Complement C3d enables cell-mediated immunity capable of distinguishing spontaneously transformed from non-transformed cells

Authors: Jeffrey L Platt[1,2*], Chong Zhao[1,2], Jeffrey Chicca[1], Matthew Pianko[3], Joshua Han[1,2], Stephanie The[4,5], Arvind Rao[4,5], Evan Keller[6], Mayara Garcia de Mattos Barbosa[1], Lwar Naing[1,8], Tracy Pasieka-Axenov[1], Lev Axenov[1], Simon Schäfer[9], Evan Farkash[7], and Marilia Cascalho[1,2,*].

Extended data figures:

SI Appendix, Figure S1 A. C3d peptide sequence and purification. The C3d peptide sequence encompasses amino acids 1024 to 1320 of murine C3, numbered as C3d residues (1-299) in the figure. The sequence includes a N-terminal protein transduction domain (RKKRRQRRR) and Cys (C) to Ser(S) substitution at position 5 (1028 of C3) to avoid the presence of a free sulfhydryl in the recombinant peptide. The C3d sequence was cloned into a baculovirus expression vector and used to transduce Sf9 insect cells. Large-scale production for purification was done in High-five infected at an MOI of 2. Purification was done using NTA Agarose from Qiagen and eluted with imidazole in PBS. The samples were analyzed on a 4-20% gradient SDS-PAGE. The monomer was detected in the soluble (shown) and in the insoluble fractions (not shown). Aliquots of 1mg/ml were prepared and mice were injected with 20 microliters, containing 20 micrograms of peptide in PBS intra tibia. **B**. Detection of M protein. Vk-MYC* mice were injected with C3d or PBS. Two weeks later, mice treated with C3d were either injected with anti-CD8 antibodies or with isotype control three times a week starting two weeks after C3d injection and continued for 6 weeks until experiment end. M protein was detected by SPEP before and at the end of the experiment. M protein seen as a single narrow band in the gamma region of the gel decreases in mice treated with C3d. Instead, a broad band of polyclonal Ig appears. In untreated mice and in mice treated with C3d followed by anti-CD8 antibodies the narrow band increases in intensity relative to the albumin band and remains narrow. **C.** Graph shows the relative change in M protein intensity, following each treatment and relative to albumin. M protein density was measured using image J.

SI Appendix, Figure S2: C3d decreases tumor load without changing other cell populations in the BM. Vk*myc mice older (late disease) and younger (early disease) than one year of age, with M-protein, were injected with 20 microgram C3d peptide intra-tibia two months before analysis. Vk*myc were immunized twice, 2 months apart before treatment with C3d, M protein was detected by SPEP. **A and B.** Tsne plots of mass cytometry (CYTOF) indicating plasma cell clusters in purple. **C.** The graph depicts frequencies of plasma cells in the treated or control mice divided according to age at beginning of treatment assessed by mass cytometry. **D.** Analysis of the frequency of non-plasma cell bone marrow cell populations by mass cytometry. Frequencies of the total number of cells are indicated. Data

analysis was by Mann-Whitney test. Comparisons that yielded $P<0.05$ are noted by **, P=0.0022; *, P=0.0286. Statistics analysis was by Prism software (v9 for mac).

SI Appendix, Figure S3: UMAP plots of sc RNA gene expression depicting the different cell populations. A. UMAP plots of every sample considered in the analysis. B. Gene expression of markers used to differentiate between cell populations. Y-axis depicts the names of the cell types; the X-axis depicts the gene identity. Expression is noted according to the number of cells that express the marker above background (diameter of the circle increases with the fraction of cells expressing the marker). The average intensity is represented by a color gradient varying from white (no expression) to purple (mid intensity expression) and to blue (maximum expression).

SI Appendix, Figure S4: **IgVH gene frequencies in B cells in the bone marrow of Vk-MYC* mice treated (red) or not (black) with C3d.** Differences in frequency according to treatment did not reach significancy by multiple Mann-Whitney tests.

SI Appendix, Figure S5: C3d treatment induces expression of genes encoding ribosomal proteins and LncRNAs *in vivo* and in B cells in culture. B cell lymphoma cells (18.81) (1) (10^5/ml) were incubated with 20mg C3d, with PBS for 24h, 48h and 72h and RNA obtained and sequenced. For immunofluorescence, 10^6 cells were incubated with 200 mg of C3d with an internalization domain (ID) or with C3d lacking ID. **A.** Immunofluorescence analysis of 18.81 B cells in cytospin slides stained with goat anti-mouse C3d followed by Texas-red labelled rabbit anti-goat antibodies. Bar indicate 50 microns. C3d with ID is incorporated by lymphoma B cells after one hour incubation while C3d without ID accumulates on the cell surface. Included are graphs depicting measurements of cell fluorescence intensity (raw integrated intensity, as the areas were the same), upper graph, and of fluorescence skewness that reflects the distribution of the pixel intensities within each cell, lower graph. Analysis by Mann-Whitney test revealed that while overall fluorescence intensity did not differ between the two groups. However, cells treated with C3d without ID showed increased skewness relative to cells transduced with C3d +ID (P=0.0083). The elevated skewness means the image contains more low-intensity pixels (darker regions), with fewer higher-intensity pixels (brighter regions) and this is consistent with an accumulation of C3d on the surface of the cell. **B.** C3d induced differential expression of genes in plasma cells of bone marrow of treated mice relative to untreated mice. Figure shows DEGs that that are also induced by C3d on 18.81 cells relative to untreated cells, after 48 hours incubation. Values are expressed as Log_2 Fold Change. Negative values mean decreased expression. Positive values mean increased expression. Changes in expression considered had an adjusted $P<0.001$ (ST6 and ST7). **C.** Volcano plot analysis of differential gene expression by 18.81 B cells incubated with C3d for 48 hours. In red are depicted all the differentially expressed lncRNAs. C3d changed the expression of 1093 genes with $P<0.001$ relative to untreated cells (ST7). **D.** Top 20 lncRNAs increased in 18.81 cells incubated with C3d for 48 hours in comparison with untreated cells.

SI Appendix, Figure S6: E2f1 Pathway is induced by C3d. Lymphoma 18.81 cells (1) (10^5/ml) were incubated with 20µg C3d or with PBS for 24h, 48h and 72h. Following incubation cells were washed, RNA produced, sequenced and processed at the University of Michigan Advanced Genomics Core. **A.** Schematics of E2f1 pathway analyzed by String (https://string-db.org). Shown are genes differentially expressed by C3d relative to controls following C3d treatment of 18.81 B cells for 48 hours culture (P>0.05). Network nodes represent proteins. Colored nodes refer to first shell of interactors. Edges represent protein interactions, known from curated data bases or experimental data; predicted by gene neighborhood, co-occurrence, homologies and/or co-expression. Software available at string-db.org/cgi run at the highest confidence (0.900) to show a maximum of 5 interactions and analysis with DBSCAN clustering to show a maximum of three clusters (String, version 12.0, String consortium 2023, used under a Creative Commons BY 4.0 license). **B.** Heat map of E2f1 pathway differentially expressed by 18.81 cells incubated with C3d for 48 hours relative to controls. **(C)** RT-PCR analysis of E2f1 pathway gene expression in RNA obtained from lymphoma cell cultured with or without C3d for 24 or 48 hours. We compared gene expression obtained at 24 hours with gene expression obtained after 48 hours incubation with C3d relative to controls in RT-PCR experiments by paired T test. The exact P value was 0.0067 (**).

SI Appendix, Figure S7: C3d treatment increases MHC class-I expression by 18.81 cells. A-B. MHC class I gene expression at 24h, 48h and 72h after incubation with C3d or in control cultures. H-2Kd expression was analyzed by FACS. Plots depict live cells gated on lymphocytes. Shown in **(A)** is H-2Kd MFI at 48h of culture in non-dividing cells. Shown in **(B)** is the graphic representation of the MFIs of three replicate cultures with C3d or untreated after, 24 or after 48 and 72 hours. Statistical analysis was by ANOVA followed by Kruskal-Wallis test (Prism for Mac v10.2.3).

SI Appendix, Figure S8: Differential gene expression in the bone marrow of mice treated with C3d relative to untreated mice. Volcano Plot on the left depicts the top DEG expressed by bone marrow cells annotated according to the cell of origin. Volcano plot on right depicts DEG expressed by Plasma Cells (PCs) in the bone marrow cells of mice treated with C3d or left untreated. Positive fold changes correspond to genes that increase in expression by C3d while negative fold changes correspond to genes that decrease in expression following C3d treatment. Changes in expression that had an adjusted P <0.05 were included. Data plotted can be found in ST6.

SI Appendix, Supplementary Data Tables:

SI Appendix, ST-primers. Taq Man primers and probes used to detect E2f1 pathway gene expression.

SI Appendix, ST 1. Single cell RNA gene expression per cluster comparing each cluster to all the others to differentiate clusters, indicating the Log_2 of the fold change and adjusted P in all the mice.

SI Appendix, ST 2. V(D)J clonotype frequencies. A clonotype was defined as >80% CDR 3 AA sequence overlap with same VH and JH heavy chain genes.

SI Appendix, ST 3. Expression of malignancy associated genes by each cell to define malignant and non-malignant clones.

SI Appendix, ST 4. Principal Component Analysis based on expression of malignancy associated genes.

SI Appendix, ST 5. Malignancy and non-malignancy scores per cell.

SI Appendix, ST 6. Single cell RNA differential gene expression by every cell cluster comparing cells obtained from the bone marrow of C3d treated mice with cells obtained from untreated mice. Genes that are differentially expressed with an Adj P<0.001 are highlighted in yellow. A volcano plot shown in figure S8 illustrates differentially expressed genes in every cluster noting the 20 top C3d-induced genes. A separate plot is shown for genes differentially expressed by Plasma Cells (Figure S8).

SI Appendix, ST 7. Differentially expressed genes after 48 hours incubation of 18.18 pre-B cells with C3d or glycerol. Positive fold change values mean increased expression in C3d treated cells; negative values mean decreased expression in C3d treated cells. Genes that are differentially expressed with an Adj P<0.001 are highlighted in yellow.

SI Appendix, ST 8. Differential LncRNA gene expression after 48 hours incubation of 18.18 pre-B cells with C3d or glycerol.

SI Appendix, ST 9-10. Single Cell RNA differential gene expression by T cells from C3d treated mice in comparison to T cells obtained from the bone marrow of mice not treated with C3d. Log_2 transformed fold change and adjusted P are indicated. Positive values mean increased expression in C3d treated mice; negative values mean decreased expression in C3d treated mice. ST10 data is formatted for Volcano plot display.

SI Appendix, ST 11. Single Cell RNA differential gene expression by granulocytes obtained from the bone marrow of C3d treated mice in comparison to granulocytes obtained from the bone marrow of mice not treated with C3d. Log_2 transformed fold change and adjusted P are indicated. Positive values mean increased expression in C3d treated mice; negative values mean decreased expression in C3d treated mice.

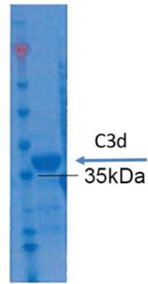
SI Appendix, ST 12. Volcano plot display of Single Cell RNA differential gene expression by granulocytes obtained from the bone marrow of C3d treated mice in comparison to granulocytes obtained from the bone marrow of mice not treated with C3d. Y axis depicts the Log_2 transformed fold change and the Y-axis depicts adjusted P values.

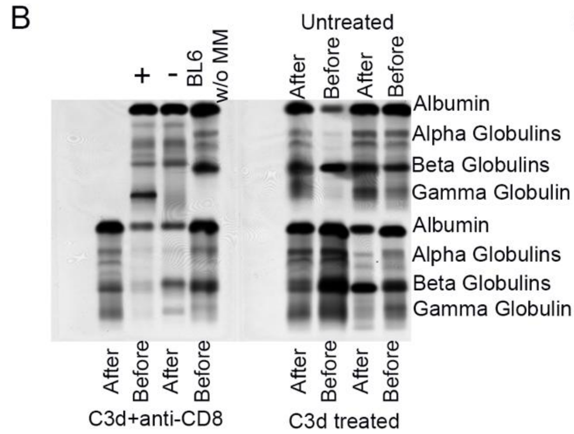
A

Purified recombinant C3d attached to a protein transduction domain (RKKRRQRRR)

```
1   TPAGSGEQNMIGMTPTVIAVHYLDQTEQWEKFGIEKRQEALELIKKGYTQQLAFKQPSSA 60
61  YAAFNNRPPSTWLTAYVVKVFSLAANLIAIDSHVLCGAVKWLILEKQKPDGVFQEDGPVI 120
121 HQEMIGGFRNAKEADVSLTAFVLIALQEARDICEGQVNSLPGSINKAGEYIEASYMNLQR 180
181 PYTVAIAGYALALMNKLEEPYLGKFLNTAKDRNRWEEPDQQLYNVEATSYALLALLLLKD 240
241 FDSVPPVVRWLNEQRYYGGGYGSTQATFMVFQALAQYQTDVPDHKDLNMDVSFHLPSRS  299
```

MW = 35.1 kDa
1mg/ml in PBS





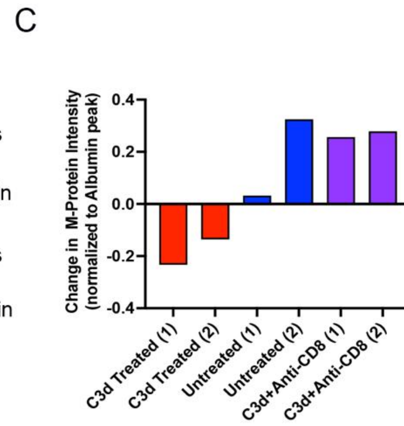


Figure S1

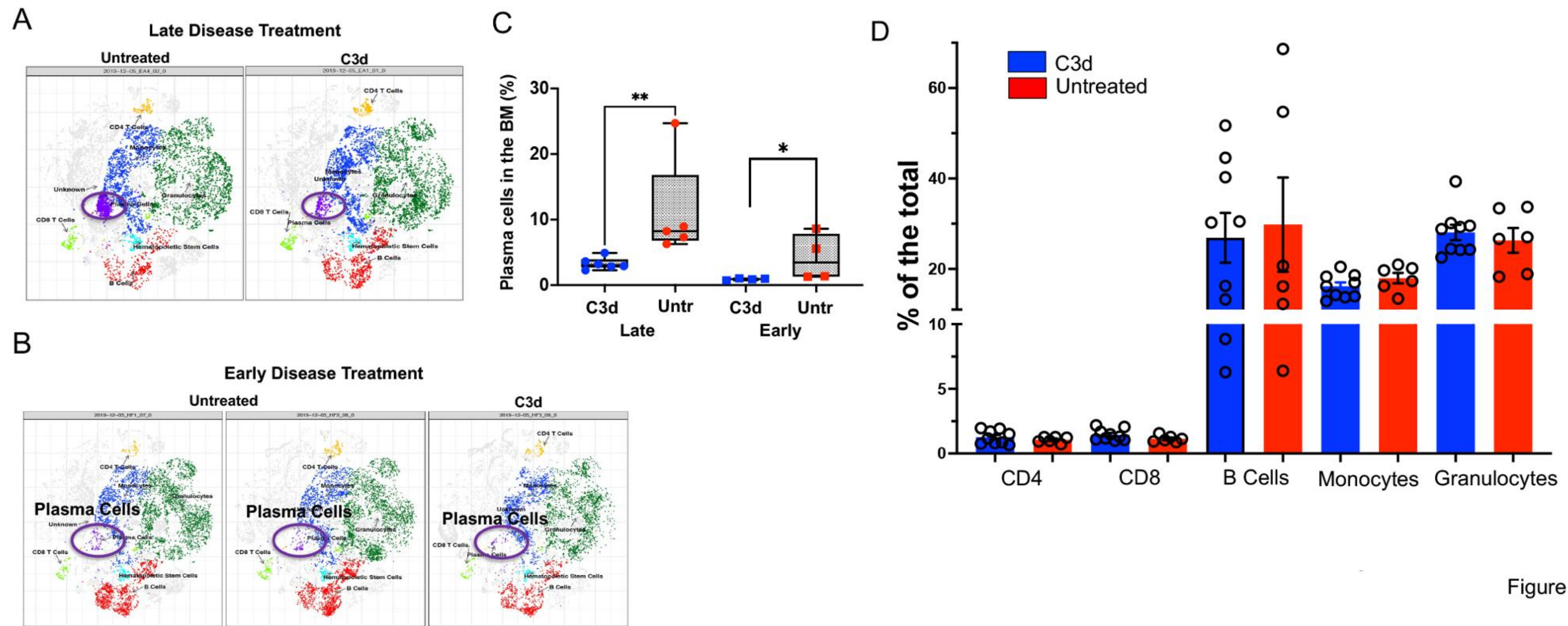


Figure S2

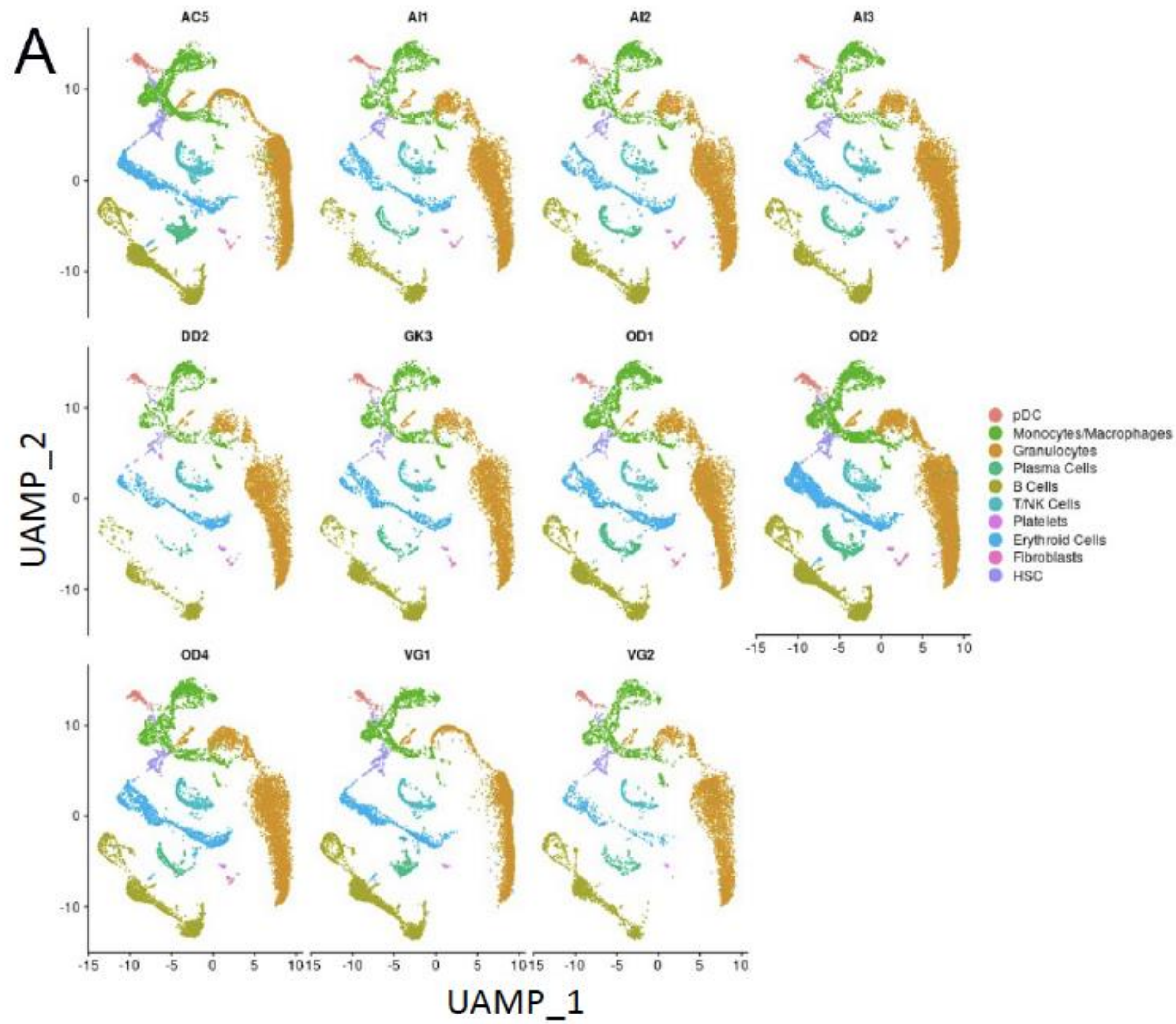


Figure S3A

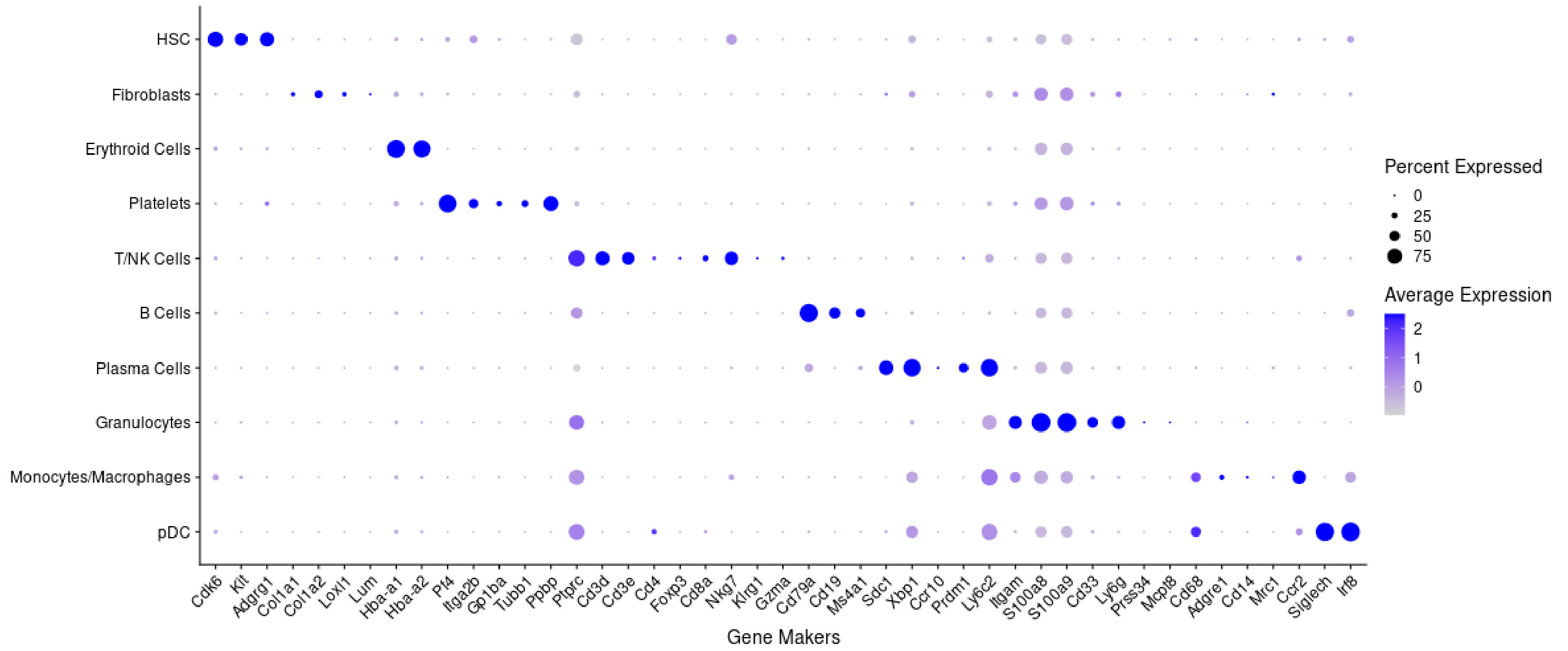


Figure S3B

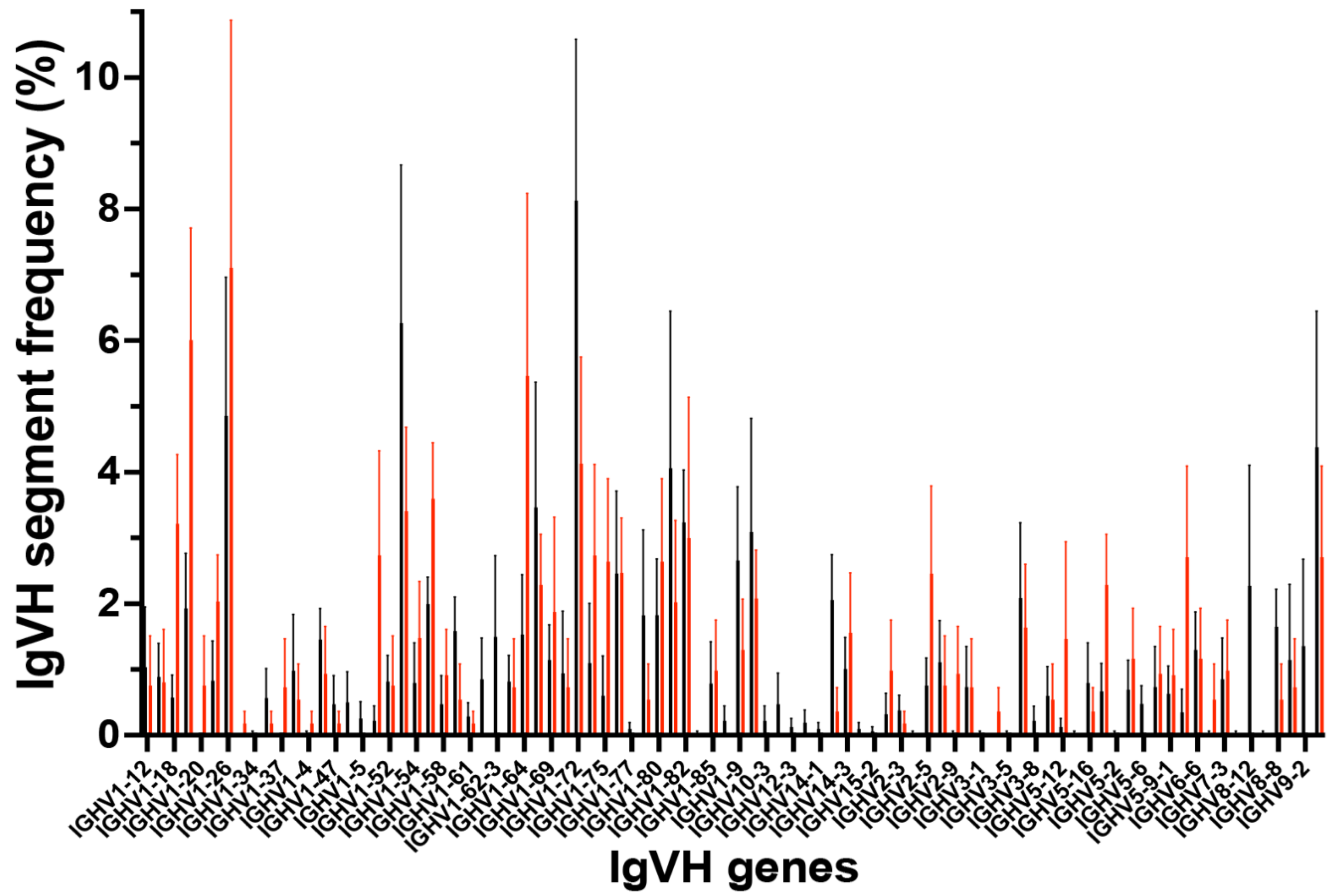
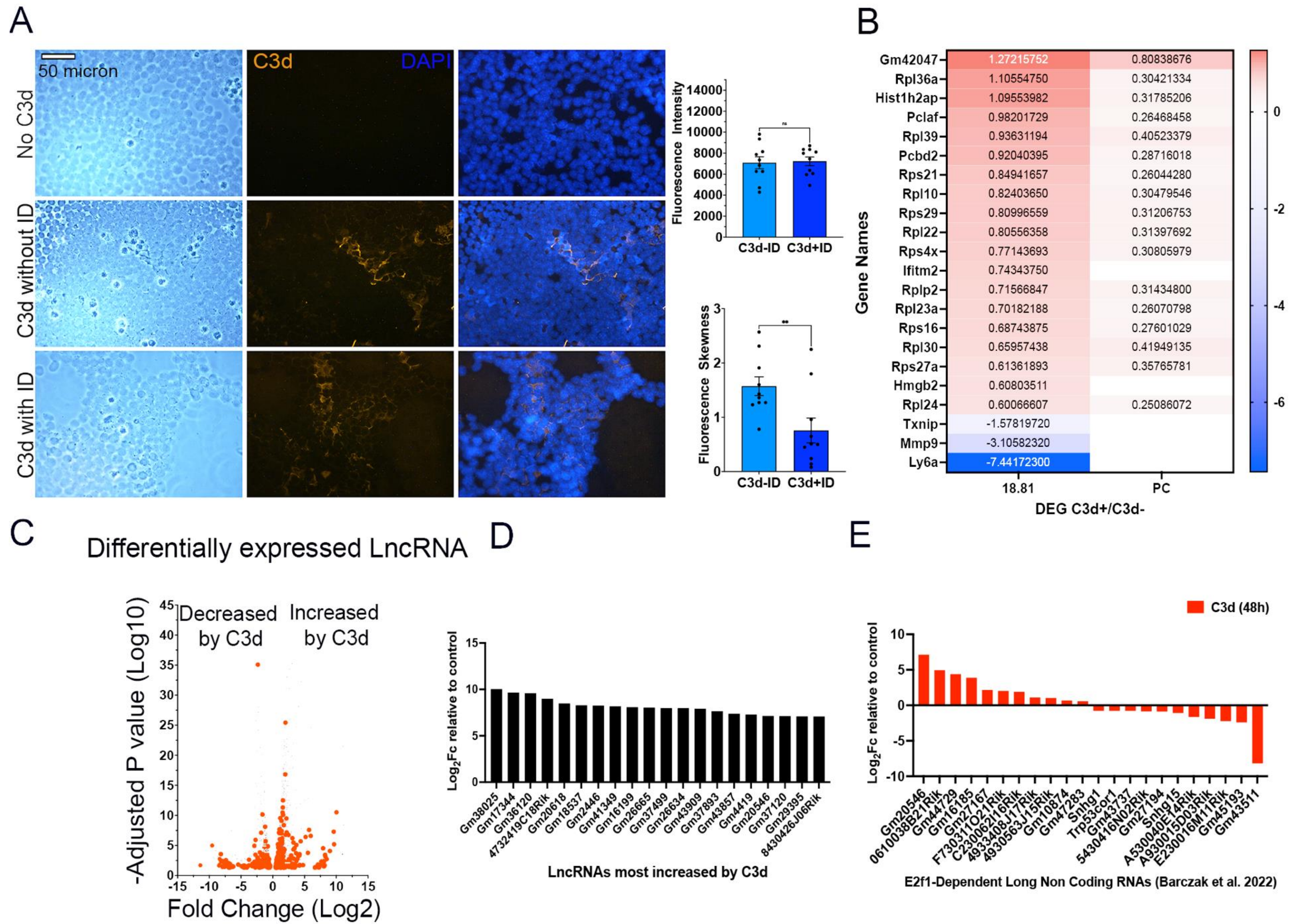


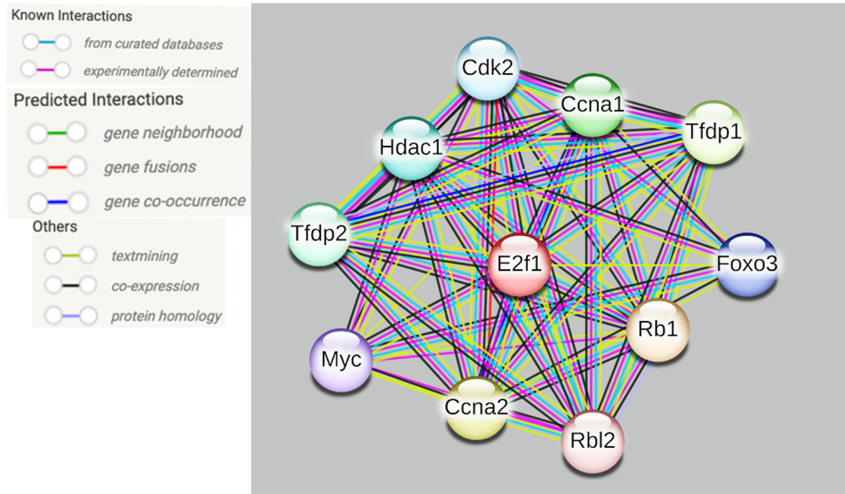
Figure S4



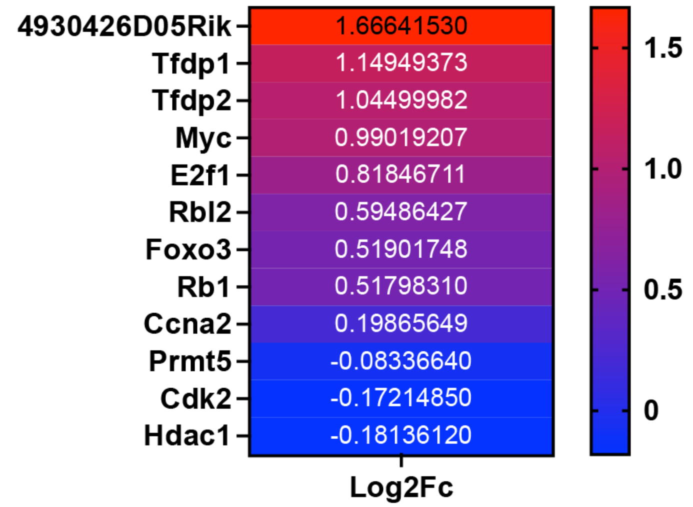
Gene Names	18.81	PC
Gm42047	1.27215752	0.80838676
Rpl36a	1.10554750	0.30421334
Hist1h2ap	1.09553982	0.31785206
Pclaf	0.98201729	0.26468458
Rpl39	0.93631194	0.40523379
Pcbd2	0.92040395	0.28716018
Rps21	0.84941657	0.26044280
Rpl10	0.82403650	0.30479546
Rps29	0.80996559	0.31206753
Rpl22	0.80556358	0.31397692
Rps4x	0.77143693	0.30805979
Ifitm2	0.74343750	
Rplp2	0.71566847	0.31434800
Rpl23a	0.70182188	0.26070798
Rps16	0.68743875	0.27601029
Rpl30	0.65957438	0.41949135
Rps27a	0.61361893	0.35765781
Hmgb2	0.60803511	
Rpl24	0.60066607	0.25086072
Txnip	-1.57819720	
Mmp9	-3.10582320	
Ly6a	-7.44172300	

Figure S5

A



B



C

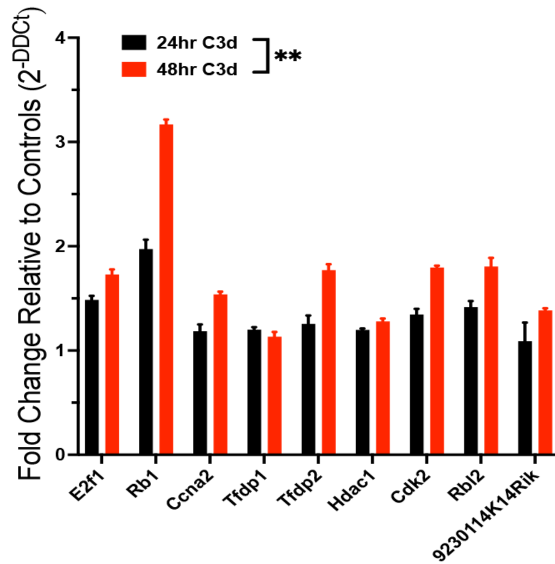
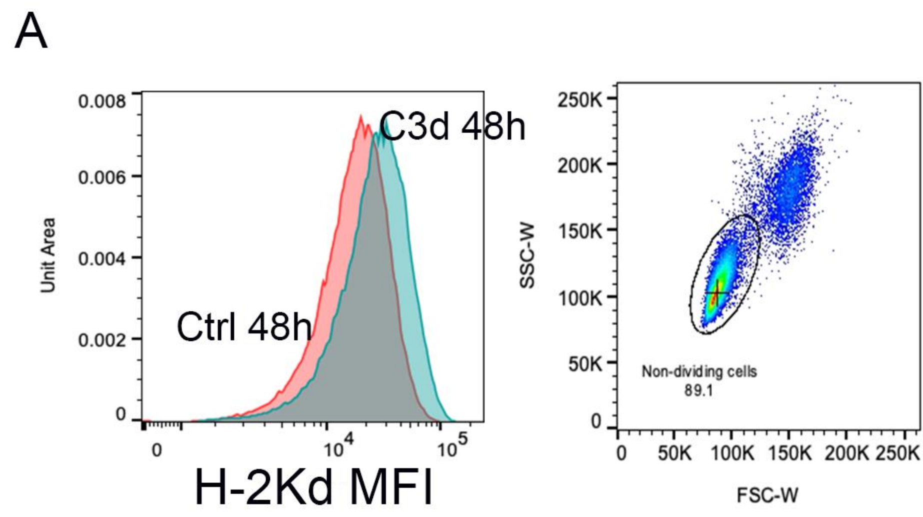


Figure S6



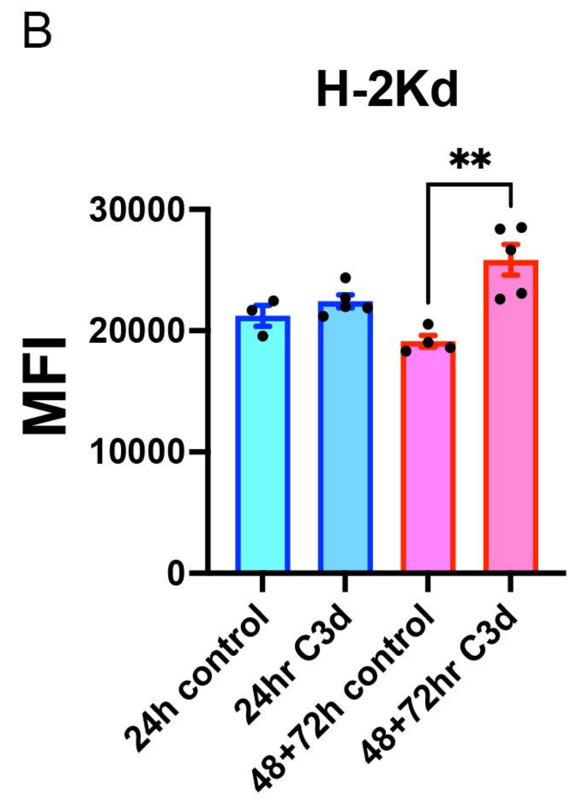
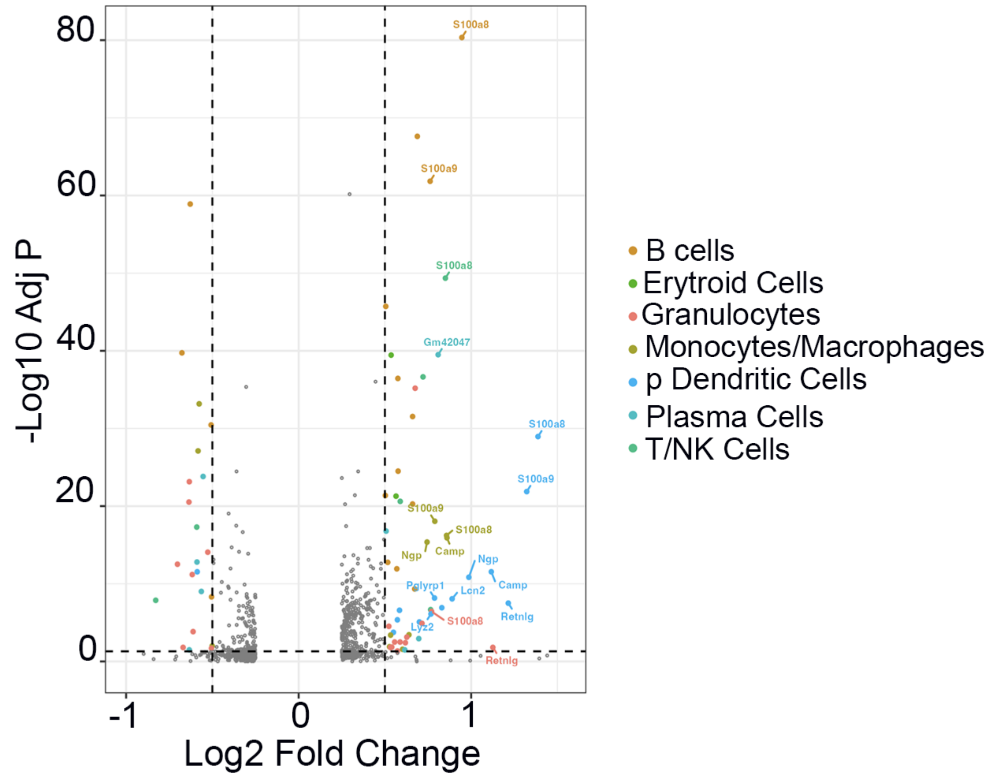


Figure S7



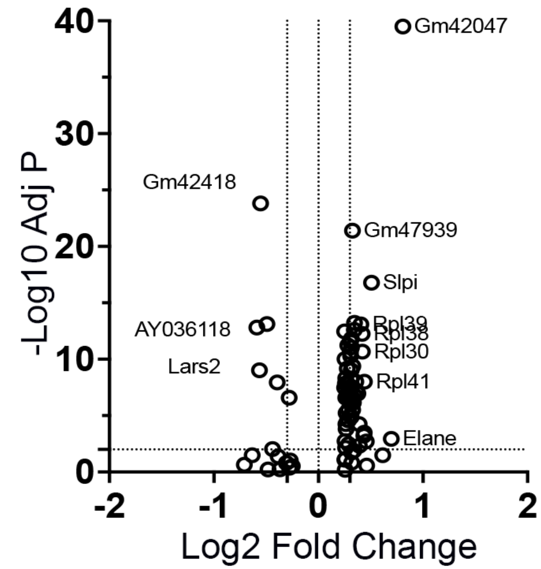


Figure S8

References Cited:

1. Wabl M, Gabain A, Steinberg C. Hypermutation at the immunoglobulin heavy chain locus in a pre B cell line. Proc Natl Acad Sci. 1985;82:479-82.
2. Barczak W, Carr SM, Liu G, Munro S, Nicastri A, Lee LN, et al. Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response. Nat Commun. 2023;14(1):1078.