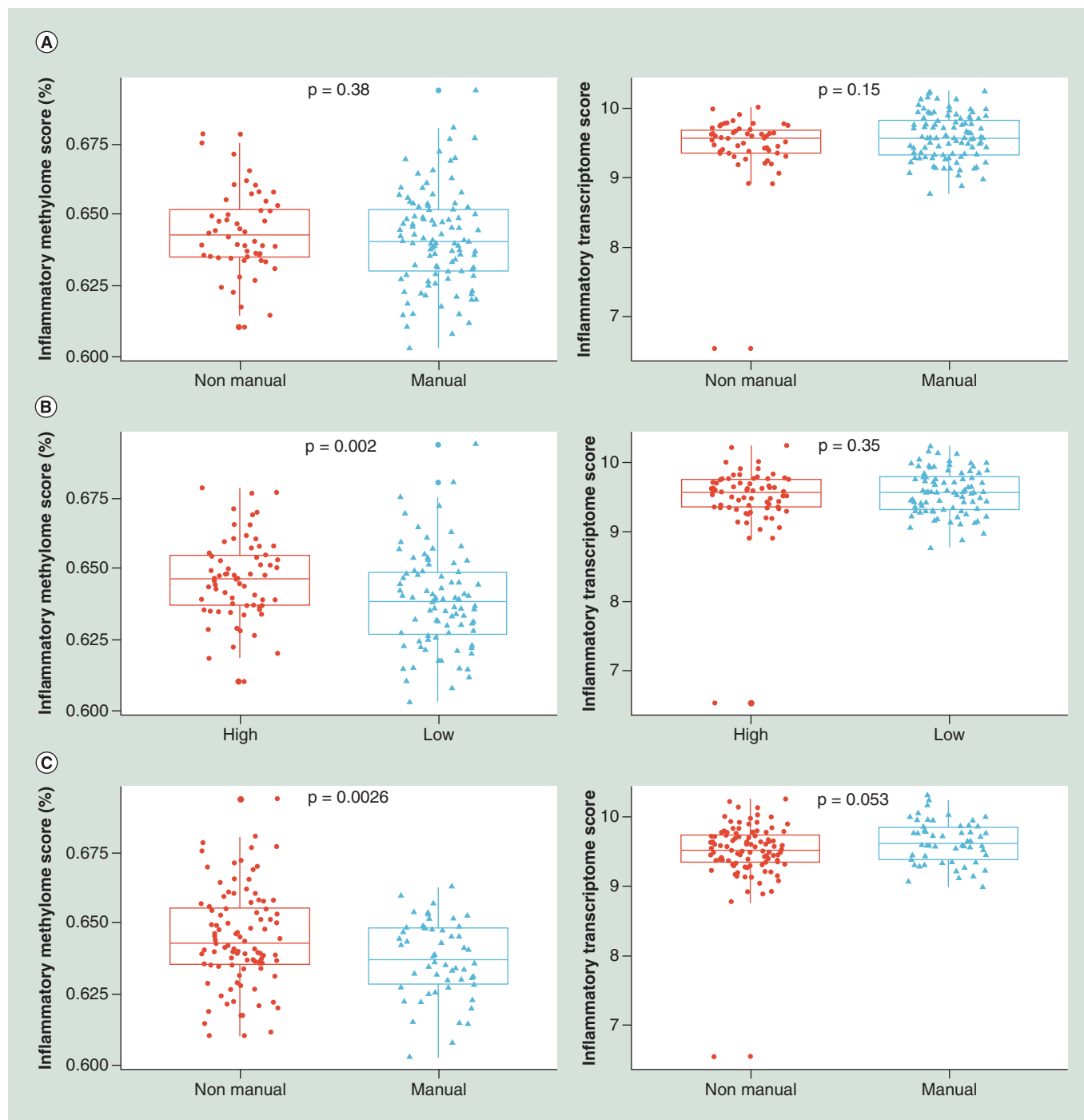


Figure 2. Boxplot summarizing the distribution of the inflammatory methylome (on the left) and corresponding inflammatory transcriptome (on the right) score by life course socioeconomic position. (A) Father's occupational position; (B) participant's education and (C) the highest household occupational position from 155 EPIC-Italy participants. t-test mean comparison p-values are reported for each score by SEP.

EPIC-Italy: European Prospective Investigation into Cancer and Nutrition-Italy; SEP: Socioeconomic position.

those risk factors within a life course approach using mediation-based methods. Additionally, we cannot rule out: the possibility that other factors may be potentially involved in social differences in DNA methylation including infection, environmental or psychosocial exposures, but these exposures were not available in the study; the presence of time-varying confounding. Conventional statistical methods in the presence of time-varying confounding can

result to an overadjustment bias by removing part of the effect of the past exposure and/or collider stratification bias.

We considered only the CpG sites in *cis*, although some informative sites may also act in *trans*. As a sensitivity analyses we ran all analyses restricting our definition of the inflammatory methylome on the 48 CpGs sites inversely associated with transcripts levels and results were mostly unchanged (Supplementary Tables 7 & 8). Another limitation is in scores measurement because there is no standard measure for building scores, however, we proposed several scores to approach the inflammatory methylome and our results were consistent. Finally, DNA methylation was assessed from whole blood and could be confounded by cellular heterogeneity [40], however, our results were only marginally affected after correction for cell composition.

Strengths of our study include the use of widely used measures of life course SEP as well as measurement of inflammation at multiple molecular scales (methylation level, gene expression and protein concentration) together with a wide range of information about lifestyle factors that allows us to control for important potential mediators. Moreover, to our knowledge, no other studies have examined associations of life course SEP with inflammation using an integrated approach of multiple omics levels.

Several recent studies have supported the possibility of differential DNA methylation profiles in peripheral blood cells associated with individuals social factors [41] and other risk factors [42]. In a previous study using participants from EPIC-Italy, we reported that indicators of life course SEP were associated with DNA methylation levels in genes (N = 17 genes, corresponding to 403 CpG sites) involved in inflammation [29]. Numerous studies have investigated associations between childhood disadvantage and DNA methylation pattern later in life [22,24,25,28,43]. A genome wide methylation scan was performed in 40 adult men from the 1958 British Birth Cohort and identified differential methylation of 1252 gene promoters [24]. Another study showed association of childhood SEP with methylation of SAT2 repetitive elements in more than 80 participants from the New York site of the Collaborative Perinatal Project [22]. A gene candidate approach was performed in more than 1000 participants from the Multi-Ethnic Study of Atherosclerosis focusing on 18 candidate genes related to stress reactivity and inflammation demonstrating that disadvantaged childhood SEP was associated with adulthood DNA methylation in three stress-related genes and two inflammation-related genes [28]. Evidence is accumulating suggesting an association between childhood socioeconomic disadvantage and methylation patterns.

Our study did not find support for the hypothesis that disadvantaged SEP in early life was associated with the inflammatory methylome independent of later in life SEP, which might be partly due to our limited sample size and the selection of the study population. However, differences in our conclusions may also reflect contextual and methodological aspect including the definition of SEP and/or the methods used to analyze DNA methylation level. Nevertheless, our study underlines the influence of education on inflammatory methylome in accordance to another study showing a consistent association between educational attainment and inflammation later in life measured using the CRP [44].

Conclusion

Our study suggests that low SEP, particularly education, is associated with a global hypomethylation of the inflammatory methylome in peripheral blood cells, which potentially increased expression of the corresponding inflammatory genes, and may potentially contribute in the long term to the onset and progression of inflammatory-related disorders like cardiovascular diseases, diabetes and cancers.

More work is needed to examine the functional consequences of the observed inflammatory load. A deeper understanding of the mechanisms by which social environment influences inflammation could have profound implications in treatment and more specifically in prevention, by potentially identifying modifiable environmental factors that affect physiological health.

Future perspective

In the upcoming years, researchers should strive to design longitudinal studies in which participants from disadvantaged populations from early life are included and well characterized along the life course in term of exposures and blood collection to better understand the role of epigenetic variation in response to environmental and psychosocial stressors on adverse health outcomes. Further functional studies will be particularly useful to investigate the downstream consequences of epigenomic variations. Similar efforts should be undertaken to better characterize the complex interplay between environmental stress, genotype and epigenetic to characterize to which extent this complex interplay contribute to explain social inequalities in health.

Summary points

- Chronic inflammation has emerged as a mechanism by which socioeconomic position (SEP) over the life course confers a risk for many health outcomes.
- We adopted a gene–candidate approach to explore the inflammatory response to life course SEP.
- We defined a functionally relevant inflammatory methylome based on the identification of CpG sites associated with inflammatory-related genes and proteins.
- When each CpG sites were examined individually, no significant associations with life course SEP were found.
- The use of an inflammatory methylome score revealed significant associations with SEP.
- Participants with a less-advantaged SEP in young adulthood and in adulthood had a lower inflammatory methylome score compared with their advantaged counterparts.
- These associations were independent of the behavioral factors considered and BMI and robust to the score definition.
- The protein-based definition of the inflammatory methylome showed weakened but consistent results.
- Integrative analyses suggest that the lower inflammatory methylome was correlated to a modest increase of the corresponding inflammatory transcriptome or proteome.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/epi-2019-0261

Acknowledgments

The authors thank all the centers that took part in the study and the additional members of the EnviroGenoMarkers Consortium. For further information on the EnviroGenoMarkers Consortium, please visit <http://www.envirogenomarkers.net/>

Financial & competing interests disclosure

This work was supported by the Institut National du Cancer (SHSESP14–082 to R Castagné). This study was supported by the European Commission Horizon 2020 grant number 633666 to P Vineis and the 7th European Framework Programme (FP7) for the Envirogenomarkers (226756 to SA Kyrtopoulos) projects. The Centre for Environment and Health is supported by the Medical Research Council and Public Health England (MR/L01341X/1). M Chadeau-Hyam acknowledges support from Cancer Research UK, Population Research Committee Project grant ('Mechanomics' project #22184 to M Chadeau-Hyam). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

For all human investigations, appropriate institutional review board approval was obtained following the principles outlined in the Declaration of Helsinki. In addition, informed consent has been obtained from the participants involved.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Stringhini S, Carmeli C, Jokela M *et al.* Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *Lancet Lond. Engl.* 389(10075), 1229–1237 (2017).
- **An important study showing that disadvantaged socioeconomic position reduced life expectancy and compared this reduction with the effect of major risk factors.**
2. Adler NE, Ostrove JM. Socioeconomic status and health: what we know and what we don't. *Ann. NY Acad. Sci.* 896, 3–15 (1999).
3. Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet* 372(9650), 1661–1669 (2008).
4. Stringhini S, Dugravot A, Shipley M *et al.* Health behaviours, socioeconomic status, and mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. *PLoS Med.* 8(2), e1000419 (2011).

5. Gallo V, Mackenbach JP, Ezzati M *et al.* Social inequalities and mortality in Europe – results from a large multi-national cohort. *PLoS ONE* 7(7), e39013 (2012).
6. McEwen BS. Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiol. Rev.* 87(3), 873–904 (2007).
- **A very important review providing a conceptual framework to study the complex stress response to social exposures.**
7. Yang EV, Glaser R. Stress-induced immunomodulation and the implications for health. *Int. Immunopharmacol.* 2(2–3), 315–324 (2002).
8. Nazmi A, Victora CG. Socioeconomic and racial/ethnic differentials of C-reactive protein levels: a systematic review of population-based studies. *BMC Public Health* 7, 212 (2007).
9. Liu RS, Aiello AE, Mensah FK *et al.* Socioeconomic status in childhood and C reactive protein in adulthood: a systematic review and meta-analysis. *J. Epidemiol. Community Health* 71(8), 817–826 (2017).
10. Muscatell KA, Brosso SN, Humphreys KL. Socioeconomic status and inflammation: a meta-analysis. *Mol. Psychiatry*(2018) (Online ahead of print).
- **A recent meta-analysis on the social patterning of the CRP and IL-6.**
11. Chen E, Miller GE, Walker HA, Arevalo JM, Sung CY, Cole SW. Genome-wide transcriptional profiling linked to social class in asthma. *Thorax* 64(1), 38–43 (2009).
12. Chen E, Miller GE, Kobor MS, Cole SW. Maternal warmth buffers the effects of low early-life socioeconomic status on pro-inflammatory signaling in adulthood. *Mol. Psychiatry* 16(7), 729–737 (2011).
13. Levine ME, Cole SW, Weir DR, Crimmins EM. Childhood and later life stressors and increased inflammatory gene expression at older ages. *Soc. Sci. Med.* 130, 16–22 (2015).
14. Gaye A, Gibbons GH, Barry C, Quarells R, Davis SK. Influence of socioeconomic status on the whole blood transcriptome in African Americans. *PLoS ONE* 12(12), e0187290 (2017).
15. Cole SW, Hawkey LC, Arevalo JM, Sung CY, Rose RM, Cacioppo JT. Social regulation of gene expression in human leukocytes. *Genome Biol.* 8(9), R189 (2007).
- **One of the first study investigating the relationship between social isolation and gene expression in human.**
16. Cole SW, Hawkey LC, Arevalo JMG, Cacioppo JT. Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes. *Proc. Natl Acad. Sci. USA* 108(7), 3080–3085 (2011).
17. Miller GE, Chen E, Fok AK *et al.* Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling. *Proc. Natl Acad. Sci. USA* 106(34), 14716–14721 (2009).
18. Cole SW. Social regulation of human gene expression: mechanisms and implications for public health. *Am. J. Public Health* 103(S1), S84–S92 (2013).
19. Cole SW. Human social genomics. *PLoS Genet.* 10(8), e1004601 (2014).
20. Simanek AM, Auer PL. Early life socioeconomic disadvantage and epigenetic programming of a pro-inflammatory phenotype: a review of recent evidence. *Curr. Epidemiol. Rep.* 5, 407–417 (2018).
21. McGuinness D, McGlynn LM, Johnson PCD *et al.* Socio-economic status is associated with epigenetic differences in the pSoBid cohort. *Int. J. Epidemiol.* 41(1), 151–160 (2012).
- **The aim of this study was to investigate the relationship between socioeconomic position and global DNA methylation in human.**
22. Tehranifar P, Wu H-C, Fan X *et al.* Early life socioeconomic factors and genomic DNA methylation in mid-life. *Epigenetics* 8(1), 23–27 (2013).
23. Subramanyam MA, Diez-Roux AV, Pilsner JR *et al.* Social factors and leukocyte DNA methylation of repetitive sequences: the multi-ethnic study of atherosclerosis. *PLoS ONE* 8(1), e54018 (2013).
24. Borghol N, Suderman M, McArdle W *et al.* Associations with early-life socio-economic position in adult DNA methylation. *Int. J. Epidemiol.* 41(1), 62–74 (2012).
25. Lam LL, Emberly E, Fraser HB *et al.* Factors underlying variable DNA methylation in a human community cohort. *Proc. Natl Acad. Sci. USA* 109(Suppl. 2), 17253–17260 (2012).
26. Beach SRH, Lei M-K, Brody GH *et al.* Parenting, socioeconomic status risk, and later young adult health: exploration of opposing indirect effects via DNA methylation. *Child Dev.* 87(1), 111–121 (2016).
27. van Dongen J, Bonder MJ, Dekkers KF *et al.* DNA methylation signatures of educational attainment. *NPJ Sci. Learn.* 3, 7 (2018).
28. Needham BL, Smith JA, Zhao W *et al.* Life course socioeconomic status and DNA methylation in genes related to stress reactivity and inflammation: the multi-ethnic study of atherosclerosis. *Epigenetics* 10(10), 958–969 (2015).
29. Stringhini S, Polidoro S, Sacerdote C *et al.* Life-course socioeconomic status and DNA methylation of genes regulating inflammation. *Int. J. Epidemiol.* 44(4), 1320–1330 (2015).
30. McDade TW, Ryan C, Jones MJ *et al.* Social and physical environments early in development predict DNA methylation of inflammatory genes in young adulthood. *Proc. Natl Acad. Sci. USA* 114(29), 7611–7616 (2017).

31. Alfano R, Guida F, Galobardes B *et al.* Socioeconomic position during pregnancy and DNA methylation signatures at three stages across early life: epigenome-wide association studies in the ALSPAC birth cohort. *Int. J. Epidemiol.* 48(1), 30–44 (2018).
32. Smith JA, Zhao W, Wang X *et al.* Neighborhood characteristics influence DNA methylation of genes involved in stress response and inflammation: the multi-ethnic study of atherosclerosis. *Epigenetics* 12(8), 662–673 (2017).
33. Castagné R, Delpierre C, Kelly-Irving M *et al.* A life course approach to explore the biological embedding of socioeconomic position and social mobility through circulating inflammatory markers. *Sci. Rep.* 6, 25170 (2016).
34. Castagné R, Kelly-Irving M, Campanella G *et al.* Biological marks of early-life socioeconomic experience is detected in the adult inflammatory transcriptome. *Sci. Rep.* 6, 38705 (2016).
35. Hebels DGAJ, Georgiadis P, Keun HC *et al.* Performance in omics analyses of blood samples in long-term storage: opportunities for the exploitation of existing biobanks in environmental health research. *Environ. Health Perspect.* 121(4), 480–487 (2013).
36. Palli D, Berrino F, Vineis P *et al.* A molecular epidemiology project on diet and cancer: the EPIC-Italy Prospective Study. Design and baseline characteristics of participants. *Tumori.* 89(6), 586–593 (2003).
37. Hallmans G, Agren A, Johansson G *et al.* Cardiovascular disease and diabetes in the Northern Sweden Health and Disease Study Cohort – evaluation of risk factors and their interactions. *Scand. J. Public Health Suppl.* 61, 18–24 (2003).
38. Georgiadis P, Liampa I, Hebels DG *et al.* Evolving DNA methylation and gene expression markers of B-cell chronic lymphocytic leukemia are present in pre-diagnostic blood samples more than 10 years prior to diagnosis. *BMC Genomics* 18(1), 728 (2017).
39. Chadeau-Hyam M, Vermeulen RCH, Hebels DG *et al.* Prediagnostic transcriptomic markers of Chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis. *Ann. Oncol.* 25(5), 1065–1072 (2014).
40. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics* 9(5), 757–768 (2017).
41. Esposito EA, Jones MJ, Doom JR, MacIsaac JL, Gunnar MR, Kober MS. Differential DNA methylation in peripheral blood mononuclear cells in adolescents exposed to significant early but not later childhood adversity. *Dev. Psychopathol.* 28(4 Pt 2), 1385–1399 (2016).
42. Terry MB, Delgado-Cruzata L, Vin-Raviv N, Wu HC, Santella RM. DNA methylation in white blood cells. *Epigenetics* 6(7), 828–837 (2011).
43. Bush NR, Edgar RD, Park M *et al.* The biological embedding of early-life socioeconomic status and family adversity in children's genome-wide DNA methylation. *Epigenomics* 10(11), 1445–1461 (2018).
44. Berger E, Castagné R, Chadeau-Hyam M *et al.* Multi-cohort study identifies social determinants of systemic inflammation over the life course. *Nat. Commun.* 10(1), 773 (2019).