OMTN, Volume *31*

Supplemental information

Enzalutamide-induced signatures revealed by epigenetic plasticity using single-cell multi-omics sequencing in prostate cancer

Huihui Fan, Jinze Li, Astrid M. Manuel, and Zhongming Zhao

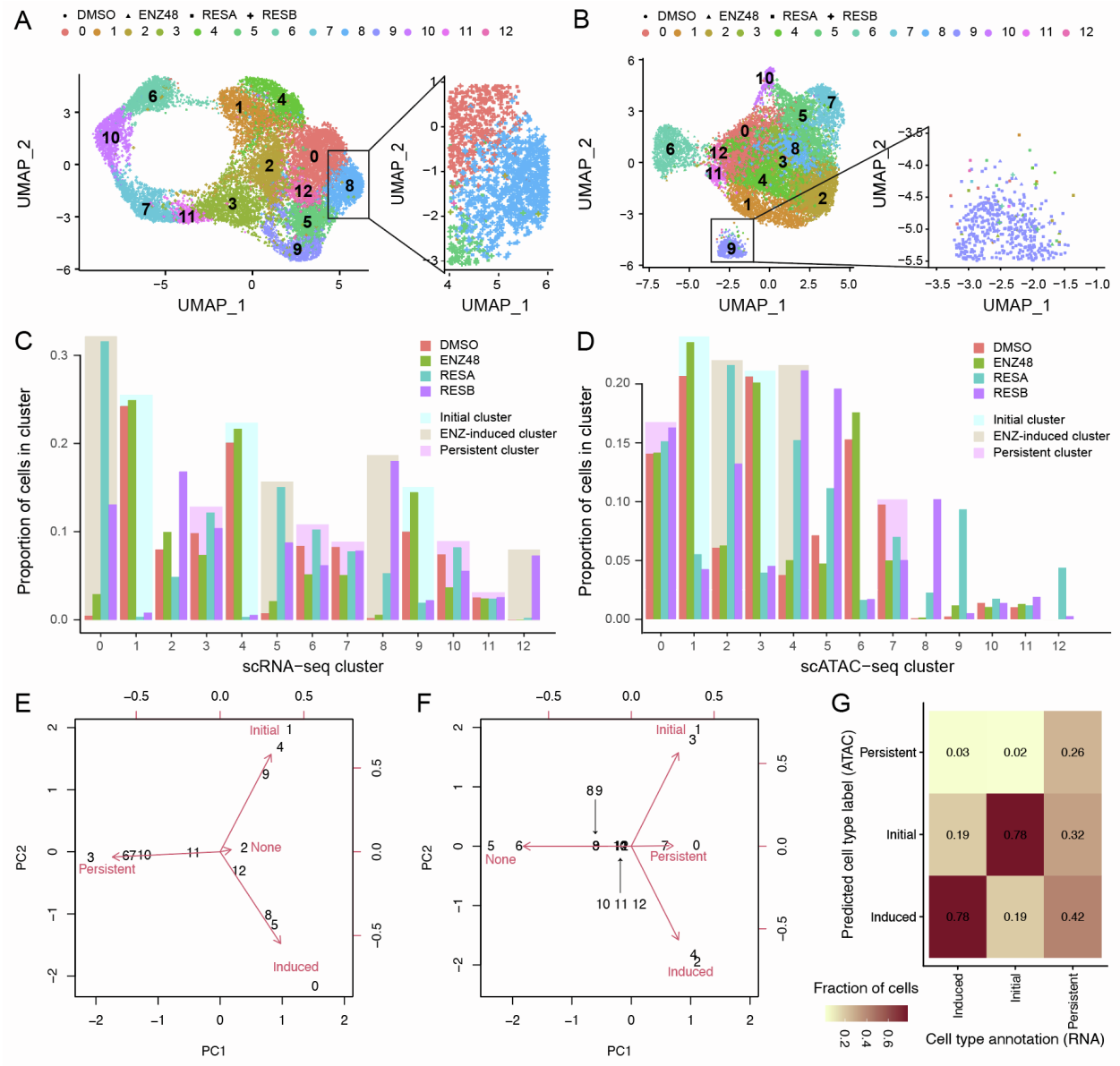

Figure S1. UMAP (Integrated uniform manifold approximation and projection) plots for scRNA-seq (A) and scATAC-seq (B) of DMSO, LNCaP-ENZ48, RES-A, and RES-B cell lines. Each dot represents a cell and colored by cluster identity. Zoom-in section shows the shapes of each dot which represents four different cell lines. Barplot showing cellular compositions of cell lines in scRNA-seq (C) and scATAC-seq (D) clusters. These clusters are further colored according to their cellular phenotypes: initial, ENZ-induced, and pre-existing ENZ-resistant. Principal component analysis (PCA) using pseudo-bulk read count tables per cluster was shown for both scRNA-seq (E) and scATAC-seq (F) data. Three cellular phenotypes are labeled in red. (G) Fraction of consistent cells being assigned to the same phenotype using label transfer technique. scRNA-seq is used as the reference, as shown on the x-axis.

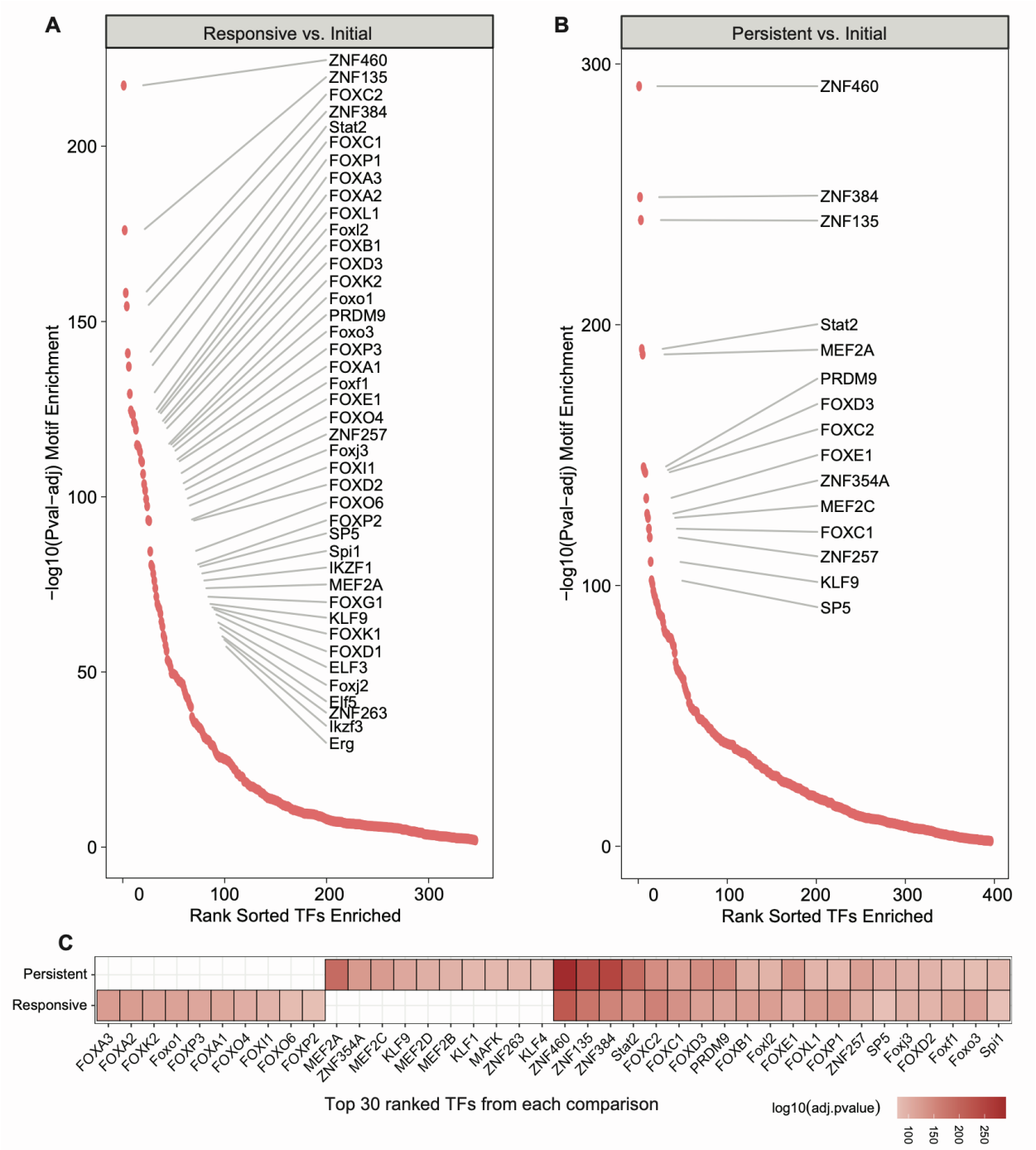

Figure S2. Transcriptional regulatory landscape in both comparisons. Transcription factor (TF) enrichment analysis using differentially accessible regions in comparison between early responsive and initial phenotype (A), and that between persistent and initial phenotype (B). Highly ranked TFs are labeled on the right side of the plot. Heatmap showing the top 30 enriched TFs per comparison. Each column is a TF, while each row is a comparison. Enrichment significant level is used as the gradient color.

Table S1. Quality control thresholds and metrics for the scRNA-seq and scATAC-seq samples.

QC thresholds for scRNA-seq							
Sample	Number of detected genes	Total number of molecules detected	Percentage of reads arising from the mitochondrial genome	Number of cells prior to QC	Number of cells after QC	Number of cells filtered out	
LNCaP	> 3000 and < 7000	> 16000 and < 50000	< 15	2358	1782	576	
LNCaP-ENZ48	> 1500 and < 5000	> 5000 and < 25000	< 15	4812	4315	497	
RES-A	> 1500 and < 5000	> 5000 and < 25000	< 17	5156	4569	587	
RES-B	> 1500 and < 5000	> 5000 and < 25000	< 20	4907	4061	846	
QC thresholds for scATAC-seq							
Sample	Total number of fragments in peaks	Fraction of fragments in peaks	Strength of nucleosome binding pattern	Transcription start site enrichment score as defined	Number of cells prior to QC	Number of cells after QC	Number of cells filtered out
LNCaP	> 2000 and < 20000	> 30	< 9	> 2	4436	3284	1152
LNCaP-ENZ48	> 1000 and < 20000	> 30	< 9	> 2	3376	3115	261
RES-A	> 2000 and < 20000	> 40	< 8	> 2	4407	3823	584
RES-B	> 2000 and < 25000	> 30	< 8	> 2	4747	3227	1520

Table S2. Differentially expressed genes identified from both comparisons.

Table S3. Differentially accessible regions identified from both comparisons.

Table S4. Signature gene sets with bi-directional epigenetic regulations.

Table S5. Gene set enrichment analysis using signature gene sets.

Table S6. Gene set enrichment analysis showing only Gene Ontology and KEGG pathway terms.

Table S7. Survival analysis showing all the genes both significant and not significant.