SUPPLEMENTAL FIGURES AND LEGENDS

Figure S1 | Reprogramming TFs MNase Tag- Enriched Target Sites Associated with Cell Identity Genes. Related to Figure 1.

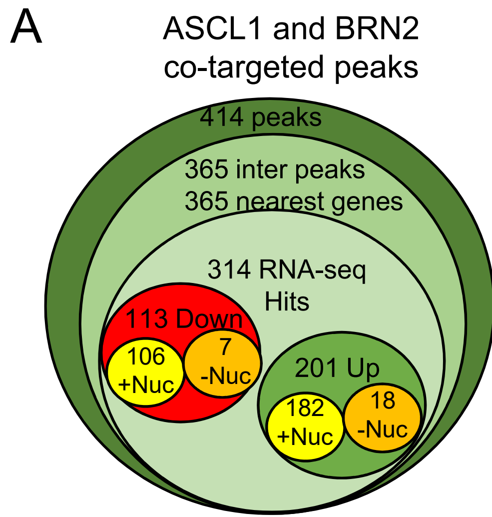
(A-B) Genomic data integration for ASCL1 and BRN2 nucleosomal target site identification. (A) Venn diagram of TFs ChIP-seq peaks association with upregulated/downregulated nearest genes and their nucleosome occupancy as determine by MNase-seq tag enrichment or depletion at 48 hr ASCL1, BRN2 and MYT1L ectopic expression in fibroblast. (B) Gene ontology analysis of nearest genes associated with TFs ChIP-seq peaks in 48 hr ASCL1, BRN2 and MYT1L ectopic expression in fibroblast.

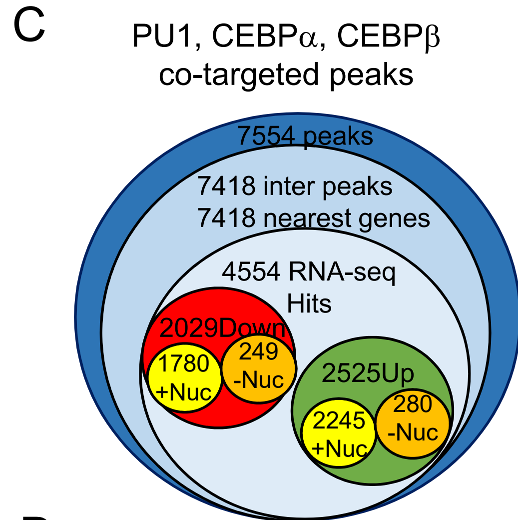
(C-D) Venn diagram of PU1, CEBPα and, CEBPβ ChIP-seq peaks associated with upregulated/downregulated nearest genes in macrophages and nucleosome occupancy in fibroblasts determined by MNase-seq tag enrichment at expected target sites, according to macrophage PU1, CEBPα and, CEBPβ ChIP-seq. (D) Gene ontology analysis of nearest genes associated with for PU1, CEBPα, and CEBPβ ChIP-seq peaks in macrophages.

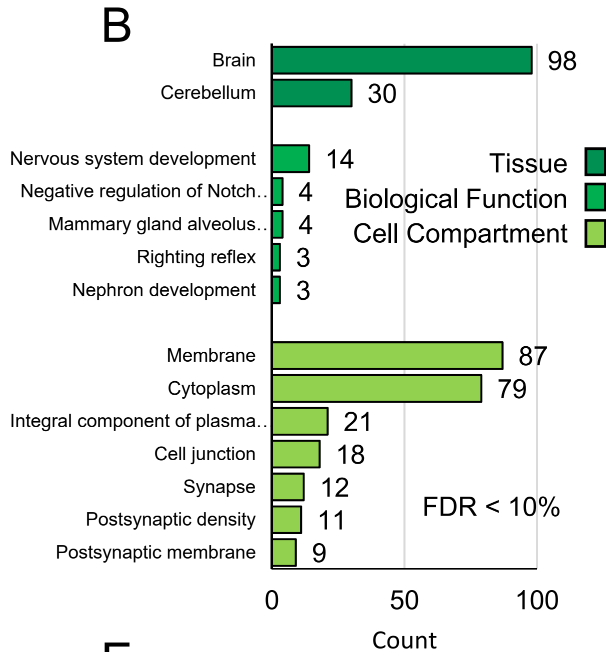
(E) Schematic diagram showing 3' end enzymatic Cy5 labeling of DNA with Klenow^{-exo}.

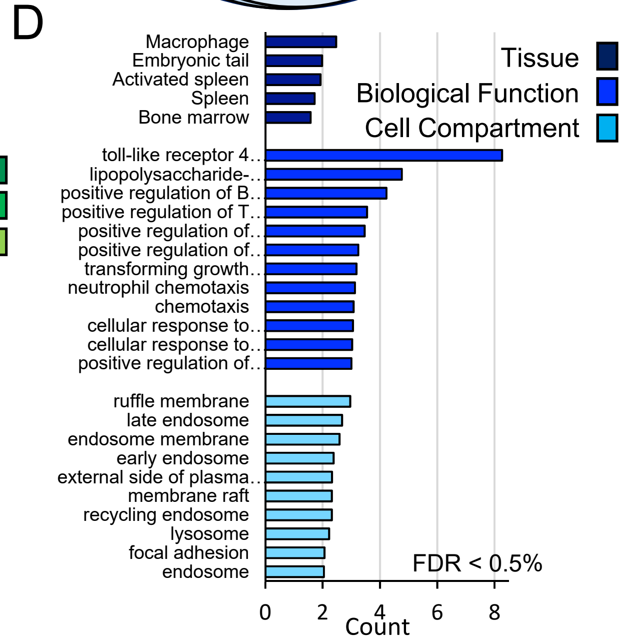
(F) Nucleosome reconstitution of *ALBN1-NUC*, *CX3CR1-NUC* and *NRCAM-NUC* (white arrows) compared to free DNA (black arrows).

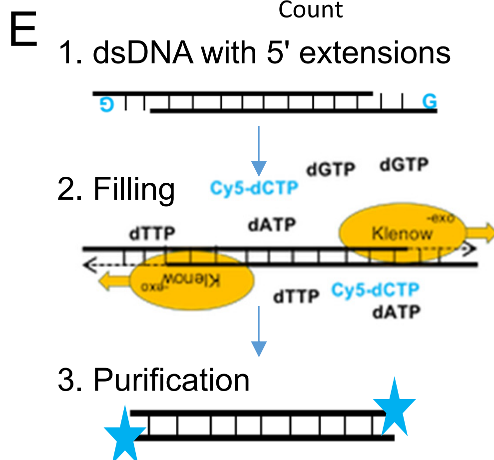
Fernandez Garcia- Zaret Supp. Figure 1











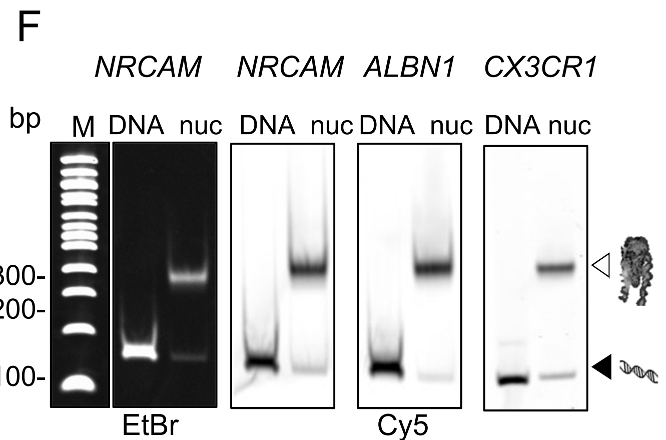


Figure S2 | Recombinant Purified Full-Length TFs Bind Their Canonical Motifs on Short dsDNA. Related to Figure 2.

(A) TF identity verification by Western blot analysis.

(B) Cy5 detection of labeled short double stranded DNA containing canonical binding motifs for reprogramming TFs. Base pair length of each DNA shown at the bottom.

(C-D) TF binding curves showing the affinity of *E. coli*. expressed full-length TFs to Cy5 labelled DNA probes containing or lacking respective canonical binding sites, (C) specific DNA and (D) non-specific DNA. TFs were titrated at 0, 0.1, 0.3, 1, 3 and 9 nM concentrations with 1 nM DNA. DNA sequences of the Cy5-labelled probes are in the Star Methods and equations for total and specific dissociation constant calculations (K_d^{Total} , $K_d^{Specific}$ are described in (Soufi et al., 2015b).

(E) Representative EMSA showing the affinity of increasing amounts of purified ASCL1, ASCL1/E12α, and E12α homodimers proteins to Cy5- MCK-DNA containing an ASCL1 E-box binding motif.

(F) Binding curves or ASCL1, ASCL1/E12α and, E12α to Cy5-MCK DNA based on (E).

(G) ASCL1 STRING functional association network of protein-protein interactions.

Fernandez Garcia- Zaret Supp. Figure 2

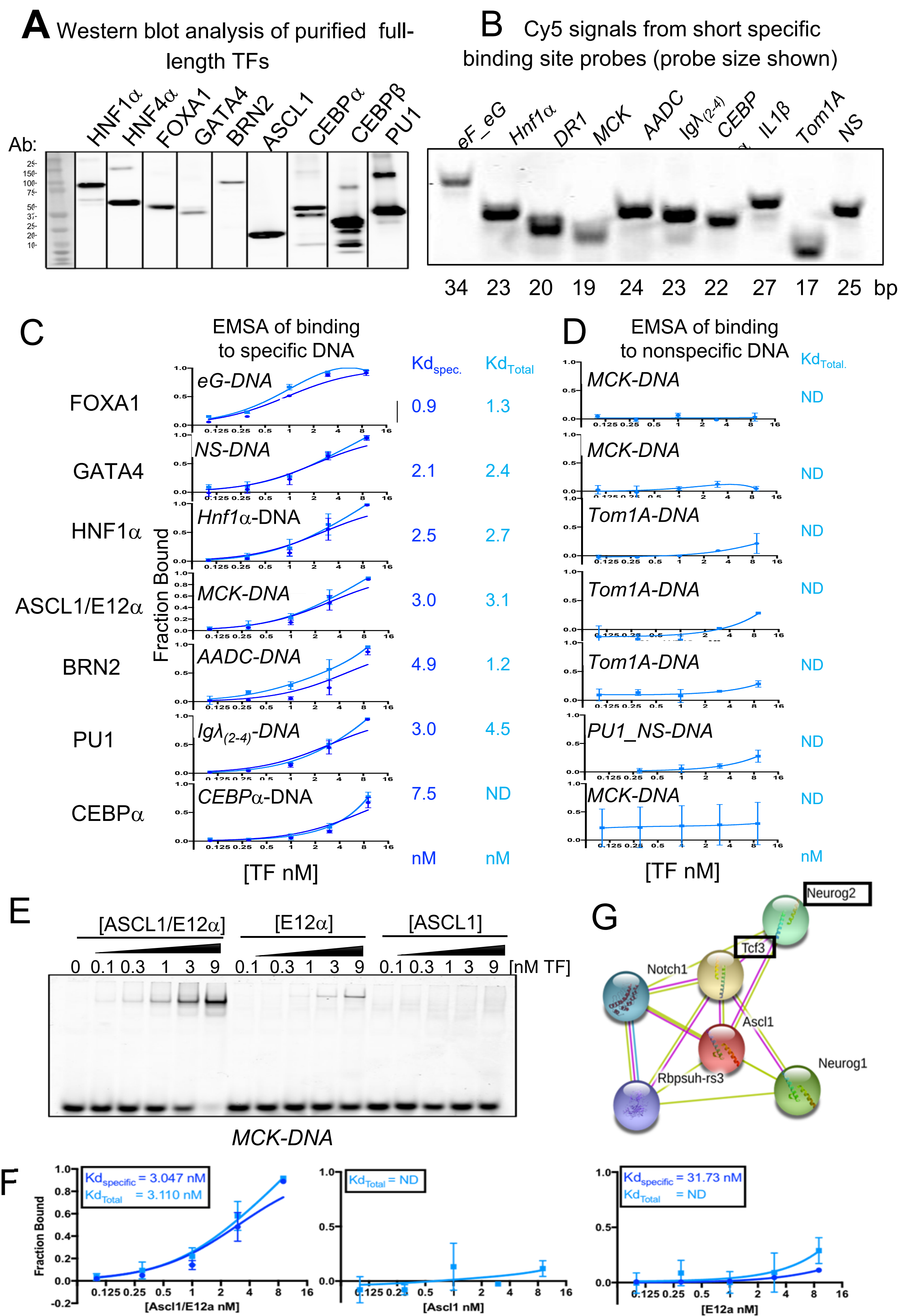


Figure S3 | Recombinant Purified Full-Length TFs Bind with Specificity to DNA and Nucleosomes. Related to Figures 2 and 3.

(A-D) TF binding curves related to EMSAs fraction bound quantification of Figures 2B-G showing the affinity of TFs to Cy5-160bp-DNA (left panels) or Cy5-160bp-NUC (right panels). TFs were titrated on 1 nM DNA or NUC.

(E) Recombinant purified full-length CEBPβ WT, CEBPβ-T163, and CEBPβ-S163D, T167D phosphomimetic mutants analyzed by SDS-PAGE and Coomassie staining. The TF bands run at the expected sizes when compared to the sizes of protein standards. All proteins in the same gel.

(F) FOXA1, ASCL1/E12α, PU1, BRN2, and CEBPα binding to *ALBN1-DNA*, *NRCAM-DNA,* and *CX3CR1-DNA* in the presence of 20-, 40-, and 80-fold molar excess of specific competitor (“s” lanes) or non-specific competitor (“ns” lanes) or absence of competitor (“-” lanes).

Fernandez Garcia- Zaret Supp. Figure 3

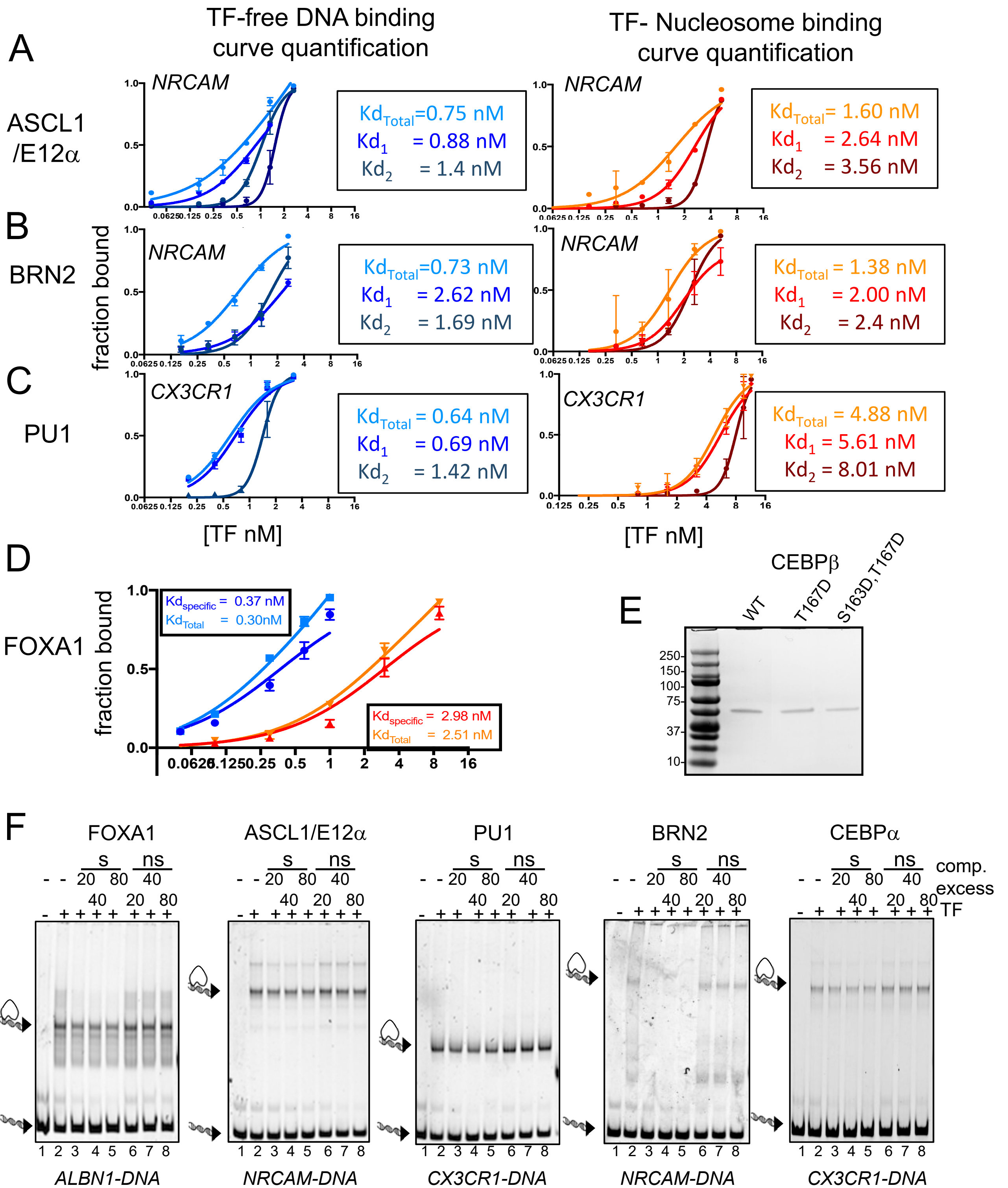


Figure S4 | Pipeline for Protein Microarray Data Processing. Related to Figure 5.

(A) Schematic diagram showing protein microarray data processing for the identification novel nucleosome binding TFs.

(B) Representative protein microarray chip showing printed protein auto-fluorescence on the chip edge.

(C) Representative protein microarray zoom showing relative printed protein amounts per spots differences by fluorescence detection with primary anti-GST-Tag antibody and Cy5-conjugated secondary antibody.

(D) Removal of proteins with low concentration per spot to yield positive signal. Scatter plot showing protein amounts of all protein spots in all experiments (left panel), and scatter plot showing proteins removed from the analysis due to low estimated concentration (right panel).

(E) Representative protein microarrays probed with *Cy5 labeled ALBN1*, *NRCAM,* and *CX3CR1* DNA or NUC showing the difference in background intensity.

(F) Examples of spots duplicates showing the range of fluorescence intensities (shown above each spot) detected when probed with Cy5 labeled *ALBN1*, *NRCAM* and *CX3CR1* DNA or NUC. A threshold for positive hits (binding) and negative hits (not binding) was determined to be F-B of 80 (black arrow).

Fernandez Garcia- Zaret Supp. Figure 4

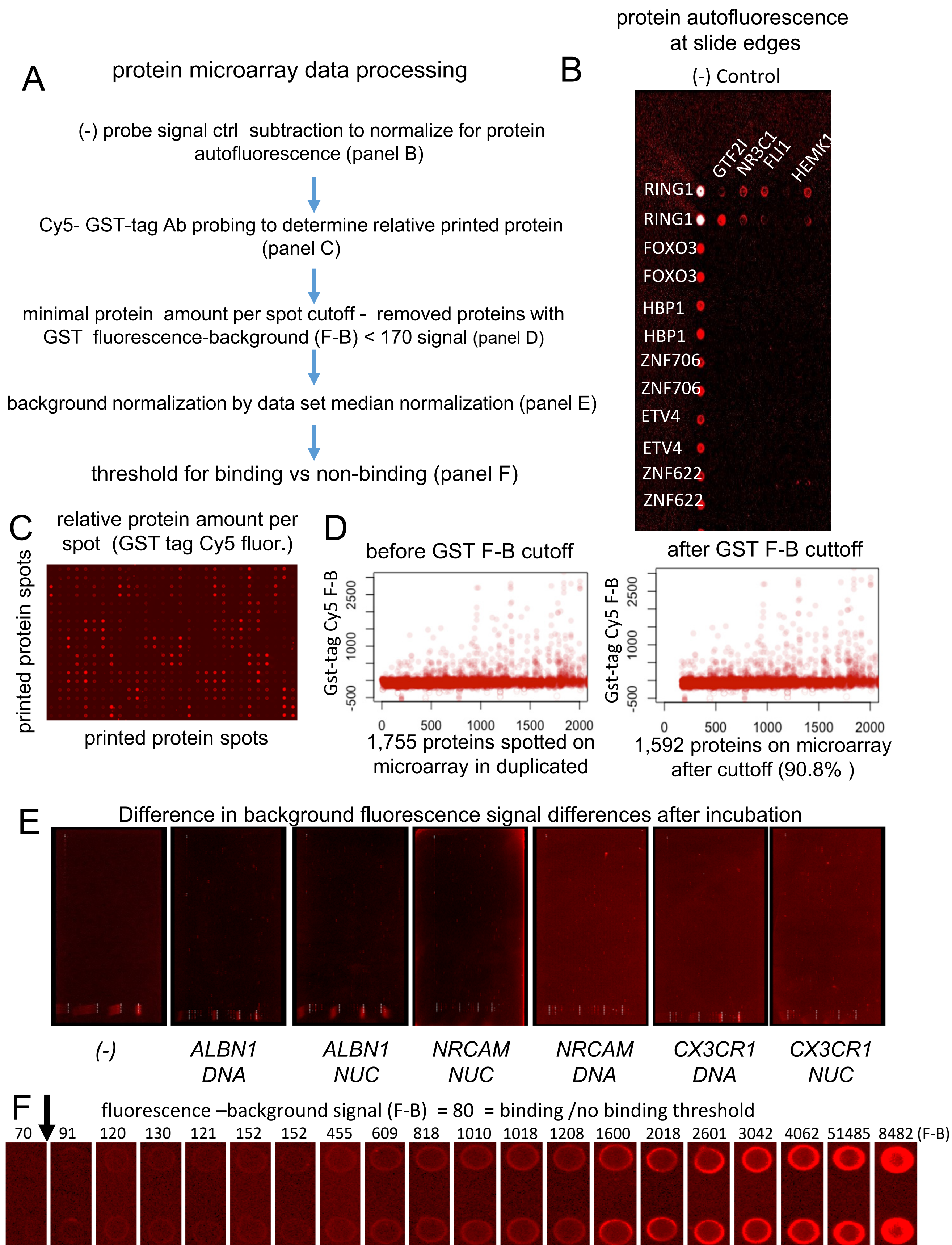


Figure S5 | 3D Representation of DNA Sequences Showing Positioning of TFs Binding Motifs of Tested Factors. Related to Figure 5.

For each TF tested by EMSA (see Figure 5D), the DNA and nucleosome target sequence is shown with the consensus motif for the factor, the motif sequence within the tested DNA sequence, the strand location of the motif, and a 3D representation of tested sequences showing TF motif (yellow) localization within the sequence.

Fernandez Garcia- Zaret Supp. Figure 5

	TF Sequence	Canonical motif	Motif of sequence	Strand	3D DNA-motif representation
BTF3	ALBN1	N/A	N/A	N/A	
GATA1	ALBN1		TGATA	+	
RBPJ	ALBN1		CCAGGGAATG	+	
FOXA1	ALBN1		TGTTTG; TGTTTGC	+;+	
HNF1A	ALBN1		TTAAAT	+	
HMGA1	ALBN1	minor groove AT-rich	AAT, TTT, TT, TTAAATA, ATxATxATT, TTT, TT, TTTT, ATxTAxAT,AAAxAAxTTT	+	
NHLH2	NRCAM	NKGNMKCAGCT GCGYCMN	TCCCAGCAGCTGCTGCCT	+	
ASCL1	NRCAM		CAGCTG; CAGCTG	+;+	
E12-alpha	NRCAM	CANNTG	CAGCTG; CAGCTG	+;+	
MYOG	NRCAM		AGCAGCTG; GACAGCTG	+;+	
TFAP2A/B/C	NRCAM		GGCCACAGGCA	-	
ZNF250	NRCAM	NTAGGCCTAN	TGTGGCCCAC	+	
ELK1	NRCAM		CAGGAAG	-	
Brn2 (POU3F2)	NRCAM		ATGCA, ATACAAAAGT	-;+	
PKNOX	NRCAM		TGCATGACAGC		
TBX20	NRCAM	NAGGTGTGAAN	ACATCACATCCTGTG	-	
CREM	NRCAM	NRTGAYGTCAYN	ATCACATCC	+	
SPDEF	NRCAM		CACAGGATGTG	-	
Brn5 (POU6F1)	NRCAM		ATACAAAAGT	+	
HMGN1	CX3CR1	N/A	N/A	N/A	
HMGN5	CX3CR1	N/A	N/A	N/A	
ESRRG	CX3CR1		CTGAAGATCAGCAG	-	
CEBPa	CX3CR1		GTTGCT; GTTGCA; AGCAAT; TAGCAAT	+;+:-;+	
CRX	CX3CR1		AGTCTTAG	+	
NFATC1	CX3CR1	NTTTCCGCGGAAAN	GGGGAATCGGGAAGTATCCAT	-	
IRF3	CX3CR1		ATCGGGAAGTA	-	
ELF2	CX3CR1	NACCAGGAAGTN	AATCGGGAAGTATC	-	
PU1	CX3CR1		GGAAGT, GGAATC	-	
T	CX3CR1		TCACAAAAATAGGTCA	+	

Figure S6 | HMG TFs Interact with Nucleosomes. Related to Figures 5 and 6.

(A) Heatmap representation of TFs' DNA bound fraction showing tested TFs (nM) concentration on DNA and nucleosomes. Concentrations were determined experimentally.

(B-E) Representative EMSA of identified microarray hits for nucleosome interaction with (B) *NRCAM-DNA*, (C,E) *CX3CR1-DNA,* and (C) *ALBN1-DNA* (black arrows) or corresponding nucleosomes (white arrows). (E) Representative EMSA showing DNA and nucleosome binding of HMG TFs.

(F-G) EMSA of DNA-only binders (F) IRF3 and (G) CREM1 to *ALBN1* and *CX3CR1* DNA and nucleosomes containing TFs binding motifs, respectively.

(H) Sequence alignment of SOX2, SOX5, and SOX9 with HMGN1, HMGN2, and HMGN5 TFs showing high conservation between SOX group B homology domain and HMGN C-term nucleosome binding domain.

Fernandez Garcia- Zaret Supp. Figure 6