



Supplementary Materials for

Polarization of the Effects of Autoimmune and Neurodegenerative Risk Alleles in Leukocytes

Towfique Raj, Katie Rothamel, Sara Mostafavi, Chun Ye, Mark N. Lee, Joseph M. Replogle, Ting Feng, Michelle Lee, Natasha Asinovski, Irene Frohlich, Selina Imboywa, Alina Von Korff, Yukinori Okada, Nikolaos A. Patsopoulos, Scott Davis, Cristin McCabe, Hyun-il Paik, Gyan P. Srivastava, Soumya Raychaudhuri, David A. Hafler, Daphne Koller, Aviv Regev, Nir Hacohen, Diane Mathis, Christophe Benoist,* Barbara E. Stranger,* Philip L. De Jager*

*Corresponding author. E-mail: christophe_benoist@hms.harvard.edu (C.B.); bstranger@medicine.bsd.uchicago.edu (B.E.S.); pdejager@partners.org (P.L.D.J.)

Published 2 May 2014, *Science* **344**, 519 (2014)
DOI: 10.1126/science.1249547

This PDF file includes:

Materials and Methods
Figs. S1 to S25
Tables S1 to S3, S10, S15, and S18
Full Reference List

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/344/6183/519/suppl/DC1)

Tables S4 to S9, S11 to S14, S16, S17, and S19

Materials and Methods

Sample collection, study populations and demographics

Healthy subjects between the ages of 18 and 50 were recruited from the Greater Boston Area by comprehensive advertisement strategies. As an initial step, 696 subjects who self-reported as healthy with no autoimmune, neurologic, metabolic, or chronic infectious diseases consented to be a part of the PhenoGenetic study at the Brigham and Women's Hospital, a living biobank of 1749 subjects who can be recalled over five years based on demographics or genotype. These subjects were then successfully recalled to be a part of the Immune Variation (ImmVar) project. The cohort specifically consists of 162 African American subjects of European and African ancestry (AA), 155 East Asian subjects of Chinese, Japanese, or Korean ancestry (EA), and 377 Caucasian subjects of European ancestry (EU). The median age of the subjects is 24. There are 414 females and 282 males. The following demographic information was recorded for each participant: age, race, sex, smoking, weight, height, BMI, self-reported ethnicity, blood pressure, and menstrual cycle.

PBMC isolation, cell sorting, and RNA preparation

Over the course of 18 months, fresh blood samples from these 696 participants were collected following a rigorous, standardized set of procedures (SOP). All participant clinic visits occurred between 7:30 AM and 8:30 AM to minimize circadian rhythm fluctuations. 15ml of blood was collected in Vacutainer tubes coated in EDTA and was used to isolate PBMCs by density gradient centrifugation on Ficoll-Hypaque. PBMC isolation and processing began promptly at 8:30am in order to reduce processing variables. 15 ml of blood were diluted 2-fold with PBS/2mM EDTA (Gibco, pH=8.0) at room temperature (20°-25°C), and slowly layered with a pipette aid over 15ml of Ficoll-Hypaque (GE Healthcare) solution in a 50ml conical tube. Samples were centrifuged for 20 mins at 900g and 25°C with no brake. Upon centrifugation, the mononuclear cell layer at the interface was extracted, washed with 12ml PBS/2mM EDTA, and centrifuged for 4 mins at 400 g. Remaining red blood cells were lysed with 1 ml of ACK lysis buffer for 90 seconds at room temperature, washed and centrifuged for 4 mins at 400 g. PBMC processing time was strictly kept at 1.25 hour for all samples. For staining, cells were resuspended in ice-cold 100 µl staining media (90% DMEM/10% FBS) at a concentration of 10^7 to 2×10^7 . Cells were stained with appropriate antibodies for 15 mins on ice: CD62L 605NC, CD25 PE, CD3 FITC, CD4 Pe Cy7, CD8 APC Cy7, CD16 APC, CD14 PerCP Cy5.5, CD19 APC (all from eBioscience). After washing, cells were sorted on a Becton Dickinson Aria beginning at 10:00am, and with a total sort time for all the sample

of less than 3 hours. Cell population of interest, CD14⁺CD16⁻ and CD4⁺CD62L⁺, underwent a two-step sorting strategy in order to achieve cell purity >99%. After the first sort, cells were directly sorted into 500 µl of ice-cold Trizol, and stored at -80°C. For the cell samples sorted from frozen PBMCs, mononuclear cells were prepared as above, washed and resuspended in 1 ml 10% DMSO in FBS, and placed in a Corning Cryogenic vial. The tubes were placed in NalGene CryoFreezing container at -80°C for 24-48 hours, and then moved into a liquid nitrogen tank for long-term storage. Frozen samples were then thawed in a 37°C water bath for 1 minute, washed in PBS/2mMEDTA for 4 minutes at 400 g, and then stained as above prior to sorting. Sort settings were essentially the same as for fresh cells, and cell viability remained good throughout the process, although the total cell yield was routinely 30% lower for sorts from frozen samples.

Gene expression quality control, quantification, and normalization

mRNAs from these cells were profiled on Affymetrix GeneChip Human Gene ST 1.0 microarrays. Raw data CEL files were processed using the Robust Multichip Average (RMA) algorithm in Affymetrix PowerTools. Dynamic range (DR), the ratio between the highest and lowest signal values in a single data set, was the primary metric of quality for individual expression profiles. To avoid confounding by single outliers, the dynamic range (DR) was calculated for each data set by dividing the 95th by the 5th percentile. Samples with DR below 40 were noted as low quality and were excluded from association analysis. Of the 536 and 501 arrays profiled in CD4⁺ T-cells and Monocytes, 499 and 485 in CD4⁺ and monocytes, respectively, passed the expression quality control. Of which, 479 and 457 in CD4⁺ T cell and monocytes are unique samples. Of the 33,297 probesets of the ST 1.0 array, 6,049 were removed due to the following filters (i) all array features with a single nucleotide polymorphism (SNP) at minor allele frequency (MAF) greater than 0.1 in any of the 1000 Genomes populations were removed. We used a comprehensive compendium of common SNPs on the 1000 Genome Project (9) to remove 31,906 features (out of 764,885 distinct 25-mer oligonucleotide “features”) found in regions with SNPs or SNPs in LD ($r^2 > 0.8$) from our analysis to minimize any potential confounding effects. (ii) All probesets with more than 4 features removed or with more than 25% of features removed were removed entirely. (iii) 4,468 probesets that do not map to the human genome were removed. In addition, we removed a group of 572 probesets known to exhibit a high degree of technical variability, as determined by profiling in technical duplicates of a panel of 15 RNA samples from sorted blood CD4⁺ T-cells. Finally, a total of 2,778 potential cross-hybridization probesets, as provided by Affymetrix, were flagged and removed prior to the *trans*-eQTL association analysis. Finally, we included in our eQTL analyses only those probesets that mapped to a GENCODE v.12 gene (18), and we excluded probesets mapping to the X and Y chromosome. The resulting sets of 26,062 autosomal probesets corresponding to 19,114 unique autosomal GENCODE v.12 annotated genes were analyzed for the association analyses.

The data for each cohort were first internally normalized by dividing the expression values for each gene in individuals of that cohort by the mean expression value across the

cohort, with the assumption that inter-batch differences on normalized data are much lower than those on raw expression values. These normalized values for the three cohorts were assembled and \log_2 -transformed.

To account for non-genetic factors such as batch effects, age, gender, technical artifacts in gene expression data, we used Principal Component Analysis (PCA) (19, 20) (**fig. S19**). Following (20), PCs were estimated separately from the gene expression matrix for each population and cell type. The optimal numbers of PCs for association analysis were determined based on the PC that resulted in maximum number of *cis*-eQTLs (**fig. S20-S21**). This procedure identified 20, 10, and 14 PCs in EU, EA and AA monocytes and 20, 12 and 12 PCs in EU, EA and AA T-cells. The number of *cis*-associations detected along with optimal number of PCs for each dataset is shown in **fig. S20**. We regressed out these PCs from the original gene expression levels, and used the residuals as phenotypes for all association analyses

Genotyping and Imputation

Each subject was genotyped using the Illumina Infinium Human OmniExpress Exome BeadChips, which includes genome-wide genotype data as well as genotypes for rare variants from 12,000 exomes as well as common coding variants from the whole genome. In total, 951,117 SNPs were genotyped, of which 704,808 SNPs are common variants (Minor Allele Frequency [MAF] > 0.01) and 246,229 SNPs are exomic variants. The genotype success rate was greater than or equal to 97%. We applied rigorous subject and SNP quality control (QC) that includes (1) gender misidentification (2) subject relatedness (3) Hardy-Weinberg Equilibrium testing (4) use concordance to infer SNP quality (5) genotype call rate (6) heterozygosity outlier (7) subject mismatches. In the European population, we excluded 1,987 SNPs with a call rate < 95%, 459 SNPs with Hardy-Weinberg equilibrium *p*-value < 10^{-6} , and 63,781 SNPs with MAF < 1% from the 704,808 common SNPs (a total of 66,461 SNPs excluded). In the African-American population, we excluded 2,161 SNPs with a call rate < 95%, 298 SNPs with Hardy-Weinberg equilibrium *p*-value < 10^{-6} , and 17,927 SNPs with MAF < 1% from the 704,808 common SNPs (a total of 20,436 SNPs excluded). In the East-Asian population, we excluded 1,831 SNPs with a call rate < 95%, 213 SNPs with Hardy-Weinberg equilibrium *p*-value < 10^{-6} , and 84,973 SNPs with MAF < 1% from the 704,808 common SNPs (a total of 87,064 SNPs excluded) (**fig. S22-S24**).

All samples were tested for population stratification using the EIGENSTRAT v3.0 software (21), which performs a principal components analysis (PCA) on SNP data. The PCA was performed on a subset of 100K SNPs selected from all genotyped SNPs with MAF > 0.05, omitting regions of high LD. Population outliers were identified by combining ImmVar genotypes with genotypes from HapMap data from the CEU, YRI, CBT, JPT and GIH populations. We used iterative outlier detection approach to remove 16 ancestry outliers. After QC, 52 subjects across all three populations and approximately 18,000 – 88,000 SNPs in each population were filtered out from our analysis (**fig. S25**).

After genotype and expression QC, total number individuals analyzed are 461, of which 401 have monocyte data and 407 have CD4+ T cell data, respectively (**table S1**).

We used the BEAGLE software (version: 3.3.2) to impute the post-QC genotyped markers using reference Haplotype panels from the 1000 Genomes Project (The 1000 Genomes Project Consortium Phase I Integrated Release Version 3) that contain a total of 37.9 Million SNPs in 1,092 individuals with ancestry from West Africa, East Asia, and Europe. For subjects of European and East-Asian ancestry, we used haplotypes from Utah residents (CEPH) with Northern and Western European ancestry (CEU), and combined panels from Han Chinese in Beijing (CHB) and Japanese in Tokyo (JPT), respectively. For imputing genotypes from African-American subjects, we used a combined haplotype reference panel consisting of CEU and Yoruba in Ibadan, Nigeria (YRI). After genotype imputation, we filtered out SNPs with MAF < 0.01 and $r^2 < 0.4$, which resulted in 7,760,136, 6,640,213, and 12,721,435 common variants in European, East Asian, and African-American subjects, respectively.

Association mapping

The primary eQTL analysis was done separately for each cell type and population. We performed association of SNP genotype (coded as 0, 1, or 2) or imputed allele dosage (ranging from 0 to 2) with PCA adjusted expression residual phenotype (10-20 PCs depending on the population; fig. S20) and genetic ancestry PCs (1-2 PCs) using a non-parametric Spearman rank correlation (SRC) as previously described (22). For the *cis* analysis, we considered only SNPs in a +/- 1 Mb window around transcription start site (TSS) of each gene. We used 1 Mb window around TSS because the longest known human enhancer is located approximately 1 Mb away from the TSS. For *trans*-eQTLs, we analyzed SNPs that are located ≥ 1 Mb from the TSS of a target gene, or the target gene is on a different chromosome.

The nominal p-value for the test of association and Spearman's rho are reported for each SNP-gene pair for each population and cell type combination.

Multiple testing, permutations and FDR estimation for eQTL analysis

Significance of the nominal *p*-values was determined by comparing the distribution of the most significant *p*-values generated by permuting residual expression phenotypes 10,000 times independently for each gene as previously described (22-24). For each population in each cell type, we performed 10,000 permutations of expression phenotypes relative to genotypes. We keep the most significant SNP to gene association p-value from each permutation round. Based on the null p-value distribution, we take the most stringent p-value for each gene and keep the genes that passed Benjamini and Hochberg False Discovery Rate (FDR) of 5%.

To account for multiple testing for *trans*-association, we used a conservative Bonferroni-corrected *p*-value of 3×10^{-12} to account for 17.9M SNPs x 19,114 genes independent tests.

Meta-analysis of *cis*- and *trans*- eQTL associations

For meta-analysis, we used METASOFT (12) to perform meta-analysis using a random effects (RE) model. The effect size and standard error for each SNP-gene pair for each dataset were used as input to METASOFT. Three separate meta-analysis were performed across the three population for each cell type: (1) Fixed effects (FE) model based on inverse-variance-weighted effect size; (2) random effect model (RE), a conventional random effects model based on inverse-variance-weighted effect size; (3) Random effect model (RE2): new random effects model optimized to detect associations under heterogeneity. The RE2 model statistics is reported here.

To assess significance, we performed conditional permutation in a pooled (EA, EU, and AA) dataset. We performed 10,000 permutations for each gene by ensuring expression values from an individual of a given population were only assigned to another individual of the same population. As has been previously noted, conditional permutations masks inflated associations while revealing the relevant association signals (22-26).

Proportion of sharing and specificity of associations

Bayesian regression and hierarchical model. We used a Bayesian framework to jointly analyze data across cell types and population (10). The Bayesian statistical method integrates recently developed meta-analysis methods that allow for heterogeneity of effects among subgroups (i.e., cell types or population). The C++ package eQtlBma was used to run the Bayesian regression. We used “Step 3” tests for association in each subgroup separately and also in all subgroups jointly. The PCA adjusted residual phenotypes and the corresponding genotypes for each subgroup were used as input to eQtlBma. The settings used for the configuration files are: *step 3*, *outs*, *outraw*, *qnorm*, *maf 0.01*, *bfs all*, *mvrl*, *nperm 10000*, *trick*, and *pbf all*.

To estimate the extent of *cis*-eQTL sharing among groups (i.e., cell types and populations), we used a hierarchical model that borrows information across genes to estimate weights associated with different types of heterogeneity (10). To estimate the proportion of sharing between the two cell types, we ran eQtlBma separately for each population with a total of three possible configurations (CD4⁺ T-cells, CD14⁺CD16⁻ Monocytes, and shared). To estimate the proportion of sharing between the populations, we ran eQtlBma separately for each cell type with total of seven possible configurations (EU, AA, EA, EU-AA, EU-EA, EA-AA, EU-AA-EU). The genome-wide raw Bayes Factors (BF) for each subgroup were used to run the hierarchical model.

Proportion of true positives. To quantify eQTL sharing in a continuous way, we used the QVALUE software (R package qvalue 1.20.0) with the default settings (11). The QVALUE program takes a list of all the p-values and computes the proportion of eQTLs that are truly null (π_0). The method assumes that p-values of nulls will be uniformly distributed among [0,1], while truly significant tend to be close to zero. The quantity $\pi_1 = 1 - \pi_0$ estimates the proportion of true positives. The proportion of sharing between two

data sets (cell type or population) is reported as the proportion of π_1 estimated from the p -value distribution of eQTLs.

Stepwise regression

To determine whether there are independent *cis*-regulatory effects for each gene, we applied a stepwise regression model. For each gene that had a significant *cis*-eQTL at the permutation threshold of 0.01, we regressed out the effect on the expression levels of the effect of the most-significant SNP, then iteratively ran the Spearman Rank Correlation analysis on the remaining significant *cis*-eQTL SNPs using the resulting expression residuals (I). At each iteration, we retained the SNPs that were more significant than the gene's permutation p -value. This iterative process was repeated until no SNPs were found to be more significant than the gene's permutation p -value.

GWAS Catalog

We obtained data from published GWAS via the National Human Genome Research Institute (NHGRI) GWAS catalog available online at (<http://www.genome.gov/gwastudies>; April 2013 version). We used SNPs reported to be associated with complex traits with a p -value of at most 10^{-8} . To compile a list of independent GWAS SNPs, we performed LD pruning using PLINK with the following parameters: window size in SNPs = 100, number of SNPs to shift the window at each step = 5. The most significant SNP per each LD block was kept.

We manually curated GWAS SNPs for eleven autoimmune diseases including Ankylosing spondylitis (AS), Crohn's disease (CD), Ulcerative colitis (UC), Celiac disease (CeD), Multiple sclerosis (MS), Type 1 diabetes (T1D), Rheumatoid arthritis (RA), Primary biliary cirrhosis (PBC), Systemic lupus erythematosus (SLE), systemic sclerosis (SS) and Psoriasis (PS), and two neurodegenerative diseases, Alzheimer's disease (AD) and Parkinson's disease (PD). The manual curation of each GWAS SNP ensured that the reported SNPs were replicated and genome-wide significant in each respective GWAS.

Relative Trait Concordance Score

We used the Relative Trait Concordance (RTC) method to integrate QTL and GWAS data to detect disease-causing *cis*-regulatory effects as previously described in (15).

GWAS – eQTL Enrichment analysis

Trait-associated cis-eQTL enrichment relative to other SNPs. We performed simulations to test for an enrichment of *cis*-eQTLs among SNPs associated with complex traits relative to 10,000 SNP sets sampled at random from bins matched for minor allele frequency (MAF), number of LD partners ($r^2 > 0.8$), and distance from transcription start site (TSS).

Enrichment for cell type specificity for trait-associated cis-eQTLs. To assess the enrichment of disease-associated *cis*-eQTLs for cell type specificity, we selected 1,000 SNP sets sampled at random from bins matched for MAF, number of LD partners ($r^2 > 0.8$), distance from TSS and gene expression levels of the nearest *cis*-eQTL gene. For each random set of matched SNPs, we ran the Bayesian hierarchical model using the eQTLBma software. This generated a background distribution of proportion of *cis*-eQTLs that are CD4⁺ T-cell specific, CD14⁺CD16⁻ monocyte specific, and shared. For each trait group, we compared the null distribution of proportion of cell-specificity to the observed proportion of cell-specificity.

Graphics packages

Graphs were generated using R packages and ggplot2. The regional association plots were generated with Locus Zoom (27).

Supplementary Figures

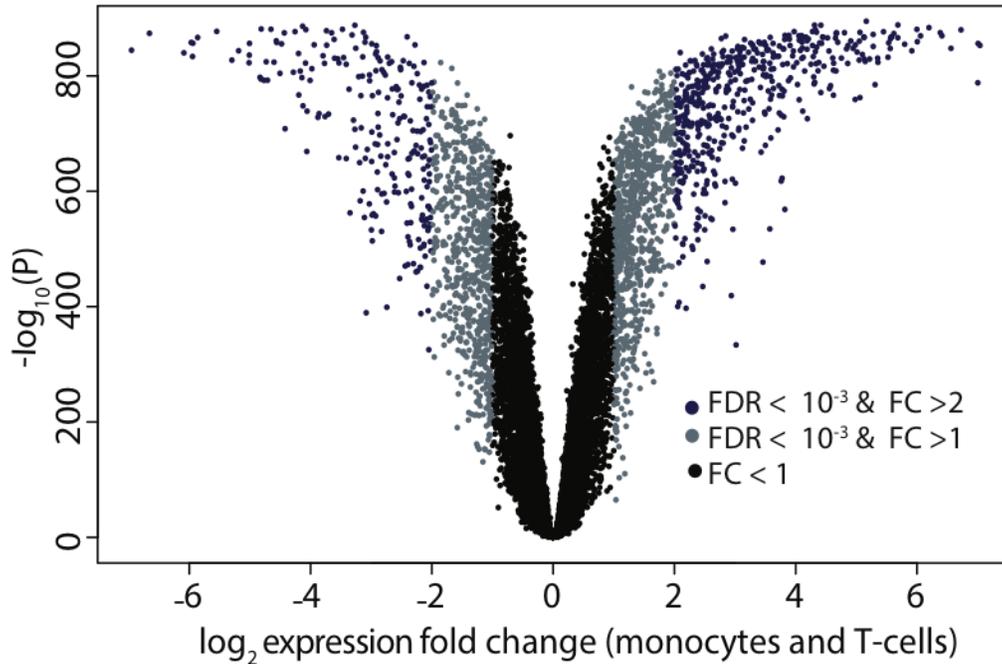


Figure S1: Differential gene expression between CD4⁺ T and Monocytes. Volcano plot showing differentially expressed genes between CD14⁺CD16⁻ monocytes and CD4⁺ T-cells in all three human populations. The FDR-adjusted $-\log_{10}(p\text{-values})$ test the null hypothesis that there is no difference in mean expression levels between monocytes and T-cells for each gene. The p -values (y-axis) are plotted against the \log_2 fold changes (FC) in expression (x-axis). There are 740 genes (out of 19,114) with $|\text{FC}| > 2$ and $\text{FDR} < 10^{-3}$ (in blue).

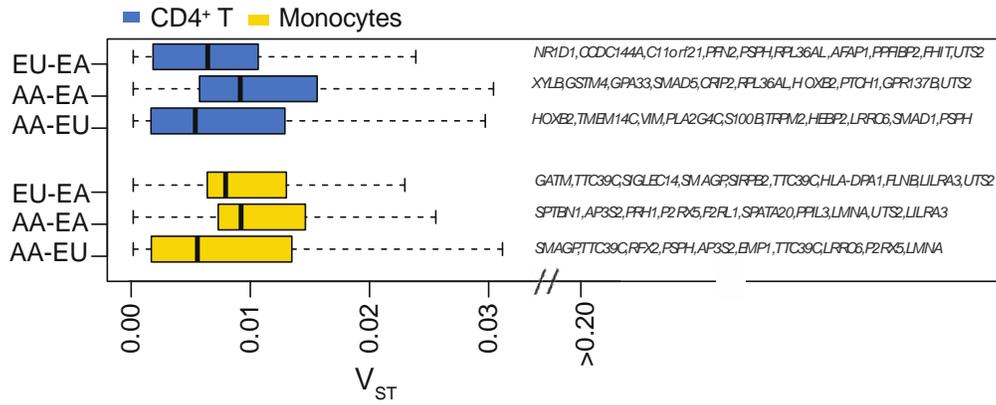


Figure S2: Transcriptome variation among human populations. V_{ST} distribution for each pairwise comparison between the African-American (AA), East Asian (EA) and European-American subjects (EU) in the two cell types. The top 1% of differentially expressed genes ($V_{ST} > 0.2$) are listed for each pairwise combination.

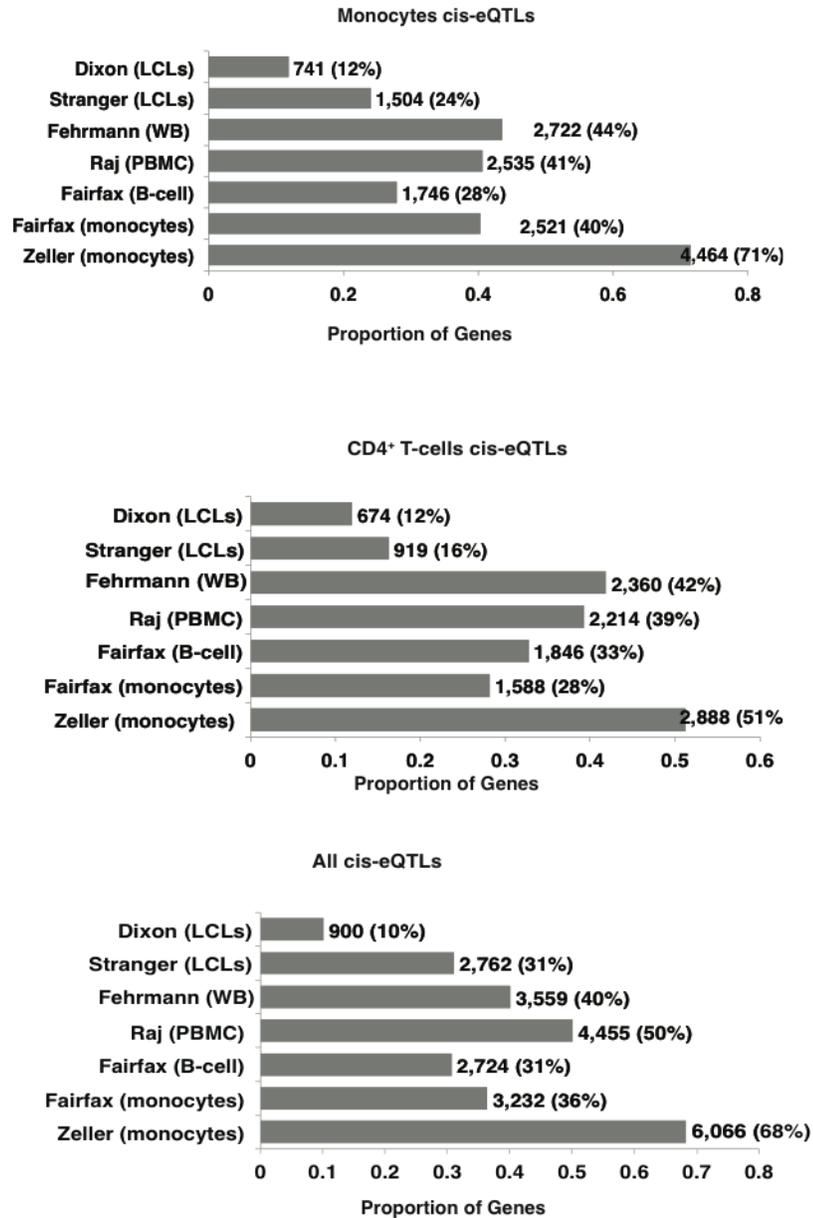


Figure S3: Proportion of cis-eQTL genes observed in this study with evidence of replication in previously published eQTL studies. The significant genes with cis-eQTLs were compared with seven publically available eQTL datasets (monocytes: Zeller et al. 2010 (28); Fairfax et al. 2012 (4); B-cells: Fairfax et al. 2012 (4); whole blood: Fehrmann et al. 2011 (29); lymphoblastoid cell lines (LCLs): Dixon et al. 2007 (30); Stranger et al. 2012 (1); peripheral blood mononuclear cells (PBMC): Raj et al. 2013 (31)). The interpretations of these results need to be carefully considered since the analysis methods, expression platforms, significance decisions, *cis* distance, etc, are all different in the various studies.

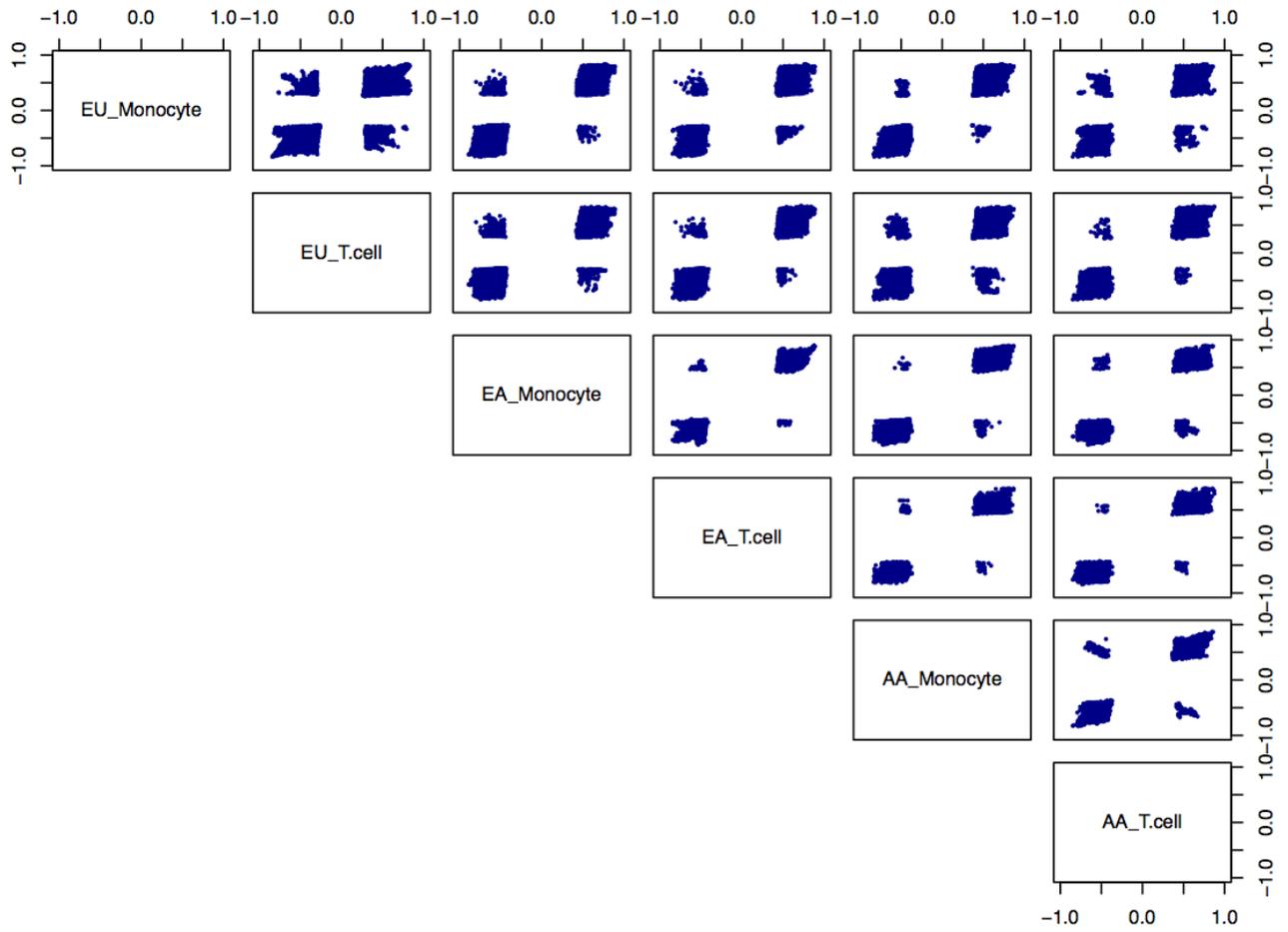


Figure S4: Direction of allelic effect of all significant cis-eQTLs. The direction of allelic effect (Spearman's rho) for all shared cis-eQTLs at FDR 0.05 across pairs of populations and cell types.

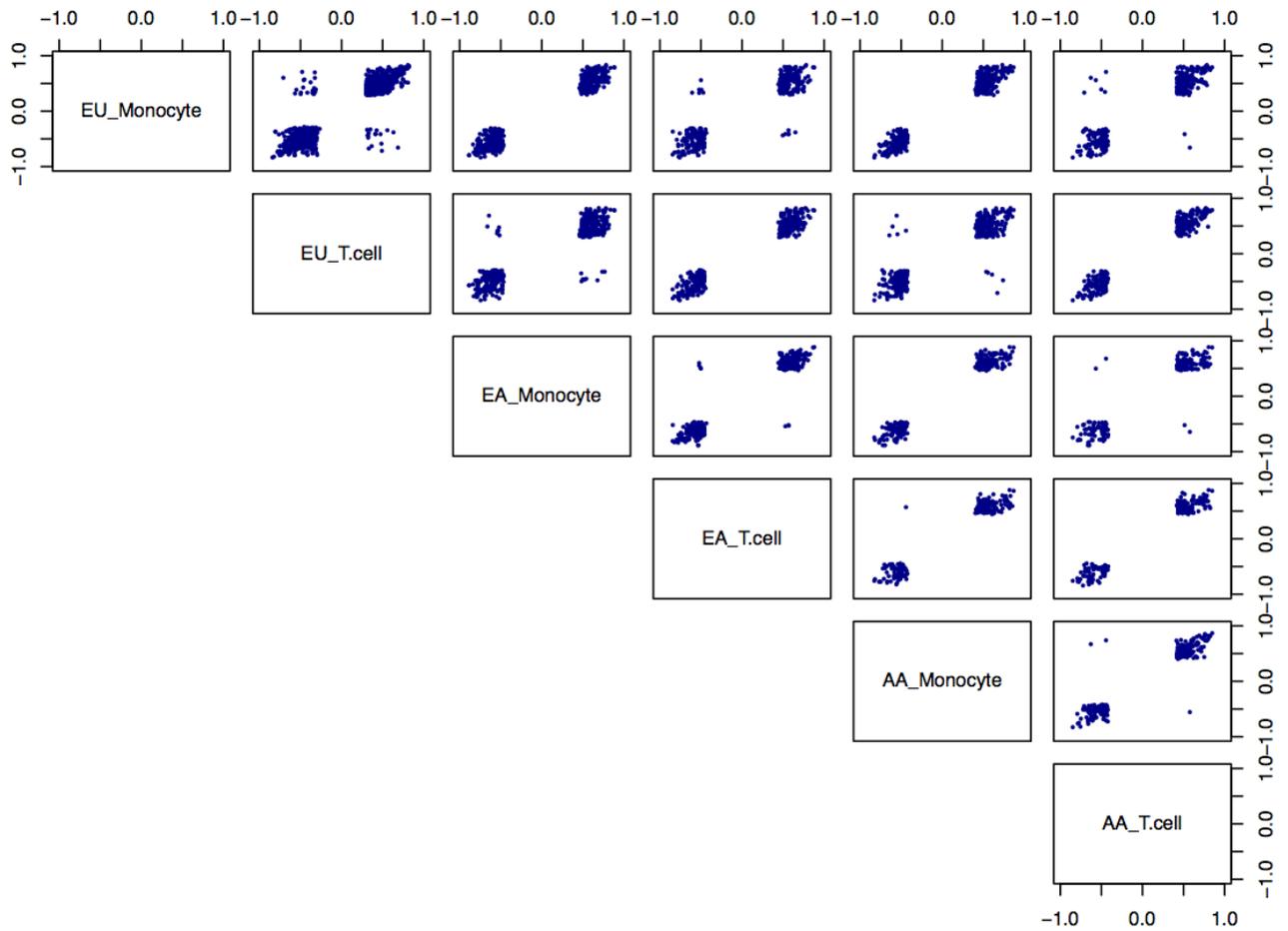


Figure S5: The direction of allelic effect (Spearman's rho) for the most significant SNP per gene. Pairwise Spearman's rho for shared cis-QTLs (top SNP per gene) across populations and cell types. When the top SNP per gene is shared across pairs of populations within the same cell type and the *cis*-eQTL is significant at FDR 0.05, we find that the allelic direction is 100% concordant, suggesting that the causal regulatory variation affects expression in the same direction across populations.

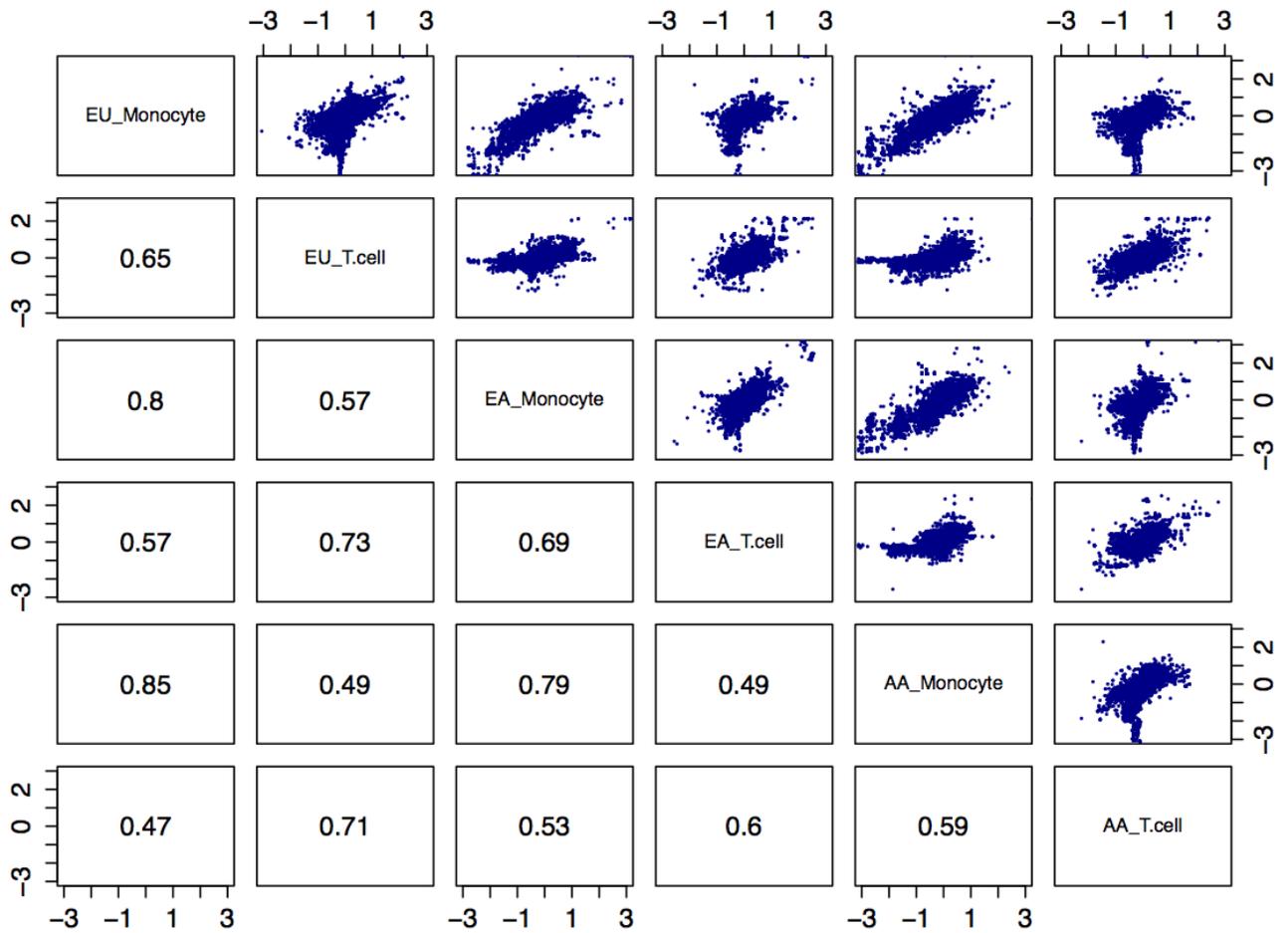


Figure S6: The effect size of shared *cis*-eQTLs at FDR 0.05 across populations and cell types. The effect size is quantified by the median expression level fold-change differences between major homozygote and heterozygote genotype groups. These plots show the correlation between effect sizes for each pairwise combination of populations and cell-types. The pairwise Pearson's correlation coefficients are shown on the lower left panel.

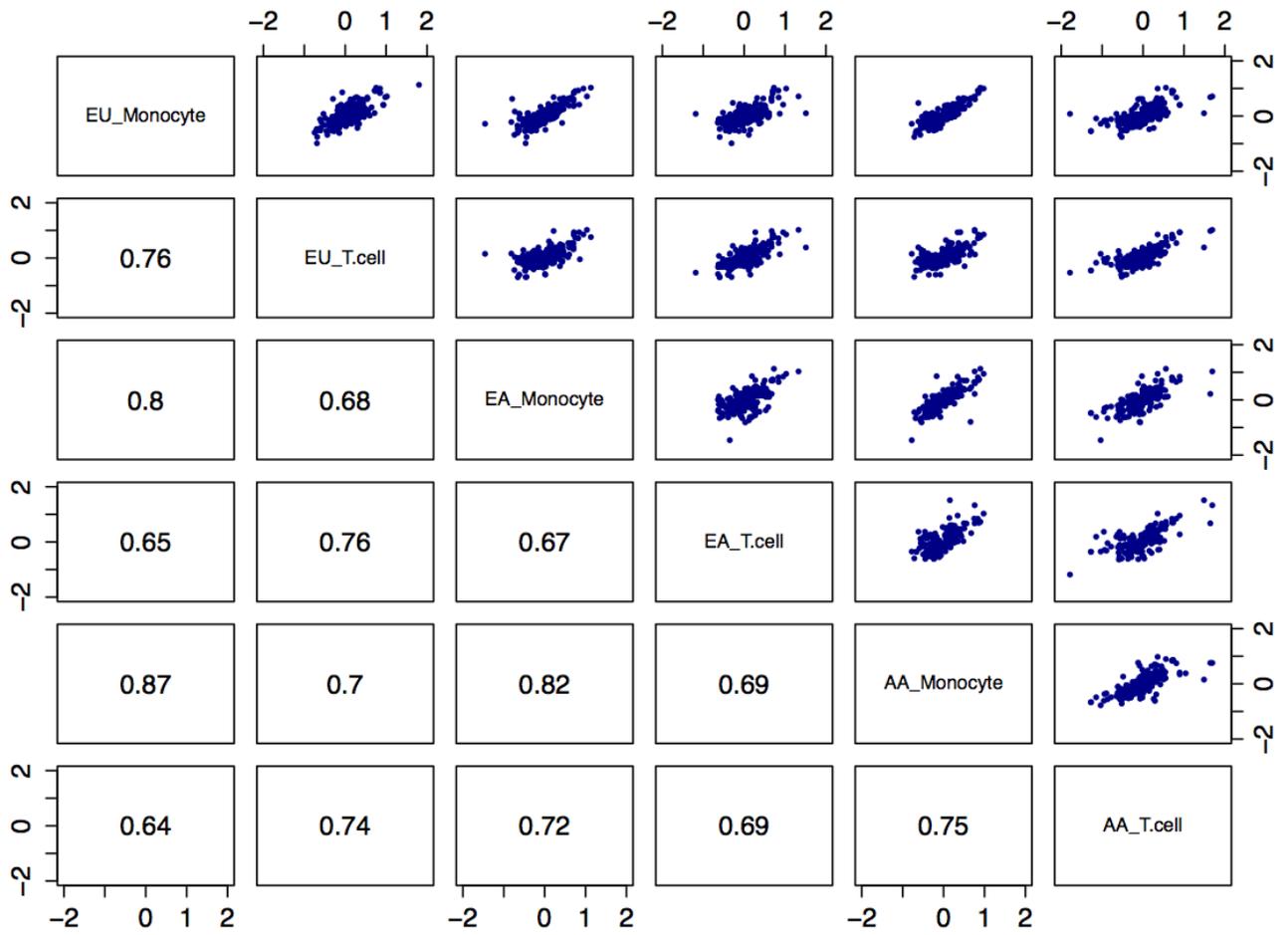


Figure S7: The effect size of shared cis-eQTLs (most significant SNP per gene) at FDR 0.05. The effect size is quantified by the median expression level fold-change differences between major homozygote and heterozygote genotype groups. The pairwise Pearson's correlation coefficients are shown on the lower left panel.

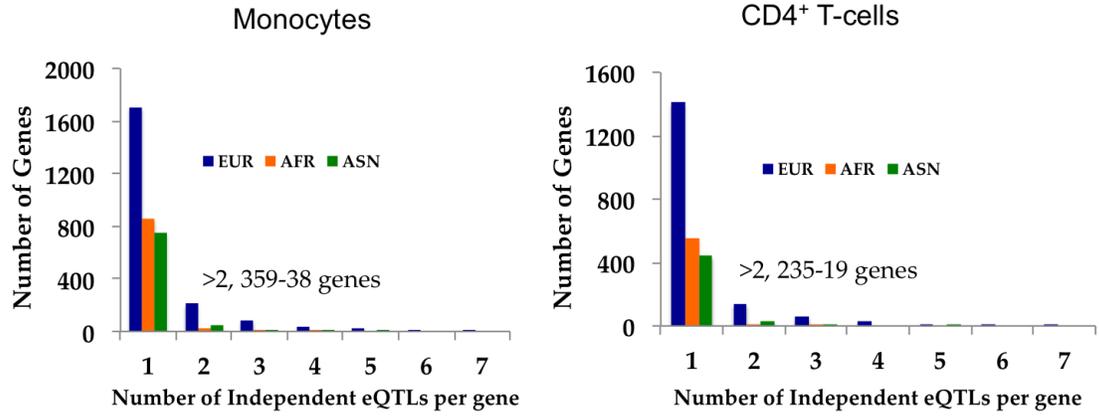


Figure S8: Number of *cis*-eQTL genes with independent *cis*-eQTL effects in the three populations. Up to 17% of genes with *cis*-eQTLs have multiple independent effects in the two cell-types.

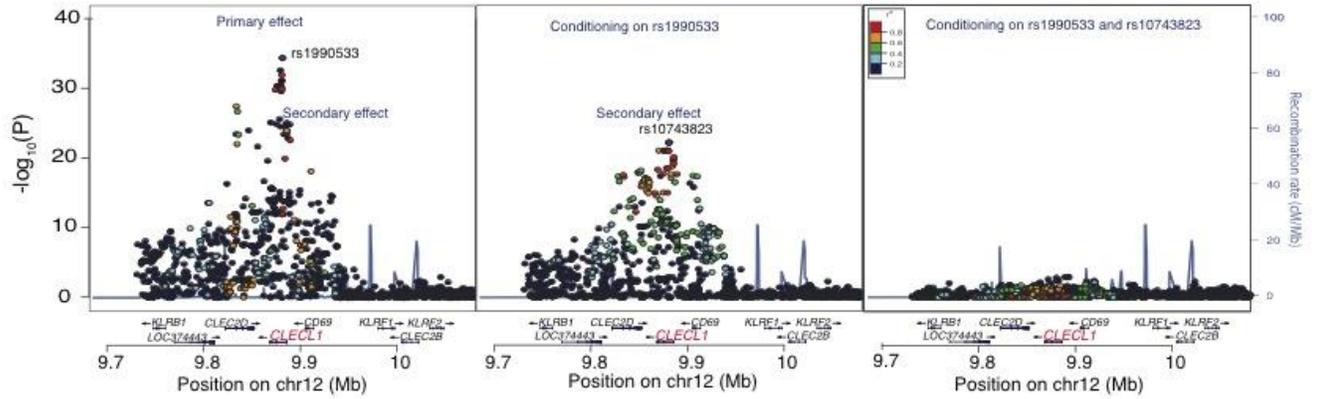


Figure S9: Independent cis-regulatory effects at *CLECL1* locus. Example of a *cis*-eQTL with two independent effects at *CLECL1* locus. *Left Panel:* Primary effect is rs1990533 ($P= 3.42 \times 10^{-35}$); *Middle panel:* Conditioning on the top SNP (rs1990533) revealed a secondary effect (rs10743823; $P= 5.06 \times 10^{-23}$); *Right Panel:* No additional effects are observed after conditioning on the two independent effects (rs1990533 and rs10743823).

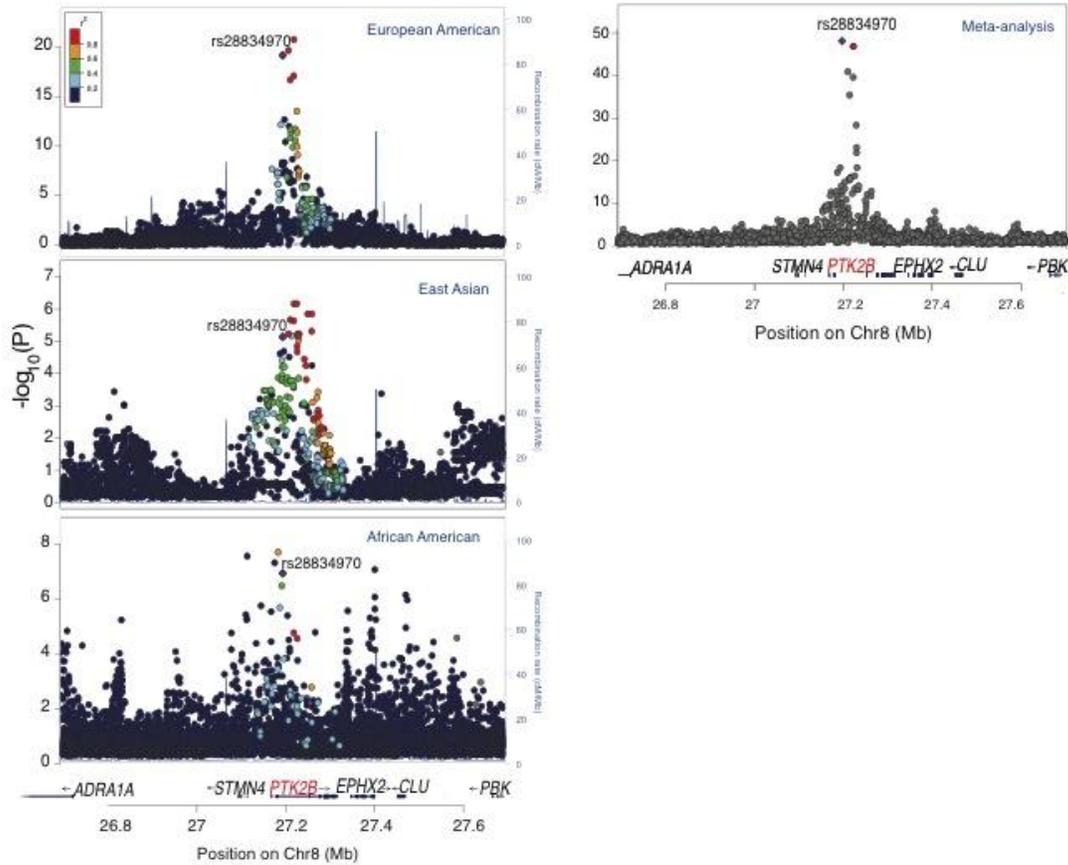


Figure S10: Regional association plot for a monocyte-specific *cis*-eQTL shared across the three populations. *Left panel:* Association signal at the *PTK2B* locus in EU (top), EA (middle) and AA subjects (bottom). *Right panel:* Cross population meta-analysis substantially narrows down the number candidate functional variants in the region. With meta-analysis of the three populations, the number of significant variants distills to just two SNPs (rs28834970 $P_{\text{meta-monocytes}}=7.59 \times 10^{-49}$; rs17057043 $P_{\text{meta-monocytes}}=1.40 \times 10^{-47}$; EU $r^2=0.93$), one of which is an Alzheimer's disease GWAS index SNP (rs28834970).

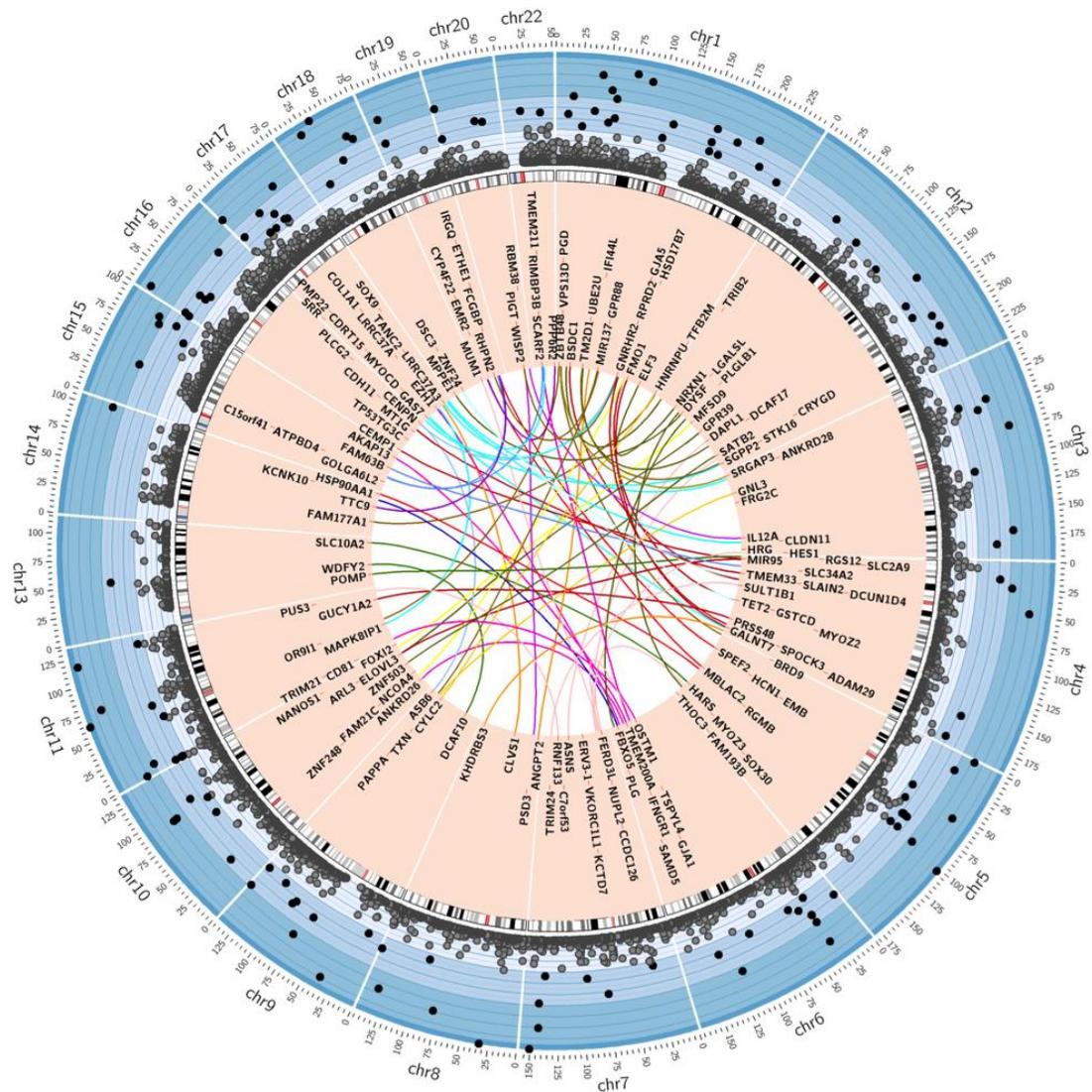


Figure S11: Circos plots for significant *trans*-eQTLs in monocytes. The outermost rim shows a Manhattan plot for *cis*-eQTLs (top SNP per gene in meta-analysis) in monocytes and the second (inner) rim shows significant *trans* eQTLs ($P_{\text{META}} < 3 \times 10^{-12}$). The lines (colored by chromosomes) connect regulatory SNP (shown as *cis*-gene if significant or the closest gene) and their *trans*-regulated genes.

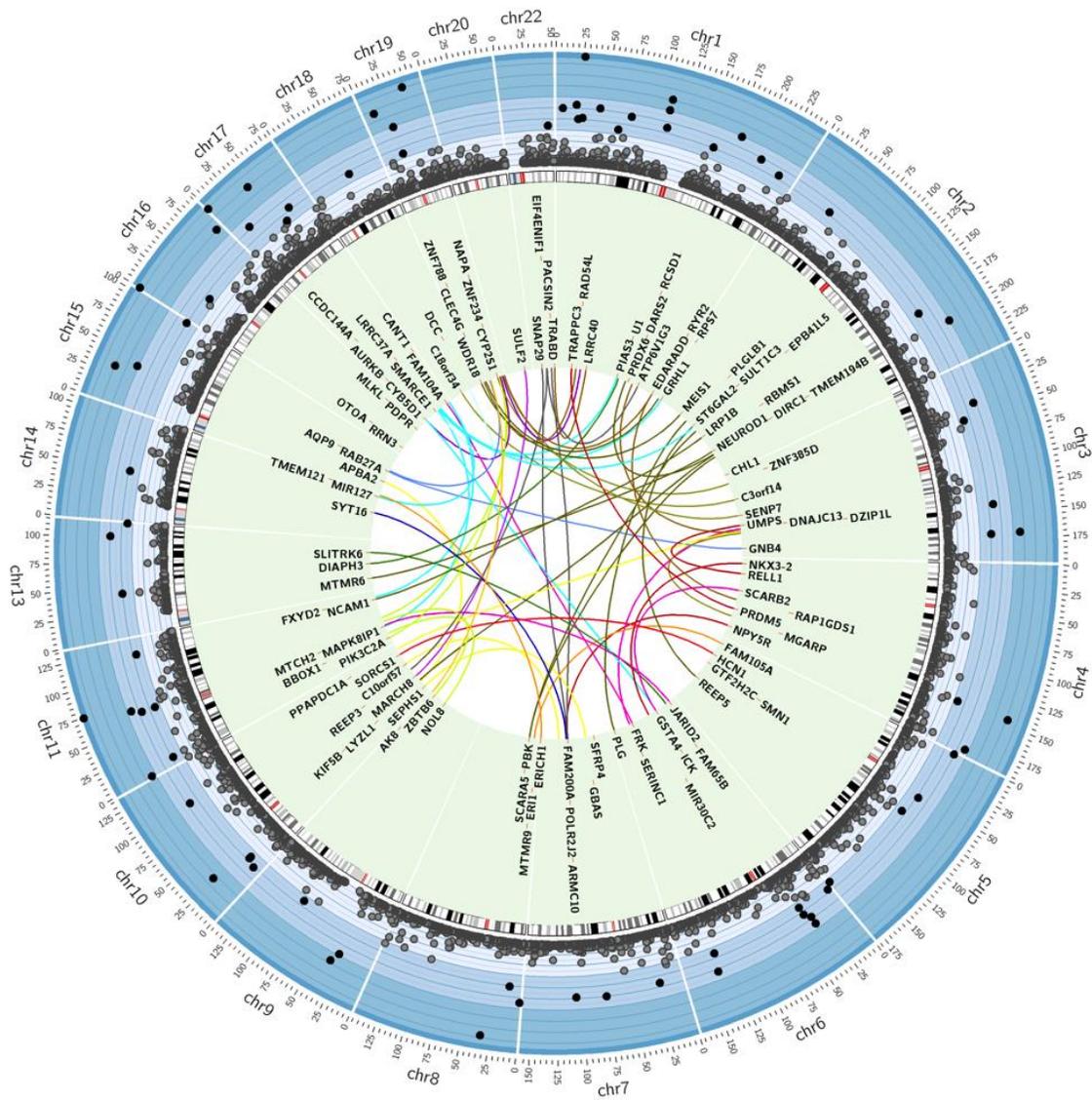


Figure S12: Circos plots for significant *trans*-eQTLs in CD4⁺ T-cells. The outermost rim shows a Manhattan plot for *cis*-eQTLs (top SNP per gene in meta-analysis) in CD4⁺ T-cells and the second (inner) rim shows significant *trans* eQTLs ($P_{META} < 3 \times 10^{-12}$). The lines (colored by chromosomes) connect regulatory SNP (shown as *cis*-gene if significant or the closest gene) and their *trans*-regulated genes.

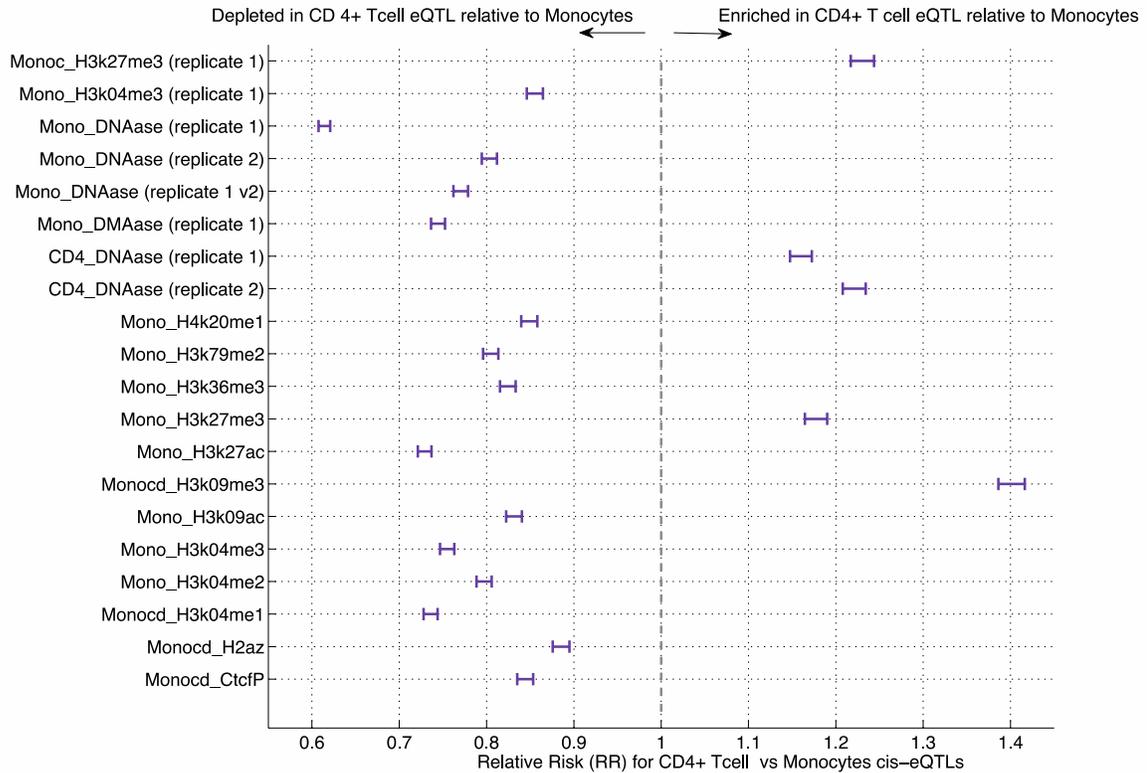


Figure S13. Relative enrichment of CD4+ T cell *cis*-eQTLs compared to Monocytes *cis*-eQTLs in regulatory marks derived in the two cell types. Twenty genomic features (regulatory marks) were identified from ENCODE and Epigenome Roadmap that matched the cell types studied here (Monocyte or a CD4+ T cell cell type) (listed in the table below). Note that only DNAase hypersensitivity data was available for CD4+ T cell. For each of these genomic features, the relative enrichment of CD4+ T cell -specific versus Mono-specific *cis*-eQTLs was quantified by computing the Relative Risk ($RR = \frac{p(feature|cd4+ T eQTL)}{p(feature|Mono eQTL)}$). Error bars mark the 99.7% CI (to account for multiple testing; 20 features tested).

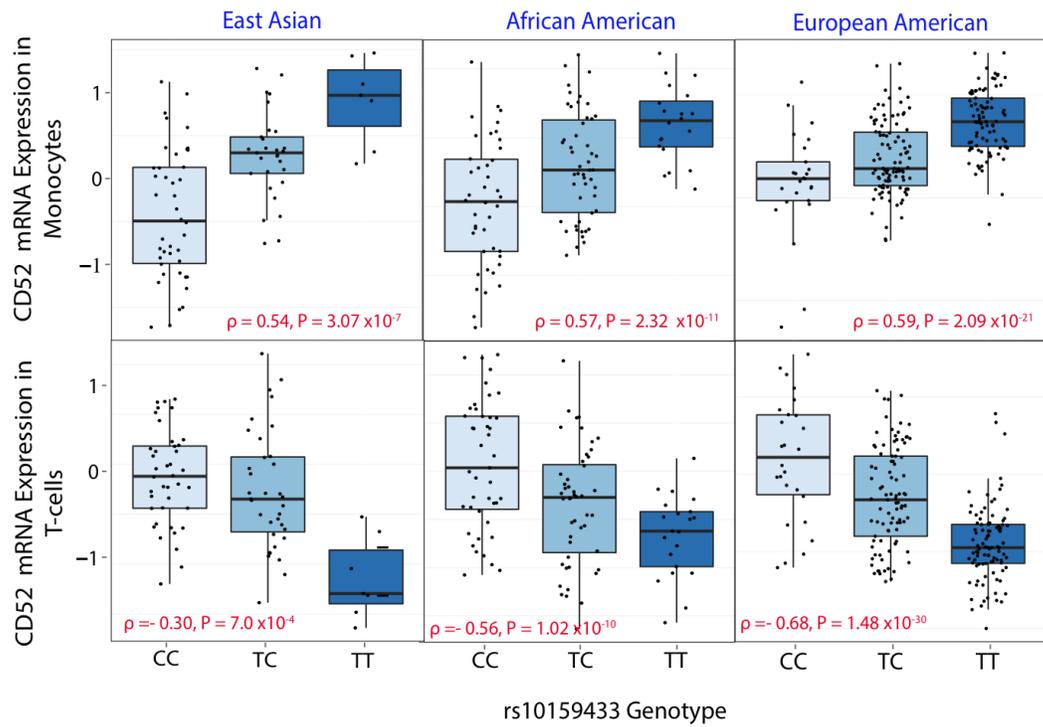


Figure S14: An example of discordant cis-eQTL (CD52-rs10159433). The opposite allelic effects are replicated across the three populations.

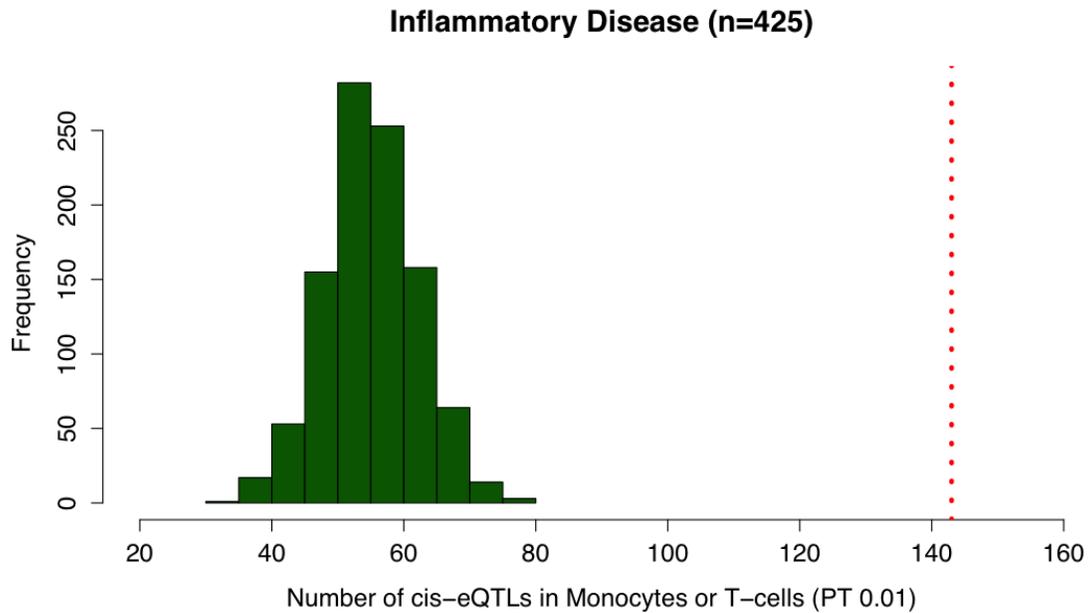


Figure S15: Inflammatory disease associated GWAS SNPs are more likely to be *cis*-eQTLs than randomly matched sets of SNPs. Among the 425 SNPs associated with at least one of eleven inflammatory diseases (NIH GWAS catalog, April 2013; LD-pruned to $r^2 > 0.4$), we find that 143 have significant *cis*-eQTL effects on 182 genes in monocytes and/or T-cells. Permuting 10,000 SNP sets sampled at random from bins matched for minor allele frequency (MAF), number of LD partners ($r^2 > 0.8$), and distance from transcription start site (TSS), we find that the p -value of this enrichment is $< 1 \times 10^{-4}$.

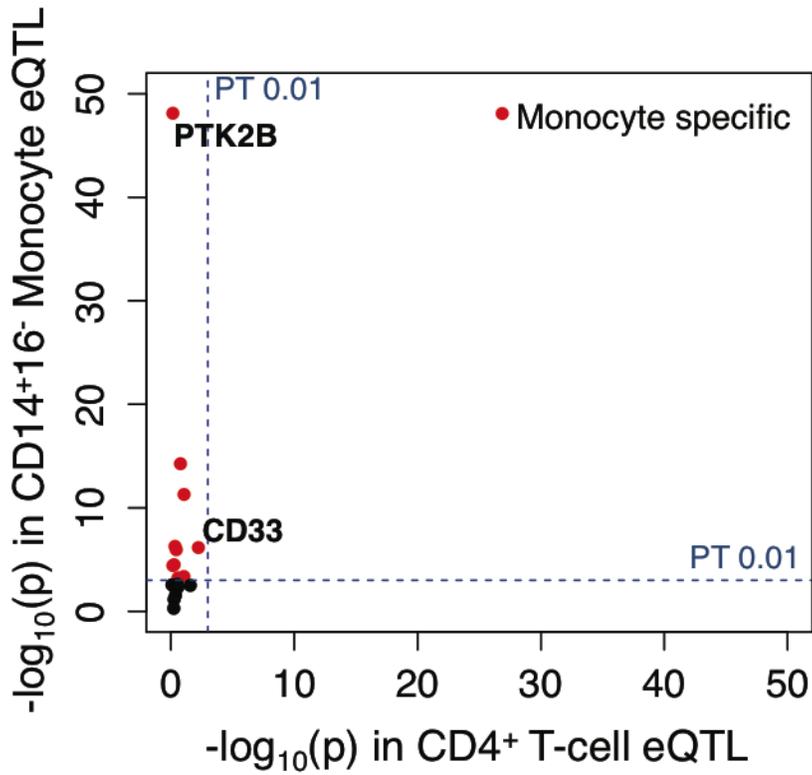


Figure S16: eQTL association signal for Alzheimer’s disease associated SNPs in monocytes and T-cells. Shown here are *cis*-eQTL $-\log_{10}(P\text{-value})$ of all AD GWAS SNPs. The AD susceptibility SNPs have significant *cis*-eQTL effects only in monocytes.

Alzheimer's Disease GWAS $P < 10^{-4}$

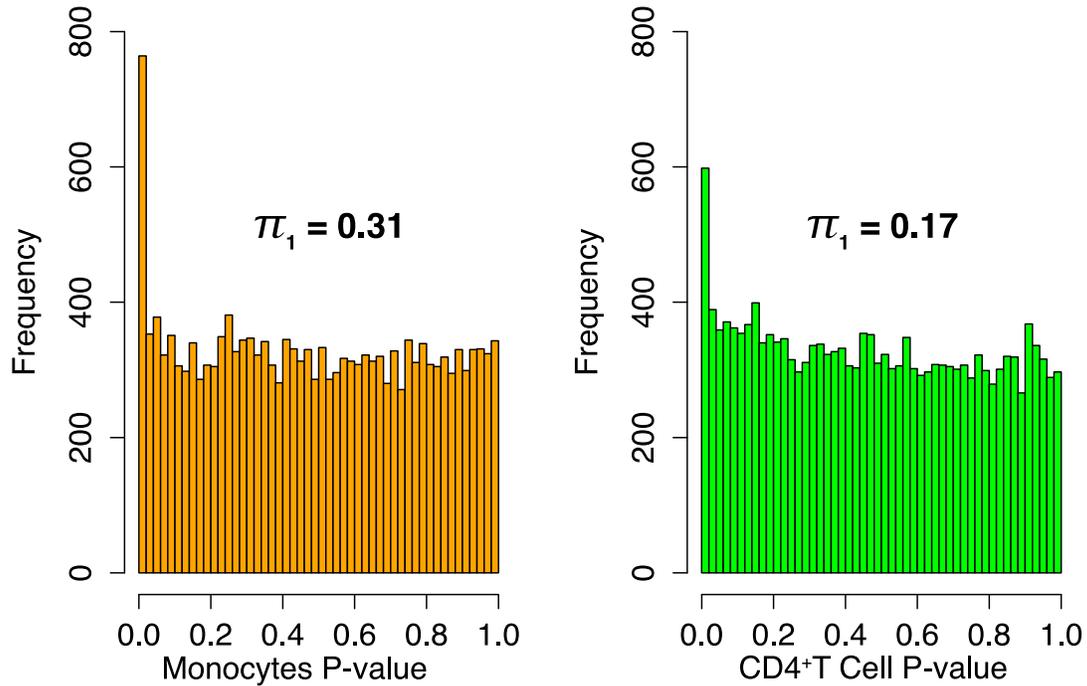


Figure S17: Enrichment of *cis*-eQTLs in Alzheimer's disease GWAS. Left: Distribution of *cis*-eQTL p-values for AD GWAS SNPs $p < 10^{-4}$ (Naj et al. 2010) in Monocytes. Right: Distribution of *cis*-eQTL p-values for AD GWAS SNPs $p < 10^{-4}$ in T-cells. Storey's Π_1 is listed in each figure. Π_1 is higher in monocytes compared to T-cells, suggesting that when considering non-significant GWAS $P < 4$ (SNPs below GWAS threshold of 5×10^{-8}) we observe over-representation of monocyte specific *cis*-eQTLs.

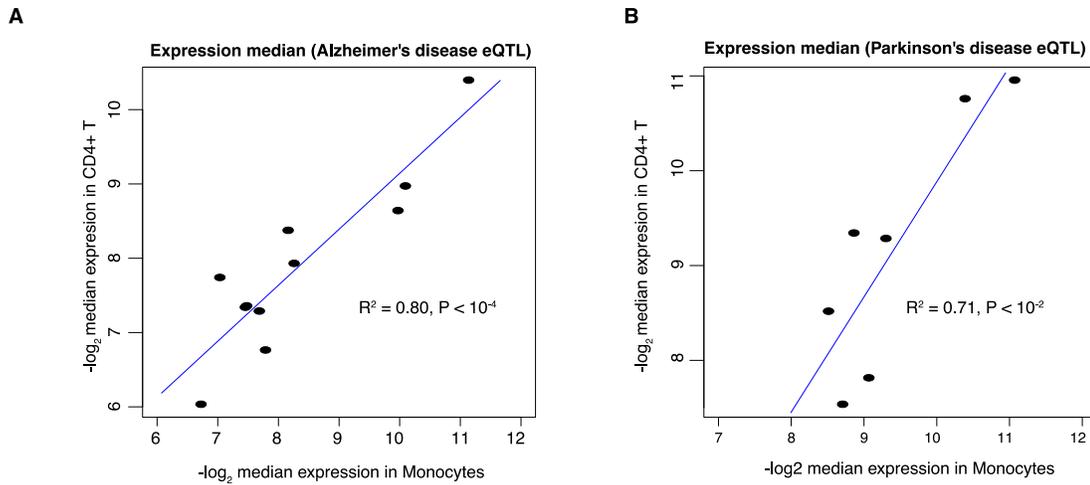
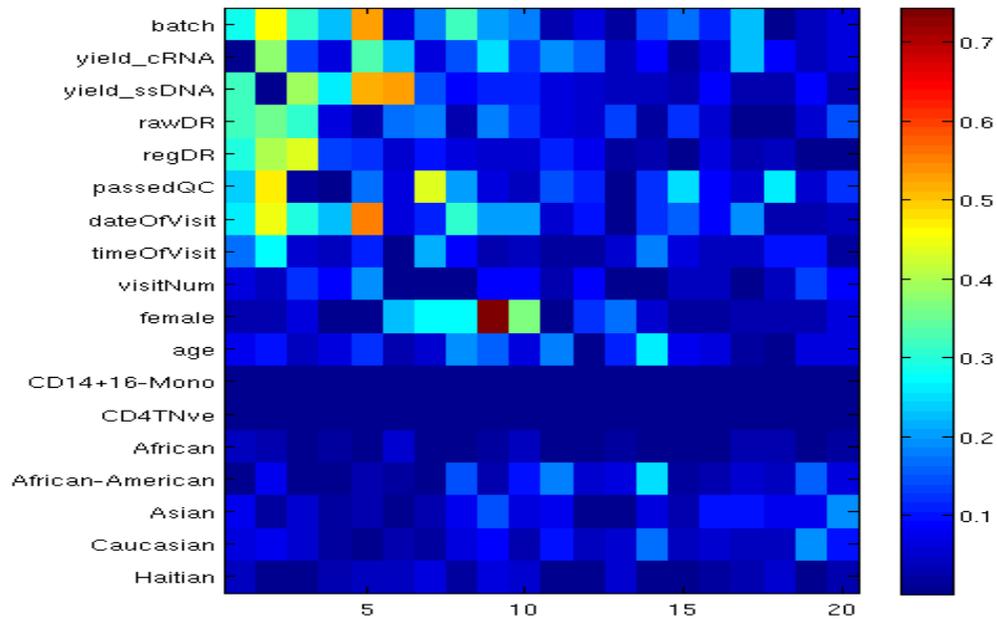


Figure S18: Comparison of Monocytes (x-axis) versus CD4+ T-cells (y-axis) expression medians. \log_2 expression medians were compared across the cell types for (A) Alzheimer's disease eQTL-genes (B) Parkinson's disease eQTL-genes. With the exception of AD-associated *cis*-eQTL gene *CD33* (expression medians= 11.5 in Monocytes, 6.2 in T-cells), and PD-associated *cis*-eQTL gene *LRRK2* (expression medians= 10.65 in Monocytes, 3.73 in T-cells), the correlation between \log_2 expression levels in monocytes and T-cells was very high ($r^2 = 0.80$ and $p < 10^{-4}$ for AD *cis*-eQTL genes; $r^2 = 0.71$, $p < 10^{-2}$ for PD *cis*-eQTL genes), suggesting that expression level differences between cells are not the primary reasons for observed monocyte-specific eQTL discovery.



Principal Components (# of factors)

Figure S19: Principal component analysis (PCA) of known non-genetic factors affecting gene expression levels. Shown here is a heatmap of Pearson's r for each eigenvalues associated with each component or factors. As expected, batch effects, technical artifacts, age and gender has significant effect on gene expression.

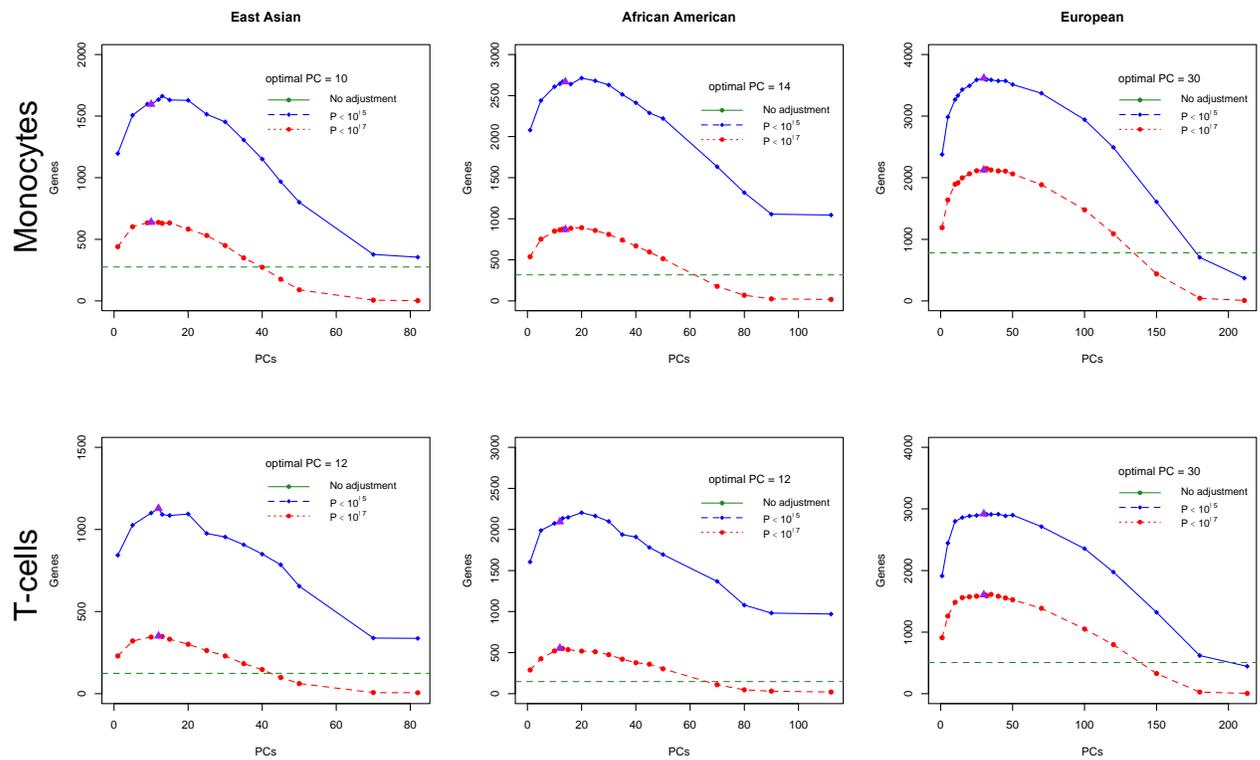


Figure S20: Principal component analysis corrects expression data for non-genetic factors. To account for non-genetic factors such as batch effects, age, gender, technical artifacts in gene expression data, we used Principal Component Analysis (PCA) as in (4). PCs were estimated separately from the gene expression matrix for each population and cell-type. For association analysis, we used the number of PCs that maximized the number of significant *cis*-eQTLs at nominal significance thresholds of p -value $< 10^{-5}$ and p -value $< 10^{-7}$. This procedure identified 20, 10, and 14 PCs in EU, EA, and AA monocytes and 20, 12, and 12 PCs in EU, EA, and AA T-cells (optimal number of PCs are shown as purple triangle in each plot). These factors were regressed out of the gene expression data before running the association tests.

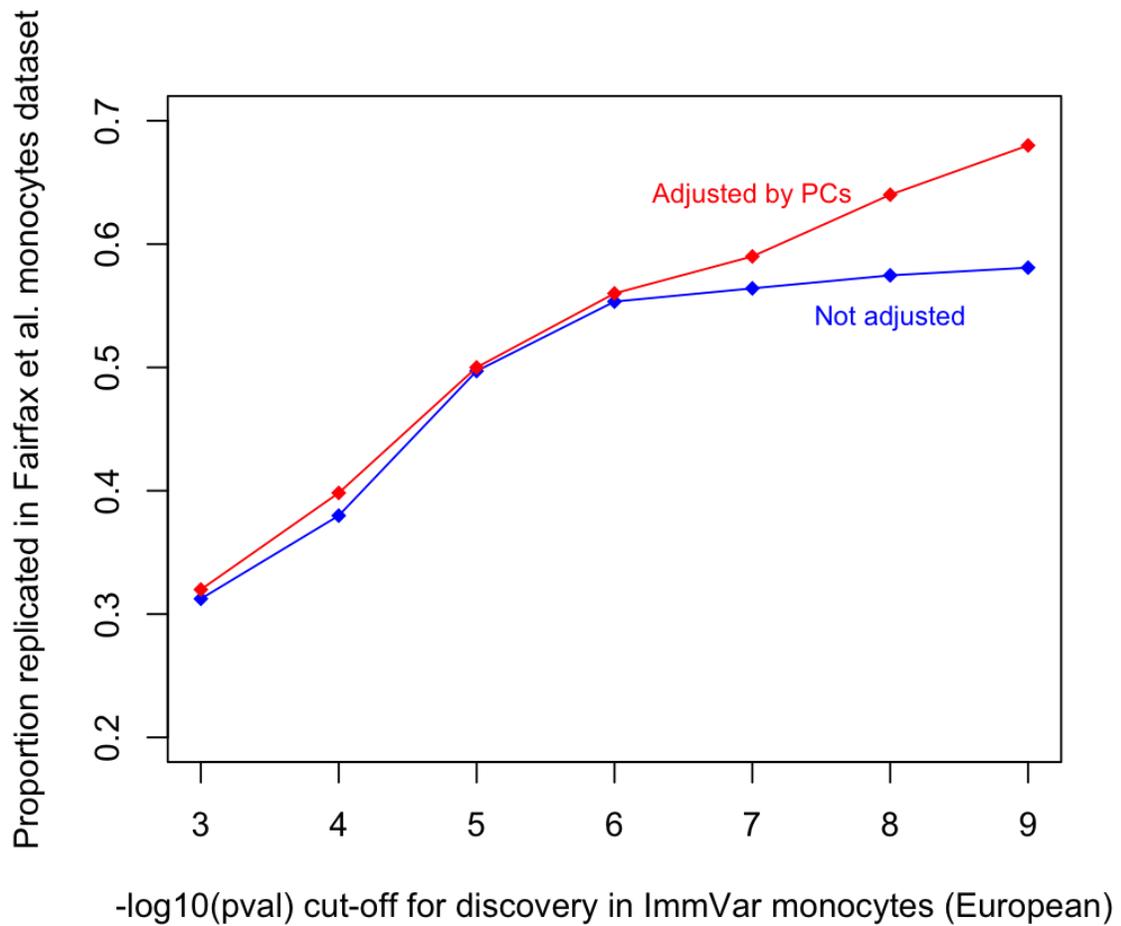


Figure S21: PCA covariate analysis improves replication of *cis*-eQTLs. Shown here are proportions of EU monocyte *cis*-eQTLs replicated in the Fairfax et al. 2012 (18) in different *p*-value threshold. The improvement of replication rate for *cis*-eQTLs using PC-adjusted data has been previously shown in Stranger et al. 2012 (14).

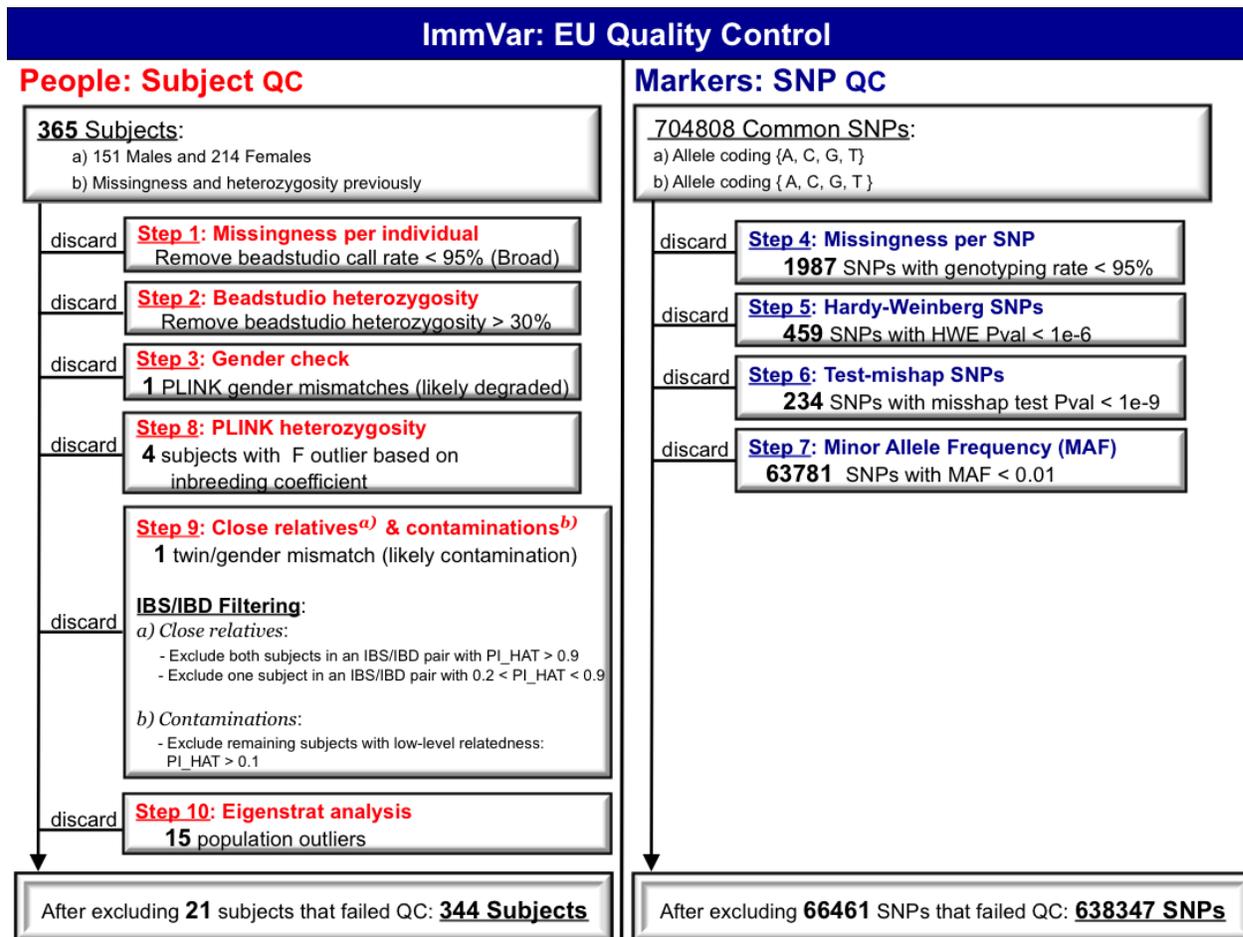
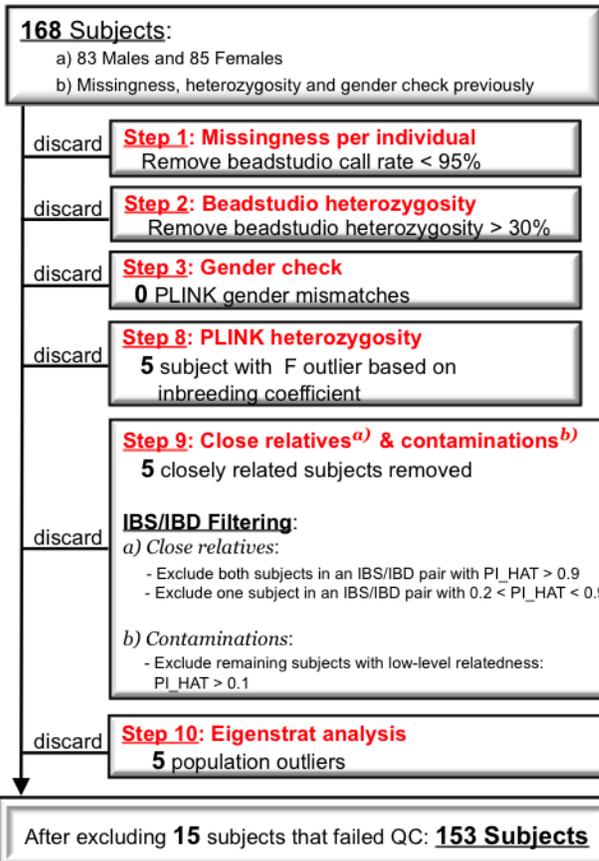


Figure S22: European-American genotyping quality control. See methods for a complete description of our quality control procedures.

ImmVar: AA Quality Control

People: Subject QC



Markers: SNP QC

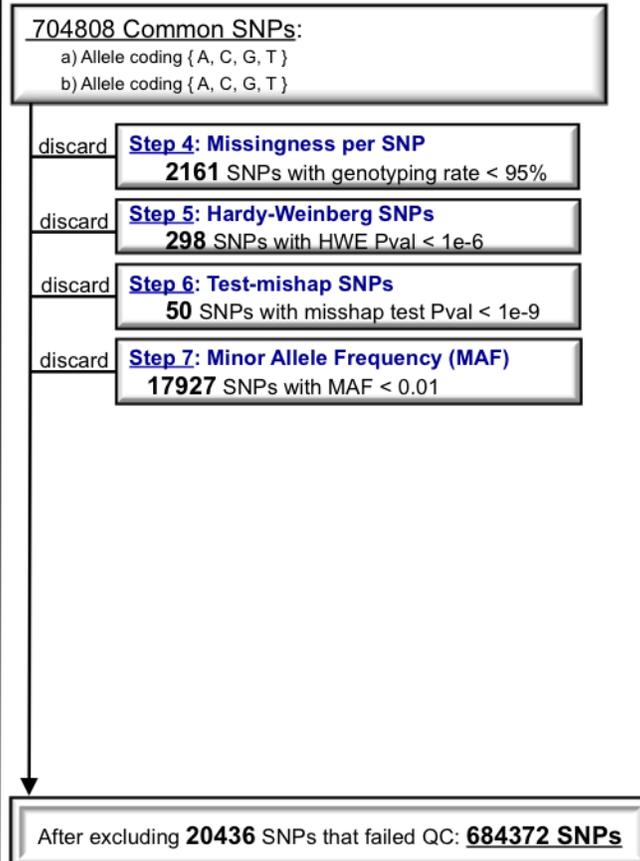


Figure S23: African-American genotyping quality control. See methods for a complete description of our quality control procedures.

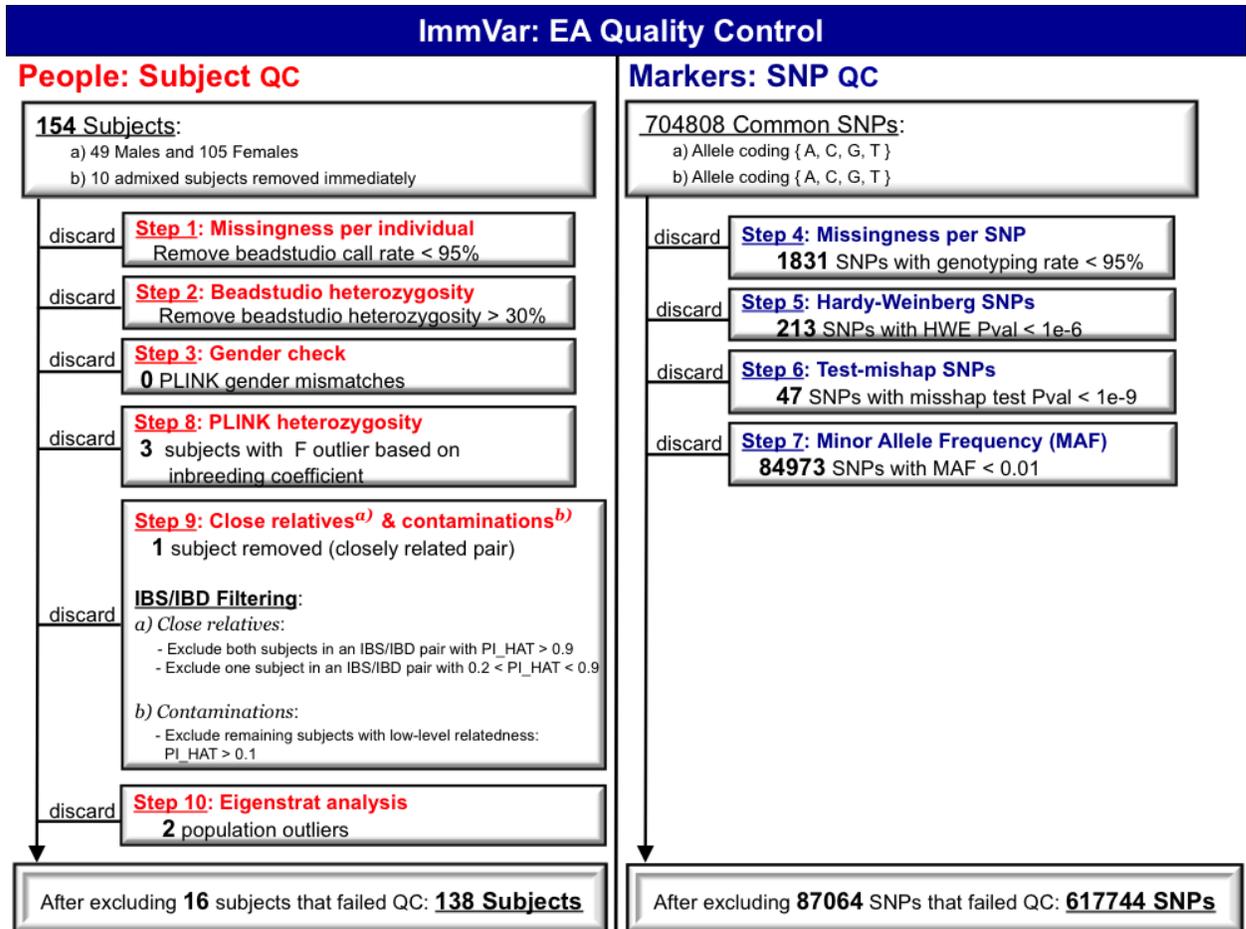


Figure S24: East Asian genotyping quality control. See methods for a complete description of our quality control procedures.

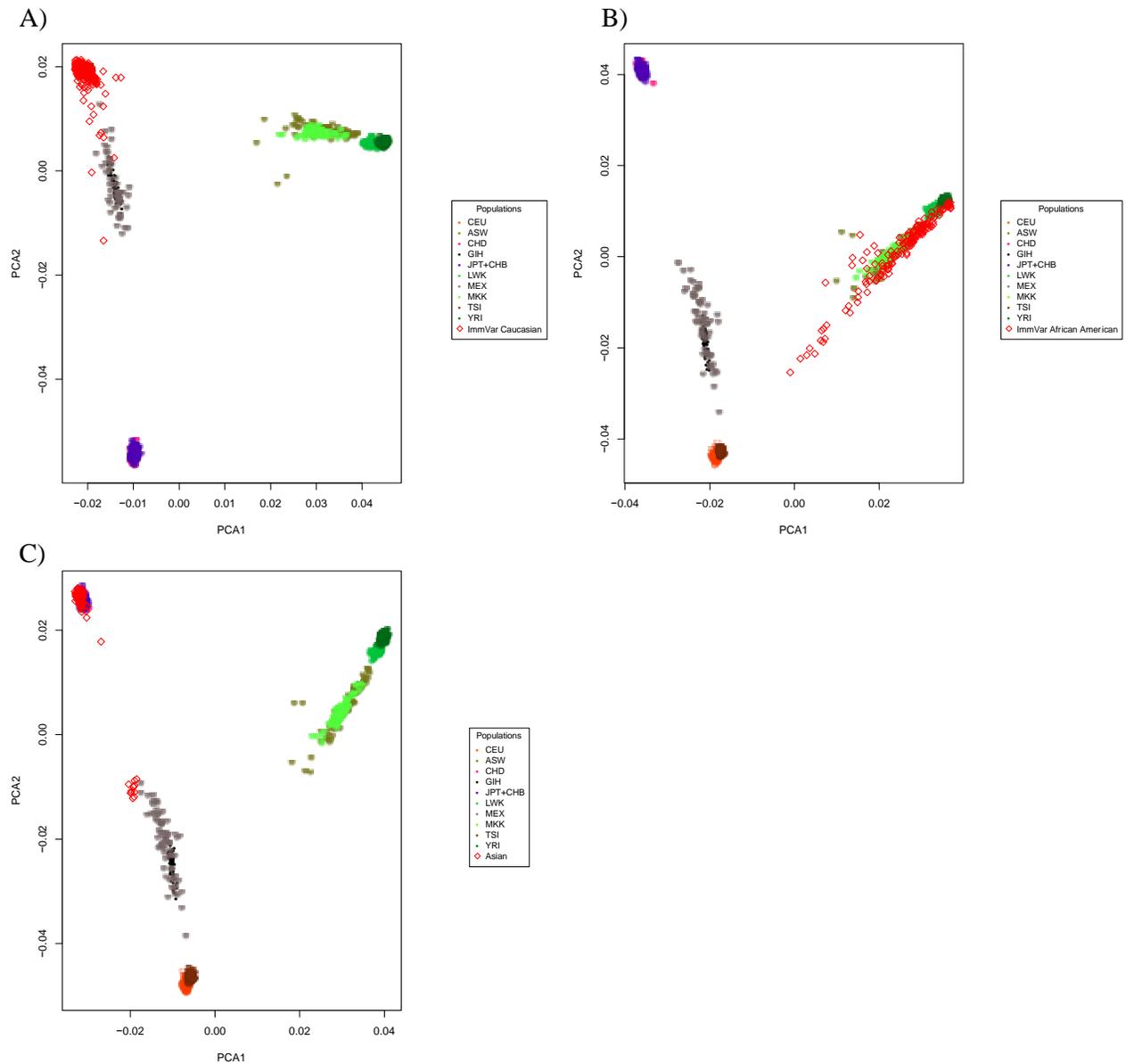


Figure S25: Principal component analysis (PCA) of genotypes to identify population outliers and correct for population stratification. EIGENSTRAT analysis (5) was performed on the (a) EU, (B) AA, and (C) EA cohorts using a subset of 100,000 SNPs randomly selected from all LD-pruned ($r^2 < 0.1$ at a window size of 100bp and step of 1 SNP) genotyped SNPs with MAF > 0.05. Population outliers were identified by combining ImmVar genotypes with HapMap Phase III genotypes from the CEU, YRI, CBT, JPT and GIH populations. After excluding an admixed cluster with East Asian and European ancestry, we used an iterative outlier detection approach, smartpca with a standard deviation of 6, to remove an additional 16 ancestry outliers.

Supplementary Tables

See individual tab-delimited text files for Tables S4-S9, S11-S14, S16-S17 and S19.

Table S1: Demographic characteristics of subjects used in this study. The healthy subjects were drawn from the PhenoGenetic Project at Brigham & Women's Hospital, Boston, MA. Total number individuals analyzed in this study are 461, of which 401 have monocyte data and 407 have CD4+ T cell data, respectively.

	# Post-QC Subjects	Gender	Median Age	Median BMI
Monocytes				
African-American (AA)	112	M: 59 F: 53	35	27.1
European-American (EU)	211	M:91 F:120	26	25
East Asian (EA)	78	M: 24 F: 54	25	23.5
CD4+ T-cells				
African-American (AA)	112	M:55 F: 57	35	27.5
European-American (EU)	213	M: 91 F: 122	26	25
East Asian (EA)	82	M: 25 F: 57	24	23.2

Table S2: Highly differentiated genes ($V_{ST} > 0.2$; top 1% of all V_{ST} scores) with respect to expression in monocytes between human population pairs.

Monocytes	Vst		
	EU-AA	EU-EA	AA-EA
UTS2	0.033	0.550	0.433
FLNB	0.081	0.328	0.122
SPTBN1	0.014	0.219	0.303
SMAGP	0.253	0.260	0.002
EMP1	0.345	0.171	0.072
TTC39C	0.352	0.271	0.008
PSPH	0.325	0.248	0.004
SPATA20	0.127	0.091	0.345
PPIL3	0.190	0.096	0.366
F2RL1	0.096	0.071	0.345
LRRC6	0.385	0.096	0.182
RFX2	0.297	0.044	0.139
LMNA	0.460	0.009	0.430
P2RX5	0.408	0.018	0.331
PRH1	0.081	0.078	0.328
SIGLEC14	0.042	0.257	0.118

GATM	0.006	0.248	0.285
LILRA3	0.071	0.368	0.488

Table S3: Highly differentiated genes ($V_{ST} > 0.2$; top 1% of all V_{ST} scores) with respect to expression in CD4⁺ T-cell between human population pairs.

CD4 ⁺ T	Vst		
	EU-AA	EU-EA	AA-EA
UTS2	0.043	0.506	0.324
CRIP2	0.104	0.078	0.278
NR1D1	0.060	0.166	0.043
C11orf21	0.015	0.173	0.103
VIM	0.170	0.005	0.217
TRPM2	0.197	0.006	0.146
TMEM14C	0.170	0.000	0.184
RPL36AL	0.004	0.219	0.292
PTCH1	0.160	0.018	0.304
PSPH	0.285	0.209	0.012
HOXB2	0.169	0.029	0.302
HEBP2	0.214	0.005	0.144
GPR137B	0.104	0.086	0.316
AFAP1	0.035	0.241	0.188
CCDC144A	0.095	0.245	0.081
FHIT	0.048	0.342	0.190
PPFIBP2	0.006	0.264	0.237
GSTM4	0.028	0.122	0.247

Table S4-S6: Significant *cis*-eQTLs at FDR 0.05 in CD4⁺T-cells of European-American, East Asian and African-American subjects. All significant SNPs (within 1MB of TSS of a gene), SNP chromosome, gene chromosome, physical position of each SNP, TSS of each gene (largest transcript), distance from SNP to TSS, Spearman's rho, $-\log_{10}$ (P-value) and P-value are reported.

Table S7-S9: Significant *cis*-eQTLs at FDR 0.05 in monocytes of European-American, East Asian and African-American subjects. All significant SNPs (within 1MB of TSS of a gene), SNP chromosome, gene chromosome, physical position of each SNP, TSS of each gene (largest transcript), distance from SNP to TSS, Spearman's rho, $-\log_{10}$ (P-value) and P-value are reported.

Table S10: The proportion of true positives estimated from enrichment of low p-values (π_1) (13) between pairs of human population. We observe a high degree of pairwise population sharing of *cis*-eQTLs (70%-90%).

Reference	Secondary	Monocytes (π_1)	CD4⁺ T-cells (π_1)
EU	AA	0.70	0.73
	EA	0.74	0.72
AA	EU	0.93	0.91
	EA	0.86	0.84
EA	EU	0.89	0.84
	AA	0.76	0.71

Table S11: Meta-analysis *cis*-eQTLs at FDR 0.05 in Monocytes. The best *cis*-SNP associated with each gene expression is reported, including effect size estimates, standard error of the effect size and P-values.

Table S12: Meta-analysis *cis*-eQTLs at FDR 0.05 in CD4⁺ T-cells. The best *cis*-SNP associated with each gene expression is reported, including effect size estimates, standard error of the effect size and P-values.

Table S13: Significant *trans*-eQTLs in monocytes from meta-analysis. All SNP-gene pairs at Bonferroni corrected P-value is reported, including effect size estimates, standard error of the effect size and P-values.

Table S14: Significant *trans*-eQTLs in CD4+T from meta-analysis. All SNP-gene pairs at Bonferroni corrected P-value is reported, including effect size estimates, standard error of the effect size and P-values.

Table S15: The proportion of true positives estimated from enrichment of low p-values (π_1) (13) between the cell types.

		SRC analysis π_1		
Reference	Secondary	EU	AA	EA
Monocytes	T-cells	0.69	0.68	0.62

Table S16: GWAS SNPs with significant *cis*-regulatory effects in monocytes. GWAS index SNP, effect size estimates (β), standard error of the effect size, meta-analysis P-value and disease/trait.

Table S17: GWAS SNPs with significant *cis*-regulatory effects in CD4+ T-cells. GWAS index SNP, effect size estimates (β), standard error of the effect size, meta-analysis P-value and disease/trait.

Table S18: Eleven Autoimmune diseases analyzed in this study. Listed here are total number of GWAS SNPS, number of SNP-gene, SNP and genes with *cis*-eQTL effects at meta-analysis FDR 0.05.

Disease	# of GWAS SNPs	# of SNP-gene(s) (eQTL)	# of GWAS SNPs (eQTL)	# of Genes (eQTL)
Ankylosing spondylitis (AS)	16	17	10	14
Crohn's disease (CD)	90	54	29	52
Ulcerative colitis (UC)	58	34	19	34
Celiac disease (CeD)	80	15	13	13
Multiple sclerosis (MS)	83	22	20	22
Type 1 diabetes (T1D)	53	30	17	29
Rheumatoid arthritis (RA)	70	22	17	22
Primary biliary cirrhosis (PBC)	19	5	5	5
Systemic lupus erythematosus (SLE)	27	12	7	12
Systemic sclerosis (SS)	18	5	4	5
Psoriasis (PS)	54	25	15	22
Total	568	241	156	230
Total (LD-pruned, top SNP per LD-block, n.r.)	425		143	164

n.r.: Non-redundant.

Table S19: Autoimmune disease-associated SNPs with significant *cis*-regulatory effects. GWAS index SNP, best proxy SNP (LD: $r^2 > 0.8$), effect size estimates (β), standard error of the effect size, meta-analysis P-value, disease, and cell type.

References and Notes

1. B. E. Stranger, S. B. Montgomery, A. S. Dimas, L. Parts, O. Stegle, C. E. Ingle, M. Sekowska, G. D. Smith, D. Evans, M. Gutierrez-Arcelus, A. Price, T. Raj, J. Nisbett, A. C. Nica, C. Beazley, R. Durbin, P. Deloukas, E. T. Dermitzakis, Patterns of *cis* regulatory variation in diverse human populations. *PLoS Genet.* **8**, e1002639 (2012). [Medline](#) [doi:10.1371/journal.pgen.1002639](https://doi.org/10.1371/journal.pgen.1002639)
2. T. Lappalainen, M. Sammeth, M. R. Friedländer, P. A. 't Hoen, J. Monlong, M. A. Rivas, M. González-Porta, N. Kurbatova, T. Griebel, P. G. Ferreira, M. Barann, T. Wieland, L. Greger, M. van Iterson, J. Almlöf, P. Ribeca, I. Pulyakhina, D. Esser, T. Giger, A. Tikhonov, M. Sultan, G. Bertier, D. G. MacArthur, M. Lek, E. Lizano, H. P. Buermans, I. Padioleau, T. Schwarzmayr, O. Karlberg, H. Ongen, H. Kilpinen, S. Beltran, M. Gut, K. Kahlem, V. Amstislavskiy, O. Stegle, M. Pirinen, S. B. Montgomery, P. Donnelly, M. I. McCarthy, P. Flicek, T. M. Strom, H. Lehrach, S. Schreiber, R. Sudbrak, A. Carracedo, S. E. Antonarakis, R. Häsler, A. C. Syvänen, G. J. van Ommen, A. Brazma, T. Meitinger, P. Rosenstiel, R. Guigó, I. G. Gut, X. Estivill, E. T. Dermitzakis; The Geuvadis Consortium, Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506–511 (2013). [Medline](#) [doi:10.1038/nature12531](https://doi.org/10.1038/nature12531)
3. A. Battle, S. Mostafavi, X. Zhu, J. B. Potash, M. M. Weissman, C. McCormick, C. D. Haudenschild, K. B. Beckman, J. Shi, R. Mei, A. E. Urban, S. B. Montgomery, D. F. Levinson, D. Koller, Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res.* **24**, 14–24 (2014). [Medline](#) [doi:10.1101/gr.155192.113](https://doi.org/10.1101/gr.155192.113)
4. B. P. Fairfax, S. Makino, J. Radhakrishnan, K. Plant, S. Leslie, A. Dilthey, P. Ellis, C. Langford, F. O. Vannberg, J. C. Knight, Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* **44**, 502–510 (2012). [Medline](#) [doi:10.1038/ng.2205](https://doi.org/10.1038/ng.2205)
5. H. J. Westra, M. J. Peters, T. Esko, H. Yaghoobkar, C. Schurmann, J. Kettunen, M. W. Christiansen, B. P. Fairfax, K. Schramm, J. E. Powell, A. Zernakova, D. V. Zernakova, J. H. Veldink, L. H. Van den Berg, J. Karjalainen, S. Withoff, A. G. Uitterlinden, A. Hofman, F. Rivadeneira, P. A. 't Hoen, E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A. B. Singleton, D. G. Hernandez, M. A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Völker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dacey, S. A. Gharib, D. A. Enquobahrie, T. Lumley, G. W. Montgomery, S. Makino, H. Prokisch, C. Herder, M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R. C. Jansen, P. M. Visscher, J. C. Knight, B. M. Psaty, S. Ripatti, A. Teumer, T. M. Frayling, A. Metspalu, J. B. van Meurs, L. Franke, Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013). [Medline](#) [doi:10.1038/ng.2756](https://doi.org/10.1038/ng.2756)
6. D. L. Nicolae, E. Gamazon, W. Zhang, S. Duan, M. E. Dolan, N. J. Cox, Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. *PLoS Genet.* **6**, e1000888 (2010). [Medline](#) [doi:10.1371/journal.pgen.1000888](https://doi.org/10.1371/journal.pgen.1000888)

7. G. Trynka, C. Sandor, B. Han, H. Xu, B. E. Stranger, X. S. Liu, S. Raychaudhuri, Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat. Genet.* **45**, 124–130 (2013). [Medline doi:10.1038/ng.2504](#)
8. M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P. Reynolds, R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kuttyavin, S. Stehling-Sun, A. K. Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen, S. Neph, P. J. Sabo, S. Heimfeld, A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, J. A. Stamatoyannopoulos, Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012). [doi:10.1126/science.1222794](#)
9. G. R. Abecasis, A. Auton, L. D. Brooks, M. A. DePristo, R. M. Durbin, R. E. Handsaker, H. M. Kang, G. T. Marth, G. A. McVean; The 1000 Genomes Project Consortium, An integrated map of genetic variation from 1092 human genomes. *Nature* **491**, 56–65 (2012). [Medline doi:10.1038/nature11632](#)
10. T. Flutre, X. Wen, J. Pritchard, M. Stephens, A statistical framework for joint eQTL analysis in multiple tissues. *PLoS Genet.* **9**, e1003486 (2013). [Medline doi:10.1371/journal.pgen.1003486](#)
11. J. D. Storey, R. Tibshirani, Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9440–9445 (2003). [Medline doi:10.1073/pnas.1530509100](#)
12. B. Han, E. Eskin, Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am. J. Hum. Genet.* **88**, 586–598 (2011). [Medline doi:10.1016/j.ajhg.2011.04.014](#)
13. B. E. Bernstein, E. Birney, I. Dunham, E. D. Green, C. Gunter, M. Snyder; The ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012). [Medline doi:10.1038/nature11247](#)
14. L. A. Hindorff, P. Sethupathy, H. A. Junkins, E. M. Ramos, J. P. Mehta, F. S. Collins, T. A. Manolio, Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9362–9367 (2009). [Medline doi:10.1073/pnas.0903103106](#)
15. A. C. Nica, S. B. Montgomery, A. S. Dimas, B. E. Stranger, C. Beazley, I. Barroso, E. T. Dermitzakis, Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genet.* **6**, e1000895 (2010). [Medline doi:10.1371/journal.pgen.1000895](#)
16. E. M. Bradshaw, L. B. Chibnik, B. T. Keenan, L. Ottoboni, T. Raj, A. Tang, L. L. Rosenkrantz, S. Imboywa, M. Lee, A. Von Korff, M. C. Morris, D. A. Evans, K. Johnson, R. A. Sperling, J. A. Schneider, D. A. Bennett, P. L. De Jager; Alzheimer Disease Neuroimaging Initiative, CD33 Alzheimer's disease locus: Altered monocyte function and amyloid biology. *Nat. Neurosci.* **16**, 848–850 (2013). [Medline doi:10.1038/nn.3435](#)
17. C. K. Glass, K. Saijo, B. Winner, M. C. Marchetto, F. H. Gage, Mechanisms underlying inflammation in neurodegeneration. *Cell* **140**, 918–934 (2010). [Medline doi:10.1016/j.cell.2010.02.016](#)

18. J. Harrow, A. Frankish, J. M. Gonzalez, E. Tapanari, M. Diekhans, F. Kokocinski, B. L. Aken, D. Barrell, A. Zadissa, S. Searle, I. Barnes, A. Bignell, V. Boychenko, T. Hunt, M. Kay, G. Mukherjee, J. Rajan, G. Despacio-Reyes, G. Saunders, C. Steward, R. Harte, M. Lin, C. Howald, A. Tanzer, T. Derrien, J. Chrast, N. Walters, S. Balasubramanian, B. Pei, M. Tress, J. M. Rodriguez, I. Ezkurdia, J. van Baren, M. Brent, D. Haussler, M. Kellis, A. Valencia, A. Reymond, M. Gerstein, R. Guigó, T. J. Hubbard, GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012). [Medline doi:10.1101/gr.135350.111](#)
19. J. T. Leek, J. D. Storey, Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLOS Genet.* **3**, e161 (2007). [Medline doi:10.1371/journal.pgen.0030161](#)
20. O. Stegle, L. Parts, R. Durbin, J. Winn, A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLOS Comput. Biol.* **6**, e1000770 (2010). [Medline doi:10.1371/journal.pcbi.1000770](#)
21. A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006). [Medline doi:10.1038/ng1847](#)
22. B. E. Stranger, M. S. Forrest, M. Dunning, C. E. Ingle, C. Beazley, N. Thorne, R. Redon, C. P. Bird, A. de Grassi, C. Lee, C. Tyler-Smith, N. Carter, S. W. Scherer, S. Tavaré, P. Deloukas, M. E. Hurles, E. T. Dermitzakis, Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007). [doi:10.1126/science.1136678](#)
23. B. E. Stranger, M. S. Forrest, A. G. Clark, M. J. Minichiello, S. Deutsch, R. Lyle, S. Hunt, B. Kahl, S. E. Antonarakis, S. Tavaré, P. Deloukas, E. T. Dermitzakis, Genome-wide associations of gene expression variation in humans. *PLOS Genet.* **1**, e78 (2005). [Medline doi:10.1371/journal.pgen.0010078](#)
24. R. W. Doerge, G. A. Churchill, Permutation tests for multiple loci affecting a quantitative character. *Genetics* **142**, 285–294 (1996). [Medline](#)
25. S. I. Lee, D. Pe'er, A. M. Dudley, G. M. Church, D. Koller, Identifying regulatory mechanisms using individual variation reveals key role for chromatin modification. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14062–14067 (2006). [Medline doi:10.1073/pnas.0601852103](#)
26. J. C. Almlöf, P. Lundmark, A. Lundmark, B. Ge, S. Maouche, H. H. Göring, U. Liljedahl, C. Enström, J. Brocheton, C. Proust, T. Godefroy, J. G. Sambrook, J. Jolley, A. Crisp-Hihn, N. Foad, H. Lloyd-Jones, J. Stephens, R. Gwilliam, C. M. Rice, C. Hengstenberg, N. J. Samani, J. Erdmann, H. Schunkert, T. Pastinen, P. Deloukas, A. H. Goodall, W. H. Ouwehand, F. Cambien, A. C. Syvänen, Powerful identification of *cis*-regulatory SNPs in human primary monocytes using allele-specific gene expression. *PLOS ONE* **7**, e52260 (2012). [Medline doi:10.1371/journal.pone.0052260](#)
27. R. J. Pruim, R. P. Welch, S. Sanna, T. M. Teslovich, P. S. Chines, T. P. Gliedt, M. Boehnke, G. R. Abecasis, C. J. Willer, LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010). [Medline doi:10.1093/bioinformatics/btq419](#)

28. T. Zeller, P. Wild, S. Szymczak, M. Rotival, A. Schillert, R. Castagne, S. Maouche, M. Germain, K. Lackner, H. Rossmann, M. Eleftheriadis, C. R. Sinning, R. B. Schnabel, E. Lubos, D. Mennerich, W. Rust, C. Perret, C. Proust, V. Nicaud, J. Loscalzo, N. Hübner, D. Tregouet, T. Münzel, A. Ziegler, L. Tiret, S. Blankenberg, F. Cambien, Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLOS ONE* **5**, e10693 (2010). [Medline doi:10.1371/journal.pone.0010693](https://doi.org/10.1371/journal.pone.0010693)
29. R. S. Fehrmann, R. C. Jansen, J. H. Veldink, H. J. Westra, D. Arends, M. J. Bonder, J. Fu, P. Deelen, H. J. Groen, A. Smolonska, R. K. Weersma, R. M. Hofstra, W. A. Buurman, S. Rensen, M. G. Wolfs, M. Platteel, A. Zernakova, C. C. Elbers, E. M. Festen, G. Trynka, M. H. Hofker, C. G. Saris, R. A. Ophoff, L. H. van den Berg, D. A. van Heel, C. Wijmenga, G. J. Te Meerman, L. Franke, Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLOS Genet.* **7**, e1002197 (2011). [Medline doi:10.1371/journal.pgen.1002197](https://doi.org/10.1371/journal.pgen.1002197)
30. A. L. Dixon, L. Liang, M. F. Moffatt, W. Chen, S. Heath, K. C. Wong, J. Taylor, E. Burnett, I. Gut, M. Farrall, G. M. Lathrop, G. R. Abecasis, W. O. Cookson, A genome-wide association study of global gene expression. *Nat. Genet.* **39**, 1202–1207 (2007). [Medline doi:10.1038/ng2109](https://doi.org/10.1038/ng2109)
31. T. Raj, M. Kuchroo, J. M. Replogle, S. Raychaudhuri, B. E. Stranger, P. L. De Jager, Common risk alleles for inflammatory diseases are targets of recent positive selection. *Am. J. Hum. Genet.* **92**, 517–529 (2013). [Medline doi:10.1016/j.ajhg.2013.03.001](https://doi.org/10.1016/j.ajhg.2013.03.001)