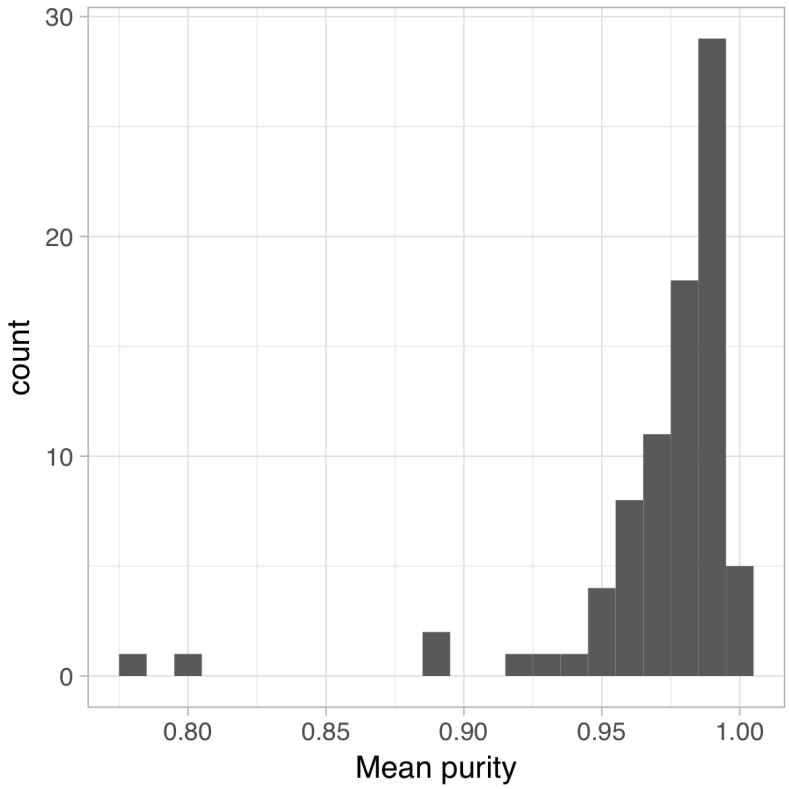
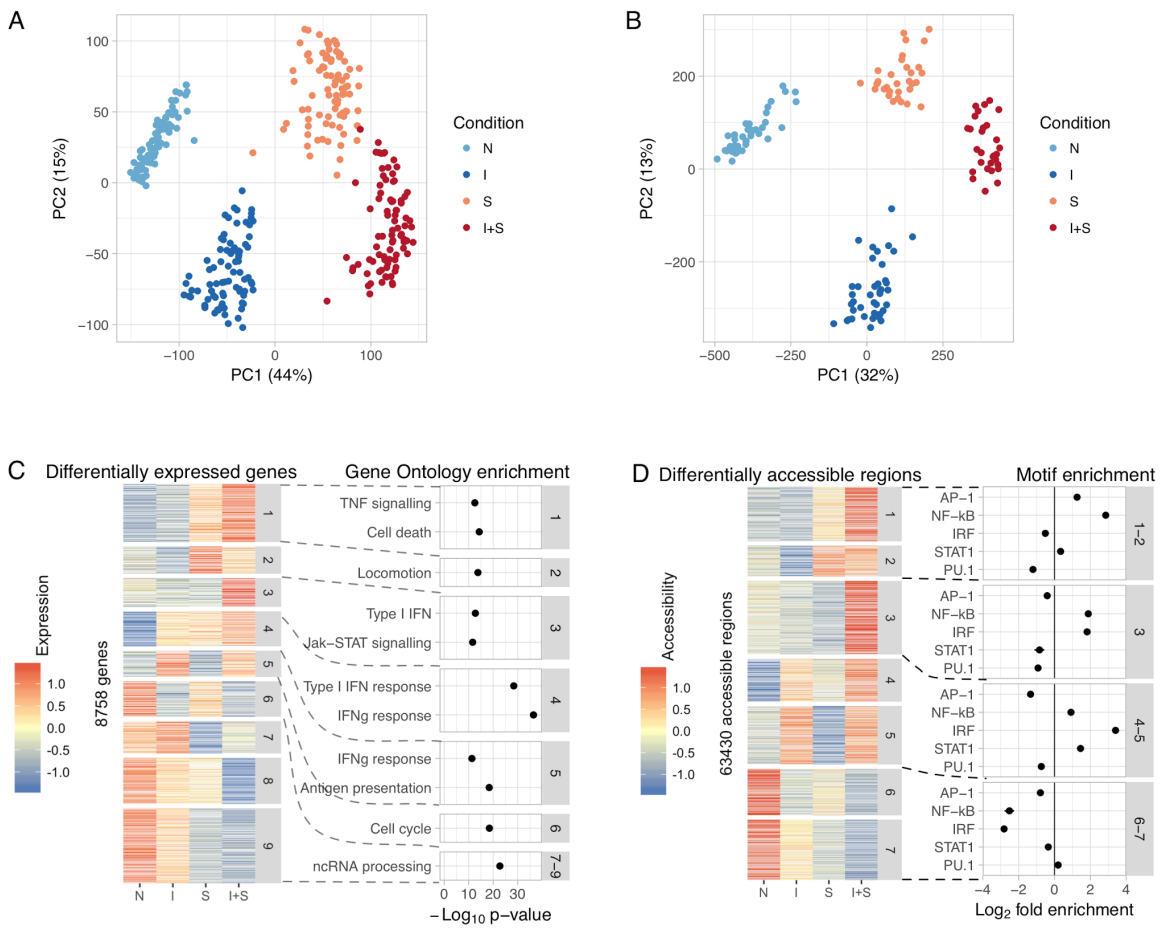
Supplementary Figures

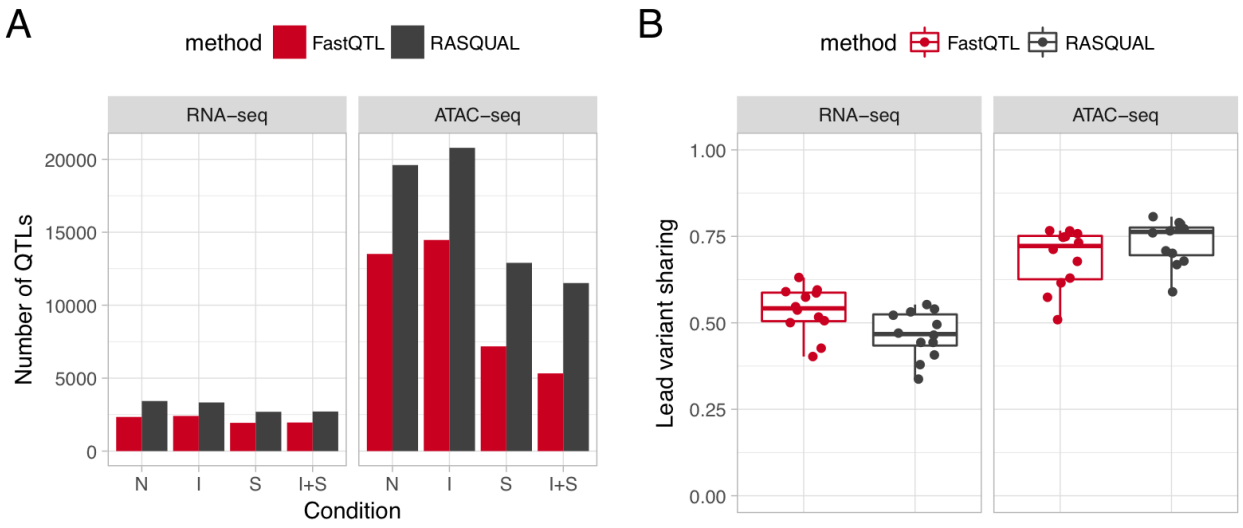


Supplementary Figure 1. Histogram of the purity estimates of iPSC-derived macrophages. Macrophages were stained with antibodies for CD14, CD16 and CD206 and flow cytometry was used to estimate the proportion of cells stained positive for each of the three markers (see Supplementary Note). Since the purity estimates from all three markers were highly correlated, we decided to calculate their mean as an estimate of sample purity. The values for each sample are presented in Supplementary Table 1.

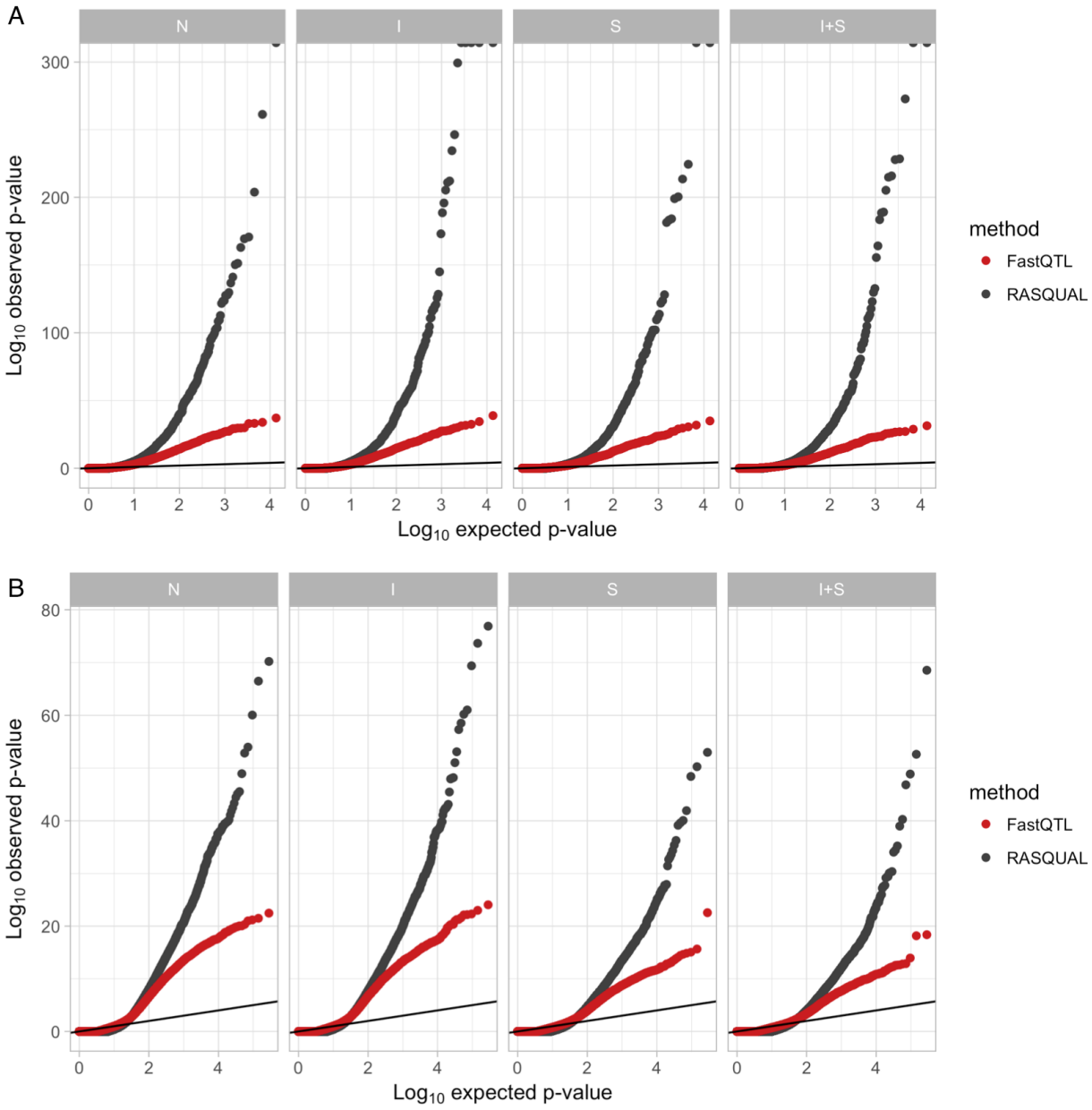


Supplementary Figure 2. Differential gene expression and chromatin accessibility in macrophage immune response. **(A)** Principal component analysis of the gene expression data, n = 84 independent donors in each condition. **(B)** Principal component analysis of the chromatin accessibility data. The number of independent donors in each condition was n = 42 (N), n = 41 (I) and n = 31 (S and I+S). **(C)** Left panel: 8,758 differentially expressed genes clustered into nine distinct expression patterns (n = 84 unique donors across four conditions). Right panel: Selection of Gene Ontology terms enriched in each cluster. Only enrichments with $p < 1\times10^{-8}$ are shown in the figure. Enrichment p-values were calculated using g:Profiler[1]. Differential gene expression patterns closely recapitulated known aspects of macrophage immune response. For example, genes upregulated by *Salmonella* (cluster 1) were enriched for tumor necrosis factor (TNF) signalling and cell death pathways whereas genes upregulated by IFNɣ (cluster 5) were enriched for IFNɣ response and antigen presentation pathways. **(D)** Left panel: heatmap of 63,350 differentially accessible regions clustered into seven distinct patterns (n = 16 high quality donors across four conditions (see Supplementary Note)). Right panel: enrichment of TF motifs in four groups of differentially accessible clusters relative to all open chromatin regions. The points represent fold enrichment calculated using two-sided Fisher's exact test. Due to the large number of differentially accessible regions, the 95% confidence intervals from Fisher's exact test are too narrow to be visible on the plot. Similarly to the gene expression data (panel C), open chromatin regions in clusters 1 and 2 that became accessible after *Salmonella* infection were specifically enriched for NF-κB and AP-1 motifs, two main TFs activated downstream of

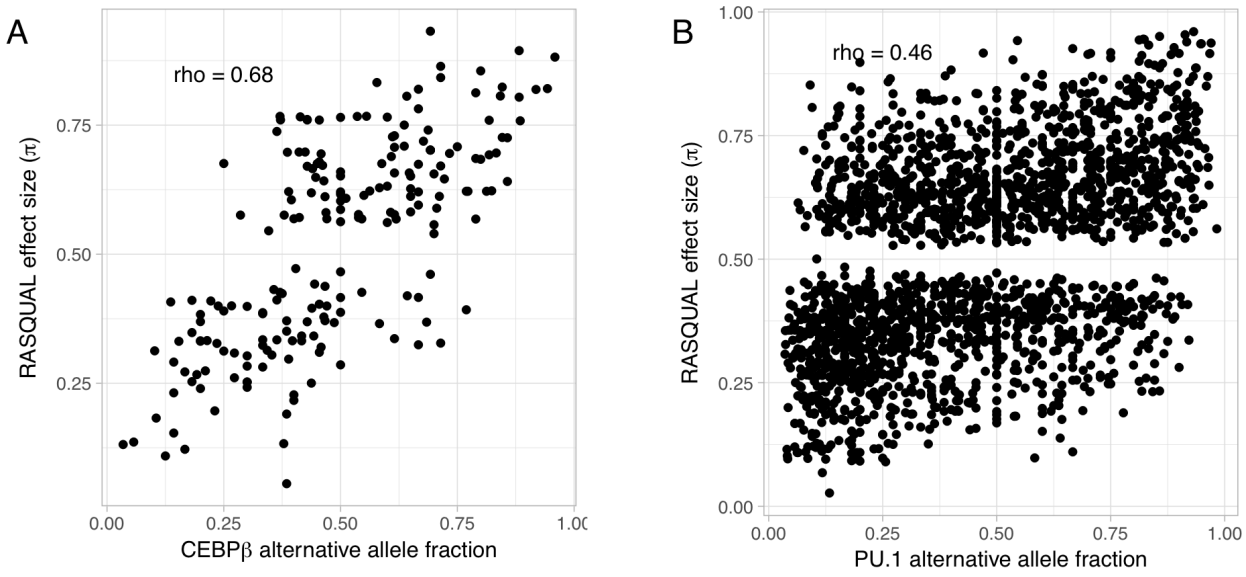
toll-like receptor 4 (TLR4) signalling[2]. In contrast, clusters 4 and 5 showed increased accessibility after IFNɣ stimulation and were enriched for IRF and STAT1 motifs, consistent with the activation of STAT1 and IRF1 downstream of IFNɣ signalling[3].



Supplementary Figure 3. Genetic effects on gene expression and chromatin accessibility across conditions. **(A)** Number of genes and open chromatin regions for which we detected at least one significant QTL (FDR < 10%) using either allele-specific model (RASQUAL) or standard linear regression (FastQTL). The number of significant QTLs was counted in each condition separately. The RNA-seq data consisted of n = 84 independent donors in all four conditions. The number of independent donors in the ATAC-seq data was n = 42 (N), n = 41 (I) and n = 31 (S and I+S), depending on the condition. (**B**) Proportion of lead QTL variants shared (R^2 > 0.8, Pearson correlation) between all pairs of conditions (n = 12) for both eQTLs and caQTLs. Box plots show the median (center line) and the 25th and 75th percentiles (box edges), with whiskers extending to 1.5 times the interquartile range.

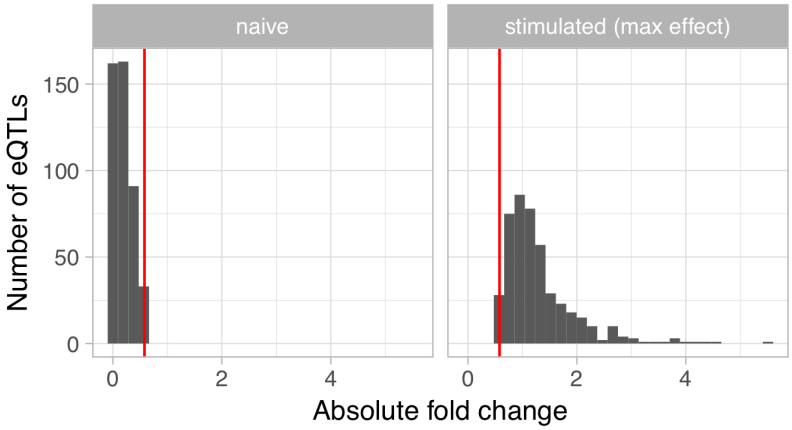


Supplementary Figure 4. Quantile-quantile plots for eQTLs (**A**) and caQTLs (**B**) detected using either linear model (FastQTL) or RASQUAL. Each point represents the p-value of the lead eQTL (caQTL) variant for each gene (region). The -$\log_{10}$ p-values were calculated using either FastQTL or RASQUAL as indicated. The p-values have been adjusted with eigenMT[4] to account for multiple variants tested per gene or accessible chromatin region. The RNA-seq data consisted of n = 84 independent donors in all four conditions. The number of independent donors in ATAC-seq data was n = 42 (N), n = 41 (I) and n = 31 (S and I+S), depending on the condition.

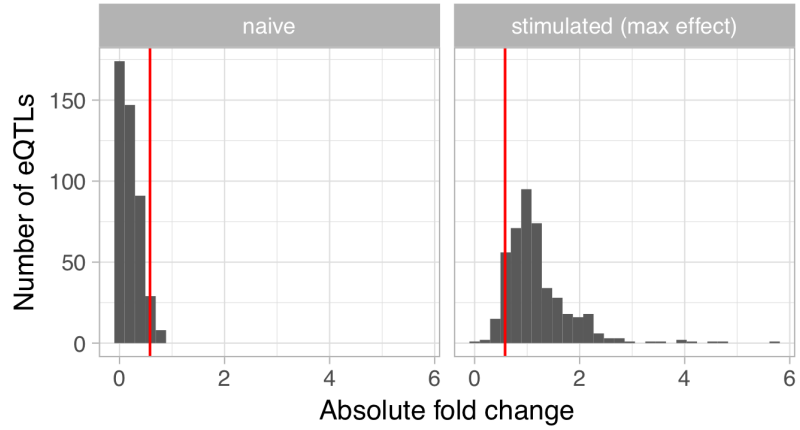


Supplementary Figure 5. Correlation between caQTL effect size and TF binding**.** Only credible set variants overlapping the regulated caQTL region were included in the analysis. Furthermore, each variant was required to have at least ten overlapping reads in the ChIP-seq data with at least two reads supporting the lower binding allele to exclude potential homozygotes. See Supplementary Note for more details. (**A**) Spearman's correlation (rho) between allele-specific binding of CEBPβ and RASQUAL caQTL effect size (n = 42 independent donors) at 202 heterozygous loci. (**B**) Spearman's correlation between allele-specific binding of PU.1 and RASQUAL caQTL effect size (n = 42 independent donors) at 2159 heterozygous loci.

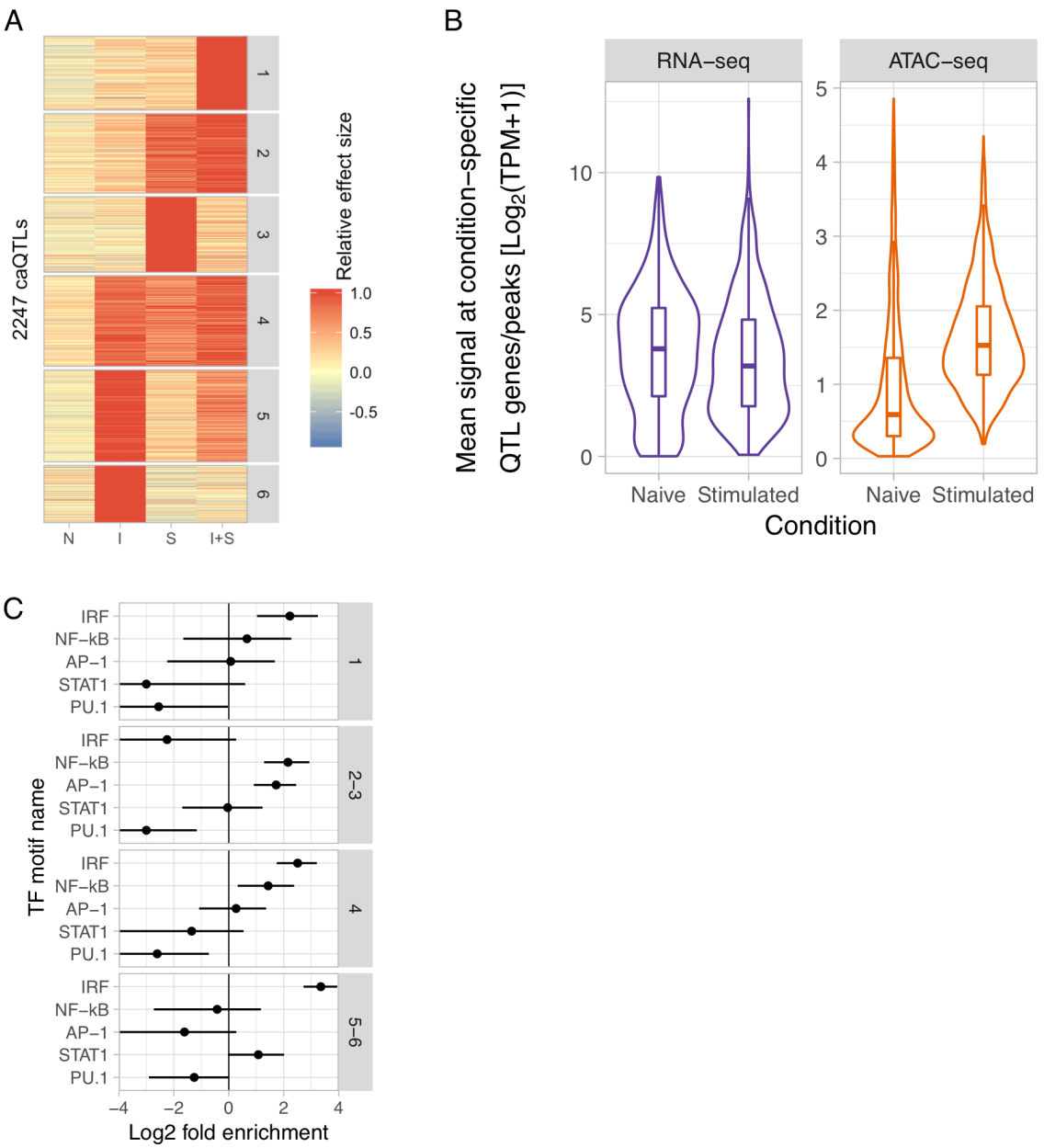
A Fold change distributions in 84 samples



B Fold change distributions for the same eQTLs in 228 samples

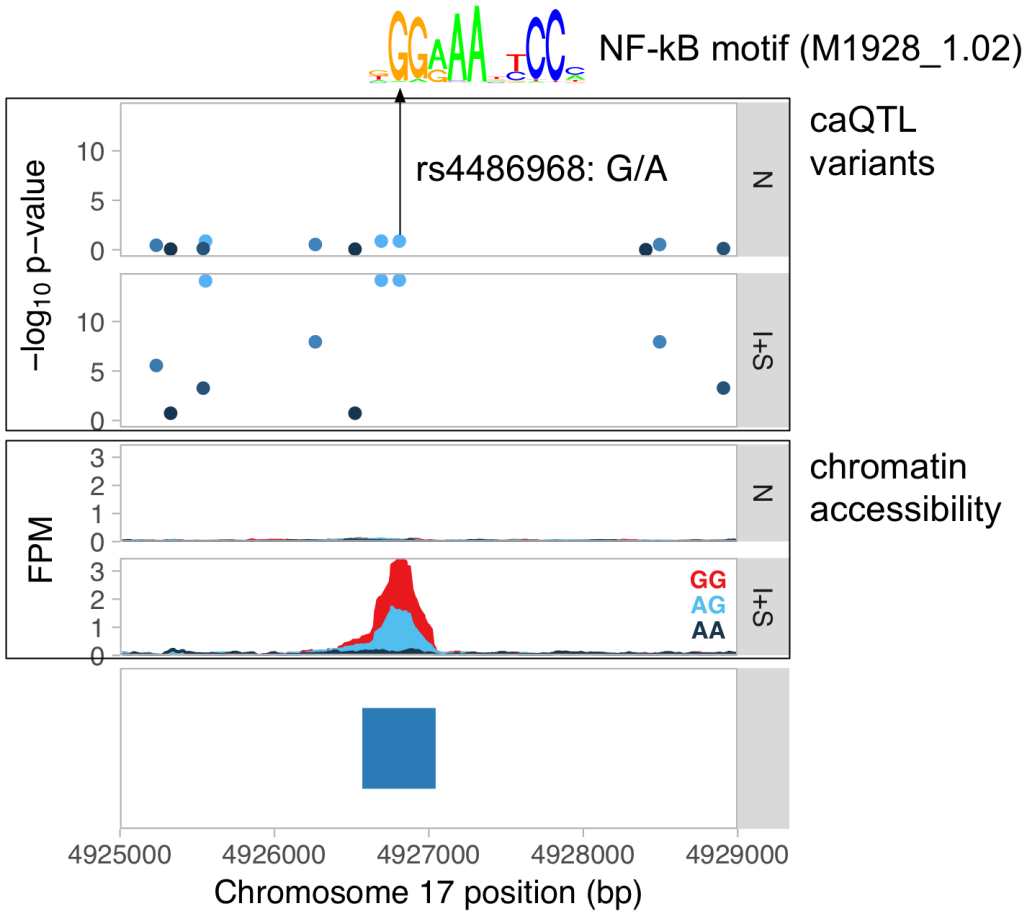


Supplementary Figure 6. Quantifying the effect of small sample size on the accuracy of response eQTL effect size estimates. (**A**) Response eQTL effect size distribution in the discovery sample of 84 individuals in the naive condition as well as in the stimulated condition with the largest effect size. (**B**) Effect size of the same response eQTLs in a much larger replication sample of 228 individuals.

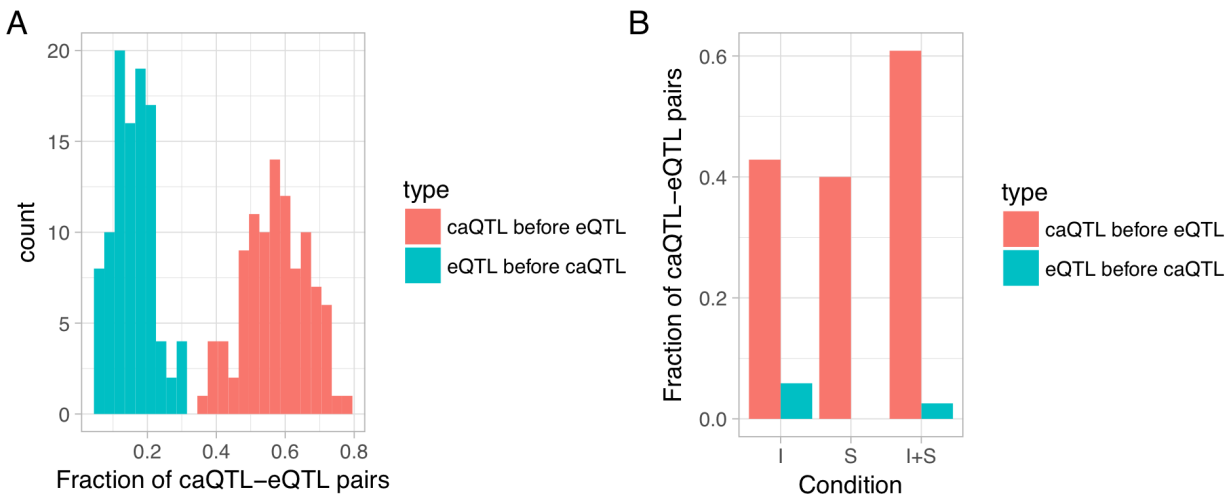


Supplementary Figure 7. Characterisation of response caQTLs. **(A)** Response caQTLs clustered into six groups using k-means clustering. The number of unique donors in the ATAC-seq data was n = 42 (N), n = 41 (I) and n = 31 (S and I+S), depending on the condition. **(B)** Distribution of mean gene expression (n = 387 genes with response eQTLs) (or chromatin accessibility (n = 2247 regions with a response caQTL)) in the naive condition and in the condition where the response QTL had the largest effect size. Condition-specific caQTLs were inaccessible in the naive cells and became accessible only in the condition in which the genetic effect appeared whereas mean expression of genes with response eQTLs did not differ between conditions. TPM, transcripts per million. Box plots show the median (center line) and the 25th and 75th percentiles (box edges), with whiskers extending to 1.5 times the interquartile range. Mean gene expression was calculated from n = 84 independent donors in each condition and mean chromatin accessibility was calculated from n = 16 high quality donors

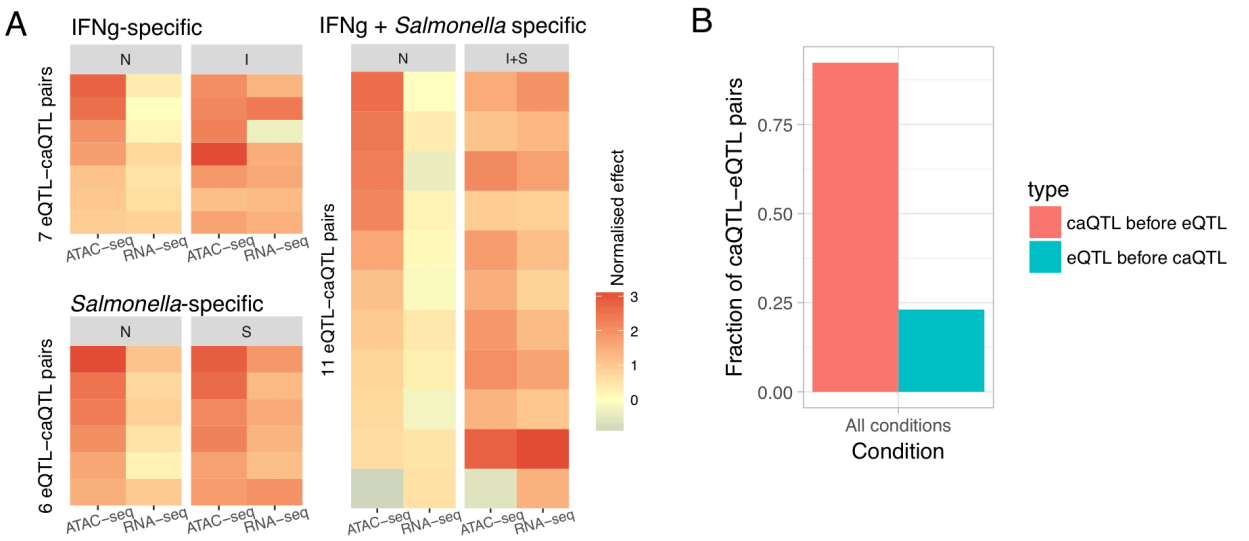
(Supplementary Note) in each condition. **(C)** Enrichment of TF motif disruption by response caQTLs relative to all other caQTLs. The numbers on the left correspond to clusters from panel (A). The points represent $\log_2$ fold enrichment calculated using two-sided Fisher's exact test and the solid lines show 95% confidence intervals. *Salmonella*-specific caQTLs (clusters 2 and 3 from panel A) are enriched for disrupting NF-kB and AP-1 TF motifs whereas IFNɣ-specific caQTLs (clusters 5 and 6 from panel A) are enriched for disrupting the IRF motif.



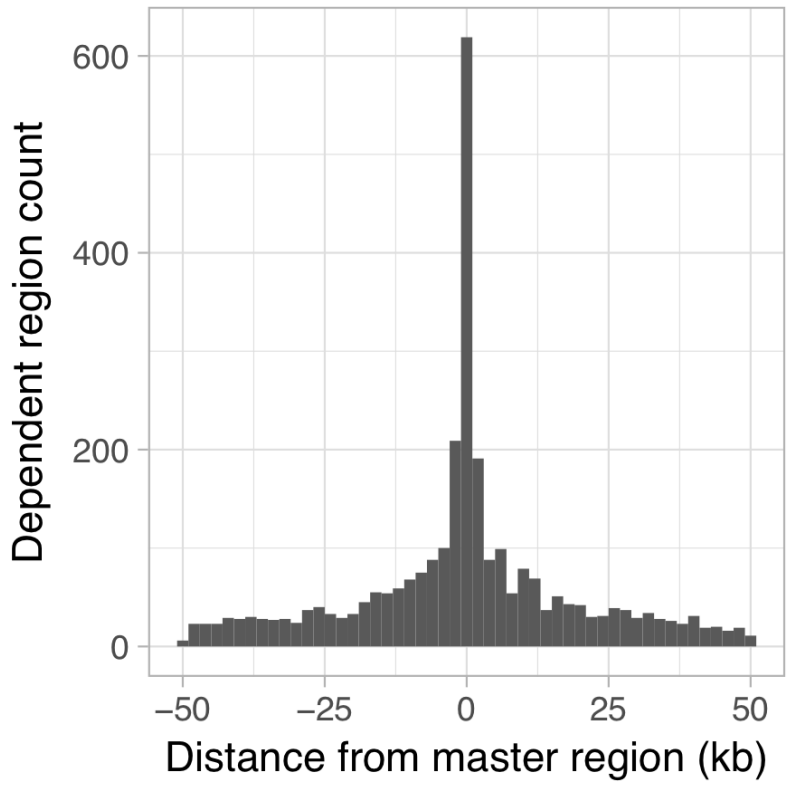
Supplementary Figure 8. Fine mapping the causal variant for the caQTL upstream of GP1BA. The -$\log_{10}$ p-values were calculated with RASQUAL using either n = 42 (N condition) or n = 31 (I+S condition) independent donors. The rs4486968 variant located within the accessible region is predicted to disrupt NF-κB binding motif (M1928_1.02 from CIS-BP) by changing high affinity G (92.6% relative binding score) at position 2 to low affinity A (88.3% relative binding score).



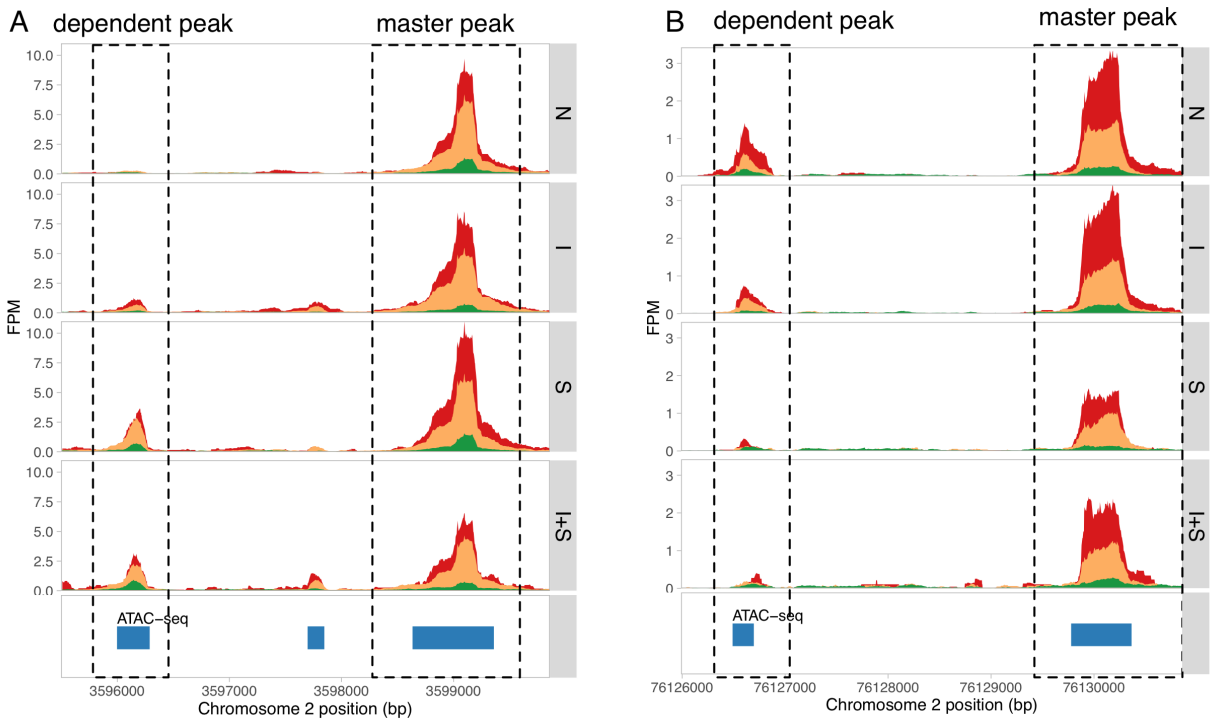
Supplementary Figure 9. Quantifying the robustness of the enhancer priming analysis. **(A)** Estimating the effect of the imbalance between the number of eQTLs and caQTLs. We randomly sampled 5383 caQTL regions (equal to the number of unique eQTL genes) 100 times and redid the enhancer priming analysis using the original fold change > 1.5 threshold. **(B)** Enhancer priming analysis using a more stringent (fold change > 2) threshold. Comparison of our estimated rate of enhancer priming (caQTL appears before response eQTL) to a negative control (eQTL appears before a response caQTL). This is the same analysis as presented on Fig. 2c except for a more stringent fold change threshold. Across conditions, 26/52 caQTLs linked to response eQTLs were present in the naive condition whereas only 4/100 eQTLs appeared before a response caQTLs.



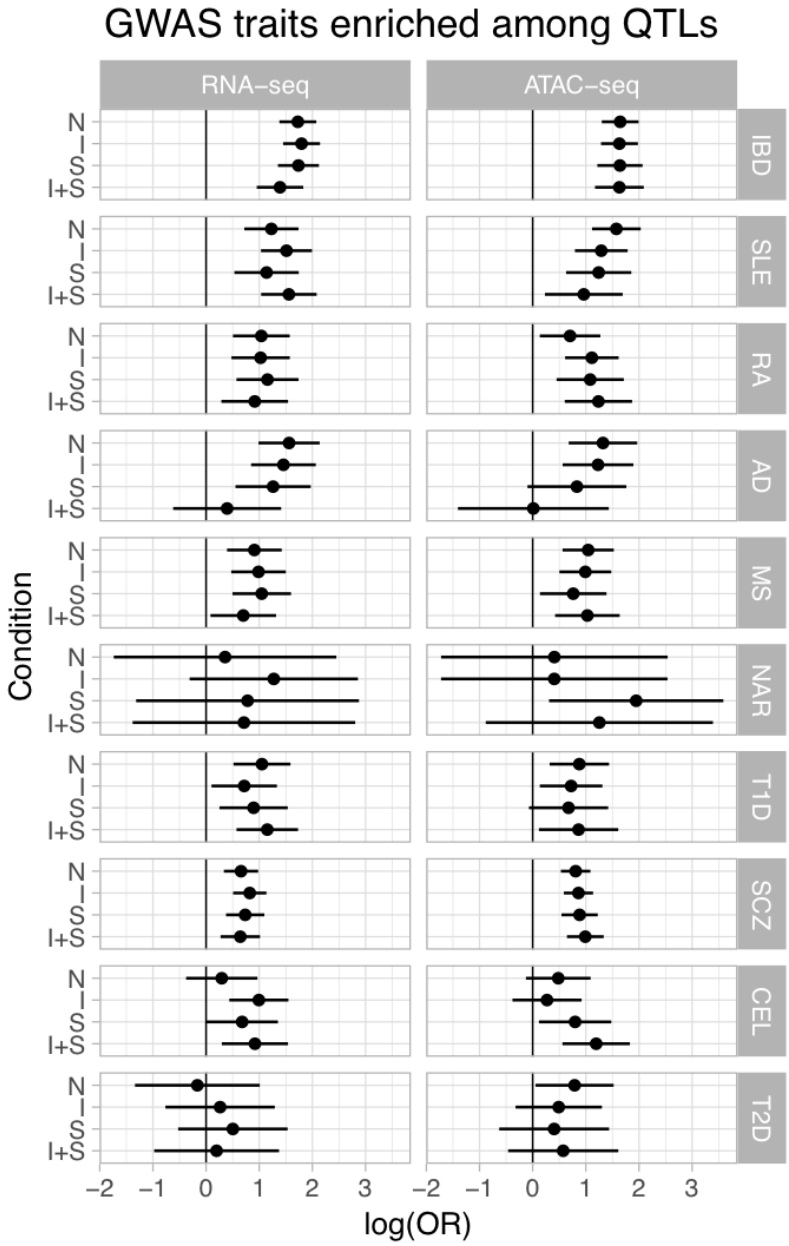
Supplementary Figure 10. We performed colocalisation between eQTLs and caQTLs using the same cis-window size (200 kb) and posterior probability cutoffs that were used for the GWAS colocalisation analysis (see Online Methods). **(A)** Effect sizes of eQTL-caQTL pairs in naive and stimulated conditions. The pairs are grouped by the condition in which the eQTL had the largest effect size (I, S or I+S). The heat maps are sorted by caQTL effect size in the naive condition (first column). Approximately 90% of the caQTLs linked to response eQTLs using colocalisation are present in the naive condition (fold change > 1.5). Only a single eQTL-caQTL pair in the IFNɣ condition shows discordant effect size direction. **(B)** Comparison of our estimated rate of enhancer priming (caQTL precedes response eQTL) to a negative control (eQTL precedes a response caQTL) on colocalised eQTL-caQTL pairs.



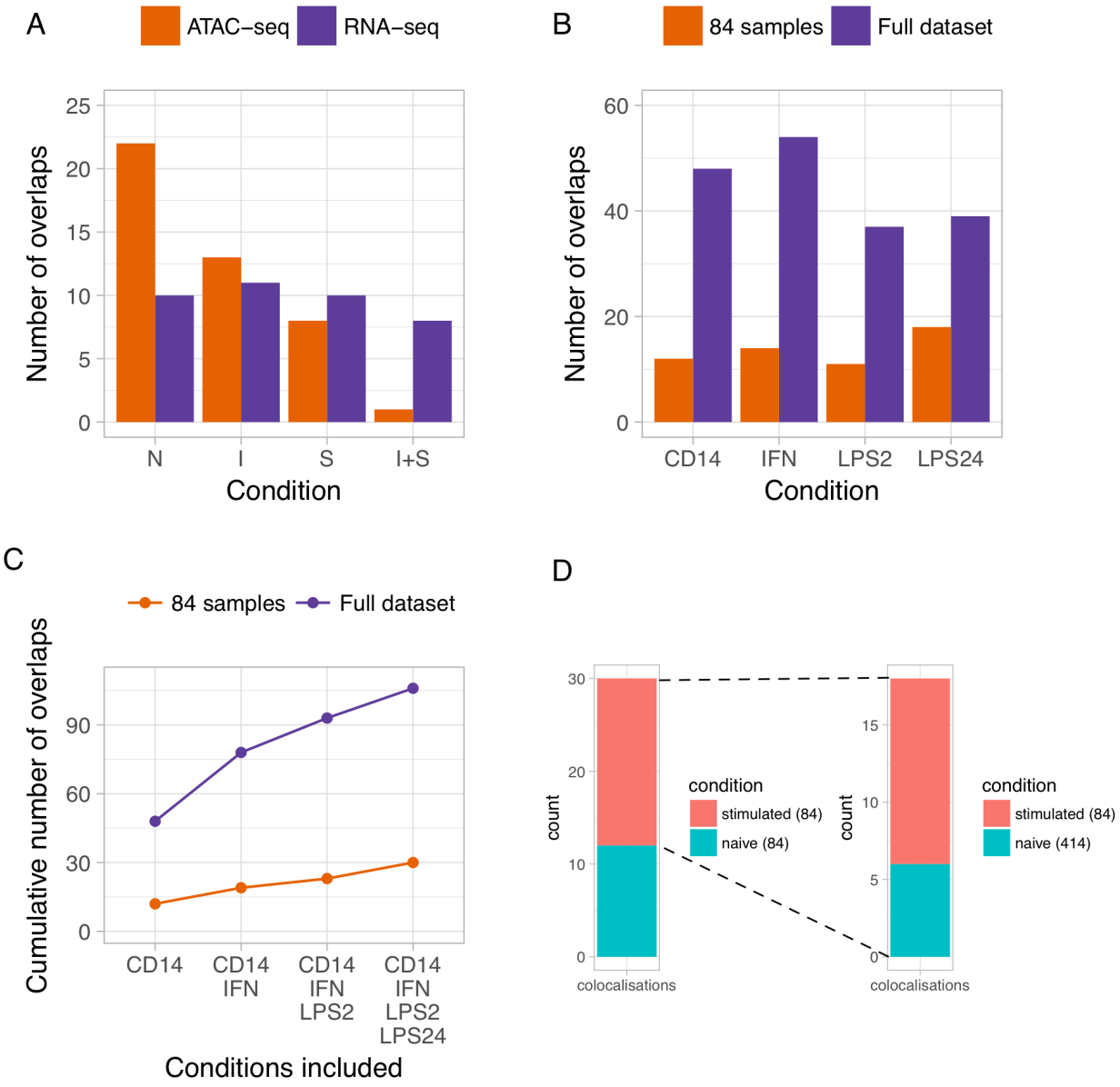
Supplementary Figure 11. Histogram of the distances between master and dependent caQTL regions.



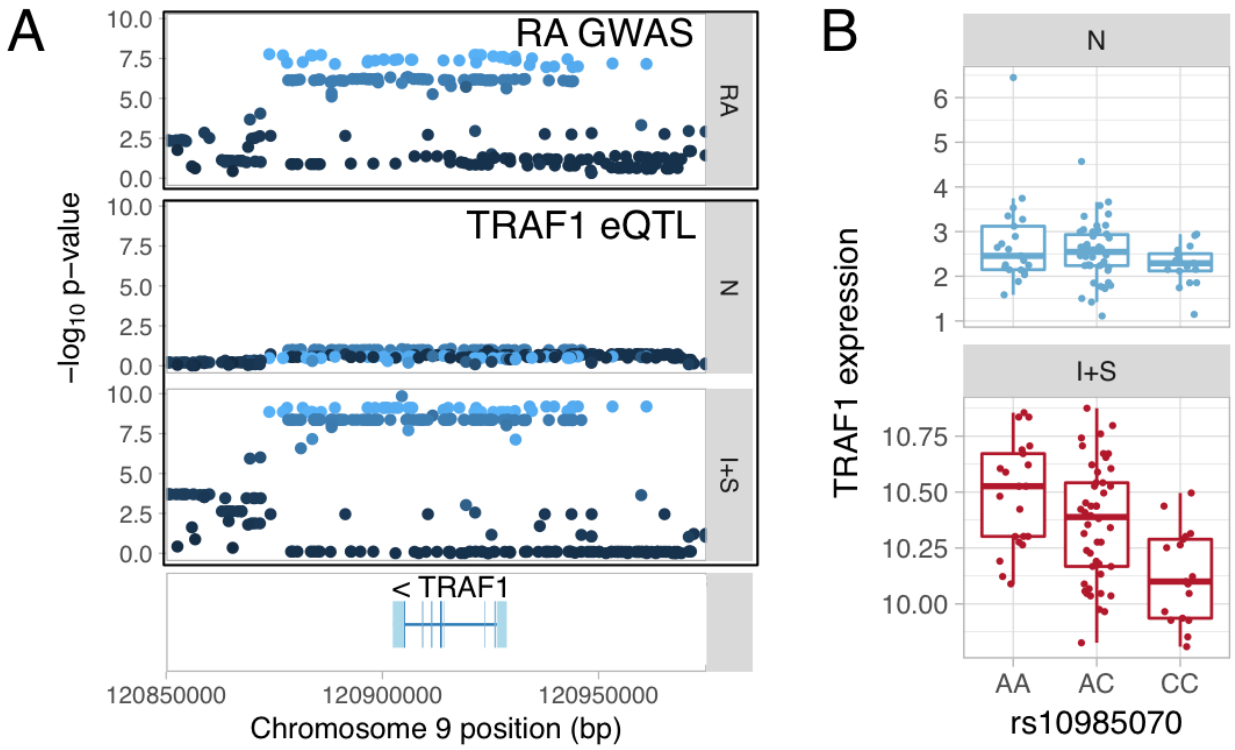
Supplementary Figure 12. Two examples of condition-specific dependent peaks. (**A**) Dependent peak appears after *Salmonella* infection. (**B**) Dependent peak disappears after *Salmonella* infection.



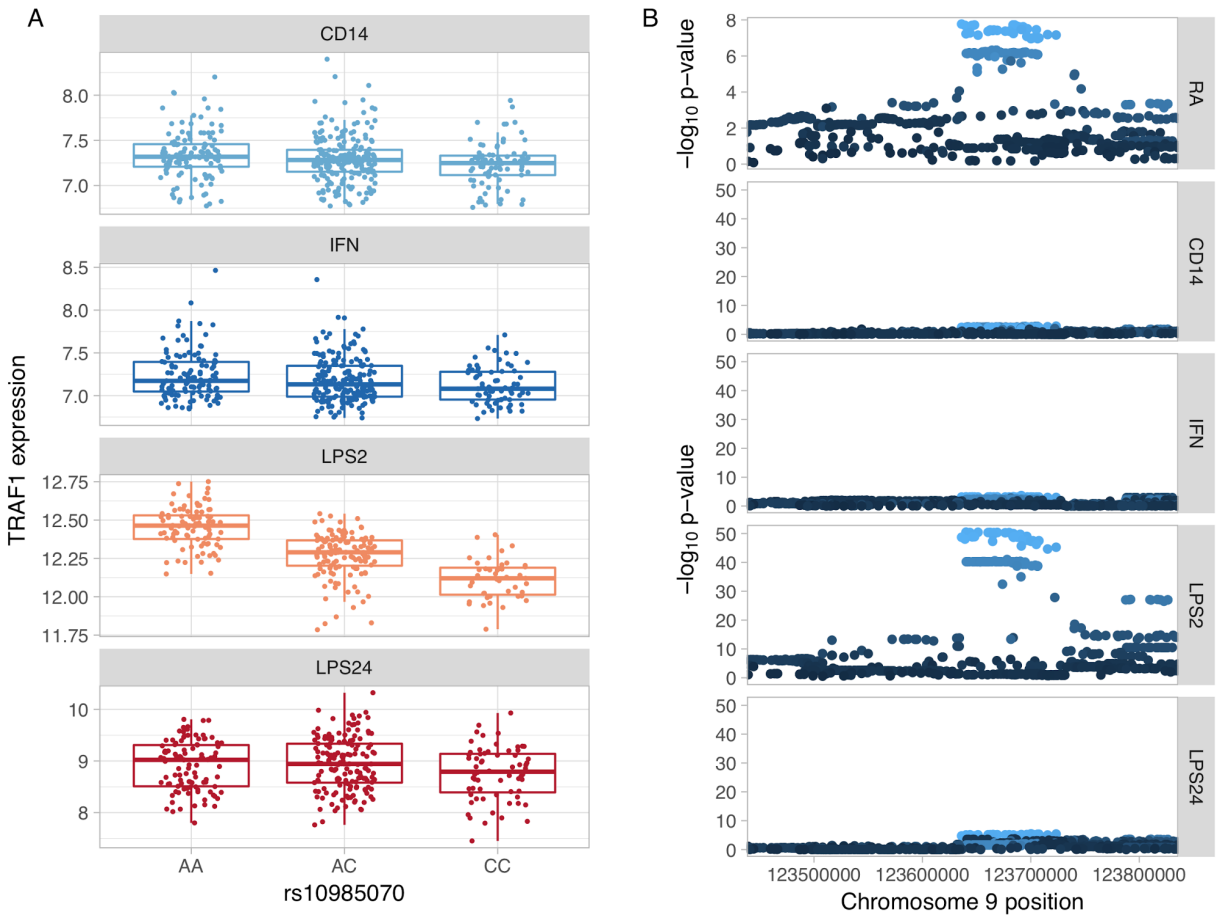
Supplementary Figure 13. GARFIELD[5] enrichment of eQTLs and caQTLs among GWAS summary statistics from 10 complex diseases. The points represent the logarithm of enrichment odds ratios (OR) and error bars show the 95% confidence intervals (see Supplementary Note for more details). The raw data underlying this figure, including the number of loci considered for each trait and condition pair, is presented in Supplementary Table 4. Disease acronyms: IBD, inflammatory bowel disease; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; AD, Alzheimer's disease; SCZ, schizophrenia; T2D, type 2 diabetes; MS, multiple sclerosis; NAR, narcolepsy; CEL, celiac disease; T1D, type 1 diabetes.



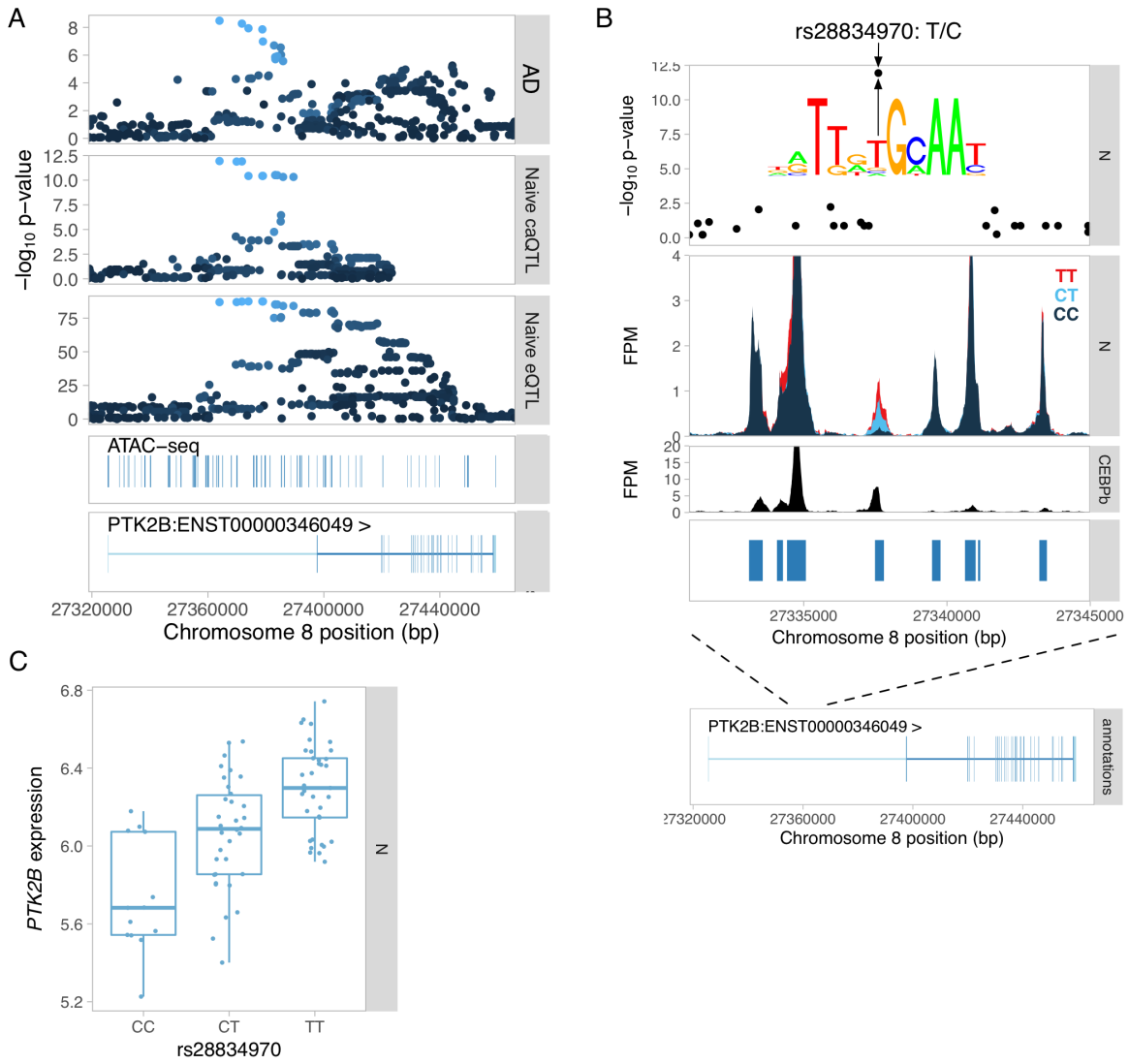
Supplementary Figure 14. Robustness of the colocalisation analysis. (**A)** Number of GWAS hits colocalised with caQTLs and eQTLs in each experimental condition. (**B**) Number of GWAS hits colocalised with eQTLs in each condition in either the full Fairfax dataset (up to 414 donors) or a subsample of 84 individuals. (**C**) Cumulative number of colocalisation detected in the Fairfax dataset across the four conditions. The analysis was performed either on the Full dataset (up to 414 donors per condition) or a random subsample of 84 donors. (**D**) In the subsample analysis, 18/30 colocalisations were detected only in the stimulated conditions and 6/18 of those 'condition-specific' colocalisations were detected in the unstimulated condition (CD14, 414 donors) in the full dataset.



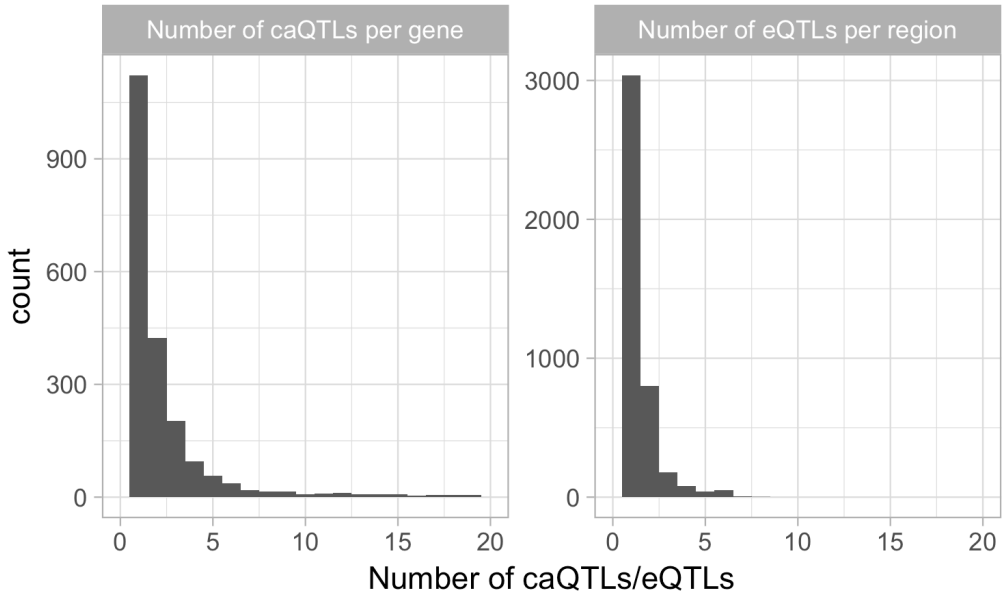
Supplementary Figure 15. Colocalisation between a response eQTL and a GWAS hit. **(A)** IFNɣ + *Salmonella* specific eQTL for TRAF1 colocalised with a GWAS hit for rheumatoid arthritis (RA). The eQTL -$\log_{10}$ p-values were calculated using FastQTL (n = 84 independent donors in both conditions). **(B)** TRAF1 expression in naive and IFNɣ + *Salmonella* conditions stratified by the lead GWAS variant (n = 84 unique donors in both conditions). Box plots show the median (center line) and the 25th and 75th percentiles (box edges), with whiskers extending to 1.5 times the interquartile range.



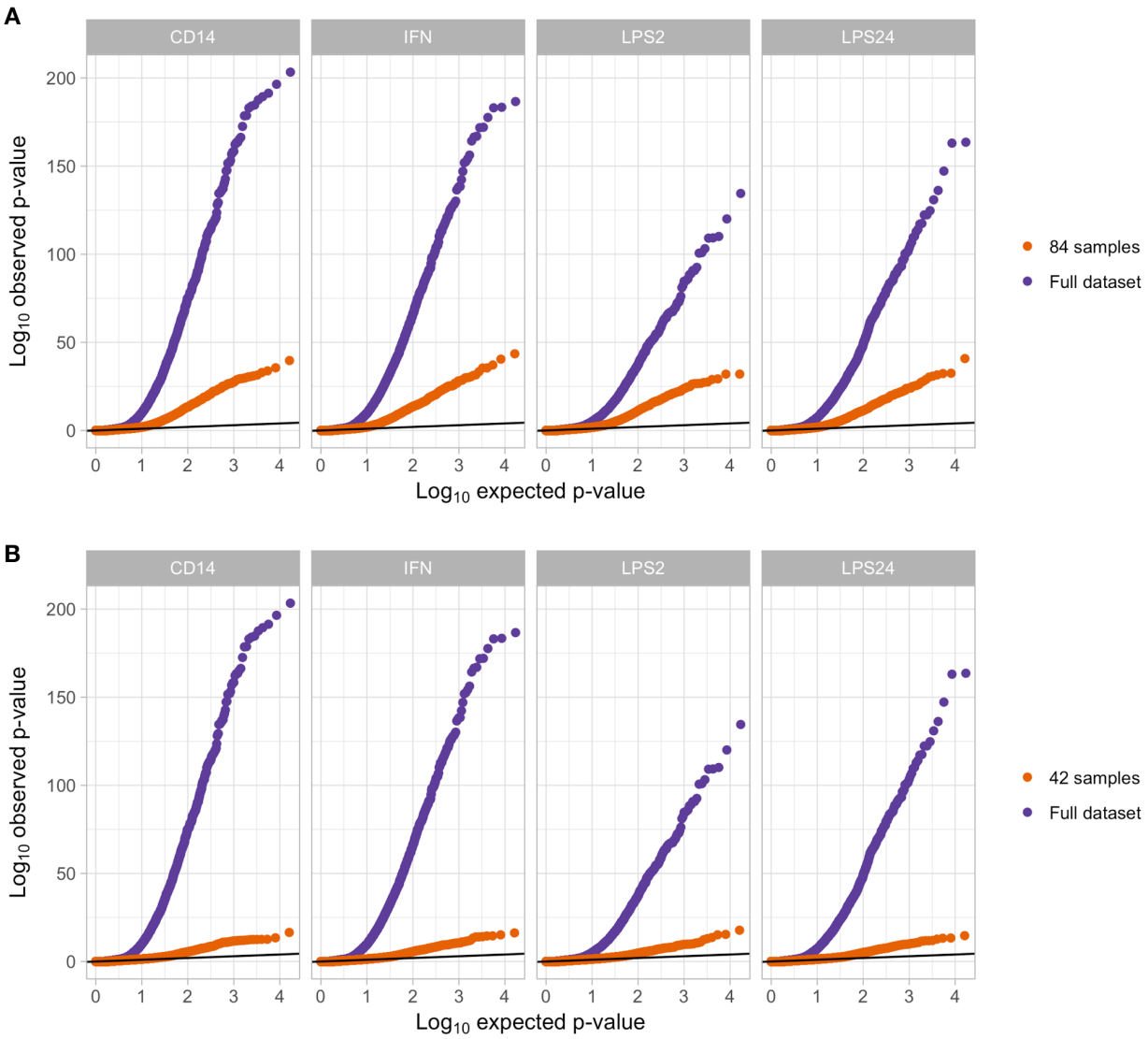
Supplementary Figure 16. TRAF1 eQTL in the Fairfax dataset. (**A**) Association between TRAF1 expression and rs10985070 variant in the four conditions. Box plots show the median (center line) and the 25th and 75th percentiles (box edges), with whiskers extending to 1.5 times the interquartile range. (**B**) Manhattan plots of the RA GWAS signal (top panel) and TRAF1 eQTL in the four conditions. The -$\log_{10}$ p-values were calculated with QTLtools using either 414 (CD14 condition), 367 (IFN condition), 261 (LPS2 condition) or 322 (LPS24 condition) independent donors.



Supplementary Figure 17. Dissecting the Alzheimer's disease (AD) causal variant at the PTK2B locus. (**A**) Manhattan plots for the AD GWAS hit (top panel), colocalised caQTL (second panel) and colocalised eQTL for PTK2B gene (third panel). The bottom two tracks show all ATAC-seq peaks in the region as well as exons of the PTK2B gene. (**B**) ATAC-seq fragment coverage stratified by the lead caQTL and GWAS variant rs28834970. The lead variant was the only associated variant lying within the regulated caQTL peak, suggesting that this is the most likely causal variant. The lead variant rs28834970 is T/C polymorphism and the alternative C allele decreases the relative binding score of the CEBPα TF motif (M1925_1.02 in CIS-BP[6]) from 0.87 to 0.84. This is consistent with the decreased chromatin accessibility at the C allele (middle panel) as well as decreased expression of the PTK2B gene (panel C). The variant also overlaps experimental CEBPβ ChIP-seq peak in primary human macrophages (bottom panel)[7]. (**C**) Expression of the PTK2B gene stratified by the rs28834970 genotype. The caQTL (n = 42 independent donors) and eQTL (n = 84 independent donors) -$\log_{10}$ p-values were calculated using RASQUAL. Box plots show the median (center line) and the 25th and 75th percentiles (box edges), with whiskers extending to 1.5 times the interquartile range.



Supplementary Figure 18. Left: Histogram of the number of caQTLs associated with each eQTL gene (Pearson correlation $R^2 > 0.8$ between lead variants). **Right:** Histogram of the number of eQTLs associated with each caQTL region (Pearson correlation $R^2 > 0.8$ between lead variants).



Supplementary Figure 19. Quantile-quantile plots for the eQTL detected in the full Fairfax dataset and two subsets of either 84 independent donors (**A**) or 42 independent donors (**B**)**. T**he full dataset consisted of either 414 (CD14 condition), 367 (IFN condition), 261 (LPS2 condition) or 322 (LPS24 condition) independent donors. Each point represents the p-value of the lead eQTL variant for each gene. The -$\log_{10}$ p-values were calculated with QTLtools. The p-values have been adjusted with eigenMT[4] to account for multiple variants tested per gene.

Supplementary Tables

Supplementary Table 1. Metadata for all iPSC to macrophage differentiation attempts.

Supplementary Table 2. Estimating the fraction of eQTL that are likely to be false positives due to small sample size. The eQTL mapping was performed with QTLtools (see Supplementary Note).

Discovery sample size	Replication sample size	Condition	Fraction of discovery genes significant in replication
42	228	CD14	0.94
		IFN	0.94
		LPS2	0.93
		LPS24	0.94
84	228	CD14	0.93
		IFN	0.93
		LPS2	0.91
		LPS24	0.93

Supplementary Table 3. Estimating the fraction of response eQTLs that are likely to be false positives due to small sample size. The eQTL mapping was performed with QTLtools and response eQTLs were identified using a linear model interaction test implemented in R (see Supplementary Note).

Discovery sample size	Replication sample size	Fraction of discovery genes significant in replication
42	228	0.825
84	228	0.934

Supplementary Table 4. Enrichment of caQTLs and eQTLs among GWAS hits. GARFIELD was used to calculate fold enrichments and p-values for 10 different GWAS traits.

Supplementary Table 5. Colocalisation between eQTLs and GWAS associations. The coloc package was used to test for colocalisation between GWAS summary statistics for ten immune-mediated traits and eQTLs identified in all four conditions (n = 84 independent donors).

Supplementary Table 6. Colocalisation between caQTLs and GWAS associations. The coloc package was used to test for colocalisation between GWAS summary statistics for ten immune-mediated traits and caQTLs identified in all four conditions. The number of independent donors in the ATAC-seq data was n = 42 (N condition), n = 41 (I condition) and n = 31 (S and I+S conditions).

Supplementary Table 7. Metadata for the RNA-seq samples.

Supplementary Table 8. Metadata for the ATAC-seq samples.

Supplementary Table 9. List of published software packages used in this study.

Name	**Website**
bedtools (v2.17.0)[8]	http://bedtools.readthedocs.io/
BWA (v0.7.12)[9]	http://bio-bwa.sourceforge.net/
skewer (v0.1.127)[10]	https://github.com/relipmoc/skewer
MACS2 (v2.1.0)[11]	https://github.com/taoliu/MACS
Mfuzz (v.2.28)[12]	http://bioconductor.org/packages/Mfuzz/
coloc (v2.3-1)[13]	https://github.com/chr1swallace/coloc
TFBSTools (v1.10.4) [14]	http://bioconductor.org/packages/TFBSTools/
RASQUAL[15]	https://github.com/dg13/rasqual
FastQTL[16]	http://fastqtl.sourceforge.net/
ASEReadCounter[17]	https://software.broadinstitute.org/gatk/
limma voom (v3.26.3)[18]	https://bioconductor.org/packages/limma/

DESeq2 (v1.10.0)[19]	https://bioconductor.org/packages/DESeq2/
FIMO[20]	http://tools.genouest.org/tools/meme/
bcftools	http://www.htslib.org/
SNPRelate[21]	https://bioconductor.org/packages/SNPRelate/
samtools[22]	http://www.htslib.org/
verifyBamID (v1.1.2)[23]	http://genome.sph.umich.edu/wiki/VerifyBamID
PEER[24]	https://github.com/PMBio/peer
STAR (v2.4.0j)[25]	https://github.com/alexdobin/STAR
cqn[26]	http://bioconductor.org/packages/cqn/
GARFIELD[5]	http://www.ebi.ac.uk/birney-srv/GARFIELD/
GAT[27]	http://gat.readthedocs.io/
Picard	https://broadinstitute.github.io/picard/
eigenMT[4]	http://joed3.github.io/eigenMT/
CrossMap (v0.1.8)[28]	http://crossmap.sourceforge.net/
featureCounts (v.1.5.0)[29]	http://subread.sourceforge.net/
LiftOver	http://genome.sph.umich.edu/wiki/LiftOver
ggplot2[30]	http://ggplot2.org/
wiggleplotr[31]	https://bioconductor.org/packages/wiggleplotr/
lme4[32]	https://github.com/lme4/lme4/
QTLtools[33]	https://qtltools.github.io/qtltools/

aFC[34]	https://github.com/secastel/aFC

Supplementary Note

HIPSCI consortium members

Helena Kilpinen[1], Angela Goncalves[2], Andreas Leha[2,10], Vackar Afzal[3], Kaur Alasoo[2], Sofie Ashford[4], Sendu Bala[2], Dalila Bensaddek[3], Francesco Paolo Casale[1], Oliver J Culley[5], Petr Danecek[2], Adam Faulconbridge[1], Peter W Harrison[1], Annie Kathuria[5], Davis McCarthy[1,9], Shane A McCarthy[2], Ruta Meleckyte[5], Yasin Memari[2], Nathalie Moens[5], Filipa Soares[6], Alice Mann[2], Ian Streeter[1], Chukwuma A Agu[2], Alex Alderton[2], Rachel Nelson[2], Sarah Harper[2], Minal Patel[2], Alistair White[2], Sharad R Patel[2], Laura Clarke[1], Reena Halai[2], Christopher M Kirton[2], Anja Kolb-Kokocinski[2], Philip Beales[8], Ewan Birney[1], Davide Danovi[5], Angus I Lamond[3], Willem H Ouwehand[2,4,7], Ludovic Vallier[2.6], Fiona M Watt[5], Richard Durbin[2,11],Oliver Stegle[1], Daniel J Gaffney[2]

[1]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom.
[2]Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kingdom.
[3]Centre for Gene Regulation & Expression, School of Life Sciences, University of Dundee, DD1 5EH, United Kingdom.
[4]Department of Haematology, University of Cambridge, Cambridge, United Kingdom.
[5]Centre for Stem Cells & Regenerative Medicine, King's College London, Tower Wing, Guy's Hospital, Great Maze Pond, London SE1 9RT, United Kingdom.
[6]Wellcome Trust and MRC Cambridge Stem Cell Institute and Biomedical Research Centre, Anne McLaren Laboratory, University of Cambridge, CB2 0SZ, United Kingdom.
[7]NHS Blood and Transplant, Cambridge Biomedical Campus, Cambridge, United Kingdom.
[8]UCL Great Ormond Street Institute of Child Health, University College London, London WC1N 1EH, United Kingdom.
[9]St Vincent's Institute of Medical Research, Fitzroy Victoria 3065, Australia.
[10]Department of Medical Statistics, University Medical Center Göttingen, Humboldtallee 32, 37073 Göttingen, Germany.
[11]Department of Genetics, University of Cambridge, Cambridge, United Kingdom

Cell culture

Feeder-free iPSC culture

Feeder-free iPSCs were grown on tissue culture treated plates coated with vitronectin (VTN-N) (Gibco, cat. no. A14700) in Essential 8 (E8) medium (Gibco). The cells were dissociated from the plates using Gentle Cell Dissociation Buffer (Stemcell Technologies, cat. no. 07174) and passaged every 3-5 days. Prior to macrophage differentiation, the feeder-free iPSCs were first transferred to feeder-dependent media and propagated for at least two passages. This step was necessary because multiple attempts to differentiate macrophage directly from feeder-free iPSCs with our protocol failed.

Feeder-dependent iPSC culture

Feeder-dependent iPSCs were grown on irradiated CF-1 mouse embryonic fibroblast (MEF) feeder cells (AMS Biotechnology) in Advanced DMEM-F12 (Gibco) supplemented with 20% Knock-Out Serum Replacement (KSR) (Gibco), 2mM L-glutamine (Sigma), 50 IU/ml penicillin (Sigma), 50 IU/ml Streptomycin (Sigma) and 50μM β-Mercaptoethanol (Sigma M6250). The media was supplemented with 4 ng/ml recombinant human fibroblast growth factor (rhFGF) basic (R&D, 233-FB-025) to maintain pluripotency and was changed daily. MEFs were seeded on 0.1% gelatine-coated tissue-culture treated plates (Corning 6-well or 10 cm plates) 24 hours prior to passaging iPSCs at a cell density of 2 million cells/6-well or 10-cm plate in Advanced DMEM-F12 supplemented with 10% FBS (labtech), 2mM L-glutamine (Sigma), 50 IU/ml Penicillin and 50 IU/ml Streptomycin (Sigma). Prior to passaging or embryoid body formation, iPSCs were dissociated from the plates using 1:1 mixture of collagenase (1 mg/ml) and dispase (1 mg/ml) (both Gibco).

Macrophage differentiation protocol

iPSCs were differentiated into macrophages using a previously published protocol[35] involving 3 stages: i) embryoid body (EB) formation, ii) generation of monocyte-like myeloid progenitors from the EBs and iii) terminal differentiation of the progenitors into macrophages. For EB formation, iPSC colonies were treated with 1:1 mixture of collagenase (1 mg/ml) and dispase (1 mg/ml) and intact colonies were transferred to low-adherence plates (Sterilin). The colonies were cultured in feeder-dependent iPSC medium without rhFGF for 3 days. On day 3, the EBs were harvested and transferred to gelatinised tissue-culture treated 10 cm plates in serum-free haematopoietic medium (Lonza X-VIVO 15), supplemented with 2 mM L-glutamine (Sigma), 50 IU/ml penicillin, 50 IU/ml streptomycin (Sigma), 50 μM β-Mercaptoethanol (Sigma M6250), 50 ng/ml macrophage colony stimulating factor (M-CSF) (R&D) and 25 ng/ml interleukin-3 (IL-3) (R&D). EBs were maintained in these plates with media changes every 3-5 days for 4-6 weeks until the progenitor cells appeared in the supernatant. Progenitor cells were harvested from the supernatant, filtered through a 40μm cell strainer (BD 352340), centrifuged at 1200 rpm for 5 minutes, counted, and plated in RPMI 1640 (Gibco) supplemented with 10% FBS (labtech), 2mM L-glutamine (Sigma) and 100 ng/ml hM-CSF (R&D) at a cell density of 150,000 cells per well on a 6-well plate and differentiated for another 7 days.

Salmonella infection and IFNɣ stimulation

Two wells of a 6-well plate were used per condition to ensure sufficient amount of RNA. On day 6 of macrophage differentiation, medium was changed for all wells with half of the wells receiving macrophage differentiation media (with M-CSF) and the other half of the cells receiving macrophage differentiation media supplemented with 20 ng/ml IFNɣ (R&D) and M-CSF. After 18 hours, cells from two wells of the naive and IFNɣ conditions were harvested for RNA extraction. The remaining two wells from each condition were additionally infected with *Salmonella enterica* serovar Typhimurium SL1344 (hereafter *Salmonella*) for 5 hours.

Two days before infection, *Salmonella* culture was inoculated in 10 ml low salt LB broth and incubated overnight in a shaking incubator (200 rpm) at 37°C. Next morning, the culture was diluted 1:100 into 10 ml of fresh LB broth and incubated again in a shaking incubator. In the afternoon the culture was diluted once more 1:100 into 45 ml of LB broth and kept overnight in a static incubator. In the morning before infection, the culture was centrifuged at 4000 rpm for 10 minutes, washed once with 4°C PBS and resuspended in 30 ml of PBS. Subsequently, optical density at 600 nm was measured and *Salmonella* was diluted in macrophage differentiation media (without MCSF) at multiplicity of infection (MOI) 10 assuming 300,000 cells per well. To infect the cells, old media was removed and replaced with 1 ml of media containing *Salmonella* for 45 minutes. Subsequently, the cells were washed twice with PBS and replaced in fresh medium with 50 ng/ml gentamicin (Sigma) to kill extracellular bacteria. After 45 minutes the medium was changed once again to fresh medium containing 10 ng/ml gentamicin.

For RNA extraction, cells were washed once with PBS and lysed in 300 ul of RLT buffer (Qiagen) per one well of a 6-well plate. Lysates from two wells were immediately pooled and stored at -80°C. RNA was extracted using RNA Mini Kit (Qiagen) following manufacturer's instructions and eluted in 35 µl nuclease-free water. RNA concentration was measured using NanoDrop and RNA integrity was measured on Agilent 2100 Bioanalyzer using RNA 6000 Nano total RNA kit.

Data acquisition

RNA-seq

All of the RNA-seq libraries were constructed using poly-A selection. The first 120 RNA-seq libraries from 30 donors were constructed manually using the Illumina TruSeq stranded library preparation kit. The TruSeq libraries were quantified using Bioanalyzer and manually pooled for sequencing. For the remaining 216 samples, we used an automated library construction protocol that was based on the KAPA stranded mRNA-seq kit. The KAPA libraries were quantified using Quant-iT plate reader and pooled automatically using the Beckman Coulter NX-8. The first 16 samples were sequenced on Illumina HiSeq 2500 using V3 chemistry and multiplexed at 4 samples/lane. All of the other samples were sequenced on Illumina HiSeq 2000 using V4 chemistry and multiplexed at 6 samples/lane. All four RNA samples from a single donor were sequenced in the same sequencing batch. Sample metadata is presented in Supplementary Table 7.

ATAC-seq

We used an adapted version of the original ATAC-seq protocol[36]. Approximately 150,000 cells were seeded into 1 well of a 6-well plate and treated identically to the RNA-seq samples. After stimulation, cells were washed once with ice-cold Dulbecco's phosphate buffered saline without calcium and magnesium and incubated for 12 minutes on ice in 500 µl freshly-made sucrose buffer (10 mM Tris-Cl pH 7.5, 3 mM $CaCl_2$, 2mM $MgCl_2$, 0.32 M sucrose). After 12 minutes, 25

µl of 10% Triton-X-100 (final concentration 0.5%) was added and the cells were incubated for another 6 minutes to release the nuclei. The lysate was transferred to 1.5 mL microfuge tube, vortexed briefly, and centrifuged at 300 *g* for 8 minutes at 4°C. All traces of the sucrose lysis buffer were removed before immediately resuspending the nuclei pellet in 50 µL of Nextera tagmentation master mix (Illumina FC-121-1030), comprising 25 µL 2x Tagment DNA buffer, 20 µL nuclease-free water and 5 µL Tagment DNA Enzyme 1. The tagmentation reaction mixture was immediately transferred to a 1.5 mL low adherence microfuge tube and incubated at 37°C for 30 minutes. The tagmentation reaction was stopped by the addition of 250 µL Buffer PB (Qiagen) and PCR was performed as described previously[15]. The tagmented chromatin was then purified using the MinElute PCR purification kit (Qiagen 28004), according to the manufacturer's instructions, eluting in 10 µL of buffer EB (Qiagen). Finally, size selection was performed using agarose gel and SPRI beads[15]. Five samples were pooled per lane and 75 bp paired end reads were sequenced on Illumina HiSeq 2000 using the V4 chemistry.

Flow cytometry

We used flow cytometry to measure the cell surface expression of three canonical macrophage markers: CD14, CD16 (FCGR3A/FCGR3B) and CD206 (MRC1). Macrophages were cultured in 10 cm tissue-culture treated plates and detached from the plates by incubation in 6 mg/ml lidocaine-PBS solution (Sigma L5647) for 30 minutes followed by gentle scraping. From each cell line we harvested between 300,000-500,000 cells. Detached cells were washed in media, centrifuged at 1200 rpm for 5 minutes and resuspended in flow cytometry buffer (2% BSA, 0.001% EDTA in D-PBS) and split into two wells of a 96-well plate. Nonspecific antibody binding sites were blocked by incubating cells with Human TruStain FcX (Biolegend) for 45 minutes and washing with flow cytometry buffer. Half of the cells were stained for 1 hour with the PE-isotype control (BD 555749) antibody. The other half of the cells were co-stained for 1 hour with following three antibodies: CD14-Pacific Blue (BD 558121), CD16-PE (BD 555407), CD206-APC (BD 550889). After staining, the cells were washed three times. Resuspended cells were filtered through cell-strainer cap tubes (BD 352235) and measured on the BD LSRFortessa Cell Analyzer.

Differential gene expression and chromatin accessibility

Differential gene expression

We included 15,797 genes whose mean expression in at least one of the conditions was greater than 0.5 TPM into our differential expression analysis. For each gene, we used likelihood ratio test implemented in DESeq2[19] v1.10.0 (test = "LRT") to test if a model that allowed different mean expression in each condition explained the data better than a null model assuming the same mean expression across conditions. Overall, 8758 genes with Benjamini-Hochberg FDR <

1% and fold change between naive and any one of the stimulated conditions greater than 2 were identified as differentially expressed.

To identify differentially expressed genes with specific expression patterns, we calculated mean quantile-normalised expression level in each condition and standardised the mean expression values across conditions to have zero mean and unit variance. Subsequently, we used c-means fuzzy clustering implemented in MFuzz v.2.28[12] package with parameters 'c = 9, m = 1.5, iter = 1000' to assign the genes into 9 clusters. The number of clusters was chosen iteratively by trialing different numbers and observing which ones led to stable clustering results from independent runs. We ranked the genes in each cluster by their fold change between naive and highest expression conditions and used g:Profiler[1] R package with 'max_set_size = 3000, ordered_query = TRUE, exclude_iea = TRUE' options to identify pathways and Gene Ontology (GO) categories enriched in each cluster.

Differentially chromatin accessibility

We used limma voom v3.26.3[18] with TMM normalisation to identify 63,430 regions that were more than 4-fold differentially accessible (FDR < 0.01) between naive and any one of the stimulated conditions. We only included high quality samples from 16 independent donors (64 samples) (Supplementary Table 8) in the analysis, because we noticed that limma voom was sensitive to additional noise in the lower quality samples. Subsequently, we quantile-normalised the region accessibility data using cqn[26], calculated the mean accessibility of each region in each condition and used Mfuzz v.2.28[12] to cluster the regions into seven distinct activity patterns.

Motif enrichment

We downloaded the CIS-BP[6] human TF motif database from the MEME website and used FIMO[20] to identify the occurrences of all TF motifs within the ATAC consensus peaks with FIMO threshold p-value < 10^{-5}. We also performed the same motif scan for 2 kb promoter sequences upstream of 21,350 human genes (downloaded from the PWMEnrich[37] R package) and used this as the background set. We used Fisher's exact test to identify motifs that occurred significantly more often in macrophage open chromatin regions compared to the background promoter sequences. Because the CIS-BP database contains many redundant motifs, we manually selected 21 representative motifs for downstream analysis corresponding to the major TFs important in macrophage biology: AP-1, IRF-family, ETS-family (PU.1, ELF1, FLI1), NF-κB, CEBPα, CEBPβ, ATF4, CTCF, STAT1, MAFB, MEF2A and USF1. We used Fisher's exact test to identify motifs that were specifically enriched in each cluster of differentially accessible peaks compared to the background of all macrophage ATAC-seq peaks.

ChIP-seq data analysis

Human macrophage PU.1 ChIP-seq data[38] (75 bp single-end reads) and C/EBPβ ChIP-seq data[7] (50 bp paired-end reads) were downloaded from GEO (accessions GSE66594 and

GSE54975, respectively). Macrophage H3K27ac ChIP-seq data[39] (50bp single-end reads) was also downloaded from GEO (GSE43036). Single-end datasets were aligned to the GRCh38 reference genome using bwa aln v0.7.12[9] and paired-end dataset was aligned to the same reference using bwa mem v0.7.12 with the -M flag set. For the paired-end dataset, only properly paired reads were used for downstream analysis. Duplicate reads were removed with Picard v1.134 MarkDuplicates with the 'REMOVE_DUPLICATES=true' parameter set. We used bedtools v2.17.0[8] to construct genome-wide read (single-end) or fragment (paired-end) coverage tracks in BigWig format. For the PU.1 and C/EBPβ data we used MACS2 v2.1.0 with '-q 0.01' option to call narrow peaks. For the H3K27Ac we used MACS2 to call both broad and narrow peaks at 1% FDR. We first removed broad peaks that were detected in only one biological replicate or did not overlap any narrow peaks in the same condition. We then defined the union of broad peaks identified in each condition as the consensus set of H3K27ac peaks that we used in downstream analysis.

Characterisation of QTLs

Sharing of QTL lead variants between conditions

To quantify how often QTL lead variants were shared between two conditions (e.g. A and B), we first identified all features (genes or peaks) with the nominal p-value of the lead variant < 10^{-6} in condition A. Next, for each feature, we took the lead variant in condition B and counted how often were the two lead variants in high LD with each other ($R^2 > 0.8$). We did not impose any p-value threshold for the lead variant in condition B, reasoning that if there is no association in condition B then the lead variant is picked randomly and is therefore unlikely to be in high LD with the lead variant in condition A.

Allele-specific binding at caQTL regions

We used the published PU.1 and C/EBPβ ChIP-seq data described above to quantify the extent of allele-specific binding of these two factors at caQTL regions. We used ASEReadCounter[17] together with the same VCF file that was used in the rest of the analysis to count the number of reads overlapping each allele. Since genotype data was not available for the ChIP-seq samples, it was not possible to distinguish extreme allele-specific binding from homozygous sites. Thus, we restricted our analysis to variants that overlapped at least ten reads in the ChIP-seq dataset, of which at least two supported the allele with lower binding affinity. To calculate correlation with caQTL effect size (π) estimated by RASQUAL (Supplementary Fig. 5), we further limited the analysis to credible set variants that lied within the regulated accessible region.

caQTL overlap with annotated promoters and enhancers

To estimate the overlap with promoters, we counted the number of caQTL regions that were within 500 bp from an annotated transcription start site for either a protein coding gene or a lincRNA in Ensembl 79. To estimate the overlap with putative enhancer elements, we counted

the number of caQTL regions that overlapped the consensus set of macrophage H3K27ac peaks defined above.

caQTL overlap with topologically associated domains

Since we were not aware of a Hi-C dataset from human macrophages, we used annotated topologically associated domains (TADs) from the GM12878 lymphoblastoid cell line[40] as a proxy to test if master-dependent region pairs were more likely to occur within the same TAD than random pairs of regions. First, we obtained the list of chromatin loops identified by Hi-C from GEO (GSE63525_GM12878_primary+replicate_HiCCUPS_looplist.txt file in GSE63525 experiment). We then counted how often at least one loop endpoint was located between master and dependent region pairs, suggesting that a given pair might span two different TADs. We repeated this analysis for randomly selected pairs of accessible chromatin regions that were matched by distance to each other. We found that that true master-dependent pairs were less likely to span multiple TADs compared to random pairs (7.7% (156/2023) vs 11.5% (543/4688), Fisher's exact test $p = 1.26 \times 10^{-6}$).

Overlap with genome-wide association studies

Data preprocessing

We obtained full summary statistics for ten immune-mediated disorders: inflammatory bowel disease (IBD) including ulcerative colitis (UC) and Crohn's disease (CD)[41], Alzheimer's disease (AD)[42], rheumatoid arthritis (RA)[43], systemic lupus erythematosus (SLE)[44], type 1 diabetes (T1D)[45], schizophrenia (SCZ)[46], multiple sclerosis (MS)[47], celiac disease (CEL)[48] and narcolepsy (NAR)[49]. In addition, we used summary statistics from type 2 diabetes (T2D)[50] as a negative control for a trait that should not be specifically enriched in macrophages. Summary statistics for T1D, CEL, IBD, RA, AD, MS and SLE were downloaded in 2015. SCZ, T2D and NAR were downloaded in 2016. T2D summary statistics were converted from GRCh36 to GRCh37 coordinates using the LiftOver tool, all of the other summary statistics already used GRCh37 coordinates.

Summary statistics for Alzheimer's disease were downloaded from International Genomics of Alzheimer's Project (IGAP)[42]. IGAP is a large two-stage study based upon genome-wide association studies (GWAS) on individuals of European ancestry. In stage 1, IGAP used genotyped and imputed data on 7,055,881 single nucleotide polymorphisms (SNPs) to meta-analyse four previously-published GWAS datasets consisting of 17,008 Alzheimer's disease cases and 37,154 controls (The European Alzheimer's disease Initiative – EADI the Alzheimer Disease Genetics Consortium – ADGC The Cohorts for Heart and Aging Research in Genomic Epidemiology consortium – CHARGE The Genetic and Environmental Risk in AD consortium – GERAD). In stage 2, 11,632 SNPs were genotyped and tested for association in an independent set of 8,572 Alzheimer's disease cases and 11,312 controls. Finally, a meta-analysis was performed combining results from stages 1 & 2.

Enrichment analysis

In each of the four conditions, we took all variants that were associated with either gene expression or chromatin accessibility with nominal p-value < 10^{-5} and used that set of variants as a custom annotation track. We then used GARFIELD[5] to test if these variants were collectivity enriched for GWAS hits for the ten traits described above. We excluded the MHC region (GRCh37: 6:20,000,000-40,000,000) from the GWAS summary statistics prior to enrichment testing, because this region was found to significantly inflate estimates of fold enrichment. We reported fold enrichment at 10^{-5} GWAS significance threshold (Supplementary Fig. 13, Supplementary Table 4).

Analysis of the Fairfax dataset

We obtained processed microarray gene expression data from the Fairfax study[51] from ArrayExpress (E-MTAB-2232). In the Fairfax dataset the gene expression of primary human monocytes was profiled in four conditions (naive, n = 414; 24 hours IFNɣ stimulation, n = 367; 2 hours LPS stimulation, n = 261; 24 hours LPS stimulation, n = 322). Data from all four conditions were available for 228 individuals. Out of the 47,231 probes in the dataset, we retained 26,543 probes that mapped uniquely to 17,755 genes in Ensembl Release 89. We then obtained the GRCh37 start and end coordinates for those genes from Ensembl 74 for eQTL analysis.

The raw genotype data were downloaded from the EGA (EGAS00000000109) and subsequently imputed using the Sanger Imputation Service (https://imputation.sanger.ac.uk/) with Human Reference Consortium v1.1 reference panel. We used 5,368,367 variants with minor allele frequency >= 0.05, INFO score >= 0.4 and Hardy-Weinberg equilibrium chi-square p-value > $1x10^{-6}$ for downstream eQTL and colocalisation analyses.

eQTL mapping

We used QTLtools[33] with the '--grp-best' and '--permute 10000' options to identify gene level eQTLs in the +/-500 kb region around each gene and used Benjamini-Hochberg correction to identify all significant QTLs at the 10% FDR level. We performed eQTL mapping in each condition of the full dataset (up to 414 samples per condition) as well as three different subsets of the data: (i) 228 donors profiled in all four conditions, (ii) a random sample of 84 donors profiled in all four conditions and (iii) a random sample of 42 donors profiled in all four conditions. The quantile-quantile plots of the association p-values are shown on Supplementary Fig. 19. Finally, we used the aFC[34] package to estimate the allelic fold change of each lead eQTL variant in each condition.

Estimating the fraction of false positive eQTLs

To test if the relatively small sample size in our study could lead to inflated rate of false positives, we first identified eQTLs in two small subsamples (n = 42 and n = 84 independent donors) and then asked what fraction of those eQTLs remained significant when we considered the full dataset of 228 shared individuals. We found that consistent with our 10% FDR, over 90% of the eQTLs detected in the small subsamples remained significant in the full dataset (Supplementary Table 2).

Next, we performed the same analysis for response eQTLs. We used identical linear mixed model with first six principal components as covariates that we used for the primary eQTL analysis. We found that 93% of the response eQTLs that were detected with 84 individuals were also detected in the full data set of 228 individuals (Supplementary Table 3). The rate of replication was a bit lower for the smaller subsample consisting of 42 individuals (83%). Thus, false positives caused by small sample size do not seem to seem to be a significant concern in our analysis.

Quantifying the impact of false negative colocalisations

We sought to estimate the proportion of condition-specific colocalisations that were missed in the naive cells due to low power (false negatives). For this purpose, we performed colocalisation analysis with the 84 donors in four conditions. We identified a total of 30 overlaps with GWAS, 18 of which were not detected in the naive condition. We then performed the same colocalisation analysis on the full naive dataset (414 individuals, 5-fold larger) and found that 6/18 colocalisations were now detected in the naive state (Supplementary Fig. 14).

References

1. Reimand, J. *et al.* g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* **44,** W83–W89 (2016).
2. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140,** 805–820 (2010).
3. Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon-gamma: an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75,** 163–189 (2004).
4. Davis, J. R. *et al.* An Efficient Multiple-Testing Adjustment for eQTL Studies that Accounts for Linkage Disequilibrium between Variants. *Am. J. Hum. Genet.* **98,** 216–224 (2016).
5. Iotchkova, V. *et al.* GARFIELD - GWAS Analysis of Regulatory or Functional Information

Enrichment with LD correction. *bioRxiv* 085738 (2016).

6. Weirauch, M. T. *et al.* Determination and Inference of Eukaryotic Transcription Factor Sequence Specificity. *Cell* **158,** 1431–1443 (2014).
7. Reschen, M. E. *et al.* Lipid-induced epigenomic changes in human macrophages identify a coronary artery disease-associated variant that regulates PPAP2B Expression through Altered C/EBP-beta binding. *PLoS Genet.* **11,** e1005061 (2015).
8. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
9. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).
10. Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15,** 182 (2014).
11. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
12. Kumar, L. & E Futschik, M. Mfuzz: a software package for soft clustering of microarray data. *Bioinformation* **2,** 5–7 (2007).
13. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* **10,** e1004383 (2014).
14. Tan, G. & Lenhard, B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32,** 1555–1556 (2016).
15. Kumasaka, N., Knights, A. J. & Gaffney, D. J. Fine-mapping cellular QTLs with RASQUAL and ATAC-seq. *Nat. Genet.* **48,** 206–213 (2016).
16. Ongen, H., Buil, A., Brown, A. A., Dermitzakis, E. T. & Delaneau, O. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* **32,** 1479–1485 (2016).
17. Castel, S., Levy-Moonshine, A., Mohammadi, P., Banks, E. & Lappalainen, T. Tools and best practices for data processing in allelic expression analysis. *Genome Biol.* **16,** 195 (2015).

18. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15,** R29 (2014).
19. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).
20. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27,** 1017–1018 (2011).
21. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28,** 3326–3328 (2012).
22. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
23. Jun, G. *et al.* Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am. J. Hum. Genet.* **91,** 839–848 (2012).
24. Stegle, O., Parts, L., Piipari, M., Winn, J. & Durbin, R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* **7,** 500–507 (2012).
25. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
26. Hansen, K. D., Irizarry, R. A. & Wu, Z. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* **13,** 204–216 (2012).
27. Heger, A., Webber, C., Goodson, M., Ponting, C. P. & Lunter, G. GAT: a simulation framework for testing the association of genomic intervals. *Bioinformatics* **29,** 2046–2048 (2013).
28. Zhao, H. *et al.* CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30,** 1006–1007 (2014).
29. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30,** 923–930 (2014).
30. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer, 2009).

31. Alasoo, K. *wiggleplotr: Make read coverage plots from BigWig files*. (Bioconductor, 2017). doi:10.18129/B9.bioc.wiggleplotr

32. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67,** 1–48 (2015).

33. Delaneau, O. *et al.* A complete tool set for molecular QTL discovery and analysis. *Nat. Commun.* **8,** 15452 (2017).

34. Mohammadi, P., Castel, S. E., Brown, A. A. & Lappalainen, T. Quantifying the regulatory effect size of cis-acting genetic variation using allelic fold change. *Genome Res.* **27,** 1872–1884 (2017).

35. van Wilgenburg, B., Browne, C., Vowles, J. & Cowley, S. A. Efficient, long term production of monocyte-derived macrophages from human pluripotent stem cells under partly-defined and fully-defined conditions. *PLoS One* **8,** e71098 (2013).

36. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10,** 1213–1218 (2013).

37. Stojnic, R. & Diez, D. *PWMEnrich: PWM enrichment analysis. R package version 4.8.2*. (Bioconductor, 2015). doi:10.18129/B9.bioc.PWMEnrich

38. Schmidt, S. V. *et al.* The transcriptional regulator network of human inflammatory macrophages is defined by open chromatin. *Cell Res.* **26,** 151–170 (2016).

39. Qiao, Y. *et al.* Synergistic activation of inflammatory cytokine genes by interferon-γ-induced chromatin remodeling and toll-like receptor signaling. *Immunity* **39,** 454–469 (2013).

40. Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159,** 1665–1680 (2014).

41. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47,** 979–986 (2015).

42. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45,** 1452–1458 (2013).
43. Okada, Y. *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506,** 376–381 (2014).
44. Bentham, J. *et al.* Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.* **47,** 1457–1464 (2015).
45. Onengut-Gumuscu, S. *et al.* Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat. Genet.* **47,** 381–386 (2015).
46. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511,** 421–427 (2014).
47. International Multiple Sclerosis Genetics Consortium (IMSGC) *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.* **45,** 1353–1360 (2013).
48. Trynka, G. *et al.* Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat. Genet.* **43,** 1193–1201 (2011).
49. Faraco, J. *et al.* ImmunoChip study implicates antigen presentation to T cells in narcolepsy. *PLoS Genet.* **9,** e1003270 (2013).
50. Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44,** 981–990 (2012).
51. Fairfax, B. P. *et al.* Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* **343,** 1246949 (2014).