Supporting Information

Piasecka et al. 10.1073/pnas.1714765115

SI Materials and Methods

Members of the Milieu Intérieur Consortium. The Milieu Intérieur Consortium* is composed of the following team leaders: Laurent Abel (Hôpital Necker), Andres Alcover, Hugues Aschard, Kalle Aström (Lund University), Philippe Bousso, Pierre Bruhns, Ana Cumano, Darragh Duffy, Caroline Demangel, Ludovic Deriano, James Di Santo, Françoise Dromer, Gérard Eberl, Jost Enninga, Jacques Fellay (EPFL, Lausanne), Magnus Fontes, Antonio Freitas, Odile Gelpi, Ivo Gomperts-Boneca, Serge Hercberg (Université Paris 13), Olivier Lantz (Institut Curie), Claude Leclerc, Hugo Mouquet, Etienne Patin, Sandra Pellegrini, Stanislas Pol (Hôpital Côchin), Antonio Rausell (INSERM UMR 1163 - Institut Imagine), Lars Rogge, Anavaj Sakuntabhai, Olivier Schwartz, Benno Schwikowski, Spencer Shorte, Vassili Soumelis (Institut Curie), Frédéric Tangy, Eric Tartour (Hôpital Européen George Pompidou), Antoine Toubert (Hôpital Saint-Louis), Marie-Noëlle Ungeheuer, Lluis Quintana-Murci**, Matthew L. Albert**.

The Milieu Intérieur Cohort. The Milieu Intérieur Project (https://clinicaltrials.gov; identifier: NCT01699893) includes 1,000 healthy donors (500 men and 500 women) aged 20–69 y old equally distributed across five decades of life (200 individuals per decade). Donors were selected on the basis of stringent inclusion and exclusion criteria as previously described (1). Donors could be included in this cohort only if they had no signs or history of neurological or psychiatric disorders or severe/chronic/recurrent pathological conditions. Other exclusion criteria included seropositivity for major chronic viral infections (HIV, hepatitis B virus, hepatitis C virus), abnormal laboratory test results, history or evidence of alcohol abuse, recent use of illicit drugs, and recent vaccine administration. We eliminated any effect of hormonal fluctuations in women during the perimenopausal phase by including only pre- or postmenopausal women. To avoid population stratification in the cohort, we recruited only individuals of western European descent (i.e., French citizens for whom the last three generations of ancestors were from mainland France).

Whole-Blood TruCulture Stimulation. TruCulture tubes were prepared in two batches (A and B) with the stimulus indicated: heat-killed *Escherichia coli* (O111:B4), live BCG, *Staphylococcus aureus*, enterotoxin SEB (Bernhard Nocht Institute), *Candida albicans* (Invivogen), or live stocks of attenuated H1N1 IAV (Charles River). We also prepared a nonstimulated control. Each tube contained 2 mL buffered media and was maintained at −20 °C until use. We performed stimulation experiments with 1 mL whole blood for 22 h as previously described (2). The time point of 22 h was chosen based on a detailed kinetic analysis that was performed for selected stimuli, the results of which were recently described (3). Because of the logistics of recruiting 1,000 donors in a highly standardized manner, a single time point was selected that was in the maximal plateau for the majority of immune responses examined.

Gene Expression Analysis. RNA samples were processed as previously described (4). Briefly, we used a specific chloroform-free one-step protocol based on a modified version of the NucleoSpin 96 RNA tissue kit protocol (Macherey-Nagel) and adapted for use with the Freedom EVO integrated vacuum system. RNA concentration was estimated with the Qubit RNA HS Assay Kit (Life Technologies), and RNA integrity was assessed with the Standard RNA Reagent Kit on a LabChip GX (Perkin-Elmer). The RNA Quality Score (RQS) was calculated with LabChip System software, and all samples with an RQS > 4 were processed for gene expression analysis. The NanoString nCounter system, a hybridization-based multiplex assay, was used for the digital counting of transcripts. We hybridized 100 ng of total RNA from each sample according to the manufacturer's instructions with the Human Immunology v2 Gene Expression CodeSet, which contains 594 endogenous gene probes, 8 negative control probes (NEG A to NEG H), and 6 positive control probes (POS A to POS F) designed against six in vitro-transcribed RNA targets premixed with the CodeSet at a range of concentrations (from 128 to 0.125 fM). We used three batches of the nCounter Legacy formulation to measure the gene expression induced by *E. coli*, *S. aureus*, and BCG and that in the nonstimulated control. The expression induced by *C. albicans*, IAV, and SEB was measured with a single batch of the nCounter XT formulation. We accounted for the slight differences in performance between these two formulations in the normalization procedure.

Mapping NanoString Probes to the Human Genome. NanoString probes were derived from cDNA sequences and often spanned multiple exon/intron junctions. We, therefore, mapped them against the human genomic sequence (GRCh37/hg19) with Genomic Short-Read Nucleotide Alignment Program (GSNAP) (5), a splice-aware aligner. We ran GSNAP with two flags to detect splice junctions: -N for the detection of novel splice sites and -S for the detection of known splice sites. We found that 573 of 594 probes mapped onto the genome with 100% identity. Twelve probes mapped with one to two mismatches in the middle of the sequence, eight probes were misaligned in the first/last 1–9 bp, and one probe did not map at all (PECAM1 located on HG183 PATCH). The misaligned probes were realigned manually with BLASTN against the Ab-initio cDNA database (grch37.ensembl.org/Homo_sapiens/Tools/Blast). Only the KIR_Activating_Subgroup_1 probe remained unaligned over its first 9 bp and was removed from the analysis; 15 of 594 NanoString probes mapped to more than one genomic location, and 4 of these probes mapped to different chromosomes (CCRL1, ITGB1, EEF1G, and TUBB) and were removed from the analysis. We then used Bioconductor biomaRt package (6), version 2.24.0, to query Ensembl (release 75) and retrieve exonic variants mapping to the same regions as the NanoString probes. For each of these SNPs, we retrieved the chromosomal position, strand, possible alleles, and minor allele frequency (MAF) for all individuals from the 1,000 Genomes Project (1000G) Phase 1. We then queried the 1000GENOMES-phase_1_EUR.vcf file (downloaded from ftp://ftp.ensembl.org/pub/release-75/variation/vcf/homo_sapiens/) to retrieve the MAFs of the corresponding SNPs in Europeans from the 1000G cohort (1000G MAFs were highly consistent with MAFs of the Millieu Intérieur cohort). In total, 47 probes had at least one common SNP (i.e., MAF ≥ 0.05) in their sequence. Five probes with at least three SNPs (HLA-DQB1, HLA-DQA1, HLA-DRB1, HLA-B, and C8G with 12, 9, 7, 3, and 3 SNPs, respectively) were removed from subsequent analyses. Finally, we identified the genomic location of the genes covered by the NanoString probes. We used the biomaRt package to query Ensembl (release 75) and to retrieve the start/end positions and HUGO Gene Nomenclature Committee symbols of the genes in the regions covered by NanoString probes. If genes were present on both strands, only the strand consistent with GSNAP mapping was considered. Full annotation of the NanoString probes is provided in Dataset S25.

*Unless otherwise indicated, partners are located at Institut Pasteur, Paris.

**Joint coordinators of the Milieu Intérieur Consortium. Additional information can be found at: www.milieuinterieur.fr/en.

Quality Control of the NanoString Data. Quality control for our data involved checking the following metrics: fields of view counted (flag if <0.75), binding density (flag if not in the 0.05–2.75 range), linearity of positive controls (flag if $R^2 < 0.9$), and limit of detection for positive controls (flag if 0.5 fM positive control <2 SDs above the mean of the negative controls). For calculations of the mean and SD of negative controls, we excluded two probes (NEG B and NEG H) that varied considerably between conditions, probably because of cross-reaction with bacterial nucleic acids present in the TruCulture stimulation systems. In addition to the quality control metrics proposed by the NanoString, we also determined total counts (flag if not in the 100,000–1,910,000 range). Three samples were removed because of very low total counts (<100,000).

Calibration Between Chemistries and Positive Control Normalization. Three different batches of Legacy chemistry reagents were used to assess the expression induced by *E. coli*, *S. aureus*, and BCG and that in the nonstimulated control, whereas a single batch of XT chemistry reagents was used to assess the expression induced by *C. albicans*, IAV, and SEB. However, for 25 samples, expression levels were measured with both Legacy (lot 1) and XT formulations in all seven conditions, making it possible to calculate calibration factors for the samples analyzed with XT chemistry.

We first performed positive control normalization on 25 × 7 samples analyzed with Legacy chemistry. NanoString provides six positive control probes to make it possible to control for differences between experimental variables (e.g., hybridization, purification, or binding efficiency). Counts of positive control probes are expected to vary between samples but independently of the stimulus. However, for all positive control probes, counts were higher in the absence of stimulation than in stimulated conditions (Fig. S8). This may reflect higher levels of expression for the targeted genes in the presence than in the absence of stimulation, leading to stronger competition between the positive control probes and gene probes for binding to the surface of the slide. The differences in positive control probe binding between stimulated and nonstimulated conditions were specific to the control RNA concentration [i.e., the difference in binding was largest for probe POS A (128) and smallest for probe POS F (0.125)].

The normalization method recommended by NanoString does not correct for this aspect. We, therefore, developed an alternative method. We log2-transformed the probe counts of the 25 × 7 samples assessed with the Legacy chemistry and then normalized them as follows: for five positive control probes (POS A to POS E; POS F was excluded, as the signal was too low), we calculated the median counts in the 25 nonstimulated samples. This resulted in five reference values for the Legacy chemistry. Then, for each sample, we fitted a linear regression model with the reference values as dependent variables and the five positive control counts as independent variables. This defined the mapping from original to normalized counts. Finally, we projected each probe count (both control and gene probes) onto the regression line and obtained the corresponding normalized values.

We performed an independent analogous normalization for 25 × 7 samples analyzed with XT chemistry. The five reference values in this case were calculated from the probe counts obtained with XT chemistry. Finally, we combined the normalized samples to obtain a set of 25 × 7 × 2 (350) samples. We calculated the calibration factors for XT chemistry by fitting the following linear model for each gene separately:

$$Y^i = {\alpha_0}^i + {\alpha_1}^i \cdot \mathrm{CHEM} + \sum {\alpha_j}^i \cdot \mathrm{STIM_j},$$

where Y^i, CHEM, and $\mathrm{STIM_j}$ are 350-element vectors, in which each element corresponds to the value of the variable for one sample measured with one type of chemistry under one condition. Y^i elements are the log2-transformed probe counts (after positive control normalization) for gene i; CHEM elements are equal to one if XT chemistry was used to measure gene expression and to zero otherwise. $\mathrm{STIM_j}$ elements are equal to one if the jth stimulus was used to induce the gene expression and to zero otherwise (where j = {nonstimulated control, *E. coli*, BCG, *S. aureus*, SEB, *C. albicans*, IAV}); ${\alpha_1}^i$ is the calibration factor for gene i used to account for differences resulting from the use of XT chemistry.

Normalization Procedure for the nCounter Legacy and XT Formulations. We performed positive control normalization for all samples analyzed with Legacy chemistry using the procedure described above (we used five reference values calculated for Legacy chemistry). In addition, we verified whether code set input was too low for any of the samples (e.g., because of evaporation) by, for each sample in a given condition, (*i*) calculating the mean count of all gene probes and determining whether this mean count was more than 2 SDs below the average of all means within the condition and (*ii*) calculating the mean count of the positive control probes (${\mu_{\mathrm{Pos}}}^i$), and determining whether the difference between the average of all means [$(\Sigma\ {\mu_{\mathrm{Pos}}}^i)/n$] within the condition and the sample mean (${\mu_{\mathrm{Pos}}}^i$) was bigger than one. We excluded two samples that simultaneously fulfilled both conditions *i* and *ii*.

Finally, we corrected for differences in RNA sample input by calculating the mean of all gene probe counts for each sample. A scaling factor for each sample was calculated as the difference between the average across the means within the conditions to which the sample belonged and the mean of the sample. For each sample, we added its corresponding scaling factor to each gene count. This approach was based on the assumption that total count is the same for samples within one condition but potentially different between conditions. Eleven samples with a very low RNA input (i.e., with a scaling factor greater than two) were excluded from the analysis. We retained 986 donors for whom none of the four samples were excluded in the quality control and normalization steps. We then adjusted counts for each gene for the effect of the CodeSet batch used to measure the level of expression. There were three batches, and we designated batch 1 as the reference batch. We estimated the effects of batches 2 and 3 by fitting the following linear model for each gene separately:

$$Y^i = {\beta_0}^i + {\beta_2}^i \cdot \mathrm{Lot2} + {\beta_3}^i \cdot \mathrm{Lot3} + \sum {\beta_j}^i \cdot \mathrm{STIM}_j,$$

where Y^i, Lot2, Lot3, and $\mathrm{STIM_j}$ are (986 × 4)-element vectors, in which each element corresponds to the value of the variable for one donor measured in one condition. Y^i elements are the normalized probe counts (log2 scale) for gene i. Lot2 elements are equal to one if batch 2 was used to measure gene expression and zero otherwise. Lot3 elements are equal to one if batch 3 was used and zero otherwise. $\mathrm{STIM_j}$ elements are equal to one if the jth stimulus was used to induce gene expression and zero otherwise (where j = {nonstimulated control, *E. coli*, BCG, *S. aureus*}). For each sample run with batch 2 (or batch 3), we subtracted β_2 (β_3 for batch 3) from the count for gene i. This procedure rendered gene expression levels comparable between the three batches.

For normalization of the samples analyzed with the nCounter XT formulation, we performed positive control normalization as described above (using the five reference values calculated for XT chemistry). We also corrected for RNA sample input differences and determined, as described above, whether any of the samples had too low CodeSet or RNA input (no samples were filtered out). Finally, we subtracted the corresponding chemistry calibration factor (${\alpha_1}^i$) from the level of expression for each gene.

Gene Probe Filtering. We estimated the background level for each sample as the mean plus 2 SDs of the six negative probe counts, excluding NEG B and NEG H, for which we observed significant differences in counts between conditions as previously explained (4). Rather than subtracting the background from gene counts, we flagged the genes as present or absent if expression levels were above or below the background level, respectively. We removed 24 gene probes (C1R, C7, C8B, C9, CCL16, CCL26, CD1A, CD34, CDH5, CLU, CX3CL1, DEFB103B, IFNB1, IL1RL2, IL22RA2, IL29, IL9, ITLN2, KIR3DL3, MASP1, MBL2, PLA2G2A, THY1, VCAM1) for which expression was absent from more than 90% of the samples in each condition. Some genes were absent in more than 90% of samples in one condition but not in the others. We facilitated the identification of genes absent in a high proportion of samples for a given condition by calculating for each gene in each condition the percentage of samples with expression levels above the background (the gene score). In this study, results are presented exclusively for genes with a gene score >0.5. The scores for each gene in each condition are presented in Dataset S26.

Flow Cytometry Analysis. $CD45^+$ cell populations from whole-blood samples were analyzed with an eight-color flow cytometry panel as previously described (7). We first identified $CD45^+$ cells and then excluded doublets. We then used expression of CD19 to identify B cells. T cells were identified among the $CD19^-$ cells on the basis of their CD3 expression and were analyzed for the expression of CD4 and CD8β. Within the $CD3^-$ cell population, NK cells were identified as $CD56^+$. In the population of $CD56^-$ cells, monocytes were identified as $CD16^+SSC^{low}$, and neutrophils were identified as $CD16^+SSC^{hi}$ [figure 2 in Hasan et al. (7)]. We excluded from the analyses 50 individuals whose immune cell proportions could not be quantified by flow cytometry because of staining and/or cell lysis issues. In total, 936 donors were retained for additional analyses.

Differential Expression Analysis. For each gene in each stimulated condition, we performed a paired *t* test to compare the expression in stimulated and nonstimulated states [the effect of TruCulture batches was first eliminated with the ComBat function of the R Surrogate Variable Analysis (SVA) package (8)]. We controlled globally for FDR. Gene expression change was considered significant if FDR < 0.01. We estimated the proportion of true null hypotheses (π_0) to be much less than one; thus, we used the Bioconductor qvalue package (9) to compute the FDR.

Linear Regression Models. For each gene, we built seven multiple regression models to estimate the effects of age and sex on gene expression in the absence of stimulation and in the six stimulation conditions. We adjusted the models for the proportions of immune cell populations to account for interindividual differences and for the batch of TruCulture tubes to correct for possible differences in the manufacturing process. The full model was as follows:

$$Y^i = \gamma_0{}^i + \gamma_1{}^i \cdot \mathrm{AGE} + \gamma_2{}^i \cdot \mathrm{SEX} + \gamma_3{}^i \cdot \mathrm{TC} + \sum \gamma_j{}^i \cdot X_j,$$

where Y^i, AGE, SEX, TC, and X_j are 936-element vectors, in which each element corresponds to the value of the variable for one donor. Y^i elements are equal to log2(counts) for gene i. AGE elements are the donor's age in years. SEX elements are equal to one for men and zero for women. TC elements are equal to one if a TruCulture tube from batch B was used for the stimulation and zero otherwise, and X_j elements are equal to the jth cell count proportions (i.e., $\#cell_j/\#CD45^+$, where $cell_j$ = {$CD19^+$, NK, $CD4^+$, $CD8^+CD4^+$, $CD8^+$, $CD4^-CD8^-$, monocytes, neutrophils}). Proportions of $CD45^+$ cell populations were preadjusted for age and sex effects to obtain estimates of the total effects of the covariates of interest on gene expression. We ensured that there were no hidden technical batch effects (an unknown source of noise) other than the known effects already included in the model for each stimulus by performing an SVA analysis [with the sva package of R (8) that provides functions for identifying and building surrogate variables for high-dimensional datasets]. No latent variables were identified for any of the stimuli. We controlled for FDR separately for age and sex but globally for conditions (*P* values for all conditions were combined, and then, FDR was calculated). Only effects with an FDR < 0.01 were considered significant. We estimated the proportion of true null hypotheses (π_0) to be much less than one for both age and sex; thus, we used the Bioconductor qvalue package (9) to compute the FDR.

ANOVA for Age Groups. For each gene in each condition, we used ANOVA to detect differences in expression between five age groups (20–29, 30–39, 40–49, 50–59, 60–69). We controlled for FDR separately for each condition, and effects with an FDR < 0.01 were considered significant. For each gene for which a significant effect was detected, we checked whether its expression was also associated with age in linear regression analysis. Only for IAV stimulation, we observed a large number of genes not associated with age in linear regression analysis (i.e., nonlinear associations between gene expression and age). For these genes, we used the Tukey Honest Significant Differences (HSD) test to identify the pairs of age groups for which the difference between mean expression levels was most significant. Dataset S4 shows the genes with a Tukey HSD $P < 0.05$ and those for which the most significant difference in expression was that between the 20- to 29- and 30- to 39-y-old age groups. For IAV stimulation, we confirmed the relevance of the selected age groups by a sliding window analysis. For each gene and each age (from 25 to 65 y old), we performed a *t* test to assess the difference in expression means between the 5 preceding years and the 5 following years. For example, for an age of 25 y old, we compared the gene expression of 20- to 24-y-old donors with that of 25- to 29-y-old donors. We controlled for FDR separately for each age. Fig. S3 reports the number of genes with an FDR < 0.01 for each age.

Mediation Analysis. We used structural equation modeling (i.e., a multivariate statistical method used to analyze complex relationships between independent and dependent variables) (10) to estimate (*i*) the direct effects of age and sex on gene expression and (*ii*) the mediated effects by which age and sex indirectly affect gene expression through the proportions of the eight immune cell populations (i.e., indirect effects). By contrast to the other analyses, the proportions of $CD45^+$ cell populations were not preadjusted for age and sex, because this would have prevented the detection of indirect effects. Path diagrams representing the structural equation models used for the analyses are shown in Fig. S5. We estimated the direct and indirect effects of age independently for each condition and only for genes with a significant total effect detected by regression analysis. Bonferroni correction was applied separately for each condition. Dataset S6 shows Bonferroni-corrected levels of significance for each condition. We then selected the genes with at least one significant age effect (direct or indirect). For these genes, we calculated the proportions of genes with direct and indirect effects. Note that the number of genes with at least one significant age effect in structural equation modeling was smaller than the number of genes found to be associated with age in regression analysis. This discrepancy was caused by differences in the multiple testing correction method used (Bonferroni and FDR). Similar analyses were performed for the sex effects.

DNA Genotyping and Imputation. All individuals were genotyped at 719,665 SNPs on a HumanOmniExpress-24 BeadChip (Illumina).

The coverage of rare functional variants was increased by genotyping an additional 245,766 SNPs on a HumanExome-12 BeadChip (Illumina). After strict quality control filters and merging the two datasets, we retained a total of 723,341 SNPs. Next, we performed genotype imputation with IMPUTE v.2 (11) using the 1000G Project imputation reference panel (Phase1 v3.2010/11/23). After quality control filters, we obtained a dataset of 5,699,237 high-quality SNPs with an MAF > 0.05. Finally, all variants with multicharacter allele codes were excluded [with the snps-only option in PLINK 1.9 (12)], yielding a final dataset of 5,265,361 SNPs that were used for eQTL association analysis.

Expression eQTL Mapping. The eQTL analysis was performed with a linear mixed model implemented in the GenABEL package (13). Genetic relatedness matrices (GRMs) were estimated for each nonsex chromosome separately using the remaining 21 chromosomes to exclude potentially associated SNPs from the GRM [the so-called "leave-one-chromosome-out" approach (14)]. For local eQTL mapping, we considered SNPs located within 1 Mb distance from the start or end of a given gene (2,173,933 tests for 542 genes; 18 of 560 genes were located on sex chromosomes and were not considered in this analysis). For *trans*-eQTL mapping, we considered all of the remaining SNPs (2,946,428,227 tests for 560 genes). In each condition, we tested the associations between SNPs and gene expression [log2(counts)]. We adjusted the linear mixed model for age, sex, proportions of $CD45^+$ cell populations, and TruCulture batch. The proportions of $CD45^+$ cell populations were preadjusted for age and sex. Bonferroni correction for multiple testing was applied separately for each condition. Local eQTLs were considered significant if $P < 2.3 \times 10^{-8}$ (0.05/2,173,933), and *trans*-eQTLs were considered significant if $P < 1.7 \times 10^{-11}$ (0.05/2,946,428,227). All statistically significant local and *trans*-genetic associations are reported in Dataset S22. If the distance between two significant *trans* associations was less than 500 kb, we considered them to belong to a single *trans*-acting locus (Table S2). All eQTLs detected for genes for which expression was assessed with probes hybridizing to a region containing one or two common SNPs (Dataset S25) should be interpreted with caution.

To verify that the detected *trans*-eQTLs were true master regulators of gene expression and not quantitative trait loci of specialized immune cell subpopulations that were not among the eight major immune cell types included in our model, we tested if *trans*-eQTL SNPs were associated with any of the 105 cell proportions measured by flow cytometry in the Milieu Intérieur cohort. Only 4 of 15 *trans*-eQTLs were significantly associated with a flow cytometric measure (linear mixed model-adjusted $P < 0.05$). We then performed eQTL mapping for all genes affected *in trans* by the corresponding SNPs, adjusting for the associated proportions of immune cell subpopulations in addition to the eight major cell populations initially included. Adjusting for these additional cell proportions did not affect the summary statistics of the four *trans*-eQTLs, providing strong statistical support that they have true *trans*-acting effects on gene expression.

SNP × SEX and SNP × AGE Interactions. We estimated the SNP × SEX and SNP × AGE interaction effects on gene expression using ProbABEL v.0.4.5 from the GenABEL suite of programs (15). The analysis was performed separately for each of seven conditions. For each gene, we ran the polygenic function of GenABEL, including age, sex, proportions of $CD45^+$ cell populations, and TruCulture batch as covariates. We then ran the palinear function of ProbABEL separately for SNP × SEX and SNP × AGE using polygenic residuals (h2.obj$residualY, where h2.obj is the object returned by the GenABEL polygenic function) as the phenotype and the inverse variance–covariance matrix (h2.obj$InvSigma) as the –mscore argument with the –interaction option. Only SNPs located within 1-Mb distance from the gene were tested. We used a Bonferroni-corrected level of significance, $\alpha = 2.3 \times 10^{-8}$ (0.05/2,173,933). For the estimation of SNP × AGE interaction effects, we created a binary variable for AGE, with zero corresponding to young (20–39) and one corresponding to old (50–69) individuals. Individuals in the 40- to 49-y-old age group were excluded to maintain equal group sizes and to maximize the possible differences between the groups (assuming a monotonic effect of age).

Power Calculation. We used simulations to evaluate the statistical power of the analysis to detect SNP × SEX and SNP × AGE interactions. We assumed total sample sizes of 936 (for sex) and 775 (for age) individuals, which we divided equally between men and women and between young and old subjects, respectively. We set three different effect sizes for women (or young individuals), and we paired them with a range of effect sizes for men (or old individuals). The choice of effect sizes for women (or young individuals) was based on the standardized main effects detected in local eQTL mapping in the absence of stimulation (Dataset S8). We selected effect sizes for women (or young individuals) corresponding to the 5th, 50th, and 95th quantiles of the observed effect sizes (i.e., $\beta = 1$, 0.5, and 0.3, respectively). For each of the three selected effect sizes, we chose a candidate gene (i.e., *ATM* for $\beta = 0.3$, *IKBKAP* for $\beta = 0.5$ and *CTSC* for $\beta = 1$). For each of these genes, we assigned female (or young) donors into three groups on the basis of genotype for the most significant SNP. In each group, we simulated the expression of the gene with a normal distribution with a mean equal to the observed mean for expression of the gene in the given group and an SD equal to the mean of SDs for the three groups. We did the same for men after multiplying the means for the three groups by 0, 0.2, 0.4, 0.6, and 0.8. This resulted in total of three × five different SNP × SEX (or SNP × AGE) interaction effects. For each interaction effect, we assumed to run a total of 2,173,933 tests corresponding to the number of SNPs located in the vicinity (±1 Mb) of the 542 genes. After Bonferroni correction for multiple testing, the significance level α was equal to 2.3×10^{-8} (0.05/2,173,933). We calculated the power of our test assuming six different MAFs (i.e., 0.05, 0.1, 0.2, 0.3 0.4, 0.5). For each combination of MAF and interaction effects, we repeated the simulation 1,000 times. The power was then calculated as the ratio between the number of times that the interaction effect was considered significant and the total number of simulations (i.e., 1,000).

Annotation of eQTLs with Genome-Wide Association Study Hits. For all sets of eQTLs listed in Datasets S8–S21, we considered those identified in the absence of stimulation and the response eQTLs, and we explored their implication in human diseases and traits using hits from genome-wide association studies (GWAS) from the January 30, 2017 version of the EBI-NHGRI GWAS Catalog. Only GWAS signals with $P < 5 \times 10^{-8}$ were used. A set of n independent eQTLs was considered to be enriched in GWAS SNPs if the proportion of eQTLs that were GWAS SNPs or in linkage disequilibrium with GWAS SNPs ($r^2 > 0.8$) in this set was larger than that in 95% of 10,000 randomly sampled sets of n SNPs. For local eQTLs, random SNPs were drawn from 92,133 independent SNPs selected from the 1,163,918 SNPs tested for local associations [with the –indep pairwise 100 5 0.5 function of PLINK 1.9 (12)]. For *trans*-eQTLs, random SNPs were drawn from 407,845 independent SNPs selected from all 5,265,361 SNPs (with –indep pairwise 100 5 0.5).

Decomposition of the Proportion of Variance Explained. For each gene in each condition, we included age, sex, significant local and *trans*-eQTLs, and proportions of $CD45^+$ cell populations in a single linear model [the effect of TruCulture batch was first eliminated with the ComBat function of the sva package of R (8)]. The proportion of the expression variance explained by each

variable was calculated by averaging the sums of squares in all orderings of the variables in the linear model using the lmg metric in the relaimpo package of R (16). We reported the number of genes found to be associated with age, sex, and local and *trans*-eQTLs (in linear regressions and linear mixed models) and the variance explained averaged across genes. For whole-blood cell composition, we reported the number of genes for which at least one cell population type explained at least 1% of the expression variance. For these genes, we calculated the mean of the overall variance explained by all eight cell population types. Dataset S24 reports the proportions of variance explained by each variable for each gene and each condition.

Data Availability. Genotype data are available in the European Genome-Phenome Archive under the accession code EGAS00001002460. Gene expression data can be found in Dataset S1. All of the reported results can be explored and mined in the accompanying web application (misage.pasteur.fr/).

1. Thomas S, et al.; Milieu Intérieur Consortium (2015) The Milieu Intérieur study—an integrative approach for study of human immunological variance. *Clin Immunol* 157: 277–293.
2. Duffy D, et al.; Milieu Intérieur Consortium (2014) Functional analysis via standardized whole-blood stimulation systems defines the boundaries of a healthy immune response to complex stimuli. *Immunity* 40:436–450.
3. Bisiaux A, et al.; Milieu Intérieur Consortium (2017) Deconvolution of the response to Bacillus Calmette-Guérin reveals NF-κB-induced cytokines as autocrine mediators of innate immunity. *Front Immunol* 8:796.
4. Urrutia A, et al.; Milieu Intérieur Consortium (2016) Standardized whole-blood transcriptional profiling enables the deconvolution of complex induced immune responses. *Cell Rep* 16:2777–2791.
5. Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 26:873–881.
6. Durinck S, Spellman PT, Birney E, Huber W (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 4:1184–1191.
7. Hasan M, et al.; Milieu Intérieur Consortium (2015) Semi-automated and standardized cytometric procedures for multi-panel and multi-parametric whole blood immunophenotyping. *Clin Immunol* 157:261–276.
8. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28:882–883.
9. Storey JD, et al. (2007) Gene-expression variation within and among human populations. *Am J Hum Genet* 80:502–509.
10. Gunzler D, Chen T, Wu P, Zhang H (2013) Introduction to mediation analysis with structural equation modeling. *Shanghai Jingshen Yixue* 25:390–394.
11. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 5: e1000529.
12. Chang CC, et al. (2015) Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 4:7.
13. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM (2007) GenABEL: An R library for genome-wide association analysis. *Bioinformatics* 23:1294–1296.
14. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL (2014) Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet* 46:100–106.
15. Aulchenko YS, Struchalin MV, van Duijn CM (2010) ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* 11:134.
16. Grömping U (2006) Relative importance for linear regression in R: The package relaimpo. *J Stat Softw* 17:1–27.

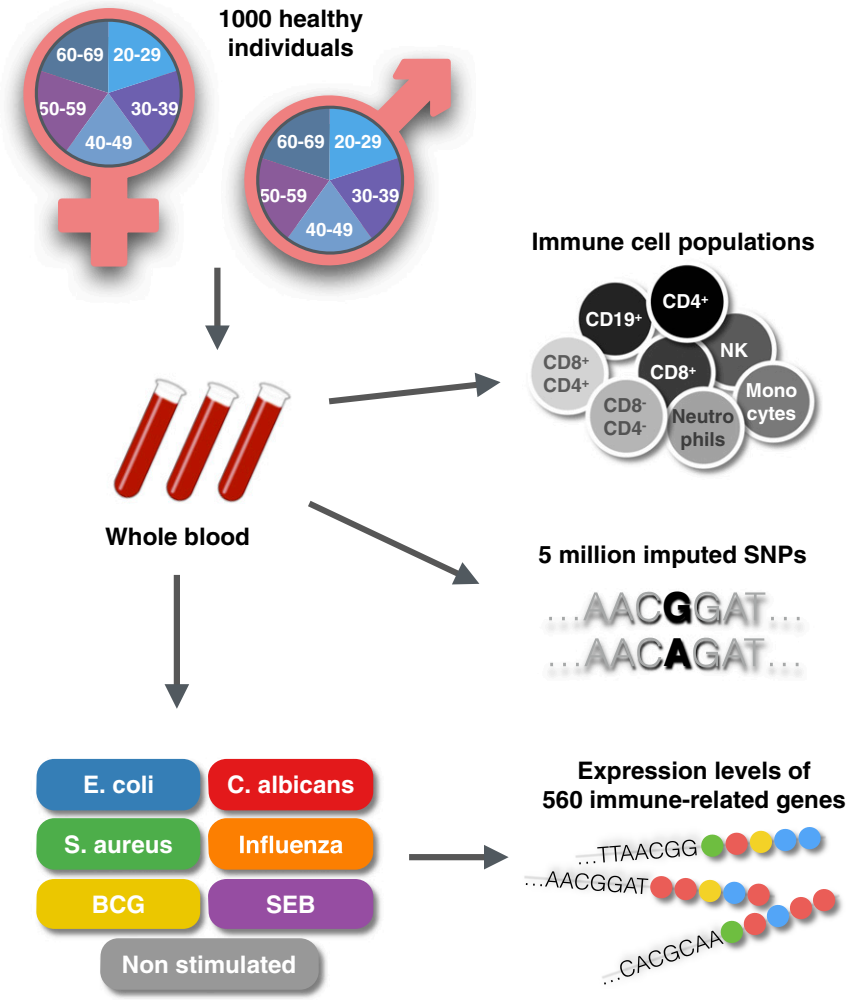

Fig. S1. Schematic overview of the study. Whole-blood samples from 1,000 healthy individuals of the Milieu Intérieur cohort stratified into groups of equal size for age and sex. The genome-wide genetic diversity in all individuals was defined with the HumanOmniExpress and HumanExome BeadChips, which after imputation, yielded a final dataset of ~5 million SNPs per individual. In parallel, eight major leukocyte populations ($CD45^+$ cells) were quantified in all individuals by flow cytometry, and whole-blood stimulation experiments were performed with three bacteria, a virus, a fungus, and a superantigen using the TruCulture system. The induced expression profiles of 560 immune-related genes were assessed with NanoString hybridization arrays.

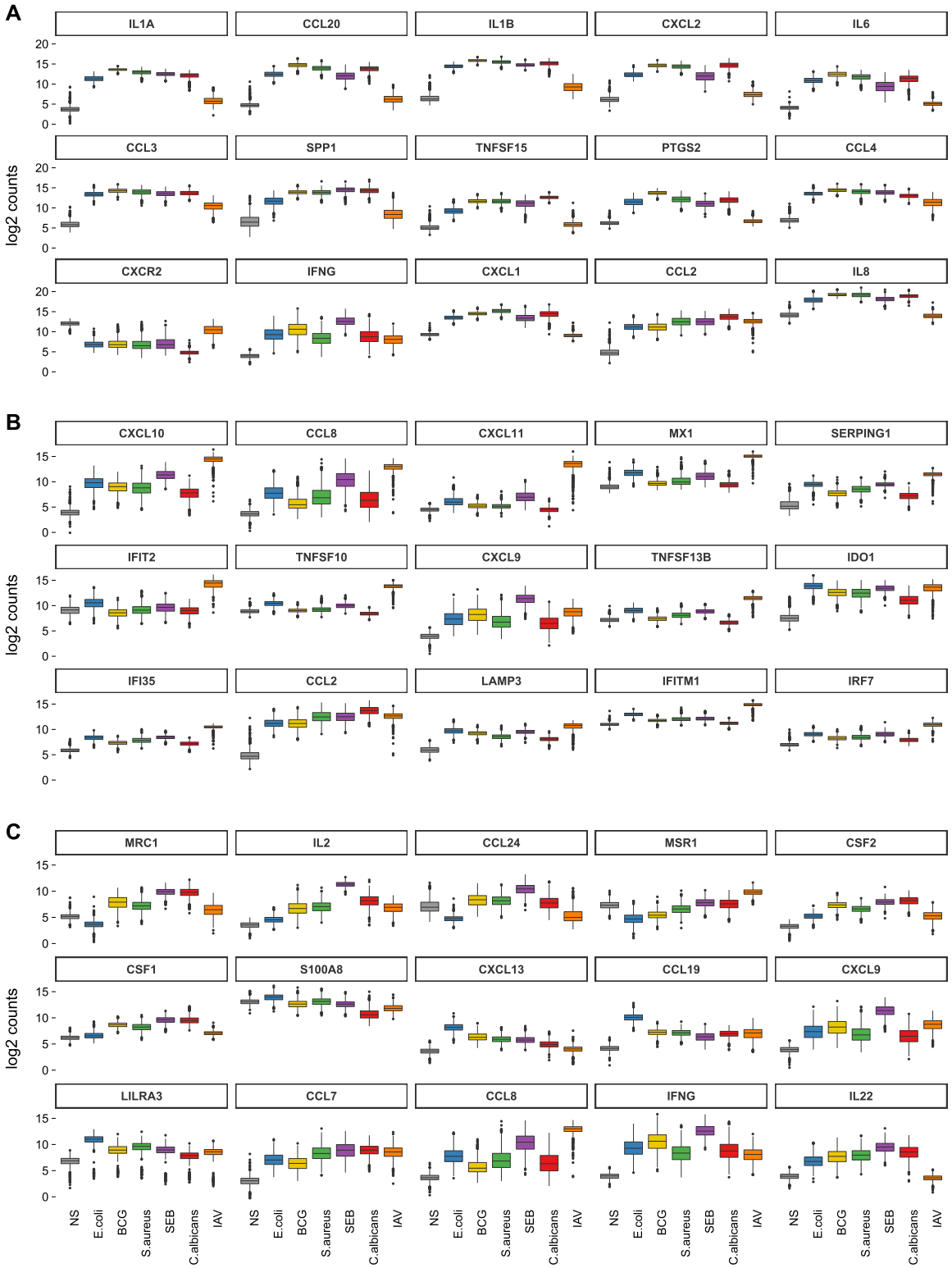

Fig. S2. Expression of genes contributing the most to the first three PCs. (*A*) Top 15 genes for PC1. (*B*) Top 15 genes for PC2. (*C*) Top 15 genes for PC3. NS, nonstimulated control.

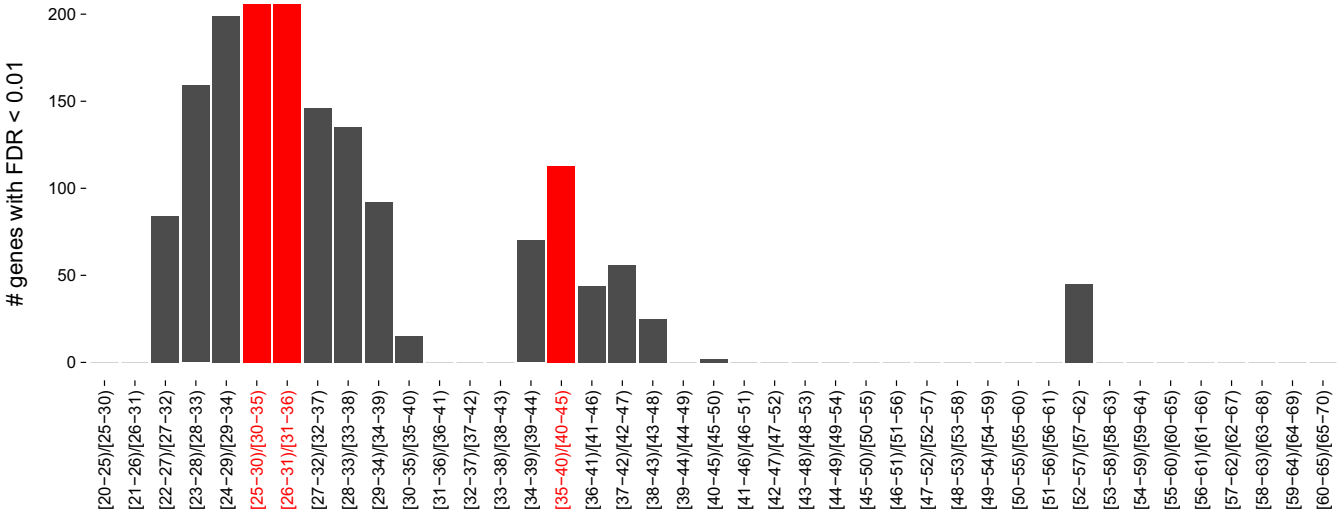

Fig. S3. Sliding window analysis for age groups relevance in response to IAV stimulation. On the *x* axis are shown the consecutive pairs of age groups compared for expression differences. On the *y* axis are shown the numbers of genes with significant *t* test after controlling for FDR at 0.01. The highest differences were observed for ages of 30 and 40 y old (in red), which support decade-long age groups division.

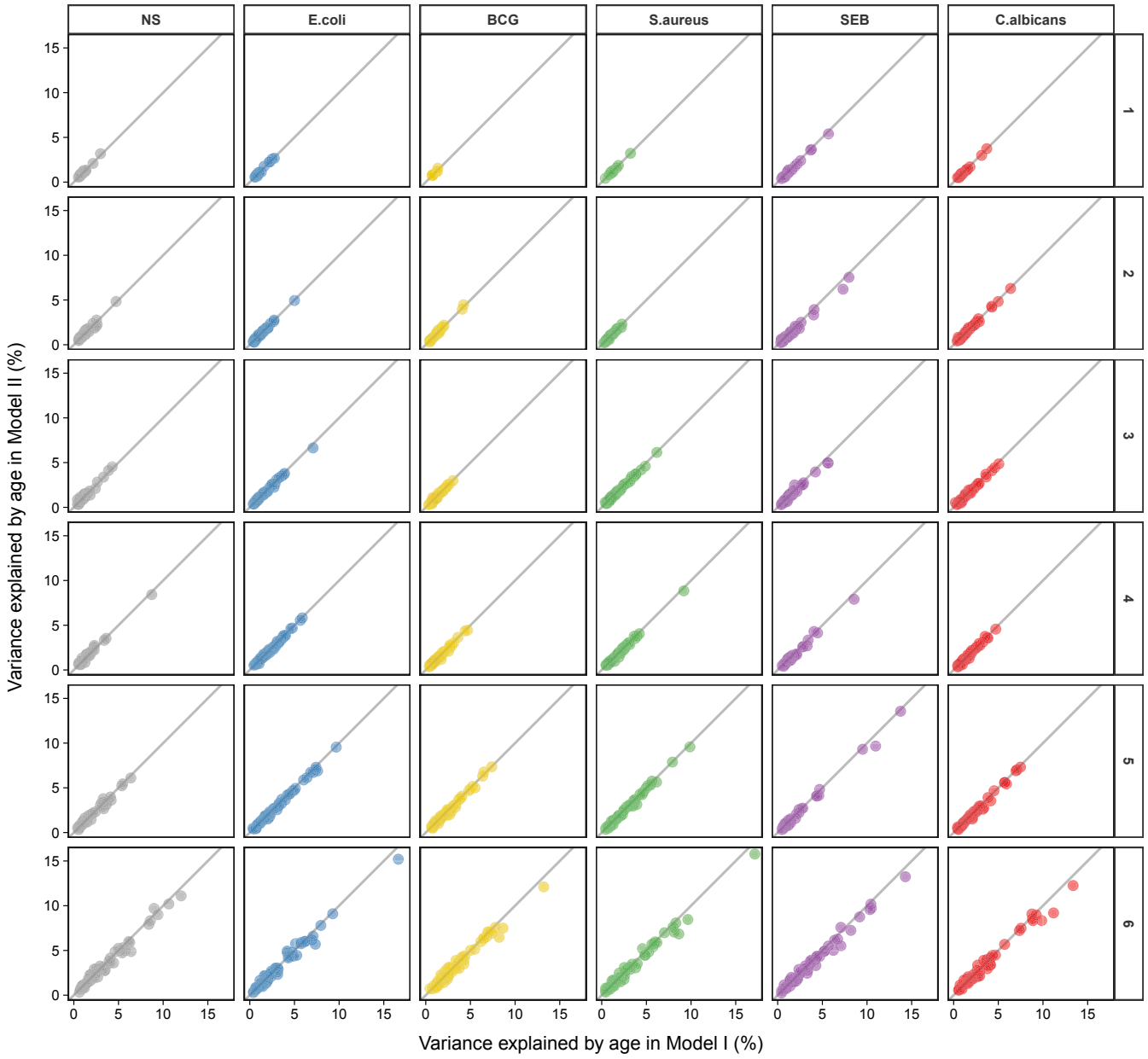

Fig. S4. Comparison of proportion of expression variance explained by age between two regression models: model I ($Y^i = \gamma_0{}^i + \gamma_1{}^i \cdot AGE + \gamma_2{}^i \cdot SEX + \gamma_3{}^i \cdot TC + \Sigma\gamma_j{}^i \cdot X_j$) and model II ($Y^i = \gamma_0{}^i + \gamma_1{}^i \cdot AGE + \gamma_2{}^i \cdot CMV + \gamma_3{}^i \cdot SEX + \gamma_4{}^i \cdot TC + \Sigma\gamma_j{}^i \cdot X_j$,). Y^i, AGE, SEX, CMV, TC, and X_j are 936-element vectors, in which each element corresponds to the value of the variable for one donor. Y^i elements are equal to log2(counts) for gene i in a given condition. AGE elements are the donor's age in years. CMV elements are equal to one if the donor was infected with CMV and zero otherwise. SEX elements are equal to one for men and zero for women. TC elements are equal to one if a TruCulture tube from batch B was used for the stimulation and zero otherwise, and X_j elements are equal to the jth cell count proportions (i.e., $\#cell_j/\#CD45^+$, where $cell_j$ = {$CD19^+$, NK, $CD4^+$, $CD8^+CD4^+$, $CD8^+$, $CD4^-CD8^-$, monocytes, neutrophils}). Results are presented separately for each stimulus and each age specificity group (from 1 to 6). NS, nonstimulated control.

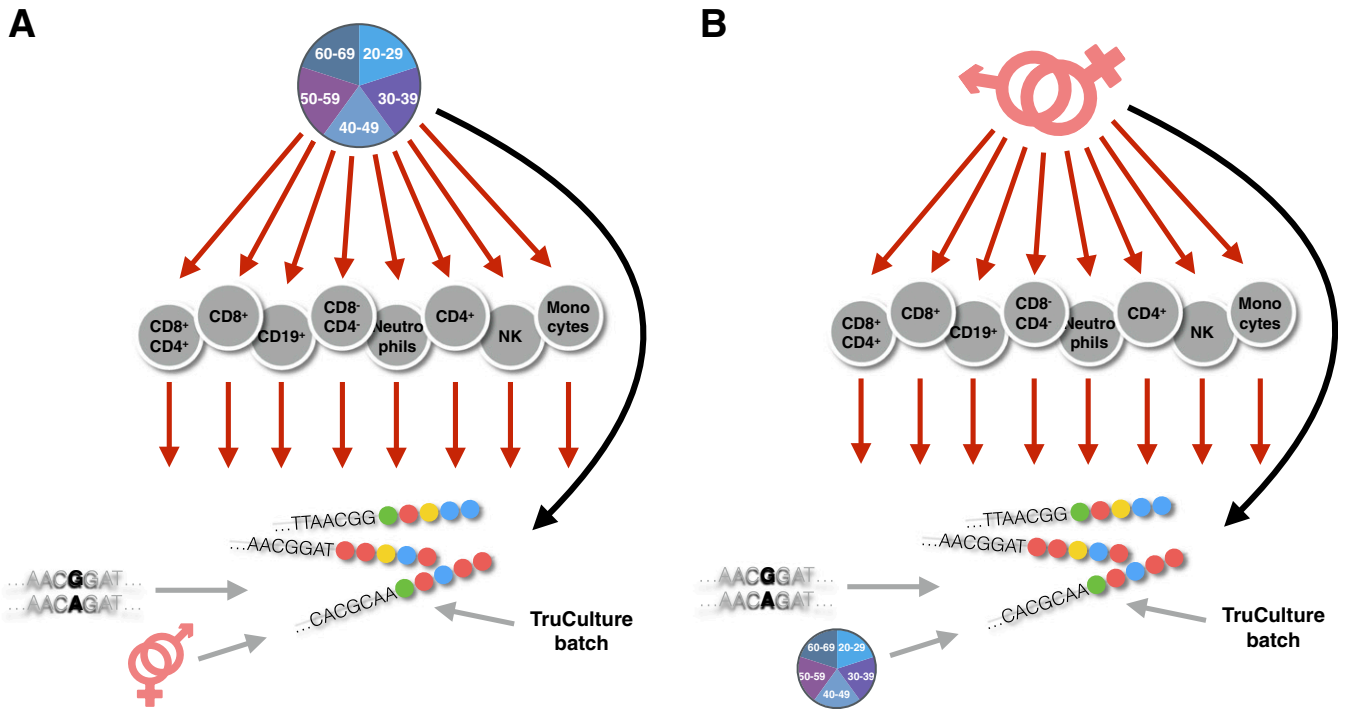

Fig. S5. Path diagrams representing structural equation models used to test mediation of (*A*) age and (*B*) sex effects on gene expression by the proportions of $CD45^+$ cell populations. Black arrows indicate direct effect, red arrows indicate indirect (mediation) effects, and gray arrows indicate additional (adjusting) variables used in the model.

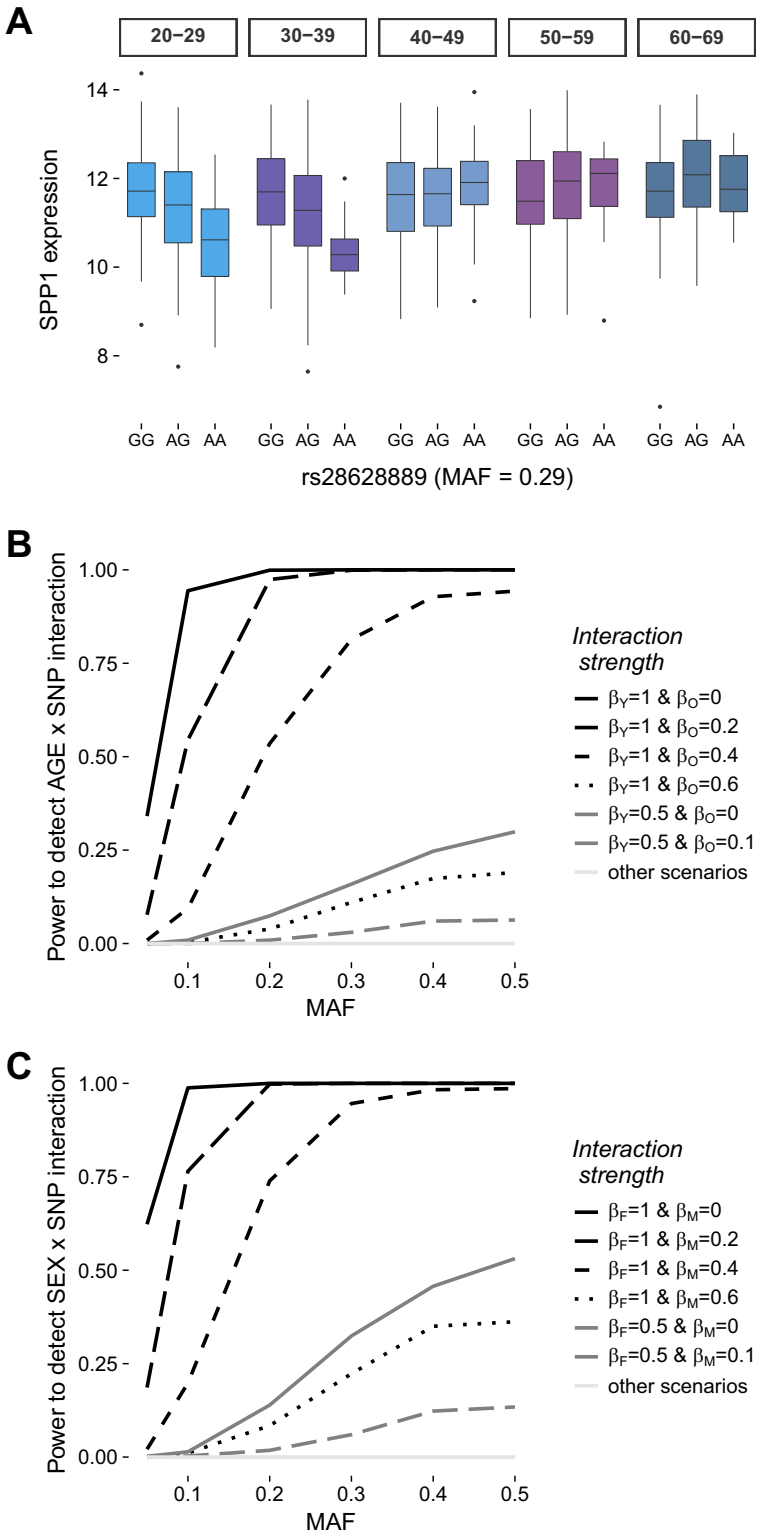

Fig. S6. Effect of age and sex on genetic control of transcriptional response to immune stimulation. (*A*) eQTL presenting age-specific effect in response to *E. coli* stimulation (AGE × SNP interaction). To test for interaction, we compared the effect of rs28628889 on *SPP1* expression within 20- to 39-y-old ("young") and 50- to 69-y-old ("old") groups. (*B*) Power of the tests to detect AGE × SNP interactions. Y stands for young (i.e., 20–39 y old), and O stands for old (i.e., 50–69 y old). Other scenarios include $\beta_Y = 1$, $\beta_O = 0.8$; $\beta_Y = 0.5$, $\beta_O \geq 0.2$; $\beta_Y = 0.3$ and $\beta_O \geq 0$. (*C*) Power of the tests to detect SEX × SNP interactions. F stands for female, and M stands for male. Other scenarios include $\beta_F = 1$, $\beta_M = 0.8$; $\beta_F = 0.5$, $\beta_M \geq 0.2$; and $\beta_F = 0.3$, $\beta_M \geq 0$.

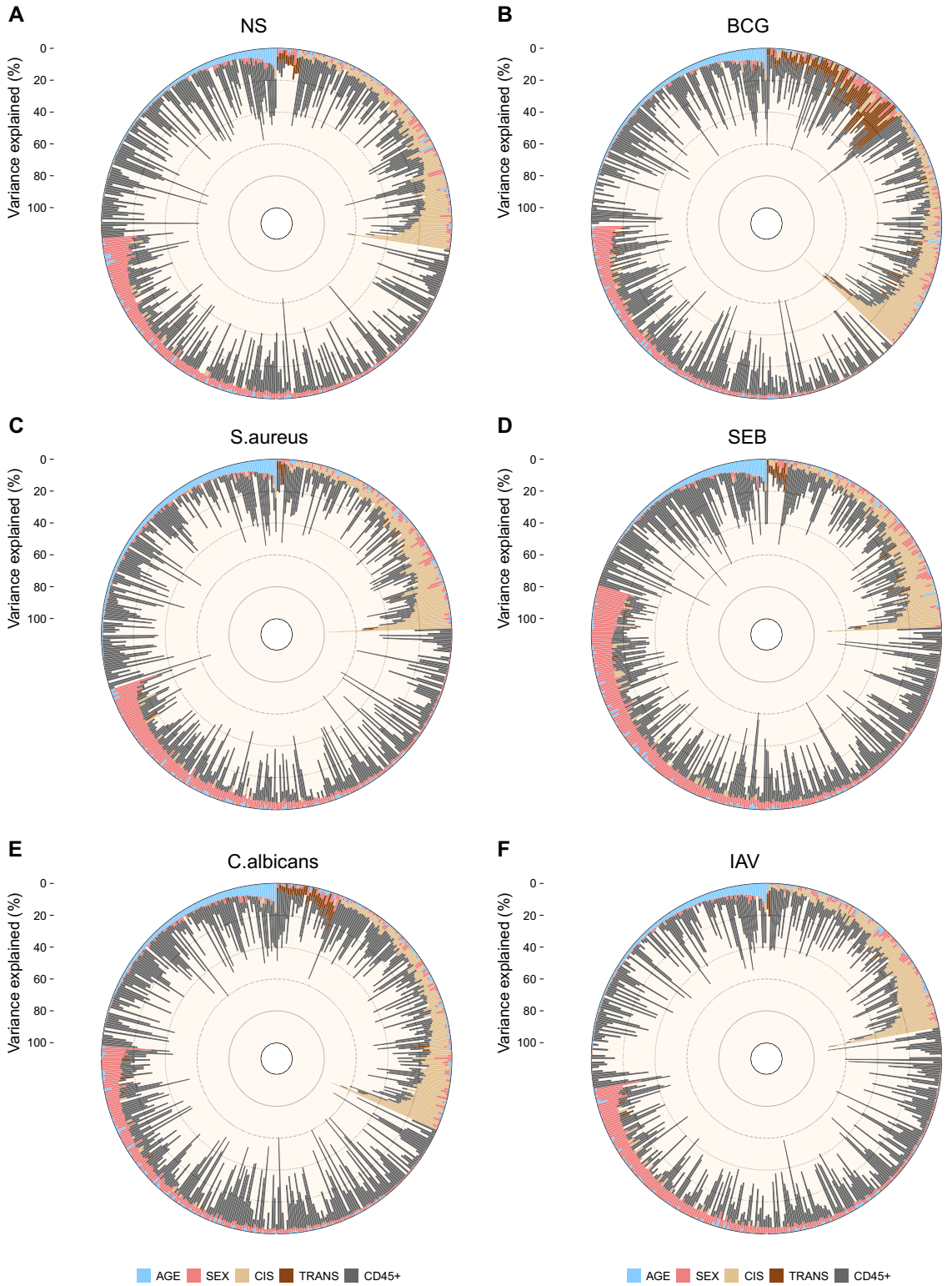

Fig. S7. Expression variance explained by age, sex, genetics, and proportions of CD45^{+} cell populations for all genes in response to nonstimulated (*A*) and five stimulated conditions (*B*–*F*). NS, nonstimulated control.

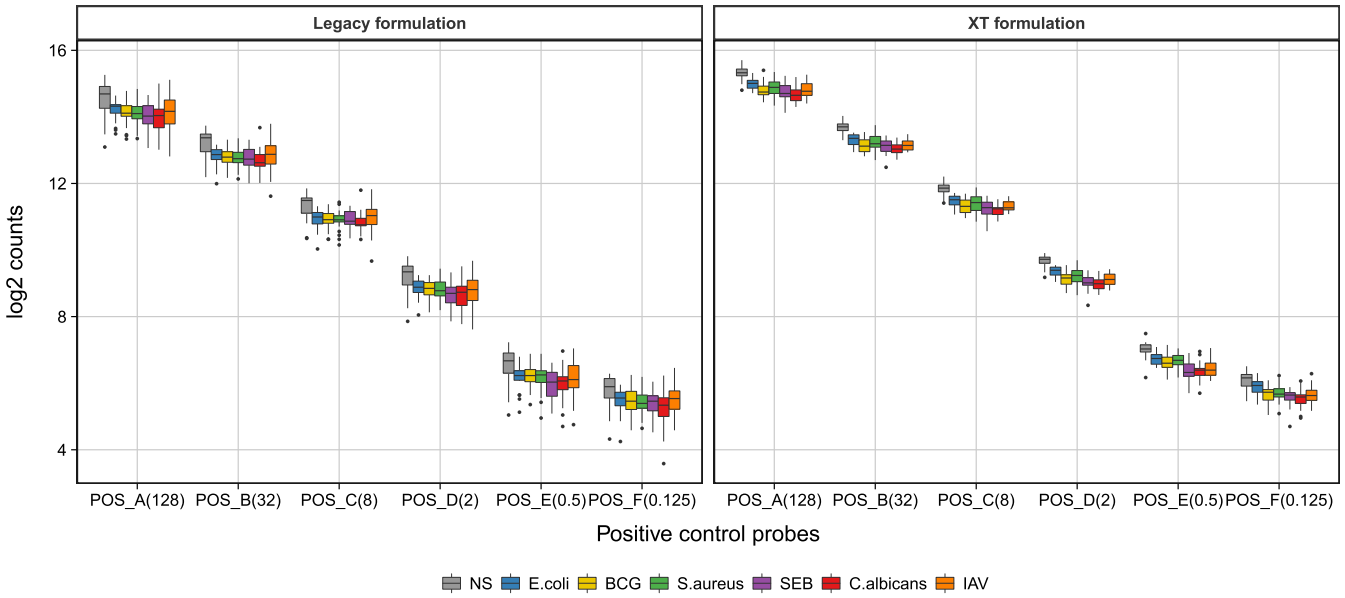

Fig. S8. Differences in counts for positive control probes measured in experiments performed with two chemistries (Legacy and XT formulations) in non-stimulated and six stimulated conditions. NS, nonstimulated control.

Table S1. Number of genes with sex- and age-dependent expression in nonstimulated (NS) and six stimulated conditions

Intrinsic factors	NS	*E. coli*	BCG	*S. aureus*	SEB	*C. albicans*	IAV
Age	206	277	252	270	227	252	217
Sex	328	336	343	348	372	347	347

Table S2. Number of genes with expression regulated by local and *trans*-eQTLs in the nonstimulated (NS) and six stimulated conditions

Stimuli	Genes with local eQTLs	*Trans*-acting eQTLs	*Trans*-regulated genes (maximum)
NS	135	4	5
E. coli	132	7	105
BCG	151	7	80
S. aureus	134	6	7
SEB	146	8	13
C. albicans	136	6	34
IAV	125	5	2

Table S3. GWAS enrichment statistics for lists of eQTLs detected in nonstimulated (NS) state and eQTLs detected after stimulations

eQTLs list	Observed	Expected	*P* value	Excess
Best.eqtls.cis.NS	11	3.1	0.037	3.5
Best.reqtls.cis.E.coli	6	1.5	0.045	4.0
Best.reqtls.cis.BCG	10	1.7	0.022	5.8
Best.reqtls.cis.S.aureus	7	1.4	0.027	5.2
Best.reqtls.cis.SEB	8	1.8	0.032	4.4
Best.reqtls.cis.C.albicans	11	1.5	0.017	7.3
Best.reqtls.cis.IAV	5	1.0	0.037	5.1
Best.eqtls.trans.NS	1	0.056	0.035	17.8
Best.reqtls.trans.E.coli	2	0.066	0.0098	30.3
Best.reqtls.trans.BCG	1	0.056	0.036	17.8
Best.reqtls.trans.S.aureus	2	0.046	0.0069	43.5
Best.reqtls.trans.SEB	2	0.06	0.008	33.3
Best.reqtls.trans.C.albicans	1	0.062	0.042	16.1
Best.reqtls.trans.IAV	0	0.036	1.0	0.0

Other Supporting Information Files

Dataset S1 (TXT)
Dataset S2 (XLSX)
Dataset S3 (XLSX)
Dataset S4 (XLSX)
Dataset S5 (XLSX)
Dataset S6 (XLSX)
Dataset S7 (XLSX)
Dataset S8 (XLSX)
Dataset S9 (XLSX)
Dataset S10 (XLSX)
Dataset S11 (XLSX)
Dataset S12 (XLSX)
Dataset S13 (XLSX)
Dataset S14 (XLSX)
Dataset S15 (XLSX)
Dataset S16 (XLSX)
Dataset S17 (XLSX)
Dataset S18 (XLSX)
Dataset S19 (XLSX)
Dataset S20 (XLSX)
Dataset S21 (XLSX)
Dataset S22 (TXT)
Dataset S23 (XLSX)
Dataset S24 (XLSX)
Dataset S25 (TXT)
Dataset S26 (TXT)