T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection

Eoin F. McKinney¹,², James C. Lee¹,², David R. W. Jayne³, Paul A. Lyons¹,² & Kenneth G. C. Smith¹,²

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. Here 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 both by persistence of antigen and by a lack of accessory 'help' signals. In autoimmunity, we find that where evidence of CD4 T-cell co-stimulation is pronounced, that of CD8 T-cell exhaustion is reduced. We can reproduce the exhaustion signature by modifying the balance of persistent stimulation of T-cell antigen receptors and specific CD2-induced co-stimulation 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 co-stimulation is 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 co-stimulation/exhaustion signatures in independent data sets, we confirm an association with good clinical outcome or response to therapy in infection (hepatitis C virus) and vaccination (yellow fever, malaria, influenza), but poor outcome in autoimmune and inflammatory disease (type 1 diabetes, anti-neutrophil cytoplasmic antibody-associated vasculitis, systemic lupus erythematosus, idiopathic pulmonary fibrosis and dengue haemorrhagic fever). 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. However, the biology underlying this observation was not clear. We therefore applied weighted gene co-expression network analysis (Extended Data Fig. 1) to the transcriptomes of purified CD4 and CD8 T cells isolated from a prospective cohort of 44 patients with idiopathic pulmonary fibrosis and dengue haemorrhagic fever. By way of validation, we repeated this analysis using an independent co-expression network algorithm that similarly demonstrated association between a CD4 co-stimulation 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 modelling (Extended Data Fig. 3b–e) and was only apparent during active disease (Extended Data Fig. 3f and Supplementary Discussion).

During chronic viral infection, CD8 T-cell memory responses are exquisitely dependent on CD4 T-cell co-stimulation, which can lead to the resolution of chronic infection both in mice and in humans. When antigen persists in the absence of co-stimulation, CD8 T cells become 'exhausted', 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. Although mice lacking inhibitory receptors have an increased incidence and severity of autoimmunity, 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) to test for altered expression of transcriptional signatures reflecting T-cell exhaustion (and other T-cell-related phenotypes) between subgroups of patients defined by the CD8 modular analysis, who go on to develop relapsing or quiescent autoimmunity (Fig. 2a). Using this approach, we observed that genes specifically downregulated in exhausted CD8 T cells during chronic murine lymphocytic choriomeningitis virus (LCMV) infection (but not altered in memory, naive or effector cells) were co-correlated (Fig. 1e) despite being mutually exclusive. A similar analysis using a cohort of 23 patients with systemic lupus erythematosus (SLE) also presenting with active, untreated disease (Supplementary Table 3) 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 (IFN) response signature was associated with disease activity but not with long-term outcome (Extended Data Fig. 2f), consistent with previous reports.

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 co-stimulation (Extended Data Figs 1f, g and 3a and Supplementary Tables 2 and 3). By way of validation, we repeated this analysis using an independent co-expression network algorithm that similarly demonstrated association between a CD4 co-stimulation 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 modelling (Extended Data Fig. 3b–e) and was only apparent during active disease (Extended Data Fig. 3f and Supplementary Discussion).

During chronic murine LCMV infection, CD8 T-cell memory responses are exquisitely dependent on CD4 T-cell co-stimulation, which can lead to the resolution of chronic infection both in mice and in humans. When antigen persists in the absence of co-stimulation, CD8 T cells become ‘exhausted’, 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. Although mice lacking inhibitory receptors have an increased incidence and severity of autoimmunity, 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) to test for altered expression of transcriptional signatures reflecting T-cell exhaustion (and other T-cell-related phenotypes) between subgroups of patients defined by the CD8 modular analysis, who go on to develop relapsing or quiescent autoimmunity (Fig. 2a). Using this approach, we observed that genes specifically downregulated in exhausted CD8 T cells during chronic murine lymphocytic choriomeningitis virus (LCMV) infection (but not altered in memory, naive or effector cells; Supplementary Table 6 (ref. 8)) were similarly downregulated in CD8 T cells from patients at low risk of subsequent relapse (Fig. 2b and Extended Data Figs 3g–i and 4).

During chronic murine LCMV infection, T-cell exhaustion is driven by coordinate upregulation of multiple co-inhibitory receptors that signal synergistically to produce a state of generalized

¹Department of Medicine, University of Cambridge School of Clinical Medicine, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 0QQ, UK. ²Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, Cambridge CB2 0XY, UK.

©2015 Macmillan Publishers Limited. All rights reserved
immunosuppression. 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 co-inhibitory receptors (Fig. 2c). Although a divergence from the murine LCMV model, T-cell exhaustion accompanied by expression of a limited subset of co-inhibitory receptors (Fig. 2c) is similar to that described in intratumoural CD8 T cells, which are a target for checkpoint therapy (Extended Data Fig. 4i–l).

To confirm whether exhaustion was associated with clinical outcome, we used the murine CD8 T-cell exhaustion signature (Supplementary Table 6 (ref. 8)) to perform unsupervised hierarchical clustering of three independent cohorts of patients with distinct diseases (AAV, Fig. 2d–f; SLE, Fig. 2g–i; inflammatory bowel disease (IBD), Fig. 2j–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 favourable 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 favourable long-term outcome in multiple autoimmune and inflammatory diagnoses.

CD8 T-cell exhaustion is characterized by high expression of co-inhibitory receptors (such as PD-1 (ref. 12)) and low expression of nascent memory markers (such as interleukin (IL)-7R17) and is promoted both by the persistence of antigen by and a lack of accessory co-stimulation. To understand signals driving exhaustion and outcome in autoimmunity, we attempted to recreate the outcome-associated transcriptional signatures using variable T-cell antigen
receptor (TCR) signal duration and co-stimulation of primary human cells in vitro. We stimulated purified human CD8 T cells using a magnetic bead conjugated with antibodies targeting co-stimulatory molecules (Fig. 3a) and measured expression of IL-7R and PD-1 expression (Fig. 3b–i and Extended Data Fig. 5a–g) as markers indicating an exhausted phenotype. Comparison between persistent (6 days) and transient (36 h) TCR stimulation showed that IL-7R 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 co-stimulatory molecules, identified from the CD4 T-cell network analysis described above (Fig. 1i and Extended Data Figs 3a and 5h–k), could overcome the effect of persistent TCR stimulation during in vitro differentiation. We found that specific co-stimulation with anti-CD2 (Fig. 3e, h), but not with other stimuli such as IFN-γ or anti-CD40,
resulted in maintained IL-7R expression, limited upregulation of PD-1 and enhanced cell survival (Fig. 3e and Extended Data Fig. 5).

While CD8 exhaustion is known to limit viral control during chronic infection, exhausted cells may be restored to useful function by blocking inhibitory signalling through PD-1 (ref. 19). Enhancing co-inhibitory signals is therefore a logical therapeutic strategy in autoimmune disease, aiming to facilitate exhaustion despite high levels of co-stimulation that would otherwise be predicted to result in an aggressive relapsing disease course. To test this concept, primary human CD8 T cells were co-stimulated during persistent TCR signalling as above (Fig. 3e) in the presence or absence of a bead-bound Fc-chimaeric version of the principal PD-1 ligand, PDL-1 (Fig. 3a, f). When added to CD2-co-stimulated CD8 T-cell cultures, increased PD-1/PDL-1 signalling suppressed differentiation of a non-exhausted IL-7Rhi subpopulation (Fig. 3f, i).

To define the phenotype of T-cell exhaustion more robustly, as small numbers of surface markers are insufficient, we analysed the transcriptome of CD8 T cells exposed to persistent stimulation with and without CD2 signalling (Supplementary Table 7). This CD2 response signature characterized exhausted cells but not effectors 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 (Figs 2d, g, j and 3m–o). Thus, CD2 signalling during persistent TCR stimulation of primary human CD8 T cells prevents the development of transcriptional changes characteristic of exhaustion, recreating transcriptional signatures associated with opposing outcomes in viral infection and autoimmunity.

To confirm that the transcriptional signatures reflected the development of functional exhaustion in vitro, we showed that cells appearing exhausted (IL-7RloPD-1hi) also expressed markers of apoptotic resistance (BCL2lo), characteristic cytokine patterns (IFN-γ, IL-10) and showed diminished survival on re-stimulation (Extended Data Fig. 6a–e). There was no evidence of preferential accumulation of CD8 T-cell subsets after CD2-induced co-stimulation (Extended Data Fig. 6f–h). These data highlight the importance of CD2 signalling in limiting the development of CD8 T-cell exhaustion in the face of persistent TCR stimulation, 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 co-stimulation and CD8 exhaustion with clinical outcome using published data sets. Most of these profile unseparated peripheral blood mononuclear cells (PBMCs), in which T-cell intrinsic signatures are not readily apparent owing to the confounding influence of expression from other cell types. We therefore used a classification algorithm (randomforests) to identify optimal surrogate markers of co-stimulation/exhaustion modules in PBMC from autoimmune patients taken prospectively from children at high risk of disease but before diagnosis (Extended Data Fig. 7). As the CD8 exhaustion and CD4 co-stimulation signatures were themselves correlated (Extended Data Fig. 3g–i), it became easier to detect their combined signal in PBMC using surrogate markers (Fig. 4a and 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 and to increase cellular resistance to cytotoxic compounds. These characteristics, along with its high expression in memory and T- follicular helper and natural killer (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 both by technical replication (using the same samples run on an independent array platform) and by 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 subgroups of patients varied slightly in different clinical circumstances (Fig. 4c–k and Extended Data Fig. 9).

Combined IFN and ribavirin therapy may result in increased virus-specific T-cell responses in chronic hepatitis C virus (HCV) infection, although such eradication therapy is successful in only 50% of cases and in some no change in endogenous immune response is observed. In a cohort of patients with hepatitis C receiving combination therapy, KAT2B expression was progressively induced and showed significantly greater induction in those ultimately responding to therapy (Fig. 4c and Extended Data Fig. 10a). In a clinical trial of malaria vaccination, high KAT2B expression identified a subgroup with response rates of 78%, almost twice that of the low KAT2B expression group (Fig. 4d and Extended Data Fig. 10b–d). Moreover, response to vaccination for either influenza (Fig. 4e and Extended Data Fig. 10e–f) or yellow fever (Fig. 4f) could be predicted by stratifying recipients on the basis of their expression of KAT2B after exposure to vaccine. Dengue viral infection can result in a wide range of clinical manifestations ranging from asymptomatic infection or self-limiting fever (uncomplicated dengue) to dengue haemorrhagic 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).

We next asked whether surrogate detection of T-cell co-stimulation/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. In a cohort of 75 patients with IPF, 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 type 1 diabetes (T1D), with levels rising sharply during the T-cell initiation phase, long before the onset of diabetic hyperglycaemia (Extended Data Fig. 10g). In a cohort of samples taken prospectively from children at high risk of disease but before its onset, 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 co-stimulatory and co-inhibitory signals that shape T-cell exhaustion coincides 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.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 February 2014; accepted 10 April 2015.
Published online 29 June 2015.
1. Wherry, E. J. T cell exhaustion. Nature Immunol. 12, 492–499 (2011).
2. McKinney, E. F. et al. A CD8 T cell transcription signature predicts prognosis in autoimmune disease. Nature Med. 16, 586–591 (2010).

©2015 Macmillan Publishers Limited. All rights reserved
et al. Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. J. Clin. Invest. 121, 4170–4179 (2011).

3. Lee, J. C. et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc. Natl. Acad. Sci. USA 100, 2610–2615 (2003).

4. Baechler, E. C. et al. Antigen-specific CD4 T-cell help rescues exhausted CD8 T cells during chronic viral infection. Proc. Natl. Acad. Sci. USA 108, 21182–21187 (2011).

5. West, E. E. et al. Tight regulation of memory CD8+ T cell exhaustion during chronic viral infection. Immunology 135, 285–298 (2011).

6. Aubert, R. D. et al. Antigen-specific CD4 T-cell help rescues exhausted CD8 T cells during chronic viral infection. Proc. Natl. Acad. Sci. USA 108, 21182–21187 (2011).

7. Urbani, S. et al. Outcome of acute hepatitis C is related to virus-specific CD4 function and maturation of antiviral memory CD8 responses. Hepatology 44, 126–139 (2006).

8. Wherry, E. J. et al. Molecular signature of CD8+ T cell exhaustion during chronic viral infection. Immunity 27, 670–684 (2007).

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

10. Francisco, L. M., Sage, P. T. & Sharpe, A. H. The PD-1 pathway in tolerance and autoimmunity. Immunol. Rev. 236, 219–242 (2010).

11. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550 (2005).

12. Blackburn, S. D. et al. Coregulation of CD8+ T cell exhaustion by multiple inhibitory receptors during chronic viral infection. Nature Immunol. 10, 29–37 (2009).

13. Sevilla, N. et al. Immunosuppression and resultant viral persistence by specific viral targeting of dendritic cells. J. Exp. Med. 192, 1249–1260 (2000).

14. Virgin, H. W., Wherry, E. J. & Ahmed, R. Redefining chronic viral infection. Nature Med. 18, 1394–1400 (2012).

15. Lang, K. S. et al. Inverse correlation between IL-7 receptor expression and CD8 T cell exhaustion during persistent antigen stimulation. Eur. J. Immunol. 35, 738–745 (2005).

16. Gubin, M. M. et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. Nature 515, 577–581 (2014).

17. Baitsch, L. et al. Exhaustion of tumor-specific CD8+ T cells in metastases from melanoma patients. J. Clin. Invest. 121, 2350–2360 (2011).

18. Muller, S. N. & Ahmed, R. High antigen levels are the cause of T cell exhaustion during chronic viral infection. Proc. Natl. Acad. Sci. USA 106, 8623–8628 (2009).

19. Barber, D. L. et al. Restoring function in exhausted CD8 T cells during chronic viral infection. Nature 439, 682–687 (2006).

20. Lyons, P. A. et al. Microarray analysis of human leucocyte subsets: the advantages of positive selection and rapid purification. BMC Genom. 8, 64 (2007).

21. Taylor, M. W. et al. Changes in gene expression during pegylated interferon and ribavirin therapy of chronic hepatitis C virus distinguish responders from non-responders to antiviral therapy. J. Virol. 81, 3391–3401 (2007).

22. Lauer, G. M. et al. Full-breadth analysis of CD8 T-cell responses in acute hepatitis C virus infection and early therapy. J. Virol. 79, 12979–12988 (2005).

23. Vahey, M. T. et al. Expression of genes associated with immunoproteasome processing of major histocompatibility complex peptides is indicative of protection with adjuvanted RTS,S malaria vaccine. J. Infect. Dis. 201, 580–589 (2010).

24. Nakaya, H. I. et al. Systems biology of vaccination for seasonal influenza in humans. Nature Immunol. 12, 786–795 (2010).

25. Herazo-May, J. D. et al. Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis. Sci. Transl. Med. 5, 205ra136 (2013).

26. Kodama, K. et al. Disease induction and progression in NOD mice. Clin. Immunol. 129, 195–201 (2008).

27. Ello, L. L. et al. Early suppression of immune response pathways characterizes children with prediabetes in genome-wide gene expression profiling. J. Autoimmun. 35, 70–76 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by The National Institute of Health Research (NIHR), Cambridge Biomedical Research Centre and funded by the Wellcome Trust (project 094227/Z/10/Z and program grants 083650/Z/07/Z), a Lister Prize Fellowship (K.G.C.S) and the Lupus Research Institute (Distinguished Innovator Award, K.G.C.S). E.F.M. is a Wellcome–Beit Research Fellow supported by the Wellcome Trust and Beit Foundation (104064/Z/14/Z). The Cambridge Institute for Medical Research is in receipt of a Wellcome Trust Strategic Award (079895). We thank A. Kaser and J. Todd for reviewing the manuscript, staff of the NIHR Cambridge Biomedical Research Centre Cell Phenotyping hub and the patients who provided samples.

Author Contributions E.F.M. collected patients, analysed data and performed the in vitro experiments. E.F.M. wrote the manuscript with P.A.L. and K.G.C.S. E.F.M., K.G.C.S. and requests for materials should be addressed to K.G.C.S (kgcs2@cam.ac.uk) or K.G.C.S (kgcs2@cam.ac.uk).
No statistical methods were used to predetermine sample size.

Ethical approval. Ethical approval for this study was obtained from the Cambridge Local Research Ethics Committee (reference numbers 04/023, 08/H0306/21, 08/H0308/176) and informed consent was obtained from all subjects enrolled.

Patients with AAV. Fifty-nine patients with AAV 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 the BVAS and the clinical impression that induction immunosuppression would be required. Prospective disease monitoring was undertaken monthly with serial BVAS disease scoring, and full biochemical, haematological 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 disease categories: (1) flares: defined as at least one major or three minor BVAS criteria; (2) low-grade activity: no major and one or two minor BVAS criteria; (3) no activity (no major or minor BVAS criteria).

All disease flares were cross-checked 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 (E.F.M.) 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 (C-reactive protein, erythrocyte sedimentation rate) and a reduction in immunosuppressive therapy.

Patients with SLE. The SLE cohort comprised 23 patients attending or referred to the Addenbrooke’s Hospital specialist vasculitis unit between July 2004 and May 2008 who met at least four American College of Rheumatology SLE criteria, presenting with active disease (defined below) and in whom immunosuppressive therapy was to be commenced or increased. After treatment with an immunosuppressant, patients were followed up monthly. Disease monitoring was undertaken with serial British Isles Lupus Assessment Group (BILAG) disease scoring and full biochemical, haematological 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 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.

Patients with IBD. Patients with active Crohn’s disease and ulcerative colitis were recruited from a specialist IBD clinic at Addenbrooke’s Hospital, before starting treatment. Disease activity was defined as follows: (1) flare: definition as at least one major or three minor BVAS criteria; (2) low-grade activity: no major and one or two minor BVAS criteria; (3) no activity (no major or minor BVAS criteria).

Follow-up analysis. Comparisons of outcome and associated clinical variables between subgroups were analysed using a Kaplan–Meier log-rank test and a non-parametric Mann–Whitney U-test or a χ² test as appropriate. Correction for multiple testing was applied using the Bonferroni method or FDR (Benjamini and Hochberg method) where appropriate as indicated.

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. PBMCs, CD4 and CD8 T cells were isolated from 110 ml of whole blood by centrifugation over ficoll and, for T cells, by positive selection using magnetic beads as previously described. The purity of separated cell subsets was determined by flow cytometry and included as a covariate in downstream correlation and network analyses (for example, Fig. 1a, i). Total RNA was extracted from each cell population using an RNaseasy mini kit (Qiagen) with quality assessed using an Agilent BioAnalyser 2100 and RNA quantification performed using a NanoDrop ND-1000 spectrophotometer.

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

Affymetrix Human Gene ST microarray. Aliquots of total RNA (200 ng) were labelled using Ambion WT Sense Target labelling 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 and GeneTitan scanner (Affymetrix) as appropriate.

Published data sets. Published data sets were accessed through either National Center for Biotechnology Information (NCBI) Gene Expression Omnibus or ArrayExpress, imported into R using the Bioconductor package GEOquery and analysed as described. Search criteria incorporated the name of individual diseases and were filtered to human data sets but not by platform used. Studies were only included if they met the following criteria. (1) Similar quality control 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 co-stimulation and exhaustion signatures in autoimmunity were identified; that is, samples taken during active disease without concurrent immunosuppressive therapy. (3) Clinical outcome data were available.

It was not feasible to build a unified predictive model across all available data sets 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 titre decrease greater than 3.5 log10 (international units per millilitre (IU ml⁻¹)) and a poor response as an HCV titre decrease less than 1.5 log10 (IU ml⁻¹) by day 28 after commencing combined therapy with ribavirin and pegylated IFN-α. For the malaria vaccine trial used in Fig. 4d, ‘protection’ was defined as delayed or complete protection from subsequent confirmed Plasmodium falciparum infection. All peptide-pulsed (five bites) compared with infectivity control subjects. For the influenza data used in Fig. 4e, protection was defined as at least one high response to at least one of (three) included strains. A high response was defined as at least a fourfold increase in haemagglutination inhibition titre at day 28 and a titre at least 1:40 as per US Food and Drug Administration guidelines.

All gene expression data used have been deposited in publicly available repositories (National Center for Biotechnology Information (NCBI) Gene Expression Omnibus and ArrayExpress): AAV, SLE (E-MTAB-2452, E-MTAB-157, E-MTAB-145), IBD (E-MTAB-331), LCMV (GSE18323), influenza vaccination (GSE18323, influenza vaccination (GSE29619), yellow fever vaccination (GSE3486), dengue fever (GSE29501), IFP (GSE28211), T1D (GSE28666), NOD (GSE28666), target-disease arthritis (GSE5258, GSE3377), in vitro CD8 stimulation (E-MTAB-3470).

Data preprocessing and quality control. For HSmedianta 25k arrays, raw image data were extracted using Koaarray version 2.4.2 software (Koasa 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 in Bioconductor. Further analysis was then performed in R and only data demonstrating a strong negative correlation (r² > 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. Quality control was performed using the Bioconductor package arrayQualityMetrics with outlying samples removed from downstream analyses. Correction for batch variation was performed using the Bioconductor package ComBat 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 Treeview56, using GenePattern57, or directly in R using hclust in the statistics package.

Differential expression. Differentially expressed genes were identified using linear modelling and an empirical Bayes method58 using an FDR threshold of 0.05 as indicated to determine significance.

Weighted gene co-expression network analysis. Highly correlated genes in immune cell subsets were identified and summarized with a modular eigengene profile using the weighted gene co-expression network analysis (WGCNA) bioconductor package in R50. Normalized, log-transformed expression data were variance-filtered using the inlet Cox point of a ranked list of median absolute deviation values for all probes. A soft thresholding power was chosen on the basis of the criterion of approximate scale-free topology59. 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 and results, and the resulting correlation matrix visualized as a heat map (Extended Data Fig. 1). As each module by definition comprises highly correlated genes, their combined expression may be usefully summarized by eigengene profiles60, effectively the first principal component of a given module (for example, Fig. 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 modular eigengenes with clinical traits (for example, Fig. 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 Table 3). The independent association of a given host module eigengene in gene expression profile (for example, KATZ2B) 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.

Hierarchical ordered partitioning and collapsing hybrid analysis. For validation purposes, highly correlated genes were independently partitioned into discrete modules using a second algorithm, hierarchical ordered partitioning and collapsing hybrid (HOPACH)62 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, 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 data set into clusters, each cluster is reiteratively subdivided until the median split silhouette is maximized, thereby producing an optimal segregation into maximally discrete modules.

Knowledge-based network generation and pathway analysis. The biological relevance of gene groups composing modules identified by co-expression analysis was further investigated using the Ingenuity Pathway Analysis Platform50. GSEA, GSEA was used to further assess whether specific biological pathways or signatures were significantly enriched between subgroups of patients 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) shows statistically significant cumulative changes in gene expression between phenotypic subgroups (such as patients with relapsing or quiescent disease). In brief, all genes are ranked on the basis of their differential expression between two groups then an enrichment score is calculated for a given gene set as the ratio of how often its members appear at the top or bottom of the ranked differential list. One thousand random permutations of the phenotypic subgroups were used to establish a null distribution of enrichment score against which a normalized enrichment score and FDR-corrected $q$ values were calculated. GSEA was run with a focused subgroup of gene signatures (as in Figs 2b and 3k)11 to test the null hypothesis that different CD8 T-cell phenotypes were not significantly enriched in subgroups of patients 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 random forests classification algorithm (Fig. 4a). Although signatures apparent in purified T-cell transcriptome data correlate with clinical outcome, they cannot be similarly detected in data derived from PBMC owing to the confounding influence of expression from other cell types; nor can the same genes be used to predict outcome in PBMC20. 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 might be feasible. The availability of both separated T-cell and PBMC from the same patients at the same time facilitated a supervised search for surrogate markers of the co-stimulation/exhaustion signatures in PBMC. Expression data derived both from CD4 T cells and from PBMC were available for a cohort of $n = 37$ patients (AAV and SLE) after quality control and hybridization to the HsMediate 25k custom microarray platform and constituted a training cohort. Normalized, log-transformed expression data were analysed using the MLInterfaces Bioconductor package in R52. Using PBMC-level expression, data samples were classified into subgroups showing either high or low expression of the co-stimulation/exhaustion signature (as illustrated in Extended Data Fig. 5h, i) and probes were subsequently ranked using the variable importance metric on the basis of their ability to predict allocation to either group. The variable importance for a given gene reflects the change in accuracy of classification (percentage 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 labelled and hybridized on an alternative microarray platform (Affymetrix Gene 1.0 ST) as a technical validation set (Fig. 4b, left panel). PBMC samples from an independent cohort of $n = 47$ cases not included in the training cohort were labelled and hybridized to the Affymetrix Gene 1.0 ST platform as an independent test set (Fig. 4b, right panel). For both technical validation and independent test sets, expression of the optimal biomarker identified in Fig. 4a (KATZ2B) was used to bisect the cohort relative to the median expression and clinical outcome was compared in patients with KATZ2Bhi and KATZ2Blow.

Linear models. Linear modelling was performed in R using the statistics package. This took the form of

$$\text{fit} = -\text{lin}(y \sim x_{1} + x_{2} + x_{3}, \text{data} = \text{mydata}),$$

where $y$ (the response variable) was selected as normalized flare rate (flares per number of days follow-up, and $x_{1} \sim x_{3}$ (the test variables) were selected to include both the clinical features and laboratory markers of inflammation), quantification of circulating leucocyte subsets (lymphocytes, neutrophils) and concurrent measurements of autoantibody titre where relevant. Test variables also included a biomarker profile (for example, exhaustion signature or KATZ2B expression). The significance and magnitude (regression coefficient, reflecting change in response variable (flares per number of 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 each model generated.

T-cell culture. Primary human CD8 T cells were separated from leucocyte cones obtained from NHS Blood and Transplant (Addenbrooke’s Hospital) by centrifugation over ficoll and positive selection using magnetic beads as previously described29. The purity of separated cell subsets was determined by three-colour flow cytometry. Purified T cells were labelled with 10 $\mu$M CFSE (Invitrogen) and
32. Tan, E. M.
33. Isenberg, D. A.

Human PD-L1 Fc Chimera (Life Technologies, 1

In some experiments, additional co-stimulation was provided by the addition of
magnetic iBead construct was removed after 36 h in some instances as indicated.

For re-stimulation experiments, cells were harvested on day 6 after stimulation
and sorted into IL-7Rhi and IL-7Rlo populations (Extended Data Fig. 6d) using a

Flow cytometry. Immunophenotyping was performed using an LSR Fortessa
analyser (BD Biosciences), and data were analysed using FlowJo software (Tree
Star). Reactions were standardized with multicolour calibration particles (BD
Biosciences), and live/dead discrimination performed using an AquaFluorescent
amine-reactive dye (Invitrogen). Cell numbers were normalized and were re-suspended in complete RPMI 1640 (2 × 10^6 per millilitre, Sigma-Aldrich) and ‘rested’ in a sterile, U-bottomed culture plate (Greiner) for 6 days.

Flow cytometry. Immunophenotyping was performed using an LSR Fortessa
analyser (BD Biosciences), and data were analysed using FlowJo software (Tree
Star). Reactions were standardized with multicolour calibration particles (BD
Biosciences), and live/dead discrimination performed using an AquaFluorescent
amine-reactive dye (Invitrogen). Cell numbers were normalized and were re-suspended in complete RPMI 1640 (2 × 10^6 per millilitre, Sigma-Aldrich) and ‘rested’ in a sterile, U-bottomed culture plate (Greiner) for 6 days (37°C, 5% CO2) before being re-stimulated (anti-CD2/3/28 1:2 bead:cell ratio, Milteny MACSiBead) for a further 6 days in the presence of IL-2 (10 ng ml^-1, 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.

For re-stimulation experiments, cells were harvested on day 6 after stimulation
and sorted into IL-7Rhi and IL-7Rlo populations (Extended Data Fig. 6d) using a

Flow cytometry. Immunophenotyping was performed using an LSR Fortessa
analyser (BD Biosciences), and data were analysed using FlowJo software (Tree
Star). Reactions were standardized with multicolour calibration particles (BD
Biosciences), and live/dead discrimination performed using an AquaFluorescent
amine-reactive dye (Invitrogen). Cell numbers were normalized and were re-suspended in complete RPMI 1640 (2 × 10^6 per millilitre, Sigma-Aldrich) and ‘rested’ in a sterile, U-bottomed culture plate (Greiner) for 6 days (37°C, 5% CO2) before being re-stimulated (anti-CD2/3/28 1:2 bead:cell ratio, Milteny MACSiBead) for a further 6 days in the presence of IL-2 (10 ng ml^-1, 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
analyser (BD Biosciences), and data were analysed using FlowJo software (Tree
Star). Reactions were standardized with multicolour calibration particles (BD
Biosciences), and live/dead discrimination performed using an AquaFluorescent
amine-reactive dye (Invitrogen). Cell numbers were normalized and were re-suspended in complete RPMI 1640 (2 × 10^6 per millilitre, Sigma-Aldrich) and ‘rested’ in a sterile, U-bottomed culture plate (Greiner) for 6 days (37°C, 5% CO2) before being re-stimulated (anti-CD2/3/28 1:2 bead:cell ratio, Milteny MACSiBead) for a further 6 days in the presence of IL-2 (10 ng ml^-1, 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
analyser (BD Biosciences), and data were analysed using FlowJo software (Tree
Star). Reactions were standardized with multicolour calibration particles (BD
Biosciences), and live/dead discrimination performed using an AquaFluorescent
amine-reactive dye (Invitrogen). Cell numbers were normalized and were re-suspended in complete RPMI 1640 (2 × 10^6 per millilitre, Sigma-Aldrich) and ‘rested’ in a sterile, U-bottomed culture plate (Greiner) for 6 days (37°C, 5% CO2) before being re-stimulated (anti-CD2/3/28 1:2 bead:cell ratio, Milteny MACSiBead) for a further 6 days in the presence of IL-2 (10 ng ml^-1, 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
analyser (BD Biosciences), and data were analysed using FlowJo software (Tree
Star). Reactions were standardized with multicolour calibration particles (BD
Biosciences), and live/dead discrimination performed using an AquaFluorescent
amine-reactive dye (Invitrogen). Cell numbers were normalized and were re-suspended in complete RPMI 1640 (2 × 10^6 per millilitre, Sigma-Aldrich) and ‘rested’ in a sterile, U-bottomed culture plate (Greiner) for 6 days (37°C, 5% CO2) before being re-stimulated (anti-CD2/3/28 1:2 bead:cell ratio, Milteny MACSiBead) for a further 6 days in the presence of IL-2 (10 ng ml^-1, 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.
transcriptome segregated into co-expressed modules, each module represented by a distinct colour

black module

module eigengene

eigengene profiles summarise the expression within each module

each module represented by hierarchically inter-related eigengenes

eigengene profiles correlated with clinical phenotype

gene modules

correlation / Pearson r

- 0.7 0.7

FDRp-value (Bonferroni)

0.001

FDRp > 0.05

costimulation (black) module v normalised flare-rate, P = 0.0005

purified cell mRNA

gene modules
Extended Data Figure 1 | Overview of weighted gene co-expression analysis.

a, Messenger RNA derived from purified leucocyte subsets sampled during active, untreated autoimmune disease is labelled and hybridized to a microarray platform (both HsMediante 25k and Affymetrix Gene 1.0 ST used here). Genes are then combined into modules (b, coloured 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 constituting 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 heat map shown here. Coloured blocks represent individual modules, defined as in a. Modules are aligned in identical order on x and y axes with heat-map colour representing the correlation between each. Note that the diagonal (top left to bottom right) therefore represents the correlation of each eigengene profile with itself, and is always 1. Distance metric is the 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 heat map. Pearson correlation, r. f, Heat map showing gene expression modules (y axis) correlated against clinical variables (x axis) for the CD4 transcriptome in AAV. Pearson correlation, r. g, Heat map illustrating significance of correlations identified in f. P value threshold at Bonferroni-corrected P < 0.05. Colour bar indicates actual P value of correlations deemed significant; grey shading, corrected P > 0.05. Significance for co-stimulation (black) module from Fig. 1 is also shown (P = 0.0005).
Extended Data Figure 2 | Weighted gene co-expression network analysis of the T-cell transcriptome and its correlation with clinical phenotype in SLE.

a, e. Heat maps illustrating the correlation of co-expression modules (coloured 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 co-expression modules, along with the distribution of a random signature of equivalent size, shown to the right of a (overlap = signature genes/module genes, as a percentage). Overlap of the CD4 T-cell co-stimulation 'black' module (defined in Fig. 1) shown to the right of e, with a randomly derived module and a type 1 IFN response signature previously shown to associate with active SLE4. Overlap shown as percentage 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; coloured lines (red, blue), module eigengenes. c, Correlation of SLE CD4 T-cell co-stimulation module eigengene (x axis, blue) against SLE CD8 T-cell prognostic signature (y axis, red). Pearson correlation, r, with P = two-tailed significance. f, Expanded detail from e, illustrating that modules corresponding to type 1 IFN response and co-stimulation signatures correlate with disease activity and outcome respectively but not vice versa.
# of documented pathway connections

## a

### b
CD8 turquoise module

- Normalised Flare Rate

| negLOG10p | regression coefficient |
|-----------|------------------------|
| 2.0       | AAV                    |
| 1.0       | AAV/SLE                |

## c
CD4 black module

- Normalised Flare Rate

| negLOG10p | regression coefficient |
|-----------|------------------------|
| 2.0       | CD4                    |
| 1.0       | CD4                    |

## d
CD8 exhaustion signature

- Normalised Flare Rate

- Test variable other variables

- sig threshold (P = 0.05)

## e

### f
CD8 exhaustion signature

- Time / months

- non-exhausted (0 months)
- exhausted (0 months)

- e0 v E0, P < 0.0001
- e12 v E12, P = 0.78

## g

### h
AAV

- SLE

- r = 0.6, P < 0.0001
- r = 0.6, P = 0.009

## i
IBD

- r = 0.2, P = 0.2

©2015 Macmillan Publishers Limited. All rights reserved
Extended Data Figure 3 | Identification and validation of genes involved in CD4 co-stimulation that correlate with clinical outcome, and how that relationship changes after treatment. a, A knowledge-based network analysis of 336 probes composing the ‘black’ expression module (Fig. 1e) identifies a network of co-stimulation signalling (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 signalling, inducible T-cell co-stimulator and its ligand (ICOS–ICOSL) signalling and CD28 signalling are all significantly enriched in this module (FDR $P < 0.05$).

b–e, Scatter plots showing the outcome of multiple linear regression models testing the association of four signatures (red symbols) as indicated, directly compared with clinical markers of disease activity (black symbols); x axis, magnitude of association (regression coefficient, change in normalized flare rate (flares per number of 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 co-stimulation (black) module eigengene in AAV, (d, e) CD8 exhaustion signature (Supplementary Table 6) in AAV/SLE (d) and IBD (e). Clinical variables incorporated vary owing to differing relevance in each case but include some of the following: disease activity score (BVAS/BILAG/CDAI/Harvey–Bradshaw score), C-reactive protein, autoantibody titre (PR3/MPO, dsDNA), lymphocyte count, neutrophil count, platelet count, IgG, IgA, IgM, erythrocyte sedimentation rate, age. f, Line plot showing mean expression of a CD8 T-cell exhaustion signature in 38 patients with AAV measured at presentation during active, untreated disease ($t_0$) and 12 months later when disease activity was quiescent and patients were on maintenance immunosuppressive therapy ($t_{12}$). 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 $t_{12}$ and $t_0$ values. The difference between the groups that is easily apparent at enrolment with active, untreated disease ($t_0$) is no longer apparent when disease is treated and quiescent 12 months later ($t_{12}$).

g–i, Scatter plots showing inverse correlation between individual eigenvalues of the CD4 co-stimulation 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. Pearson correlation, $r^2$, two-tailed significance.
AAV

Regulation of activation

Exhaustion

Activation

Anergy

Senescence

SLE

Exhaustion

Regulation of activation

Exhaustion

Activation

Anergy

Senescence

IBD

Exhaustion

Regulation of activation

Exhaustion

Activation

Anergy

Senescence

Melanoma TILN CD8 v circulating CD8

Exhaustion

Regulation of activation

Exhaustion

Activation

Anergy

Senescence

4-1BB
CD244
CTLA4
PDCD1
PTGER2
LILRB4
LAG3
CD160
CD86
KLRG1
BCL2
IL7R
KLRC1
BTLA

down in exhausted group
up in exhausted group
no change
Extended Data Figure 4 | Wind rose plots showing relative GSEA enrichment of immune signatures in autoimmune disease and melanoma. Wind rose plots showing relative enrichment (GSEA FDR q value) of distinct immune signatures between subgroups of patients (as defined as in Fig. 2). 

a, b, AAV; c, d, SLE; e, f, IBD. 

a, c, e, Enrichment of immune signatures from selected CD8 T-cell phenotypes; b, d, f, enrichment of signatures specifically up-/downregulated by CD8 T-cell subsets derived from the LCMV model of T-cell exhaustion (acute LCMV Armstrong versus chronic LCMV Cl13 (ref. 8)). Detailed information on genes included in each signature is provided in Supplementary Table 6. 

g, h, Wind rose plots showing relative enrichment (GSEA FDR q value) of distinct immune signatures between CD8 T cells from patients with melanoma, comparing CD8 from tumour-infiltrated lymph node with circulating CD8 T cells26. 

Enrichment of immune signatures from selected CD8 T-cell phenotypes; h, enrichment of signatures specifically up-/downregulated by CD8 T-cell subsets derived from the LCMV model of T-cell exhaustion (acute LCMV Armstrong versus chronic LCMV Cl13 (ref. 8)). Specific enrichment is seen for genes downregulated by exhausted cells but not for all genes upregulated by exhausted cells. 

Heat map showing differential expression of selected canonical co-inhibitory receptors (as for Fig. 2c (ref. 12)) in the LCMV exhaustion model, between prognostic subgroups identified in Fig. 2d, g, j and between exhausted CD8 T cells from melanoma-infiltrated lymph node compared with circulating tumour-specific CD8 T cells26. Blue, up in exhausted; red, up in non-exhausted; grey, no significant change (FDR P < 0.05).
Extended Data Figure 5 | T-cell co-stimulation with CD2, but not type 1 IFN or anti-CD40, prevents development of an exhausted IL-7R<sup>lo</sup>PD1<sup>hi</sup> phenotype during prolonged anti-CD3/28 T-cell stimulation.

a–d. Representative scatter plots showing IL-7R expression (y axis) by cell division (CFSE dilution, x axis) in (a) unstimulated cells and following each of three different co-stimulation cultures: b, anti-CD3/CD28 alone; c, anti-CD2/3/28; d, anti-CD40/3/28. IL-7R<sup>hi</sup>-expressing subset indicated in black gate with the percentage of live cells shown. e–g. Line and scatter plots showing absolute number of IL-7R<sup>hi</sup> cells (e), PD-1 expression (f) and cell death (g) (death = AquaFluorescent dye<sup>+</sup>) during CD8 T-cell differentiation (x axis, number of divisions undergone by day 6 of culture measured by CFSE dilution) after anti-CD3/28 (blue) or anti-CD2/3/28 (red) stimulation. P = paired t-test, n = 5 paired samples. h, i. Hierarchical clustering of 44 patients with AAV (left panels) and 23 patients with SLE (right panels) using 336 genes composing a CD4 T-cell co-stimulation module (black module, Fig. 1) identifies two subgroups of patients (high co-stimulation, red; low co-stimulation, blue) in CD4 T-cell expression data defined by the first major division in the patient dendrogram. j, k. Scatter plots illustrating selected co-stimulatory and co-inhibitory receptors for the subgroups identified in h and i. Selected receptors were chosen on the basis of their inclusion in networks derived from the co-stimulation and exhaustion signatures as illustrated in Extended Data Fig. 3a. l, m. Line and scatter plots showing absolute number of IL-7R<sup>hi</sup> cells (y axis) by number of divisions undergone at day 6 (x axis) after polyclonal stimulation with anti-CD3/28 (blue) or anti-CD3/28 plus anti-CD40 (I, green) or IFN-α (m, green) co-stimulation. n. Line and scatter plot showing extent of proliferation occurring (percentage of live cells on day 6 having undergone each of zero to four divisions) after polyclonal stimulation of primary human CD8 T cells with CD3/28 alone (blue) or with additional anti-CD2 co-stimulation (red), confirming no difference in the extent of live cell proliferation between groups. o. Absolute live (AquaFluorescent Dye<sup>+</sup>) cell counts (y axis) by the number of divisions undertaken (x axis) by day 6 after polyclonal stimulation of primary human CD8 T cells with CD3/28 alone (blue) or with additional anti-CD2 co-stimulation (red), illustrating increased cell survival with CD2 co-stimulation despite equivalent proliferation. P values = two-way ANOVA of four paired stimulations.
Extended Data Figure 6 | CD2 co-stimulation results in functionally distinct subpopulations showing enhanced survival after in vitro re-stimulation 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 the percentage 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, Scatter plots showing cytokine levels (y axis, picograms per millilitre) 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, Scatter plots illustrating populations sorted after polyclonal anti-CD3/28 (left panel) and anti-CD2/3/28 (right panel) stimulation of primary CD8 T cells. e, Percentage of live cells (AquaFluorescent dye−) remaining 7 days after re-stimulation of each sorted subpopulation of CD8 cells. Cells were rested for 6 days in complete RPMI1640 medium without IL-2 before being re-stimulated with anti-CD2/3/28 for a further 7 days. P = Mann–Whitney; error bars, mean ± s.e.m. f, Representative scatter plot illustrating CD8 T-cell memory populations isolated by flow cytometric sorting and stimulated in g, h. g, Scatter plot showing absolute number of IL-7Rhi cells (y axis) on day 6 after 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, Scatter plots showing percentage CD8 T-cell memory subsets (y axis) resulting from stimulation of purified central memory (Tcm), naive (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 Figure 7 | Top PBMC surrogate markers reflect expression of CD4 co-stimulation/CD8 exhaustion modules within CD4 and CD8 data respectively. Top PBMC-level predictors (n = 13) were selected as indicated in Fig. 4a, and data are shown comparing expression of the optimal predictor (KAT2B, a, e) and of each other top predictor gene (d, h) in PBMC data compared with expression of the CD4 co-stimulation 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, Scatter plots 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 with clinical markers of disease activity (black symbols); x axis, magnitude of association (regression coefficient, change in normalized flare rate (flares per number of 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 were disease activity score (BVAS), C-reactive protein, lymphocyte count, neutrophil count, IgG. c, g, Heat maps reproduced from Fig. 1a, i, respectively, showing overlap of top PBMC-level predictors with the modular analysis presented for CD4 (c) and CD8 (g) data in Fig. 1. As expected, surrogate markers showed stronger correlation with the CD4 than the CD8 signature as the algorithm was trained to detect the CD4 co-stimulation module.
** P < 0.001 correlation with KAT2B expression across all subsets
Extended Data Figure 8 | Immune cell subset expression pattern of top PBMC-level surrogate markers of CD4 co-stimulation/CD8 exhaustion signatures. Dot plots showing expression (median ± s.e.m.) of KAT2B (a) and for each of 12 other top PBMC-level surrogate predictors of CD4 co-stimulation/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 Figure 9 | Hierarchical clustering of multiple data sets using 13 top PBMC-level surrogate markers of CD4 co-stimulation/CD8 exhaustion modules identifies subgroups of patients with distinct clinical outcomes. Replication of association between surrogate markers of CD4 co-stimulation/CD8 exhaustion signatures and clinical outcome (as shown in Fig. 4c–k) but using all top 13 PBMC-level surrogates rather than KAT2B alone. a, c, e, g, i, k, m. Heat maps showing hierarchical clustering of gene expression data of 13 top PBMC-level surrogate predictors of CD4 co-stimulation/CD8 exhaustion signatures (from Fig. 4a) in patients with chronic HCV (a), during malaria vaccination (c), influenza vaccination (e), yellow fever vaccination (g), dengue fever infection (i), IPF (k) and pre-T1D (m). Subgroups were defined using a major division of the cluster dendrogram and group 1 allocated on the basis of KAT2B expression (highest in group 1). Clinical outcome associated with each subgroup identified is shown in b (HCV, percentage of responders to IFN-α/ribavirin therapy), d (percentage showing protection versus no protection from malaria vaccine), f (percentage response to influenza vaccination), h (yellow fever antibody-titre after vaccination), j (percentage progression to DHF), l (percentage of patients progressing to need for transplantation or death) and n (percentage of samples from patients with previous or subsequent progression to islet-cell antibody seroconversion or to a diagnosis of T1D).
**Research Letter**

**Figure a**
- HCV viral clearance
- IFN response signature (log2)
- pegIFNα
- Ribavirin
- Time / days
- P > 0.05

**Figure b**
- Gene Ontology: Inflammatory response
- T0 T1 T2 T3 T4
- Vaccine:
  - T0: Challenge
  - T1-T4: Immune response
- T2 protection v none
- 76 v 41%, P = 0.04

**Figure c**
- Gene Ontology: Inflammatory response
- T0 T1 T2 T3 T4
- T2 KAT2Bhi
- T2 KAT2Blo

**Figure d**
- Gene Ontology: Inflammatory response
- KAT2B expression
- T0 T1 T2 T3 T4

**Figure e**
- Influenza vaccination (TIV), n=28
- HAI titre
- D0 D3 D7 D28

**Figure f**
- Gene Ontology: Inflammatory response
- KAT2Bhi vs KAT2Blo
- D0 D3 D7

**Figure g**
- KAT2B expression (log2 NOD/B10 ratio)
- T0 T1 T2 T3 T4
- T cell initiation
- Insulitis
- Hyperglycaemia

**Figure h**
- Flare-free survival / %
- IBD
- Above median KAT2B
- Below median KAT2B
- P = 0.33

**Figure i**
- % Patients responding to anti-TNF therapy
- Above median KAT2B
- Below median KAT2B
- P = 0.76

**Figure j**
- % Patients responding to anti-TNF therapy
- Above median KAT2B
- Below median KAT2B
- P = 1

©2015 Macmillan Publishers Limited. All rights reserved
Extended Data Figure 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 patients with IBD and rheumatoid arthritis. a, Expression of a type 1 IFN response signature (average eigenvalue of type 1 IFN response signature plotted for each response group at each time point, A, signature as defined in ref. 4) in a cohort of 54 patients during treatment of chronic HCV infection with pegylated IFN-α and ribavirin (as described in ref. 53 and Fig. 4c), including 28 showing a marked response (red line, HCV titre decrease >3.5 log₁₀(U ml⁻¹) by day 28) and 26 a poor response (HCV titre decrease <1.5 log₁₀(U ml⁻¹) by day 28). *P* = two-way ANOVA. b, Schematic representation of the vaccination (black) and transcriptome profiling (red) schedule for the adjuvanted RTS,S malaria vaccine trial23 (as shown in Fig. 4d). b–d, Heat map (b) and line plots (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 after booster vaccination as this equates to the period of most ‘active’ immune response. Plots are mean ± s.e.m. 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 vaccine24 as shown in Fig. 4e) with response assessed at day 28 by haemagglutination inhibition titre (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 after vaccination; *y* expression log₂; *x*, time-point, days after vaccination; *P* = two-way ANOVA. g, Linear plot showing ratio of Kat2b expression in peripheral blood of NOD mice (*y* axis, *n* = 37 mice in total across six time points) before 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 log₂ ratio versus B10 controls29. h, Kaplan–Meier censored survival curve showing flare-free survival (*y* axis) during follow-up (*x* axis) of *n* = 58 patients with IBD stratified by KAT2B expression (red, above median; blue, below median). *P* = log-rank test. i, j, Box plots showing clinical response (percentage responders) 3 months after treatment with anti-TNF therapy in two independent cohorts (I54 and J55) of patients with rheumatoid arthritis (RA). *P* = Fisher’s exact test. Linear plots show mean ± s.e.m. throughout.