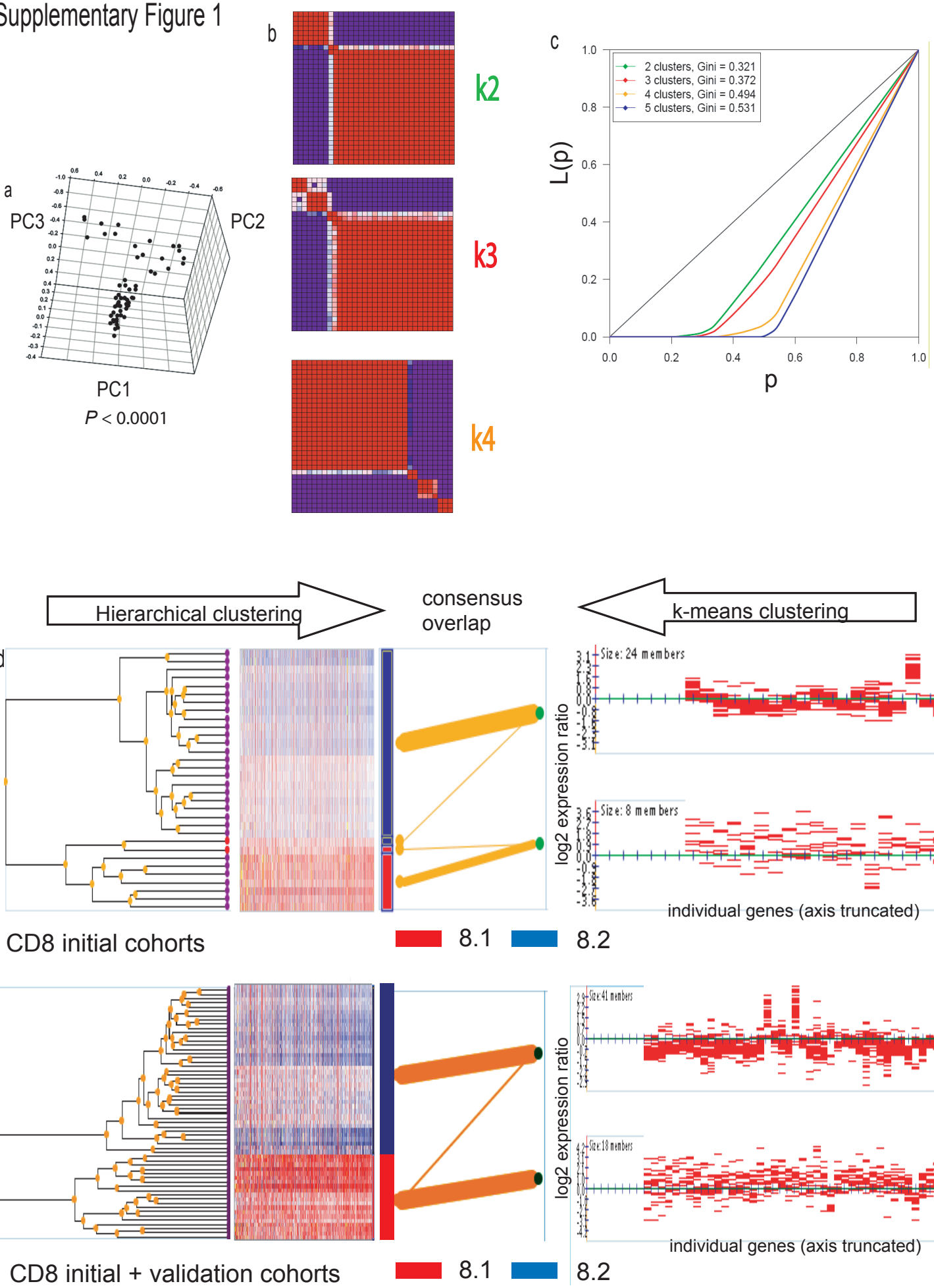
Supplementary Figure 1

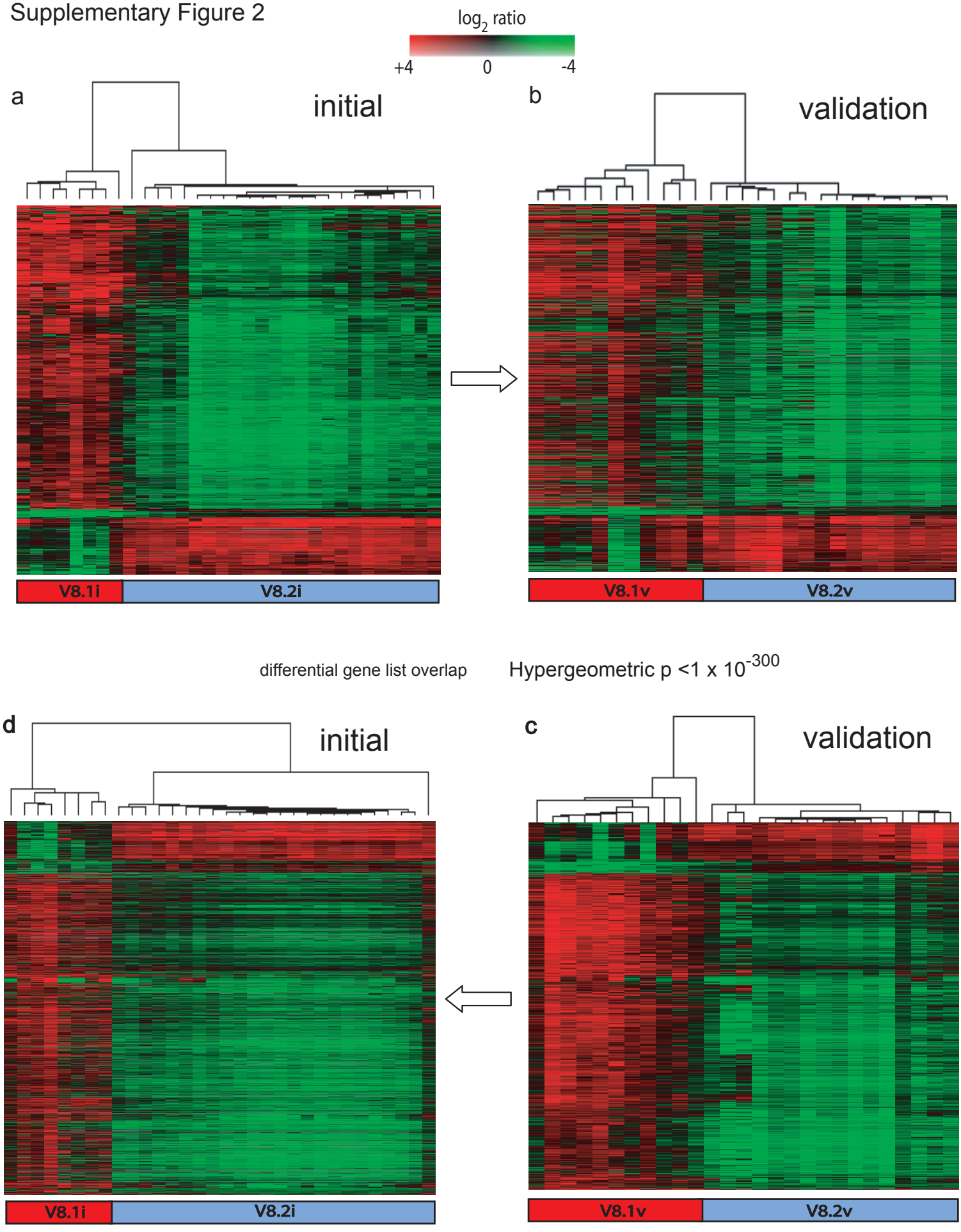


Supplementary Fig. 1: Gene expression data from vasculitis patients shows significant substructure with an optimal division into 2 subgroups.

Scatterplot (**a**) showing the top 3 principal components of expression data in CD8 T cells. A global test of clustering was used to compare the dataset with a multivariate unimodal distribution using multidimensional scaling (BRB-ArrayTools)[1]. This confirms a significant degree of 'substructure' in the dataset as a whole, $p<0.0001$. (**b, c**) Consensus heatmaps showing the proportion of times each patient clusters next to all others across multiple subsampling runs of unsupervised hierarchical clustering [2]. The heatmap spectrum ranges from 1 (red) to 0 (blue) while patients are arranged in identical order along x and y axes (thus giving proportion = 1 along a diagonal axis, top left to bottom right, each patient by definition clustering with themselves 100% of the time). The relative stability of splitting the expression dataset into different numbers of subgroups (k) can be assessed by visual inspection of consensus heatmaps produced for each division (**b**, k=2-4) and by deriving the Gini coefficient from the Lorenz plot relating to each division [2] (**c**). In this way similar samples are seen to cluster together (in red), separated from dissimilar samples (in blue) with colours of intermediate intensity reflecting less stable subdivisions of the data. (**d, e**) Gene expression data from CD8 T cells for 59 patients with vasculitis (n = 32 initial cohort, and n = 27 validation cohort) was clustered using two independent clustering techniques (hierarchical clustering, left panel, and k-means clustering, truncated, right panel). Consensus subgroups (shown centrally as red and blue vertical bars) were defined using a comparative algorithm (ClusterComparison)[3] which matches subdivisions produced by each technique. (**d**) AAV initial cohort and (**e**) initial and validation cohorts combined. For hierarchical clustering

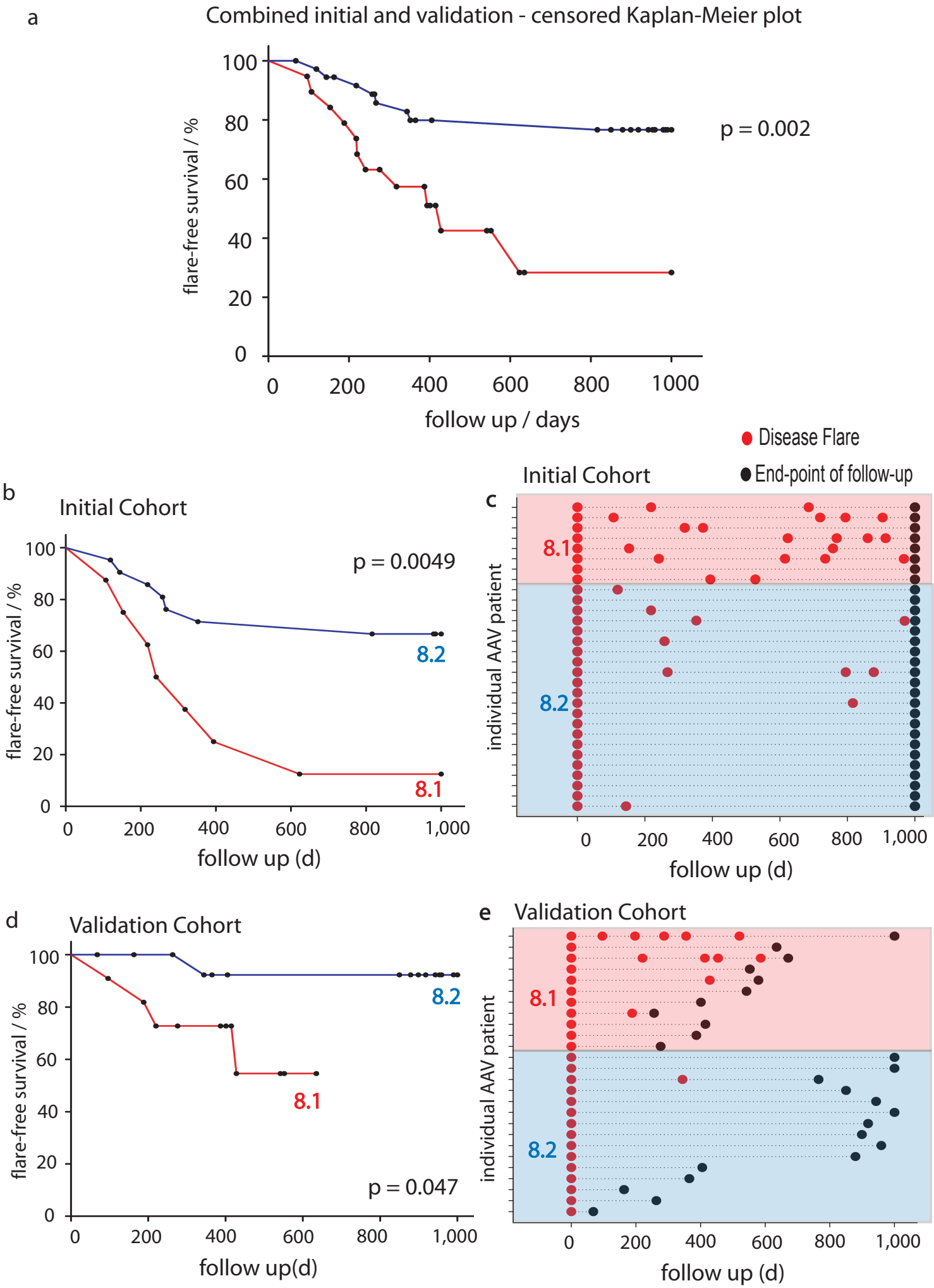
plot, red = upregulated expression, blue = downregulated exrpression. For k-means clustering y axis = $\log_2$ expression ratio, x-axis (truncated) showing expression of individual genes by each patient (red bars).

Supplementary Figure 2



Supplementary Fig. 2: Analysis of CD8 T cell expression data from two independent cohorts of vasculitis patients identifies identical subgroups. (**a**) Supervised hierarchical clustering of CD8 samples from the initial cohort of vasculitis patients (n = 32) demonstrates the two subgroups as defined previously. (**b**) Unsupervised hierarchical clustering of CD8 samples from an independent validation cohort of vasculitis patients (n = 27) using the same genes as in panel **a** also produced two subgroups with similar expression profiles to those seen in the initial cohort. (**c**) An independent unsupervised analysis of CD8 samples from the validation cohort separated the patients into identical subgroups to those obtained previously (in panel **b**). (**d**) Reclustering of CD8 samples from the initial cohort with the genes derived from the validation subgroups (used in panel **c**) faithfully reproduced the two subgroups seen initially (in panel **a**). The differentially expressed genes defining the CD8 subgroups in both cohorts (panels **a** and **d)** showed a highly significant degree of overlap (n = 490/925 genes for the initial and 490/1523 genes for the validation cohort at fold-change >2, $p < 1 \times 10^{-300}$). Hierarchical clustering was performed using an uncentered correlation metric and average-linkage analysis. Significant overlap between lists of differentially expressed genes was determined using hypergeometric probability.

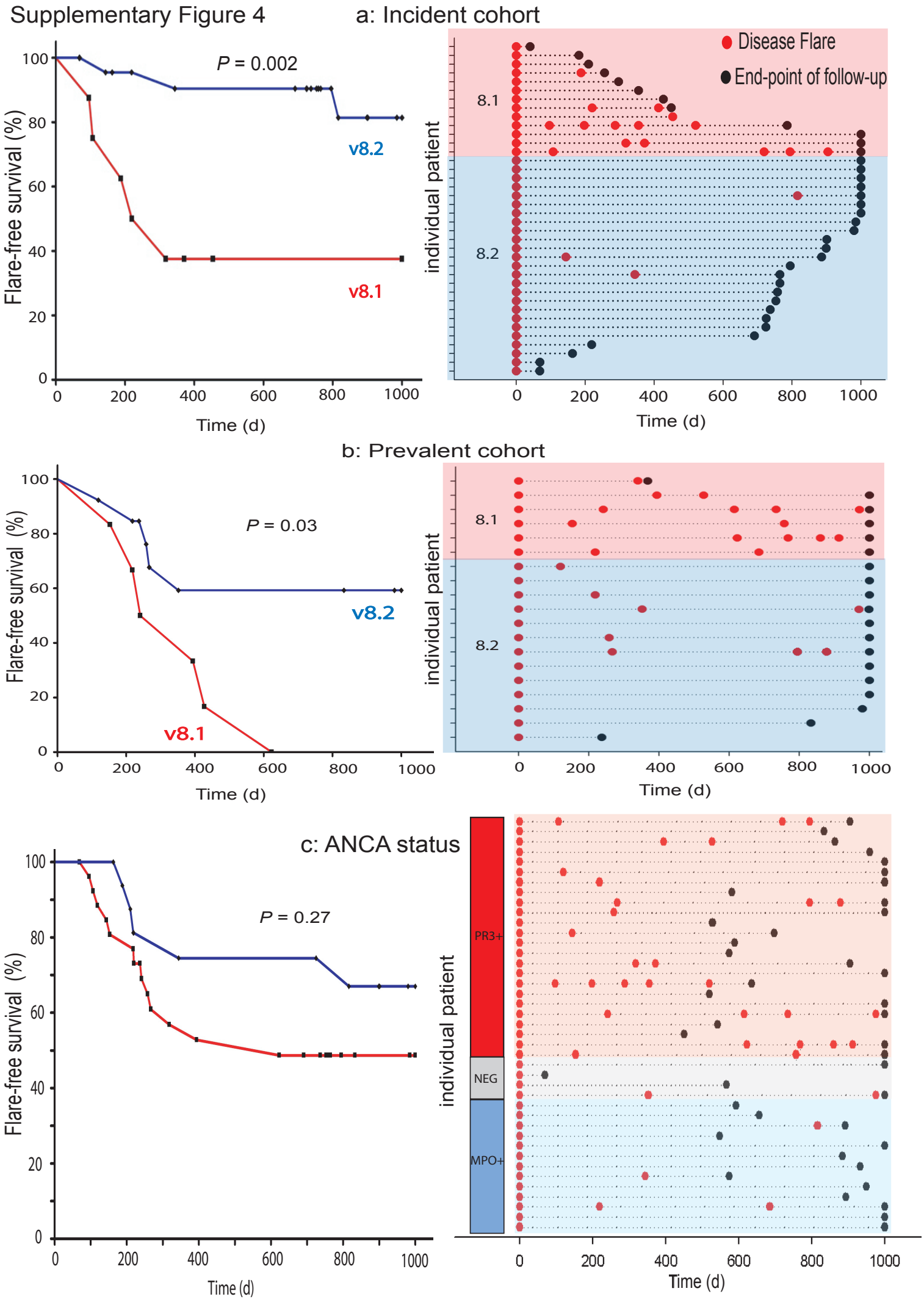
Supplementary Figure 3



Supplementary Fig. 3: The CD8 T cell expression signature identifies an AAV patient subgroup with poor prognosis in both initial and validation cohorts.

When applied to both (**a**) or either (**b-e**) of two independent prospective cohorts of vasculitis patients the CD8 T cell signature identifies a subgroup with poor prognosis, showing both shorter time to first flare (**b, d**) and increased flare-rate over time (**c, e** mean flare-rate normalized to duration of follow-up = 1 flare / year in 8.1 initial v 0.25 flares / year in 8.2 initial and 0.68 flares / year in 8.1 validation v 0.05 flares / year in 8.2 validation). (**a, b** and **d**) Kaplan-Meier survival curves censored for duration of follow-up, p-values determined by a log-rank test of significance. (**c** and **e**).

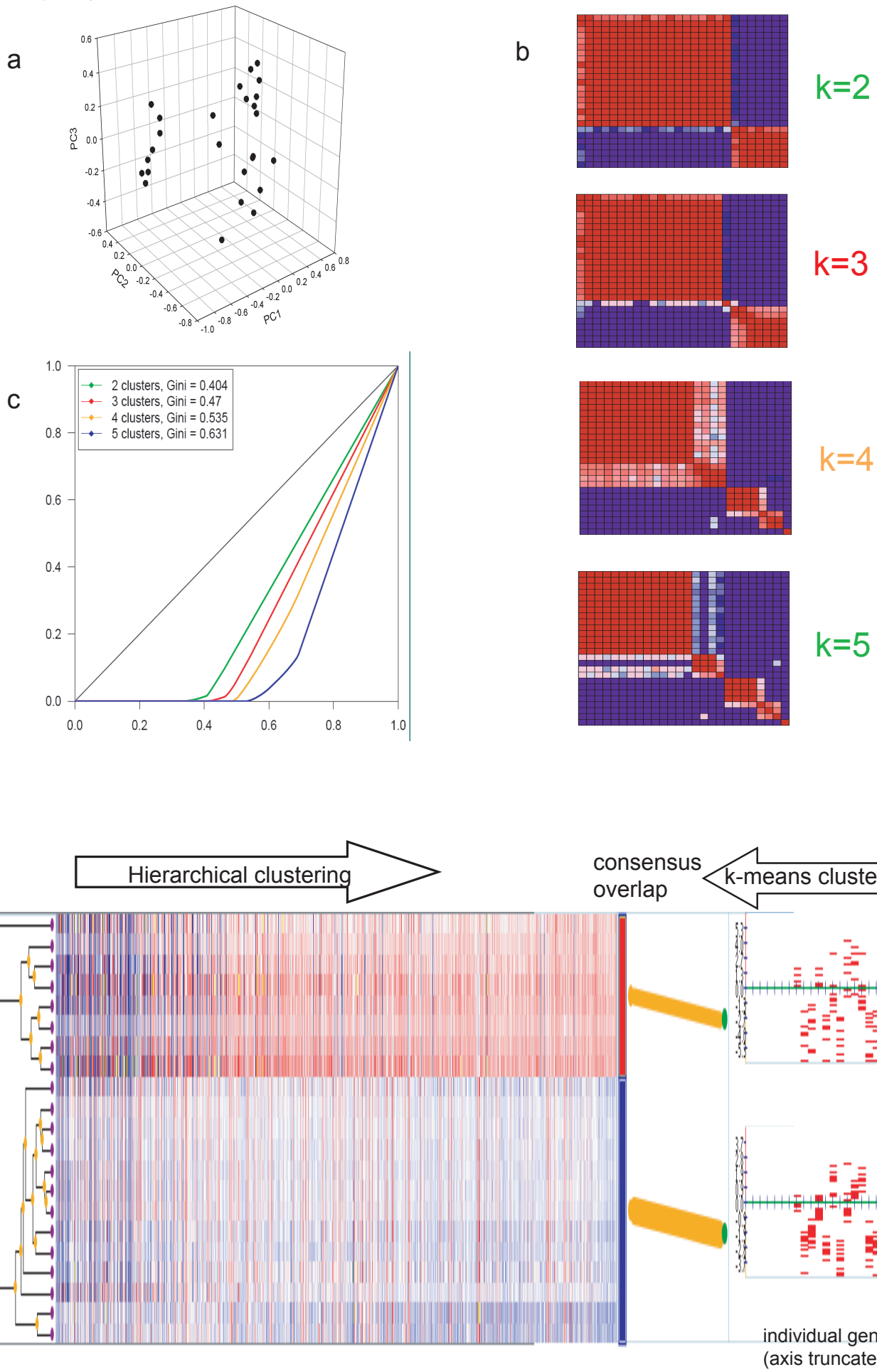
Supplementary Figure 4



Supplementary Fig. 4: The CD8 T cell expression signature can identify an AAV patient subgroup with poor prognosis in both incident and prevalent populations while risk stratification by ANCA status does not.

When applied to a subpopulation of AAV patients presenting for the first time with no previous episodes of disease (incident population) the CD8 T cell signature identifies a subgroup of patients with poor prognosis, showing shorter time to first flare (**a**) and increased flare-rate over time (**b**, mean flare-rate normalized to duration of follow-up = 0.66 flares / year in 8.1 v 0.06 flares / year in 8.2). Similarly, a subpopulation of AAV patients who have experienced previous disease episodes (prevalent population) - and who have therefore, expectedly, a higher rate of relapse – can also be stratified by prognostic risk using the same CD8 T cell signature. The prevalent population also show a shorter time to first flare (**c**) and an increased flare-rate over time (**d**, mean flare-rate normalized to duration of follow-up = 1.05 flares / year in 8.1 v 0.25 flares / year in 8.2). (**e**) ANCA status does not predict time to first flare. Flare-free survival shown as a Kaplan-Meier plot with significance measured using the log-rank test. (**f**) PR3-ANCA and MPO-ANCA positive patients showed similar numbers of flares when followed to 1000 days post-treatment (mean flare rate normalised to duration of follow-up = 0.23 flares / year in MPO+ group v 0.43 flares / year in PR3+ group, p=0.27).

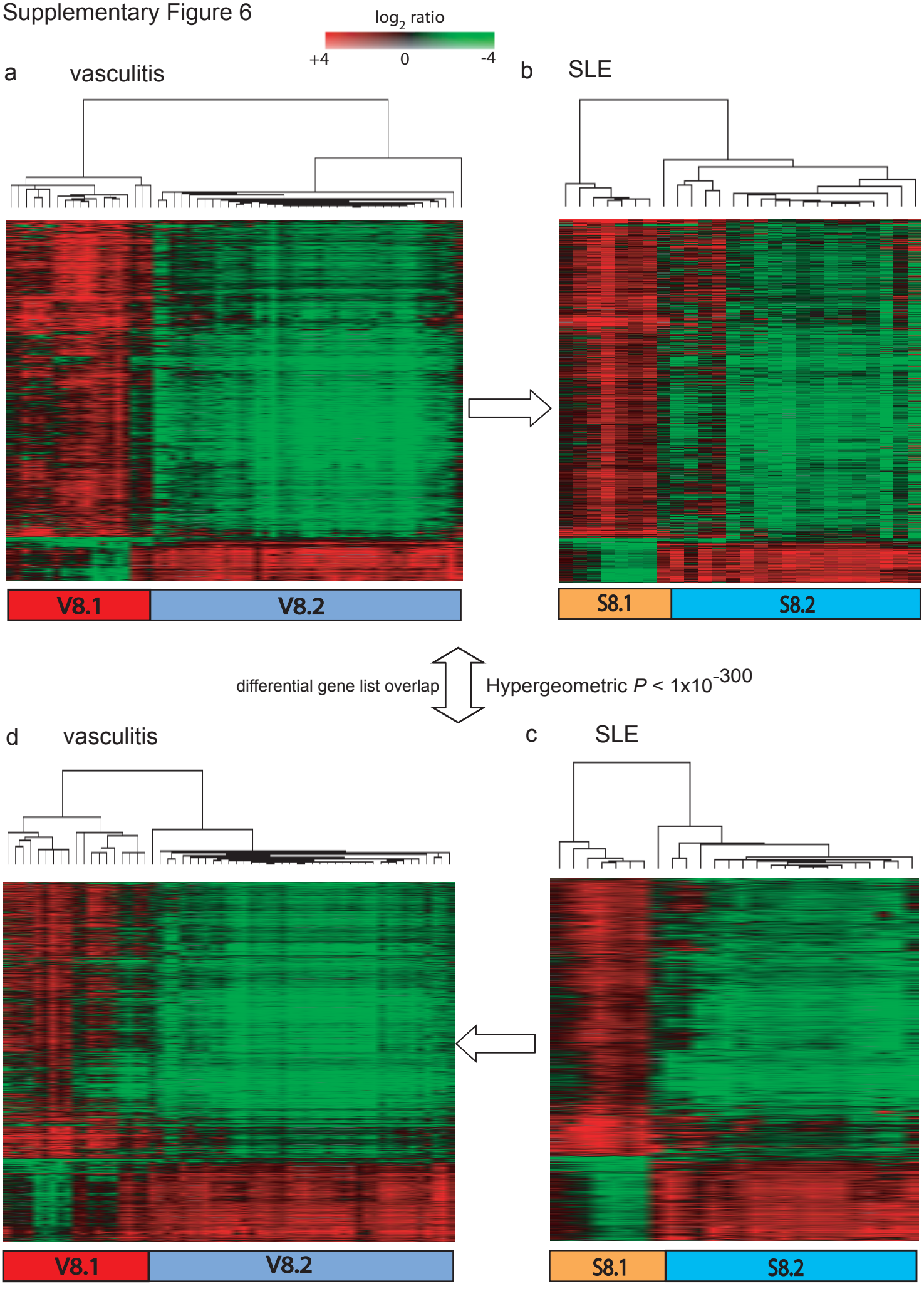
Supplementary Figure 5



Supplementary Fig. 5: Gene expression data from SLE patients shows significant substructure with an optimal division into 2 subgroups.

(**a**) 3D scatterplot showing the top 3 principal components of expression data in SLE CD8 T cells (n=26). Global test of clustering could not be performed as fewer than 30 samples are included. (**b, c**) Consensus heatmaps showing the proportion of times each patient clusters next to all others across multiple subsampling runs of unsupervised hierarchical clustering[2]. The heatmap spectrum ranges from 1 (red) to 0 (blue) while patients are arranged in identical order along x and y axes (thus giving proportion = 1 along a diagonal axis, top left to bottom right, each patient by definition clustering with itself 100% of the time). The relative stability of splitting the expression dataset into different numbers of subgroups (k) can be assessed by visual inspection of consensus heatmaps produced for each division (**b**, k=2-5) and by deriving the Gini coefficient from the Lorenz plot relating to each division (**c**)[2]. In this way similar samples are seen to cluster together (in red), separated from dissimilar samples (in blue) with colours of intermediate intensity reflecting less stable subdivisions of the data. (d) Gene expression data from CD8 T cells for 26 patients with SLE were clustered using two independent clustering techniques (hierarchical clustering, left panel, and k-means clustering, truncated, right panel). Consensus subgroups (shown centrally as blue and red vertical bars) were defined using a comparative algorithm, ClusterComparison[3] which determines the best match subgroups across both techniques. For hierarchical clustering plot, red = upregulated expression, blue = downregulated exrpression. For k-means clustering y axis = log2 expression ratio, x-axis (truncated) showing expression of individual genes by each patient (red bars).

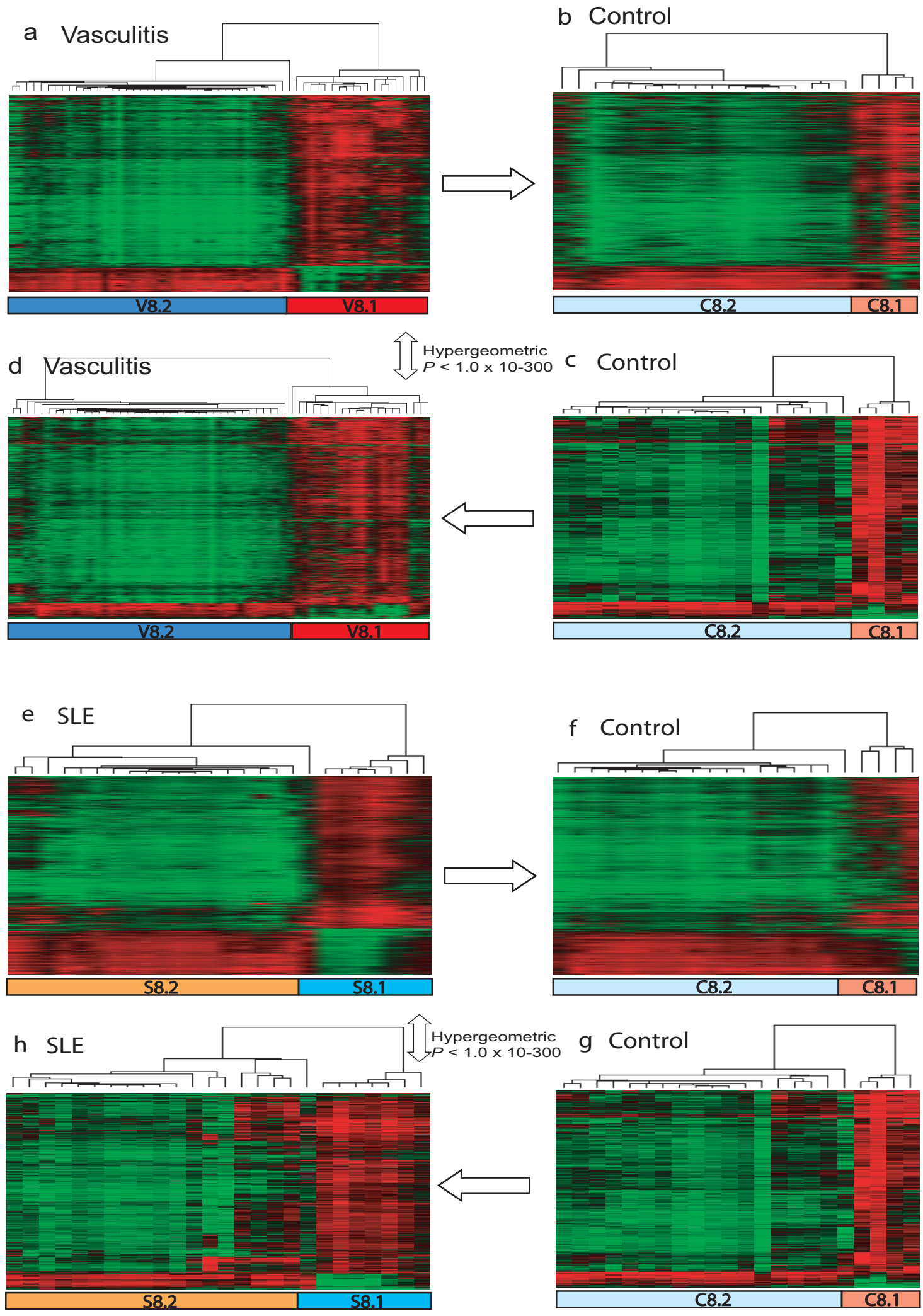
Supplementary Figure 6



Supplementary Fig. 6: Subgroups seen in vasculitis and SLE are defined by a highly similar list of differentially expressed genes.

Supervised hierarchical clustering of CD8 samples from vasculitis patients (**a**, n=59) demonstrates the two subgroups as defined previously. Reclustering of CD8 samples from SLE patients (**b**, n = 26) with the genes derived from the vasculitis subgroups (used in panel **a**) faithfully reproduces the two subgroups seen with an unsupervised approach (in panel **c**). Supervised hierarchical clustering of CD8 samples from SLE patients demonstrates the two subgroups as defined previously (**c**). Reclustering of CD8 samples from vasculitis patients (**d**) with the genes derived from the SLE subgroups (used in panel **c**) also reproduces the two subgroups seen with unsupervised clustering (in panel **a**). The differentially expressed genes defining the CD8 vasculitis and SLE subgroups (in panels **a** and **c**) showed a highly significant degree of overlap (n = 639/1228 genes in vasculitis and 639/1913 genes in SLE, $p < 1 \times 10\text{-}300$). Hierarchical clustering was performed using an uncentred correlation metric and average-linkage clustering. Significant overlap between lists of differentially expressed genes was determined using hypergeometric probability.

Supplementary Figure 7



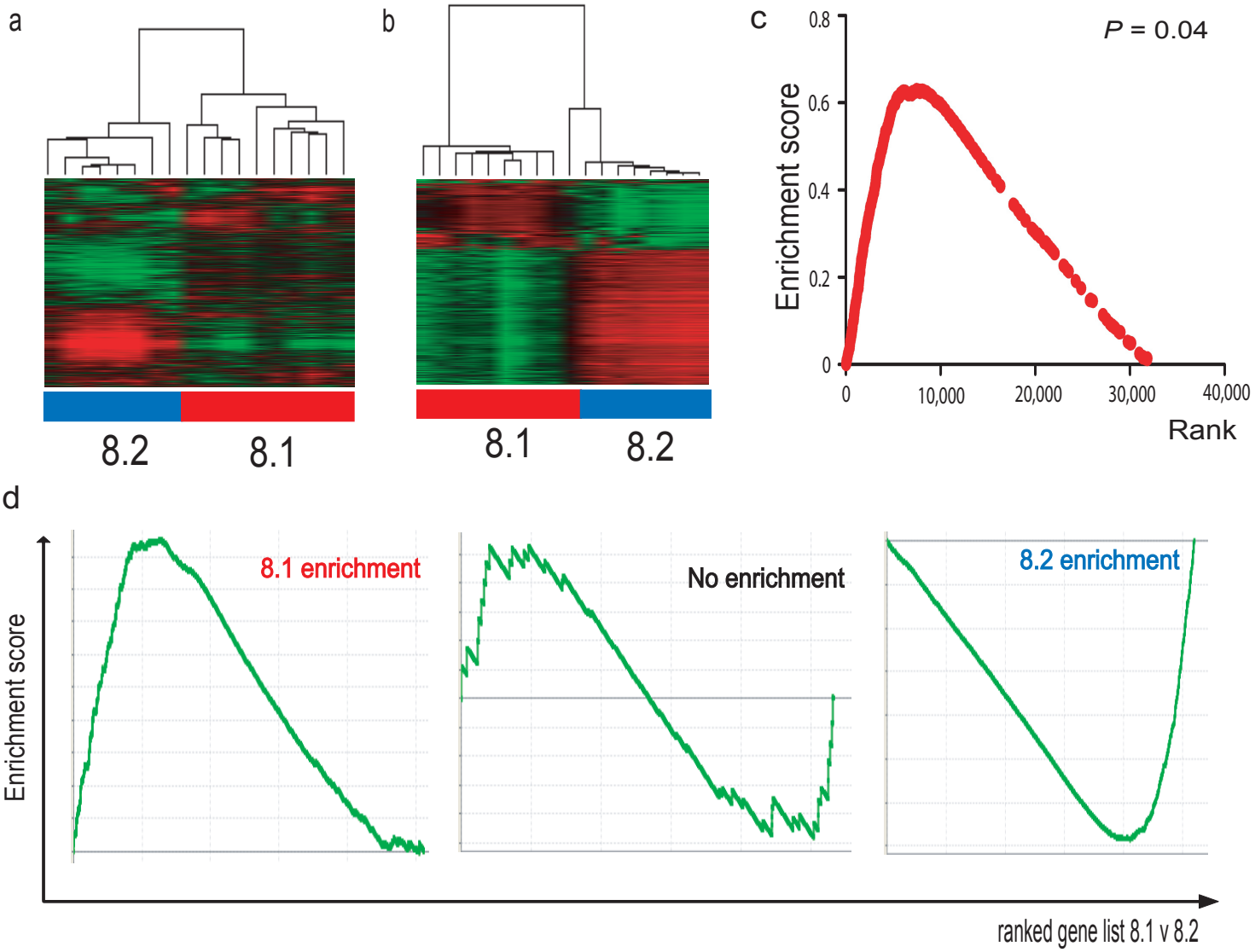
Supplementary Figure 7: Unsupervised clustering of CD8 T cell gene expression data in healthy controls identifies two similar subgroups, defined by a highly overlapping list of differentially expressed genes

Supervized hierarchical clustering of CD8 T cell samples from vasculitis patients reveals two distinct subgroups as defined previously (**a**). Supervised hierarchical clustering of CD8 T cell samples from healthy controls (**b**, n = 22) using genes derived from the vasculitis subgroups (as used in panel **a**) reveals two subgroups with similar expression profiles to those seen in the disease cohort. An independent, unsupervised analysis of CD8 T cell gene expression in healthy controls (using the methods described in Supplementary Fig.1) reveals the same 2 distinct subgroups (**c**). Reclustering of CD8 T cell samples from vasculitis patients with genes derived from the healthy control subgroups (**d**) reproduces the 2 subgroups seen initially (in panel **a**).

Supervised hierarchical clustering of CD8 T cell samples from SLE patients (n = 26) reveals two distinct subgroups as defined previously (**e**). Supervised hierarchical clustering of CD8 T cell expression data from healthy controls (**f**) using genes derived from the SLE subgroups (used in panel **e**) again reveals two subgroups with similar expression profiles to those seen in the disease group. Reclustering of CD8 T cell samples from SLE patients with genes derived from the healthy control subgroups (**g**) reproduces the 2 subgroups seen initially (panel **h**). The differentially expressed genes defining the CD8 vasculitis and healthy control subgroups (in panels **a** and **c**) showed a highly significant degree of overlap (n = 598/1228 vasculitis genes and 598/944 control genes, $p < 1 \times 10^{-300}$) as did the genes defining the CD8 SLE and healthy control subgroups (in panels **e** and **g**; n = 508/1913 SLE genes and 508/944 control genes, $p < 1 \times$

10^{-300}). Hierarchical clustering was performed using an uncentred correlation metric and average-linkage clustering. Significant overlap between lists of differentially expressed genes was determined using hypergeometric probability.

Supplementary Figure 8



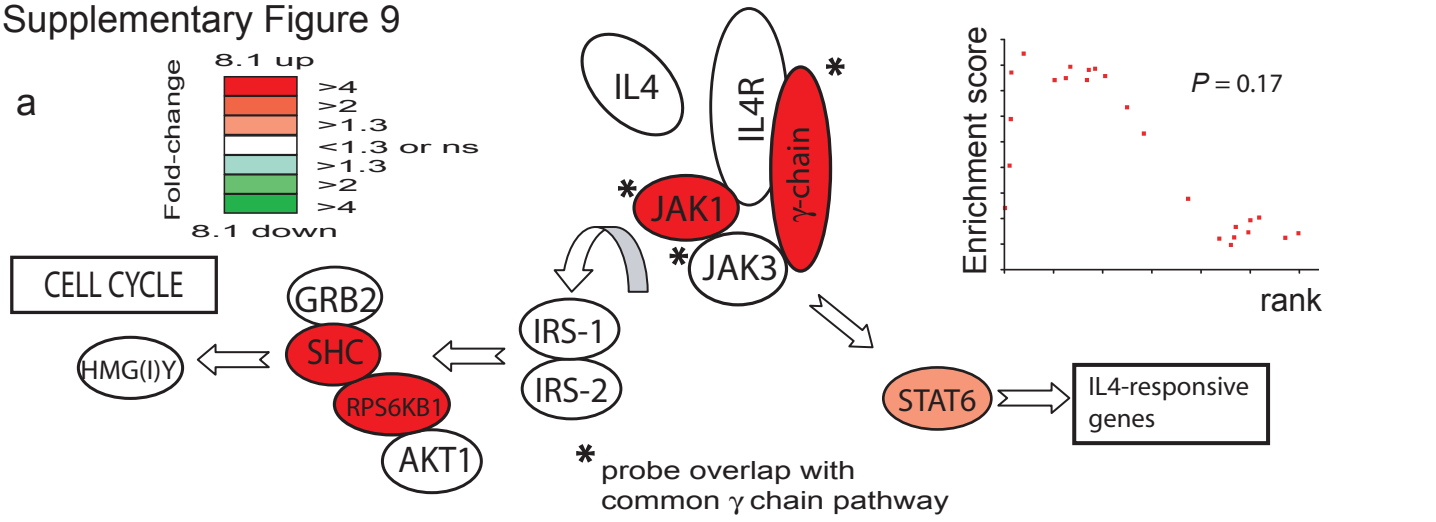
Supplementary Fig. 8: Validation of the custom oligonucleotide microarray findings on the Affymetrix Human Gene 1.0 ST Array platform.

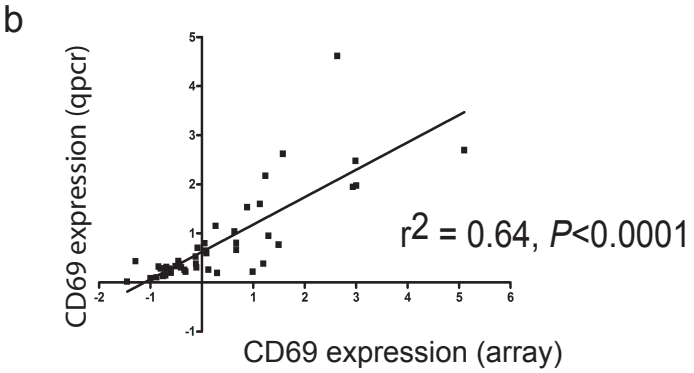
CD8 T-cell RNA from a subset of 18 patients (n = 9 SLE, n = 9 AAV) derived from both prognostic subgroups in the original cohorts (n = 8 subgroup 8.2, n = 10 subgroup 8.1) was hybridized to the Affymetrix Gene 1.0 ST expression array. The genes defining subgroups 8.1 and 8.2 (FDR<0.05, FC>2) were mapped onto Affymetrix IDs using common Entrez gene identifiers producing a signature of 1054 genes. **(a)** This translated gene signature was used to perform supervised hierarchical clustering of the Affymetrix-derived CD8 dataset, reproducing the patient subgroups seen on the original platform. (**b**) An unsupervised clustering using the most variable genes within the Affymetrix dataset from the same 18 patients again produced an identical subdivision (gene list filtered on median absolute deviation >0.6, n = 827). **(c)** Gene-set enrichment analysis confirms significant enrichment in the Affymetrix dataset of the 8.1 v 2 prognostic signature seen using the mediante custom array. (**d**) Gene-Set Enrichment Analysis model examples. Gene-Set Enrichment Analysis is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states[4], in this case subgroups 8.1 and 8.2. Genes are ranked based on differential expression between the 2 defined subgroups 8.1 and 8.2 (x-axis). The Enrichment Score (y-axis) reflects the degree of over-representation of a defined set of genes (e.g. a curated pathway or transcriptional “signature”) at either end of the ranked list. Significance of this enrichment is assessed by performing multiple random permutations of the phenotype label (i.e. 8.1 or 8.2) for each gene 1000 times, generating a null distribution of enrichment scores. The observed enrichment score for a given gene

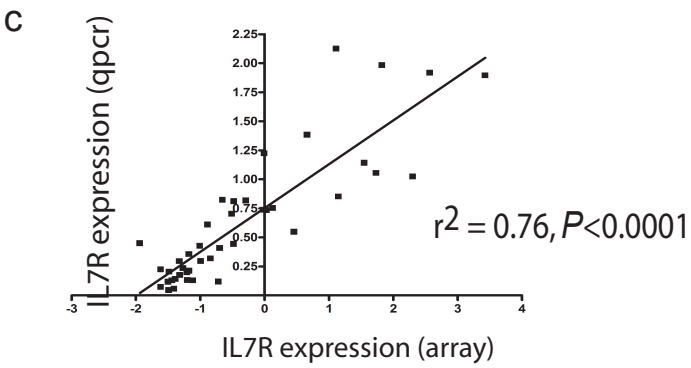
set is then compared to the derived null distribution. Model examples are shown to aid in the interpretation of other figures. (**e**) Expression density distribution templates used for matching profiles of differentially expressed genes.

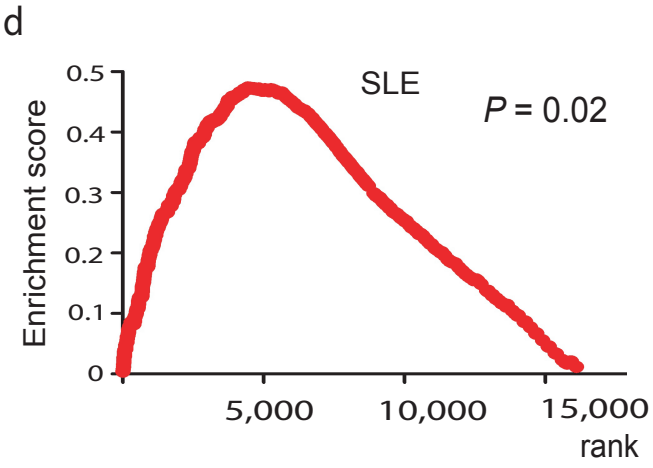
Differentially expressed gene signatures (FDR<0.05, fold-change >1.5) were derived from comparisons between both subgroups v8.1 v v8.2 and between a randomized, stratified subdivision of the vasculitis cohort (n=59). In order to compare expression profiles of individual genes in each list, they were compared to a range of idealized distribution templates illustrated here. Comparison was performed using the BioConductor expression density diagnostics package in R.

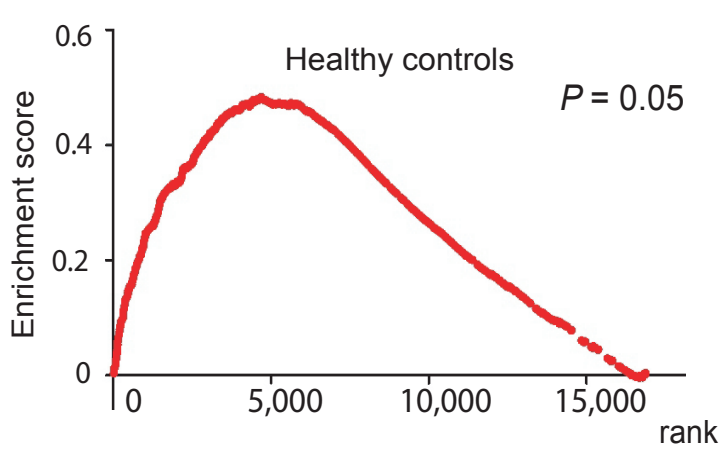
Supplementary Figure 9

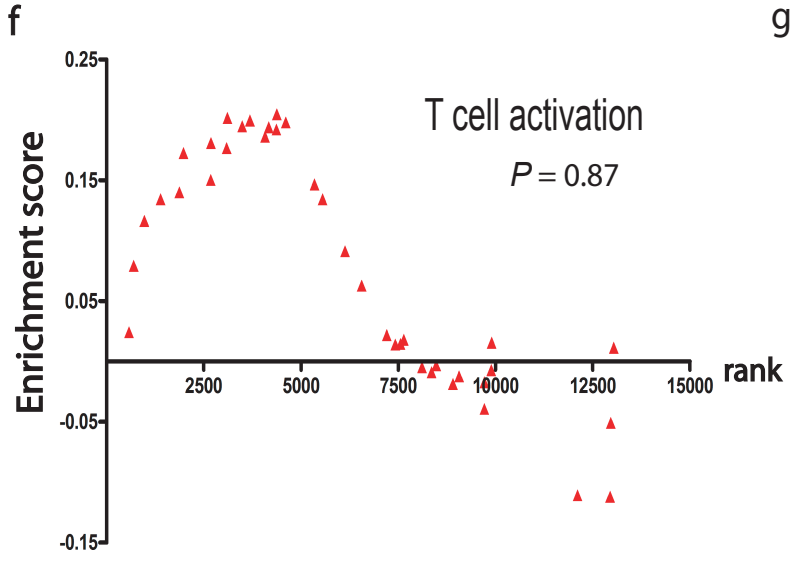


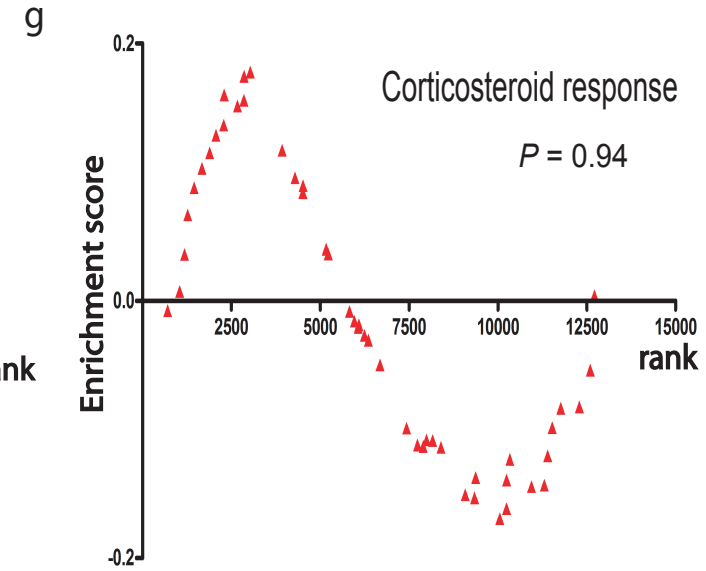








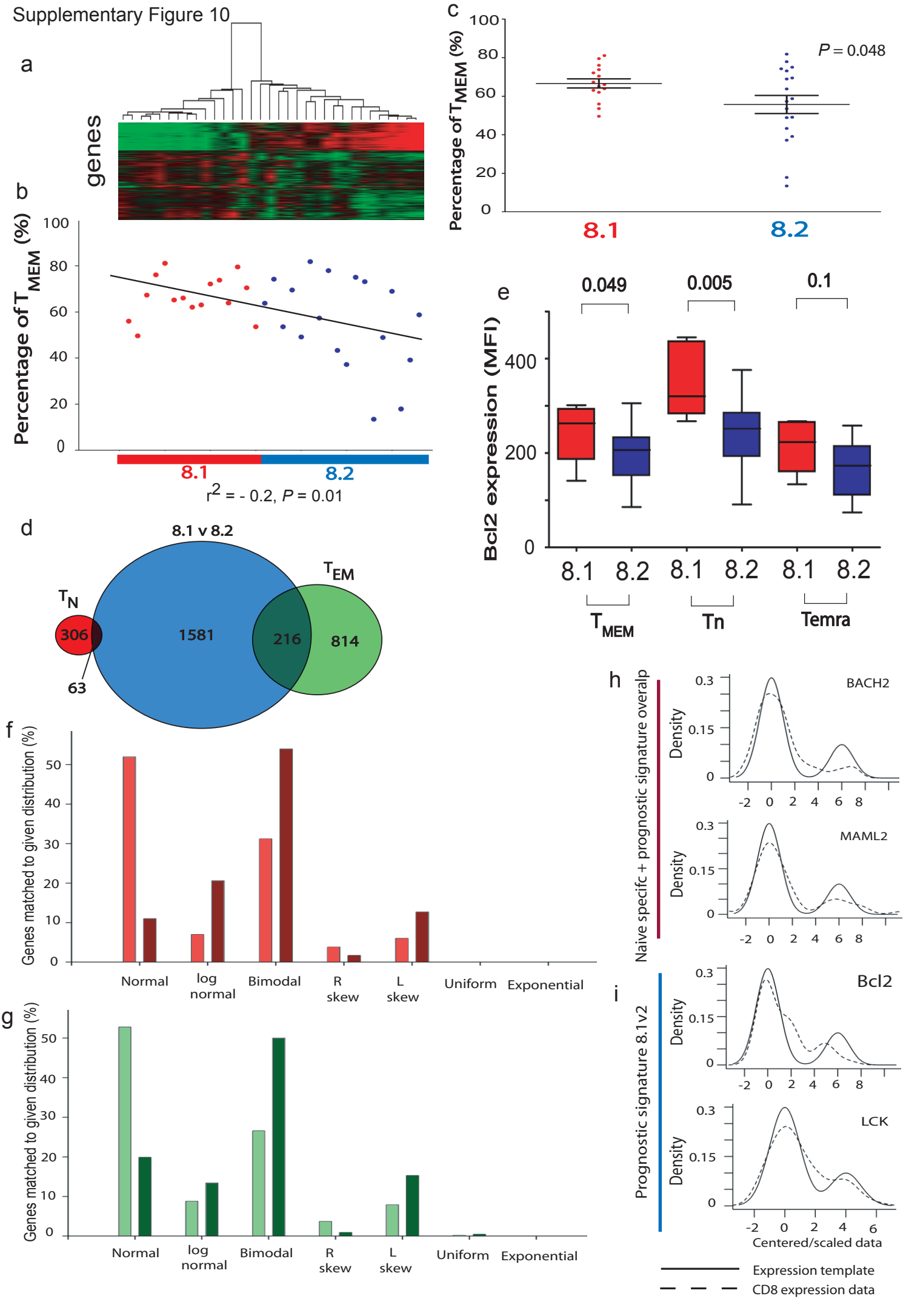




Supplementary Fig. 9: Enrichment is seen in SLE and healthcy controls, but is specific and is not seen for other signatures while qPCR amplification validates the microarray expression results.

(**a**) Genes annotated as comprising part of the IL4 signalling pathway (MSigDB)[4] were not enriched amongst the genes differentially expressed between subgroups 8.1 and 8.2 (GSEA, inset, p=0.17). Microarray and quantitative RT-PCR measurements of mRNA abundance showed a strong, statistically significant positive correlation. Example data are shown for CD69 (**b**) and IL7R (**c**). GSEA plots showing enrichment of genes comprising the T_{EM} signature in both SLE (**d,** n = 26) and healthy Caucasian controls (**e,** n = 22). GSEA plots showing lack of enrichment for genes shown to increase transcription in response to T cell activation (**f**)[5] or those varying in response to dexamethasone treatment (**g**)[6].

Supplementary Figure 10



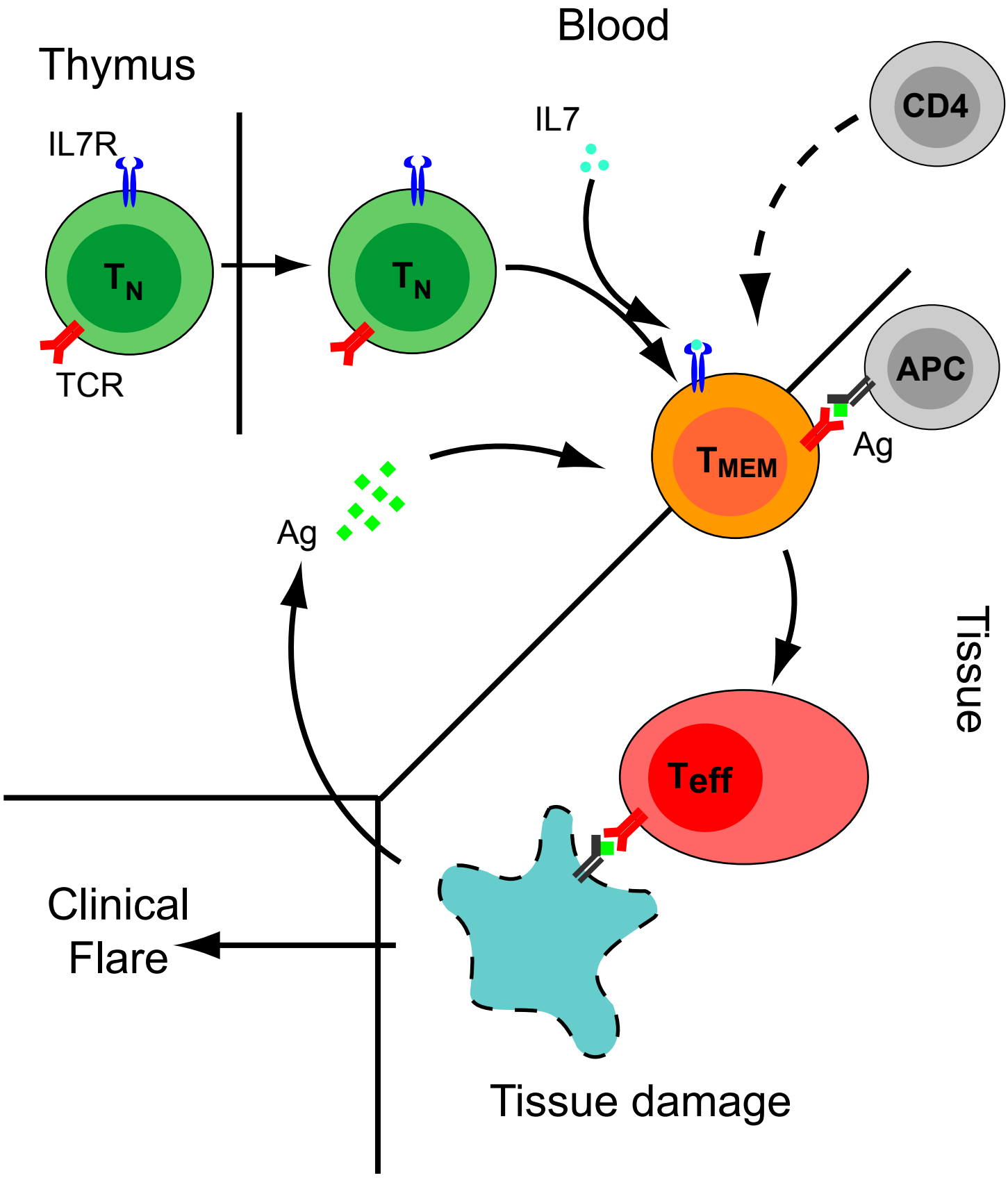
Supplementary Figure 10: Changes in both T_N and T_{MEM}-specific gene expression contribute to the signature defining the prognostic subgroups.

(**a**) Enumeration of CD8 memory populations was performed on a cohort of 24 healthy Caucasian controls who underwent concurrent gene expression profiling on Affymetrix Gene 1.0 ST arrays. Unsupervised hierarchical clustering illustrates 2 clear subgroups which are ordered by their statistical 'distance' from the dividing node in the dendrogram (distance metric = Pearson correlation with average linkage). (**b**) A positive correlation was thus seen between subgroup identity and CD8 T_{MEM} population size (p=0.01). (**c**) A significant difference in median % T_{MEM} was seen between groups 8.1 and 8.2, albeit with overlapping distributions (Mann-Whitney p = 0.048, bars = mean +/- SEM). **(d)** Venn diagram illustrating the contribution of T_N and T_{EM}-specific genes to the differentially-expressed genes (FDR<0.05, fold-change 1.5) defining subgroups v8.1 and v8.2. Red and green circles include only genes defined as characteristic of T_N or T_{EM} respectively (and therefore by definition mutually exclusive genesets) [7] with many in the unenclosed blue circle common to both. **(e)** Bcl2 protein expression in CD8 T cell populations in subgroup v8.1 (red bars) and v8.22 (blue bars), MFI = median fluorescence intensity. Significance determined using the Mann Whitney U test.

(f, g) Expression density distribution profiling was performed on genes found to overlap between the 8.1 v 8.2 signature and the T_N signature (**f,** dark red bars) or T_{EM} signature (**g**, dark green bars). Colours used are as for the Venn diagram in **d**. Overlapping genes were found to be predominantly bimodally expressed, whereas the expression of those for the total T_N (**f**, light red bars) and total T_{EM} (**g**, light green) signatures was predominantly normally distributed. **(h, i)** Representative comparison of gene expression profiles from

CD8 T cells to model distribution templates. Bimodal distributions are illustrated for genes of the 8.1v2 prognostic signature that are (**h**, *BACH2* and *MAML2*), or are not (**i**, *BCL2* and *LCK*), characteristically expressed preferentially (but not exclusively) in naive CD8 T cells relative to memory subsets. Each gene was compared to a range of possible distribution 'shapes' and matched to the best fit, as described in **Supplementary Methods**. To facilitate matching of expression data to idealized distributions, values were median-centred and scaled to unit variance prior to analysis, hence x-values represent scaled, transformed expression units, y-axis represents expression density. Solid line = template bimodal expression distribution, dashed line = actual expression distribution for a given gene.

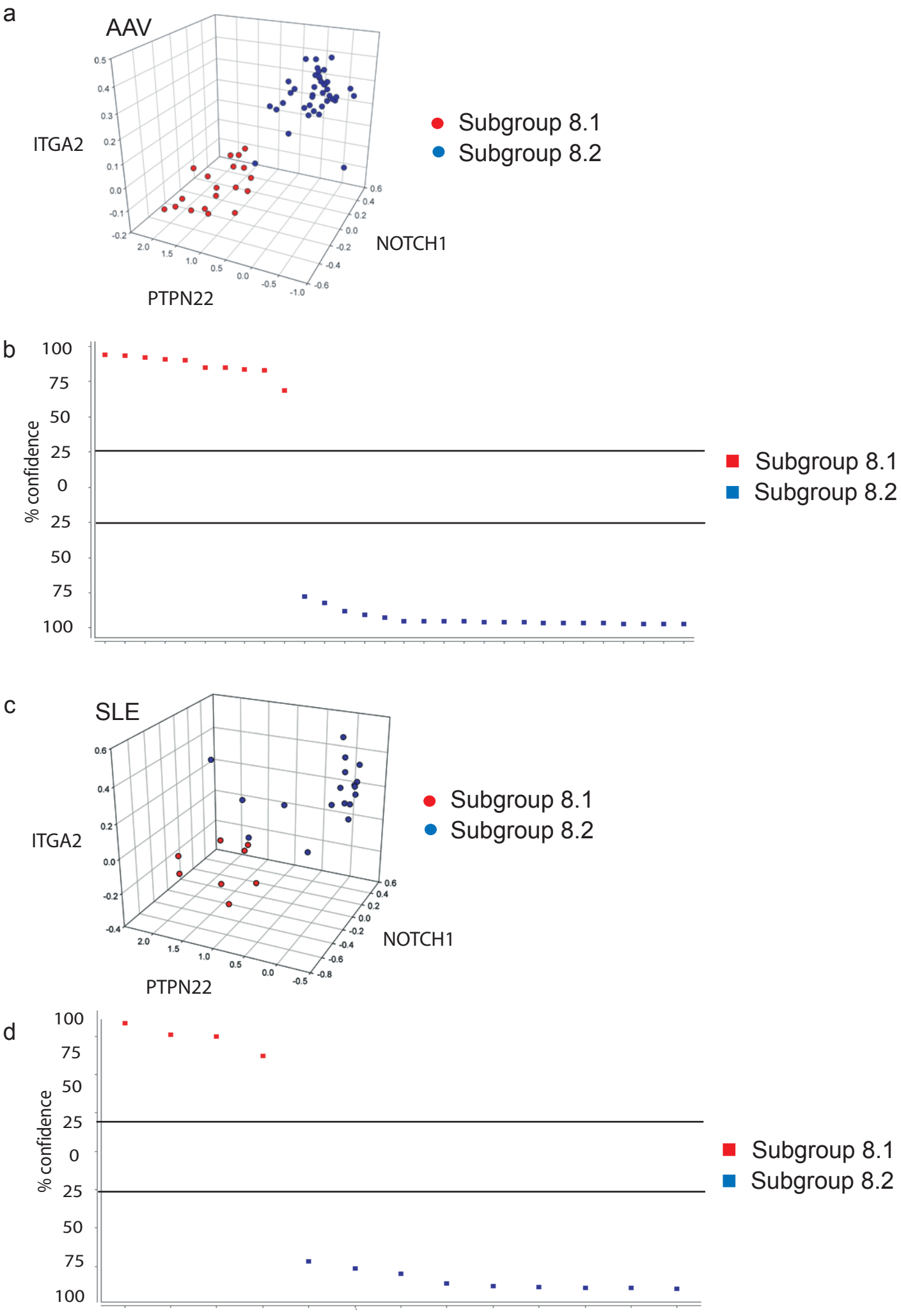
Supplementary Figure 11



Supplementary Fig. 11: Proposed model illustrating altered T_{MEM} populations, in turn driving T effector populations, disease relapses and tissue damage.

Naïve T cell (T_N) defects or interactions with other cell subsets (e.g. CD4 T cells or antigen-presenting cells, APC) may encourage the formation of a CD8 T cell memory response (T_{MEM}) on exposure to (auto)antigen, aided and defined by high levels of IL7R expression at the time of activation. Such memory cells would persist over time, expanding at lower thresholds of antigenic stimulation to produce the CD8 effector population which drives tissue damage responsible for clinical disease, while releasing further autoantigen to reinforce the memory response.

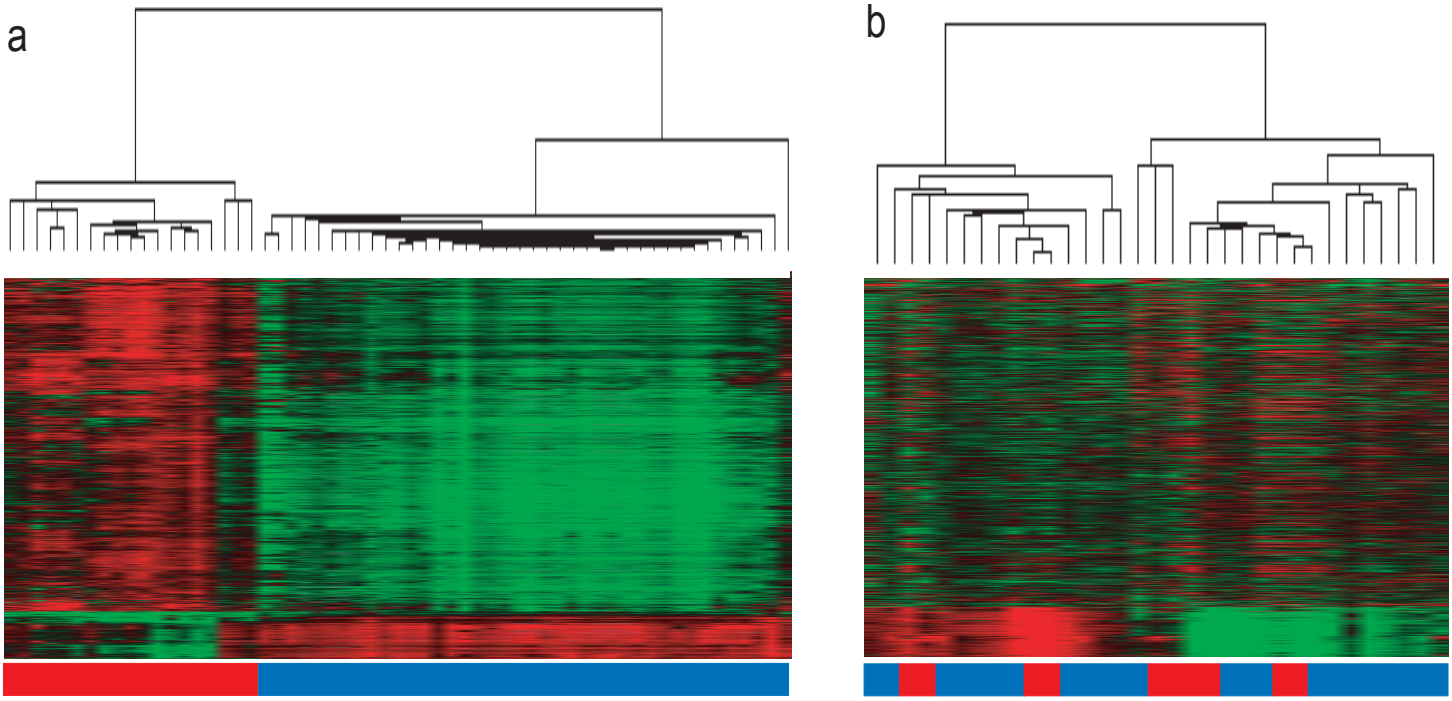
Supplementary Figure 12



Supplementary Fig. 12: A predictive model based on the expression of the same 3 genes robustly identifies prognostic subgroups in both AAV and SLE

3D scatterplots illustrating the distribution of AAV (**a**, n=59) and SLE (**c**, n=26) patients by expression of three CD8 T cell memory-related genes (ITGA2, PTPN22 and NOTCH1) which comprise an optimised predictive model. Such a model was generated on a stratified 50% subset of the total AAV cohort and applied to the remaining 50% as an independent test set using a Support Vector Machines algorithm [8]. This confirms confident, accurate prediction of both subgroups (PPV 100%, NPV 100%) in both AAV (**b**) and SLE (**d**).

Supplementary Figure 13



Acute presentation, untreated, active disease

Follow-up on treatment, inactive disease

Supplementary Fig. 13: The same prognostic subgroups are not seen in a cross-section of the same vasculitis patients with quiescent disease on maintenance immunotherapy. **(a)** Supervised hierarchical clustering of the AAV cohort (n = 59) at time of enrolment, all patients with active disease on minimal therapy as illustrated previously. **(b)** Supervised hierarchical clustering of a subset of the same patients (n = 33) at least 12 months after initial presentation with quiescent disease on maintenance immunosuppressive therapy does not reproduce the same subgroups.