Supplementary Information for

Dysregulated RNA splicing impairs regeneration in alcohol-associated liver disease

Ullas V. Chembazhi[1*], Sushant Bangru[1*], Rajesh Kumar Dutta[2], Diptatanu Das[1], Brandon Peiffer[3], Subhashis Natua[1], Katelyn Toohill[1], Aurelia Leona[4], Ishita Purwar[1], Anuprova Bhowmik[1], Yogesh Goyal[4,5,6], Zhaoli Sun[3], Anna Mae Diehl[2†], and Auinash Kalsotra[1,6,7,8,9†]

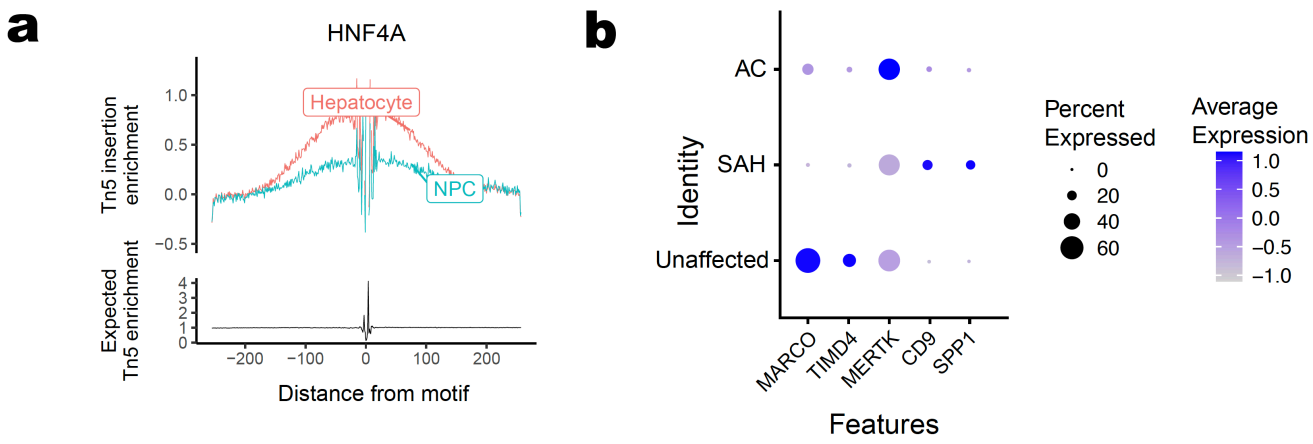
[1]Department of Biochemistry, University of Illinois, Urbana-Champaign, IL, USA; [2]Department of Medicine, Duke University, Duke University Health System, Durham, NC, USA; [3]Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [4]Department of Cell and Developmental Biology, Northwestern University, Chicago, IL, USA; [6]CZ Biohub Chicago, LLC, Chicago, IL, USA; [7]Cancer Center@ Illinois, University of Illinois Urbana-Champaign, IL, USA; [8]Division of Nutritional Sciences, University of Illinois, Urbana-Champaign, IL, USA; [9]Carl R. Woese Institute of Genomic Biology, University of Illinois Urbana-Champaign, IL, USA.*Equal contribution

[†]Address correspondence to:

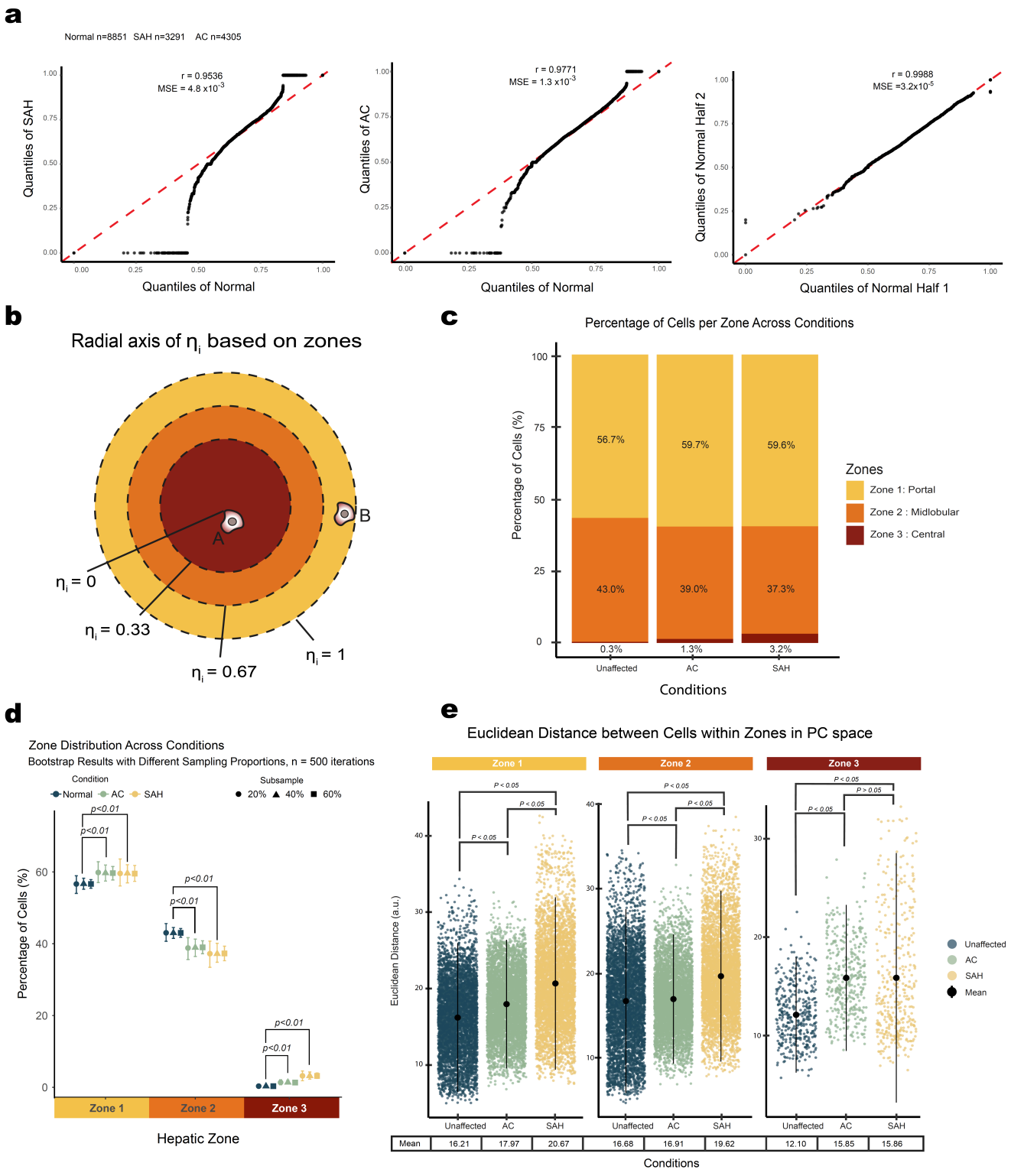
Auinash Kalsotra (lead contact), University of Illinois, 407 Medical Sciences Building, 506 South Mathews Avenue, Urbana, Illinois, USA. Phone: 217.300.7654; Email: kalsotra@illinois.edu.

Or to:

Anna Mae Diehl, 905 South LaSalle Street, GSRB I, Suite 1073, Durham, North Carolina 27710, USA. Phone: 919.684.4173; Email: annamae.diehl@duke.edu.

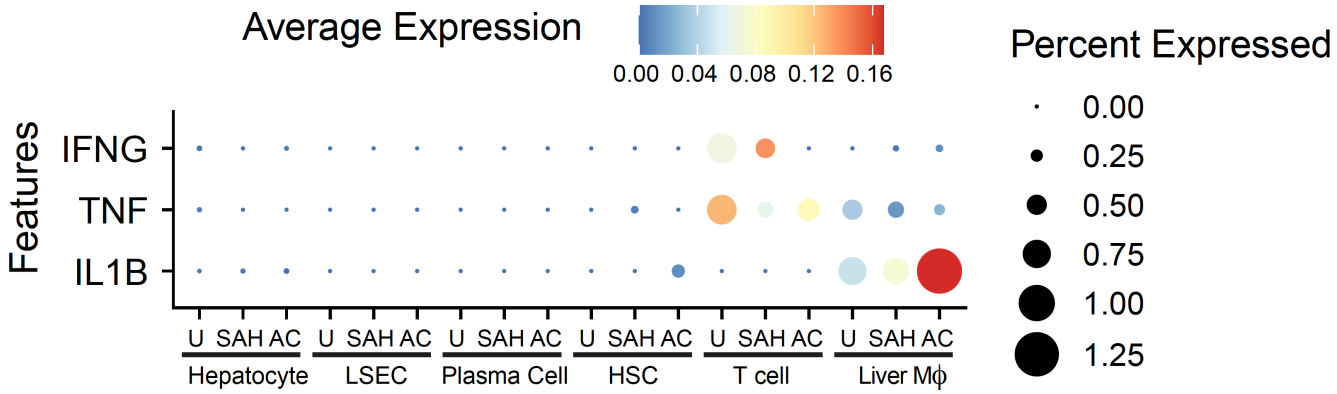


Supplementary Figure 1: (a) Genome-wide Tn5 insertion enrichment plots around motifs for HNF4α indicating their increased accessibility in hepatocytes compared to hepatocytes **(b)** Dot plot showing disease-dependent changes in expression levels of markers for Kupffer cell (MACRO, TIMD4, MERTK) and scar-associated macrophages (TREM2, CD9, SPP1).

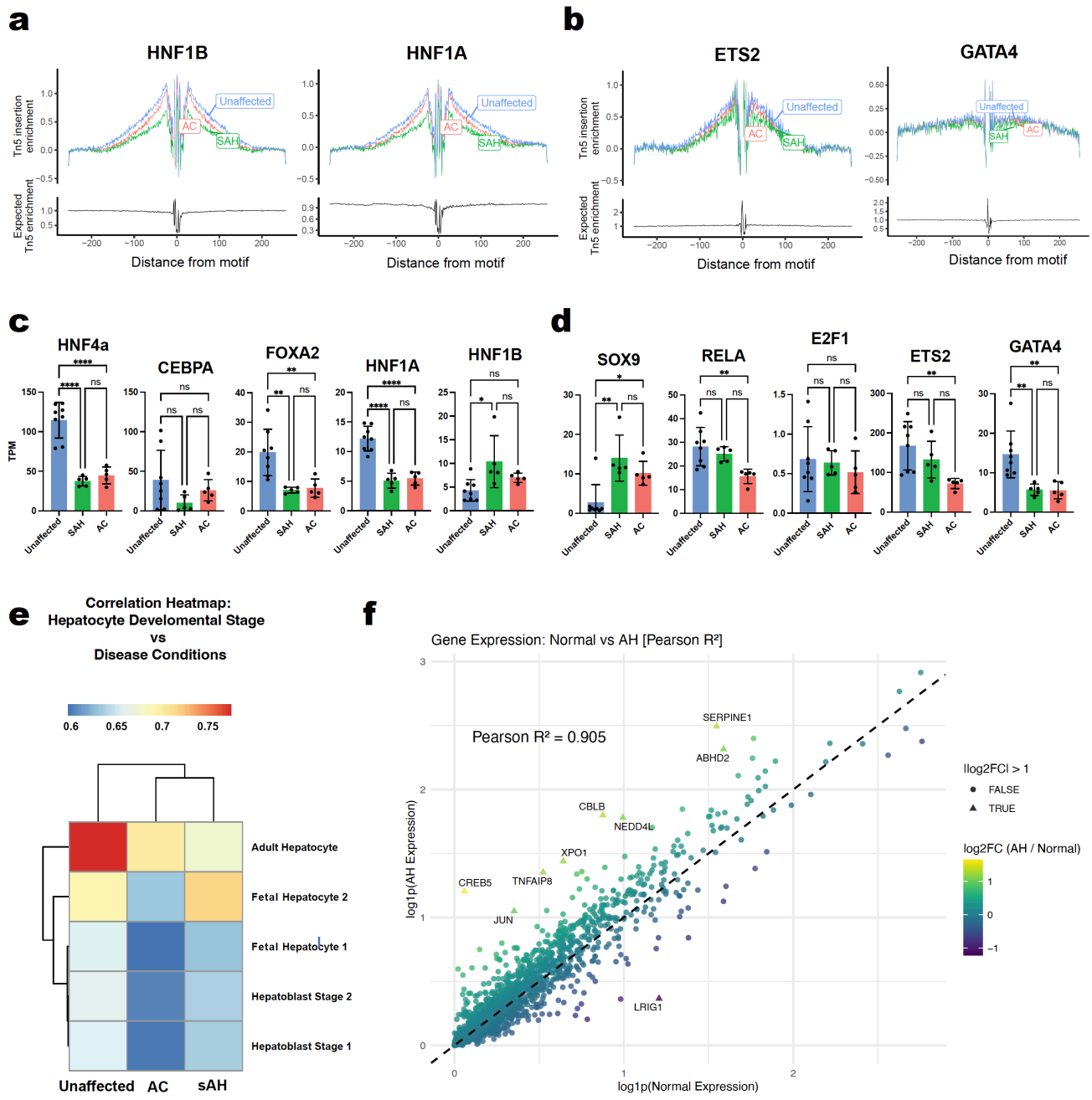


Supplementary Figure 2: (a) Quantile-Quantile plot of Unaffected-SAH (left), Unaffected-AC (middle), and Unaffected Half 1-Half 2 (right), comparing the similarity of η distributions. The control comparison of split Unaffected samples demonstrates identical distributions, with near-perfect correlation (r = 0.9988) and an MSE (3.2×10^{-5}) two orders of magnitude closer to zero than treatment comparisons. Both treatment conditions show a pronounced horizontal clustering

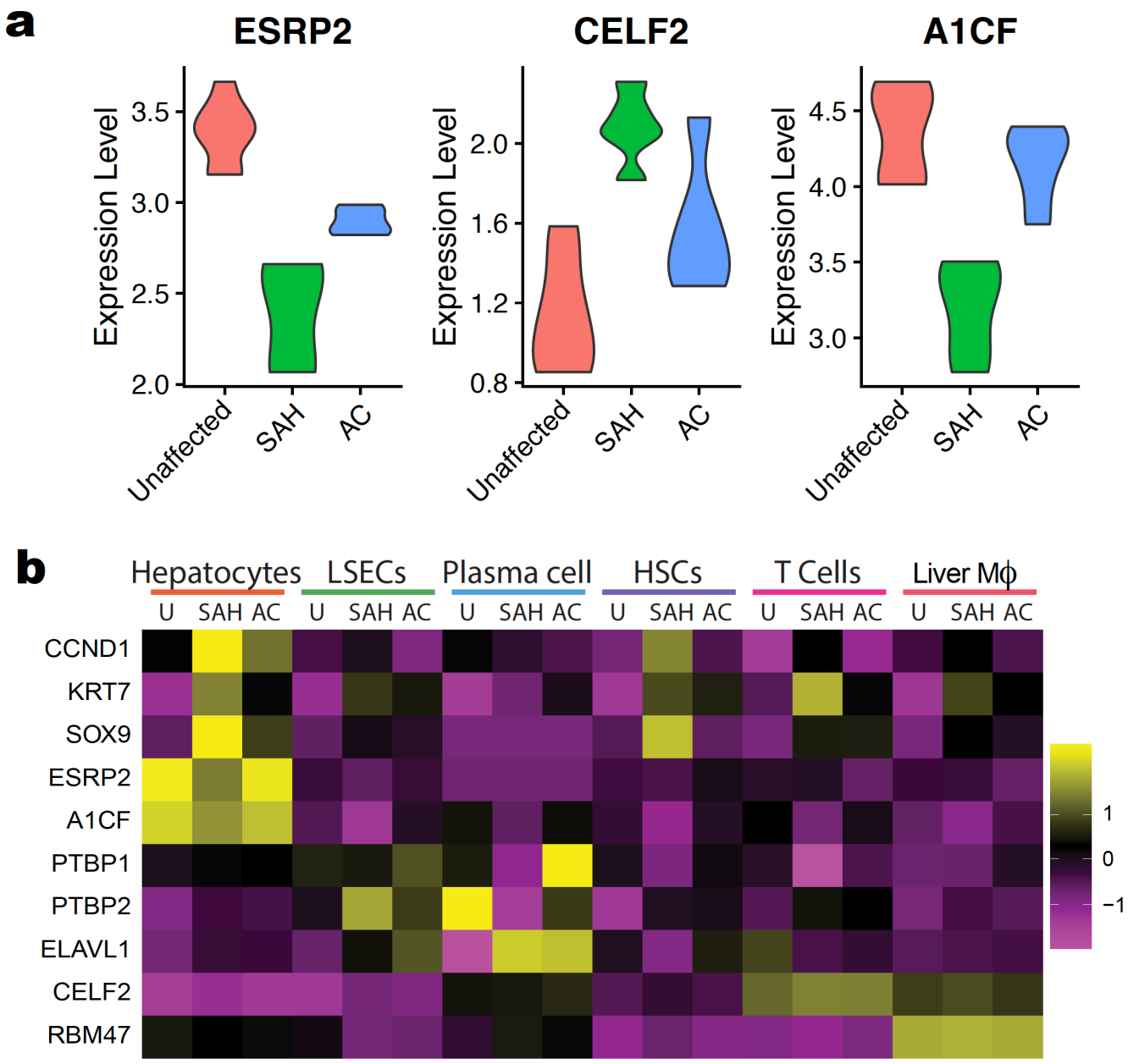
of points in the lower range (0.2-0.4) followed by deviation in the middle range (0.4-0.6), which is quantified by their larger MSE values (AC: MSE = 1.3×10^{-3}; SAH: MSE = 4.8×10^{-3}). The higher MSE in SAH reflects more pronounced clustering and deviation from Unaffected compared to AC while maintaining strong correlations (AC: r = 0.9771; SAH: r = 0.9536). **(b)** Schematic representation of the radial axis of η_i based on hepatic lobule zones. The concentric circles illustrate the equal division of the lobule into three zones, with η_i values ranging from 0 (central) to 1 (portal). Cell A is closer to the central axis and Cell B is closer to the portal axis. **(c)** Stacked bar chart displaying the percentage of cells per zone across Unaffected, SAH, and AC conditions. Zone 1 represents the Portal Axis and Zone 3 represents the Central Axis. **(d)** Bootstrapping the data with different subsampling proportions (50%, 70% and 90%) and measuring the percentage of each cells distributed across each zones (n = 500 iterations). Wilcoxon t-test was done on all percentages of cells from all iterations per subsampled condition, where $p<0.05$ shows significant differences between unaffected vs. diseased conditions. **(e)** Euclidean distances between cells within hepatic zones in principal component (PC) space. Scatter plot showing pairwise Euclidean distances between cells within each zone (1, 2, 3) for Unaffected, SAH, and AC conditions.



Supplementary Figure 3: Cell type-specific expression levels of key pro-inflammatory cytokines from snRNA-seq dataset.

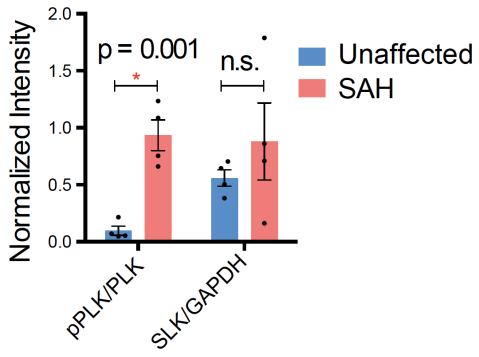


Supplementary Figure 4: Genome-wide Tn5 insertion enrichment plots for **(a)** HNF1A and HNF1B, **(b)** ETS2 and GATA4, comparing activities across Unaffected, SAH and AC livers. Expression levels of transcription factors active in **(c)** adult hepatocytes and **(d)** fetal livers from bulk RNA-seq datasets. **(e)** Single-cell gene expression of hepatocytes from diseased livers (datasets from this paper) and liver development[47] was averaged across disease conditions (columns) and developmental stages (rows), respectively. The average expression profile of each disease condition was compared to all developmental stages to calculate the Spearman correlation coefficient and plotted as a row-normalized heatmap. **(f)** Single-cell gene expression of hepatocytes was aggregated and averaged for SAH and Normal conditions, and a scatter plot was generated comparing gene expression of bonafide E2F1 transcriptional targets. Significant changes in gene expression (FindMarkers function, Test - MAST) between the two conditions were highlighted using a triangular shape (vs. non-significant, circle,) and points were colored according to the level of change with a Pearson coefficient calculated for the two conditions.

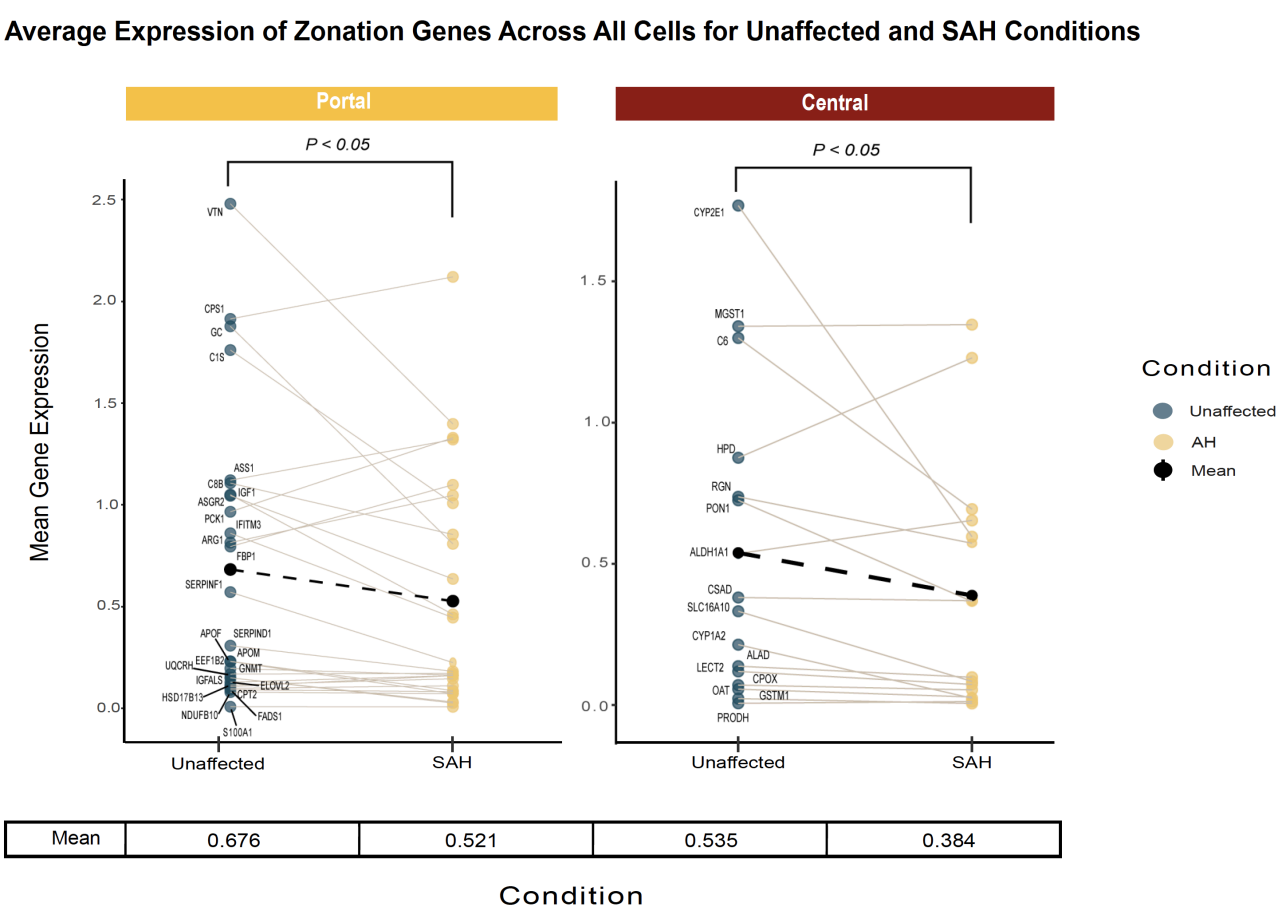


Supplementary Figure 5: (a) Violin Plot showing expression of ESRP2 and CELF2 across liver pathologies from bulk RNA sequencing data. **(b)** Heatmap from snRNA-seq module showing cell type-specific changes in expression of progenitor/fetal-like markers and RNA binding proteins from alcohol-associated liver disease (ALD).

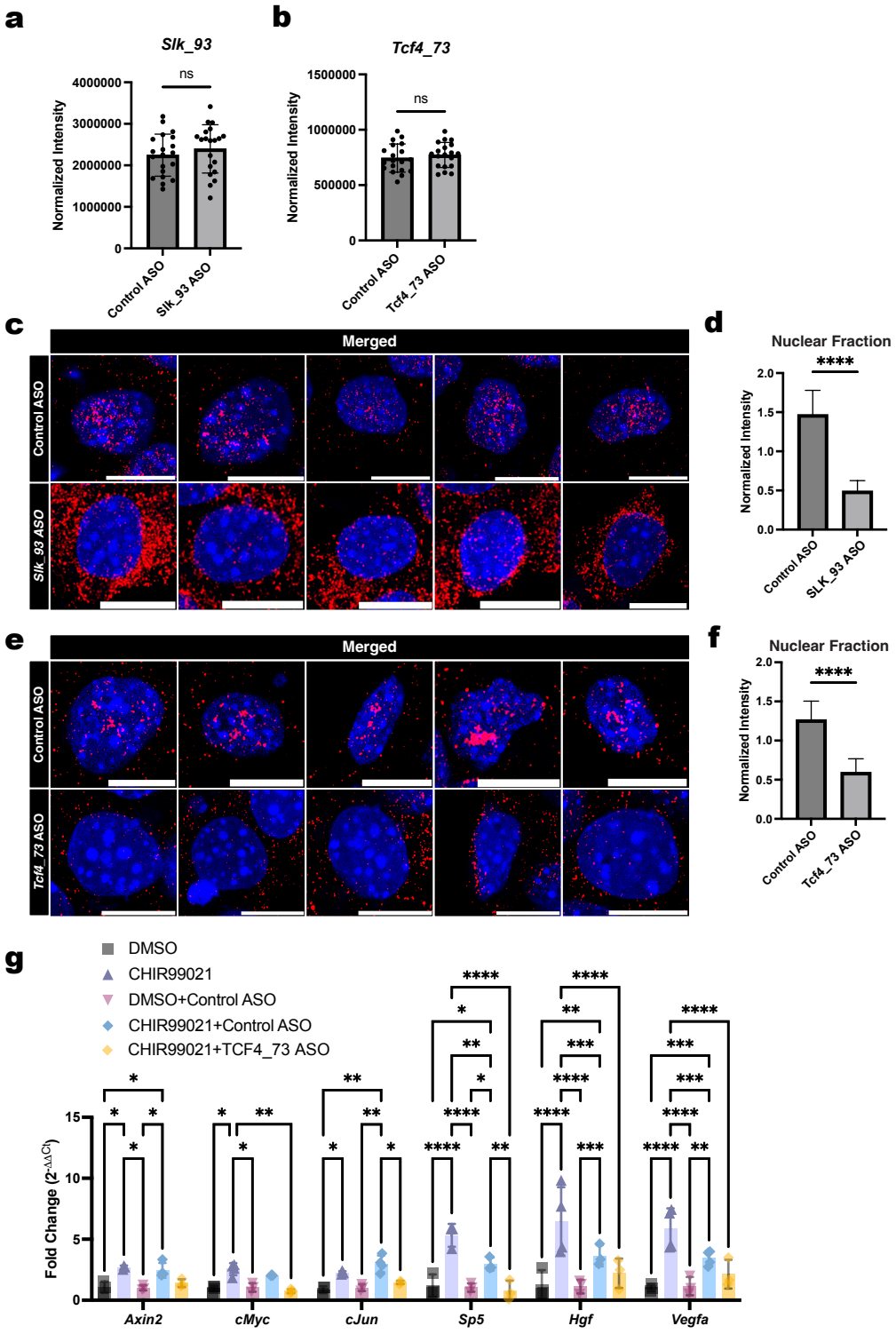
a



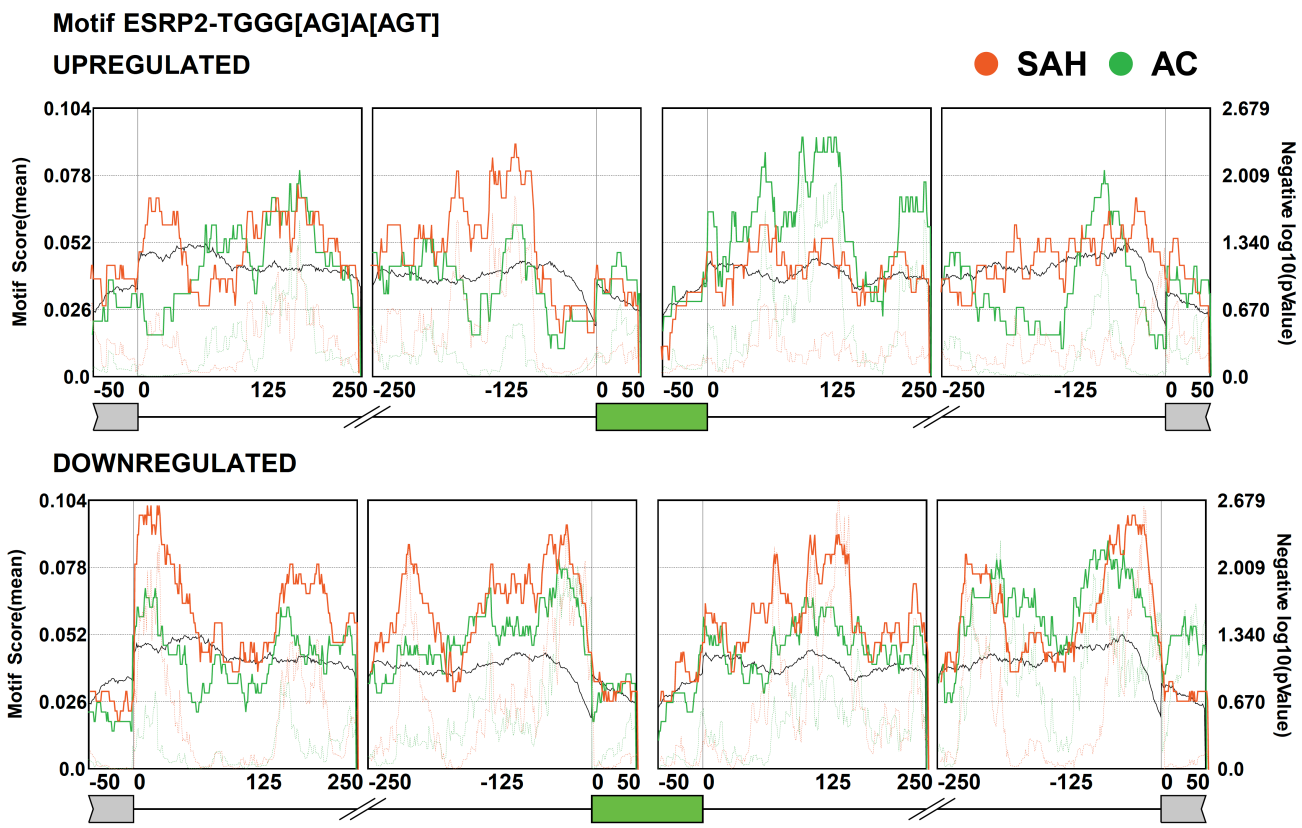
b



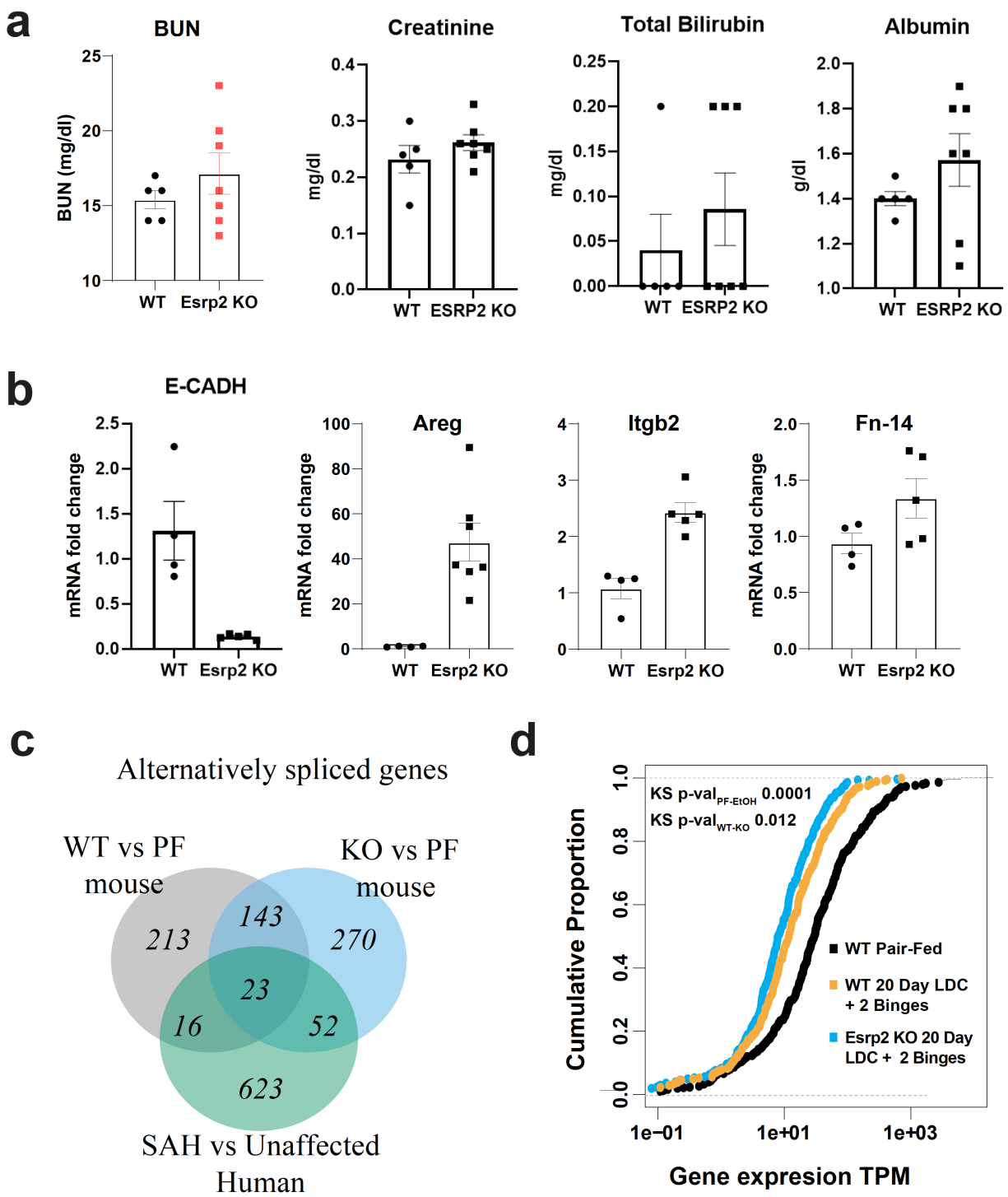
Supplementary Figure 6: (a) Quantitation for western blot analysis from human liver lysates demonstrating increased phosphorylation of Polo-like kinase (PLK) in SAH patients. GAPDH blot shown in Figure 3E is used for normalization. **(b)** Plots showing average expression of zonation Genes across all cells for unaffected and SAH conditions.



Supplementary Figure 7: Quantitation of normalized intensity from IHC staining after treatment with **(a)** SLK and **(b)** TCF4 after treatment with indicated ASOs. Examples of IHC from nuclei from cells post-ASO treatment to skip the disease-linked **(c)** 93 bp exon of SLK and **(e)** 73 bp exon of TCF4 in AML12 cells demonstrating altered nucleo-cytoplasmic localization of the protein upon exon skipping. Top: Control ASO treated cells; Bottom: ASO treated cells. Scale bar: 10µm. Quantification of normalized intensity within the nucleus for **(d)** SLK and **(f)** TCF4 after treatment with indicated ASOs. **(g)** qRT-PCR-based expression levels of WNT target genes in AML12 cells treated with control ASO, CHIR99021 and/or *Tcf4*_73-targeting ASO.



Supplementary Figure 8: Enrichment plots for ESRP2 binding motif presence in alternative exons and surrounding intronic regions that are misregulated in SAH and AC.



Supplementary Figure 9: (a) Serum levels of BUN, Creatine, total bilirubin, and Albumin in ESRP2 KO as compared to control mice after extended NIAAA diet schedule (20D+2B). **(b)** qRT-PCR based quantitation of selected genes from ESRP2 KO mice liver relative to WT C57BL6 mice, after extended NIAAA diet schedule. **(c)** Venn diagram showing overlap of genes alternatively spliced in WT and ESRP2 KO mice upon 20D+2B diet relative to pair-fed WT control mice. **(d)** Cumulative plot illustrating trends in expression of TCF4 target genes in 20D+2B fed WT and ESRP2 KO mice liver compared to pair-fed WT control mice liver.