Cell Reports, Volume *27*

Supplemental Information

Rational Reprogramming of Cellular States by Combinatorial Perturbation

Jialei Duan, Boxun Li, Minoti Bhakta, Shiqi Xie, Pei Zhou, Nikhil V. Munshi, and Gary C. Hon

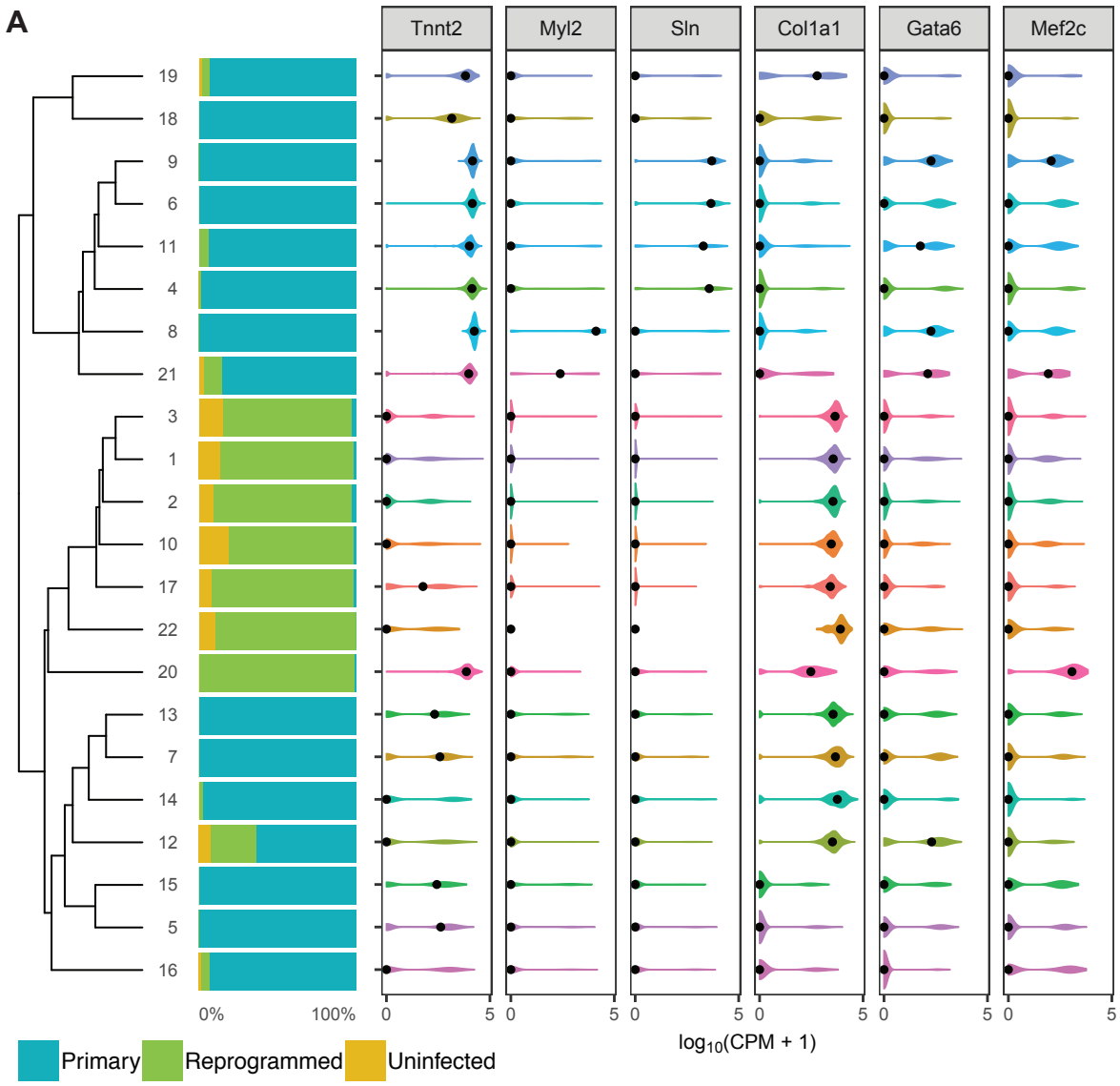
Table S1: Sequencing statistics. Related to Figure 2.

Sample	Platform	Number of cells	UMIs	
			mean	median
Whole heart	Drop-seq	2729	3306.41	2441
Left atrium	Drop-seq	699	3030.607	1960
Right atrium	Drop-seq	9127	1940.593	1276
Remainder (except LA RA)	Drop-seq	268	3644.078	2517
Remainder (except RA)	Drop-seq	2861	4691.692	3466
Reprogrammed, 48 factors pooled	Drop-seq	8730	10415.459	7337.5
Uninfected control	Drop-seq	1362	10540.609	8137
Whole heart	10x	11935	5742.818	4672
Uninfected control	10x	4914	14930.817	13713
Reprogrammed, 10 TFs pooled, day 7	10x	5291	11468.627	10534
Reprogrammed, 10 TFs pooled, day 14	10x	5639	11013.78	9724
Reprogrammed, 3 TFs pooled, day 7	10x	6785	9678.652	8805
Reprogrammed, Myod1, day 7	10x	6258	7563.493	6302
Total		**66598**		

Table S4: qPCR primers. Related to Figure 6.

Primer name	Sequence
Upk1b forward	CACTGTTCGTTGCTTCCAGG
Upk1b reverse	GCTTCGAGAAGTGGGTAAAGACT
Gpm6a forward	TGCTTCGAGTGCTGCATTAAA
Gpm6a reverse	TTGCCAACTCAAAGTAGGTCTG
Krt19 forward	GGGGGTTCAGTACGCATTGG
Krt19 reverse	GAGGACGAGGTCACGAAGC
Col3a1 forward	GCCTTCTACACCTGCTCCTGTGC
Col3a1 reverse	TGGAAGGCCAGGGTCACCATTTC

Figure S1, related to Figure 2

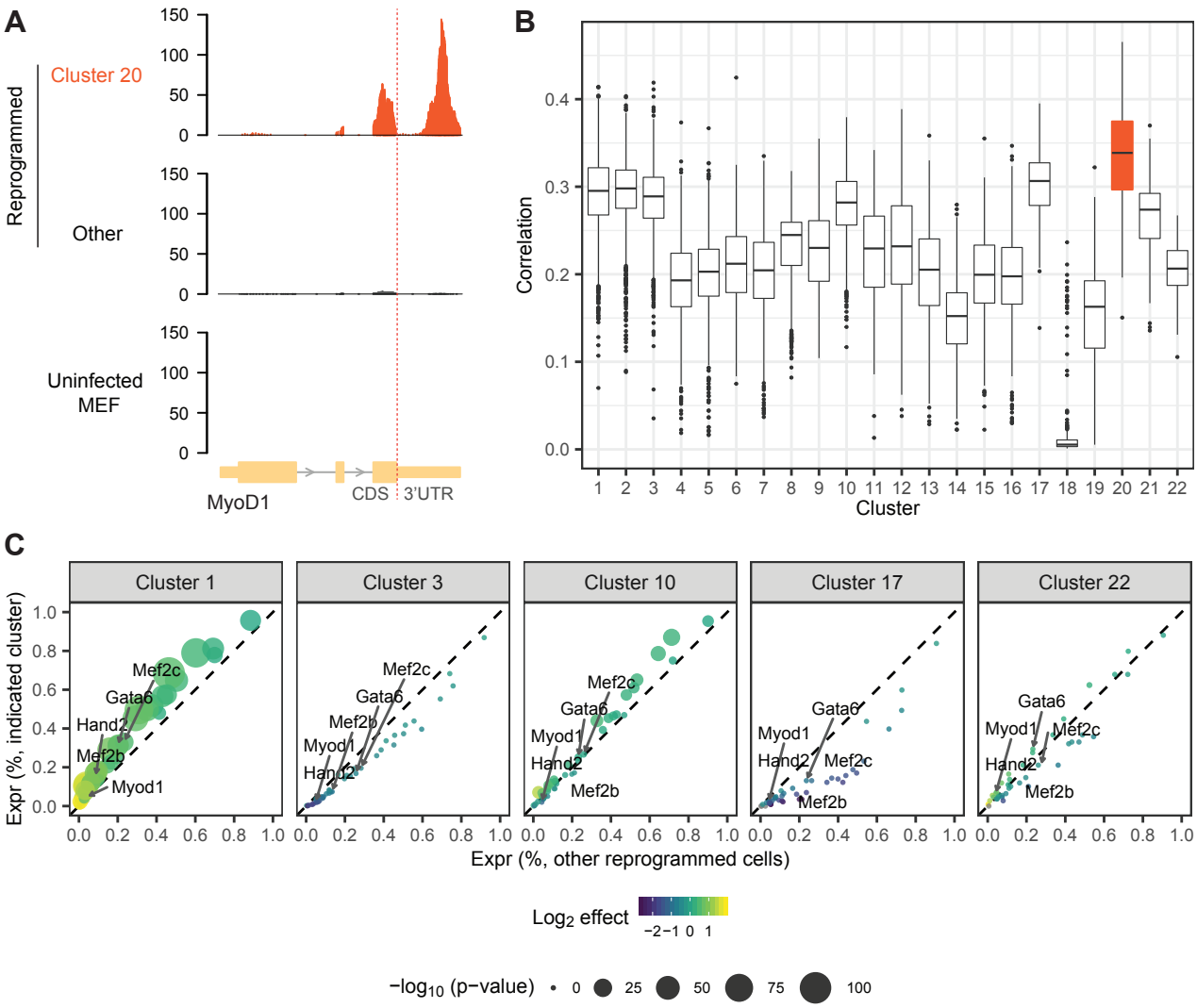


A) (left) Bar charts indicating the cellar composition of t-SNE clusters defined in Figure 1B. (right) Violin plots illustrating the expression of cardiac markers from single-cell expression data derived from P0 mouse heart and reprogrammed/uninfected MEFs.

B) Heatmap of cells in Figure 1B by the expression of (top left) the cardiomyocyte marker gene *Tnnt2*, (top right) the epicardial marker gene *Bnc1*, (bottom left) the lymphocytic marker gene *Ly6c1*, and (bottom right) cell proliferation marker gene *Cenpp*.

C) 48F reprogramming cells are highlighted.

Figure S2, related to Figure 2

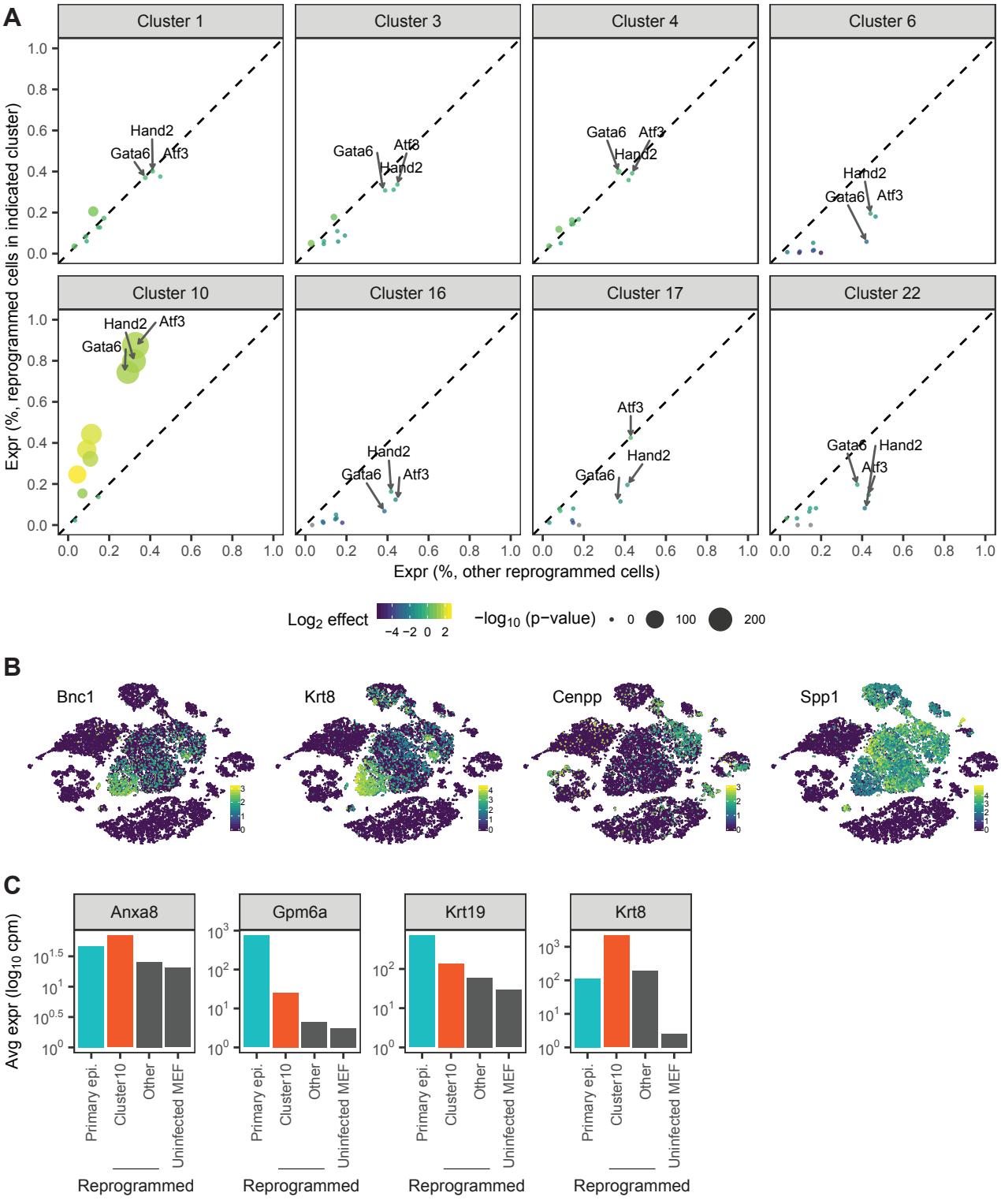


A) Genome Browser snapshot of *MyoD1* expression from Cluster 20 cells, other reprogrammed cells, or uninfected MEFs. The cloned exogenous *MyoD1* gene does not have a 3' UTR. Hence, exogenous reads are concentrated upstream of the 3' UTR, while endogenous reads are inside the 3' UTR.

B) We performed scRNA-Seq on MyoD1-reprogrammed MEFs. Shown is the gene expression correlation of MyoD1-reprogrammed MEFs with the clusters from Figure 1B.

C) Differential expression analysis of 48F in MEF-derived clusters, as compared to all other reprogrammed cells. Each plot contains 48 dots (colored by fold change and sized by p-value).

Figure S3, related to Figure 3

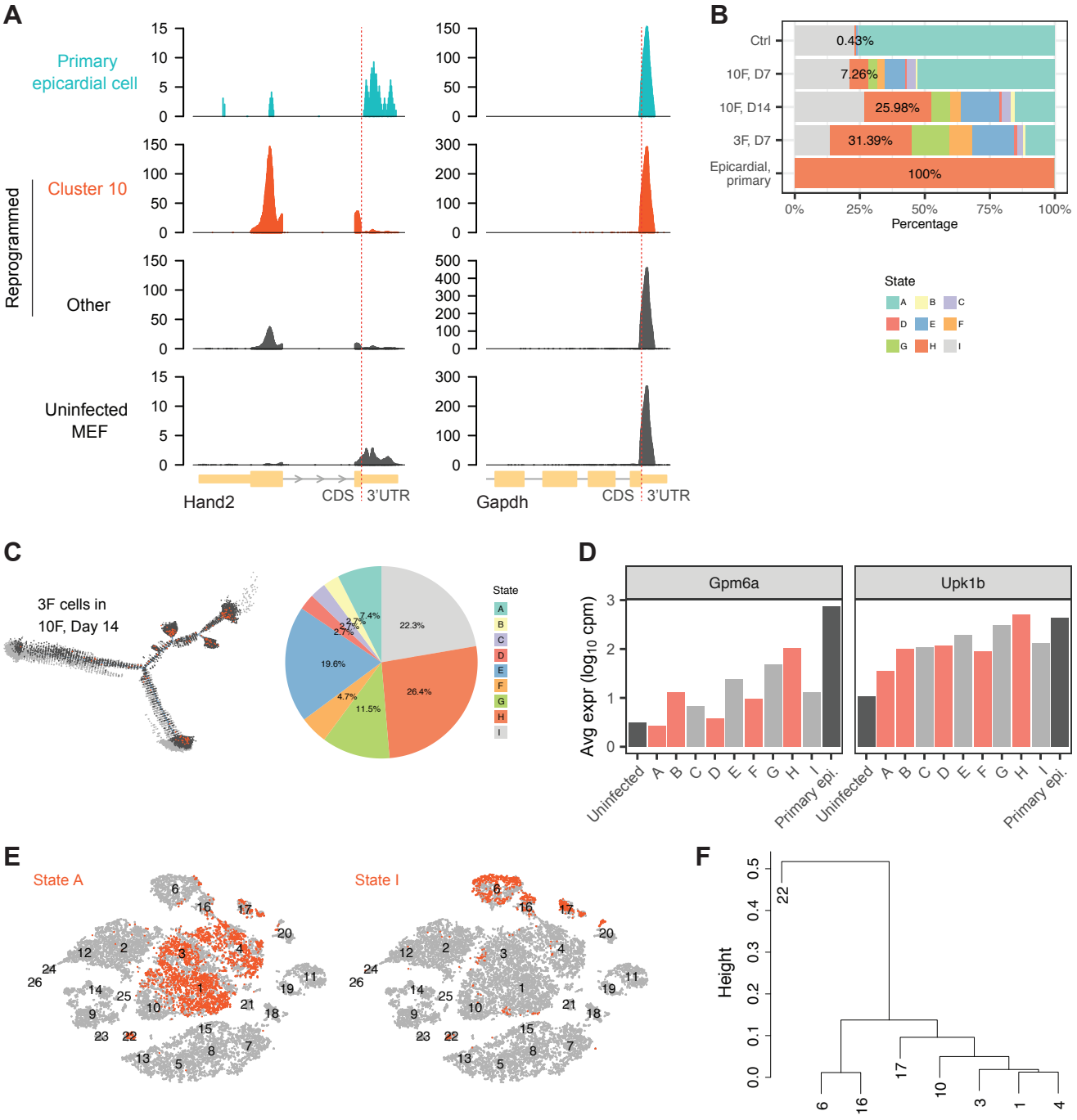


A) Enrichment of 10F in MEF-derived clusters, as compared to all other reprogrammed cells.

B) Heatmap of cells in Figure 3C by the expression of (left) the epicardial marker genes *Bnc1*, *Krt8* and cell proliferation marker gene *Cenpp*, fibroblast gene *Spp1*.

C) Expression of epicardial marker genes *Anxa8*, *Gpm6a*, *Krt19*, and *Krt8* in primary and MEF-derived cells.

Figure S4, related to Figure 4



A) Genome browser snapshots of (left) *Hand2* and (right) *Gapdh*. In primary cells and uninfected cells, *Hand2* reads are almost exclusively derived from the 3'UTR. In contrast, read enrichment in reprogrammed cells is rarely in the 3' UTR, which is indicative of exogenous expression.

B) Bar plot indicating the composition of cells in each reprogramming experiment based on pseudotime states from Figure 4C.

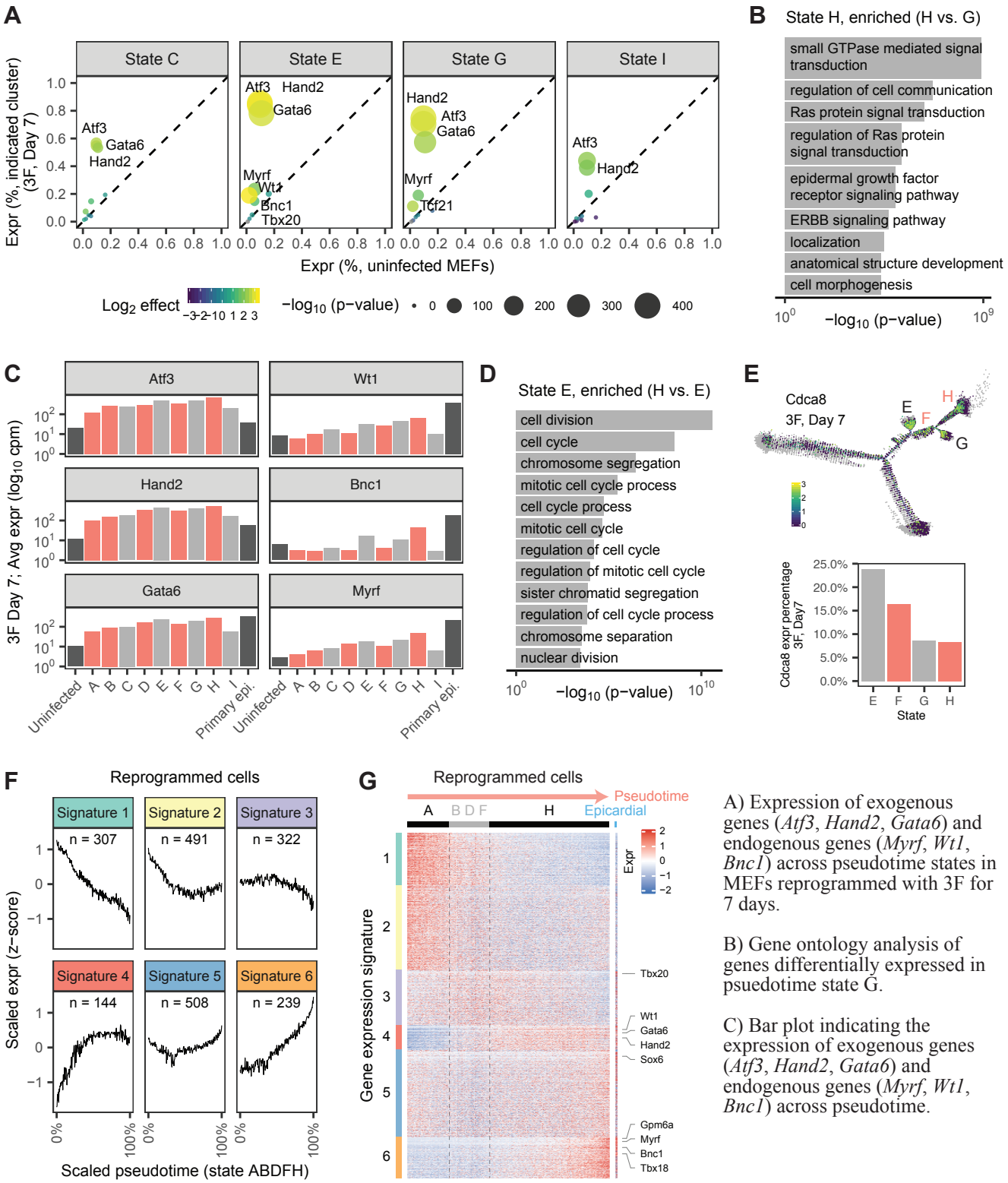
C) The distribution of 3F cells in 14-day 10F reprogramming across different pseudotime branches. Left, red: 3F cells in 10F 14-day reprogramming experiment; black: other 10F 14-day reprogramming cells; grey: all the other cells.

D) Quantified expression of *Gpm6a* and *Upk1b* in each pseudotime branch (Figure 4H).

E) Cells from t-SNE plots in Figure 3C are colored by pseudotime state A or I.

F) Dendrogram of cell clusters from Figure 3C illustrating that Cluster 10 cells are most similar to MEF-derived cells from Clusters 1, 3, and 4.

Figure S5, related to Figure 5



A) Expression of exogenous genes (*Atf3*, *Hand2*, *Gata6*) and endogenous genes (*Myrf*, *Wt1*, *Bnc1*) across pseudotime states in MEFs reprogrammed with 3F for 7 days.

B) Gene ontology analysis of genes differentially expressed in psuedotime state G.

C) Bar plot indicating the expression of exogenous genes (*Atf3*, *Hand2*, *Gata6*) and endogenous genes (*Myrf*, *Wt1*, *Bnc1*) across pseudotime.

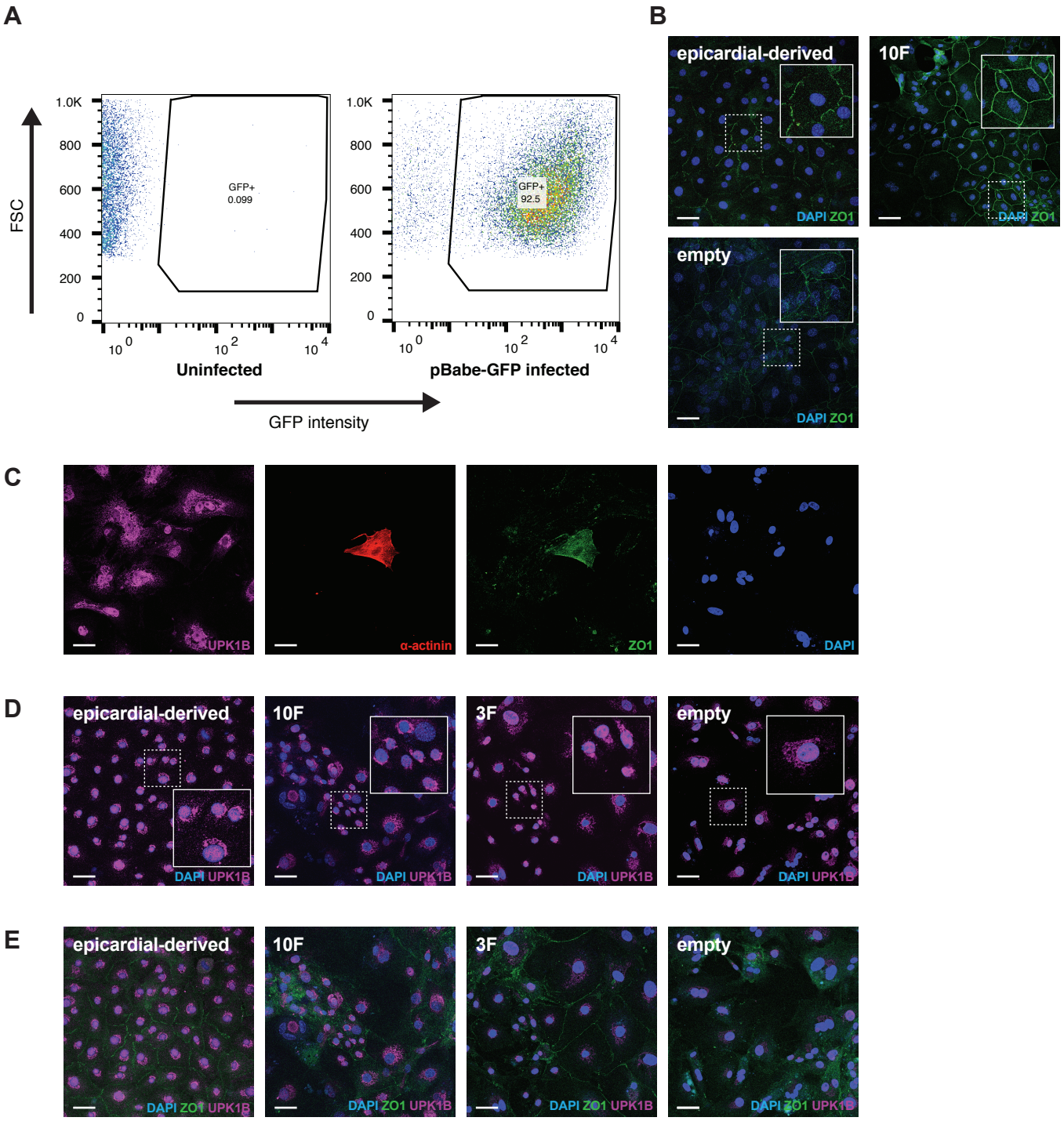
D) Gene ontology analysis of genes differentially expressed in psuedotime state E.

E) (top) Heatmap of cells by the expression of the cell division marker *Cdca8*. (bottom) Quantification of *Cdca8* expression in pseudotime states E-H.

F) Gene regulatory network during 3F reprogramming. The genes cluster into six gene expression signatures during pseudotime. Expression is log10 transformed and Z-normalized.

G) Heatmap of the six gene expression signatures. Columns indicate cells ordered by pseudotime (with primary epicardial cells at the far right), rows indicate genes. For visualization purposes, groups of 20 cells are binned.

Figure S6: Functional validation of epicardial reprogramming, related to Figure 6



A) Flow cytometry analysis of GFP+ cells in uninfected and pBabe-GFP infected MEFs at Day 5.

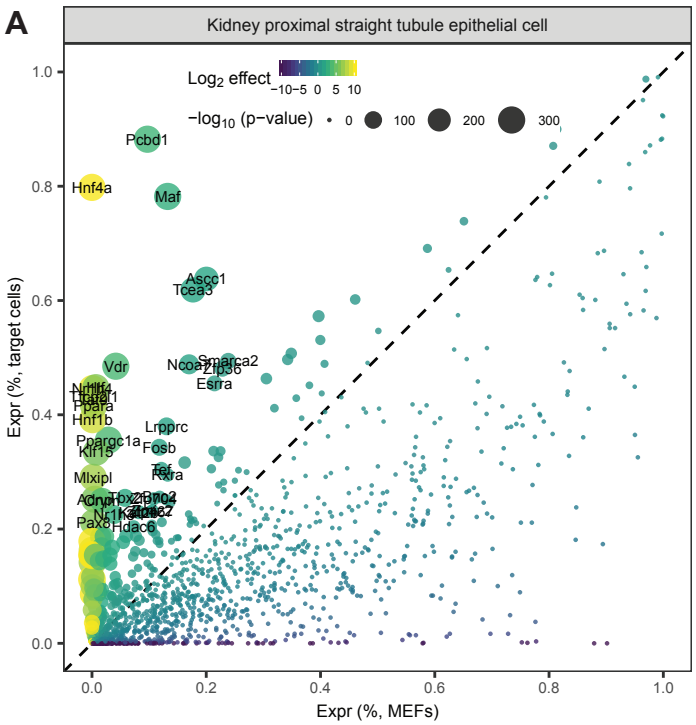
B) ICC with DAPI of ZO-1 in epicardial-derived cell line, 10F-reprogrammed, and empty control MEFs at confluency. Insets: Zoomed-in views of example cells. Scale bar indicates 50 microns.

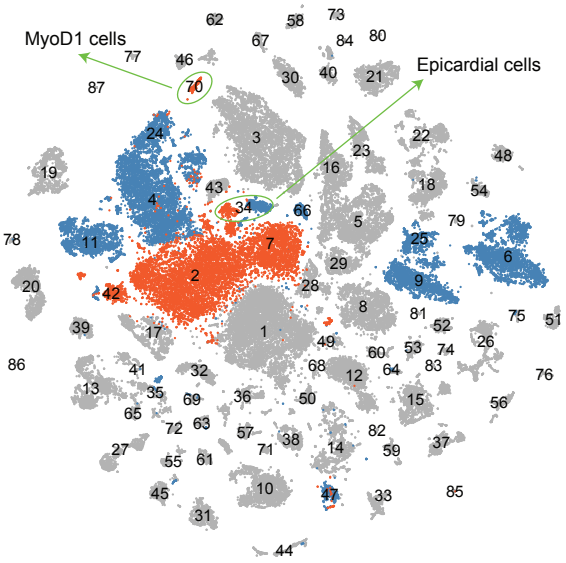
C) Immunocytochemistry (ICC) with DAPI of a-actinin, ZO-1, and Upk1b in GHMT-reprogrammed MEFs at Day 7. Scale bar indicates 50 microns.

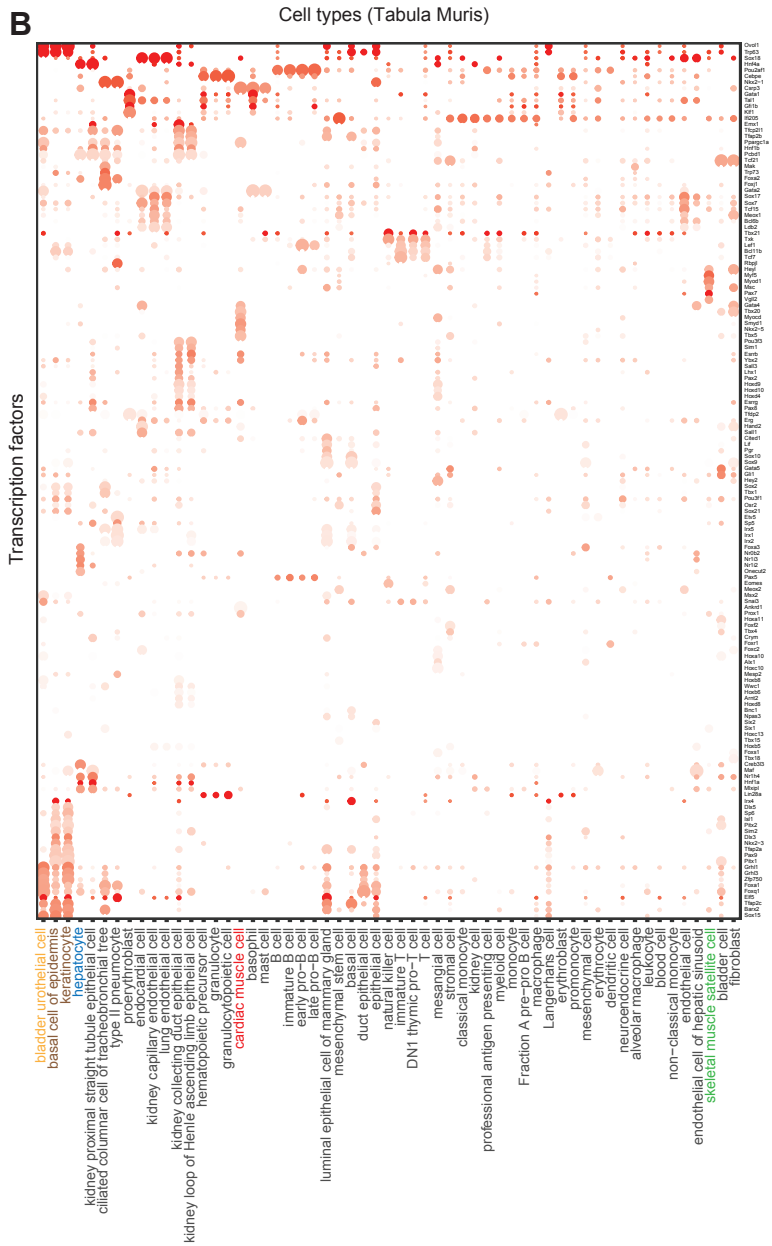
D) ICC with DAPI of Upk1b in epicardial-derived cell line, 10F-, 3F-, and empty control MEFs. Insets: Zoomed-in views of example cells. Scale bar indicates 50 microns.

E) ICC with DAPI of ZO-1 and Upk1b in epicardial-derived cell line, 10F-, 3F-, and empty control MEFs. Scale bar indicates 50 microns.

Figure S7, related to Figure 7.







A) Differential expression analysis of TFs in renal tubular epithelial cells defined by Tabula Muris, as compared to MEFs.

B) Dot plot indicating the expression of 145 low entropy TFs across cell types defined by Tabula Muris. This is an enlarged version of Figure 6D.

C) Clustering of Tabula muris data with 48F reprogrammed cells. Grey: Tabula Muris; blue: primary cells in Figure 2B; salmon: 48F-reprogrammed cells.