nature immunology

SUPPLEMENTARY INFORMATION

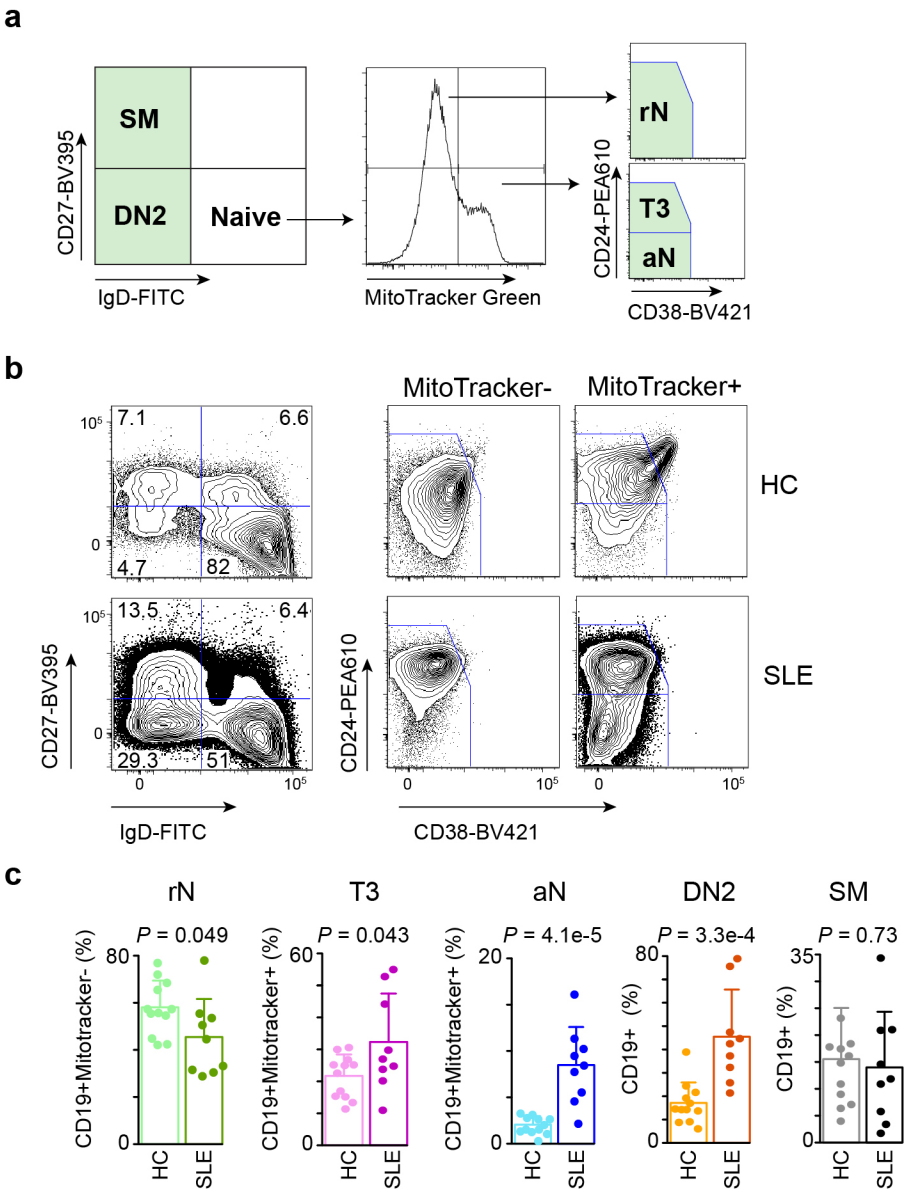
https://doi.org/10.1038/s41590-019-0419-9

In the format provided by the authors and unedited.

Epigenetic programming underpins B cell dysfunction in human SLE

Christopher D. Scharer[1], Emily L. Blalock[2,3], Tian Mi[1], Benjamin G. Barwick[4], Scott A. Jenks[2,3], Tsuneo Deguchi[2,3], Kevin S. Cashman[2,3], Bridget E. Neary[2,3], Dillon G. Patterson[1], Sakeenah L. Hicks[1], Arezou Khosroshahi[2,3], F. Eun-Hyung Lee[3,5], Chungwen Wei[2], Iñaki Sanz[2,3]* and Jeremy M. Boss[1]*

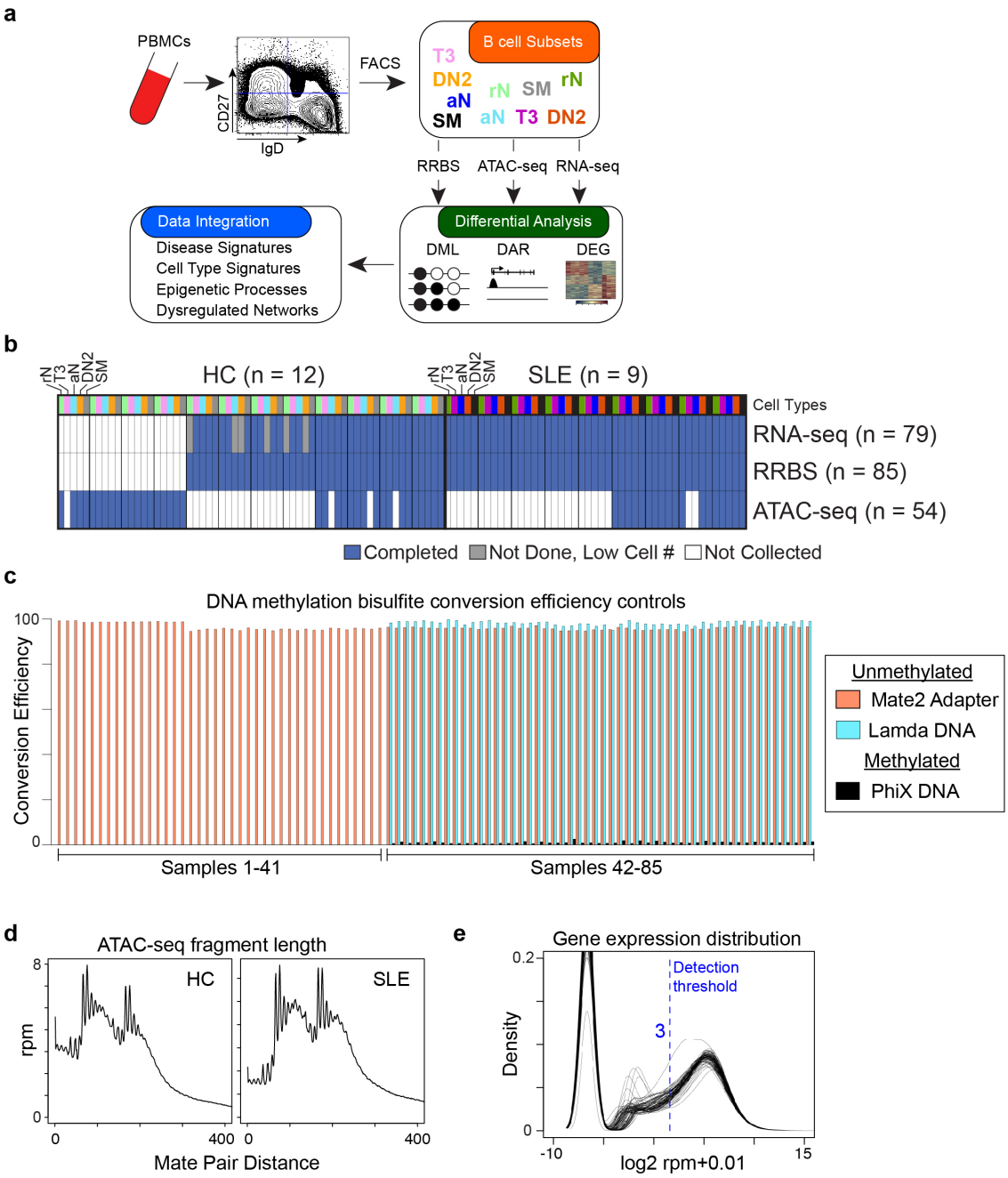
[1]Department of Microbiology and Immunology, School of Medicine, Emory University, Atlanta, GA, USA. [2]Division of Rheumatology, Department of Medicine, School of Medicine, Emory University, Atlanta, GA, USA. [3]Lowance Center for Human Immunology, School of Medicine, Emory University, Atlanta, GA, USA. [4]Department of Hematology and Medical Oncology, School of Medicine, Emory University, Atlanta, GA, USA. [5]Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Department of Medicine, School of Medicine, Emory University, Atlanta, GA, USA.
*e-mail: ignacio.sanz@emory.edu; jmboss@emory.edu



Supplementary Figure 1

aN and DN2 B cell subsets are expanded in SLE subjects.

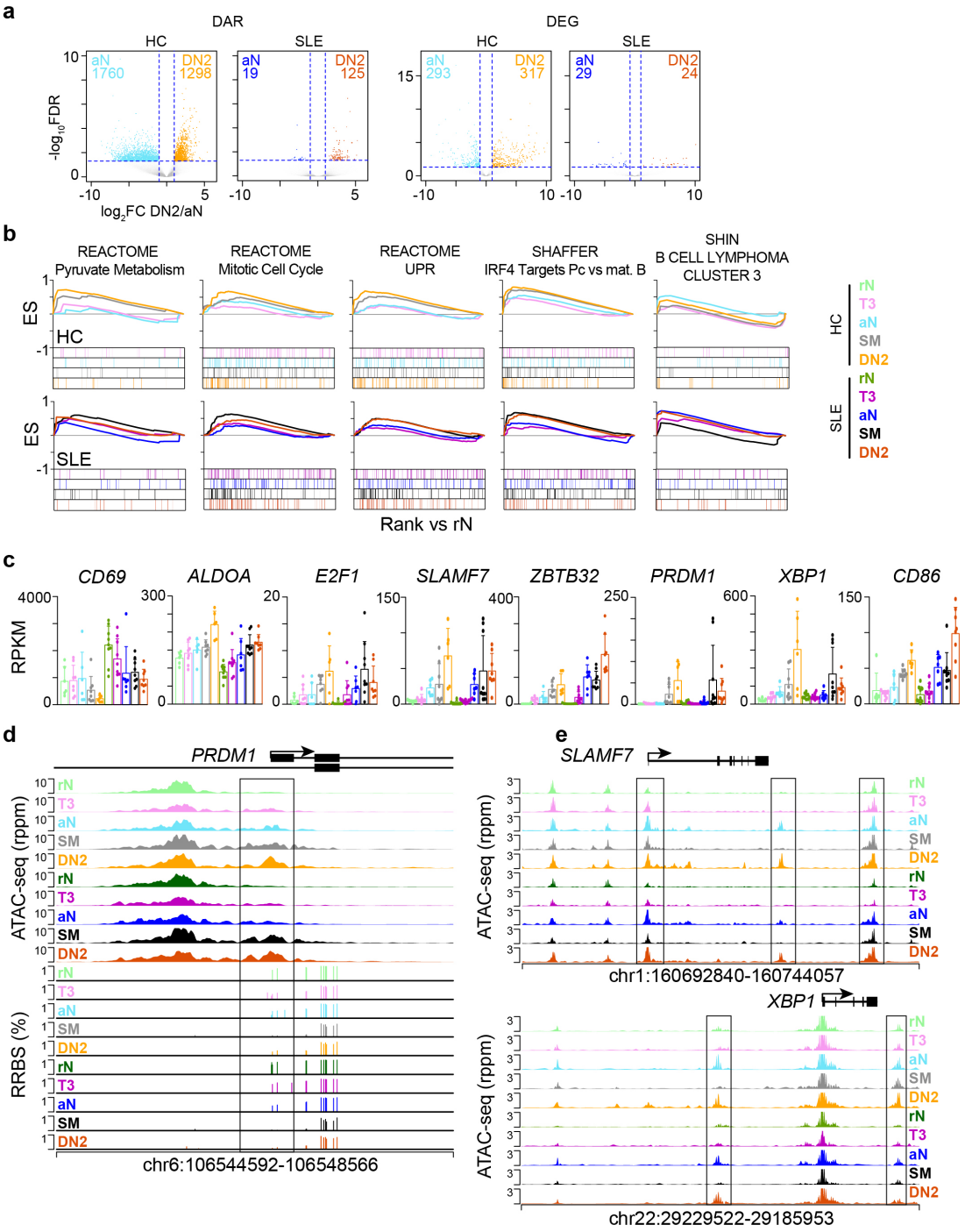
(a) Schematic showing gating strategy used to define B cell subsets. **(b)** Flow cytometry data for a representative HC and SLE subject from one experiment. Sample sizes for each cell type can be found in Supplementary Table 5. **(c)** Bar plot showing the frequency of each cell subset defined in **a** between HC and SLE subjects from one experiment. Each subject is denoted by a dot and the mean ±SD is shown. Significance determined by two-tailed Student's *t*-test.



Supplementary Figure 2

Sequencing and QC of B cell subsets.

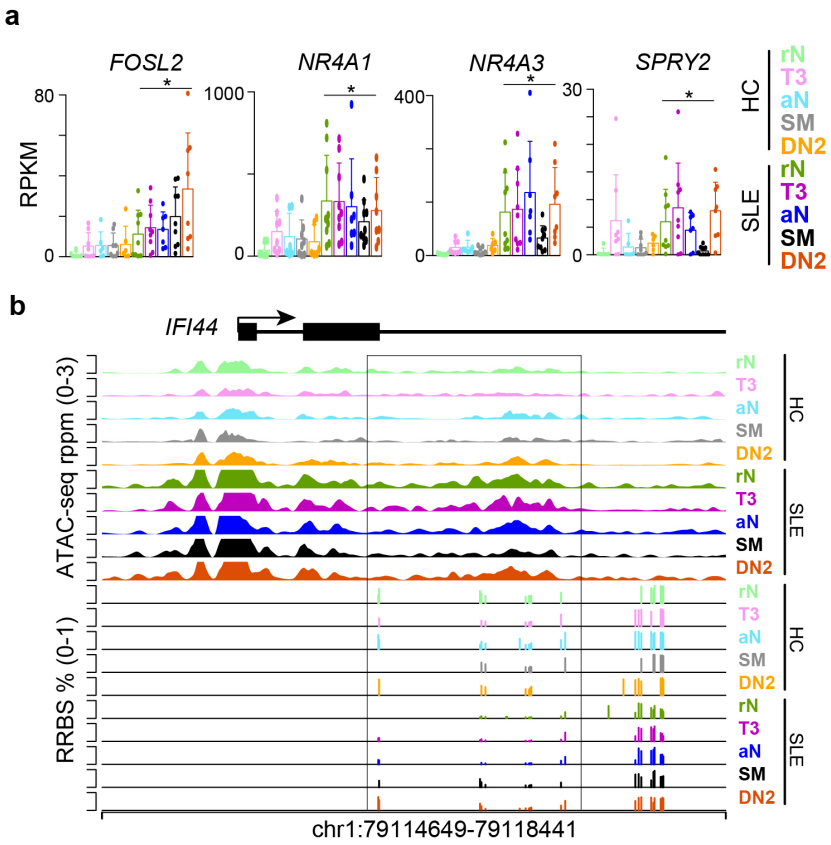
(a) Schematic and workflow of cell isolation and processing. **(b)** Annotation of data sets collected for each subject and cell type for each of the three genomic assays performed. **(c)** Bar plot showing the conversion efficiency of methylated and unmethylated DNA methylation libraries. **(d)** Representative histogram showing distance between paired-end reads for ATAC-seq data from one experiment. Similar results were obtained from all ATAC-seq samples. **(e)** Density plots of transcript expression for all RNA-seq libraries with the detection threshold annotated.



Supplementary Figure 3

Progressive upregulation of gene sets associated with B cell differentiation.

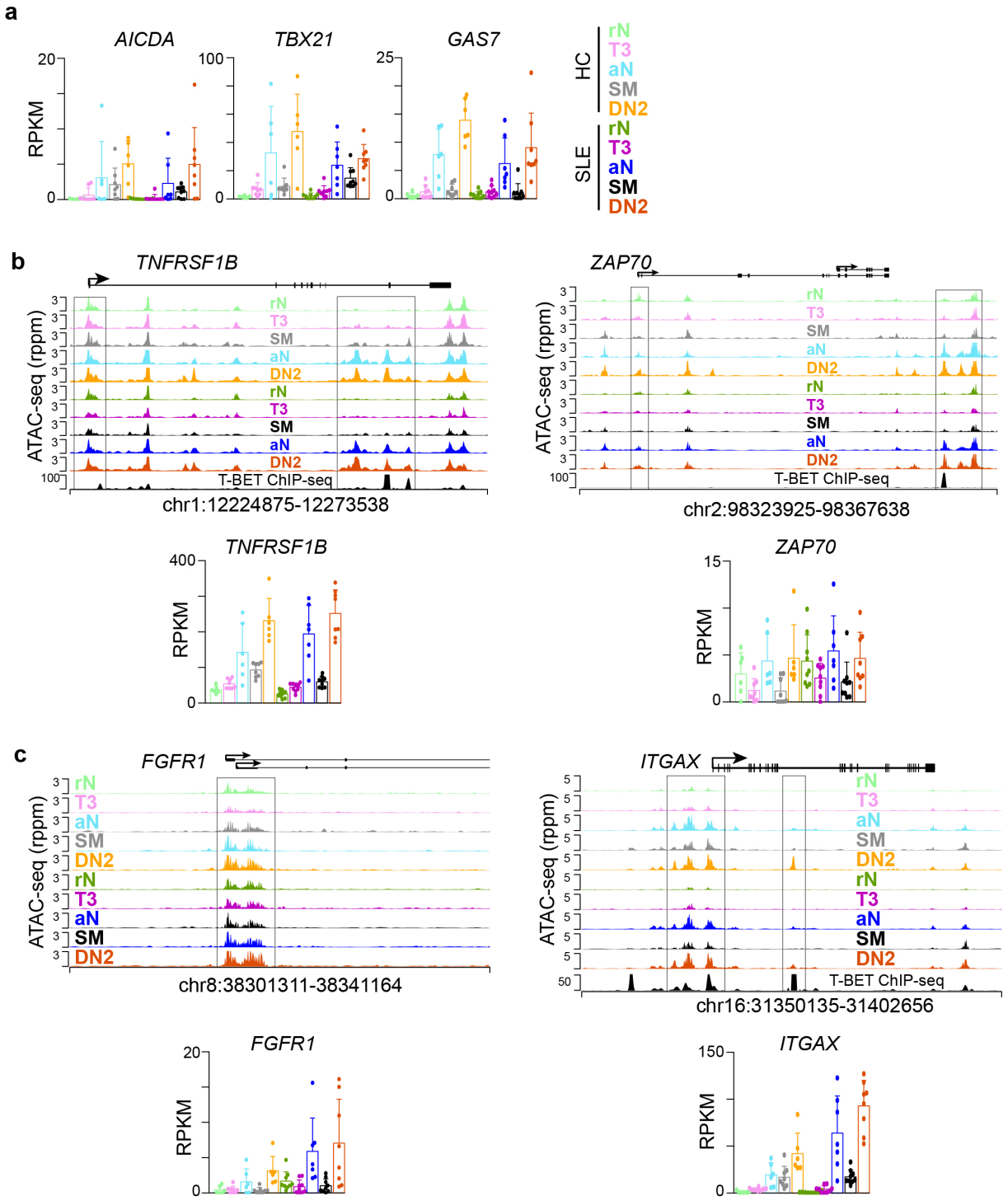
(a) Volcano plot of DAR and DEG comparing DN2 vs. aN B cells from HC (left) and SLE (right). The number of differential features is indicated. DEG and DAR represent features with >=2-fold change and FDR < 0.05 as determined by edgeR. **(b)** GSEA plots of gene sets displayed in Fig. 1d depicting the enrichment for HC (top) and SLE (bottom) cell types. **(c)** Bar plot of gene expression levels for the indicated gene. Data represent mean ±SD. **(d)** Genome plot showing the accessibility and DNA methylation levels at the *PRDM1* locus. The location of DAR and DML is highlighted with a box. **(e)** Genome plot of the indicated locus showing the accessibility pattern for each cell type. The location of DAR is highlighted with a box. Data from **d-e** represent the mean for each cell type from one experiment.



Supplementary Figure 4

Coordinated changes in accessibility and gene expression in rN B cells.

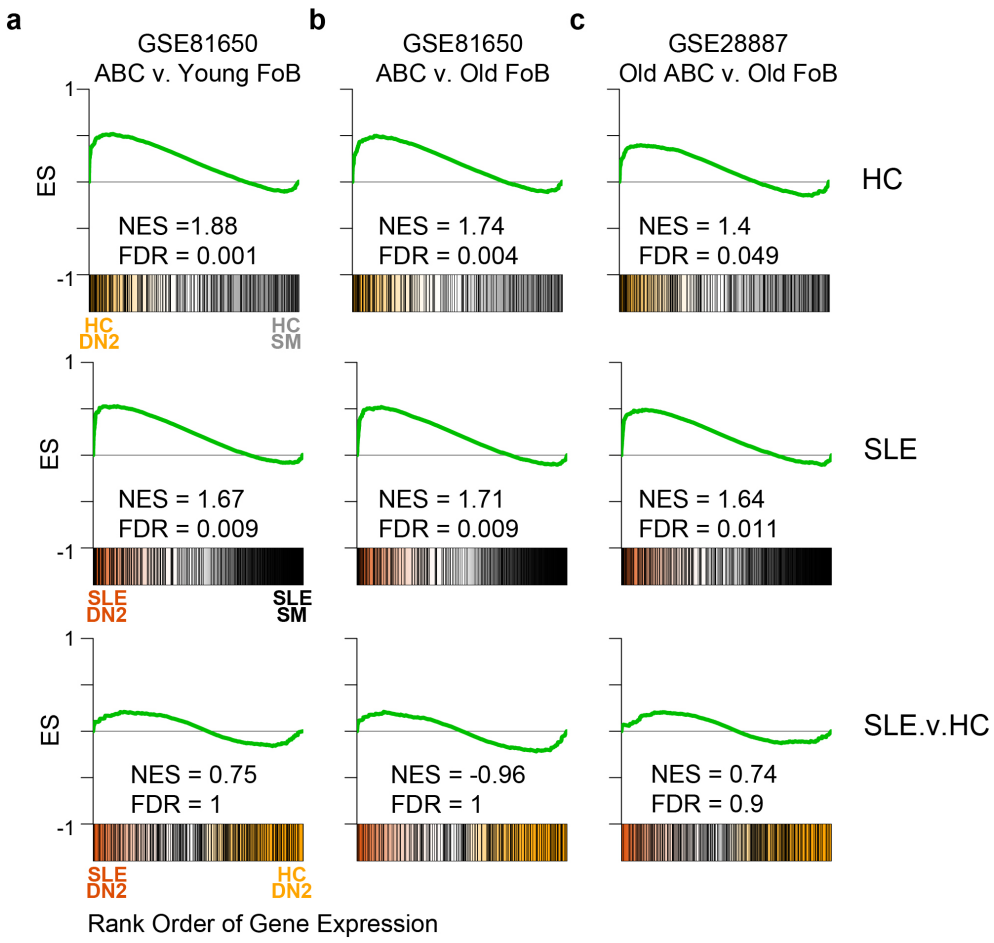
(a) Bar plot of gene expression levels for the indicated gene. Data represent mean ±SD. * indicates DEG between SLE and HC (>=2-fold change and FDR < 0.05) as determined by edgeR. **(b)** Genome plot showing the accessibility and DNA methylation levels at the *IFI44* locus. Boxed region contains a DAR and DML between SLE and HC. Data represent the mean for each cell type from one experiment. See also Fig. 2.



Supplementary Figure 5

Gene expression and chromatin accessibility changes in DN2 cells are shared with aN.

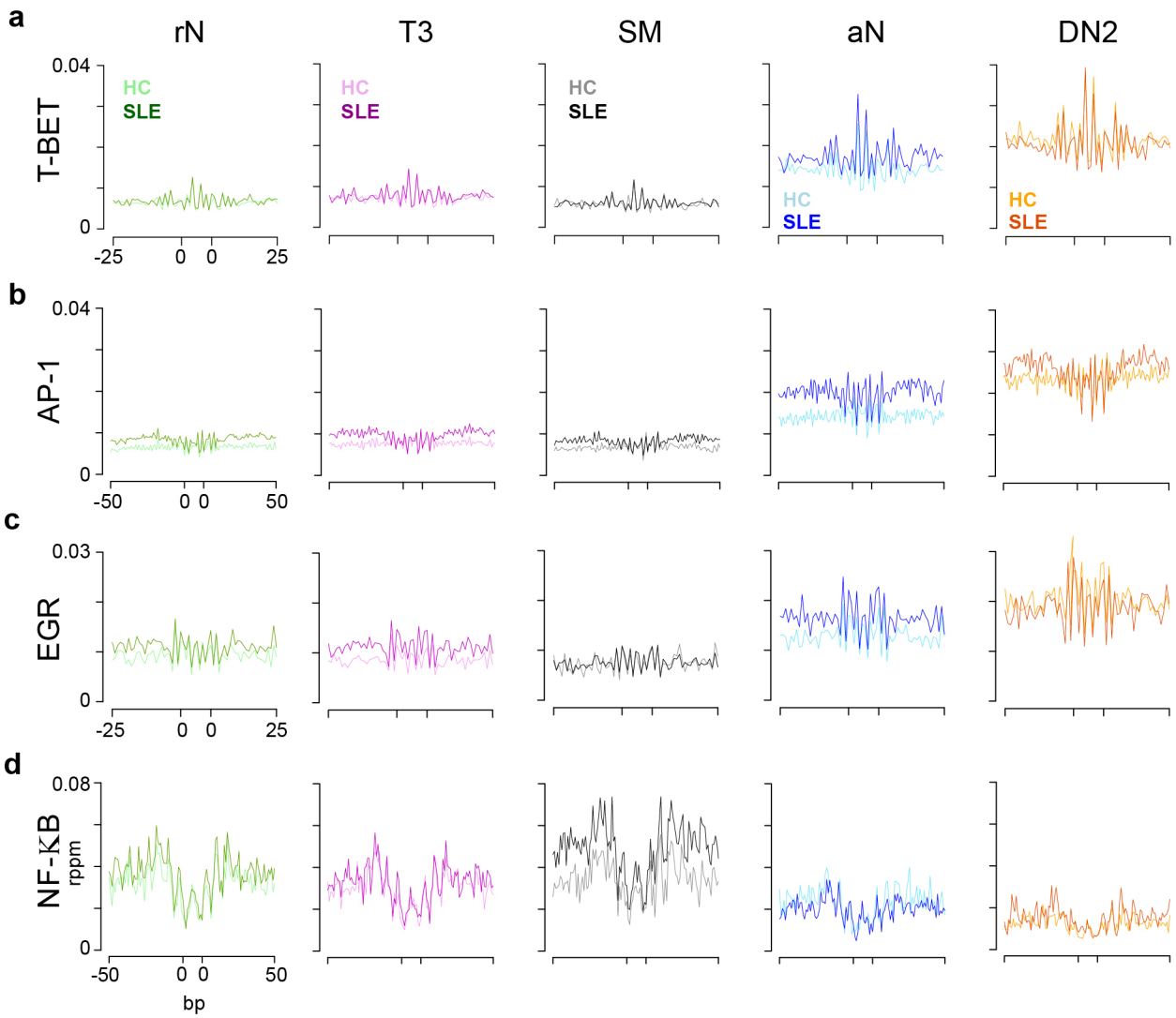
(a) Bar plot of gene expression levels for the indicated gene. Data represent mean ±SD. For each indicated gene, a genome plot (top) showing the accessibility of the locus and bar plot of gene expression (bottom) at loci that are shared with HC **(b)** or unique to SLE DN2 B cells **(c)**. DAR between DN2 and SM are highlighted in a box. Gene expression data represent mean ±SD. Genome plot data for **b-c** represent the mean for each cell type from one experiment. T-BET binding in GM12878 B cells is previously reported[1]. See also Fig. 4.



Supplementary Figure 6

The ABC signature is enriched in both SLE and HC DN2 B cells.

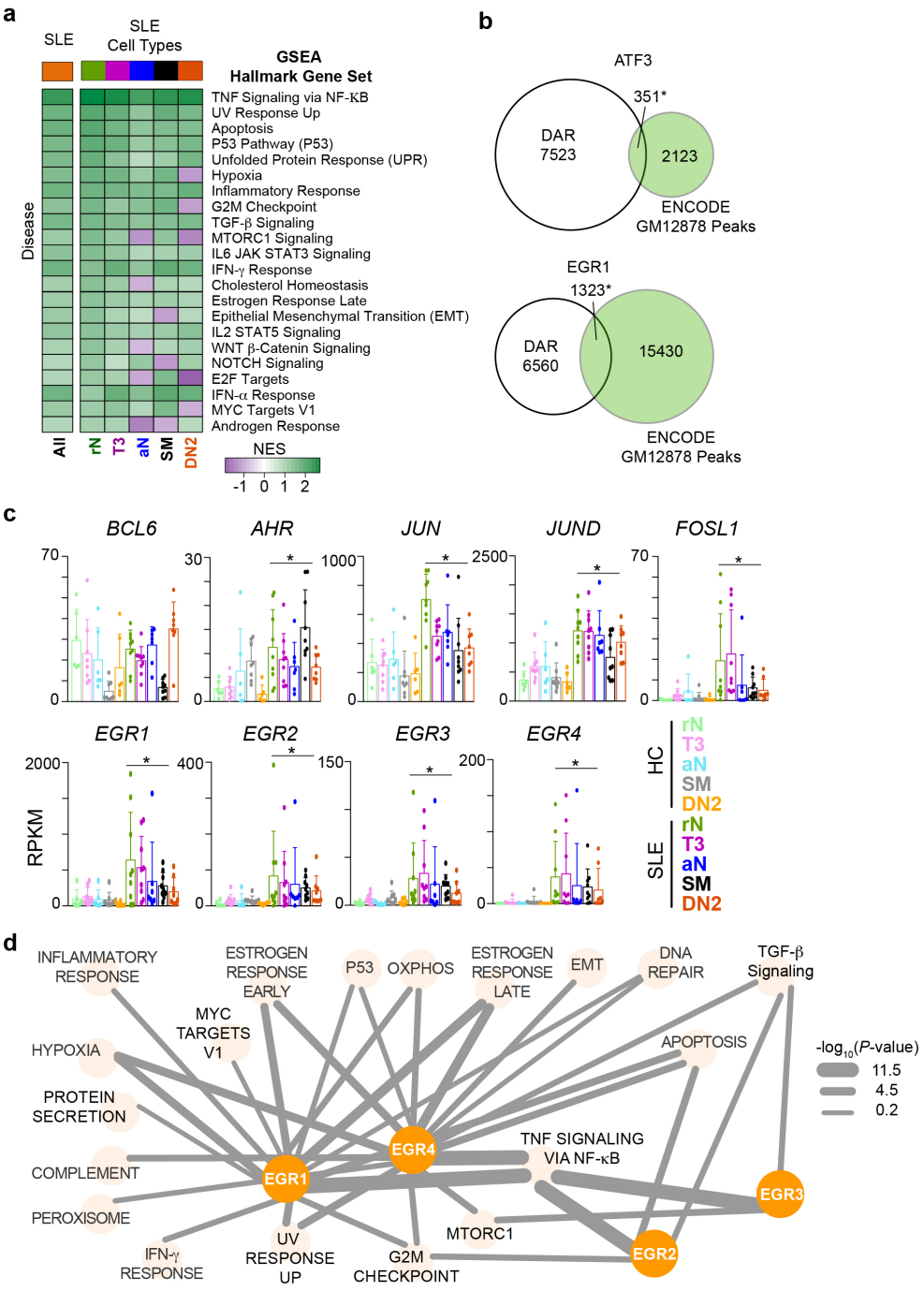
GSEA of the comparing the HC DN2 versus HC SM (top), SLE DN2 versus SLE SM (middle), or SLE DN2 versus HC DN2 (bottom) for enrichment with ABC datasets. Gene set comparing **(a)** ABC versus young follicular B cells (FoB)[2], **(b)** ABC versus old FoB [2], and **(c)** old ABC versus old FoB [3]. FDR < 0.05 was considered significant using the Benjamini-Hochberg correction on the *P*-value derived from permutation testing.



Supplementary Figure 7

DN2 and aN B cells have similar transcription factor accessibility footprints.

Histogram of accessibility for the indicated range surrounding **(a)** T-BET, **(b)** AP-1, **(c)** EGR, and **(d)** NF-κB motifs in the indicated B cell subset (columns). For each B cell subset the HC and SLE sample is shown. rppm, reads per peak per million. See also Fig. 5b.



Supplementary Figure 8

Transcription factor and gene set enrichment in SLE.

(a) Heatmap of normalized enrichment score (NES) calculated by GSEA for pathways up regulated in all SLE cell types (left) or within each cell type (right). For each gene set the NES for each cell type compared to the HC counterpart is annotated. See also Fig. 6b. **(b)** Venn diagram showing the overlap of ChIP-seq peaks for ATF3 (top) and EGR1 (bottom) from the ENCODE Consortium[1] with DAR between HC and SLE B cells. * indicates *P*-value < 0.0001 based on randomly permuting the DAR 10,000 times. **(c)** Bar plot of gene expression levels for the indicated gene. Data represent mean ±SD. * indicates DEG between SLE and HC (>=2-fold change and FDR < 0.05) as determined by edgeR. See also Fig. 6e. **(d)** Network diagram depicting the gene sets targeted by each EGR factor. Line thickness is scaled to the significance as determined by Fisher's Exact test. See also Fig. 6g.

References

1. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).

2. Russell Knode, L.M. *et al.* Age-Associated B Cells Express a Diverse Repertoire of VH and Vkappa Genes with Somatic Hypermutation. *J Immunol* **198**, 1921-1927 (2017).

3. Rubtsov, A.V. *et al.* Toll-like receptor 7 (TLR7)-driven accumulation of a novel CD11c(+) B-cell population is important for the development of autoimmunity. *Blood* **118**, 1305-1315 (2011).