A CD8$^+$ NK cell transcriptomic signature associated with clinical outcome in relapsing remitting multiple sclerosis

Eoin F. McKinney*[1,2], Iona Cuthbertson[2], Kristina M. Harris[3], Dawn E. Smilek[3], Christopher Connor[2], Giulia Manferrari[2], Edward J. Carr[2], Scott S. Zamvil[4] and Kenneth G.C. Smith[1,2].

1 Cambridge Institute for Therapeutic Immunology and Infectious Disease, Jeffrey Cheah Biomedical Centre, Cambridge, U.K.
2. Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, U.K.
3. Biomarker Discovery Research, ITN, Bethesda, Maryland, USA.
4. Department of Neurology and Program in Immunology, University of California, San Francisco, CA USA.
*e-mail: efm30@cam.ac.uk

Abstract

Multiple sclerosis (MS) is an inflammatory demyelinating disease of the central nervous system (CNS) with the majority of cases characterised by relapsing/remitting (RRMS) attacks of neurologic dysfunction followed by variable resolution. Improving clinical outcomes in RRMS requires both a better understanding of the immunological mechanisms driving recurrent demyelination and better means of predicting future disease course to facilitate early targeted therapy. Here, we apply hypothesis-generating network transcriptomics to CD8$^+$ cells isolated from patients in RRMS, identifying a signature reflecting expansion of a subset of CD8$^+$ natural killer cells (NK8$^+$) associated with favourable outcome. NK8$^+$ are capable of regulating CD4$^+$ T cell activation and proliferation in vitro, with reduced expression of HLA-G binding inhibitory receptors and consequent reduced sensitivity to HLA-G-mediated suppression. We identify surrogate markers of the NK8$^+$ signature in peripheral blood leucocytes and validate their association with clinical outcome in an independent cohort, suggesting their measurement may facilitate early, targeted therapy in RRMS.

Introduction

The process of T cell exhaustion – by which CD8$^+$ T cells acquire effector dysfunction in the face of persistent antigen stimulation – has been associated with clinical outcome, not only in multiple distinct autoinflammatory diseases but also in both chronic viral infections and anti-tumour immunity[1]. In relapsing remitting multiple sclerosis (RRMS), relapse resolution associates with favourable long-term outcome[2,3], however the frequency and manifestations of early attacks are highly variable and the immunopathology of RRMS onset, relapse and progression remains poorly understood[4]. Distinct immunological pathways are thought to drive relapse and secondary neurodegenerative progression[5] with CD4$^+$ T cells remaining the prime suspect for initiating demyelination in both human disease and animal models[5] . Natural killer (NK) cells are an innate lymphocyte subset capable of eliminating cells identified as either missing-self or altered self[6]. Missing self elimination occurs when downregulation of self major histocompatibility complexes (MHC, e.g. on virally infected cells[7]) attenuates signalling through NK inhibitory receptors. Altered self responses are driven by activating NK receptors whose ligands are upregulated on transformed, stressed or activated cells[7]. NK cells may also play an immunoregulatory role[8], as highlighted by enhanced infection-induced immunopathology in NK-deficient mice, where NK fail to control autologous T cell activation[9]. Genetic variants or alterations in NK cell function or phenotype have been repeatedly linked to multiple autoimmune diseases including Rheumatoid arthritis[10], Type 1 diabettes (T1D)[11] and

RRMS[12], although it remains unclear whether this reflects a contribution to tissue injury or failure to control immune reactivity[13].

Here we apply unsupervised network transcriptomics to circulating $CD8^+$ leucocytes isolated from patients who had experienced a first demyelinating episode (clinically isolated syndrome, CIS), were not on treatment and were at high risk for RRMS relapse in a phase II randomised, placebo-controlled trial of the immumodulatory effect of atorvatstatin (STAyCIS)[14]. Samples were taken prior to commencing therapy in the trial to ensure this did not impact on the baseline result. In the STAyCIS study we observe modules of coexpressed transcripts in this $CD8^+$ population that are associated with favourable clinical outcome, are independent of treatment effects and reflect expansion of a $CD8^+CD3^-CD56^+$ natural killer cell subset ($NK8^+$) rather than a profile of $CD8^+$ T cell exhaustion. In vitro studies indicate that $NK8^+$ can regulate autologous $CD4^+$ T cell activation and proliferation with reduced sensitivity to Human Leucocyte Antigen G (HLA-G) mediated suppression. We validate the association of an $NK8^+$ gene signature with reduced relapse risk in peripheral blood mononuclear cells from an independent cohort of 94 RRMS and CIS cases.

Results

Network coexpression analysis of early MS

Unsupervised analyses of high throughput 'omic' data facilitate the discovery of previously unsuspected associations as they are, by definition, not founded on existing knowledge[15]. We performed unsupervised transcriptomic analysis of $CD8^+$ cells isolated from patients enrolled after a first demyelinating episode into a randomized, double blind, placebo-controlled clinical trial (STAyCIS[14]) that tested the effect of atorvastatin on relapse in patients at the earliest clinical phase of MS. $CD8^+$ cells were isolated from peripheral blood mononuclear cells (PBMC) collected at the baseline trial visit (Supp Fig 1) allowing unsupervised identification of coexpressed gene 'modules' (weighted gene coexpression analysis, WGCNA[16]). 'Eigengenes' summarising modular coexpression signatures were then correlated to prospective and baseline clinical data, including the primary endpoint of the study (PEP, one clinical relapse or >=3 new T2 MRI lesions within 12 months, Fig 1A), the treatment arm (atorvastatin or placebo) and technical covariates (Supp Fig 2, Supp Data 1). Four $CD8^+$ cell-intrinsic eigengenes were specifically and significantly correlated with clinical and radiological study endpoints (Fig 1A, B). While transcripts associated with T cell exhaustion[1] were coexpressed in the RRMS dataset, they were not associated with clinical outcome (Supp Fig 2E ,F).

Biological interpretation of progression-associated transcriptional signature

Next, we sought to interpret gene modules associated with relapse risk by performing enrichment analysis against public repositories of immune signatures[17]. The largest module ('black', Supp Data 2) was strongly enriched only for NK-cell specific transcripts as defined in two distinct immune signature repositories[18,19] while smaller modules showed no clear enrichment (Fig 1C). Supervised clustering using the 'black' transcripts identified patient subgroups with distinct outcomes that were comparably enriched for NK transcripts (Fig 1D, E). This NK signature was not correlated with

basline clinical traits (Supp Data 1) and was the strongest predictor identified in a multivariate penalised Cox regression model of time to relapse, although measures of clinical severity (T1 lesion load and clinical severity) further improved prediction when used alongside it (Fig 1H). In humans, a subset of NK cells also express the CD8 coreceptor[20],albeit at lower levels than their T cell counterparts, and flow cytometry confirmed that $CD3^-CD56^+CD8^+$ NK cells ($NK8^+$) comprised a significant minority of the isolated $CD8^+$ fraction profiled by transcriptomics ($NK8^+$, Supp Fig 3A, B). Correlation of the outcome-associated 'black' signature with extensive concurrent immunophenotyping confirmed robust, specific association with an $NK8^+$ expansion in peripheral blood (Fig 1F, Supp Fig 3D-G). To further confirm $NK8^+$ cells as the source of this signature, we isolated $NK8^+$ and $NK8^-$ cells from healthy individuals, performing RNAseq to identify a signature defining the $NK8^+$ subset, and demonstrated strong differential enrichment between disparate outcome groups identified with the 'black' signature (Fig 1G, Supp Data 3). Together these data demonstrate that a transcriptional signature reflecting an expanded population of $NK8^+$ cells is associated with reduced future relapse risk following an initial demyelinating event.

An autoregulatory role for $NK8^+$ cells

We next sought to investigate potential immune mechanisms explaining the association of $NK8^+$ cells with reduced relapse risk. Cytotoxic NK cell function is controlled by a complex interplay of both activating and inhibitory signals[21]. As an $NK8^+$ expansion associated with reduced relapse risk, we hypothesized that $NK8^+$ might play an immunoregulatory role, limiting recurrent T cell driven demyelination in RRMS. To test this, we stimulated $CD4^+$ or $CD8^+$ T cells in the presence of a titrated ratio of autologous $NK8^+$ or $NK8^-$ cells (Supp Fig 4A), while taking care to prevent NK pre-activation (Supp Fig 4B, C). We found that, compared to their $NK8^-$ counterparts, $NK8^+$ cells exerted a significant suppressive effect on autologous $CD4^+$ (Fig 2A-D) but not $CD8^+$ T cell (Fig 2E-H) proliferation and activation, although we cannot exclude that higher NK:T cell ratios or reduced CD8 TCR stimulation levels may allow CD8 T cell suppression to also occur. As no differences were observed in the cytotoxic granule number or content of $NK8^+$ and $NK8^-$ cells (by either flow cytometry or electron microscopy, Supp Fig 5), we asked whether an altered balance of activating and inhibitory signals might explain the superior ability of $NK8^+$ cells to inhibit autologous $CD4^+$ T cell activation. Using a validated flow cytometry panel[22] we quantified expression of NK killer inhibitory and activating receptors (Supp Data 4) on $NK8^+$, $NK8^-$ and $CD56^{hi}$ subsets in PBMC from healthy individuals. While most receptors were equivalently expressed (Supp Fig 6), we observed increased expression on $NK8^+$ of the activating receptor NKG2D, and reduced expression of two inhibitory receptors (ILT2, KIR2DL4, Fig 3A). While KIR2DL4 may serve as an activating or inhibitory receptor[23], both ILT2 and KIR2DL4 share the same ligand, HLA-G, a non-classical HLA molecule first identified as key regulator of foetal-maternal tolerance[24] and now recognised to play a regulatory role in suppressing immunity during infection, transplantation, autoimmunity and cancer[25]. Polymorphic variation at the HLA-G locus dictates expression levels of the receptor, and has been associated with altered susceptibility to multiple autoimmune diseases[26], including MS[27]. We observed that HLA-G was upregulated on $CD4^+$ T cell activation (Supp Fig 7A), and hypothesised that $NK8^+$ may

be relatively refractory to its suppressive effect by virtue of their reduced HLA-G receptor expression. To test this, we investigated the ability of HLA-G to differentially inhibit $NK8^+$ and $NK8^-$ responses in a cytotoxicity assay. Isolated $NK8^+$ or $NK8^-$ cells were co-cultured in the presence of target cell lines with cytotoxicity measured by surface translocation of the granule protein CD107[28]. In each case we observed similar levels of induced cytotoxicity (Supp Fig 7B), but in the presence of a titrated dose range of soluble HLAG (sHLAG), we observed a reduction in cytotoxicity of $NK8^-$ but not $NK8^+$ cells (Fig. 3C). Thus $NK8^+$ cells can play an immunoregulatory role, mediated through suppression of $CD4^+$ T cell activation and proliferation, and appear relatively impervious to the inhibitory effects of HLA-G through limited expression of HLA-G receptors.

$NK8^+$ cells may therefore be capable of regulating proliferation and activation of autologous $CD4^+$ T cells, through a mechanism at least partially dependent on reduced HLA-G mediated NK suppression.

Validating NK8 prediction of clinical outcome

Having identified a potential mechanism for the $NK8^+$ gene signature associated with favourable outcome in STAyCIS, we sought to confirm the finding in an independent cohort of patients with RRMS and CIS. To facilitate this, we used a cohort of patients with anti-neutrophil cytoplasmic antibody associated vasculitis (Supp Data 6) to first identify surrogate markers of the $NK8^+$ signature that could be measured in a mixed peripheral blood mononuclear cell (PBMC) population (Supp Fig 8, Supp Data 5). to stratify a second prospective clinical RRMS study[29] (Supp Data 7) into subgroups enriched or depleted for $NK8^+$ transcripts (Fig 3D, analogous to STAyCIS subgroups shown in Fig 1D). As observed in the STAyCIS trial, favourable clinical outcome (clinical relapse-free survival) was associated with enrichment of surrogate markers of the $NK8^+$ signature (Fig 3D-F) in this dataset derived from PBMC. These data independently validate the finding that an $NK8^+$ expansion is associated with favourable prognosis, supporting its use as a predictive tool to guide early aggressive therapy in RRMS.

Discussion

Our observations are consistent with and extend previous reports of NK mediated control of autoimmunity in RRMS. Active RRMS is known to be characterised by periods of reduced NK cell number and cytotoxicity in both cerebrospinal fluid (CSF) and peripheral blood while expansion of immature $CD56^{bright}$ NK cells accompanied clinical response to immunotherapy[8,30,31]. The comparatively small population of immature $CD56^{hi}$ cells has a well described immunoregulatory role, dependent on NKG2D-mediated regulation of CD4 T cell activation[32]. Our findings indicate that $NK8^+$ cells contribute to regulation of autologous $CD4^+$ T cells, and that HLA-G expression on activated $CD4^+$ T cells may allow escape from $CD8^-$ $CD56^{dim}$ NK-mediated regulation. Therapeutic strategies promoting $NK8^+$ differentiation or altered T cell HLA-G expression may have coordinated regulatory effects on pathogenic T cells to limit demyelination in CIS/RRMS.

While clinical translation will require further validation work, data presented here suggest that measurement of an NK8+ cell associated transcriptional signature could potentially be used to facilitate prediction of clinical outcome to target therapy in RRMS, and guide early treatment decisions.

Methods

The STAyCIS trial and sample collection

The STAyCIS trial[14] is a randomized, double-blind, placebo-controlled, multicentre study evaluating the efficacy and safety of atorvastatin (Lipitor, Pfizer, 80mg/day) in patients with clinically isolated syndrome (a first demyelinating event) and at high risk of conversion to RRMS. Written informed consent was obtained from patients prior to enrollment in the STAyCIS study (NCT00094172). 82 participants were recruited within a screening phase of 90 days from the index CIS event and followed up for 18 months (12 months treatment phase) with serial clinical and radiological (MRI imaging) review. The primary combined endpoint was the development of either radiological (>=3 new T2 MRI lesions) or clinical relapse (>=1 clinical exacerbation) during the 12 month treatment phase. The trial was sponsored by NIAID in collaboration with the Immune Tolerance Network (clinicaltrials.gov NCT00094172). The study was approved by institutional review boards at 14 centers in the United States and Canada. Written informed consent was obtained from patients prior to enrollment in the STAyCIS study (NCT00094172).

Sample Collection, processing and QC

56 STAyCIS PBMC samples for which viable cells were available were isolated from peripheral blood by centrifugation over Histopaque (Invitrogen) before controlled freezing 1°C/min (with isopropyl alcohol, Nalgene) and storage in buffer (10%FBS/DMSO) in liquid nitrogen. Frozen samples were rapidly thawed in the presence of endonucleases (Benzonase, Merck) at 37ºC, stained with anti-CD8 microbeads (Miltenyi) and enriched using a MACS column (Miltenyi).

Transcriptomic data generation and QC

Aliquots of total RNA (200ng) were labelled using Ambion WT sense Target labelling kit and hybridised to Human Gene 1.0 Arrays (Affymetrix) following the manufacturer's instructions. After washing, arrays were scanned using a GS 3000 scanner (Affymetrix) and CEL files were imported into RStudio (version 3.5.1) for QC and analysis. Affymetrix raw data (.CEL) files were imported into R and subjected to variance stabilisation normalisation using the VSN package in BioConductor[33]. Quality control was performed using the Bioconductor package arrayQualityMetrics[34] with correction for batch variation performed using the Bioconductor package ComBat[35]. Differential expression was conducted using the Bioconductor package limma in R[36]. Samples identified as outliers during QC filtering were excluded resulting in 44 samples being taken forward for further network analyses (Supplementary Figure2). To generate a transcriptional signature specific to NK8+ cells (Supplemental Data 3), 50ml peripheral whole blood was collected from 5 healthy volunteers from which

PBMC were isolated by centrifugation over Histopaque (Invitrogen). $NK8^+$ ($CD3^-CD56^+CD8^+$), $NK8^-$ ($CD3^-CD56^+CD8^-$) and $CD3^+8^+$ T cells were isolated by flow sorting following staining with anti-CD56 (Clone REA196), anti-CD8a Fab (Clone BW135/80 Fab) and anti-CD3 REA (Clone REA641) antibodies. Cells were sorted into RNAlysis medium (Tempus Blood RNA Tubes, ThermoFisher) and 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. cDNA libraries were generated using the SMARTer stranded total RNAseq kit (Takara) with RNAsequencing performed on an Illumina HiSeq4000 instrument following the manufacturer's instructions. Sequences were processed, cleaned and aligned using a combination of bbsplit, hisat2 and trimgalore with count data read into R using the subRead package (Bioconductor). Normalization and transformation were undertaken using edgeR and voom packages for Bioconductor in R with pairwise differential expression analysis (FDR threshold 5%) using edgeR. Data QC was ensured using the arrayQualitymetrics (microarray) and fastqc, qorts and pcaexplorer (RNAseq) packages for Bioconductor in R.

WGCNA Analysis

Unsupervised identification of co-correlated genes was undertaken using the Weighted Gene Coexpression Network Analysis Bioconductor package in R[16]. Normalised, log-transformed expression data was variance filtered using the median absolute deviation and a soft thresholding power was chosen based on the criterion of approximate scale-free topology[37]. 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 with a specified minimum module size of n=30. Modules were summarised as a network of modular eigengenes which were then correlated with a matrix of clinical variables (Supplementary Data 1, Supplementary Figure 2). 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. Modular signatures were compared to clinical laboratory measures, supporting clinical data and the time to the primary and secondary endpoints of the STAyCIS study (Supplementary Figure 2, Figure 1). Overlap of signatures with modules derived from network analysis (Supplementary Figure 1F) used a proportional representation method defined by the following formula to allow correction for variable module size: (signature genes in module, n)/(genes in module, n) x100. Hierarchical clustering was performed using a Pearson correlation distance metric and average linkage analysis performed using Genepattern[38].

Enrichment analyses

Modular signatures were compared against cell and tissue-specific signatures as defined in the Human Gene Atlas (HGA, GSE1133) and Immune Response In Silico (IRIS, GSE22886) datasets. In brief these datasets have defined cell and/or tissue-specific patterns of gene expression comparing across data from 79 human/61 murine tissues (HGA) and 22 immune cell subsets (IRIS) respectively. Outcome

associated modular signatures were compared to all cell/tissue types to confirm specificity and radarplots shown (Figure 1C) include comparison against all tissues with at least 1 overlapping transcript. Enrichment analyses were performed using Enrichr[17] or fgsea (Bioconductor) with a significance threshold of 5% FDR.

Functional NK Assays

Primary human NK cells and $CD4^+$ and $CD8^+$ T cells were separated from leucocyte cones obtained from NHS Blood and Transplant (Addenbrooke's Hospital, Cambridge, UK) by centrifugation over ficoll and either positive selection using magnetic beads (MACS, Miltenyi) according to the manufacturers instructions or by flow cytometric sorting (NK cell subsets $CD3^-56^+$ $CD8^{+/-}$, Supp Fig 9), using anti-CD56 (Clone REA196), anti-CD8a Fab (Clone BW135/80 Fab)and anti-CD3 REA (Clone REA641). Purity of separated cell subsets was determined by three-colour flow cytometry and purified T cells were labelled with 10μM CFSE (Invitrogen) and resuspended in complete RPMI 1640 (Sigma Aldrich) in the presence of 10% FCS. Purified $CD4^+$ and $CD8^+$ T cells (>95%) were then stimulated in sterile, U-bottomed culture plates (Greiner) using microbead particles conjugated to anti-CD2/3/28 antibodies (1:2 bead:cell ratio, Miltenyi) in the presence of IL2 (10ng/ml, Gibco life technologies) for 6 days. Natural killer cells were cultured in the presence of autologous stimulated $CD4^+$ T cells, $CD8^+$ T cells or target K562 or Jurkat T cells as indicated for 6 days with or without titrated range (0.1-0.2ug/ml[39]) of soluble HLA-G (Abcam) before analysis. Flow cytometry analysis of cultured cells comprised panels including Live/Dead Blue reactive dye (Life Technologies), anti-CD56 (Clone REA196), anti-CD8a Fab (Clone BW135/80 Fab), anti-CD3 REA (Clone REA641), anti-CD107 (Biolegend, Clone H4A3, used following Brefeldin A protein transport blockade, eBiosciences), anti-HLADR (Clone L243), Granzyme B (Biolegend, clone GB11), Perforin (eBioscience, eBioOMAK-D) and anti-HLAG (Biolegend, clone 87G) as indicated.
For NK coculture with autologous T cells, it was necessary to account for variable simulation bead:target cell ratio occurring due to titration of increasing numbers of NK cells into culture. Consequently, the ratio of NK activation (CD107 MFI) seen with $NK8^+$ cells was compared to that seen on coculture with a comparable number of $NK8^-$ cells (Figure3).

Flow cytometric immunophenotyping

Heparinized peripheral blood samples were obtained for analysis of lymphocyte populations and subpopulations by flow cytometry. Whole blood was collected in sodium heparin vacutainers (Becton Dickinson) and shipped ambient overnight to the ITN Flow Cytometry Core (Roswell Park Cancer Institute). Using a stain-lyse method, cells from blinded samples were labeled with 5-color monoclonal antibody panels using anti-human CD8-PE-Cy5, CD57-FITC (clone NK-1), CD56-PE (clone NCA-1) CD14-APC (clone MϕP9), plus CD3-PE-Cy7 (clone SK7, all BD Biosciences). Following staining, data were acquired on a FACSCanto flow cytometer (BD Biosciences), and analyzed using WinList's™ (http://www.vsh.com) FCOM function[40].
A validated panel of KIR receptors[22] was used to quantify NK cell subset expression of activating and inhibitory receptors for which clone ids and reagent sources are

detailed in Supplementary Data 4. Flow analysis was undertaken using an LSR Fortessa (BD) in the NIHR Cambridge BRC flow phenotyping hub (Supp Fig 9).

Electron microscopy

Transmission electron microscopy was undertaken on purified, pooled NK cell subsets sorted as above and processed at the Cambridge Advanced Imaging Centre. Lytic granules were counted and cell size quantified in at least 10 images from each of $NK8^+$ and $NK8^-$ subsets from each of 3 biological replicates (Supplementary Figure 5).

Surrogate NK signature marker identification and validation

Optimal surrogate markers for identification of the $NK8^+$ signature in PBMC-level data was determined using a randomforests classification algorithm[41] (Supplementary Figure 8). The $NK8^+$ signature itself cannot be directly used in a mixed cell population due to the confounding influence of transcripts from other cell types[42,43]. Expression data derived from both MACS-purified $CD8^+$ cells and PBMC were available for a cohort of n=47 patients anti-neutrophil cytoplasmic antibody associated vasculitis (AAV) following QC and hybridisation to the HsMediante25k custom microarray platform and constituted a training cohort. Normalised, log- transformed expression data was analysed using the ML Interfaces Bioconductor package in R[44]. Using $CD8^+$-level expression data, AAV samples were classified into subgroups showing either high or low expression of the $NK8^+$ signature (as illustrated in Supplementary Figure 8). Subsequently, PBMC-level data from the same AAV samples was used to identify and rank probes for their ability to discriminate the $NK8^+$-defined subgroups using the variable Importance metric, reflecting the change in accuracy of classification (% change in Gini coefficient, Supplementary Data 5) when that variable is randomly permuted. Optimal probes were then used to identify analogous patient groups in an independent validation cohort of PBMC data from RRMS/CIS patients (GSE15245[29]). For the validation set, data was downloaded from GEO and imported into R using the Bioconductor package GEOquery[45] in R. The validation study represented a prospective collection of 94 PBMC samples taken from CIS/RRMS cases (n=32/62) and followed up to 3.5 years recording clinical relapses (Supplementary Data 7). Microarray data was generated using Hu133A[29].

Data Availability

The microarray data generated during and/or analysed during the current study are available in the GEO repository accession number E-MTAB-9637 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9637/).

Source data are provided with this paper.

References

1. McKinney, E.F., Lee, J.C., Jayne, D.R.W., Lyons, P.A. & Smith, K.G.C. T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection. *Nature* **523**, 612-616 (2015).

2. Ransohoff, R.M., Hafler, D.A. & Lucchinetti, C.F. Multiple sclerosis-a quiet revolution. *Nat Rev Neurol* **11**, 134-142 (2015).
3. Bsteh, G.*, et al.* Long Term Clinical Prognostic Factors in Relapsing-Remitting Multiple Sclerosis: Insights from a 10-Year Observational Study. *PLoS One* **11**, e0158978 (2016).
4. Dobson, R. & Giovannoni, G. Multiple sclerosis - a review. *Eur J Neurol* **26**, 27-40 (2019).
5. Dendrou, C.A., Fugger, L. & Friese, M.A. Immunopathology of multiple sclerosis. *Nat Rev Immunol* **15**, 545-558 (2015).
6. Abel, A.M., Yang, C., Thakar, M.S. & Malarkannan, S. Natural Killer Cells: Development, Maturation, and Clinical Utilization. *Front Immunol* **9**, 1869 (2018).
7. Biron, C.A., Nguyen, K.B., Pien, G.C., Cousens, L.P. & Salazar-Mather, T.P. Natural killer cells in antiviral defense: function and regulation by innate cytokines. *Annu Rev Immunol* **17**, 189-220 (1999).
8. Gross, C.C.*, et al.* Impaired NK-mediated regulation of T-cell activity in multiple sclerosis is reconstituted by IL-2 receptor modulation. *Proc Natl Acad Sci U S A* **113**, E2973-2982 (2016).
9. Waggoner, S.N., Cornberg, M., Selin, L.K. & Welsh, R.M. Natural killer cells act as rheostats modulating antiviral T cells. *Nature* **481**, 394-398 (2012).
10. Dalbeth, N. & Callan, M.F. A subset of natural killer cells is greatly expanded within inflamed joints. *Arthritis Rheum* **46**, 1763-1772 (2002).
11. Rodacki, M.*, et al.* Altered natural killer cells in type 1 diabetic patients. *Diabetes* **56**, 177-185 (2007).
12. Bielekova, B.*, et al.* Regulatory CD56(bright) natural killer cells mediate immunomodulatory effects of IL-2Ralpha-targeted therapy (daclizumab) in multiple sclerosis. *Proc Natl Acad Sci U S A* **103**, 5941-5946 (2006).
13. Poggi, A. & Zocchi, M.R. NK cell autoreactivity and autoimmune diseases. *Front Immunol* **5**, 27 (2014).
14. Waubant, E.*, et al.* Randomized controlled trial of atorvastatin in clinically isolated syndrome: the STAyCIS study. *Neurology* **78**, 1171-1178 (2012).
15. Libbrecht, M.W. & Noble, W.S. Machine learning applications in genetics and genomics. *Nat Rev Genet* **16**, 321-332 (2015).
16. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
17. Chen, E.Y.*, et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
18. Abbas, A.R.*, et al.* Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* **6**, 319-331 (2005).
19. Su, A.I.*, et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* **101**, 6062-6067 (2004).
20. Ahmad, F.*, et al.* High frequencies of polyfunctional CD8+ NK cells in chronic HIV-1 infection are associated with slower disease progression. *J Virol* **88**, 12397-12408 (2014).

21. Raulet, D.H. & Vance, R.E. Self-tolerance of natural killer cells. *Nat Rev Immunol* **6**, 520-531 (2006).
22. Czaja, K.*, et al.* A comprehensive analysis of the binding of anti-KIR antibodies to activating KIRs. *Genes Immun* **15**, 33-37 (2014).
23. Campbell, K.S. & Purdy, A.K. Structure/function of human killer cell immunoglobulin-like receptors: lessons from polymorphisms, evolution, crystal structures and mutations. *Immunology* **132**, 315-325 (2011).
24. Favier, B., LeMaoult, J. & Carosella, E.D. Functions of HLA-G in the immune system. *Tissue Antigens* **69 Suppl 1**, 150-152 (2007).
25. Rouas-Freiss, N., Moreau, P., Ferrone, S. & Carosella, E.D. HLA-G proteins in cancer: do they provide tumor cells with an escape mechanism? *Cancer Res* **65**, 10139-10144 (2005).
26. Rizzo, R., Bortolotti, D., Bolzani, S. & Fainardi, E. HLA-G Molecules in Autoimmune Diseases and Infections. *Front Immunol* **5**, 592 (2014).
27. Cree, B.A.*, et al.* A major histocompatibility Class I locus contributes to multiple sclerosis susceptibility independently from HLA-DRB1*15:01. *PLoS One* **5**, e11296 (2010).
28. Bryceson, Y.T.*, et al.* Functional analysis of human NK cells by flow cytometry. *Methods Mol Biol* **612**, 335-352 (2010).
29. Gurevich, M., Tuller, T., Rubinstein, U., Or-Bach, R. & Achiron, A. Prediction of acute multiple sclerosis relapses by transcription levels of peripheral blood cells. *BMC Med Genomics* **2**, 46 (2009).
30. Bielekova, B.*, et al.* Humanized anti-CD25 (daclizumab) inhibits disease activity in multiple sclerosis patients failing to respond to interferon beta. *Proc Natl Acad Sci U S A* **101**, 8705-8708 (2004).
31. Kastrukoff, L.F.*, et al.* Clinical relapses of multiple sclerosis are associated with 'novel' valleys in natural killer cell functional activity. *J Neuroimmunol* **145**, 103-114 (2003).
32. Nielsen, N., Odum, N., Urso, B., Lanier, L.L. & Spee, P. Cytotoxicity of CD56(bright) NK cells towards autologous activated CD4+ T cells is mediated through NKG2D, LFA-1 and TRAIL and dampened via CD94/NKG2A. *PLoS One* **7**, e31959 (2012).
33. 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* **18 Suppl 1**, S96-104 (2002).
34. Kauffmann, A., Gentleman, R. & Huber, W. arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. *Bioinformatics* **25**, 415-416 (2009).
35. Johnson, W.E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118-127 (2007).
36. Ritchie, M.E.*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
37. Barabasi, A.L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509-512 (1999).
38. Renshaw, B.R.*, et al.* Humoral immune responses in CD40 ligand-deficient mice. *J Exp Med* **180**, 1889-1900 (1994).

39. Marchal-Bras-Goncalves, R., *et al.* A soluble HLA-G protein that inhibits natural killer cell-mediated cytotoxicity. *Transplant Proc* **33**, 2355-2359 (2001).
40. Siebert, J.C., *et al.* Exhaustive expansion: A novel technique for analyzing complex data generated by higher-order polychromatic flow cytometry experiments. *J Transl Med* **8**, 106 (2010).
41. Breiman, L. Random Forests. *Machine Learning Journal* **45**, 5-32 (2001).
42. Lyons, P.A., *et al.* Microarray analysis of human leucocyte subsets: the advantages of positive selection and rapid purification. *BMC Genomics* **8**, 64 (2007).
43. McKinney, E.F., *et al.* A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. *Nat Med* **16**, 586-591 (2010).
44. Vince Carey, R.G., Jess Mar, and contributions from Jason Vertrees and Laurent Gatto. MLInterfaces: Uniform interfaces to R machine learning procedures for data in Bioconductor containers. *R package* **version 1.40.0**.
45. Davis, S. & Meltzer, P.S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846-1847 (2007).

Author Contributions

EFM designed the experiments and conducted them with the help of IC, CC, GM and EJC. Analysis was undertaken by EFM with review by KH, DS, SZ, GN and KGCS. The STAyCIS trial was conducted and samples were collected by the ITN STAyCIS study group with SZ as principal investigator. The manuscript was written by EFM along with IC, GM, KH, DS, SZ, GN and KGCS.

Acknowledgements

This work is supported by National Institute of Health Research (NIHR) Cambridge Biomedical Research Centre (BRC) and funded by a Medical Research Council (UK) programme award (MR/L019027). E.M. was a Wellcome –Beit Research Fellow supported by the Wellcome Trust and Beit Foundation (104064/Z/14/Z). K.G.C.S was a Lister Prize Fellow. We thank Anna Petrunkina-Harrison and Natalia Savinykh at the NIHR Cambridge Biomedical Research Cell Phenotyping Hub for extensive assistance. The authors also thank their colleagues at the Immune Tolerance Network, and their collaborators who contribute in many capacities to Immune Tolerance Network projects and perspectives. The authors are grateful to the ITN020AI STAyCIS study participants, and thank the clinical site investigators and study coordinators. Research reported in this publication was supported by the Immune Tolerance Network and sponsored by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number UM1AI109565. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the NIHR or the Department of Health. The STAyCIS clinical trial was performed as a project of the Immune Tolerance Network (ITN020AI, ITN contract number N01-AI-15416) within a clinical research consortium sponsored by the National Institute of Allergy and Infectious Diseases (NIAID).

Competing interests

The authors declare no competing interests.

Figure 1: An NK8 signature associates with relapse risk in multiple sclerosis

(A) Heatmap showing correlation of gene expression eigengenes (coloured blocks, y-axis) with clinical endpoints of the STAyCIS trial (Pearson correlation, colourbar). Phase of event = event occurring during 12m treatment phase or extended follow-up to 18m. Event leading to primary endpoint (PEP) classified as clinical or radiological. (B) Volcano plot showing modules with significant association with PEP (x = Pearson correlation of module eigengene and time to event, y = $-\log_{10}$FDR). (C) Radar plots showing enrichment ($-\log_{10}$FDR) of outcome-associated modular signatures (from panels A, B) compared to immune cell-specific signatures from the Human gene atlas (HGA, upper) and heatmaps showing 'black' module eigengene expression across all cell types in HGA and immune response in silico (IRIS) repositories. Each point on the circumference reflects a different set of cell-specific transcripts including all (of 79 total) showing any overlap with modular signatures. (D) Heatmap (upper panel, patients in columns, genes in rows red:blue =max:min expression) and dotplot (lower panel) showing unsupervised hierarchical clustering using the 'black' module (panel A) and associated time to PEP of identified patient subgroups. (E) Volcano plot showing enrichment of NK-associated signatures from the IRIS dataset in patient subgroups identified in (D). (F) Scatterplot illustrating correlation of the black module eigengene (x-axis) against %$CD56^+CD8^+$ cells (G) Enrichment of an $NK8^+$-specific signature, obtained by flow sorting/RNAseq of $CD8^+$ v $CD8^-$ NK cells from 5 healthy individuals in $CD8^+$ leukocytes from STAyCIS participants with NK8hi/lo signature at baseline as illustrated in (C). (H) Scatterplot (top) and barplot (bottom) illustrating cross-validation and feature-specific weighted coefficients respectively in a penalised Cox proportional hazards model using a naïve elastic net. Scatterplot shows cross-validated error (deviance +/- sem, y-axis) against increasing regularisation penalty ($\log\lambda$, x-axis). PEP = primary endpoint, FDR = false discovery rate, sem = standard error of the mean, FDR threshold on radarplots = 0.05 (red dashed line). Phase of event = event occurring during 12m treatment phase or extended follow-up to 18m. CIS = clinically isolated syndrome, PASAT = Paced Auditory Serial Addition Test, MSFC = Multiple Sclerosis Functional Composite. Error bars = mean +/-sem.

Figure 2: NK8 cells regulate autologous $CD4^+$ T cell activation

Representative histograms (A,E), contour plots (C,G) and summary scatter/line plots (B,D,F,H) showing flow cytometric quantification of $CD4^+$ (A-D) and $CD8^+$ (E-H) T cell proliferation (A,B,E,F, CFSE dilution) and activation (HLADR expression, C,D,G,H) when unstimulated (inset, black) or following polyclonal stimulation in vitro (anti-CD2/3/28 bead) in the presence of a titrated ratio of autologous $NK8^+$ (red) or $NK8^-$ (blue) cells as indicated. Summary $NK8^+$ results (B D, F, H) expressed as ratio relative to paired $NK8^-$ cell-enriched culture. P = 2 way ANOVA with NK:T ratio as the categorical variable. For B, D, F and H n=5 independent biological replicates per group. Error bars = mean +/-sem, red dashed line = no difference v control.

Figure 3: NK8 and sensitivity to HLA-G mediated suppression.
(A) Representative histograms (left) and scatterplots (right) of inhibitory HLA-G binding (KIR2DL4, ILT2) and activating (NKG2D) NK receptors on $NK8^+$ and $NK8^-$ subsets, * = P=0.02, Wilcoxon paired signed rank two-tailed test, n=7 independent biological replicates per group; error bars = mean +/- sem. (B) Representative histograms illustrating NK activation (CD107 expression, x-axis) on NK subset ($NK8^-$ top, $NK8^+$ bottom) on coculture with target cell lines (K562/Jurkat, columns) in the presence (red/blue) or absence (black) of sHLAG at 0.1ug/ml (C) Dot and line plot of CD107 expression (y-axis, MFI ratio v vehicle control) on coculture of $NK8^+$ (red) or $NK8^-$ (blue) with target cell lines (K562/ Jurkat) in the presence of a titrated range of soluble HLAG (x-axis). n=6 per group, P= 2 way ANOVA. (D) Heatmap showing unsupervised hierarchical clustering of 94 CIS/RRMS cases (GSE32915) by gene expression (red:blue, max:min) of an optimised panel of surrogate markers reflecting the $NK8^+$ signature in PBMC samples (Supp Fig 8). (E) Scatterplot showing bidirectional enrichment (GSEAp < 0.05) of optimal $NK8^+$ surrogates in patient subgroups identified in (D). (F) Kaplan-Meier plot showing censored relapse-free survival (y-axis, %) of patient subgroups defined in (D), P = log rank test. Error bars = mean+/- sem.

Figure 1

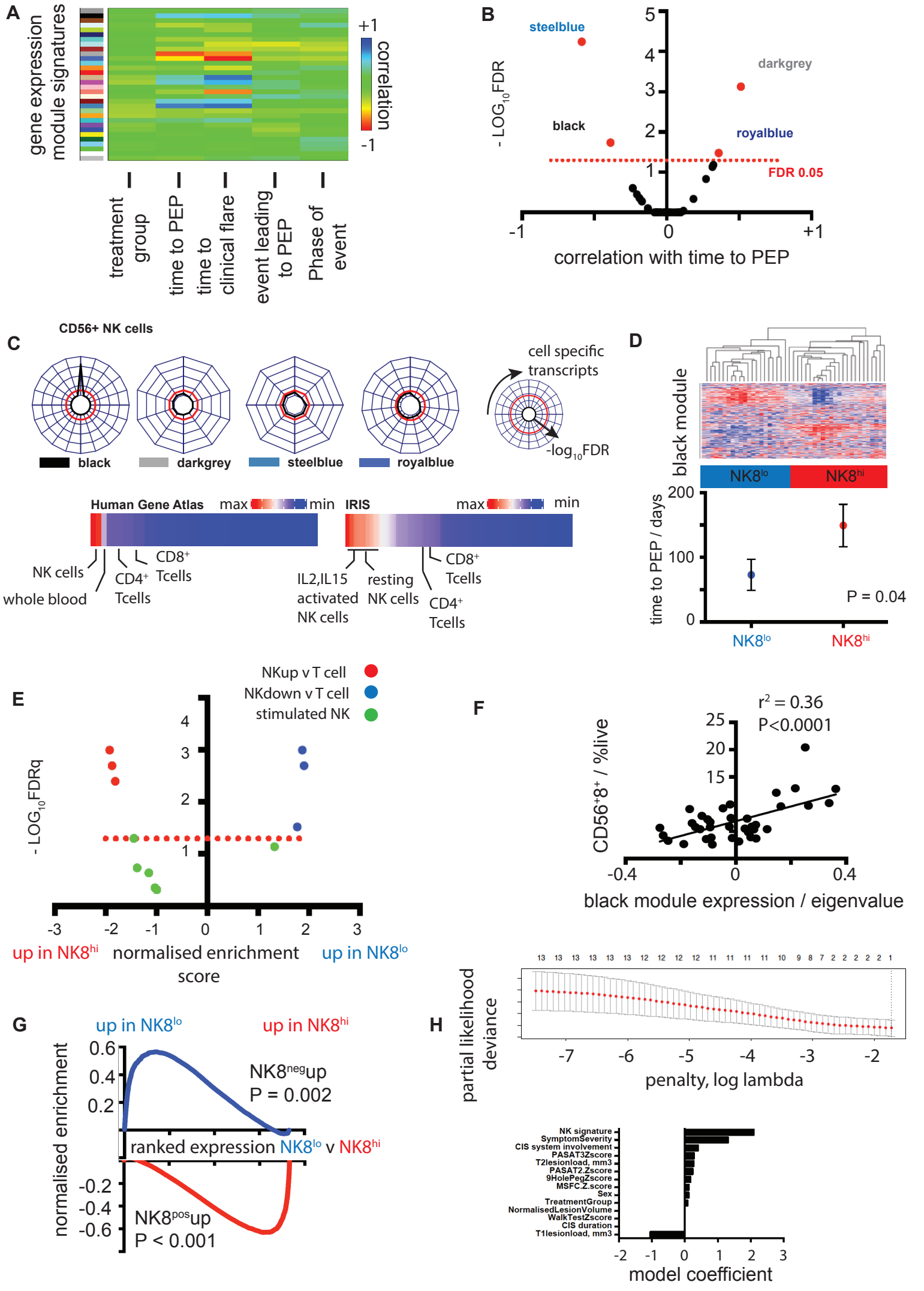

A
gene expression module signatures
+1
correlation
-1
treatment group
time to PEP
time to clinical flare
event leading to PEP
Phase of event
B
steelblue
darkgrey
black
royalblue
FDR 0.05
- LOG10FDR
correlation with time to PEP
C
CD56+ NK cells
black
darkgrey
steelblue
royalblue
cell specific transcripts
-log10FDR
Human Gene Atlas
max
min
IRIS
NK cells
whole blood
CD4+ Tcells
CD8+ Tcells
IL2,IL15 activated NK cells
resting NK cells
D
black module
NK8lo
NK8hi
time to PEP / days
P = 0.04
E
NKup v T cell
NKdown v T cell
stimulated NK
- LOG10FDRq
up in NK8hi
normalised enrichment score
up in NK8lo
F
r2 = 0.36
P<0.0001
CD56+8+ / %live
black module expression / eigenvalue
G
up in NK8lo
up in NK8hi
normalised enrichment
NK8negup
P = 0.002
ranked expression NK8lo v NK8hi
NK8posup
P < 0.001
H
partial likelihood deviance
penalty, log lambda
NK signature
SymptomSeverity
CIS system involvement
PASAT3Zscore
T2lesionload, mm3
PASAT2.Zscore
9HolePegZscore
MSFC.Z.score
Sex
TreatmentGroup
NormalisedLesionVolume
WalkTestZscore
CIS duration
T1lesionload, mm3
model coefficient

Figure 2

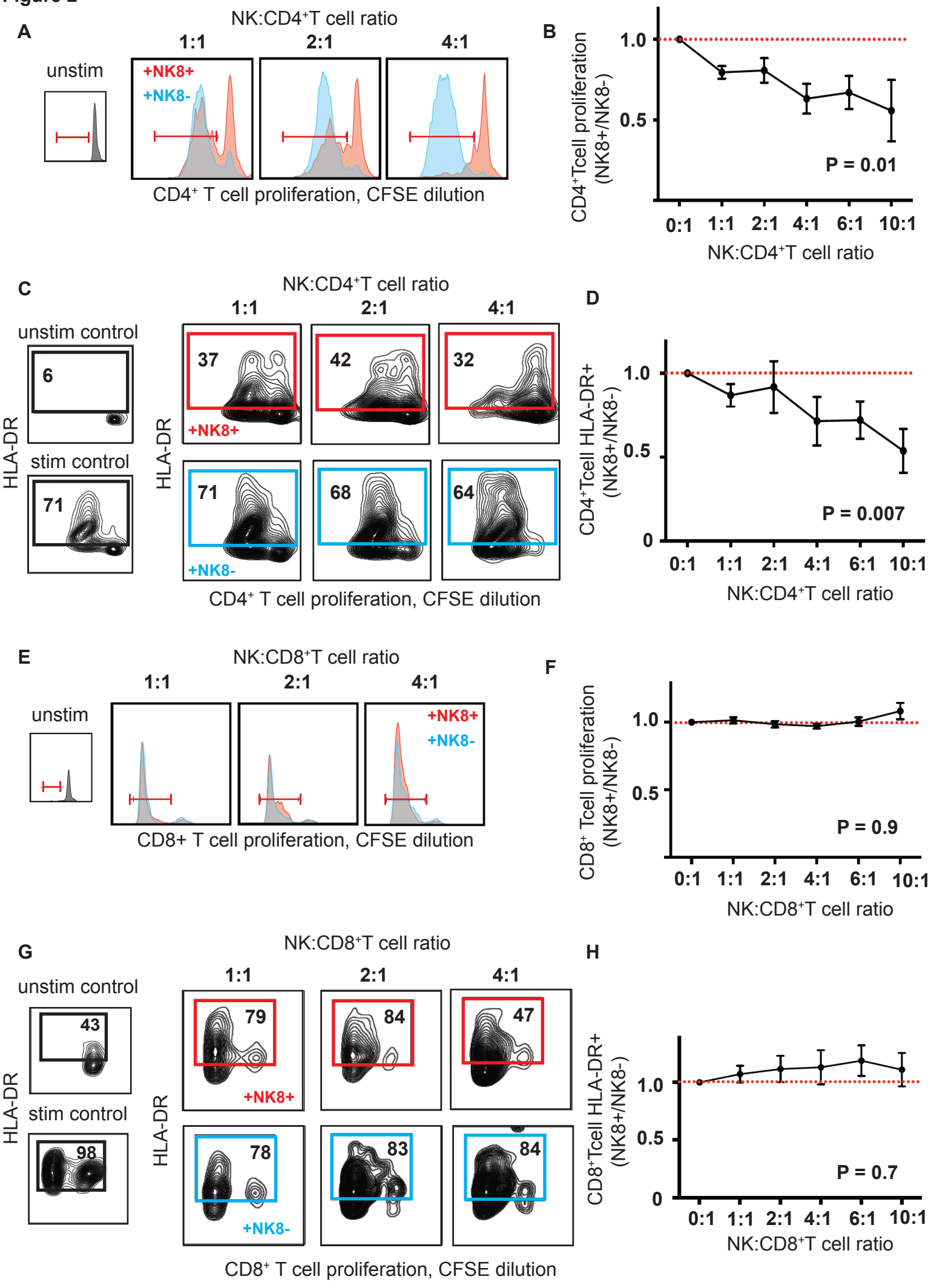
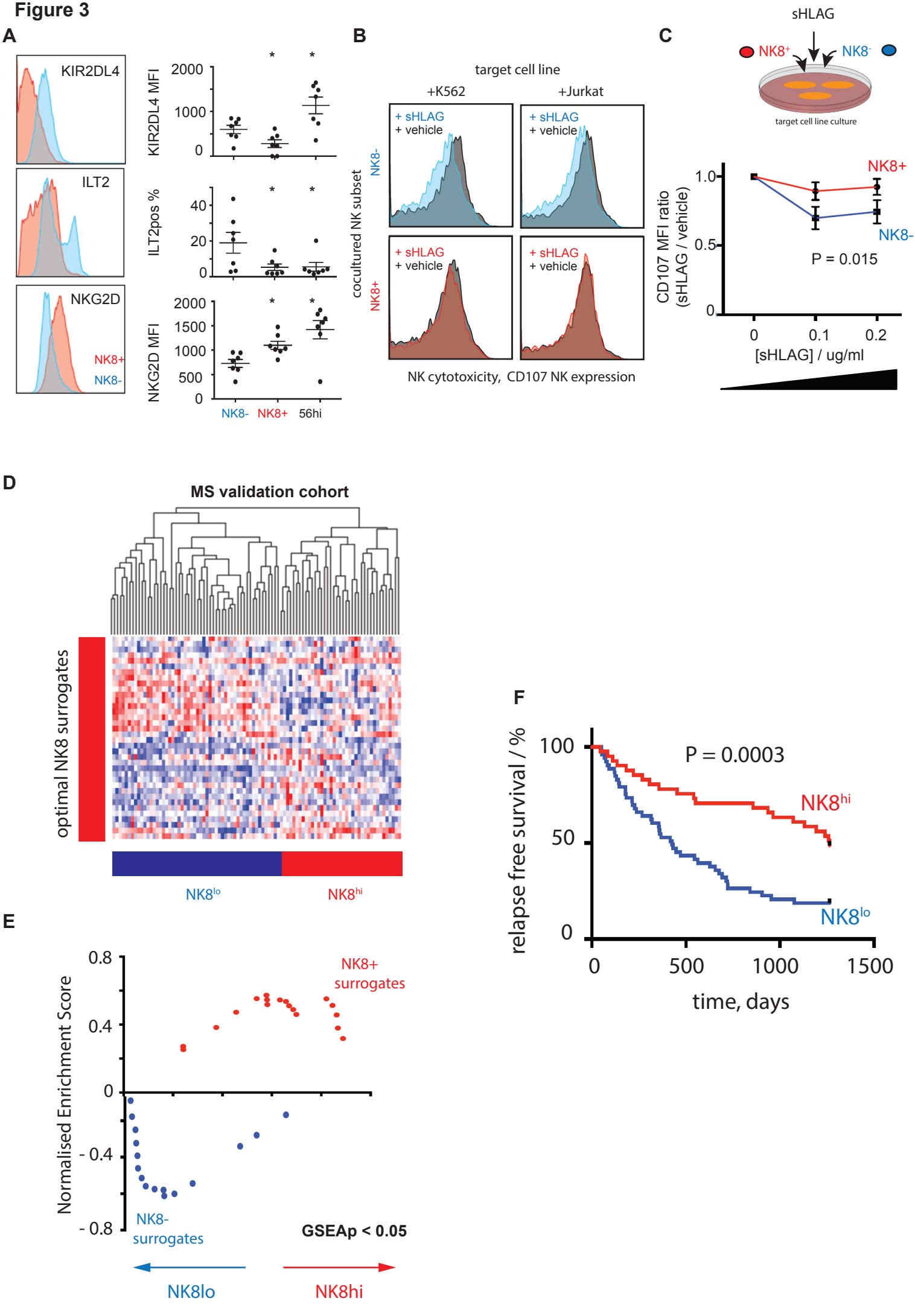
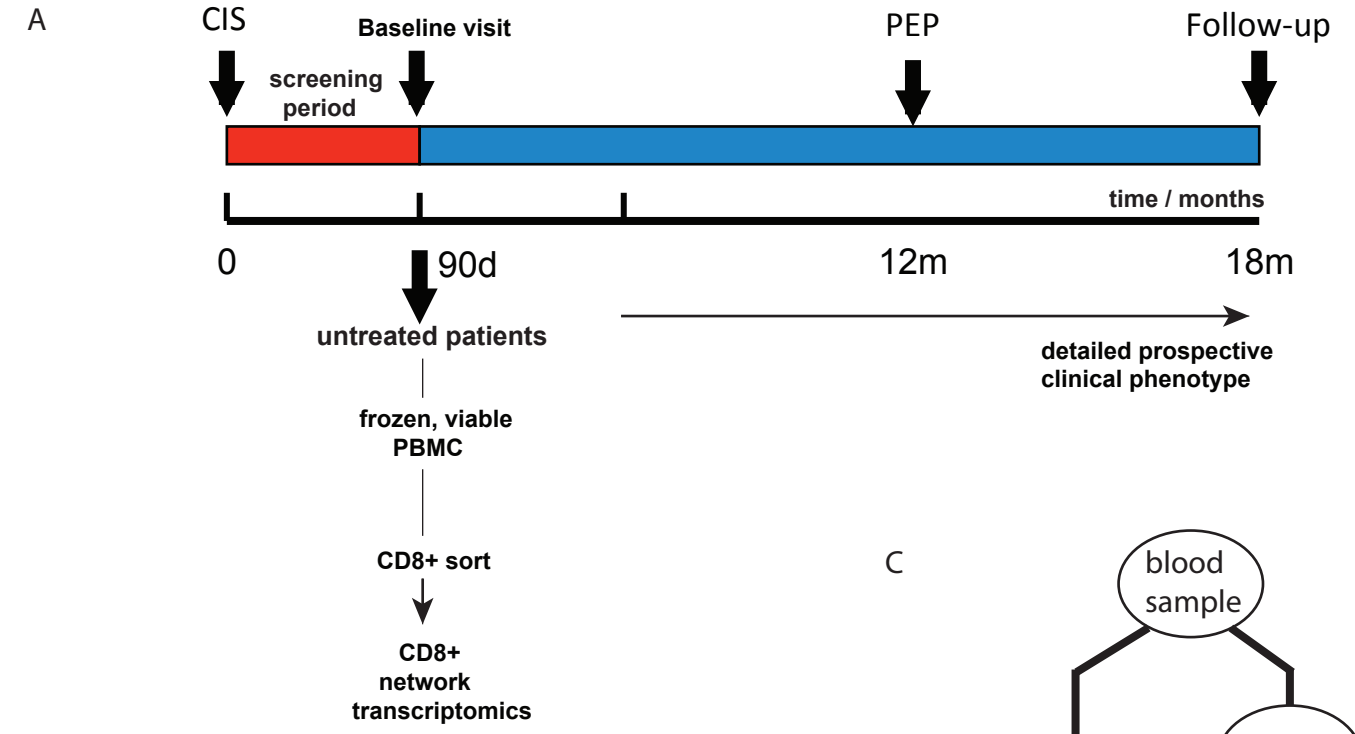
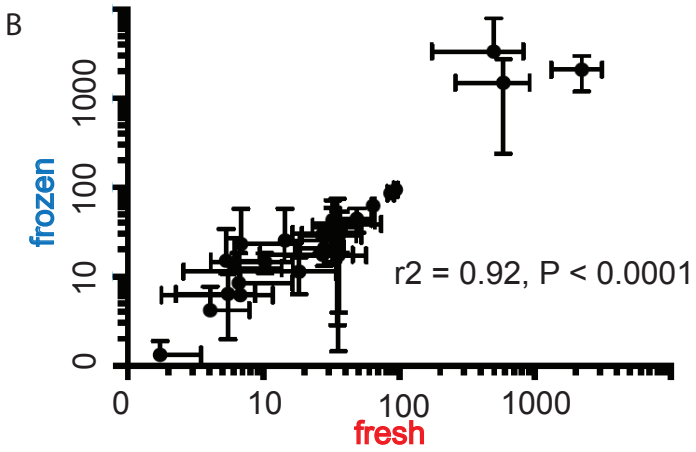

A
NK:CD4+T cell ratio
1:1
2:1
4:1
unstim
+NK8+
+NK8-
CD4+ T cell proliferation, CFSE dilution
B
CD4+Tcell proliferation (NK8+/NK8-)
1.0
0.5
P = 0.01
0:1
1:1
2:1
4:1
6:1
10:1
NK:CD4+T cell ratio
C
NK:CD4+T cell ratio
1:1
2:1
4:1
unstim control
6
stim control
71
HLA-DR
37
42
32
+NK8+
71
68
64
+NK8-
CD4+ T cell proliferation, CFSE dilution
D
CD4+Tcell HLA-DR+ (NK8+/NK8-)
1.0
0.5
0
P = 0.007
0:1
1:1
2:1
4:1
6:1
10:1
NK:CD4+T cell ratio
E
NK:CD8+T cell ratio
1:1
2:1
4:1
unstim
+NK8+
+NK8-
CD8+ T cell proliferation, CFSE dilution
F
CD8+ Tcell proliferation (NK8+/NK8-)
1.0
0.5
P = 0.9
0:1
1:1
2:1
4:1
6:1
10:1
NK:CD8+T cell ratio
G
NK:CD8+T cell ratio
1:1
2:1
4:1
unstim control
43
stim control
98
HLA-DR
79
84
47
+NK8+
78
83
84
+NK8-
CD8+ T cell proliferation, CFSE dilution
H
CD8+Tcell HLA-DR+ (NK8+/NK8-)
1.0
0.5
0
P = 0.7
0:1
1:1
2:1
4:1
6:1
10:1
NK:CD8+T cell ratio

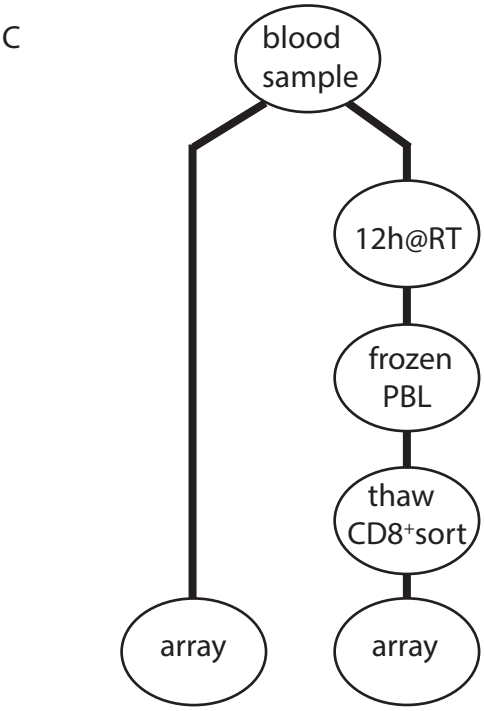
Figure 3

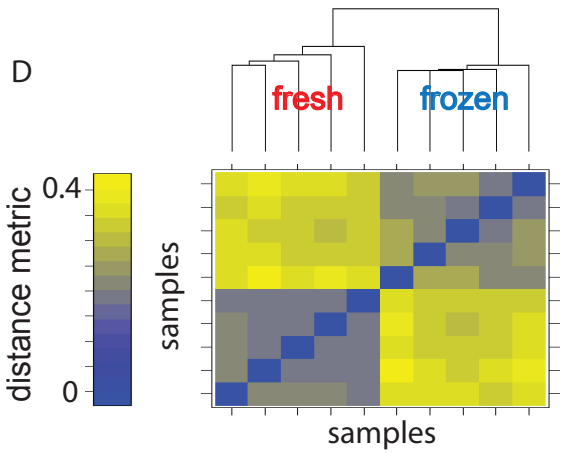

A
KIR2DL4
ILT2
NKG2D
NK8+
NK8-
KIR2DL4 MFI
ILT2pos %
NKG2D MFI
NK8-
NK8+
56hi
B
target cell line
+K562
+Jurkat
+ sHLAG
+ vehicle
cocultured NK subset
NK8-
NK8+
NK cytotoxicity, CD107 NK expression
C
sHLAG
NK8+
NK8-
target cell line culture
CD107 MFI ratio
(sHLAG / vehicle)
P = 0.015
[sHLAG] / ug/ml
D
MS validation cohort
optimal NK8 surrogates
NK8lo
NK8hi
E
Normalised Enrichment Score
NK8+
surrogates
NK8-
surrogates
GSEAp < 0.05
NK8lo
NK8hi
F
relapse free survival / %
P = 0.0003
NK8hi
NK8lo
time, days

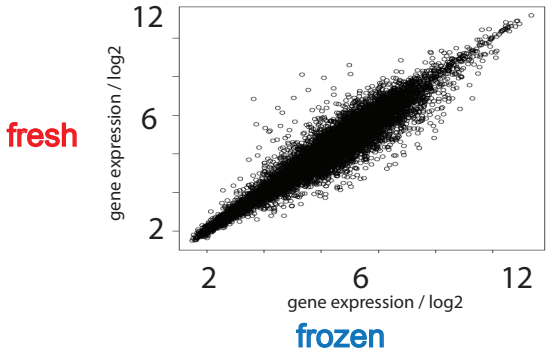
Supplementary Figure1

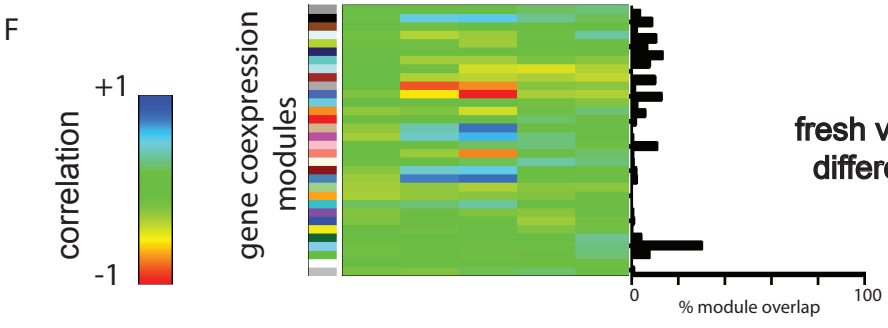






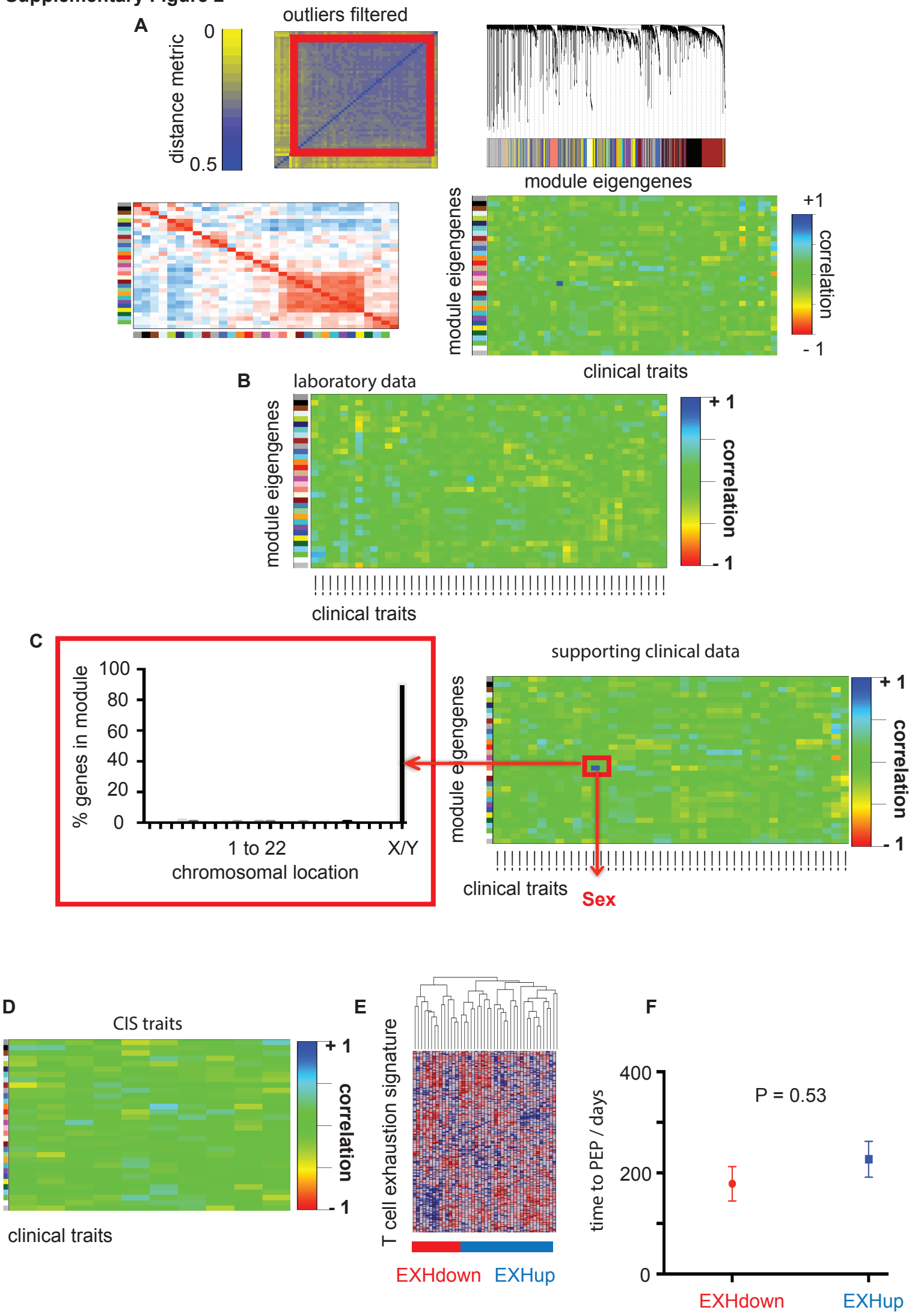






Supplementary Fig 1: Overview of the STAyCIS trial and sample processing
(A) Schematic timeline (days, d; months, m) illustrating STAyCIS trial endpoints and design. CIS = clinically isolated syndrome/1st demyelinating event, PEP = primary endpoint. Frozen, viable lymphocytes were collected as a PBMC (peripheral blood mononuclear cell) fraction at baseline enrolment (within 90 days of CIS) and frozen in liquid nitrogen. On thawing, a CD8^{+} subfraction was isolated by magnetic bead enrichment and analysed by microarray as described, comparing to detailed longitudinal clinical phenotyping (details available at www.itntrialshare.org). (B) Scatterplot showing fresh v frozen immunophenotyping, including major cell subset proportions and selected surface marker MFI for 27 immune cell subsets and traits when processed fresh (x-axis) or following freeze/thaw cycling (y-axis), values represent mean +/- SEM, n=5. Spearman correlation (r2) from linear regression with two-tailed P-value (C) Schematic outline of test protocol for fresh/frozen transcriptomic comparison. (D) Heatmap of distance matrix illustrating distinction between samples processed fresh or frozen as in C. (E) Scatterplot illustrating transcriptional differences between CD8^{+} enriched samples processed fresh (y-axis) and frozen (x-axis) as in (C), prior to transcriptomic analysis. (F) Heatmap illustrating mapping of fresh v frozen differentials as in (E) onto the transcriptomic network used (as in Fig 1A) to associate with clinical traits in the STAyCIS study.

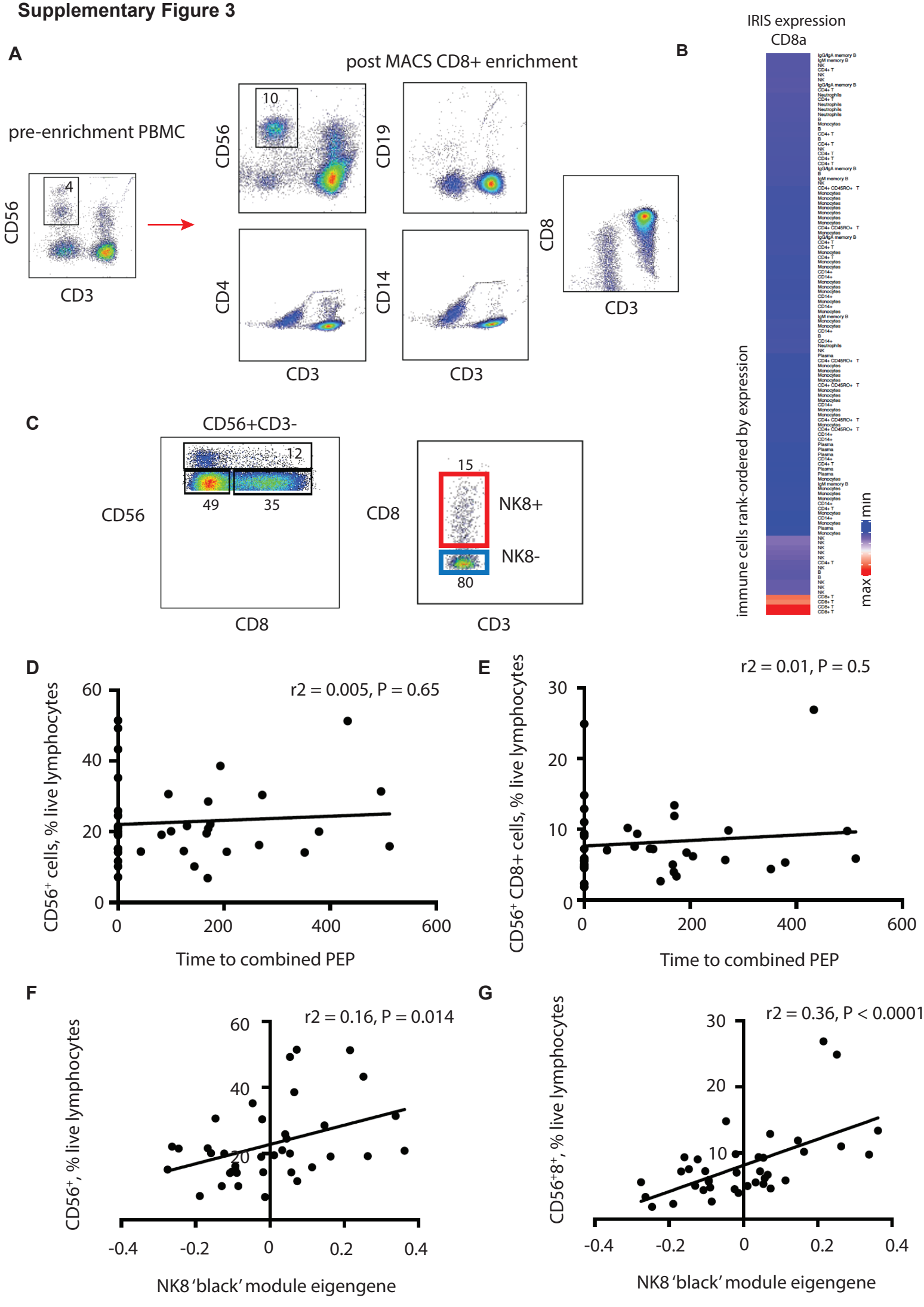
Supplementary Figure 2



Supp Fig 2: Network analysis of STAyCIS and accessory data

(A) Schematic illustration of weighted gene coexpression network analysis with the STAyCIS dataset including QC filtering (top left), module identification and correlation with clinical traits (heatmap, bottom right). Each coloured block represents a distinct module of coexpressed genes identified by the network. (B) Heatmap illustrating correlation of module eigengenes with laboratory data (ordered traits listed in Supp Table 1). (C) Heatmap illustrating correlation of supporting clinical data with modular eigengenes. Strong, significant correlation was apparent between gender and a module that almost exclusively comprised X/Y-linked transcript expression (red box). (D) Heatmap illustrating correlation of modular eigengenes (y, coloured blocks) with baseline CIS traits (E) Heatmap illustrating unsupervised hierarchical clustering of the STAyCIS $CD8^+$ transcriptional dataset by the T cell exhaustion signature previously shown to correlate with clinical outcome in multiple autoinflammatory diseases. (F) Time to primary endpoint for MS patient subgroups defined by the T cell exhaustion signature as indicated in panel (E). Mann –Whitney two-tailed test of significance, P = 0.53. PEP = primary endpoint, EXH = exhaustion. For traits included in panels B-D see Supplementary Table 1.

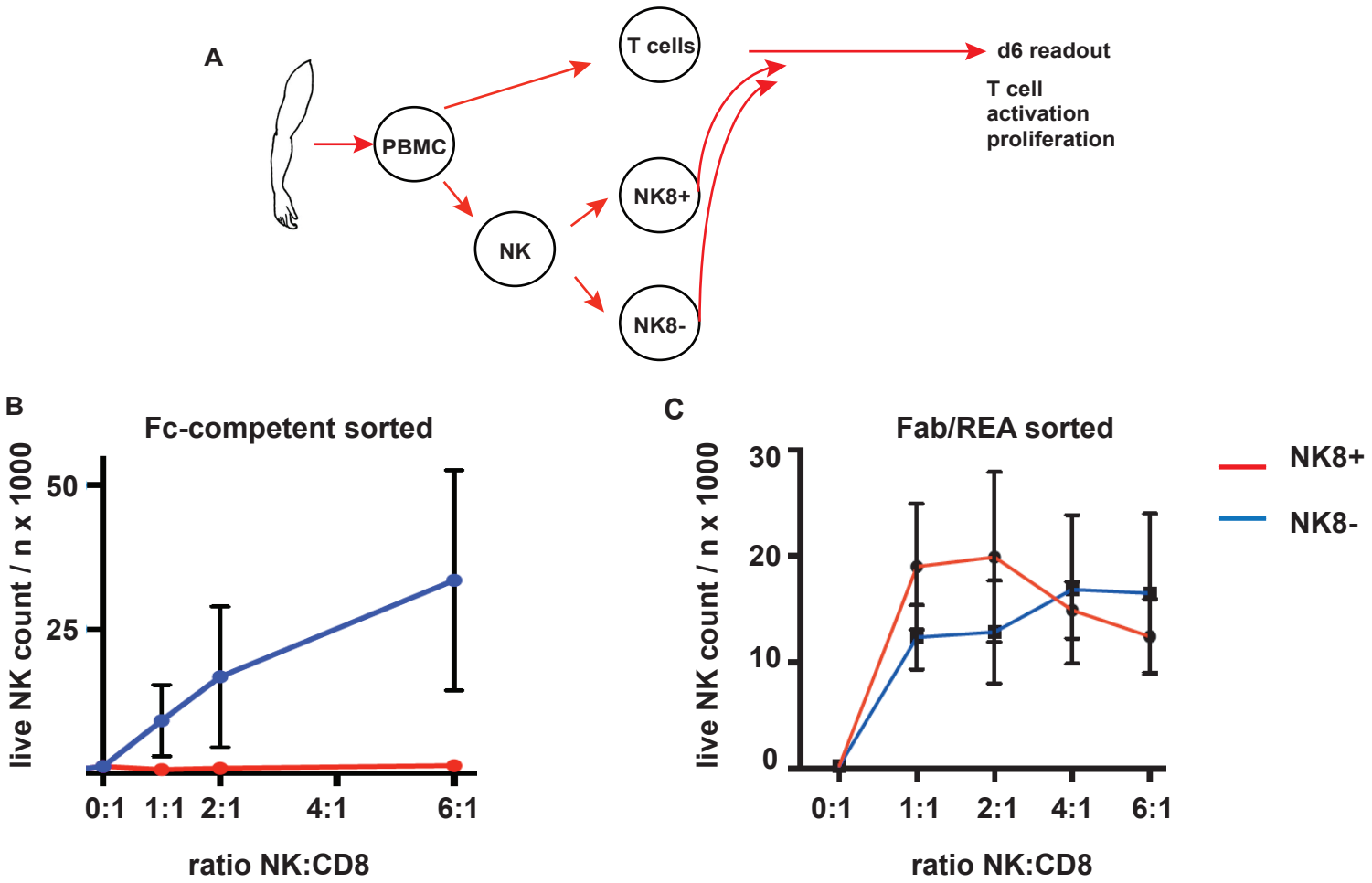
Supplementary Figure 3



Supp Fig 3: in vitro NK functional studies

(A) Representative pseudocolored scatterplots illustrating MACS enrichment of $CD8^+$ cells with $CD3^-CD56^+CD8^+$ cells as the principal non-T cell 'contaminant' population. Post enrichment plots show a representative sample gated by major cell surface markers. (B) Heatmap illustrating mRNA expression of CD8 by immune cell subtype. Y-axis ranked by CD8 expression level across all immune cell types in immune response in silico (IRIS) dataset (C) Representative pseudocolored scatterplots illustrating CD8 expression on $CD56^{bright}$ and $CD56^{dim}$ natural killer cell subsets. (D-G) Scatterplots illustrating correlation (linear regression with Pearson r^2 and two-tailed significance test) of $\%CD56^+$ total (D,F % lymphocytes) and $\%CD56^+CD8^+$ (E,G % lymphocytes) with clinical outcome directly (D, E) and with the $NK8^+$ 'black' module eigengene (F,G).

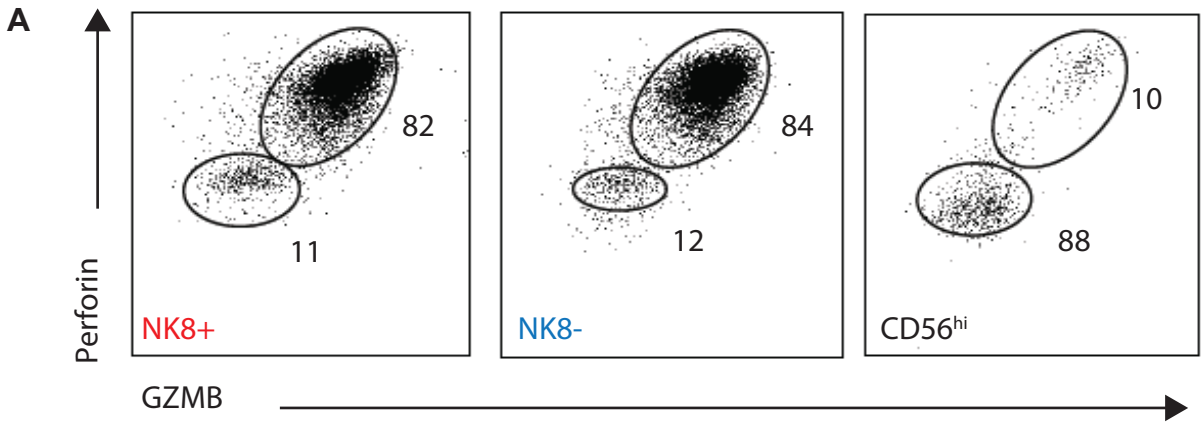
Supplementary Figure 4

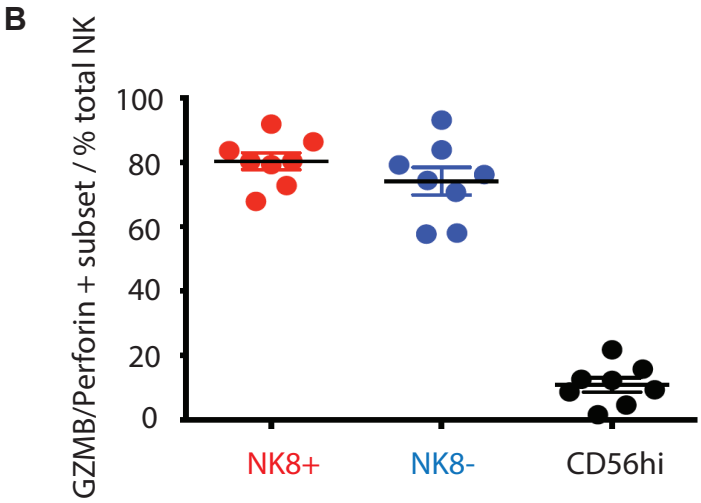


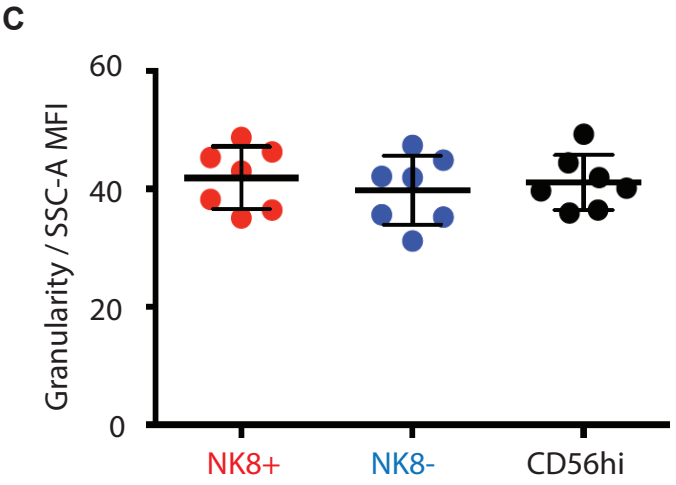
Supp Fig 4: NK/T cell co-culture

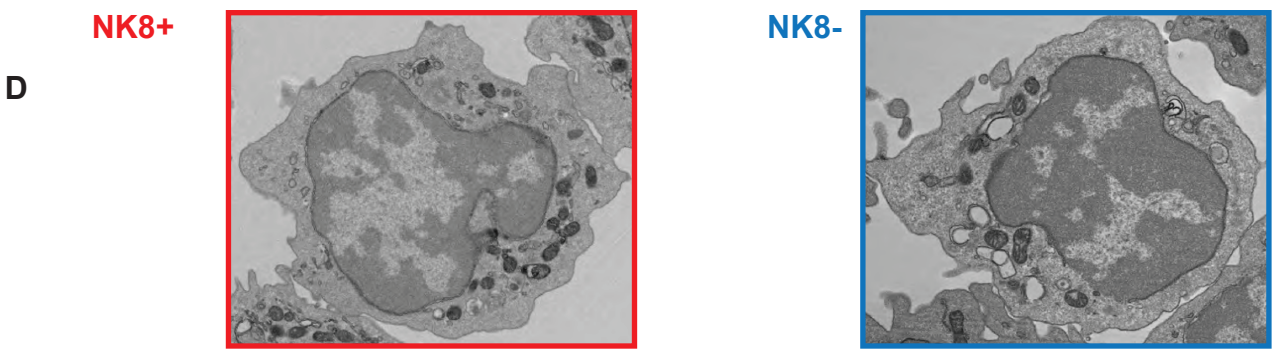
(A) Schematic illustration of autologous T cell and $NK8^{+/-}$ co-culture experiment. (B, C) Line and scatterplots illustrating $NK8^{+/-}$ survival during titrated coculture with autologous T cells following NK sorting with Fc-competent antibodies (B) or either Fc-incompetent antibodies/Fab fragments. Sorting NK with Fc-competent antibodies resulted in substantial and differential fratricide of $NK8^{+}$ cells, due to retention of $CD8^{+}$ antibodies and resultant Fc-mediated cell activation. This was not seen during functional studies reported here, which were performed using Fab and recombinantly engineered Fc-binding deficient antibodies (as in C). For B, C: n = 3, 10 independent replicates respectively; error bars = mean +/- sem.

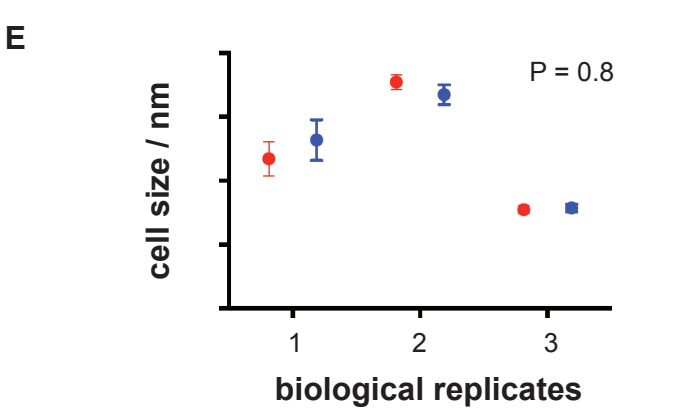
Supplementary Figure 5

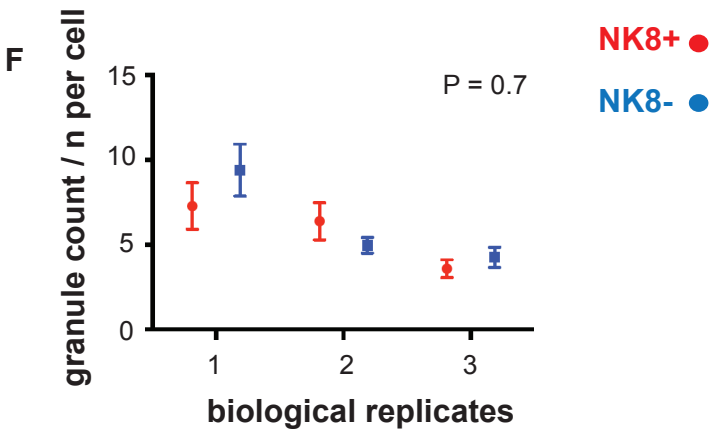








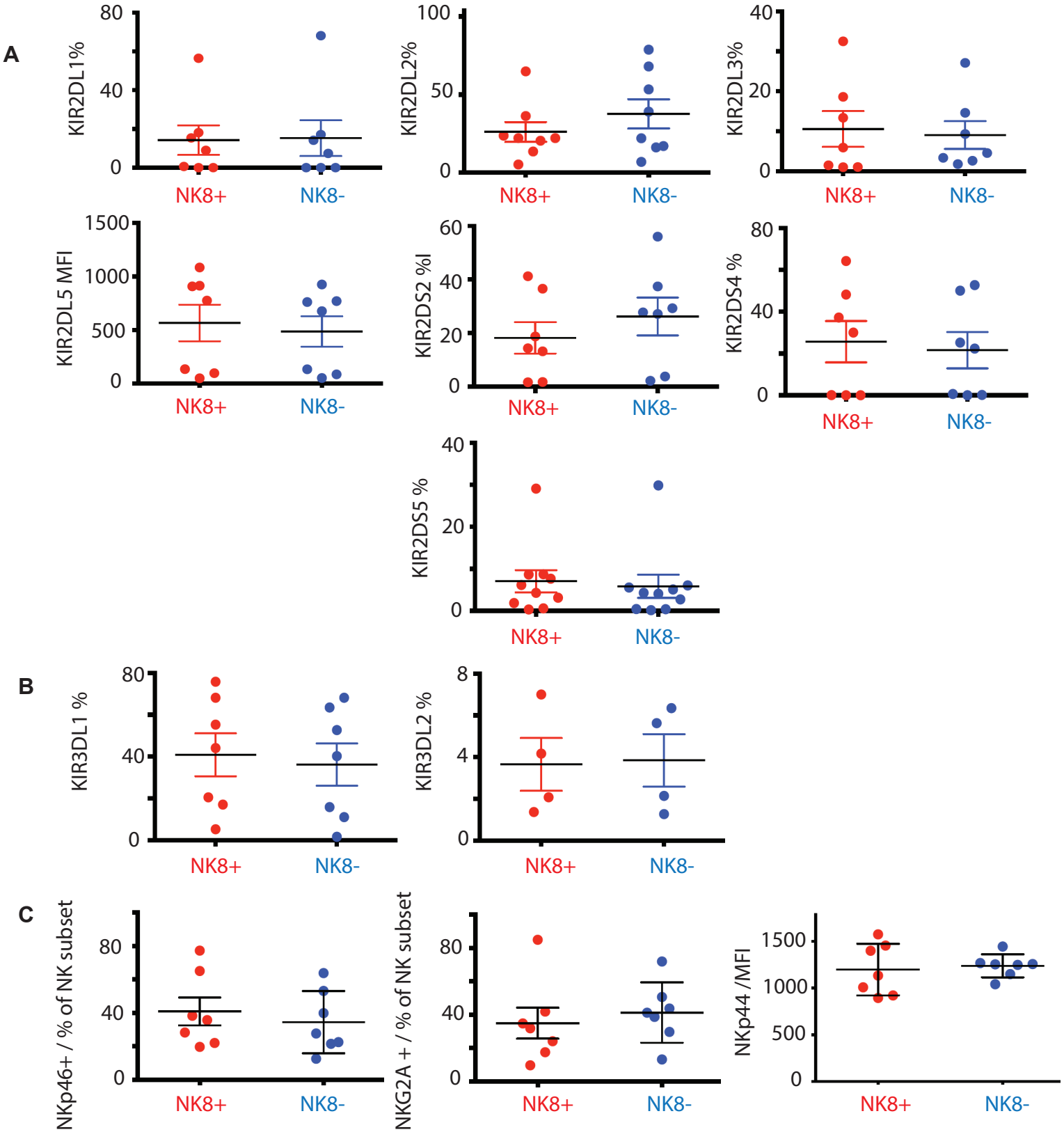




Supp Fig 5: Cytotoxic potential of NK subsets

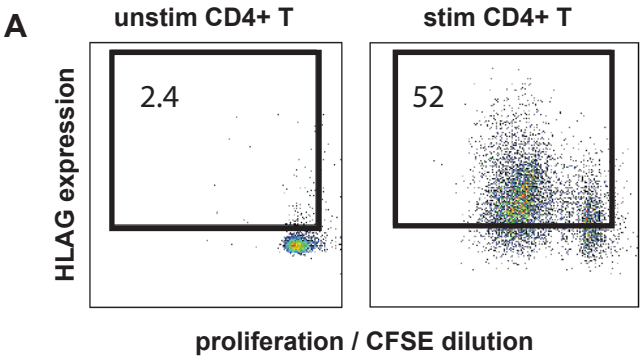
(A-C) Representative flow cytometry scatterplots (A) and summary scatterplots (B, C) illustrating cytotoxic granule protein (perforin, y axis; granzyme B, x-axis) expression and cell granularity (SSC-A MFI) in $NK8^+$ (red), $NK8^-$ (blue) and $CD56^{hi}$ (black) NK subsets. For B, C: n = 8, 7 independent biological replicates respectively (D) Representative transmission electron microscopy (TEM) images of $NK8^+$ (left, red) and $NK8^-$ (right, blue) cells. (E, F) Scatterplots summarising quantification of cell size (E) number of lytic granules per cell (F) for $NK8^+$ (red) and $NK8^-$ (blue). Mean +/- SEM for 10 TEM images from flow sorted purified NK subsets across three biological replicates (from healthy donors). P = two-tailed Mann-Whitney test.

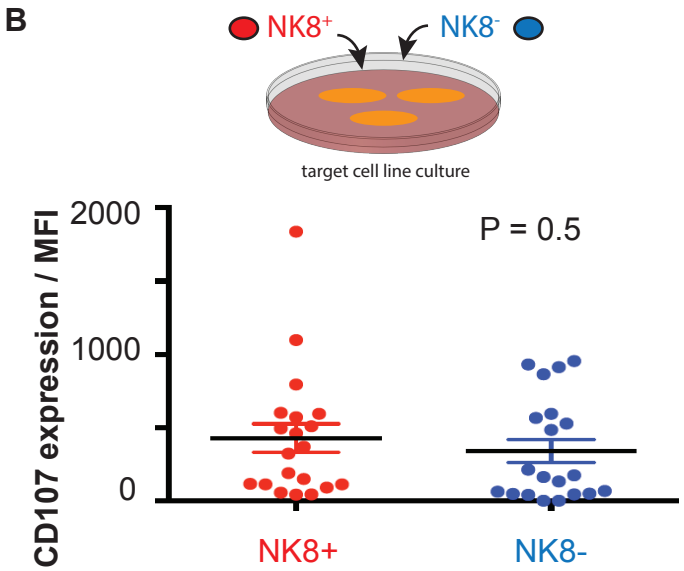
Supplementary Figure 6



Supp Fig 6: Activating/inhibitory NK cell receptor expression by $NK8^+$ and $NK8^-$ subsets
Scatterplots illustrating comparison of NK cell activating/inhibitory receptor protein expression on $NK8^+$ and $NK8^-$ cells measured by flow cytometry using a validated panel of KIR-targeting antibodies (Suppl Table 4). (A) KIR2 family expression, (B) KIR3 receptor expression, (C) selected other receptors. For A: n=7-10 replicates as indicated, B: n=4-7 and C = 7 independent biological replicates.

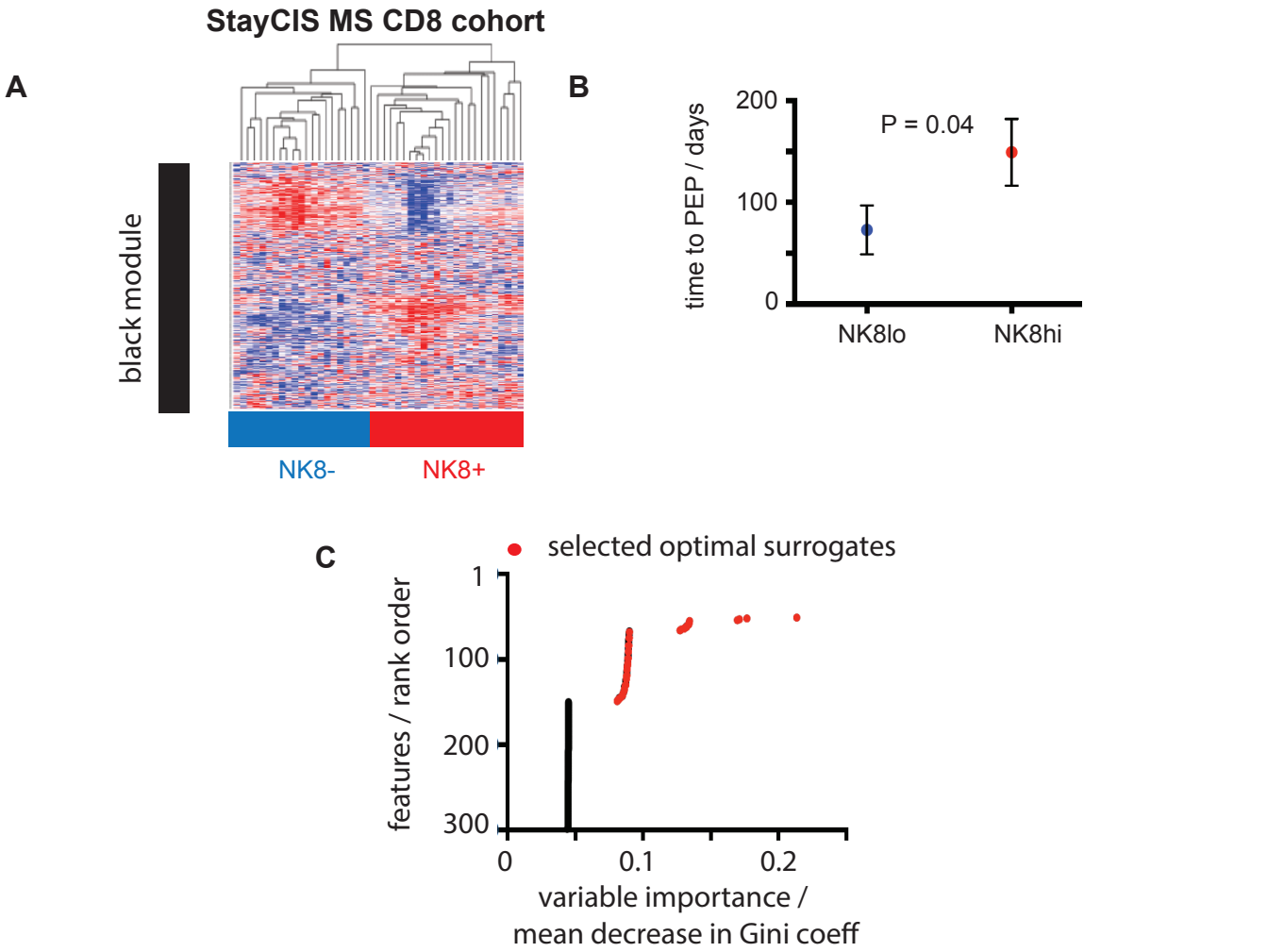
Supplementary Figure 7





Supp Fig 7: HLA-G expression on CD4 cells and induced cytotoxicity of NK subsets
(A) Representative scatterplots illustrating HLA-G expression (y-axis) on unstimulated (left) and polyclonally stimulated $CD4^+$ T cells (right). x-axis = CFSE dilution reflecting proliferation. (B) Scatterplot showing induced NK cytotoxicity (CD107 MFI, y-axis) on coculture of $NK8^+$ (red) or $NK8^-$ (blue) with tumour cell lines (K562/Jurkat). For B, n=20 independent biological replicates; error bars = mean +/- sem with Mann-Whitney two-tailed test of significance.

Supplementary Figure 8

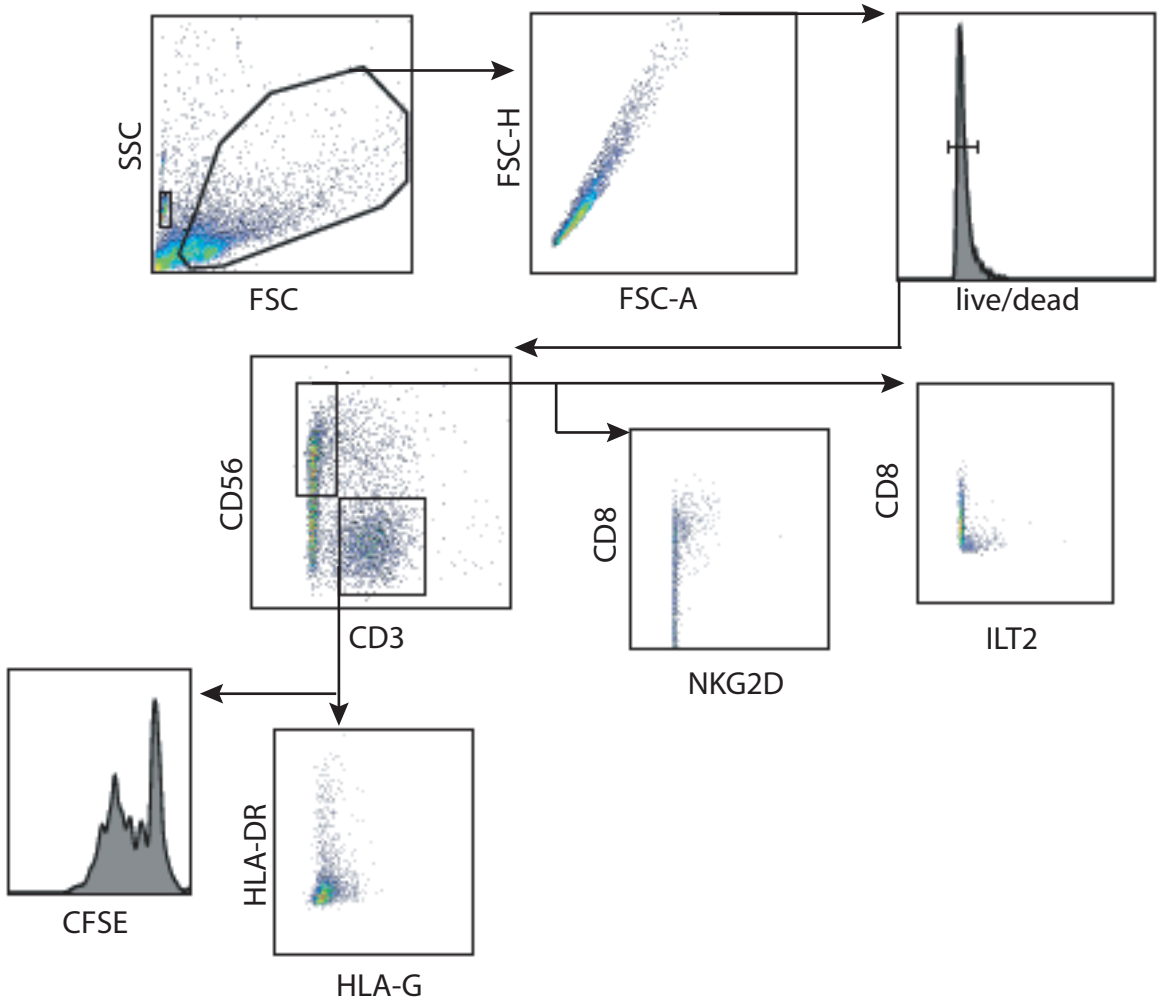


Supp Fig 8. Identification of surrogate markers of an NK8^{+} signature in PBMC samples for validation purposes

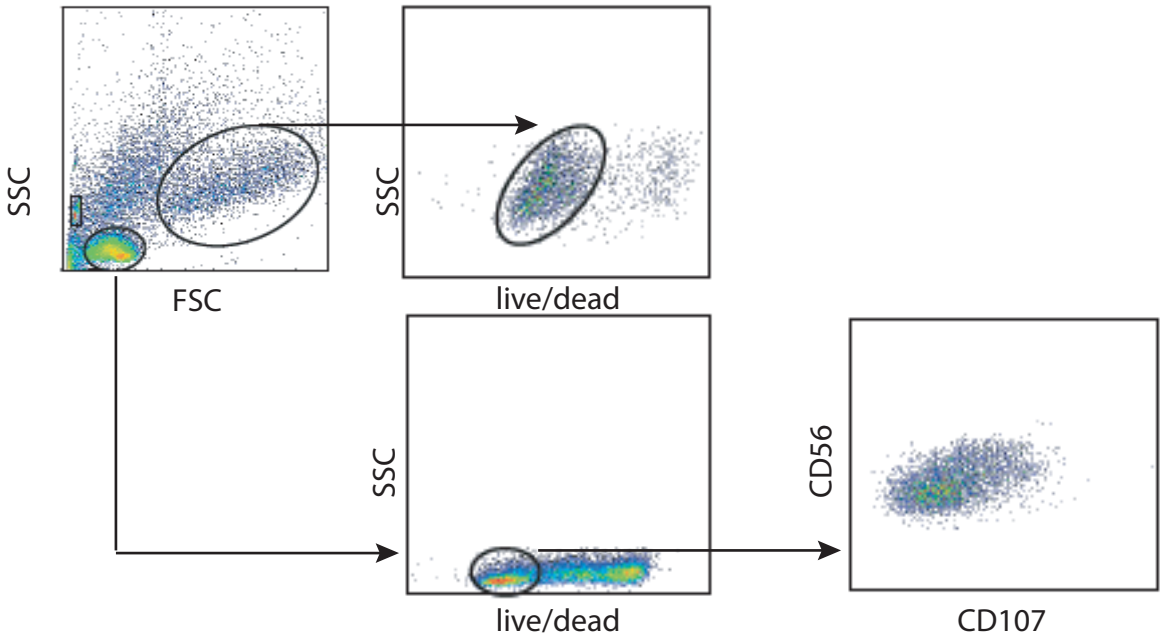
Heatmap illustrating hierarchical clustering of the STAyCIS MS cohort using the 'black' module genes (A) with associated clinical outcome of identified patient subgroups (B) as in Fig 1C. (C) Scatterplot showing ranked optimal predictive genes (y-axis) against feature importance (x-axis, decrease in Gini Coefficient for models not incorporating each feature) following randomforests feature selection identification, searching for analogous subgroups in a matched PBMC cohort of anti-neutrophil antibody associated vasculitis (AAV) patients. This set of features was then used for clustering of an independent MS cohort as shown in Fig 3D. (D) Heatmap illustrating hierarachical clustering of PBMC data from n=47 patients (columns) with AAV using the optimal surrogate markers of the MS black eigengene (rows). (E) Dot plot showing NK8lo and NK8hi subgroups from AAV cohort in panel D and associated time to relapse (mean +/- sem).

Supplementary Figure 9

A



B



Supp Fig 9. Flow cytometry gating strategies

Representative scatterplots showing flow cytometry gating strategies for (A) NK immunophenotyping (relating to Fig2, Supp Figs4-6) and (B) NK cytotoxicity assays (relating to Fig3, Supp Fig7).