T cell exhaustion, costimulation and clinical outcome in autoimmunity and infection

Eoin F. McKinney[1,2], **James C. Lee**[1,2], **David R.W. Jayne**[1], **Paul A. Lyons**[1,2], and **Kenneth G.C. Smith**[1,2]

[1]Department of Medicine, University of Cambridge School of Clinical Medicine, Addenbrooke's Hospital, Hills Road, Cambridge, United Kingdom

[2]Cambridge Institute for Medical Research, Addenbrooke's Hospital, Hills Road, Cambridge, United Kingdom

The clinical course of autoimmune and infectious disease varies greatly even between individuals with the same condition. An understanding of the molecular basis for this heterogeneity could lead to significant improvements in both monitoring and treatment. During chronic infection the process of T cell exhaustion inhibits the immune response, facilitating viral persistence[1]. We show that a transcriptional signature reflecting CD8 T cell exhaustion is associated with poor clearance of chronic viral infection, but conversely predicts better prognosis in multiple autoimmune diseases. The development of CD8 T cell exhaustion during chronic infection is driven by both persistence of antigen and a lack of accessory 'help' signals. In autoimmunity, we found that where evidence of CD4 T cell costimulation was pronounced, that of CD8 T cell exhaustion was reduced. We could reproduce the exhaustion signature by modifying the balance of persistent TCR stimulation and specific CD2-induced costimulation provided to human CD8 T cells *in vitro*, suggesting that each process plays a role in dictating outcome in autoimmune disease. The "non-exhausted" T cell state driven by CD2-induced costimulation was reduced by signals through the exhaustion-associated inhibitory receptor PD-1, suggesting that induction of exhaustion may be a therapeutic strategy in autoimmune and inflammatory disease. Using expression of optimal surrogate markers of costimulation/exhaustion signatures in independent datasets, we confirmed an association with good clinical outcome or response to therapy in infection (hepatitis C virus (HCV)), and vaccination (yellow fever, malaria, influenza) but poor outcome in autoimmune and inflammatory disease (type 1 diabetes (T1D), anti-neutrophil cytoplasmic antibody (ANCA)-associated vasculitis (AAV), systemic lupus erythematosus (SLE), idiopathic pulmonary fibrosis (IPF) and dengue hemorrhagic

fever (DHF)). Thus, T cell exhaustion plays a central role in determining outcome in autoimmune disease and targeted manipulation of this process could lead to new therapeutic opportunities.

In a complex set of data such as the transcriptome, similar measurements may be grouped together by network analysis to form discrete modules that can highlight novel pathways contributing to the pathogenesis of complex diseases. We have previously shown that a CD8 T cell transcriptional signature in patients with multiple immune-mediated diseases can predict a subsequent relapsing disease[2,3]. However, the biology underlying this observation was not clear. We therefore applied weighted gene co-expression network analysis (Extended Data Fig. 1)[3] to the transcriptomes of purified CD4 and CD8 T cells isolated from a prospective cohort of 44 AAV patients with active, untreated disease[7] (Supplementary Table 1) to further explore the mechanisms driving relapsing autoimmunity. Modules of genes (Fig. 1A, rows) were summarized as 'eigengene' profiles (Fig. 1B, F) that were correlated with clinical variables (Fig. 1A, I, columns) and visualized in the form of a heatmap (Fig. 1A, I). Modules derived from both CD8 (Fig. 1A-D) and CD4 (Fig. 1 F-I) T cell transcriptomes showed strong correlation with disease outcome but not activity, and were co- correlated (Fig. 1E) despite being mutually exclusive (Supplementary Table 2). A similar analysis using a cohort of 23 SLE patients also presenting with active, untreated disease (Supplementary Data Table 3)[2] identified analogous CD8 and CD4 T cell expression modules (Extended Data Fig. 2) that again correlated with clinical outcome but not disease activity. By contrast a type 1 interferon response signature was associated with disease activity but not with long-term outcome (Extended Data Fig. 2F), consistent with previous reports[4].

Next, we reasoned that genes within co-correlated modules in related cell types might inform the biology of relapsing disease. By selecting CD4 T cell modules showing significant, strong correlation with relapse rate and performing network enrichment analysis we identified a module corresponding to CD4 T cell costimulation (Extended Data Figs. 1F, G, 3A, Supplementary Tables 2, 4). By way of validation we repeated this analysis using an independent co-expression network algorithm that similarly demonstrated association between a CD4 costimulation module and clinical outcome (Supplementary Table 5). The independent association of modular signatures with clinical outcome (Fig. 1A, I) was confirmed using multiple linear regression modeling (Extended Data Fig. 3B-E) and was only apparent during active disease (Extended Data Fig. 3F, Extended Discussion). During chronic viral infection CD8 T cell memory responses are exquisitely dependent on CD4 T cell costimulation[5,6] which can lead to the resolution of chronic infection in both mice[1] and humans[7]. When antigen persists in the absence of costimulation CD8 T cells become 'exhausted'[1], a phenotype characterized by progressive loss of effector function, persistent high expression of inhibitory receptors and profound changes in gene expression, distinct from those seen in effector, memory or anergic T cells[8]. Although mice lacking inhibitory receptors have an increased incidence and severity of autoimmunity[9,10] a specific role for exhaustion in dictating the outcome of autoimmune responses has not been demonstrated.

We hypothesized that CD4 T cell signals may be important in limiting exhaustion towards persistent *self*-antigen during autoreactive immunity, analogous to responses during

persistent infection. We therefore used Gene Set Enrichment Analysis (GSEA[11]) to test for altered expression of transcriptional signatures reflecting T cell exhaustion (and other T cell-related phenotypes) between patient subgroups defined by the CD8 modular analysis, who go on to develop relapsing or quiescent autoimmunity (Fig2A). Using this approach, we observed that genes specifically downregulated in exhausted CD8 T cells during chronic murine LCMV infection (but not altered in memory, naïve or effector cells, Supplementary Table 6[8]) were similarly downregulated in CD8 T cells from patients at low risk of subsequent relapse (Fig2B, Extended Data Figs. 3G-I, 4).

During chronic murine LCMV infection, T cell exhaustion is driven by coordinate upregulation of multiple coinhibitory receptors[12] that signal synergistically to produce a state of generalized immunosuppression[13]. In autoimmunity, these receptors were not coordinately upregulated as a group. Instead patients with good prognosis from each disease were characterized by upregulation of a distinct subset of exhaustion-associated coinhibitory receptors (Fig2C). Although a divergence from the murine LCMV model, T cell exhaustion accompanied by expression of a limited subset of coinhibitory receptors is similar to that described in intratumoral CD8 T cells[14] which are a target for checkpoint therapy (Extended Data Fig.4I)[15,16].

To confirm whether exhaustion was associated with clinical outcome, we used the murine CD8 T cell exhaustion signature (Supplementary Table 6[8]) to perform unsupervised hierarchical clustering of three independent cohorts of patients with distinct diseases (AAV, Fig2. D-F; SLE, Fig.2 G-I; IBD, Fig. 2 J-L). In each case this identified a subgroup of patients with both early (Fig. 2E, H, K) and recurrent (Fig. 2F, I, L) relapses. Whereas CD8 exhaustion was associated with poor outcome in viral infection, in every case it predicted favorable prognosis in autoimmune and inflammatory disease (Fig.2D-L). Again, independent association with outcome was confirmed using multiple linear regression models (Extended Data Fig.3D, E). Together, these data demonstrate that a transcriptional signature of relative CD8 T cell exhaustion, similar to that determining outcome in chronic viral infection and cancer, is apparent during active, untreated disease in patients with favorable long-term outcome in multiple autoimmune and inflammatory diagnoses.

CD8 T cell exhaustion is characterized by high expression of coinhibitory receptors (such as PD-1[12]) and low expression of nascent memory markers (such as IL7R[17]) and is promoted by both the persistence of antigen[18] and a lack of accessory costimulation[6]. To understand signals driving exhaustion and outcome in autoimmunity, we attempted to recreate the outcome-associated transcriptional signatures using variable TCR signal duration and costimulation of primary human cells *in vitro*. We stimulated purified human CD8 T cells using a magnetic bead conjugated with antibodies targeting costimulatory molecules (Fig. 3A) and measured expression of IL7R and PD-1 expression (Fig. 3B-I, Extended Data Fig. 5A-G) as markers indicating an exhausted phenotype. Comparison between persistent (6 days) and transient TCR stimulation (36 hours) showed that IL7R expression returned on a proportion of cells after several divisions when the TCR stimulus was removed (Fig. 3C) but failed to do so if it persisted (Fig. 3D, G). We then systematically tested whether costimulatory molecules, identified from the CD4 T cell network analysis described above (Fig.1I, Extended Data Figs. 3A, H-K), could overcome the effect of persistent TCR

stimulation during *in vitro* differentiation. We found that specific costimulation with anti-CD2 (Fig. 3E, H), but not with other stimuli such as IFNα or anti-CD40, resulted in maintained IL7R expression, limited upregulation of PD-1 and enhanced cell survival (Fig. 3E, Extended Data Fig. 5L-O).

While CD8 exhaustion is known to limit viral control during chronic infection, exhausted cells may be restored to useful function by blocking inhibitory signaling through PD-1[19]. Enhancing coinhibitory signals is therefore a logical therapeutic strategy in autoimmune disease, aiming to facilitate exhaustion despite high levels of costimulation that would otherwise be predicted to result in an aggressive relapsing disease course. To test this concept, primary human CD8 T cells were costimulated during persistent TCR signaling as above (Fig. 3E) in the presence or absence of a bead-bound Fc-chimeric version of the principal PD-1 ligand, PDL-1 (Fig. 3A, F). When added to CD2-costimulated CD8 T cell cultures, increased PD-1/PDL-1 signaling suppressed differentiation of a non-exhausted IL7R^{hi} subpopulation (Fig.3 F, H, I).

To define the phenotype of T cell exhaustion more robustly, as small numbers of surface markers are insufficient, we analyzed the transcriptome of CD8 T cells exposed to persistent stimulation with and without CD2 signaling (Supplementary Table 7). This CD2 response signature characterized exhausted cells but not effector or memory subsets (by GSEA, Fig. 3J- L). Consistent with this, patient clusters generated using the CD2 response signature recreated subgroups similar to those generated using the murine LCMV CD8 exhaustion signature (Fig. 2D, G, J and Fig. 3M-O). Thus, CD2 signaling during persistent TCR stimulation of primary human CD8 T cells prevents the development of transcriptional changes characteristic of exhaustion, recreating transcriptional signatures associated with outcome in both viral infection and autoimmunity.

To confirm that the transcriptional signatures reflected the development of functional exhaustion *in vitro*, we showed that cells appearing exhausted by surface markers (IL7R^{lo}PD-1^{hi}) also expressed markers of apoptotic resistance, characteristic cytokine patterns and showed diminished survival on restimulation (BCL2^{lo}IFNγ^{lo}IL10^{hi}, Extended Data Fig. 6A-E). There was no evidence of preferential accumulation of CD8 T cell subsets following CD2-induced costimulation (Extended Data Fig. 6F-H). These data highlight the importance of CD2 signaling in limiting the development of CD8 T cell exhaustion in the face of persistent TCR simulation, and provide a starting point for more sophisticated attempts to therapeutically exhaust an autoimmune response in a targeted fashion.

We next aimed to independently validate the association between the balance of CD4 costimulation and CD8 exhaustion with clinical outcome using published datasets. The majority of these profile unseparated peripheral blood mononuclear cells (PBMC), in which T cell-intrinsic signatures are not readily apparent due to the confounding influence of expression from other cell types[20]. We therefore used a classification algorithm (randomforests) to identify optimal surrogate markers of costimulation/exhaustion modules in PBMC data from autoimmune patients taken concurrently with the T cells described above (Fig. 4A). As the CD8 exhaustion and CD4 costimulation signatures were themselves correlated (Extended Data Fig. 3G-I), it became easier to detect their combined signal in

PBMC using surrogate markers (Fig.4A, Extended Data Fig.7). The top-ranked candidate KAT2B is a transcriptional co-activator known to mediate an anti-apoptotic effect under conditions of metabolic stress[52] and to increase cellular resistance to cytotoxic compounds[53]. These characteristics, along with its high expression in memory and T-follicular helper and NK cells (Extended Data Fig. 8), suggest that it may mark the development of a durable, persistent T cell phenotype promoting long-lived responses in either infection or autoimmunity. The observed association was confirmed by both technical replication (using the same samples run on an independent array platform) and independent validation (Fig. 4B).

To test whether similar associations may be apparent in multiple infectious and autoimmune diseases we directly compared expression levels of KAT2B (and of the other top surrogate markers, Extended Data Fig. 9) between clinical subgroups defined within published studies for which PBMC expression and linked clinical outcome data were available. Where subgroups were not pre-specified, we compared clinical outcome in groups stratified as having either above or below-median expression of KAT2B (Fig. 4C-K). Hierarchical clustering using all top surrogate markers gave similar stratification to that seen using KAT2B alone, while as expected the separation of patient subgroups varied slightly in different clinical circumstances (Fig. 4 C-K, Extended Data Fig. 9).

Combined interferon and ribavirin therapy may result in increased virus-specific T cell responses in chronic HCV, although such eradication therapy is successful in only 50% of cases[21] and in some cases no change in endogenous immune response is observed[22]. In a cohort of hepatitis C patients receiving combination therapy, KAT2B expression was progressively induced and showed significantly greater induction in patients ultimately responding to therapy (Fig. 4C, Extended Data Fig. 10A). In a clinical trial of malaria vaccination[23] high KAT2B expression identified a subgroup with response rates of 78%, almost twice that seen in the low response group (Fig. 4D, Extended Data Fig. 10B-D). Moreover, response to vaccination for either influenza[24] (Fig. 4E, Extended Data Fig. 10E-F) or yellow fever[25] (Fig. 4F) could be predicted by stratifying recipients based on their expression of KAT2B following vaccine exposure. Dengue viral infection can result in a wide range of clinical manifestations ranging from asymptomatic infection or self-limiting fever (uncomplicated dengue, UD) to hemorrhagic fever (DHF). Consistent with our observations in autoimmunity, we observed that KAT2B expression was elevated in patients developing the excessive inflammatory response of DHF (Fig. 4G)[26].

We next asked whether surrogate detection of T cell costimulation/exhaustion modules could predict progression of other autoimmune diseases. Idiopathic pulmonary fibrosis (IPF) is a progressive interstitial lung disease characterized by both autoantibodies and autoreactive CD4 T cells[27]. In a cohort of 75 IPF patients[28] high expression of KAT2B predicted subsequent progression to transplantation or death (Fig. 4H). We also observed that PBMC Kat2b expression was elevated in the murine NOD model of T1D[29] with levels rising sharply during the T cell initiation phase, long before the onset of diabetic hyperglycemia (Extended Data Fig. 10G). In a cohort of samples taken prospectively from children at high risk of disease but prior to its onset[30] expression of KAT2B was seen to

specifically and progressively rise (Fig. 4I-K) both in patients who progressed to T1D and in those who developed islet-cell autoantibodies.

We show that the balance between costimulatory and coinhibitory signals that shape T cell exhaustion coincide with opposite clinical outcomes during autoreactive and anti-viral immunity. This at once allows prediction of outcome during infection and autoimmunity and creates the potential for targeted therapeutic exhaustion of an autoimmune response in those predicted to follow an aggressive disease course. That this association is apparent in multiple autoimmune and inflammatory diseases emphasizes the importance of signals shaping T cell exhaustion in driving risk of relapse or recurrence (prognosis) rather than disease susceptibility (diagnosis) or immediate severity (disease activity), and suggests that targeted manipulation of these processes may lead to new treatment strategies that extend beyond the conditions discussed here.

Methods

Patients

Ethical approval—Ethical approval for this study was obtained from the Cambridge Local Research

Ethics Committee (REC reference numbers 04/023, 08/H0306/21,08/H0308/176) and informed consent was obtained from all subjects enrolled.

AAV patients

59 AAV patients attending or referred to the specialist vasculitis unit at Addenbrooke's hospital, Cambridge, UK between July 2004 and May 2008 were enrolled into the present study. Active disease at presentation was defined by Birmingham vasculitis activity score (BVAS[31]) and the clinical impression that induction immunosuppression would be required. Prospective disease monitoring was undertaken monthly with serial BVAS disease scoring[31] and full biochemical, hematological and immunological profiling followed by treatment with an immunosuppressant and tapering dose steroid therapy (Supplementary Table 1). At each time-point of follow-up, disease activity was allocated into one of three categories defined as follows:

1. Flare (at least 1 major or 3 minor BVAS criteria),

2. low grade activity (0 major and 1-2 minor BVAS criteria),

3. no activity (0 major or minor BVAS criteria).

All disease flares were crosschecked against patient records to confirm clinical impression of disease activity and the need for intensified therapy as a result. Disease activity scoring was performed by a single investigator (EFM) who was blinded to gene expression data at the time of scoring. Additional flares were defined in the absence of BVAS scoring if patients attended for emergency investigation (bronchoscopy, or specialist ophthalmological or Ear/Nose/Throat surgical review) that confirmed evidence of active disease. To differentiate between discrete flares, clear improvement in disease activity was required in the form of an improvement in flare-related symptoms together with a reduction in BVAS

score, a reduction in markers of inflammation (CRP, ESR), and a reduction in immunosuppressive therapy.

SLE patients

The SLE cohort was composed of 23 patients attending or referred to the Addenbrooke's Hospital specialist vasculitis unit between July 2004 and May 2008 meeting at least four ACR SLE criteria[32], presenting with active disease (defined below) and in whom immunosuppressive therapy was to be commenced or increased. Following treatment with an immunosuppressant patients were followed up monthly. Disease monitoring was undertaken with serial BILAG disease scoring[33] and full biochemical, hematological and immunological profiling (Supplementary Table 4).

A discrete disease flare required all three of the following prospectively defined criteria:

1. new BILAG score A or B in any system

2. clinical impression of active disease by the reviewing physician

3. the intention to increase in immunosuppressive therapy as a result.

Additional flares were defined in the absence of BILAG scoring if patients were admitted directly to hospital as emergency cases for increased immunosuppressive therapy. To differentiate between disease flares clear improvement in disease activity was required in the form of diminished flare-related symptoms together with a reduction in both BILAG score and immunosuppressive therapy.

IBD patients

Patients with active CD and UC were recruited from a specialist IBD clinic at Addenbrooke's Hospital, prior to commencing treatment. Diagnosis was made using standard endoscopic, histologic, and radiological criteria[34]. Patients who had already received immunomodulators or corticosteroids were excluded. Enrolled patients were managed conventionally using a step-up strategy[3].

Assessment of disease activity was in accordance with national and international guidelines and included consideration of symptoms, clinical signs, and objective measures, including blood tests (C-reactive protein [CRP], erythrocyte sedimentation rate [ESR], hemoglobin concentration, and serum albumin), stool markers (calprotectin), and mucosal assessment (by sigmoidoscopy or colonoscopy) where appropriate. Validated scoring tools were used as another means of assessing disease activity (Harvey-Bradshaw severity index[35] or simple clinical colitis activity index[36] for CD and UC, respectively), although these were not used to guide treatment decisions. All clinicians were blinded to the microarray results.

For each disease, all patients were not included in all analyses as, for example, comparison of modular network analysis in related cell types required that samples passing QC filtering were available for all cell types for all patients. Our previous publications have shown that the sample sizes used here are adequate to detect reproducible signatures correlating with clinical traits.

Follow up Analysis

Comparisons of outcome and associated clinical variables between subgroups were analyzed using the Kaplan-Meier log-rank test and non-parametric Mann Whitney U test or the Chi-square test as appropriate. Correction for multiple testing was applied using the Bonferroni method or false discovery rate (FDR, Benjamini and Hochberg method) where appropriate as indicated.

3. Cell separation and RNA extraction

Venepuncture was performed at a similar time of day in all cases to minimize gene expression differences arising from circadian variation[37]. Peripheral blood mononuclear cells (PBMC), CD4 and CD8 T cells were isolated from 110ml of whole blood by centrifugation over ficoll and positive selection using magnetic beads as previously described[20]. The purity of separated cell subsets was determined by flow cytometry and included as a covariate in downstream correlation and network analyses (e.g. Figure. 1 A, I). Total RNA was extracted from each cell population using an RNeasy mini kit (Qiagen) with quality assessed using an Agilent BioAnalyser 2100 and RNA quantification performed using a NanoDrop ND-1000 spectrophotometer.

Microarray gene expression profiling

HsMediante25k custom spotted microarray—Total RNA (250 ng) was converted into double-stranded cDNA and labelled with Cy3- or Cy5-dCTP as previously described[20]. Appropriate Cy3- and Cy5-labelled samples were pooled and hybridized to custom spotted oligonucleotide microarrays (HsMediante25k) comprised of probes representing 25,342 genes and control features[38]. All samples were hybridized in duplicate, using a dye-swap strategy, against a common reference RNA derived from pooled PBMC samples. Following hybridization, arrays were washed and scanned on an Agilent G2565B scanner.

Affymetrix Human Gene 1.0 ST microarray—Aliquots of total RNA (200ng) were labeled using Ambion WT sense Target labeling kit and hybridized to Human Gene 1.0 or 1.1 ST Arrays (Affymetrix) as described. After washing, arrays were scanned using a GS 3000 or Gene Titan scanner (Affymetrix) as appropriate.

Published datasets

Published datasets were accessed through either NCBI-GEO or ArrayExpress, imported into R using the Bioconductor package GEOquery and analyzed as described. Search criteria incorporated the name of individual diseases and were filtered to human datasets but not by platform used. Studies were only included if they met the following criteria:

1. Similar QC filters as applied to the data produced in-house were satisfied (described below).

2. Samples were taken at an analogous time-point to those from which the costimulation and exhaustion signatures in autoimmunity were identified. i.e. samples taken during active disease without concurrent immunosuppressive therapy.

3. Clinical outcome data was available.

It was not feasible to build a unified predictive model across all available datasets as they originated from different groups and were performed on mutually incompatible microarray platforms.

For the HCV data used in Fig. 4C a marked response was defined as an HCV titer decrease > 3.5 log10iu/ml and a poor response as an HCV titer decrease <1.5 log10iu/ml by day 28 after commencing combined therapy with ribavirin and pegylated interferon-alpha. For the Malaria vaccine trial used in Fig. 4D 'protection' was defined as delayed or complete protection from subsequent confirmed *P.Falciparum* infection following standardised exposure (x5 bites) compared to infectivity control subjects. For the influenza data used in Fig. 4E protection was defined as >/= 1 high response to at least 1 (of 3) included strains. A high response was defined as >/= 4-fold increase in HAI titre at d28 and a titre >/= 1:40 as per US FDA guidelines.

All gene expression data used has been deposited in publicly available repositories (NCBI-GEO and ArrayExpress): AAV, SLE (E-MTAB-2452, E-MTAB-157, E-MTAB-145) IBD (E-MTAB-331), LCMV (GSE9650), HCV (GSE7123), malaria vaccination (GSE18323), influenza vaccination (GSE29619), yellow fever vaccination (GSE13486), dengue fever (GSE25001), IPF (GSE28221), T1D (E-TABM-666), NOD (GSE21897), RA (GSE15258, GSE33377), *in vitro* CD8 stimulation (XXXX).

Data analysis

Preprocessing and quality control (QC)—For Mediante hs25k arrays, raw image data were extracted using Koadarray v2.4 software (Koada Technology) and probes with a confidence score >0.3 in at least one channel were flagged as present. Extracted data were imported into R where log transformation and background subtraction were performed followed by within array print-tip Loess normalization and between-array quantile and scale normalization using the Limma package[39] in Bioconductor[40]. Further analysis was then performed in R and only data demonstrating a strong negative correlation (r^2>0.9) between dye swap replicates were used in downstream analyses.

Affymetrix raw data (.CEL) files were imported into R and subjected to variance stabilization normalization using the VSN package in BioConductor[41]. Quality control was performed using the Bioconductor package arrayQualityMetrics[42] with outlying samples removed from downstream analyses. Correction for batch variation was performed using the Bioconductor package ComBat[43] and batch structure was included as a covariate in downstream correlation analyses.

Clustering

Hierarchical clustering was performed using a Pearson correlation distance metric and average linkage analysis, performed either in Cluster with visualization in Treeview[44], using Genepattern[45] or directly in R using hclust in the stats package.

Differential expression

Differentially-expressed genes were identified using linear modeling and an empirical Bayes method[39] using a false discovery rate threshold of 0.05 as indicated to determine significance.

Weighted Gene Coexpression Network Analysis (WGCNA)

Highly correlated genes in immune cell subsets were identified and summarized with a modular eigengene profile using the Weighted Gene Coexpression Network Analysis (WGCNA) Bioconductor package in R[46]. Normalized, log transformed expression data was variance filtered using the inflexion point of a ranked list of median absolute deviation values for all probes. A soft thresholding power was chosen based on the criterion of approximate scale-free topology[47]. Gene networks were constructed and modules identified from the resulting topological overlap matrix with a dissimilarity correlation threshold of 0.01 used to merge module boundaries and a specified minimum module size of n=30. Modules were summarized as a network of modular eigengenes, which were then correlated with a matrix of clinical variables and the resulting correlation matrix visualized as a heatmap (Extended Data Figure 1). As each module by definition is comprised of highly correlated genes, their combined expression may be usefully summarized by eigengene profiles[48], effectively the first principal component of a given module (e.g. Figure 1B, F). A small number of eigengene profiles may therefore effectively 'summarize' the principle patterns within the cellular transcriptome with minimal loss of information. This dimensionality-reduction approach also facilitates correlation of ME with clinical traits (e.g. Figure 1A, I). Significance of correlation between a given clinical trait and a modular eigengene was assessed using linear regression with Bonferroni adjustment to correct for multiple testing (Extended Data Figure 1). Independent association of a given module eigengene or gene expression profile (e.g. KAT2B) with clinical outcome was assessed using a multiple linear regression model. Significance of each term in the linear model was plotted against its regression coefficient, as a measure of the strength of association (the regression coefficient reflecting the change in clinical outcome per unit change in modular/gene expression), for example Extended Data Fig.3B-E.

Overlap of signatures with modules derived from network analysis is shown to the right of selected module heatmaps (Figure 1A, Extended Data Figures 2A, E, F) by the following formula to allow correction for variable module size: (signature genes overlapping with module genes, n)/(genes in module, n) x100. The overlap of randomly selected signatures of equivalent size was used as a control and is shown adjacent to the above plots.

HOPACH analysis

For validation purposes, highly-correlated genes were independently partitioned into discrete modules using a second algorithm, Hierarchical Ordered Partitioning And Collapsing Hybrid (HOPACH[49]) in R. This approach differs from WGCNA in that it does not rely on a user-specified correlation threshold to define module boundaries but rather aims to maximize homogeneity of modules. Normalized, log transformed data were clustered using a hierarchical algorithm with modular boundaries defined by the median split silhouette (MSS), a measure of how well-matched a gene is to the other genes within its

current cluster versus how well-matched it would be if it were moved to another cluster. On partitioning the dataset into clusters, each cluster is reiteratively subdivided until the MSS is maximized, thereby producing an optimal segregation into maximally discrete modules.

Knowledge-based network generation and pathway analysis

The biological relevance of gene groups comprising modules identified by co-expression analysis were further investigated using the Ingenuity Pathways Analysis platform[50]. Six modules from the CD4 T cell WGCNA analysis showed significant correlation with clinical outcome in AAV after correction for multiple testing (Bonferroni method, Supplementary Table 3). We applied network and pathway enrichment analysis to genes comprising these modules to determine whether they may have any biological relevance. Briefly, for network analysis genes from a specified target set of interest are progressively linked together based on a measure of their interconnection, which is derived from described functional interactions. Additional highly interconnected genes that are absent from the target genes (open symbols) may be added to complete a network of arbitrary size (set at n = 35). Networks may be ranked by significance which reflects the probability of randomly generating a network of similar size from genes included in the full knowledge database containing at least as many target genes as in the network in question. For pathways analysis, the overrepresentation of canonical pathways (pre-defined, well-characterized metabolic and signaling pathways curated from extensive literature reviews) amongst a specified set of target genes is assessed, with significance determined by computing a Fisher's exact test with false discovery rate correction for multiple testing.

Gene Set Enrichment Analysis (GSEA)

GSEA[11] was used to further assess whether specific biological pathways or signatures were significantly enriched between patient subgroups identified by gene modules (as opposed to testing for enrichment of pathways within modules themselves as outlined in the previous section). GSEA determines whether an *a priori* defined 'set' of genes (such as a signature) show statistically significant cumulative changes in gene expression between phenotypic subgroups (such as patients with relapsing or quiescent disease). In brief, all genes are ranked based on their differential expression between two groups then an enrichment score (ES) is calculated for a given gene set based on how often its members appear at the top or bottom of the ranked differential list. 1000 random permutations of the phenotypic subgroups were used to establish a null distribution of ES against which a normalized enrichment score (NES) and FDR-corrected q values were calculated. GSEA was run with a focused subgroup of gene signatures (as in Figure 2B and Figure 3K)[11] selected to test the null hypothesis that different CD8 T cell phenotypes were not significantly enriched in patient subgroups identified by modular analysis.

Selection of optimal PBMC-level biomarkers

Optimal surrogate markers facilitating identification of the CD4 T cell co-stimulation/CD8 exhaustion signatures in PBMC-level data were determined using a randomforests classification algorithm[51] (Figure 4A). Although signatures apparent in purified T cell transcriptome data correlate with clinical outcome, they cannot be similarly detected in data

derived from PBMC due to the confounding influence of expression from other cell types nor can the same genes be used to predict outcome in PBMC[2,20]. However, as CD4 T cell co-stimulation and CD8 T cell exhaustion signatures themselves showed close correlation we hypothesized that this would amplify the signal detectable in PBMC and that detection of the combined CD4/CD8 T cell response may be feasible. The availability of both separated T cell and PBMC data from the same patients at the same time facilitate a supervised search for surrogate markers of the co-stimulation/exhaustion signatures in PBMC. Expression data derived from both CD4 T cells and PBMC were available for a cohort of n=37 patients (AAV and SLE) following QC and hybridization to the HsMediante25k custom microarray platform and constituted a training cohort. Normalized, log- transformed expression data was analyzed using the MLInterfaces Bioconductor package in R[52]. Using PBMC-level expression data samples were classified into subgroups showing either high or low expression of the costimluation/exhaustion signature (as illustrated in Extended Data Figure 5H, I) and probes were subsequently ranked using the variable importance metric based on their ability to predict allocation to either group. The variable importance for a given gene reflects the change in accuracy of classification (% increase in MSE or increase in node purity) when that variable is randomly permuted. For a poorly predictive gene, random permutation of its values will minimally influence classification accuracy. Conversely, the most robust predictors will have a comparatively large effect on classification accuracy when randomly permuted. PBMC samples from a subset of n=37 cases derived from the training cohort were labeled and hybridized on an alternative microarray platform (Affymetrix Gene ST1.0) as a technical validation set (Figure 4B, left panel). PBMC samples from an independent n=47 cases not included in the training cohort were labeled and hybridized to the Affymetrix Gene ST1.0 platform as an independent test set (Figure 4B, right panel). For both technical validation and independent test sets expression of the optimal biomarker identified in Figure 4A (*KAT2B*) was used to bisect the cohort relative to the median expression and clinical outcome was compared in *KAT2Bhi* and *KAT2Blo* patients.

Linear Models

Linear modeling was performed in R using the stats package. This took the form of

$$\text{fit} <- \text{lm}\,(\text{y} \sim \text{x1+x2+x3}, \text{data=mydata})$$

where y (the response variable) was selected as normalized flare rate (flares/days follow-up) and x_1-x_n (the test variables) were selected to include measures of disease activity (both clinical scores and laboratory markers of inflammation), quantification of circulating leucocyte subsets (lymphocytes, neutrophils) and concurrent measurements of autoantibody titer where relevant. Test variables also included a biomarker profile (e.g. exhaustion signature or KAT2B expression). The significance and magnitude (regression coefficient, reflecting change in response variable (flares/days follow-up) per unit change in each test variable included) were extracted and plotted against each other (for example, Extended Data Fig. 3B-E). Not all clinical or laboratory measures were relevant comparisons in each case and therefore were not all included in every model generated.

T cell culture

Primary human CD8 T cells were separated from leucocyte cones obtained from NHS Blood and Transplant (Addenbrooke's Hospital, Cambridge, UK) by centrifugation over ficoll and positive selection using magnetic beads as previously described[20]. The purity of separated cell subsets was determined by three-color flow cytometry. Purified T cells were labeled with 10μM CFSE (Invitrogen) and resuspended in complete RPMI 1640 (Sigma Aldrich) in the presence of 10% FCS. Purified CD8+ T cells (>95%) were then stimulated in sterile, 96-well U-bottomed culture plates (Greiner) using an 'artificial APC' consisting of MACS iBead particles (1:2 bead:cell ratio, Miltenyi) or DynaBead particles (Invitrogen) conjugated to either CD3/CD28 or CD2/CD3/CD28 as indicated in the presence of IL2 (10ng/ml, Gibco life technologies) for 6 days. The magnetic iBead construct was removed after 36h in some instances as indicated. In some experiments, additional costimulation was provided by the addition of either IFNα (10ng/ml, Abcam) or by additional conjugation of recombinant Human PD-L1 Fc Chimera (life technologies, 1μg/ml) or anti-CD40 antibody (50ng/ml, Abcam) as indicated. The nature of costimulatory signals tested was based upon the results of the network analysis of CD4 T cell modules described above (Supplementary Table 2).

For restimulation experiments cells were harvested on day 6 post-stimulation and sorted into IL7R^{hi} and IL7R^{lo} populations (Extended Data Figure 6D) using a FACSAriaIII cell sorter (BD Biosciences) with live/dead discrimination performed using an AquaFluorescent amine-reactive dye (Invitrogen). Cell numbers were normalized and were resuspended in complete RPMI 1640 (2×10^4/ml, Sigma-Aldrich) and 'rested' in a sterile, U-bottomed culture plate (Greiner) for 6 days (37C, 5% CO_2) before being restimulated (anti-CD2/3/28 1:2 bead:cell ratio, Miltenyi MACSiBead) for a further 6 days in the presence of IL2 (10ng/ml, Gibco life technologies).

Note that, as described in Extended Data Fig. 6G, human memory CD8 T cell subsets do not equivalently respond to the stimulation conditions described above. As primary whole human CD8 T cells are composed of highly variable proportions of memory subsets and whole CD8 T cells were stimulated it was necessary to perform paired tests of significance when comparing resulting T cell subsets and transcriptional profiles.

Flow cytometry

Immunophenotyping was performed using an LSR Fortessa analyzer (BD Biosciences), and data was analyzed using FlowJo software (Tree Star). Reactions were standardized with multicolor calibration particles (BD Biosciences) with saturating concentrations of the following antibodies: AquaFluorescent Live/Dead (Invitrogen), IL7Rα AF647 (BD biosciences, clone HIL-7R-M21), PDCD1 APC (eBioscience, clone MIH4). For intracellular staining, cells were fixed and permeabilized using a transcription factor staining buffer set (eBioscience) and before staining with saturating concentrations of antibody against BCL2 (BD Biosciences, clone 100).

Extended Data
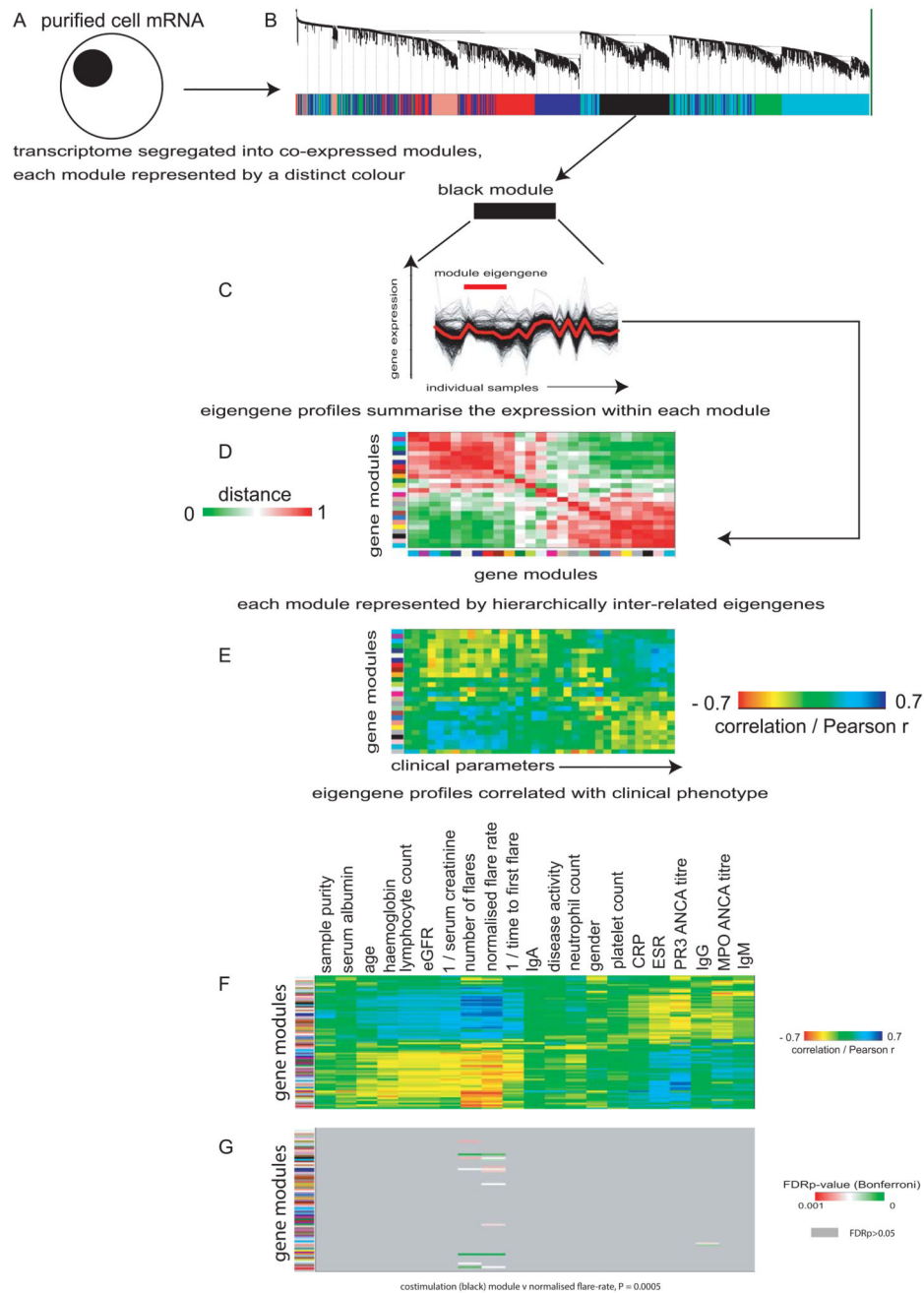
**Extended Data Figure 1. Overview of weighted gene coexpression analysis**
(A) mRNA derived from purified leucocyte subsets sampled during active, untreated autoimmune disease is labeled and hybridized to a microarray platform (both HsMediante 25k and Affymetrix Gene ST1.0 used here). Genes are then combined into modules (B, colored blocks) based on the similarity of their expression profile in all samples. (C) Detail for the 'black' module. Each horizontal black line represents expression of a single gene within the given module. y-axis = gene expression, x-axis = patient samples, red-bar =

eigengene profile which effectively summarizes the expression of all genes comprising the black module. (D) Each modular profile is related to all others in a hierarchy that can itself be visualized by plotting correlation of all module eigengenes, such as in the heatmap shown here. Colored blocks represent individual modules, defined as in (A). Modules are aligned in identical order on x and y-axes with heatmap color representing the correlation between each. Note that the diagonal (top left to bottom right) therefore represents correlation of each eigengene profile with itself, and is always 1. Distance metric = Euclidean distance. (E) As each module is summarized by a representative eigengene profile, each may then be correlated against a range of clinical variables allowing visualization of how the transcriptome relates to clinical variables, again in the form of a correlation heatmap. Correlation = Pearson, r. (F) Heatmap showing gene expression modules (y-axis) correlated against clinical variables (x-axis) for the CD4 transcriptome in AAV, correlation = Pearson, r. (G) Heatmap illustrating significance of correlations identified in (F). P-value threshold at Bonferroni-corrected P<0.05. Color-bar indicates actual P-value of correlations deemed significant, grey shading = corrected P >0.05. Significance for costimulation (black) module from Figure 1 is also shown (P = 0.0005).
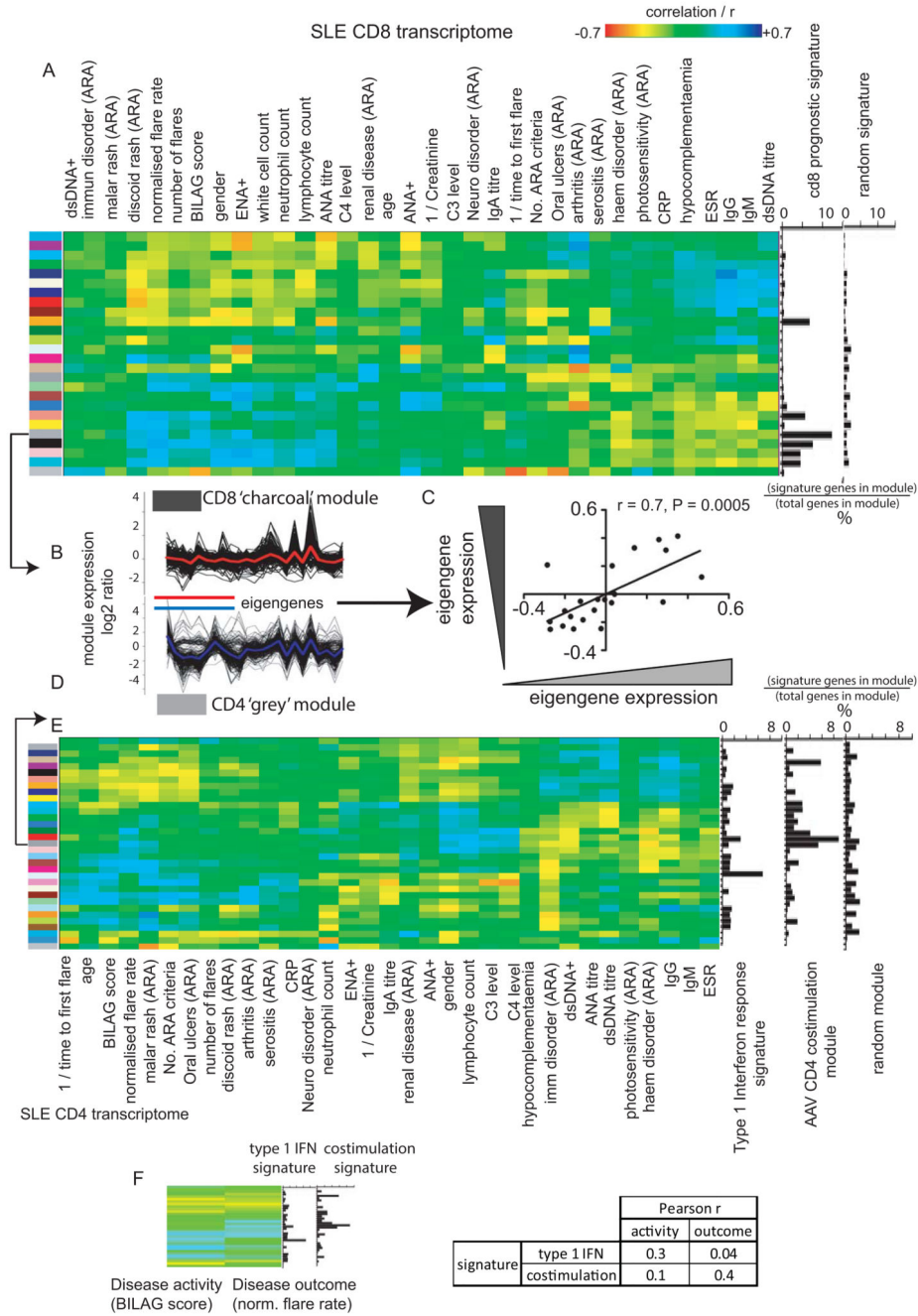
Extended Data Figure 2. Weighted gene coexpression network analysis of the T cell transcriptome and its correlation with clinical phenotype in SLE

(A, E) Heatmaps illustrating the correlation of coexpression modules (colored blocks, y-axis) derived from the CD8 (A) and CD4 (E) transcriptomes of 23 SLE patients with clinical traits (x-axis). Overlap of the previously described prognostic signature with coexpression modules, along with the distribution of a random signature of equivalent size, shown to the right of (A) (overlap = signature genes / module genes %). Overlap of the CD4 T cell costimulation 'black' module (defined in Fig.1) shown to the right of (E) along with a randomly derived module and a type 1 interferon response signature previously shown to

associate with active SLE[4]. Overlap shown as % representation of the signature within each module. (B, D) Linear plots illustrating the 'charcoal' (B) and 'grey' (D) modules in detail. y-axis = gene expression, x-axis = individual patients, colored lines (red, blue) = module eigengenes. (C) Correlation of SLE CD4 T cell costimulation module eigengene (x-axis, blue) against SLE CD8 T cell prognostic signature (y, red). Pearson correlation, r, with P = 2-tailed significance. (F) Expanded detail from (E) illustrating that modules corresponding to type 1 IFN response and costimulation signatures correlate with disease activity and outcome respectively but not vice versa.
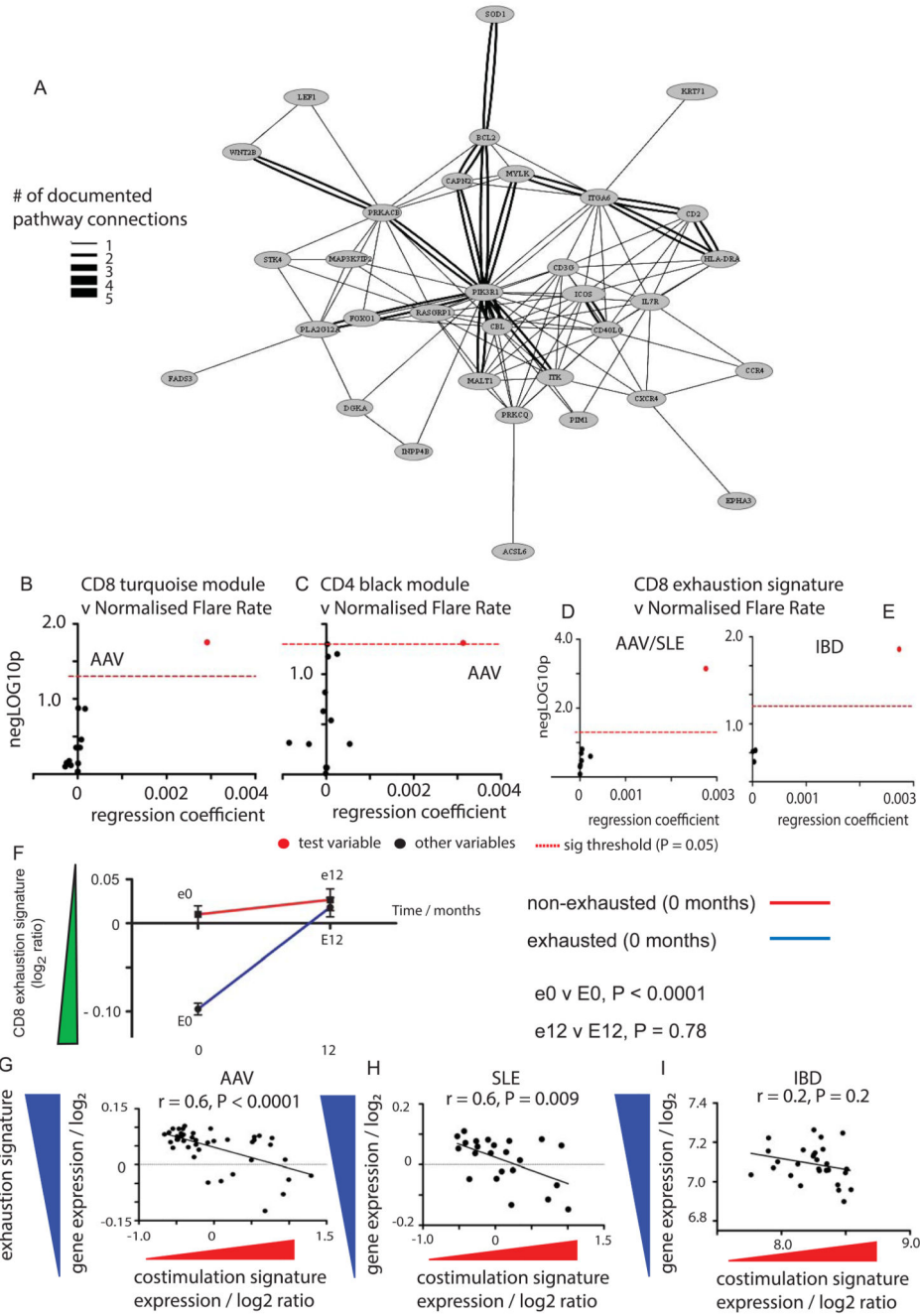
Extended Data Figure 3. Identification and validation of genes involved in CD4 costimulation that correlate with clinical outcome, and how that relationship changes after treatment

(A) A knowledge-based network analysis of 336 probes comprising the 'black' expression module (Fig.1E) identifies a network of costimulation signaling (Supplementary Table 3). Individual genes are shown in circles with the 'strength' of their connections indicated by the weight of the black bar linking them. Pathways of TCR signaling, ICOS-ICOSL signaling and CD28 signaling all significantly enriched in this module (FDR $p < 0.05$). (B-E) Scatterplots showing the outcome of multiple linear regression models testing the association of 4 signatures (red symbols) as indicated, directly compared to clinical markers

of disease activity (black symbols). x-axis = magnitude of association (regression coefficient, change in normalized flare rate (flares/days follow-up) per unit change in each variable tested). y-axis = significance of association in multiple regression model, P. significance threshold (dashed red line, P = 0.05). (B) CD8 turquoise module eigengene in AAV, (C) CD4 costimulation (black) module eigengene in AAV, (D, E) CD8 exhaustion signature (Supplementary Table 6) in AAV/SLE (D) and IBD (E). Clinical variables incorporated vary due to differing relevance in each case but include some of: disease activity score (BVAS/BILAG/CDAI/Harvey-Bradshaw score), CRP, autoantibody titer (PR3/MPO, dsDNA), Lymphocyte count, neutrophil count, platelet count, IgG, IgA, IgM, ESR, age. (F) Line plot showing mean expression of a CD8 T cell exhaustion signature in 38 AAV patients measured at presentation during active, untreated disease (t0) and again 12 months later when disease activity was quiescent and patients were on maintenance immunosuppressive therapy (t12). Patients are grouped into those falling above (red) and below (blue) median expression of the exhaustion signature eigengene at entry. P = Mann-Whitney test comparing t12 and t0 values. The difference between the groups that is easily apparent at enrolment with active, untreated disease (t0) is no longer apparent when disease is treated and quiescent twelve months later (t12). (G-I) Scatterplots showing inverse correlation between individual eigenvalues of the CD4 costimulation signature (x-axis, red) and the CD8 exhaustion signature (y-axis, blue) defined as in Fig. 2, for AAV (G), SLE (H) and IBD (I) cohorts. Correlation = Pearson, r^2, 2-tailed significance.
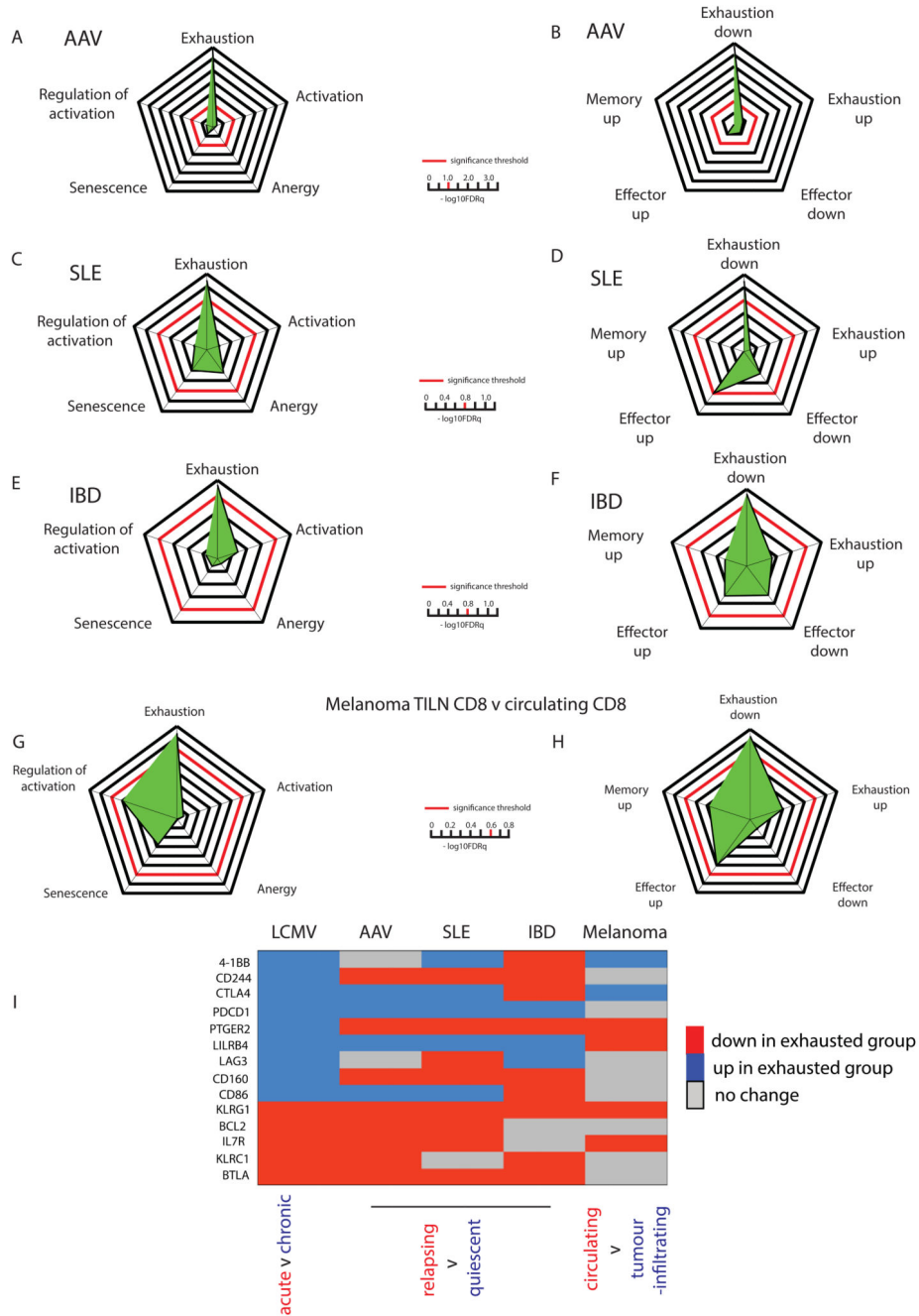
Extended Data Figure 4. Windrose plots showing relative GSEA enrichment of immune signatures in autoimmune disease and melanoma

Windrose plots showing relative enrichment (GSEA FDR q value) of distinct immune signatures between patient subgroups (as defined as in Fig2). (A, B) AAV, (C, D) SLE and (E, F) IBD. (A, C, E) enrichment of immune signatures from selected CD8 T cell phenotypes and (B, D, F) enrichment of signatures specifically up/down regulated by CD8 T cell subsets derived from the LCMV model of T cell exhaustion (acute LCMV-Armstrong v chronic LCMV-Cl13[8]). Detailed information on genes included in each signature is provided in Supplementary Table 6. (G, H) Windrose plots showing relative enrichment

(GSEA FDR q value) of distinct immune signatures between CD8 T cells from melanoma patients, comparing CD8 from tumor-infiltrated lymph node with circulating CD8 T cells[16]. (G) Enrichment of immune signatures from selected CD8 T cell phenotypes and (H) enrichment of signatures specifically up/down regulated by CD8 T cell subsets derived from the LCMV model of T cell exhaustion (acute LCMV-Armstrong v chronic LCMV-Cl13[8]). Specific enrichment is seen for genes downregulated by exhausted cells but not for all genes upregulated by exhausted cells. (C) Heatmap showing differential expression of selected canonical coinhibitory receptors (as for Fig2C[12]) in the LCMV exhaustion model, between prognostic subgroups identified in D, G, J (reproduced from Fig.2C) and also between exhausted CD8 from melanoma-infiltrated lymph node compared to circulating tumor-specific CD8 T cells[16]. Blue = up in exhausted, Red = up in non-exhausted, grey = no significant change (FDR p<0.05).
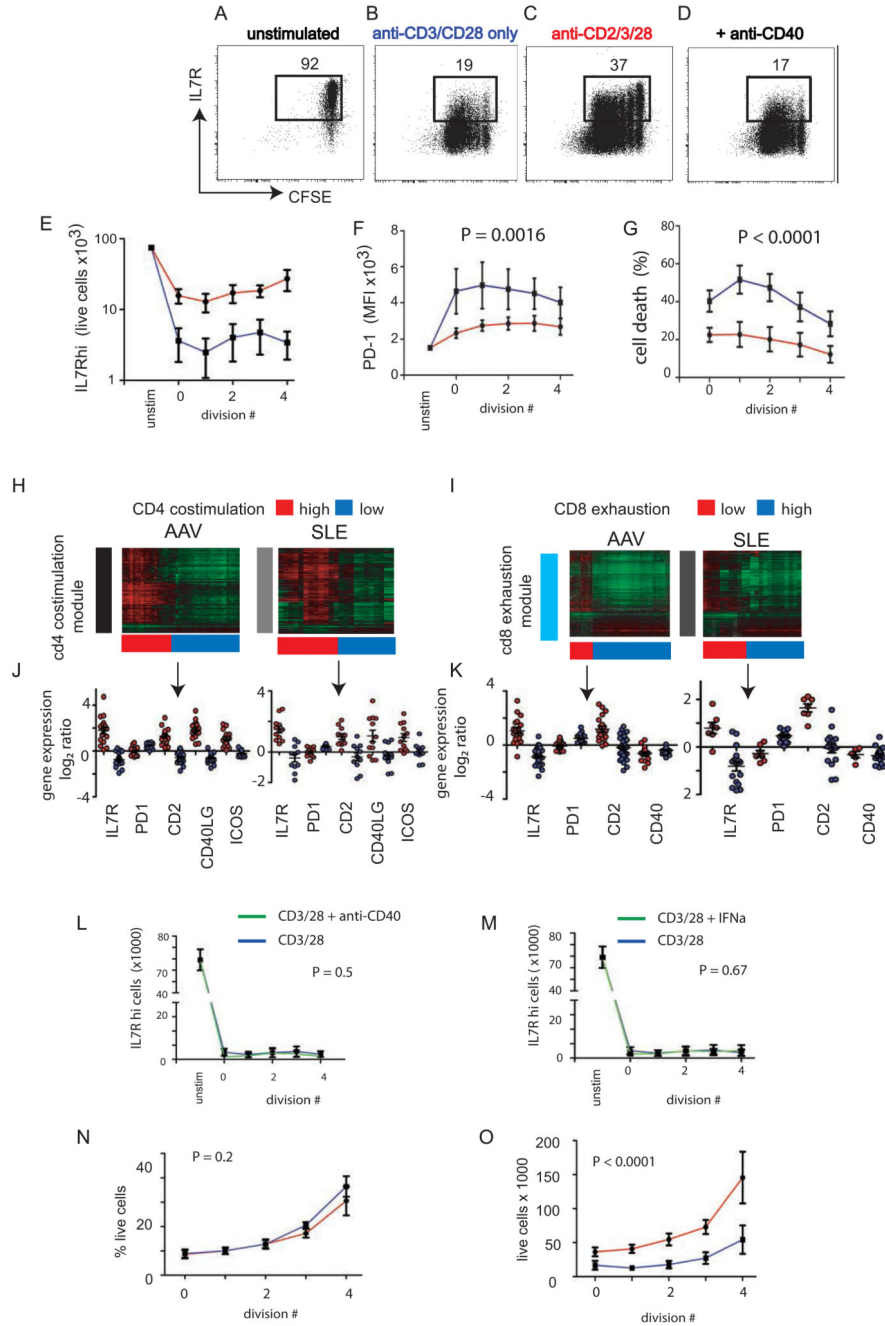
Extended Data Figure 5. T cell costimulation with CD2, but not type 1 interferon or anti-CD40, prevents development of an exhausted IL7R^{lo}PD1^{hi} phenotype during prolonged anti-CD3/28 T cell stimulation

(A-D) Representative scatterplots showing IL7R expression (y-axis) by cell division (CFSE dilution, x-axis) in (A) unstimulated cells and following each of three different costimulation cultures: (B) anti-CD3/CD28 alone, (C) anti-CD2/3/28 and (D) anti-CD40/3/28. IL7R^{hi} expressing subset indicated in black gate with % live cells shown. (E- G) Line and scatterplots showing absolute number of IL7R^{hi} cells (E), PD-1 expression (F) and cell death (G, death = AquaFluorescent dye^{+}) during CD8 T cell differentiation (x-axis, number of

divisions undergone by day 6 of culture measured by CFSE dilution) following anti- CD3/28 (blue) or anti-CD2/3/28 stimulation (red). P= paired t-test, n = 5 paired samples. (L, M) Line and scatterplots showing absolute number of IL7R^{hi} cells (y-axis) by number of divisions undergone at day 6 (x-axis) following polyclonal stimulation with anti-CD3/28 (blue) or anti-CD3/28 plus anti-CD40 (L, green) or interferon alpha (IFNα, green, M) costimulation. (N) Line and scatterplot showing extent of proliferation occurring (% of live cells on day 6 having undergone each of 0-4 divisions) following polyclonal stimulation of primary human CD8 T cells with CD3/28 alone (blue) or with additional anti-CD2 costimulation (red), confirming no difference in the extent of live cell proliferation between groups. (O) Absolute live (AquaFluorescent Dye$^-$) cell counts (y- axis) by the number of divisions undertaken (x-axis) by day 6 following polyclonal stimulation of primary human CD8 T cells with CD3/28 alone (blue) or with additional anti-CD2 costimulation (red), illustrating increased cell survival with CD2 costimulation despite equivalent proliferation. P values = 2-way ANOVA of 4 paired stimulations. (H, I) Hierarchical clustering of 44 AAV (left panels) and 23 SLE (right panels) patients using 336 genes comprising a CD4 T cell costimulation module (black module, Fig 1) identifies 2 patient subgroups (high costimulation, red, and low costimulation, blue) in CD4 T cell expression data defined by the first major division in the patient dendrogram. (J, K) Scatterplots illustrating selected costimulatory and coinhibitory receptors for the subgroups identified in (H) and (I). Selected receptors were chosen based on their inclusion in networks derived from the costimulation and exhaustion signatures as illustrated in Extended Data Figure 3A.
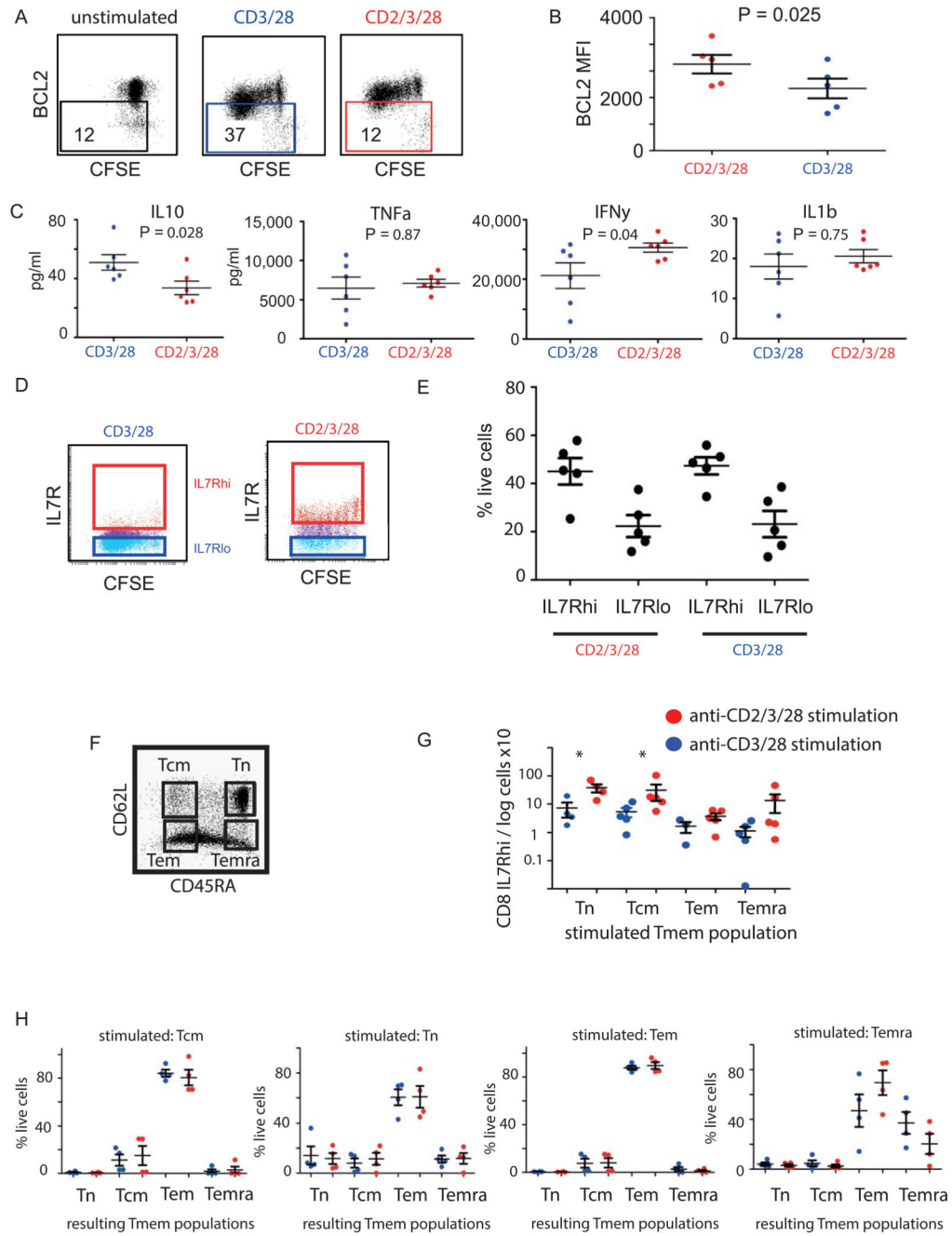
Extended Data Figure 6. CD2-costimulation results in functionally distinct subpopulations showing enhanced survival following *in vitro* restimulation but no preferential expansion of CD8 memory subsets

(A) Representative flow cytometry density plots of CD8 T cells showing BCL2 expression on day 7 after stimulation with anti-CD3/28 (blue) or anti-CD2/3/28 (red). Figures are % of total CD8 T cells. (B) Quantification of BCL2 expression in CD8 T cells stimulated as in (A). P = Mann-Whitney, n = 5 paired biological replicates per group. (C) Scatterplots showing cytokine levels (y-axis, pg/ml) measured in supernatants of CD8 T cells on day 7 after *in vitro* stimulation with either anti-CD3/28 (left column, blue) or CD2/3/28 (right

column, red). Samples represent paired stimulations of primary CD8 T cells from the same individual using either stimulation protocol, n = 6 biological replicates per group. (D) Scatterplots illustrating populations sorted following polyclonal anti-CD3/28 (left panel) and anti-CD2/3/28 (right panel) stimulation of primary CD8 T cells. (E) % live cells (AquaFluorescent dye⁻) remaining 7 days after restimulation of each sorted subpopulation of CD8 cells. Cells were rested for 6 days in complete RPMI1640 medium without IL2 before being restimulated with anti-CD2/3/28 for a further 7 days. P = Mann-Whitney, Error bars = Mean +/− SEM. (F) Representative scatterplot illustrating CD8 T cell memory populations isolated by flow cytometric sorting and stimulated in (G, H). (G) Scatterplot showing absolute number of IL7R^hi cells (y-axis) on day 6 following anti-CD3/28 (blue) or anti-CD2/3/28 (red) stimulation of purified CD8 T cell memory populations (x-axis). * = P<0.05, Mann-Whitney test. n = 5 paired biological replicates per group. (H) Scatterplots showing % CD8 T cell memory subsets (y-axis) resulting from stimulation of purified central memory (Tcm), naïve (Tn), effector memory (Tem) and effector memory-RA (Temra) populations with anti-CD3/28 (blue) or anti-CD2/3/28 (red) for 6 days, n = 4 paired biological replicates per group.
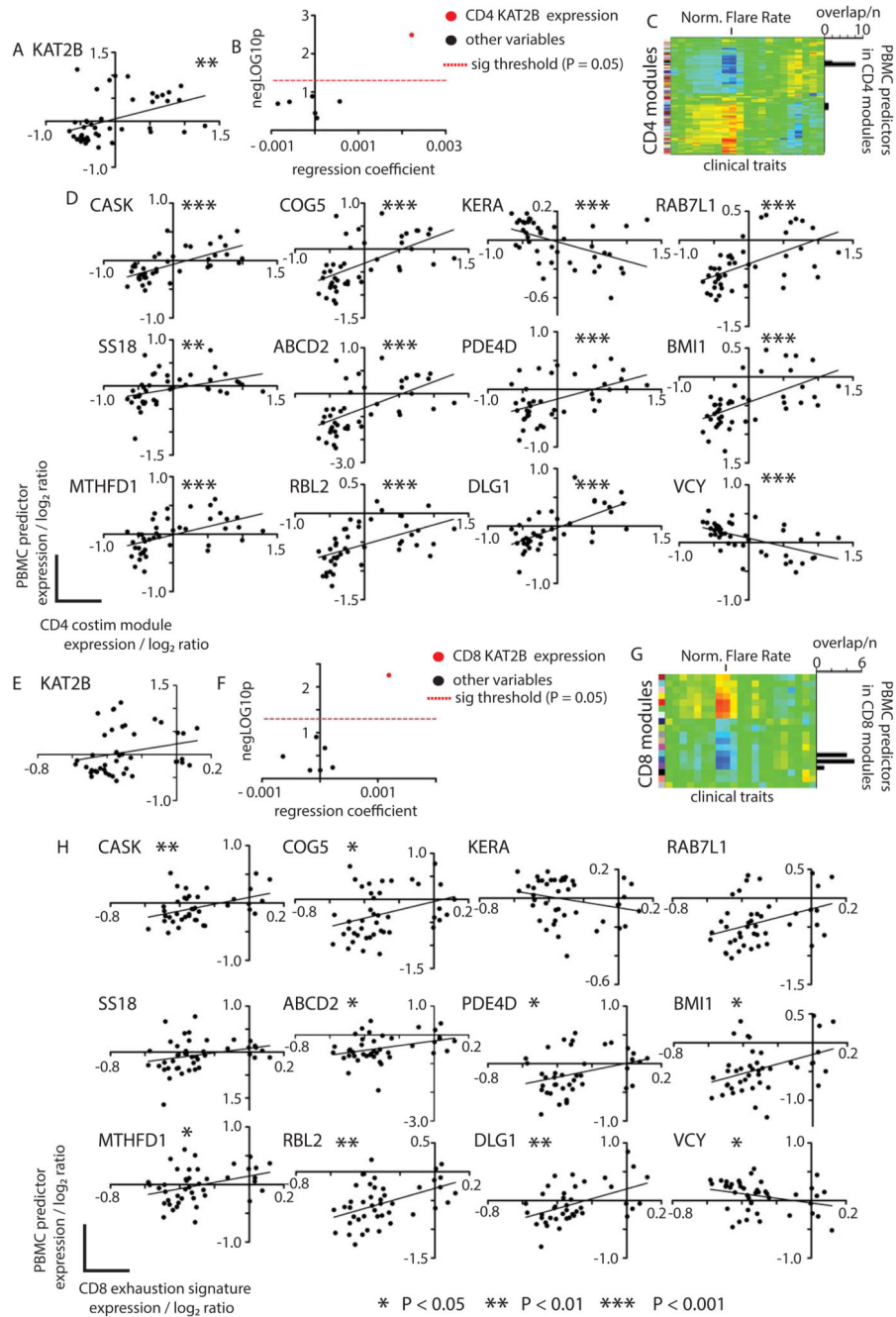
Extended Data Fig. 7. Top PBMC surrogate markers reflect expression of CD4 costimulation/CD8 exhaustion modules within CD4 and CD8 data respectively

Top PBMC-level predictors (n=13) were selected as indicated in Fig4A and data is shown comparing expression of the optimal predictor (KAT2B, A, E) and of each other top predictor gene (D, H) in PBMC data compared to expression of the CD4 costimulation module eigengene in CD4 data (A-D) and the CD8 exhaustion signature eigengene in CD8 data (E-H) for n=44 patients with AAV. Significance of correlation, *P<0.05, **P<0.01, ***P<0.001. (B, F) Scatterplots showing the outcome of multiple linear regression models testing the association of KAT2B expression in CD4 (B) and CD8 (F) data (red symbols)

directly compared to clinical markers of disease activity (black symbols). x-axis = magnitude of association (regression coefficient, change in normalized flare rate (flares/days follow-up) per unit change in each variable tested). y-axis = significance of association in multiple regression model, P. significance threshold (dashed red line, P = 0.05). Clinical variables incorporated = disease activity score (BVAS), CRP, Lymphocyte count, neutrophil count, IgG. (C, G) Heatmaps reproduced from Fig1A and I respectively, showing overlap of top PBMC-level predictors with the modular analysis presented for CD4 (C) and CD8 (G) data in Figure 1. As expected, surrogate markers showed stronger correlation with the CD4 than the CD8 signature as the algorithm was trained to detect the CD4 costimulation module.
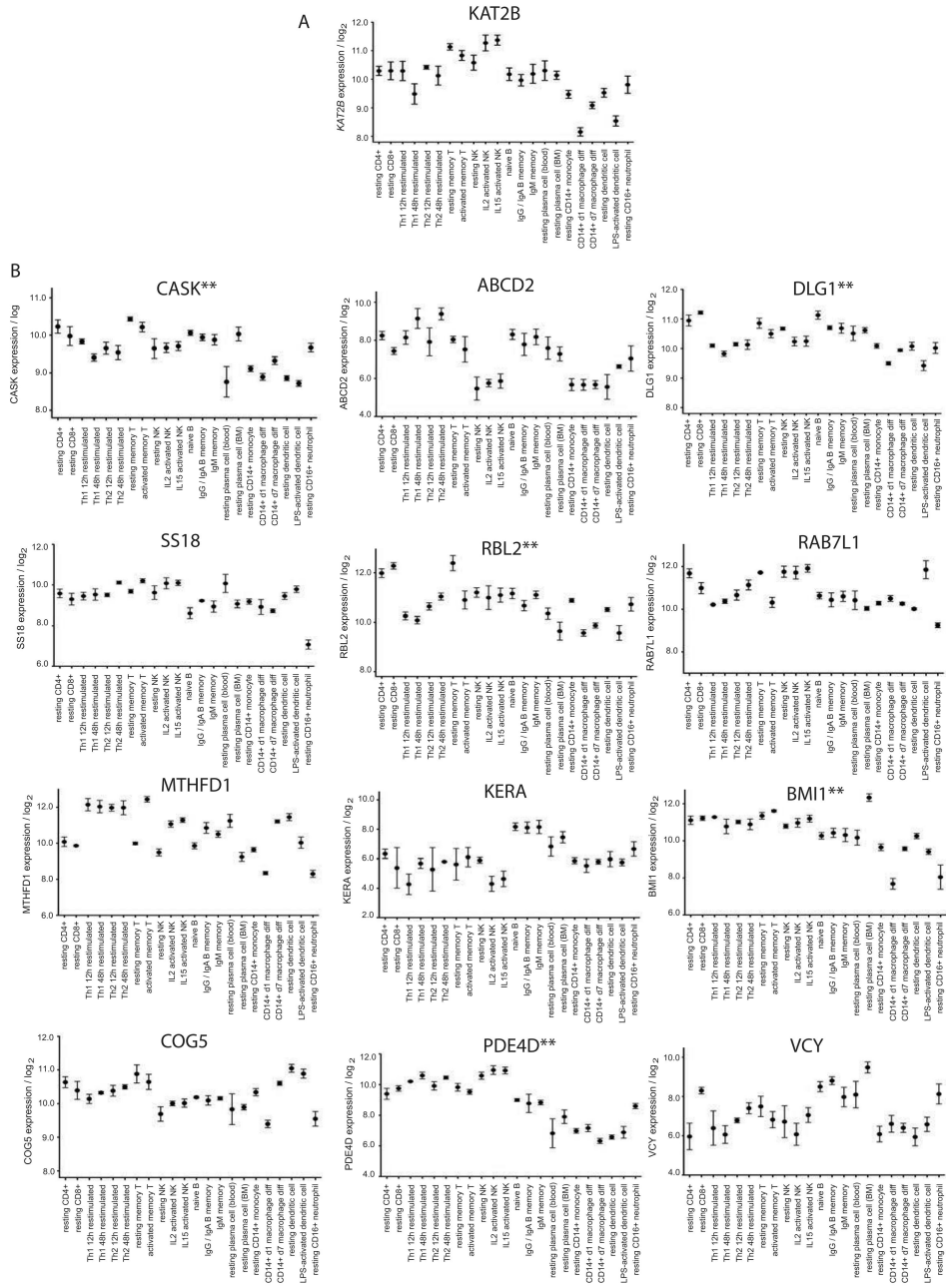
** P < 0.001 correlation with KAT2B expression across all subsets

Extended Data Fig. 8. Immune cell subset expression pattern of top PBMC-level surrogate markers of CD4 costimulation/CD8 exhaustion signatures

Dot plots showing expression (median +/− SEM) of KAT2B (A) and for each of 12 other top PBMC-level surrogate predictors of CD4 costimulation/CD8 exhaustion signatures (from Fig.4A) in a range of 22 immune cell subsets. Genes showing significant correlation of expression with KAT2B across all cell types are indicated (**P<0.001).
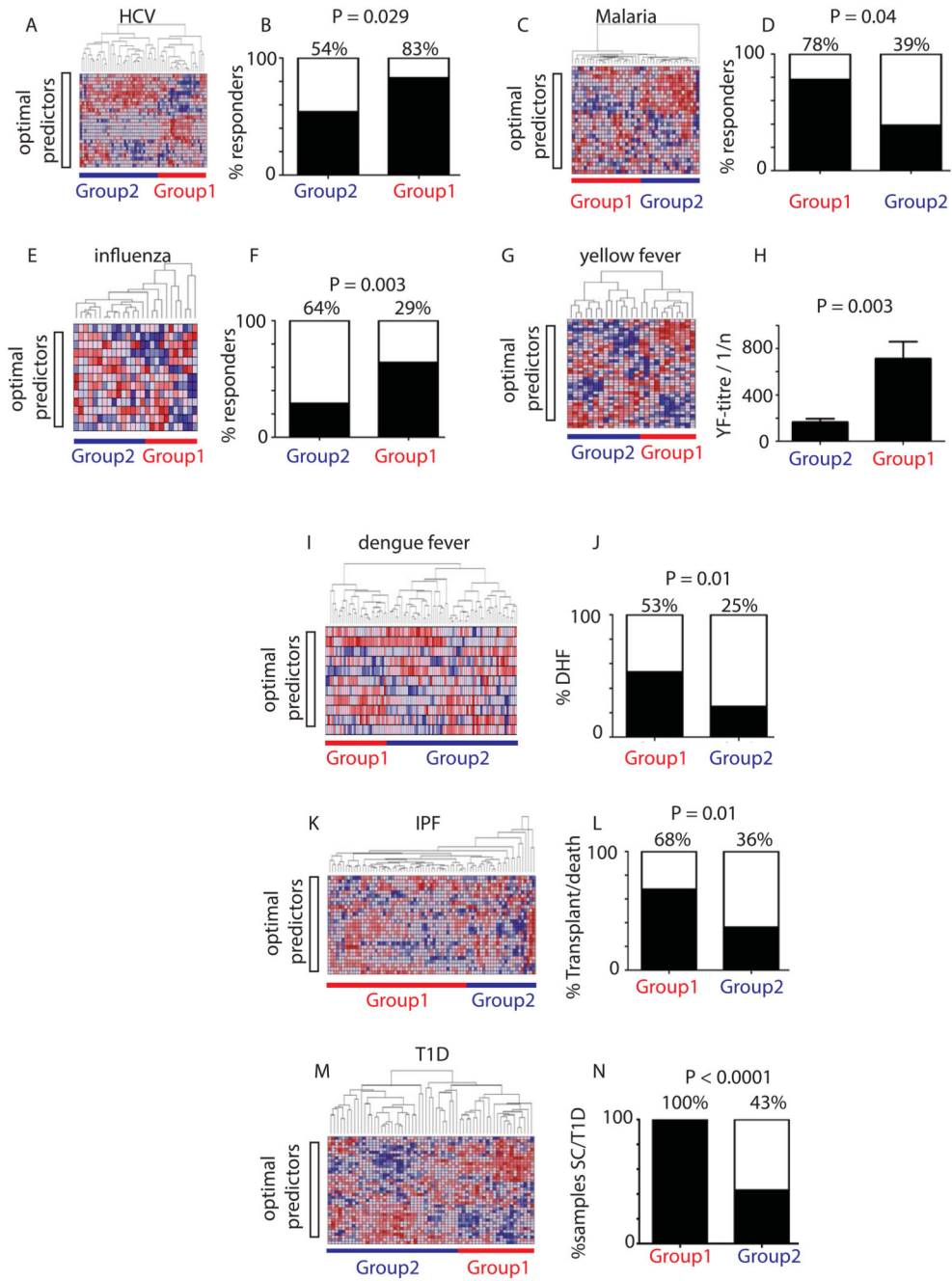
Extended Data Fig. 9. Hierarchical clustering of multiple datasets using 13 top PBMC-level surrogate markers of CD4 costimulation/CD8 exhaustion modules identifies patient subgroups with distinct clinical outcomes

Replication of association between surrogate markers of CD4costimulation/CD8 exhaustion signatures and clinical outcome (as shown in Fig4C-K) but using all top 13 PBMC-level surrogates rather than KAT2B alone. (A, C, E, G, I, K, M) Heatmaps showing hierarchical clustering of gene expression data of 13 top PBMC-level surrogate predictors of CD4 costimulation/CD8 exhaustion signatures (from Fig.4A) in patients with chronic HCV[53] (A), during malaria vaccination (C), influenza vaccination (E), yellow fever vaccination (G),

dengue fever infection (I), idiopathic pulmonary fibrosis (IPF, K) and pre-T1D (M). Subgroups were defined using a major division of the cluster dendrogram and Group1 allocated based on KAT2B expression (highest in Group 1). Clinical outcome associated with each subgroup identified is shown in B (HCV, % responders to IFNα/ribavirin therapy), D (% showing protection v no protection from malaria vaccine), F (% response to influenza vaccination), H (yellow fever antibody-titer post-vaccination), J (% progression to dengue hemhorrhagic fever, DHF), L (% patients progressing to need for transplantation or death) and N (% samples from patients with prior or subsequent progression to islet-cell antibody seroconversion or to a diagnosis of T1D).
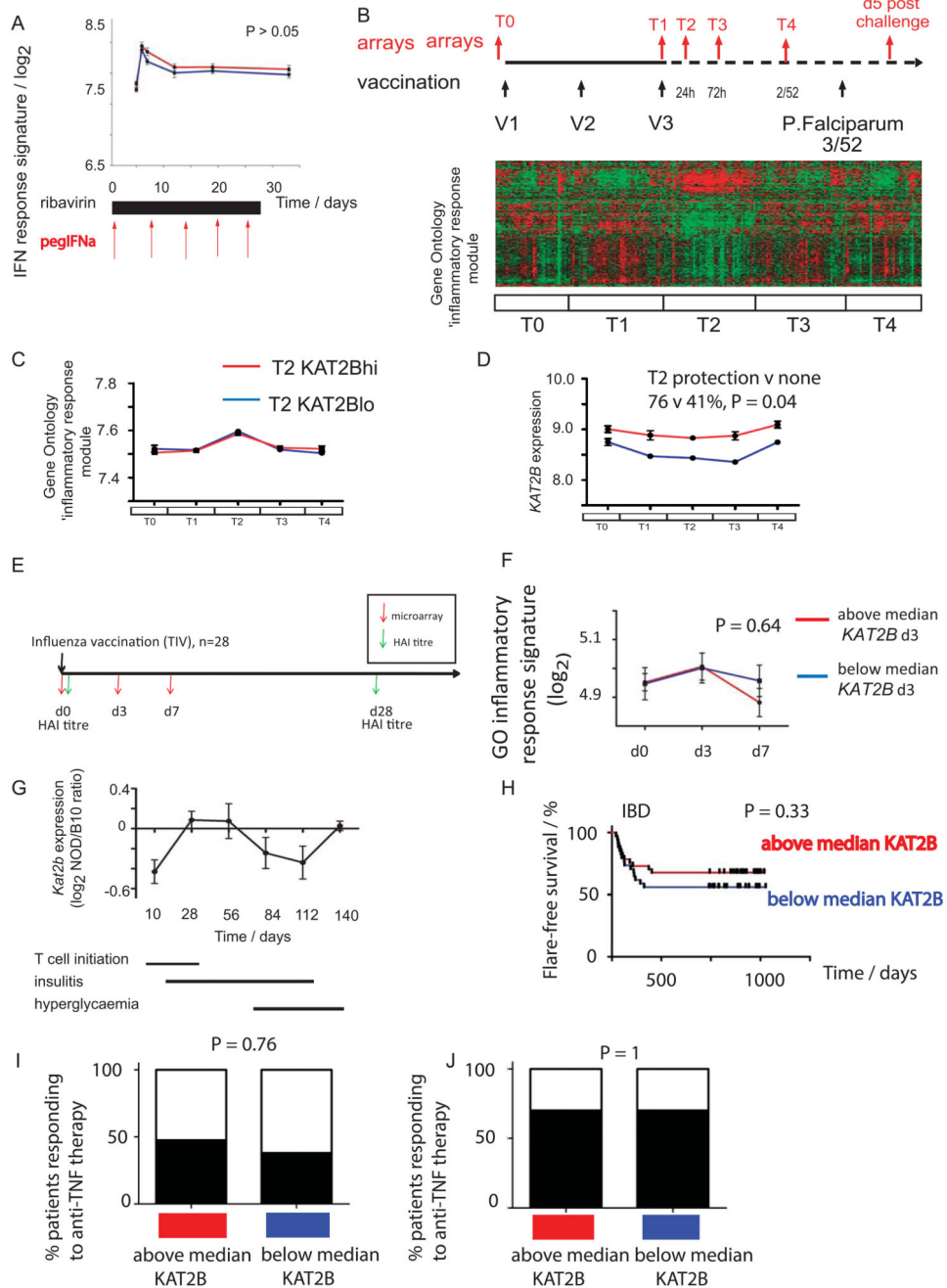
Extended Data Fig. 10. Kinetics of *KAT2B* expression during treatment of chronic HCV, malaria and influenza vaccination, during T1D development in the NOD mouse and in PBMC data from IBD and RA patients

(A) Expression of a type 1 interferon response signature (average eigenvalue of type 1 IFN response signature plotted for each response group at each timepoint, A, signature as defined in[4]) in a cohort of 54 patients during treatment of chronic HCV infection with pegylated interferon-α and ribavirin (as described in[53] and Figure 4C), including 28 showing a marked response (red line, HCV titer decrease > 3.5 $\log_{10}$iu/ml by day 28) and 26 a poor response (HCV titer decrease <1.5 $\log_{10}$iu/ml by day 28), P = 2-way ANOVA. (B) Schematic

representation of the vaccination (black) and transcriptome profiling (red) schedule for the adjuvanted RTS,S Malaria Vaccine Trial[23] (as shown in Fig4D). (B-D) Heatmap (B) and line plot (C, D) illustrating temporal changes in expression of 404 genes representing the GO 'inflammatory response' module (C) or KAT2B expression (D) at each time-point during vaccination in patients with above (red) and below (blue) median KAT2B expression throughout the vaccination schedule outlined in (B). Subgroups defined at T2, immediately following booster vaccination as this equates to the period of most 'active' immune response. Plots = Mean +/− SEM. (E) Schematic representation of the vaccination (black arrows) and transcriptome profiling (red arrows) schedule for 28 vaccinees receiving the 2008 seasonal influenza vaccination (combined trivalent inactivated influenza vaccine[24] as shown in Fig 4E) with response assessed at d28 by HAI titer (green arrow). (F) Linear plot illustrating temporal changes in expression of 404 genes representing the GO 'inflammatory response' module at each time-point during vaccination (d0-d7 corresponding to microarray bleed points in E) for patients showing above (red) or below (blue) median expression of *KAT2B* at day 3 following vaccination. y = expression, log2, x = time-point, days post-vaccination, P = 2way ANOVA. (G) Linear plot showing ratio of *Kat2b* expression in peripheral blood of NOD mice (y-axis, n=37 mice in total across 6 timepoints) prior to and during the induction and onset of insulitis and the development of overt diabetes (illustrated by black bars below). x-axis = age (days), y-axis = *Kat2b* expression $\log_2$ ratio v B10 controls[29]. (H) Kaplan-Meier censored survival curve showing flare-free survival (y-axis) during follow-up (x-axis) of n=58 IBD patients stratified by KAT2B expression (red, above median, blue, below median). P = log-rank test. (I, J) Boxplots showing clinical response (% responders) 3 months post-treatment with anti-TNF therapy in two independent cohorts (I[54] and J[55]) of rheumatoid arthritis (RA) patients. P = Fisher's exact test. Linear plots show mean+/− SEM throughout.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

1. Wherry EJ. T cell exhaustion. Nature immunology. 2011; 12:492–499. [PubMed: 21739672]

2. McKinney EF, et al. A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. Nature medicine. 2010; 16:586–591.

3. Lee JC, et al. Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. The Journal of clinical investigation. 2011; 121:4170–4179. [PubMed: 21946256]

4. Baechler EC, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:2610–2615. [PubMed: 12604793]

5. West EE, et al. Tight regulation of memory CD8(+) T cells limits their effectiveness during sustained high viral load. Immunity. 2011; 35:285–298. [PubMed: 21856186]

6. Aubert RD, et al. Antigen-specific CD4 T-cell help rescues exhausted CD8 T cells during chronic viral infection. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:21182–21187. [PubMed: 22160724]

7. Quigley M, et al. Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. Nature medicine. 2010; 16:1147–1151.

8. Wherry EJ, et al. Molecular signature of CD8+ T cell exhaustion during chronic viral infection. Immunity. 2007; 27:670–684. [PubMed: 17950003]

9. Rangachari M, et al. Bat3 promotes T cell responses and autoimmunity by repressing Tim-3-mediated cell death and exhaustion. Nature medicine. 2012; 18:1394–1400.

10. Francisco LM, Sage PT, Sharpe AH. The PD-1 pathway in tolerance and autoimmunity. Immunological reviews. 2010; 236:219–242. [PubMed: 20636820]

11. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:15545–15550. [PubMed: 16199517]

12. Blackburn SD, et al. Coregulation of CD8+ T cell exhaustion by multiple inhibitory receptors during chronic viral infection. Nature immunology. 2009; 10:29–37. [PubMed: 19043418]

13. Sevilla N, et al. Immunosuppression and resultant viral persistence by specific viral targeting of dendritic cells. The Journal of experimental medicine. 2000; 192:1249–1260. [PubMed: 11067874]

14. Virgin HW, Wherry EJ, Ahmed R. Redefining chronic viral infection. Cell. 2009; 138:30–50. [PubMed: 19596234]

15. Gubin MM, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. Nature. 2014; 515:577–581. [PubMed: 25428507]

16. Baitsch L, et al. Exhaustion of tumor-specific CD8(+) T cells in metastases from melanoma patients. The Journal of clinical investigation. 2011; 121:2350–2360. [PubMed: 21555851]

17. Lang KS, et al. Inverse correlation between IL-7 receptor expression and CD8 T cell exhaustion during persistent antigen stimulation. European journal of immunology. 2005; 35:738–745. [PubMed: 15724249]

18. Mueller SN, Ahmed R. High antigen levels are the cause of T cell exhaustion during chronic viral infection. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:8623–8628. [PubMed: 19433785]

19. Barber DL, et al. Restoring function in exhausted CD8 T cells during chronic viral infection. Nature. 2006; 439:682–687. [PubMed: 16382236]

20. Lyons PA, et al. Microarray analysis of human leucocyte subsets: the advantages of positive selection and rapid purification. BMC Genomics. 2007; 8:64. [PubMed: 17338817]

21. Taylor MW, et al. Changes in gene expression during pegylated interferon and ribavirin therapy of chronic hepatitis C virus distinguish responders from nonresponders to antiviral therapy. Journal of virology. 2007; 81:3391–3401. [PubMed: 17267482]

22. Lauer GM, et al. Full-breadth analysis of CD8+ T-cell responses in acute hepatitis C virus infection and early therapy. Journal of virology. 2005; 79:12979–12988. [PubMed: 16189000]

23. Vahey MT, et al. Expression of genes associated with immunoproteasome processing of major histocompatibility complex peptides is indicative of protection with adjuvanted RTS,S malaria vaccine. The Journal of infectious diseases. 201:580–589. [PubMed: 20078211]

24. Nakaya HI, et al. Systems biology of vaccination for seasonal influenza in humans. Nature immunology. 2010; 12:786–795. [PubMed: 21743478]

25. Querec TD, et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nature immunology. 2009; 10:116–125. [PubMed: 19029902]

26. Hoang LT, et al. The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults. Journal of virology. 2010; 84:12982–12994. [PubMed: 20943967]

27. Shum AK, et al. BPIFB1 Is a Lung-Specific Autoantigen Associated with Interstitial Lung Disease. Science translational medicine. 2013; 5:206ra139.

28. Herazo-Maya JD, et al. Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis. Science translational medicine. 2013; 5:205ra136.

29. Kodama K, et al. Tissue- and age-specific changes in gene expression during disease induction and progression in NOD mice. Clinical immunology. 2008; 129:195–201. [PubMed: 18801706]

30. Elo LL, et al. Early suppression of immune response pathways characterizes children with prediabetes in genome-wide gene expression profiling. Journal of autoimmunity. 2010; 35:70–76. [PubMed: 20356713]

METHODS REFERENCES

31. Stone JH, et al. A disease-specific activity index for Wegener's granulomatosis: modification of the Birmingham Vasculitis Activity Score. International Network for the Study of the Systemic Vasculitides (INSSYS). Arthritis and rheumatism. 2001; 44:912–920. [PubMed: 11318006]

32. Tan EM, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis and rheumatism. 1982; 25:1271–1277. [PubMed: 7138600]

33. Isenberg DA, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology. 2005; 44:902–906. [PubMed: 15814577]

34. Silverberg MS, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol. 2005; 19(Suppl A):5A–36A.

35. Harvey RF, Bradshaw MJ. Measuring Crohn's disease activity. Lancet. 1980; 1:1134–1135. [PubMed: 6103463]

36. Walmsley RS, Ayres RC, Pounder RE, Allan RN. A simple clinical colitis activity index. Gut. 1998; 43:29–32. [PubMed: 9771402]

37. Whitney AR, et al. Individuality and variation in gene expression patterns in human blood. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:1896–1901. [PubMed: 12578971]

38. Le Brigand K, et al. An open-access long oligonucleotide microarray resource for analysis of the human and mouse transcriptomes. Nucleic Acids Res. 2006; 34:e87. [PubMed: 16855282]

39. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004; 3 Article3.

40. Gentleman RC, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004; 5:R80. [PubMed: 15461798]

41. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics. 2002; 18(Suppl 1):S96–104. [PubMed: 12169536]

42. Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. Bioinformatics. 2009; 25:415–416. [PubMed: 19106121]

43. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007; 8:118–127. [PubMed: 16632515]

44. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95:14863–14868. [PubMed: 9843981]

45. Renshaw BR, et al. Humoral immune responses in CD40 ligand-deficient mice. The Journal of experimental medicine. 1994; 180:1889–1900. [PubMed: 7964465]

46. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. [PubMed: 19114008]

47. Barabasi AL, Albert R. Emergence of scaling in random networks. Science. 1999; 286:509–512. [PubMed: 10521342]

48. Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. BMC systems biology. 2007; 1:54. [PubMed: 18031580]

49. Pollard, K.S.a.v.d.L.M.J. A new algorithm for hybrid hierarchical clustering with visualization and the bootstrap. http://cran.fhcrc.org/web/packages/hopach/index.htm

50. Ingenuity Systems. http://www.ingenuity.com

51. Breiman L. Random Forests. Machine Learning Journal. 2001; 45:5–32.

52. Vince Carey, RG. R package. version 1.40.0. Jess Mar, and contributions from Jason Vertrees and Laurent Gatto. MLInterfaces: Uniform interfaces to R machine learning procedures for data in Bioconductor containers.

53. Cramp ME, et al. Hepatitis C virus-specific T-cell reactivity during interferon and ribavirin treatment in chronic hepatitis C. Gastroenterology. 2000; 118:346–355. [PubMed: 10648463]

54. Bienkowska JR, et al. Convergent Random Forest predictor: methodology for predicting drug response from genome-scale data applied to anti-TNF response. Genomics. 2009; 94:423–432. [PubMed: 19699293]

55. Toonen EJ, et al. Validation study of existing gene expression signatures for anti-TNF treatment in patients with rheumatoid arthritis. PloS one. 2012; 7:e33199. [PubMed: 22457743]
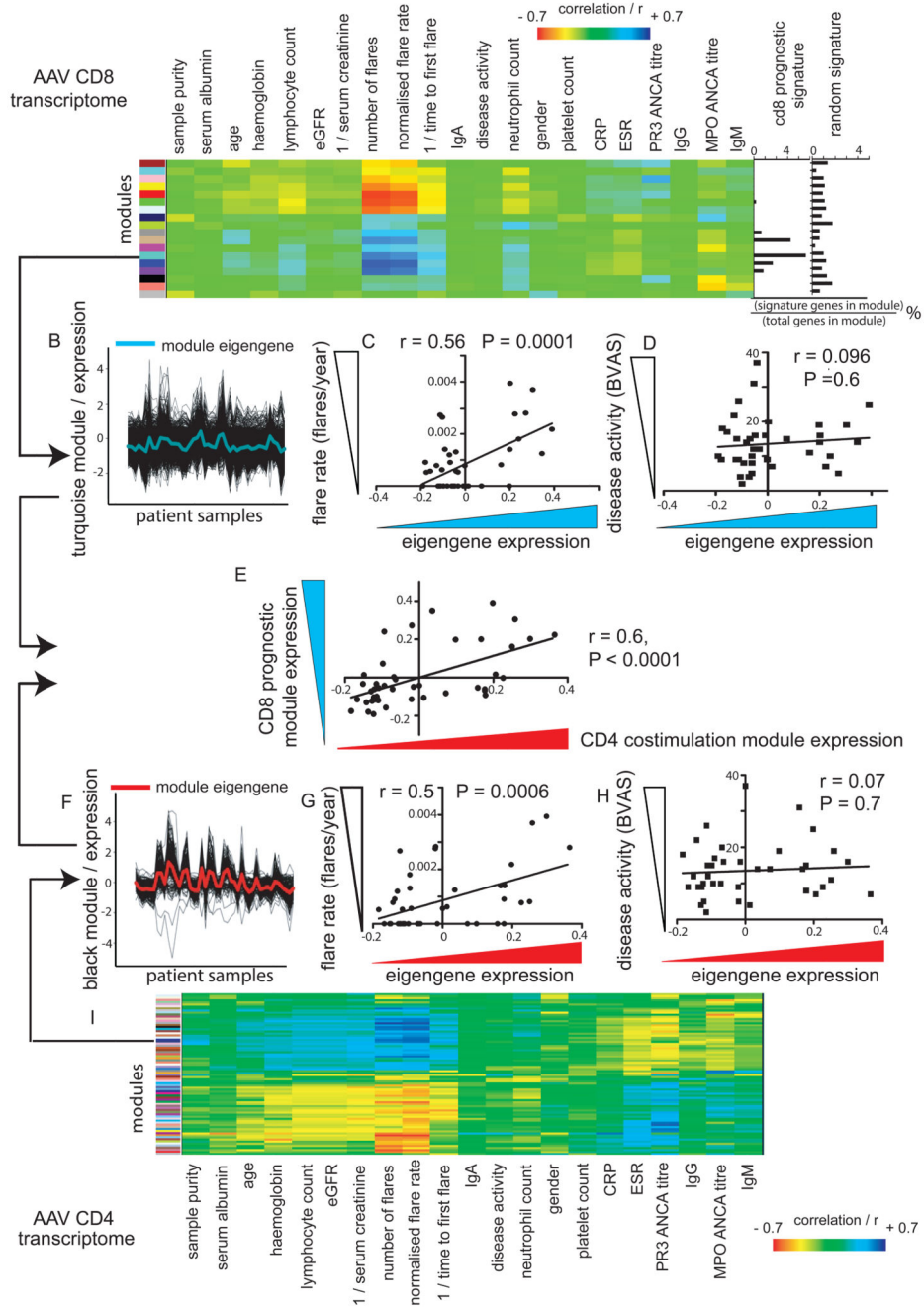
Figure 1. Weighted gene co-expression network analysis of the T cell transcriptome and its correlation with clinical phenotype in AAV

(A, I) Heatmaps illustrating the correlation of CD8 (A) and CD4 (I) co-expression modules (colored blocks, y-axis) with clinical traits in AAV (n=44). Prognostic[2] and random signature overlap with modules shown (A, right) (overlap = signature genes / module genes %). (B, F) Linear plots illustrating turquoise (B) and black (F) modules and summary eigengenes, y = expression ($\log_2$ ratio), x = samples. (C, D, G, H) Scatterplots showing normalized flare-rate (C, G) and disease activity (D, H, Birmingham Vasculitis Activity Score (BVAS), y-axis) against turquoise (C, D) or black (G, H) module eigengene

expression (x-axis). (E) Scatterplot showing correlation between CD4 T cell black (x-axis) and CD8 T cell turquoise module eigengenes (y-axis). Pearson correlation, r, with P = 2-tailed significance.
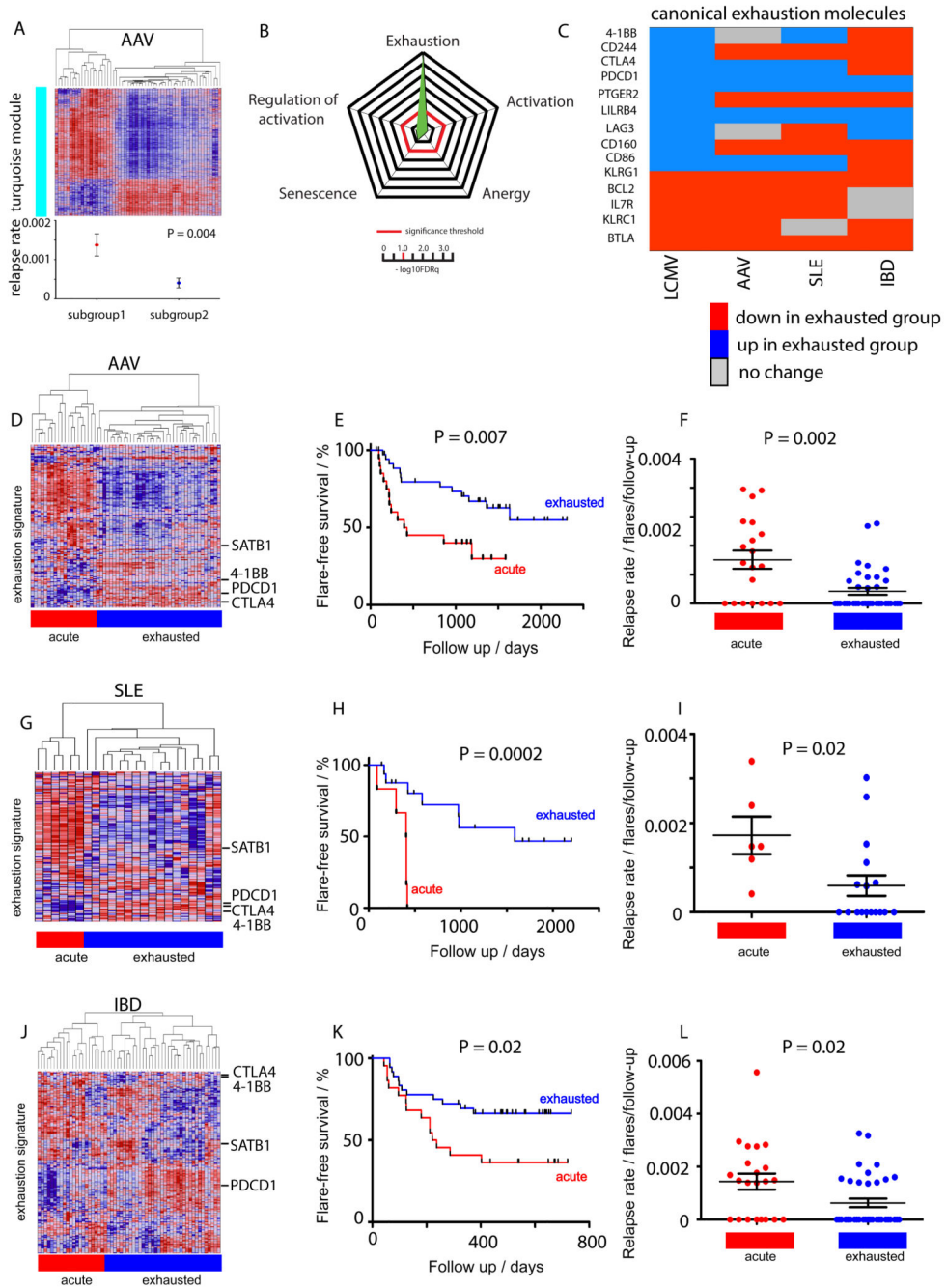
Figure 2. A gene expression signature of CD8 T cell exhaustion predicts contrasting outcomes in infection and autoimmune disease

(A) Heatmap showing hierarchical clustering of AAV patients (n=44) by expression of the turquoise module (Fig.1B) with corresponding flare rates (flares/days follow-up, y-axis). (B) Windrose plot showing GSEA significance (increasing from center, $-\log_{10}$FDRq value) of CD8 T cell signatures tested between prognostic subgroups defined in (A). (C) Heatmap showing differential expression of exhaustion-associated coinhibitory receptors between prognostic subgroups identified in D, G, J. Blue = up, red = down in exhausted, grey = no change (FDR p <0.05). (D, G, J) Heatmaps showing hierarchical clustering of CD8 T cell

expression data isolated from patients with AAV (D, n=58), SLE (G, n=23) and IBD (J, n=58) using a murine CD8 exhaustion signature[8]. 'Exhausted' (blue) and 'non-exhausted' (red) patient subgroups defined from the primary division of the cluster dendrogram. (E, H, K) Kaplan-Meier curves showing censored flare-free survival and (F, I, L) scatterplots showing normalized flare-rate against duration of follow-up for patient subgroups defined in (D, G, J) for AAV (E, F), SLE (H, I) and IBD (K, L) cohorts. (E, H, K) P = log-rank test. (A, F, I, L) P = Mann-Whitney test.
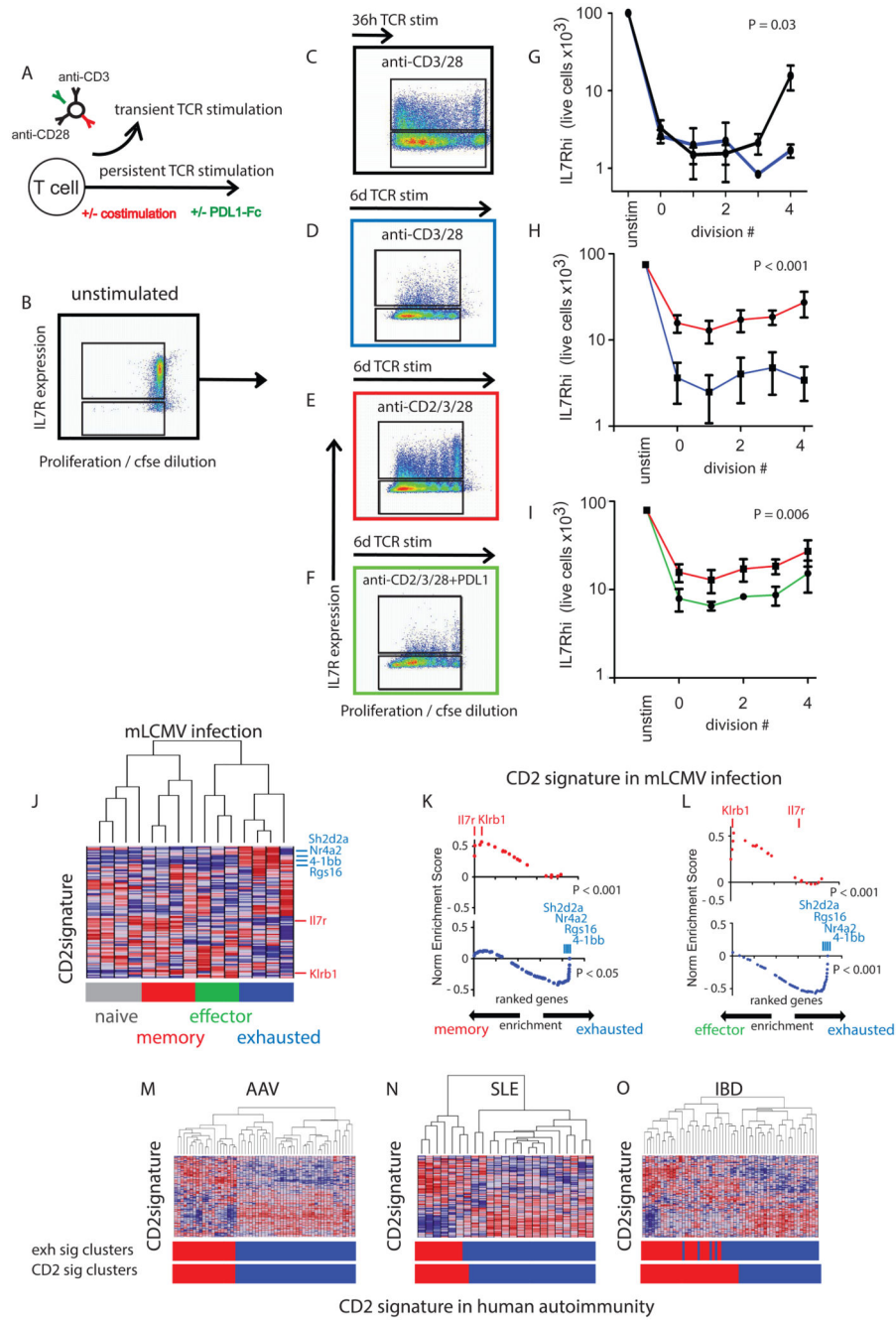
Figure 3. T cell costimulation with CD2 prevents development of an exhausted IL7R^loPD1^hi phenotype

(A) Schematic of the magnetic bead system providing variable TCR signal duration/costimulation during *in vitro* culture. (B-F) Scatterplots illustrating IL7R expression by cell division in unstimulated CD8 T cells (B) and following each of three different costimulation cultures (C-F), as indicated. (G-I) Linear plots showing IL7R^hi population resulting from (G) 36h (black line) v 6d (blue line) anti-CD3/28 stimulation, (H) 6d anti-CD2/3/28 (red line) v 6d anti-CD3/28 (blue line) and from 6d anti-CD2/3/28 with (green line) and without

(I, red line) Fc-PDL1. (J) Heatmap showing unsupervized hierarchical clustering of murine CD8 T cell gene expression data[8] before (naïve, grey), 8 days (effector, green) or 30 days (memory, red) after acute or >30 days (exhausted, blue) after chronic LCMV infection clustered by a CD2 response signature. (K, L) Scatterplots showing GSEA enrichment for genes up (red) and downregulated (blue) by CD2 in (K) memory v exhausted and (L) effector v exhausted CD8 T cells. (M-O) Heatmap showing unsupervized hierarchical clustering of AAV (M, n=58), SLE (N, n=23) and IBD (O, n=58) CD8 T cell expression data using the CD2 response signature. 'Exhausted' (blue) and 'non-exhausted' (red) subgroups were defined from the major division of the cluster dendrogram. Upper bar indicates comparison with patient subgroups produced using the murine LCMV exhaustion signature (as shown in Fig.2D, G, J). Enrichment by GSEA of CD2 signature in autoimmune subgroups < FDR q 0.1.
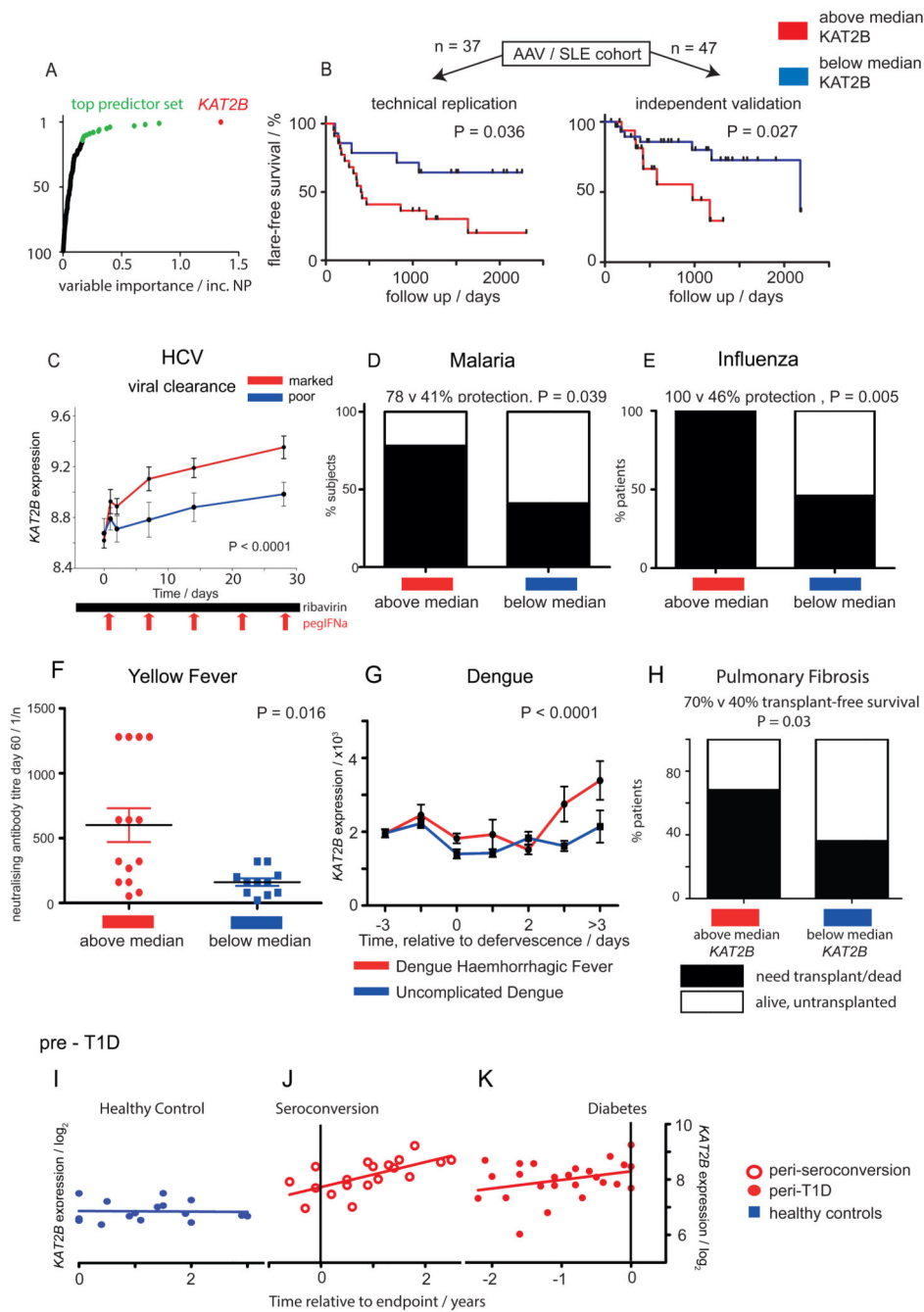
Figure 4. A surrogate marker of CD4 costimulation in PBMC gene expression data correlates with clinical outcome in chronic viral infection, vaccination, infection and autoimmunity
(A) Scatterplot showing the top 100 genes ranked by ability to identify CD4 T cell costimulation subgroups in PBMC data. x-axis = variable importance. (B) Kaplan-Meier plots showing censored flare-free survival stratified by expression of KAT2B (red = above median, blue = below median) in AAV and SLE patients (n=37, training set) replicated on Affymetrix GeneST1.0 and in an independent cohort (test set, n=47), P = log-rank test. (D) Line and scatterplots showing serial KAT2B expression (n=54) following therapy of chronic HCV infection giving a marked (red, n=28) or poor response (blue, n=26). P = 2-way

ANOVA. (E) Boxplot showing post-vaccine malaria protection in a clinical trial (n=43) stratified by KAT2B expression (red = above, blue = below median), P = Fisher's exact test. (F) Boxplot showing % protection (black) in vaccinees (n=28) following seasonal influenza vaccine stratified by KAT2B expression, P = Fisher's exact test (G) Scatterplot showing neutralizing antibody titer following YF-17D vaccination, stratified by KAT2B expression (F, G red = above, blue = below median KAT2B). P = Mann-Whitney test. (H) Line and scatterplot showing serial KAT2B expression throughout dengue infection (n=78) stratified by progression to hemorrhagic fever (DHF, n=24) or uncomplicated course (UD, n=54). x-axis = time (days) relative to defervescence. (H) Boxplot showing % IPF patients (n=75) progressing to transplantation/death (black) stratified by KAT2B expression (red = above median, blue = below median). P = Fisher's exact test. (I-K) Scatterplots showing serial KAT2B expression in healthy age, sex and HLA-matched controls (I, blue) and in pre-T1D cases (n=5, red), 2 of which seroconvert to islet-cell antibodies (J, black line) and 3 of which develop T1D (K, black line).