Supporting Information

Ferraro et al. 10.1073/pnas.1401343111

SI Materials and Methods

Blood Donors. Blood samples were obtained in the context of the Juvenile Diabetes Foundation-supported Human Sample Core at the Joslin Diabetes Center. We collected a sample set from 229 individuals with self-reported European family origins, including 83 type-1 diabetes (T1D) patients, 46 type-2 diabetes (T2D) patients, and 100 healthy controls. Recruiting was split into three consecutive cohorts over 3 y. Six donors from cohort 2 were sampled on two independent occasions. The donor's questionnaire included health and family history. Age, sex, body-mass index (BMI), age of T1D onset and duration, glycated hemoglobin (HbA1c), and insulin requirement were recorded.

For sample collection, a strict standard operating protocol (SOP) was adopted to minimize environmental and circadian variation and to standardize and minimize the time between blood collection and cell sorting. Donors were excluded if self-reporting an infectious event in the 7 d before sampling, and donors were excluded if presenting with inflammatory or autoimmune pathologies other than diabetes (celiac disease is frequent among T1D patients at the Joslin Diabetes Center and was not an exclusion criterion). All blood samples were collected from nonfasting donors between 9:00 AM and noon and immediately processed (no more than 30 min from vein puncture to loading onto Ficoll gradient).

This study was reviewed and approved by the Institutional Review Board at the Joslin Diabetes Center and Harvard Medical School (JDC/02–15 and HMS/M1479-105), and informed consent was obtained from all subjects.

Cell Preparation, Labeling, and Sorting. Peripheral blood mononuclear cells (PBMCs) were first fractionated by density-gradient centrifugation on a Ficoll-Hypaque solution. For analytical flow cytometry, PBMCs were labeled with anti-CD4 PerCP-Cy5.5 (clone OKT4; Biolegend) and anti-CD25 PE (clone BC96; Biolegend) monoclonal antibodies. For Foxp3 and Helios analysis, cells were fixed and permeabilized according to the manufacturer's instructions. Anti-Foxp3 Alexa647 (clone 259D; Biolegend) and anti-Helios Pacific Blue (clone 22F6; Biolegend) monoclonal antibodies were used. Samples were analyzed using a BD LSRII flow cytometer (BD Biosciences) and analyzed by FlowJo (Tree Star). To ensure that results of different flow experiments were comparable, we calibrated the instrument using a negative control and single mAb stainings, with instrument settings such that these controls had the same value in each experiment.

For sorting of regulatory T (Treg) and conventional $CD4^+$ T (Tconv) cells for gene expression profiling, 20×10^6 PBMCs were stained with anti-CD4 Pacific Blue (clone RPA-T4; Biolegend), anti-CD25 PE (clone BC96; Biolegend), anti-CD127 Alexa488 (clone A019D5; Biolegend), and Fc block. Cells were then incubated for 20 min in ice and washed before sorting on a BD FACSAria (BD). Just before sorting, 7AAD was added to the stained PBMCs to exclude dead cells. Sorted cells (first round) were immediately resorted (second round) and collected in 500 μL of Trizol (Invitrogen). Purity of cells after sorting was analyzed, and it was always in the 98–99% range. For each sort, 30,000–50,000 Treg and Tconv cells were collected.

Treg Suppression Assay. To functionally evaluate Treg-cell activity in different donors, we used the in vitro suppression assay as described (1). To ensure reproducible results and avoid variation in the responder cells, responder cells were frozen aliquots from a single healthy donor. Treg-depleted PBMCs from this donor were labeled with 10 μmol/L carboxyfluorescein succinimidyl ester (CFSE) (Molecular Probes) in RPMI 1640 at 10^6/mL at 37 °C for 20 min, washed in RPMI 1640 plus 5% FBS (Omega Scientific), and resuspended in X-vivo 15 medium (Lonza) and 5% pooled AB human serum (Lonza). Treg-depleted PBMCs were then cultured at ratio 1:1 with 50,000 Treg cells in a round-bottom, 96-well plate for 4 d. Stimulation was effected by addition of anti-CD3, anti-CD2, and anti-CD28–coated beads (Treg Inspector; Miltenyi Biotech) at a 1:1 ratio. Proliferation was assessed by flow-cytometric analysis of the CFSE dilution. To calculate the proliferation index, the number of cells (events) in a given cycle (division: n) was multiplied by $2\hat{}n$ and divided for the total number of cells to calculate the mean division number. To measure the percentage of inhibition, the proliferation index of stimulated $CFSE^+$ cells divided in the presence of Treg cells was compared with the proliferation index of stimulated $CFSE^+$ cells cultured alone × 100.

Microarray Profiling. Genome-wide gene expression was quantitated on Affymetrix HuGene 1.0 ST microarrays at Expression Analysis. RNA was amplified for two rounds. For cohorts 1 and 2, amplification and biotin labeling used GeneChip WT cDNA Synthesis and Amplification Kits (Affymetrix) and GeneChip WT Sense Target Labeling and Control Reagents (Affymetrix), respectively. For cohort 3, amplification was performed with the Ambion WT Expression Kit (Invitrogen) and biotin labeling with the GeneChip WT Terminal Labeling Kit (Affymetrix). Probes were purified with the RNeasy Mini kit (Qiagen) before hybridization to the arrays. Image reads were processed through Affymetrix Power Tools software to obtain raw .cel files.

Data Quality Control and Preprocessing of the Microarray Data. Raw .cel files were processed using the Robust Multichip Average algorithm in Affymetrix PowerTools (apt-probeset-summarize function). Per ImmGen SOP, the dynamic range (DR; the ratio between the 95th by the 5th quantile highest of signal values in each data set) was the primary metric of quality; samples with DR <40 were not considered.

For gene assignments, we used release 32 of the Affymetrix probeset annotations file. Of the 33,297 probesets of the ST1.0 array, 6,049 were removed according to the SOP determined for the ImmVar project. Briefly, we used three criteria for probe selection: (*i*) All features that contained a single nucleotide polymorphism (SNP) (with a minimum allele frequency >0.1) were removed before probeset summarization; (*ii*) all probesets for which more than four features were removed, or with >25% of features removed, were dropped entirely; and (*iii*) probesets that do not map to the genome were removed.

In addition, we removed from consideration in the analyses reported here probesets corresponding to the following: (*i*) genes encoded on ChrY; (*ii*) MHC-II genes, particularly at issue here because of the very tight association between T1D and MHC haplotypes, and of the particular linkage disequilibrium structure at the MHC [like others (2), we have observed that the probes for MHC class II genes on the ST1.0 arrays were poorly designed and include many features that vary between haplotypes, such that the signals denote genetic match with the sequence used for the design rather than actual expression levels; we kept the HLA-DRA probesets, whose features do not include SNPs]; (*iii*) a set of genes expressed at high levels in B lymphocytes and monocytes, for which even trace contamination might be reflected

as differential expression (*LYZ, S100A8, VCAN, CYBB, S100A12, CD36, FCGR2A, IRAK3, MPEG1, FCER1G, MS4A1, IGJ, BANK1, CD180, IGHM, IGKC, IGKV3D-11, GAPT, IGKC*); (*iv*) a group of 572 probesets known to exhibit a high degree of technical variability, as determined in the ImmVar project by profiling in technical duplicates of a panel of 15 RNA samples from sorted blood $CD4^+$ T cells; and (*v*) finally, to ensure robust expression, probesets with postnormalization expression >220 (approximately the median expression value across the three cohorts) in fewer than five individuals were removed from consideration. Altogether, these filtering steps left a total of 14,243 probesets, representing 11,753 unique annotated genes and 1,556 mappable but unannotated ("—") transcripts.

There were substantial variations in the data obtained for each of the three cohorts, resulting from usual batch effects but also compounded because samples from the third cohort were processed with a different labeling kit. Two strategies were used to alleviate this issue. First, several analyses were conducted independently in the three cohorts, and the results were compiled. Second, when the full power of the entire datagroup was needed, we generated a datagroup corrected for batch effects by the following process. (*i*) 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 interbatch differences on normalized data are much lower than those on raw expression values. (*ii*) These normalized values for the three cohorts were assembled and log-transformed, and the value for each gene across all donors was fitted to the batch in a generalized linear model (*glm* function in S-Plus; Insightful). The residuals from this fit were then converted back to pseudo-expression values for comparability with values used in intrabatch analyses (exp10, add 1, multiply by mean expression for that gene).

For some uses where other sources of interindividual variation were to be avoided, such as expression quantitative trait locus (eQTL) mapping, the same *glm* fit procedure was used, additionally including sex, age, BMI, and diagnosis as explanatory variables.

Preprocessed data were analyzed using the GenePattern suite (3) or with statistical tools and custom code written in R or S-Plus.

SNP Genotyping and Data Processing. Genomic DNA was prepared from peripheral blood of each donor by means of phenol-chloroform extraction by the Genetics Core of the Joslin Diabetes Center.

SNP genotyping was performed, for financial limitations, on a subset of 65 donors (split evenly between the three diagnoses). Each subject was genotyped using the Illumina Infinium HumanOmniExpressExome 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 are part of the exomes. We applied rigorous quality control (QC) that includes (*i*) sex misidentification, (*ii*) subject relatedness, (*iii*) Hardy–Weinberg Equilibrium testing, (*iv*) genotype call rate > 95%, (*v*) subject missingness rate >95%, (*vi*) MAF < 0.01, and (*vii*) heterozygosity outlier. To detect outliers in terms of population stratification, we performed principal component (PC) analysis using EIGENSTRAT software (4). We used PLINK v1.07 software for all of the quality control steps (5). After the QC, eight subjects and 84,461 SNPs were filtered out from our analysis.

eQTL Analysis. Associations between 606,150 SNP genotypes and adjusted RNA expression levels for 13,720 Affymetrix Human Gene 1.0 ST probes were conducted using the Spearman rank correlation (SRC) model as previously described (6). To identify *cis*-eQTLs, we tested those SNPs located within a region 1 Mb upstream or downstream of the transcription start site (TSS) of a gene. The significance of nominal *P* values was evaluated by comparison with the tail of the distribution of the most significant *P* values of 10,000 permutations per gene (6, 7). A false discovery rate (FDR) was calculated by taking the ratio of expected false positives (number of genes) at a given threshold over the number of significant associations at the same threshold. An association was considered significant if the *P* value from the analysis of the observed data (nominal *P* value) was lower than the threshold of the 0.001 tail of the distribution of the minimal *P* values from 10,000 permutations of the expression levels (6, 7). *RPS26* and *RPS23,* which give spurious association to many genomic locations, were removed.

Definition of the Treg Signature. To define a canonical signature of genes most differentially expressed in Treg vs. Tconv cells (Dataset S2), we first computed within each cohort the expression ratio between the mean of all Treg and the mean of all Tconv samples. We flagged for inclusion all probesets with Treg/Tconv FoldChange (FC) >1.75 in at least one cohort and FC >1.5 in the other two cohorts (or the reverse Tconv/Treg ratios), or with a Treg/Tonv FC >2 in at least 20% of donors (to avoid excluding genes with differential expression in only a fraction of the population), selecting 194 and 192 genes over- and underexpressed in Treg cells.

Variability Score. The variability score (VS) was derived to estimate the interindividual variability of each gene within the population, but without the component due to technical variability or short-term variation within an individual. The computation was motivated by the derivation of the Fst estimator of population differentiation in genetics, which is calculated as $(Dist_{Between} - Dist_{Within})/Dist_{Between}$, where $Dist_{Within}$ and $Dist_{Between}$ are averages of pairwise distances separating individuals within a population, or between two different populations, respectively. Here, a coefficients of variation was used as a measure of distance, and $VS^i = (|CV^i_{inter}| - |CV^i_{rep}|)/|CV^i_{inter}|$, where $|CV^i_{rep}|$ is the average of pairwise coefficients of variation from replicate samples of the same donor (drawn at 1- to 27-wk intervals) for $gene^i$, and $|CV^i_{inter}|$ is the average of pairwise coefficients of variation between samples of different donors.

Gene Ontology and Pathway Analysis. Gene Ontology (GO) analysis was performed to determine which functional processes were differentially represented in the Treg signature gene list. We used the "Generic GO Term Mapper," an online tool maintained by the Center for Quantitative Biology at Princeton University (http://go.princeton.edu/cgi-bin/GOTermMapper). Functional analysis of coregulated Treg signature genes was performed with Ingenuity Pathways Analysis (Ingenuity Systems), using the Functional Analysis of a Network function to identify the biological functions in the Ingenuity Knowledge Base that were most significant to the molecules in the network, with a Fisher's exact test to calculate a *P* value for the preponderance of each biological function in the network.

Definition of Coregulated Modules in Treg and Tconv Cells by Sequential Clustering. Modules of genes whose expression varied in concert across individuals were defined by sequential clustering. First, a correlation matrix (Pearson coefficients) was calculated for all 386 canonical Treg signature genes, from the patterns of interindividual variation in the Treg and Tconv datasets (all cohorts, glm-smoothed for batch). After removing genes with little or no connectivity (≤2 connections with a coefficient >0.5), the 316-gene matrix was used as input for partition clustering (*pam*

function in S-Plus, initial k = 12), yielding a first cluster structure. These clusters were then tested for homogeneity within the correlation matrix generated from the Tconv data (dispersion estimated by the difference between the 0.08 and 0.98 quantiles of correlation coefficients within a cluster). Clusters with comparable dispersion in Treg and Tconv datasets were left untouched; those with 20% higher dispersion in the Tconv data were split on the basis of correlation in the Tconv correlation space (*pam*, k = 2). Examination of the clusters obtained showed no association for batch, with the exception of two clusters that were clearly driven by remaining batch effects and were removed from consideration (leaving 290 genes).

Relating Gene Expression to Type-1 Diabetes. ***Single transcript association with T1D.*** To compare the expression of single genes in T1D patients and controls, a simple Welch modified *t* test was performed. Because the expression of many Treg signature genes is related to age and because T1D donors were overall younger than T2D donors and corresponding controls, we either (*i*) restricted this comparison with donors <45 y of age (from all cohorts), or (*ii*) used as input the residuals of a glm fit of expression vs. age (all donors). Both approaches yielded very similar results, as did *t* tests run individually from data of each of the three cohorts. *P* values were –log10 transformed for display.

Comparing Treg gene clusters in T1D and control patients. To compare the connectivity between coregulated gene clusters in T1D and control donors, gene–gene Pearson correlation matrices were calculated independently for each set of donors. These correlation values were then plotted.

Association between the Treg signature and T1D. To compare the expression of the Treg signature as a whole, *P* values (*t* test) and FoldChanges between Treg and Tconv cells of T1D patients and healthy controls were calculated in GenePattern. We compared T1D patients and healthy controls in the range of age 18–30 y and the T2D patients and their relative controls in the range of age 45–60 y. Moreover, because also the BMI affects the expression of the Treg signature genes, we repeated the same analysis on the residuals after a glm fit for age, sex, and BMI, which yielded similar results.

1. Ferraro A, et al. (2011) Expansion of Th17 cells and functional defects in T regulatory cells are key features of the pancreatic lymph nodes in patients with type 1 diabetes. *Diabetes* 60(11):2903–2913.
2. Weiner Lachmi K, et al. (2012) DQB1*06:02 allele-specific expression varies by allelic dosage, not narcolepsy status. *Hum Immunol* 73(4):405–410.
3. Reich M, et al. (2006) GenePattern 2.0. *Nat Genet* 38(5):500–501.
4. Price AL, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38(8):904–909.
5. Purcell S, et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559–575.
6. Stranger BE, et al. (2007) Population genomics of human gene expression. *Nat Genet* 39(10):1217–1224.
7. Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138(3):963–971.

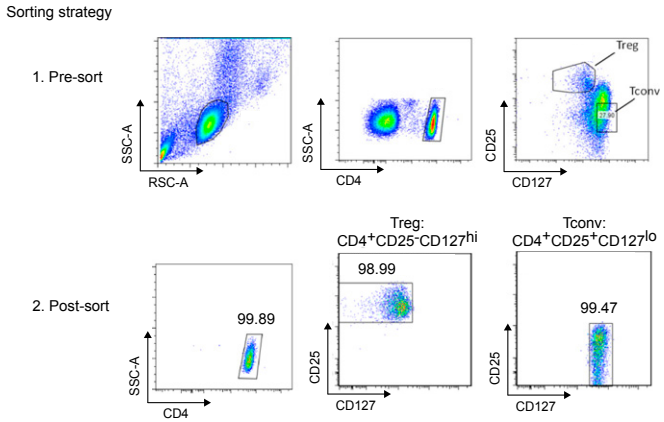

Fig. S1. Representative sorting strategy. (*Upper*) For presort, fresh blood PBMCs were stained and analyzed by flow cytometry. $CD4^+CD25^+CD127^{low/-}$ Treg and $CD4^+CD25^-CD127^+$ Tconv cells were gated as indicated (Tconv cells gated as the highest half of the CD127 profile, and the lowest 2/3 of the $CD25^{-/dull}$ population). (*Lower*) For postsort, shown are examples of populations after the two rounds of cell sorting.

Fig. S2. Comparison of Treg signature gene clusters in Treg cells from T1D patients and matched controls. Pearson correlation coefficients were calculated independently for Treg signature genes in expression datasets from T1D patients ($n = 60$) or matched controls ($n = 56$). The results are displayed as a heatmap, clustered in the same order as in Fig. 2.

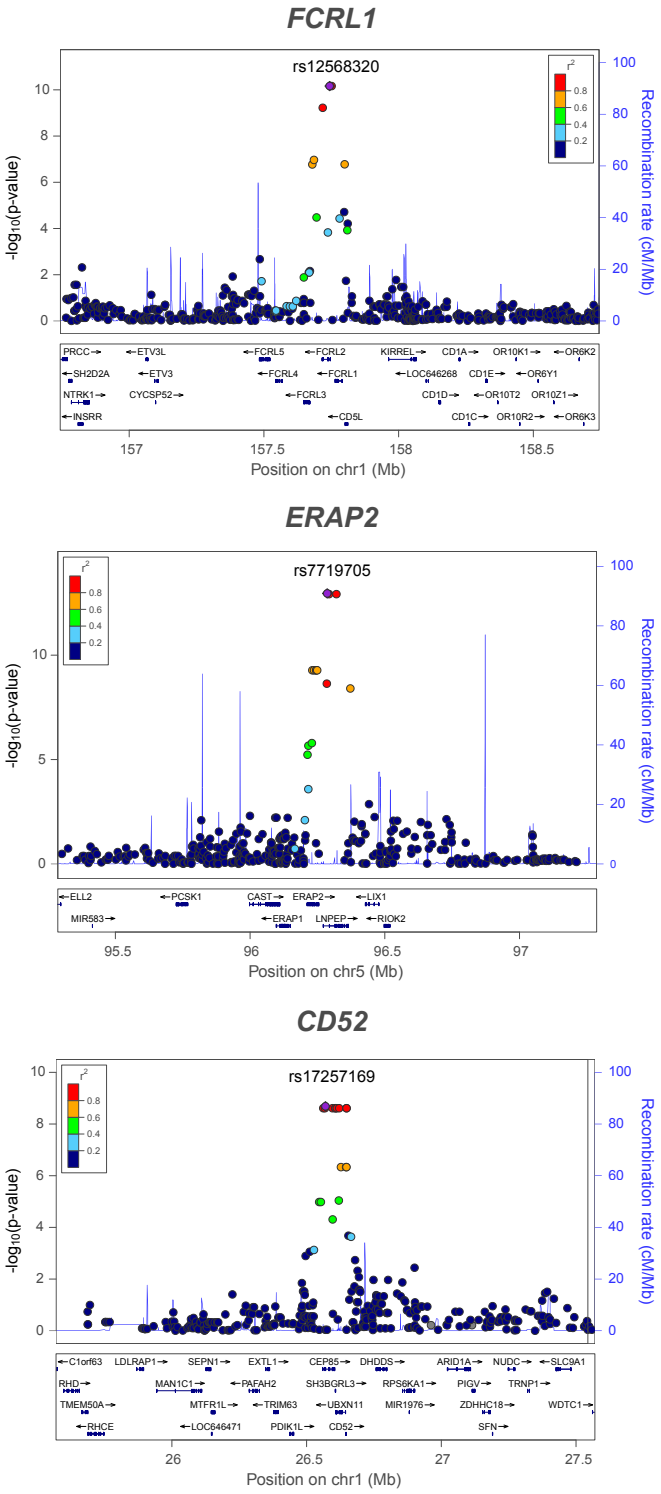

Fig. S3. eQTLs with a genome-wide significance in Treg cells. Representative regional association plots for *FCRL1*, *ERAP2*, and *CD52*. Each dot represents one SNP, positioned according to its significance of association [−log10(*P* value), left scale] and color-coded according to r^2 value. Predicted recombination sites are shown as bars (axis on right).

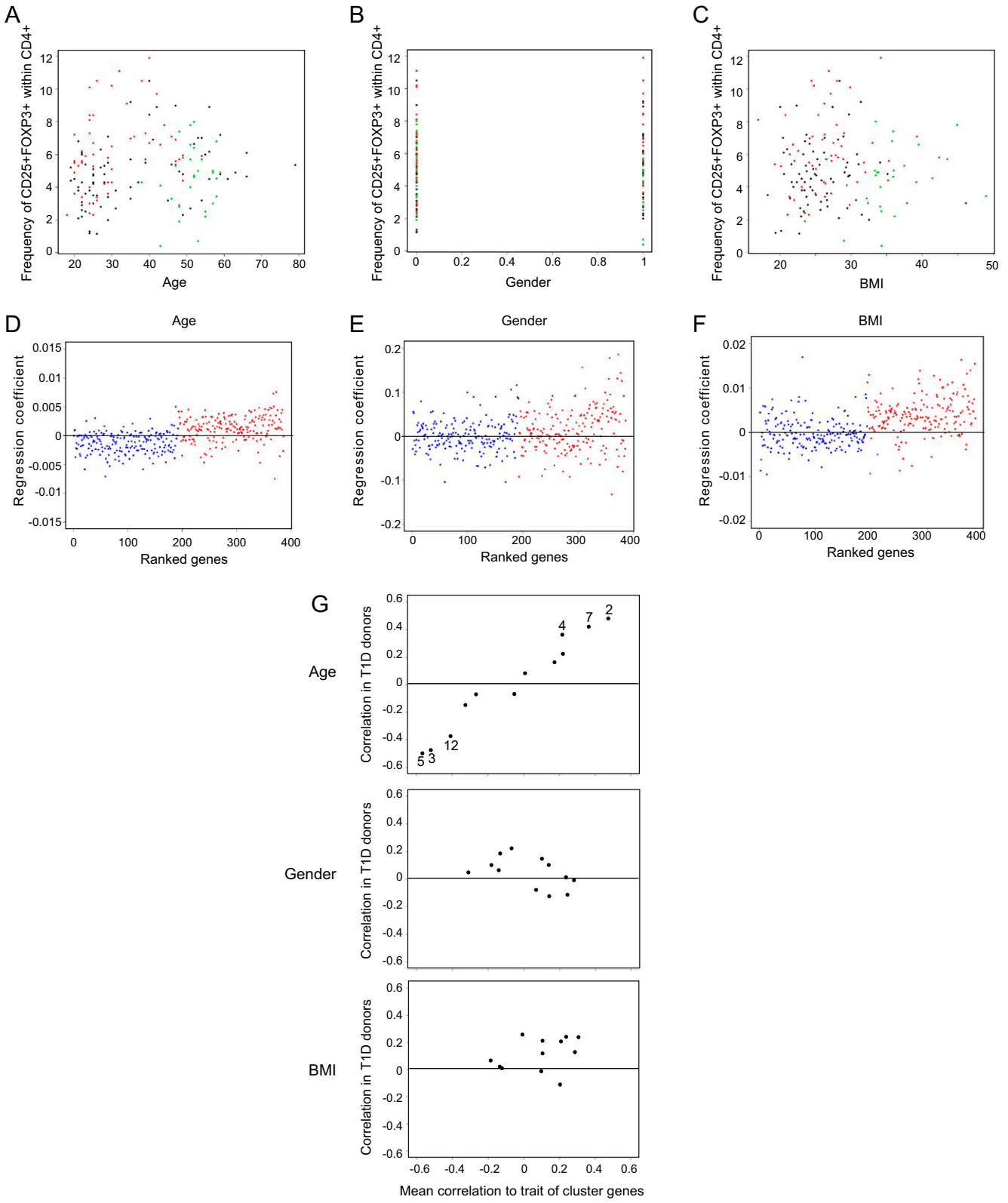

Fig. S4. Correlation of Treg gene expression with age, sex,, and BMI. (*A–C*) Frequency of Tregs (as $CD25^{+}FOXP3^{+}$ within $CD4^{+}$ T cells) is plotted relative to age, sex, and BMI in T1D (red dots), T2D (green dots), and healthy (black dots) donors. For sex, 0 = female and 1 = male. (*D–F*) Gene-expression values in the Treg dataset from healthy controls were fitted in a mixed-model with batch, donor age, sex, and BMI as explanatory variables. The coefficients of the fit for each of the Treg signature genes are plotted, ranked according to mean Treg/Tconv differential expression. (*G*) A "cluster index" was calculated for each cluster in each individual (averaging normalized expression values of all genes in each cluster) and tested for correlation to donor age, sex, and BMI. Correlation coefficients for each cluster are plotted for T1D patients and healthy controls.

Dataset S1. Donors

Dataset S1

Characteristics of the blood donors, grouped by cohort. HbA1c, glycosylated hemoglobin; HC, healthy controls; T1D, type-1 diabetic patients; T2D, type-2 diabetic patients.

Dataset S2. Treg signature genes

Dataset S2

Treg up- and down-regulated genes. Fold change, expression values, FDR *q* values, coefficients of variation, and membership of coregulated clusters are shown.

Dataset S3. Cluster composition

Dataset S3

Composition of the clusters of genes whose interindividual variation is coregulated in Treg and Tconv cells across the 168 donors. The Gene Ontology identifiers most associated with each cluster are shown at right.

Dataset S4. cis-eQTLs in Tconv cells and Treg cells

Dataset S4

eQTLs with genome-wide significance in Tconv cells (Tab 1) and Treg cells (Tab 2). SNP and gene chromosomal location, SNP distance from transcription start site (TSS), and *r* and *P* values are shown.

Dataset S5. eQTLs in Treg signature genes

Dataset S5

The local SNP (±1 Mb from TSS) for each Treg signature gene in both Treg and Tconv cells. SNP distance from transcription start site (TSS) and *P* values are shown.

Dataset S6. eQTL overlap with GWAS hits

Dataset S6

Cis-eQTLs in Treg and Tconv datasets in genes that were identified as associated to autoimmune diseases in genome-wide association studies (GWASs).

Dataset S7. Correlation to Treg number and FOXP3 expression

Dataset S7

Genes correlated to Treg cell number ($CD25^{hi}FOXP3^{+}$ among $CD4^{+}$) and FOXP3 expression (FOXP3 MFI).