



HAL
open science

Impact of socioeconomic status on healthy immune responses in humans

Anthony Bertrand, Jamie Sugrue, Tianai Lou, Nollaig Bourke, Lluís Quintana-Murci, Violaine Saint-André, Cliona O'Farrelly, Darragh Duffy

► **To cite this version:**

Anthony Bertrand, Jamie Sugrue, Tianai Lou, Nollaig Bourke, Lluís Quintana-Murci, et al.. Impact of socioeconomic status on healthy immune responses in humans. *Immunology and Cell Biology*, In press, 10.1111/imcb.12789 . pasteur-04614647

HAL Id: pasteur-04614647

<https://pasteur.hal.science/pasteur-04614647>

Submitted on 17 Jun 2024


HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Impact of socioeconomic status on healthy immune responses in humans

Anthony Bertrand^{1,2} , Jamie Sugrue¹, Tianai Lou³, Nollaig M Bourke⁴, Lluís Quintana-Murci^{5,6}, Violaine Saint-André^{1,7}, Cliona O'Farrelly^{3,8,a}, Darragh Duffy^{1,a} & the Milieu Intérieur Consortium^b

¹ Translational Immunology Unit, Institut Pasteur, Université Paris Cité, Paris, France

² Frontiers of Innovation in Research and Education PhD Program, LPI Doctoral School, Université Paris Cité, Paris, France

³ School of Medicine, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Ireland

⁴ Discipline of Medical Gerontology, School of Medicine, Trinity Translational Medicine Institute, Trinity College Dublin, Dublin, Ireland

⁵ Human Evolutionary Genetics Unit, Institut Pasteur, Université Paris Cité, CNRS UMR2000, Paris, France

⁶ Chair of Human Genomics and Evolution, Collège de France, Paris, France

⁷ Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, Paris, France

⁸ School of Biochemistry & Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Ireland

Keywords

CMV serostatus, Elo, immune response, socioeconomic status, transcriptomic

Correspondence

Cliona O'Farrelly, School of Medicine, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Ireland.

E-mail: ofarrecl@tcd.ie

Darragh Duffy, Translational Immunology Unit, Institut Pasteur, Université Paris Cité, Paris, France.

E-mail: darragh.duffy@pasteur.fr

Violaine Saint-André, Translational Immunology Unit, Institut Pasteur, Université Paris Cité, Paris, France.

E-mail: violaine.saint-andre@pasteur.fr

^aEqual contribution.

^bThe Milieu Intérieur Consortium is composed of the following team leaders presented in the Acknowledgements section.

Received 20 February 2024;

Revised 23 April and 20 May 2024;

Accepted 20 May 2024

doi: 10.1111/imcb.12789

Immunology & Cell Biology 2024; 1–12

Abstract

Individuals with low socioeconomic status (SES) are at greater risk of contracting and developing severe disease compared with people with higher SES. Age, sex, host genetics, smoking and cytomegalovirus (CMV) serostatus are known to have a major impact on human immune responses and thus susceptibility to infection. However, the impact of SES on immune variability is not well understood or explored. Here, we used data from the *Milieu Intérieur* project, a study of 1000 healthy volunteers with extensive demographic and biological data, to examine the effect of SES on immune variability. We developed an Elo-rating system using socioeconomic features such as education, income and home ownership status to objectively rank SES in the 1000 donors. We observed sex-specific SES associations, such as females with a low SES having a significantly higher frequency of CMV seropositivity compared with females with high SES, and males with a low SES having a significantly higher frequency of active smoking compared with males with a high SES. Using random forest models, we identified specific immune genes which were significantly associated with SES in both baseline and immune challenge conditions. Interestingly, many of the SES associations were sex stimuli specific, highlighting the complexity of these interactions. Our study provides a new way of computing SES in human populations that can help identify novel SES associations and reinforces biological evidence for SES-dependent susceptibility to infection. This should serve as a basis for further understanding the molecular mechanisms behind SES effects on immune responses and ultimately disease.

INTRODUCTION

Socioeconomic status (SES) is a measure of the relative social position of individuals within a population. It is determined by factors such as education, home ownership status, income, parental status and occupation.¹ SES is a strong indicator of inequalities in health, with people from lower SES backgrounds having a higher risk of multimorbidity, premature death, poorer cancer prognosis and a higher risk of infection.² Many complex social factors contribute to increased susceptibility to ill health, such as lower health care access, increased social isolation, higher infectious exposure and increased smoking and alcohol consumption rates, which makes the study of underlying biological processes challenging.³ Moreover, stress is known to be associated with differences in the immune response.^{4,5} The central nervous system, the endocrine system and the immune system are closely linked by autoregulation. The production of stress hormones modulates the production of cytokines.^{6,7} Exposition to chronic stress can lead to a conserved transcriptional response to adversity involving upregulation of proinflammatory genes and a downregulation of type I interferon responses.⁸ Similarly, adolescents with symptoms of depression showed increases in the level of expression of inflammation genes and a reduction in the expression of antiviral genes.⁹

Low SES has been strongly associated with differences in inflammation and chronic diseases, with higher levels of circulating C-reactive protein (CRP), a common inflammatory marker, consistently observed in people of low SES.^{10,11} Studies of macaques, which have strong social ranking in their societies, have also provided the strongest evidence of a direct effect of social position on immunity independently of the compounding influences in human studies.¹² For example, high SES macaques have significantly higher levels of cytotoxic T cells, and therefore, lower CD4⁺/CD8⁺ ratios compared with lower SES macaques. Interestingly, changes in SES status within the group were mirrored by changes in gene expression within T helper cells and natural killer cells, suggesting a plasticity in the effect of social status on immune responses. Although these studies are highly informative, social hierarchy in primates is determined by direct individual-to-individual interactions in contrast to SES in humans, which is determined more indirectly at a societal level and includes additional social factors as highlighted above.

This diversity of SES parameters can make the concept challenging to study in humans. To address some of these limitations, we utilized a well-established

human cohort study, the *Milieu Interieur* (MI) cohort.¹³ Critically, the cohort is balanced in terms of age and sex, and donors were defined as healthy, based on an objective list of inclusion and exclusion criteria. The cohort reflects the general French population without extremes of income.¹³ Within this cohort, we combined three SES features, namely, educational attainment, home ownership status and income, to generate an Elo ranking system of multiple pairwise comparisons that defined the relative socioeconomic position of each individual. By analyzing diverse *Milieu Intérieur* data sets according to SES, we demonstrate a significant effect of SES on immune response variability within a healthy human population.

RESULTS

Computation of the socioeconomic ranking of individuals in a cohort of 1000 healthy donors

Applying the Elo rating system (as described in the “Methods” section) to the *Milieu Intérieur* cohort highlights socioeconomic disparities regarding age. There is a peak in density for lower Elo for the 20–30-year age bracket while the other age ranges have similar distributions (Figure 1e). Because of this disparity, we excluded the 20–29-year age range in the models to avoid any bias related to age. No significant differences in Elo levels were found between the sexes (Figure 1f).

The age disparity observed in Elo can be explained by proportional differences for each of the SES variables of interest as a function of age within the *Milieu Intérieur* cohort. There is a peak in proportions for salaries between €0 and €1000/month among donors aged 20 to 29 years. For salaries between €2000 and €3000/month or €3000 and €4000/month, the proportions are higher for donors over the age of 30 years. In addition, the proportions for higher incomes (over €4000/month) increase slightly with age (Figure 2a). Similarly, for housing, there is a peak in the proportion of leasers in the 20–29-year age bracket. For donors over the age of 30 years, the proportion of homeowners is higher than that of leasers (Figure 2b). By contrast, for education levels, there is a greater disparity between donors in the 20–39-year age bracket and those over 40 years. In fact, the proportion of donors with a technical degree or third-level education is higher among donors aged between 20 and 39 years. There is also a higher proportion of people who have completed less than secondary school for donors over the age of 40 years (Figure 2c).

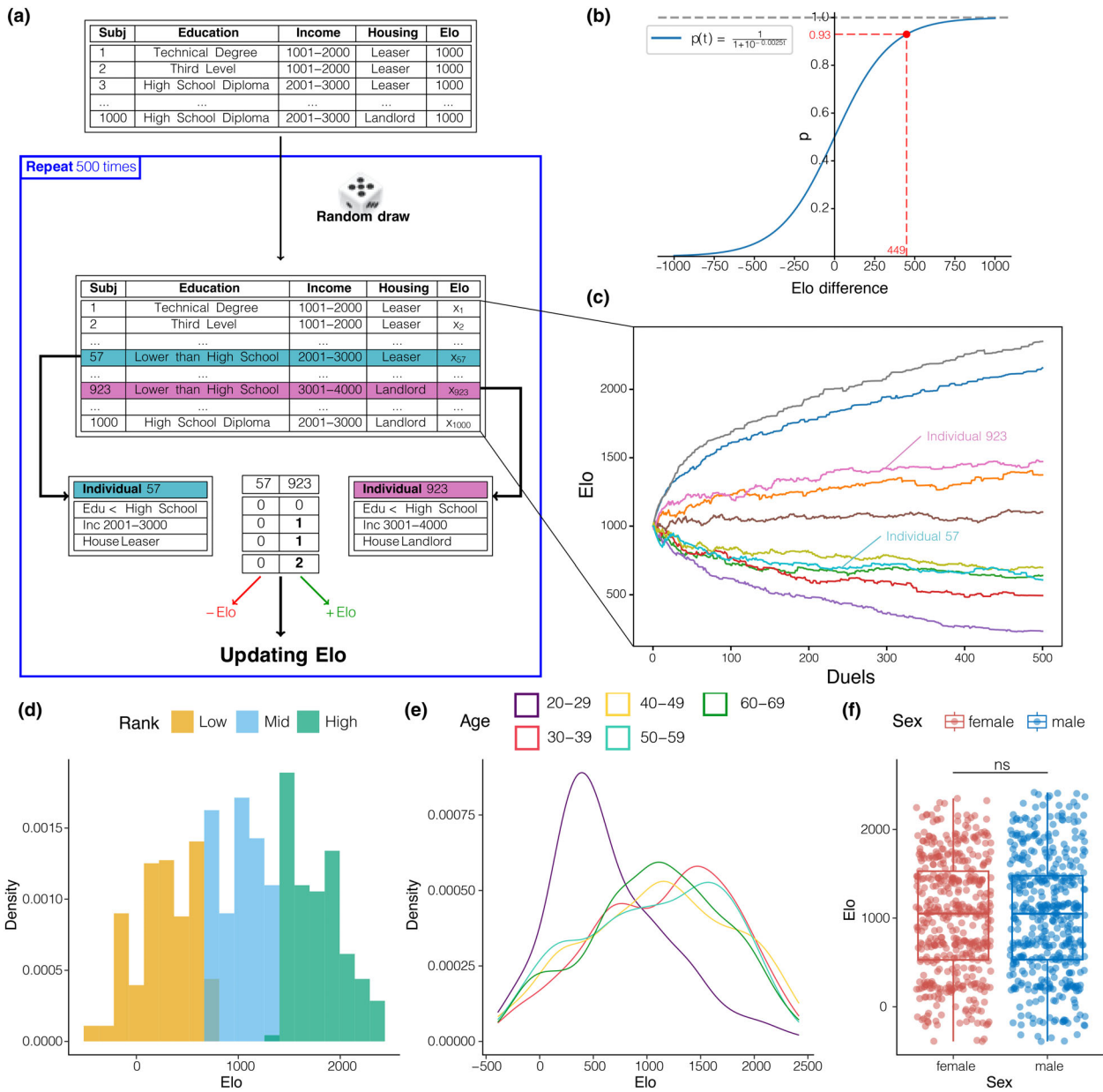


Figure 1. Measure of the socioeconomic status (SES) in the Milieu Interieur (MI) cohort by integration of educational attainment, home ownership status and income in an Elo rating system. **(a)** The process to compute Elo for the MI donors where two individuals are randomly selected to make a pairwise comparison based on three SES features. After the comparison, the individuals with the highest SES gain Elo, whereas the others lose Elo. An example of two individuals from the MI cohort (subject IDs 57 and 923) is shown. **(b)** Illustration of Elo difference calculated between two individuals. **(c)** Evolution of Elo for 10 randomly selected donors of the MI cohort across the 500 simulated duels. The evolution of the two individuals used for the example in **(b)** is displayed. **(d)** Distribution profile of the Elo of MI depending on SES rank. **(e)** Elo distribution grouped by age of MI donors. **(f)** Elo distribution according to sex. $N = 464$ females and 460 males.

Sex-specific SES associations

To assess the impact of the SES on immune responses, we first looked at the level of CRP. This is an acute-phase

protein secreted by the liver that is increased during inflammation or chronic infections. It has been shown that lower SES is associated with higher CRP levels, and therefore, a higher level of inflammation.¹⁰ However,

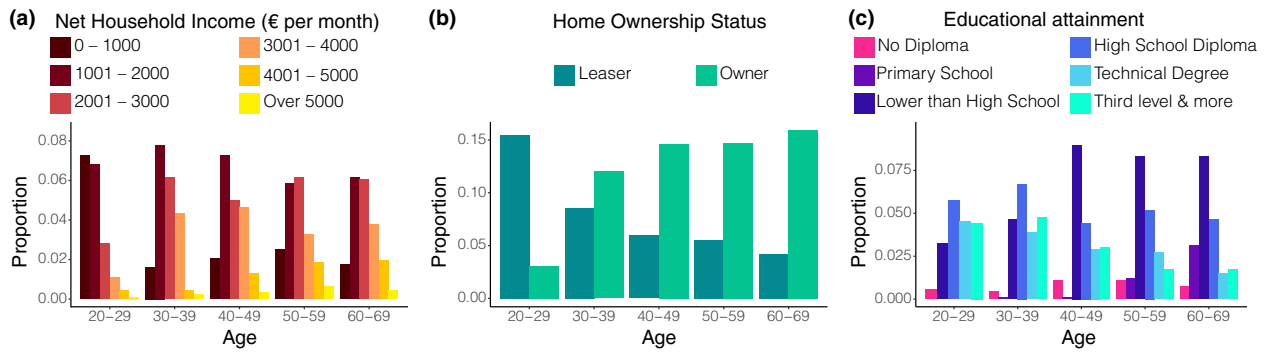


Figure 2. Socioeconomic feature distribution according to age within the *Milieu Intérieur* cohort. **(a)** Distribution of net family household income (in euros per month as reported in 2012). **(b)** Distribution of the home ownership status of the donors. **(c)** Distribution of educational attainment. $N=924$.

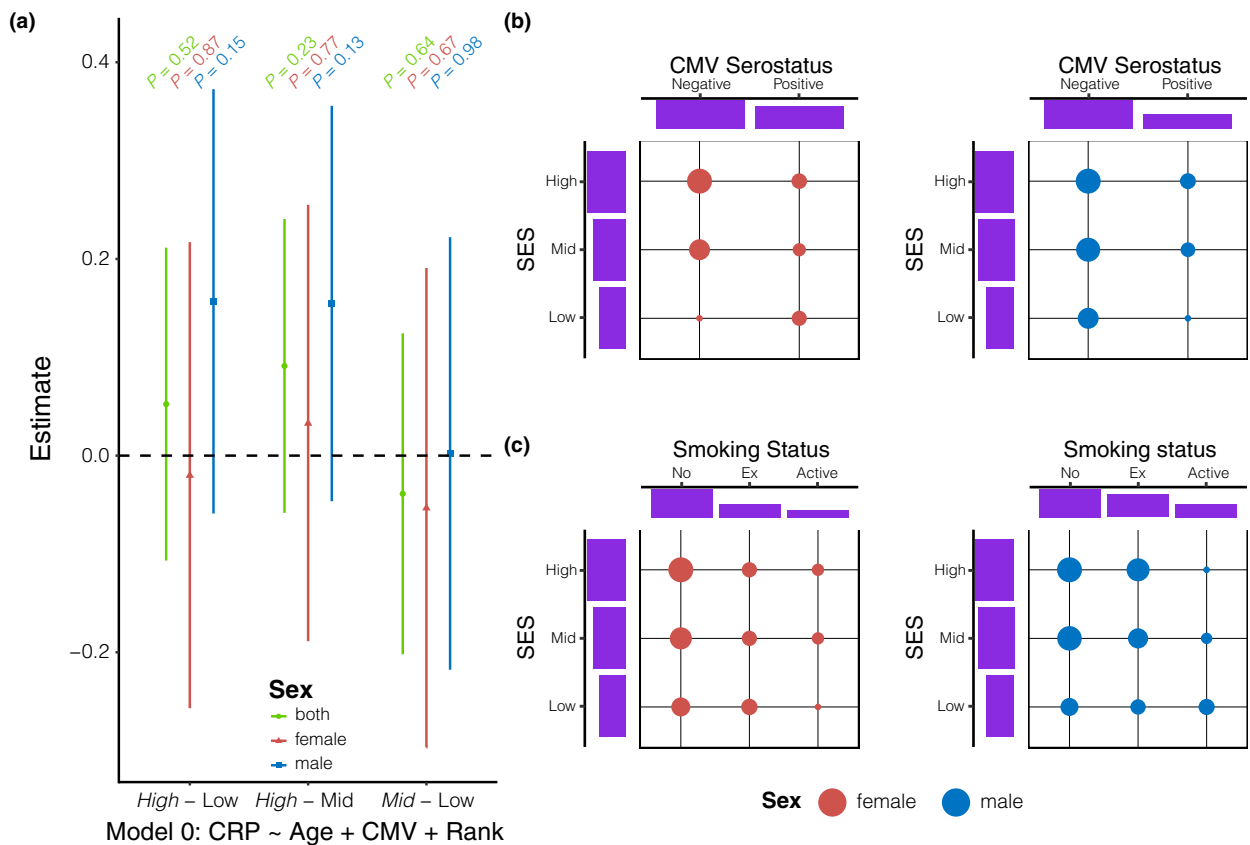


Figure 3. Sex-specific socioeconomic status (SES) associations with C-reactive protein (CRP), SES and smoking status. **(a)** Forest plot of the linear model results for the effect of SES on CRP level (SES rank used as the reference in the model is indicated in *italics*). **(b)** Comparisons of distribution between cytomegalovirus (CMV) serostatus and SES (female $P=0.026$, male $P=0.92$). **(c)** Comparisons of distribution between the smoking status of *Milieu Intérieur* donors and their SES (female $P=0.11$, male $P=0.014$). $N=376$ females and 376 males.

with the *Milieu Intérieur* cohort, such association was not observed (Figure 3a). This could be explained by the fact that this cohort had a strict inclusion/exclusion criteria.

Here the donors are healthy, and thus, the detection of differences in the level of an acute phase protein can become complex. We next looked at the frequency of

latent cytomegalovirus (CMV) infection for each SES category, given its well-described impact on immune variability.¹⁴ This analysis revealed a significantly ($P = 0.026$ Pearson's χ^2 contingency table test) higher CMV seropositive status for women of lower SES (Figure 3b). However, for men, there was no association ($P = 0.92$) between CMV seropositivity and SES (Figure 3b). Similar results were observed when adjusting for age in logistic regression models ($P = 5.1 \times 10^{-3}$ for females and $P = 0.44$ for males). We also assessed smoking status for associations with MI donors' SES. This choice was motivated by the fact that smoking status is known to influence several immune parameters in the cohort.^{15,16} We found a significant association between smoking status and SES of men ($P = 0.014$). Men with lower SES were more likely to be active smokers than those with medium or higher SES (Figure 3c). However, this association was not found for women (Figure 3c; $P = 0.11$). It is important to note that the distribution of smoking status is uneven because the majority of female donors are nonsmokers, whereas for men it is more balanced. However, these results show that SES can be associated with other environmental factors that also affect immune variability, and these associations are often sex specific.

Immune cell phenotypes and SES

To identify potential associations between immune cell phenotypes and SES, we performed linear regression models. In this analysis, among the 166 immune cell phenotypes studied, only one phenotype appeared to be significantly associated with SES, and in males only (Figure 4a, b). Specifically, the Mean Fluorescence Intensity (MFI) of CD38 in memory B cells in males was associated with the Elo variable ($\beta_{Elo} = 0.12$, $P = 0.003$; Figure 4b). However, no significant associations were found for females for both counts and MFI. These results are consistent with the differential sex effects of SES, but overall, they suggest that the SES effects on immune cell phenotypes in a healthy context are relatively limited.

Immune gene expression associated with SES

Given the lack of SES effects on baseline immune cell phenotypes, we next assessed whether SES may impact induced immune responses after standardized whole blood microbial stimulation with transcriptomic data generated using NanoString gene arrays.¹⁷ To test for potential SES effects, individual gene expression data were integrated into random forest models.¹⁸ The random forest models integrate a set of variables as described in the "Methods" section. These include cell

subset counts, which are the variables that we previously showed contribute the most to the variability of the gene expression profiles.¹⁷ Regardless of sex and the stimulation, cell count data have the highest percentage of increase in mean squared error (%IncMSE) compared with the other variables integrated into the models (Figure 5a, c, e). The only factor that showed a higher importance for explaining gene expression than cell types (CD8b⁺, B cells, monocytes and Effector Memory CD4⁺) was age in females for the influenza A virus stimulation condition (Figure 5e). After the major contribution of cell types, we found that age, CMV serostatus and CRP levels were associated with immune gene expression with associations for both males and females. These observations were consistent with previous observations because age and CMV are known to be associated with gene expression and cell composition.^{15,17}

Interestingly, SES as calculated by Elo is one of the variables most associated with gene expression with the same contribution as CRP levels or CMV serostatus in women in the null unstimulated condition. After *Escherichia coli* and influenza A virus stimulations, SES was also associated with gene expression profiles, although to a lesser extent. However, these associations were not observed for males where SES had a negative overall %IncMSE, indicating a lack of effect. These results suggested a sex-specific impact on immune responses for SES. To better understand this phenomenon, we also assessed the %IncMSE values on a gene-specific level. In the null unstimulated control condition, we found that 68 immune genes had a %IncMSE for Elo greater than 5% for women, whereas for men, we found only 16 immune genes (Figure 5b). These differences in proportions were significant through Pearson's χ^2 contingency table test ($\chi^2 P = 7.2 \times 10^{-9}$). Similarly, after *E. coli* stimulation, we found 46 immune genes for which the %IncMSE for Elo was greater than 5% for women, and 17 immune genes for men (Figure 5d). The associated χ^2 contingency table test showed statistically significant differences in these proportions between men and women ($P = 2.8 \times 10^{-4}$). Besides, upon influenza A virus stimulation, we found 36 immune genes with a %IncMSE for Elo greater than 5% for women and 20 immune genes for men ($P = 0.0041$; Figure 5f). These results suggest that specific immune genes are associated with the SES of the donors in the cohort. In addition, the difference in the proportion of genes with an importance greater than 5% between men and women suggests a sex specificity for SES effects on immunity.

DISCUSSION

In this study, we applied a systems immunology approach to integrate data from a cohort of 1000 healthy

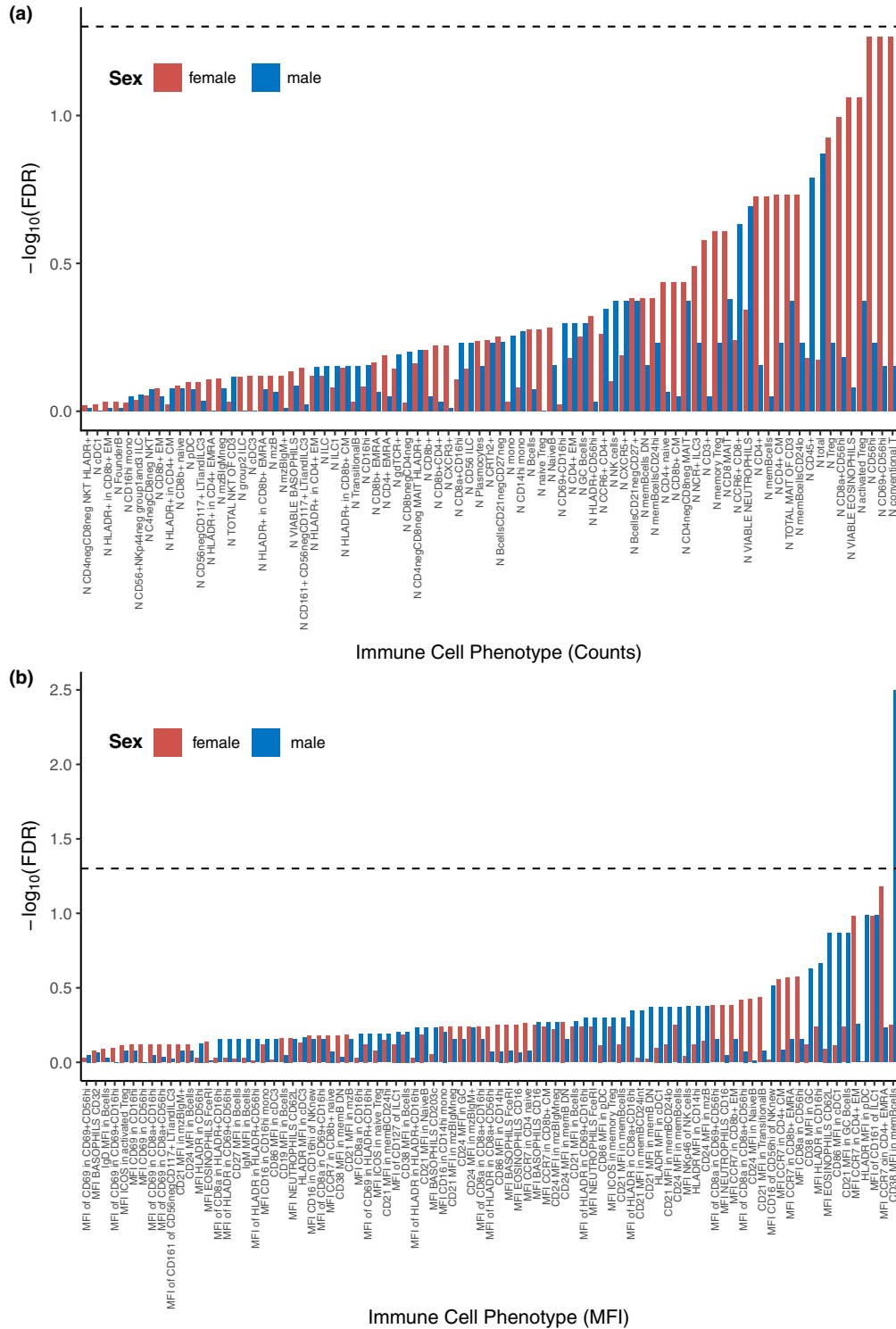


Figure 4. Immune cell phenotype associations with socioeconomic status (SES). Representation of the FDR-corrected P -values of the linear models for each of the 166 immune cell phenotypes tested. FDR-corrected P -values as **(a)** a function of the immune cell counts phenotypes and **(b)** a function of the immune cell mean fluorescence intensity phenotypes. FDR, false discovery rate.

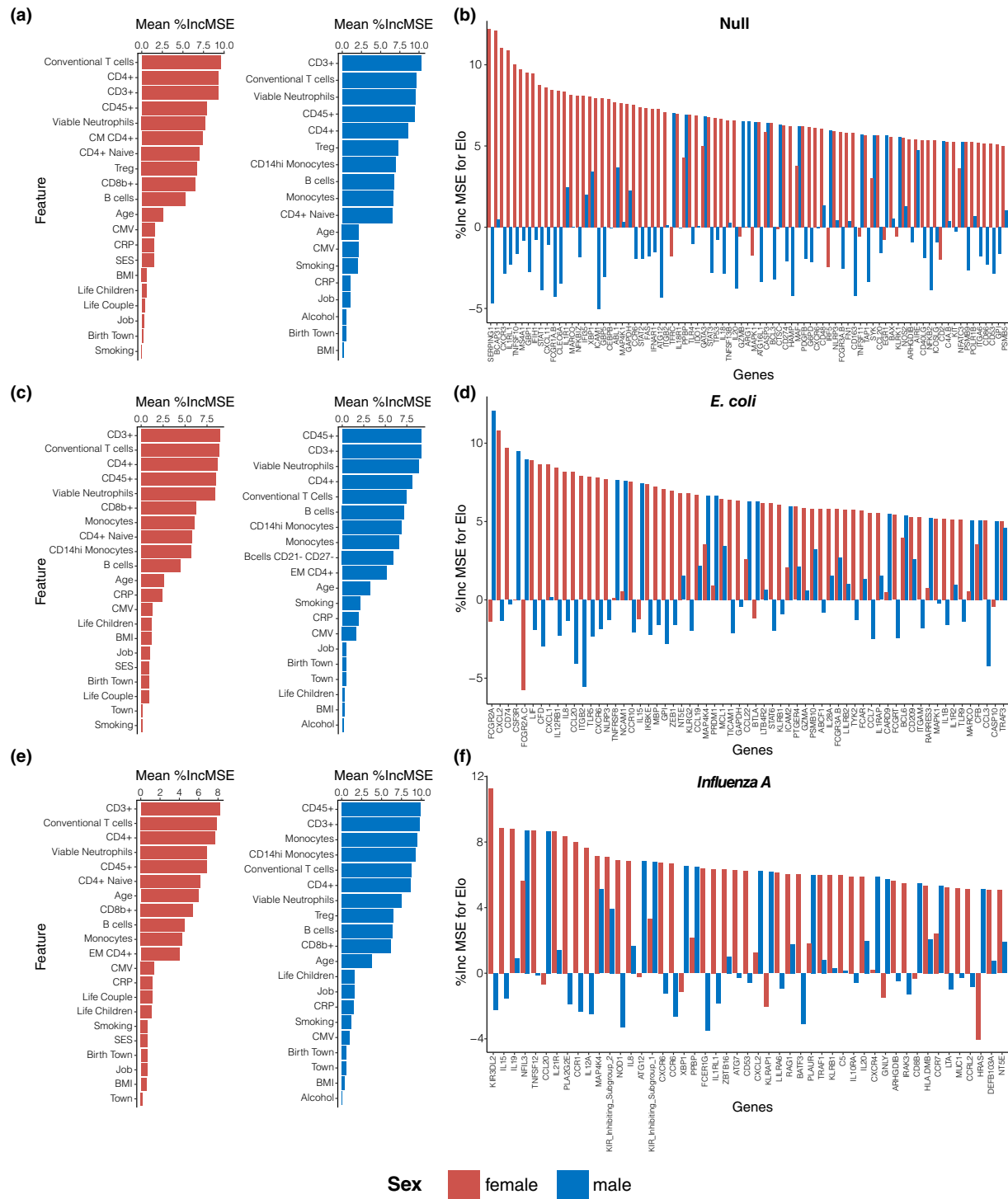


Figure 5. Sex stimuli specificity in immune response genes according to the socioeconomic status. **(a, d, e)** Overall importance (mean of % IncMSE over all the genes) of variables with a positive %IncMSE. **(a)** Null unstimulated, **(c)** *Escherichia coli* stimulation, **(e)** influenza A virus stimulation. **(b, d, f)** Importance of EIo in explaining the distribution of gene expression levels. Here, the genes for which EIo presented a % IncMSE > 5 for either men or women are shown. **(b)** Null unstimulated, **(d)** *E. coli* stimulation, **(f)** influenza A virus stimulation. Females are represented in red and males in blue. $N = 376$ females and 376 males. %IncMSE, percentage of increase in mean squared error. BMI, body mass index; CMV, cytomegalovirus; CRP, C-reactive protein; SES, socioeconomic status.

individuals to assess the potential impact of SES on immune response variability. Using a new way to compute SES in humans and integrating age, sex, CMV seropositivity and CRP levels with transcriptomic data on 560 immune genes after microbial stimulation, we identified a significant impact of SES on transcript levels after specific immune stimulations. The associations found between the genes and SES varied widely according to the sex of the donors. These results support a sex stimuli specificity in the associations between gene expression and SES. Most of the SES-associated genes identified in this study are not located on the sex chromosome. Among the stimulations, only four X chromosome genes were found to be associated with SES for females (*BCAP31*, *CD40LG*, *CXCR3* and *G6PD*). The underlying processes explaining such differences may be related to sex hormones. Indeed, it has been observed that in men, a reduced immune response was partially associated with a higher level of testosterone.¹⁹ Altogether, complex potential interactions between gender-associated social effects, sex hormones and immunity should be considered in future studies.

In line with this point, our Elo ranking system allowed us to identify significant associations between CMV seropositivity and SES in women, and between smoking status and SES in men. We found that active smokers of the MI cohort were associated with lower SES, as previously observed in other cohorts.^{20,21} The observations on CMV prevalence associations with sex and SES are also consistent with other studies identifying a higher rate of CMV infection in women and people with lower SES.²² This may reflect increased interactions between women and young children who are a likely source of transmission, especially in lower SES categories.²² As we did not observe the same gene expression differences in men, this suggests that sex-differential SES effects could be a result of either hormonal differences or other unidentified gender social effects. Moreover, associations between immune cell phenotypes and SES were only found for a very specific phenotype (MFI of CD38 in memory B cells) for males. Thus, this analysis revealed a very limited effect of SES on immune cell phenotypes in a healthy context, although the limited associations observed were sex specific.

Random forest identified several genes that are associated with SES in the *Milieu Interieur* cohort. In the null condition in females, several genes involved in the type I interferon antiviral immune response were identified as being impacted by SES, including *SERPING1*, *IFIH1*, *STAT1*, *IFI35* and *CXCL11*. These type I interferon pathway associations were not seen in males. There are well-documented sex differences in the type I interferon antiviral system which may explain the

disparity in SES–type I interferon associations between the males and females.^{23,24} In response to stimulation with the influenza A virus, many natural killer cell-related genes appear to be associated with SES in females, including *KIR3DL2* and *IL-15*. However, it should be noted that these random forest models do not provide any direction on the differences observed, thus limiting the biological interpretation of these results.

SES in humans is a complex phenomenon with many confounding variables, making it challenging to study. For example, while occupation or income is highly relevant to determining the current SES of an individual, it can also depend on previous social positions. In the same way, education and parents' occupation are also important SES features because of their strong impact on the future SES of an individual. Moreover, an indicator such as education will likely be stable after a certain age, whereas employment, home ownership status or income can vary throughout the life of an individual.^{25,26} Therefore, there is a need to combine these indicators to characterize SES at different stages of life. This variation in SES raises questions about the potential plasticity and reversibility of SES on the immune response. To answer these questions, longitudinal studies that incorporate such factors are required.

Our study contains certain limitations, for example, the *Milieu Interieur* cohort only consists of healthy donors, which may potentially create biases in the study of SES. For example, people with lower SES are more prone to illness and infection, and therefore, the selection of donors based on health status will reduce the range of SES within the cohort. However, this approach removes the potential confounding impact of illness, making it an ideal setting to assess moderate SES effects on immunity within a general population. This is supported by the novel SES immune associations identified in this study. Future studies will try to further uncover the underlying mechanisms with the integration of unbiased transcriptomic and epigenetic data sets. A deeper understanding of how low SES negatively impacts immunity could identify new approaches for improving the public health of neglected and vulnerable populations.

METHODS

Human samples

Human samples are from the *Milieu Interieur* cohort, which was approved by the Comité de Protection des Personnes-Ouest 6 on June 13, 2012, and by the French Agence Nationale de Sécurité du Médicament (ANSM) on June 22, 2012. The study is sponsored by Institut Pasteur (Pasteur ID-RCB Number: 2012-A00238-35) and was

conducted as a single-center interventional study without an investigational product. The original protocol was registered under [ClinicalTrials.gov](https://clinicaltrials.gov) (study# NCT01699893). The samples and data used in this study were formally established as the Milieu Intérieur biocollection (NCT03905993), with approvals by the Comité de Protection des Personnes – Sud Méditerranée and the Commission Nationale de l'Informatique et des Libertés (CNIL) on April 11, 2018. Donors gave written informed consent. The 1000 donors of the Milieu Intérieur cohort were recruited by Biotrial (Rennes, France) and were composed of “healthy” individuals of the same genetic background (Western European) and to have 100 women and 100 men from each decade of life, between 20 and 69 years of age. Donors were selected based on various inclusion and exclusion criteria that were previously described.¹³ Donors were required to have no history or evidence of severe/chronic/recurrent pathological conditions, neurological or psychiatric disorders, alcohol abuse, recent use of drugs, recent vaccine administration and recent use of immune-modulatory agents. To avoid the influence of hormonal fluctuations in women, pregnant and perimenopausal women were not included. The recruitment of donors was restricted to individuals whose parents and grandparents were born in Metropolitan France and who had no family relationships which minimizes genetic stratification.

Measuring socioeconomic status with Elo ranking

Using income, educational attainment and home ownership status information obtained from the electronic case report form, the SES of MI donors was measured with an Elo rating system.^{12,27} This choice was motivated by the ability of such a process to combine several SES features to establish an objective SES ranking within the cohort, instead of integrating these variables separately in models or computing an arbitrary SES score. To compute the Elo for each donor, pairwise comparisons were performed based on their educational attainment, income and home ownership status. The algorithm was encoded as follows. First, Elo is initialized at 1000 for each donor. Then, two individuals are randomly drawn without replacement until there is no longer any possibility of a match. Each pair of individuals undergoes pairwise comparisons based on the previously selected SES features. For each SES feature, the individual with the highest level gets a point and if the compared individuals have the same level, none of them get a point. After the comparisons, the individual with the most points “wins” the interaction, and therefore, gains Elo, whereas the other loses Elo. In case both individuals of the pair have the same number of points, no Elo is gained or lost. For example, in Figure 1a, the pairwise comparison shows that donor 923 has the same level of education but a higher level of income and home ownership status. This individual has 2 points, and therefore, wins the interactions and gains Elo. This whole process is repeated 500 times (Figure 1a). The gain and loss of Elo are weighted by the expected outcome of the match. The win probability function is given as follows:

$$P : t \in \mathbb{R} \mapsto \frac{1}{1 + 10^{\frac{t}{400}}} \in [0, 1] \quad (1)$$

where t is the Elo difference between the two contending individuals (Figure 1b).

Let P be the win probability of Individual_{*i*} over Individual_{*j*}, then their corresponding Elo ratings are updated as follows:

$$\text{If Individual}_i \text{ won over Individual}_j \begin{cases} \text{Elo}_i = \text{Elo}_i + (1-P)k \\ \text{Elo}_j = \text{Elo}_j - (1-P)k \end{cases} \quad (2)$$

$$\text{If Individual}_j \text{ won over Individual}_i \begin{cases} \text{Elo}_i = \text{Elo}_i - Pk \\ \text{Elo}_j = \text{Elo}_j + Pk \end{cases} \quad (3)$$

In these formulas, k is a constant adjusting the Elo winning rate. Here, we fixed $k=50$ such that in the case of Elo equality between individuals before the interaction, the gain/loss of Elo will be ± 25 . For example, if the interaction between individual 923 and 57 occurs at the 100th duel, at this step, SUBJ0923 has an Elo of 1209.4081 and SUBJ0057 has 760.1367. The difference of Elo between these two individuals equals approximately 449 (Figure 1b). Such a difference corresponds to a win probability of 0.93. Therefore, after the interaction, and according to the formulas [Equations 2 and 3], individual 923 would gain 46.5 of Elo and individual 57 would lose the same amount. The use of weighting in the calculation of the Elo allows an adaptive evolution of the Elo and decreases the slope of the Elo evolution over the course of the duels (Figure 1c). After each individual undergoes 500 interactions, a stable ranking of the Elo rating system is obtained: individuals from a higher SES have a higher Elo and vice versa. This stable ranking allows the cohort to be divided into three balanced SES categories that will be used later for Pearson's contingency χ^2 test (Figure 1d).

Pearson's χ^2 contingency table test

Pearson's χ^2 contingency table tests were performed on MI donors to determine whether there was a dependence between their CMV serostatus and SES. The contingency table contained information about the number of CMV seropositive individuals from low, mid and high SES and the number of CMV seronegative individuals from low, mid and high SES. The same approach was performed to test the dependence between the MI donors' smoking status (as self-reported by the donors in the electronic case report form) and their SES. Similarly, the contingency table included the number of nonsmokers, ex-smokers and smokers for each of the SES ranks. The Pearson's χ^2 contingency table test was also used to assess the number of immune genes associated (5% increase in the MSE threshold) or not with socioeconomic factors depending on sex. The contingency tables report the number of genes in each subcategory and Pearson's chi-squared tests report the significance of the difference between the proportions.

Logistic regression

In addition to Pearson's χ^2 contingency table tests, logistic regression models were used to compare the CMV seropositivity status and SES of the donors while correcting for age [Equation 4]. This choice was motivated by the fact that a variable such as seropositivity to a virus is age dependent. The older a person is, the greater the risk of exposure to a virus.

$$\log \frac{P(\text{CMV}_+|\text{Elo and Age})}{1-P(\text{CMV}_+|\text{Elo and Age})} = \beta_0 + \beta_{\text{Elo}}\text{Elo} + \beta_{\text{Age}}\text{Age} \quad (4)$$

Linear models

Linear models were performed to assess the dependency between CRP level and SES. The SES variables, as measured by Elo, were used in the categorical forms (low, mid and high). High SES was set as the reference for the model and was compared with both mid and low SES. In the first model, age and CMV serostatus are used as covariates alongside with the SES rank. Age is used as a numeric variable and CMV serostatus is a categorical variable, where CMV negative is used as a reference [Equation 5]. Other models were performed using, respectively, smoking status, body mass index and alcohol consumption as covariates in addition to the variables used in the first model. For smoking status, nonsmoker is taken as a reference for the models, body mass index is a numeric variable and for alcohol consumption frequency, never is taken as a reference.

$$\log(\text{CRP}) = \beta_0 + \beta_{\text{Age}}\text{Age} + \beta_{\text{CMV}_{\text{pos}}}\text{CMV}_{\text{pos}} + \beta_{\text{SES}_{\text{low}}}\text{SES}_{\text{low}} + \beta_{\text{SES}_{\text{mid}}}\text{SES}_{\text{mid}} \quad (5)$$

A similar type of model has been used to test eventual associations between SES and immune cell phenotypes. A total of 166 immune cell phenotypes were fitted with a linear regression model to SES, adjusting for age and CMV seropositivity [Equation 6].

$$\log(\text{Immune Cell Phenotype}) = \beta_0 + \beta_{\text{Age}}\text{Age} + \beta_{\text{CMV}_{\text{pos}}}\text{CMV}_{\text{pos}} + \beta_{\text{Elo}}\text{Elo} \quad (6)$$

Random Forest models

Random forest models¹⁸ were used to select features and to assess their importance regarding immune gene expression. This can be measured using the %IncMSE for a given feature.²⁸ The higher the value, the more important the feature is for explaining the distribution of the model's target variable. As applied here, we tested the importance of the SES (as measured by Elo) of *Milieu Intérieur* donors to explain immune gene expression levels. We used the random forest approach in two steps. The first one determined which immune cell types most influence the gene expression profiles.

To test for sex stimuli-specific effects, each random forest model was performed separately according to sex and for each of the nine immune stimulations. For each sex-stimulation pair, each immune gene transcript level was fitted on the count values of 75 blood cell subsets, and the 10 cell types with the highest mean of increased mean square error (%IncMSE) over all the genes were kept for the second step. In the second step, transcriptomic data were integrated with socioeconomic features (Elo, job status, parental status, marital status, place of residence and birth town), smoking status, alcohol consumption, age and sex, and the models were adjusted with biological variables (CMV serostatus, CRP level and the most contributing cell types). Then, for each sex-stimulation pair, the expression level of each gene was fitted to the variables of interest highlighted in Figure 5. The hyperparameters of the models were set according to the R documentation for the number of variables randomly sampled at each split (i.e. 25 for the first step and 7 for the second step) using the *mtry* parameter²⁹ and the number of trees was set at 1000 for both steps. Random forest models were chosen to test and catch potential nonlinear relationships between variables because it is known that upon some stimulations, genes have a nonlinear association with other variables (e.g. the effect of age on influenza A virus¹⁷).

Code availability

All the scripts related to this study were written in R/4.2.1 and Python/3.8 and have been deposited on GitLab: <https://gitlab.pasteur.fr/abertran/SESImmune>.

Immunological data sets

Flow cytometry¹⁵ and NanoString (Nanostring, Seattle, WA, USA) gene expression¹⁷ data sets have been previously described. For whole blood immune stimulation data sets (TruCulture, Rules Based Medicine, Austin, TX, USA), we focused on the influenza virus (H1N1) as a model to study viral immunity and *E. coli* to study bacterial immunity, in addition to the Null unstimulated control condition.

ACKNOWLEDGMENTS

This program is managed by the Agence Nationale de la Recherche. This work benefited from the support of the French government's Invest in the Future program (reference ANR-10-LABX-69-01). AB is a student from the FIRE PhD program funded by the Bettencourt Schueller Foundation and the EURIP Graduate Program (ANR-17-EURE-0012). We thank Cliona Ní Cheallaigh for initially inspiring us to look at the effects of SES on viral infection.

The Milieu Intérieur Consortium is composed of the following team leaders (unless otherwise indicated, partners are located at Institut Pasteur, Paris): Laurent Abel (Hôpital Necker), Andres Alcover, Hugues Aschard, Philippe Bouso, Nollaig Bourke (Trinity College Dublin), Petter Brodin (Karolinska Institutet), Pierre Bruhns, Nadine Cerf-Bensussan

(INSERM UMR 1163 – Institut Imagine), Ana Cumano, Christophe D'Enfert, Caroline Demangel, Ludovic Deriano, Marie-Agnès Dillies, James Di Santo, Gérard Eberl, Jost Enninga, Jacques Fellay (EPFL, Lausanne), Ivo Gomperts-Boneca, Milena Hasan, Gunilla Karlsson Hedestam (Karolinska Institutet), Serge Herberg (Université Paris 13), Molly A Ingersoll (Institut Cochin and Institut Pasteur), Olivier Lantz (Institut Curie), Rose Anne Kenny (Trinity College Dublin), Mickaël Ménager (INSERM UMR 1163 – Institut Imagine) Hugo Mouquet, Cliona O'Farrelly (Trinity College Dublin), Etienne Patin, Antonio Rausell (INSERM UMR 1163 – Institut Imagine), Frédéric Rieux-Laucat (INSERM UMR 1163 – Institut Imagine), Lars Rogge, Magnus Fontes (Institut Roche), Anavaj Sakuntabhai, Olivier Schwartz, Benno Schwikowski, Spencer Shorte, Frédéric Tangy, Antoine Toubert (Hôpital Saint-Louis), Mathilde Touvier (Université Paris 13), Marie-Noëlle Ungeheuer, Christophe Zimmer, Matthew L Albert (Octant Biosciences), Darragh Duffy and Lluís Quintana-Murci are co-coordinators of the Milieu Intérieur Consortium. Additional information can be found at <https://www.milieuinterieur.fr/en/>.

AUTHOR CONTRIBUTIONS

Anthony Bertrand: Data curation; formal analysis; investigation; methodology; visualization; writing – original draft; writing – review and editing. **Jamie Sugrue:** Formal analysis; supervision; validation; writing – review and editing. **Tianai Lou:** Formal analysis. **Nollaig M Bourke:** Investigation; supervision; validation; writing – review and editing. **Lluís Quintana-Murci:** Funding acquisition; project administration; supervision. **Violaine Saint-André:** Conceptualization; data curation; formal analysis; investigation; methodology; supervision; writing – original draft; writing – review and editing. **Cliona O'Farrelly:** Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; supervision; validation; writing – original draft; writing – review and editing. **Darragh Duffy:** Conceptualization; funding acquisition; investigation; methodology; project administration; supervision; validation; writing – original draft; writing – review and editing.

CONFLICT OF INTEREST

All authors declare no conflict of interest for this study.

DATA AVAILABILITY

Data are available upon request at <https://dataset.owey.io/> (Institut Pasteur data repository).

REFERENCES

- Psaki SR, Seidman J, Miller M, et al. Measuring socioeconomic status in multicountry studies: results from the eight-country MAL-ED study. *Popul Health Metr* 2014; **12**: 8.
- Mackenbach JP, Stirbu I, Roskam AJ, et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008; **358**: 2468–2481.
- Marmot MG, Shiple MJ, Rose G. Inequalities in death—specific explanations of a general pattern? *Lancet* 1984; **323**: 1003–1006.
- Reiche EM, Nunes SO, Morimoto HK. Stress, depression, the immune system, and cancer. *Lancet Oncol* 2004; **5**: 617–625.
- Dhabhar FS. Effects of stress on immune function: the good, the bad, and the beautiful. *Immunol Res* 2014; **58**: 193–210.
- Glaser R, Kiecolt-Glaser JK. Stress-induced immune dysfunction: implications for health. *Nat Rev Immunol* 2005; **5**: 243–251.
- Zefferino R, Di Gioia S, Conese M. Molecular links between endocrine, nervous and immune system during chronic stress. *Brain and Behavior* 2021; **11**: e01960.
- Cole SW. The conserved transcriptional response to adversity. *Curr Opin Behav Sci* 2019; **28**: 31–37.
- Chiang JJ, Cole SW, Bower JE, et al. Depressive symptoms and immune transcriptional profiles in late adolescents. *Brain Behav Immun* 2019; **80**: 163–169.
- Berger E, Castagné R, Chadeau-Hyam M, et al. Multi-cohort study identifies social determinants of systemic inflammation over the life course. *Nat Commun* 2019; **10**: 773.
- Muscattell KA, Brosso SN, Humphreys KL. Socioeconomic status and inflammation: a meta-analysis. *Mol Psychiatry* 2020; **25**: 2189–2199.
- Snyder-Mackler N, Sanz J, Kohn JN, et al. Social status alters immune regulation and response to infection in macaques. *Science* 2016; **354**: 1041–1045.
- Thomas S, Rouilly V, Patin E, et al. The milieu Intérieur study — an integrative approach for study of human immunological variance. *Clinic Immunol* 2015; **157**: 277–293.
- Kaczorowski KJ, Karthik S, Nkulikiyimfura D, et al. Continuous immunotypes describe human immune variation and predict diverse responses. *Proc Natl Acad Sci U S A* 2017; **114**: e6097–e6106.
- Patin E, Hasan M, Bergstedt J, et al. Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. *Nat Immunol* 2018; **19**: 302–314.
- Saint-André V, Chabrit B, Biton A, et al. Smoking changes adaptive immunity with persistent effects. *Nature* 2024; **626**: 827–835.
- Piasecka B, Duffy D, Urrutia A, et al. Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proc Natl Acad Sci U S A* 2018; **115**: e488–e497.
- Breiman L. Random forests. *Machine Learning* 2001; **45**: 5–32.
- Gassen J, White JD, Peterman JL, et al. Sex differences in the impact of childhood socioeconomic status on immune function. *Sci Rep* 2021; **11**: 9827.
- Hiscock R, Bauld L, Amos A, Platt S. Smoking and socioeconomic status in England: the rise of the never smoker and the disadvantaged smoker. *J Public Health (Oxf)* 2012; **34**: 390–396.

21. Leventhal AM, Bello MS, Galstyan E, Higgins ST, Barrington-Trimis JL. Association of Cumulative Socioeconomic and Health-Related Disadvantage with Disparities in smoking prevalence in the United States, 2008 to 2017. *JAMA Intern Med* 2019; **179**: 777–785.
22. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Rev Med Virol* 2010; **20**: 202–213.
23. Hagen SH, Henseling F, Hennesen J, *et al.* Heterogeneous escape from X chromosome inactivation results in sex differences in type I IFN responses at the single human pDC level. *Cell Rep* 2020; **33**: 108485.
24. Souyris M, Cenac C, Azar P, *et al.* TLR7 escapes X chromosome inactivation in immune cells. *Sci Immunol* 2018; **3**: eaap8855.
25. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey SG. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006; **60**: 7–12.
26. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey SG. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health* 2006; **60**: 95–101.
27. Neumann C, Duboscq J, Dubuc C, *et al.* Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. *Anim Behav* 2011; **82**: 911–921.
28. James G, Witten D, Hastie T, Tibshirani R. Tree-based methods. An Introduction to Statistical Learning. 2nd ed. New York, NY: Springer Texts in Statistics, Springer; 2021:327–361.
29. Liaw A, Wiener M. Classification and regression by randomForest. *R News* 2002; **2**: 18–22.

© 2024 The Author(s). Immunology & Cell Biology published by John Wiley & Sons Australia, Ltd on behalf of the Australian and New Zealand Society for Immunology, Inc.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.