

Perturb-seq reveals TCF7 as a transcriptional link between MAPK- and Wnt-driven gene expression

AUTHORS

Ghanem El Kassem¹, Anja Sieber², Bertram Klinger², Florian Uhlitz², David Steinbrecht^{2,3}, Mirjam van Bentum^{2,6}, Shawez Khan¹, Jasmine Hillmer¹, Jennifer von Schlichting², Reinhold Schäfer^{4,5}, Nils Blüthgen^{*# 2,3,5}, Michael Boettcher^{*# 1}

¹ Institute of Molecular Medicine, Section for Molecular Medicine of Signal Transduction, Faculty of Medicine, Martin-Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany

² Institute of Pathology, Charité - Universitätsmedizin Berlin, Charitéplatz 1, 10115 Berlin, Germany

³ Institut für Biologie, Humboldt Universität zu Berlin, Haus 18, Philippstr. 13, 10115 Berlin

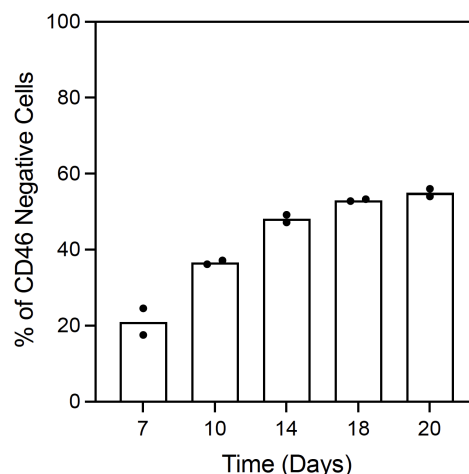
⁴ Comprehensive Cancer Center, Charité - Universitätsmedizin Berlin, Charitéplatz 1, 10115 Berlin, Germany

⁵ German Consortium for Translational Cancer Research (DKTK)

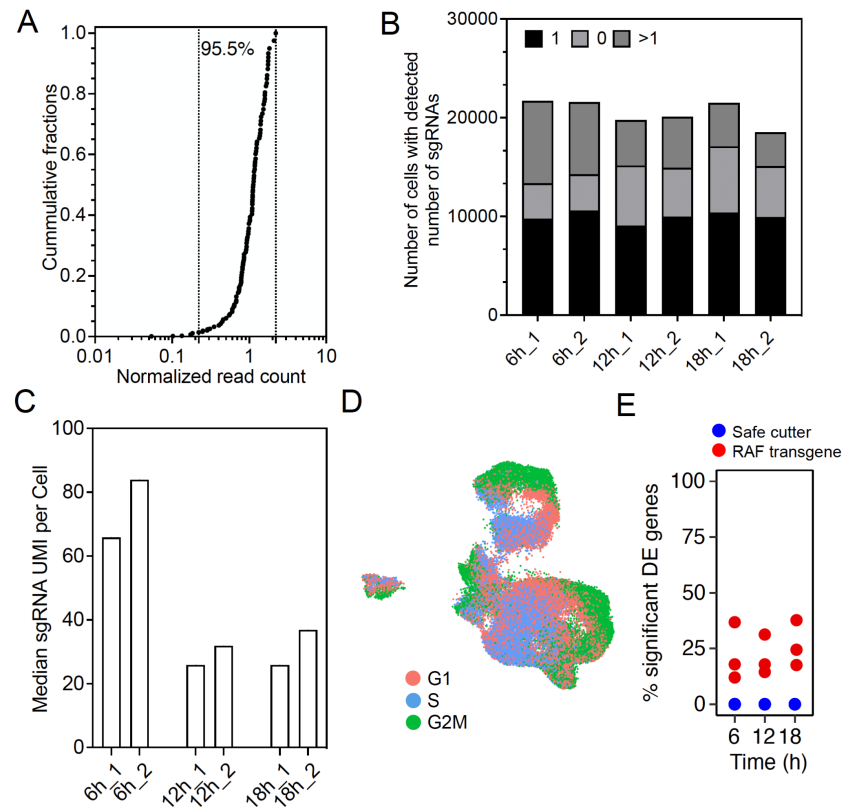
⁶ Max Delbrück Center for Molecular Medicine, Robert-Rössle-Straße 10, 13125 Berlin, Germany

* To whom correspondence should be addressed. Email: michael.boettcher@medizin.uni-halle.de. Correspondence may also be addressed to: nils.bluthgen@charite.de

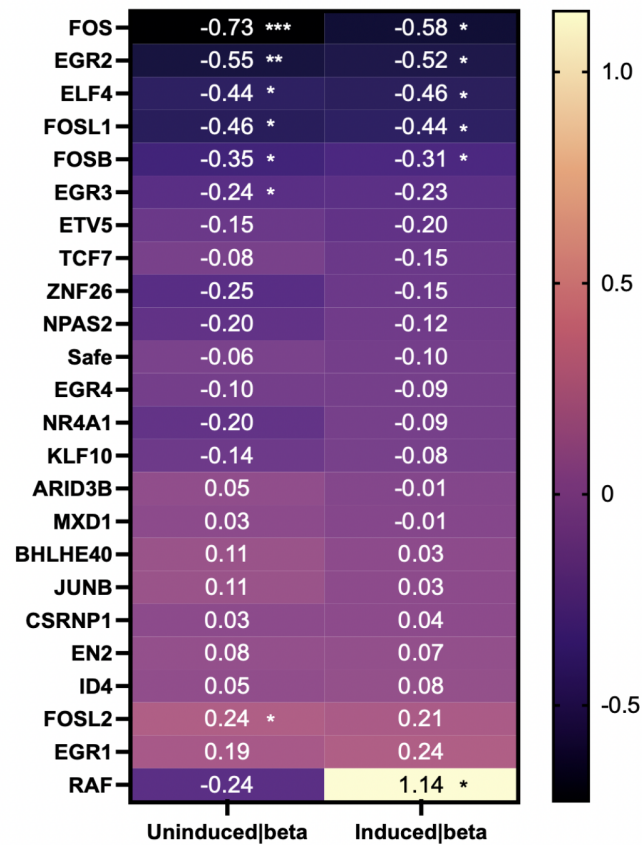
Supplementary Figures



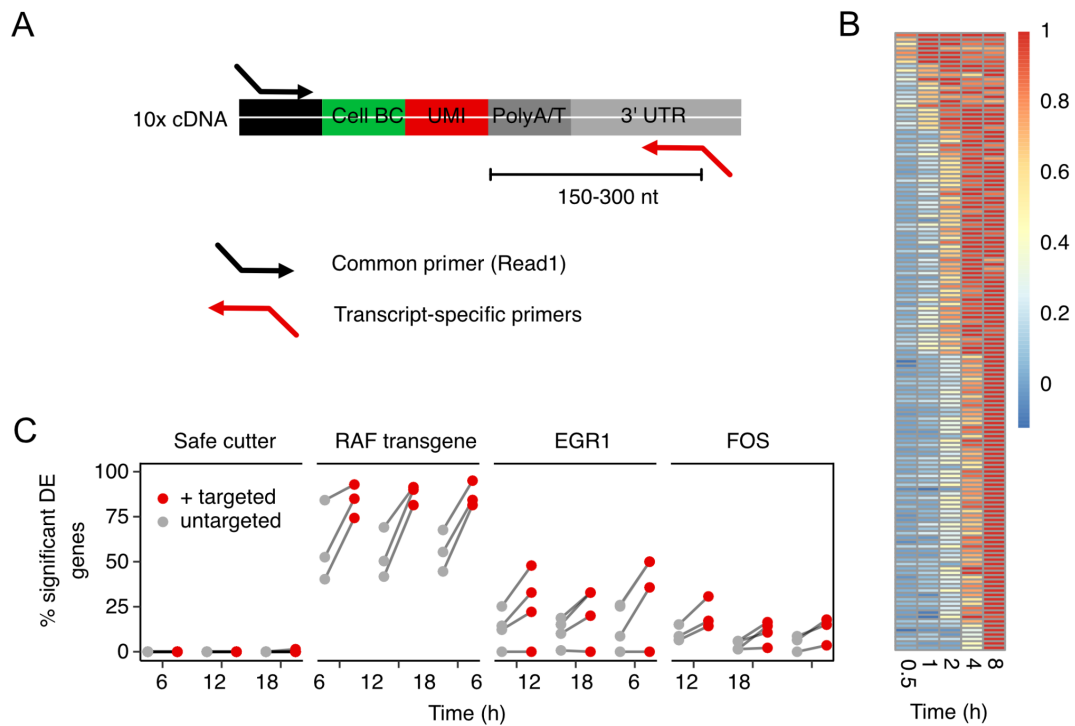
Supplementary Fig. 1 - CD46 knockout kinetics in HEK293ΔRAF1:ER cells at different time points after lentiviral infection. Percentage of CD46 knockout cells was determined via flow cytometry analysis of >10,000 cells stained with CD46 antibodies (Miltenyi). CD46 was used as a surrogate marker for perturbation efficiency because it is a ubiquitously expressed, non-essential surface protein whose loss can be quantified sensitively by flow cytometry.



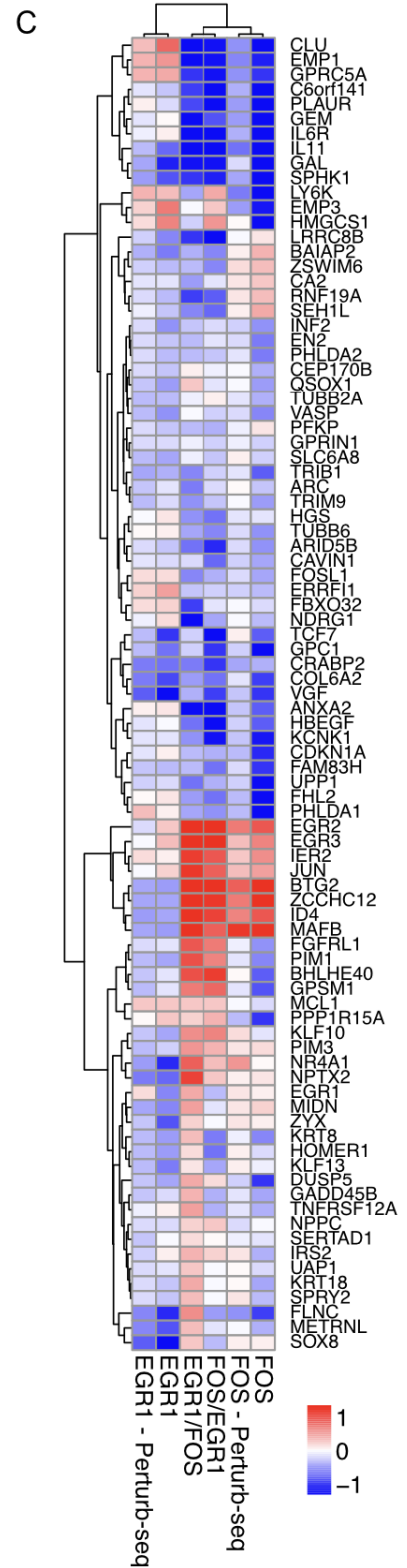
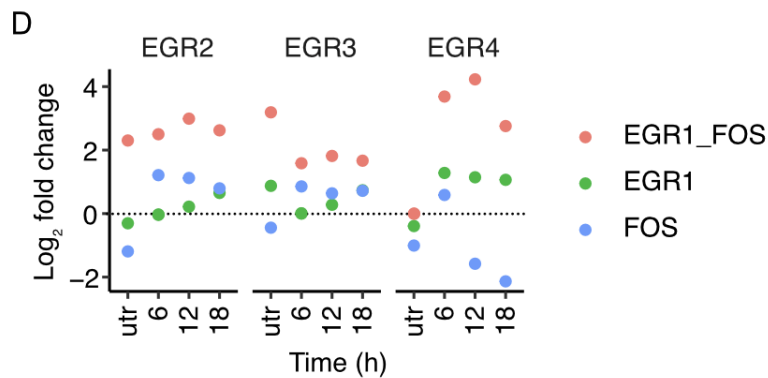
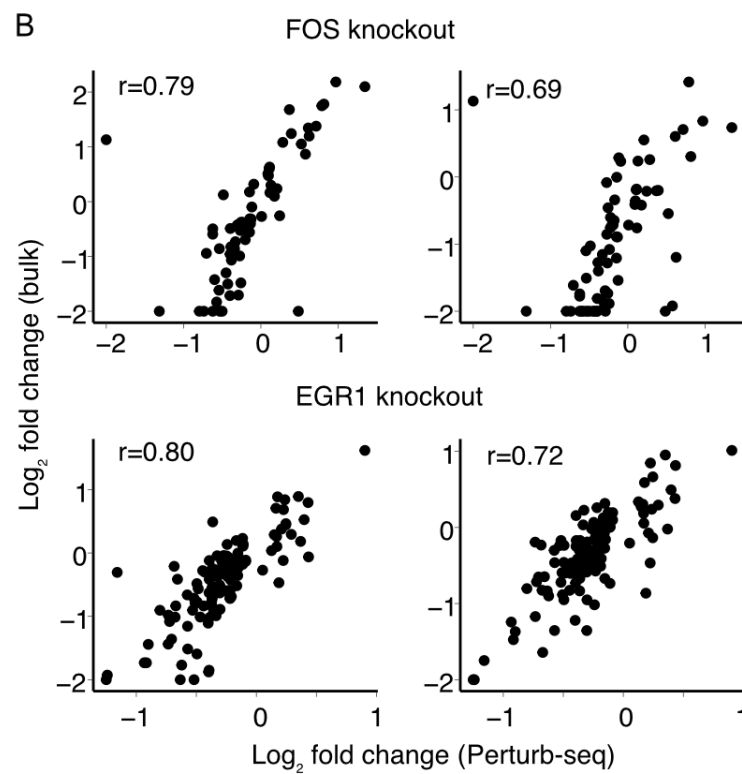
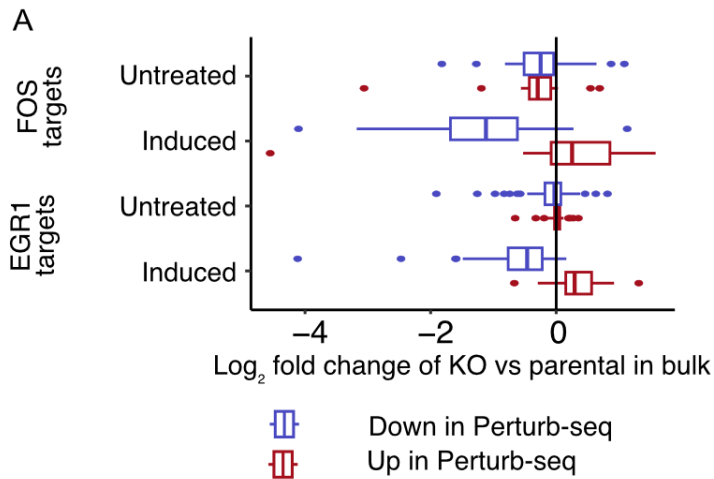
Supplementary Fig. 2 - Quality assessment of Perturb-seq screen performance. (A) Distribution of the pooled sgRNA library used for all Perturb-seq and proliferation screens. (B) Total number of recovered cells and number of cells with 0, 1, and >1 sgRNAs detected in the respective Perturb-seq samples. (C) Median sgRNA UMI counts per cell detected in the respective Perturb-seq samples. (D) Distribution of cell cycle phases G1, S, and G2M on the integrated UMAPs. (E) Percentage of significantly differentially expressed genes in the RAF1-knockout cells from the total number of genes induced by RAF activation (adjusted p-value<0.05). Red = RAF1 sgRNAs, Blue = safe cutter sgRNAs.



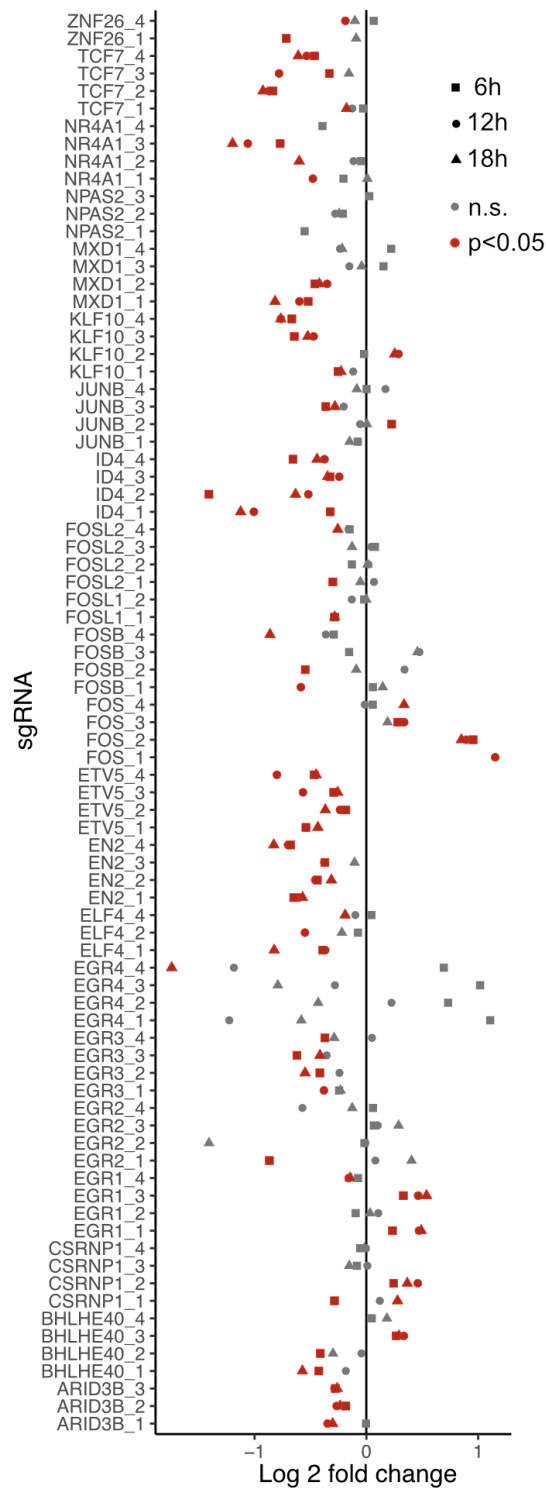
Supplementary Fig. 3 - CRISPR/Cas9 proliferation screen in HEK293ΔRAF1:ER cells. MAGECK MLE beta scores and corresponding significance are shown. * Wald FDR <0.05, ** Wald FDR <0.01, * Wald FDR <0.001**



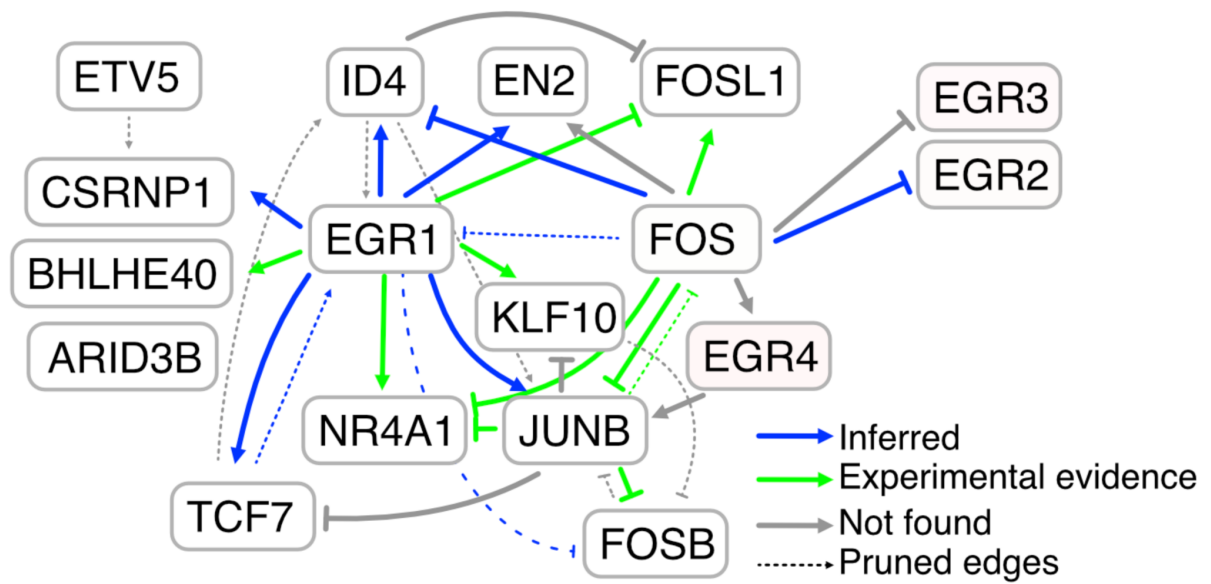
Supplementary Fig. 4 - Targeted scRNA-Seq library amplification enhances Perturb-seq sensitivity. (A) Schematic of the modified TAP-seq approach. **(B)** Selection of 140 candidate genes for modified TAP-seq. Log_2 gene expression fold changes from bulk RNA-Seq analysis of significantly induced genes after pulse induction of the RAF1 transgene for different times (FDR = 1%), normalized to gene-wise maximum Log_2 fold changes. **(C)** Difference in percentage of significant differentially expressed genes from the 140 selected genes for the modified TAP-seq between targeted and untargeted approach after knockout of RAF1 transgene, EGR1, FOS, and safe cutter control.



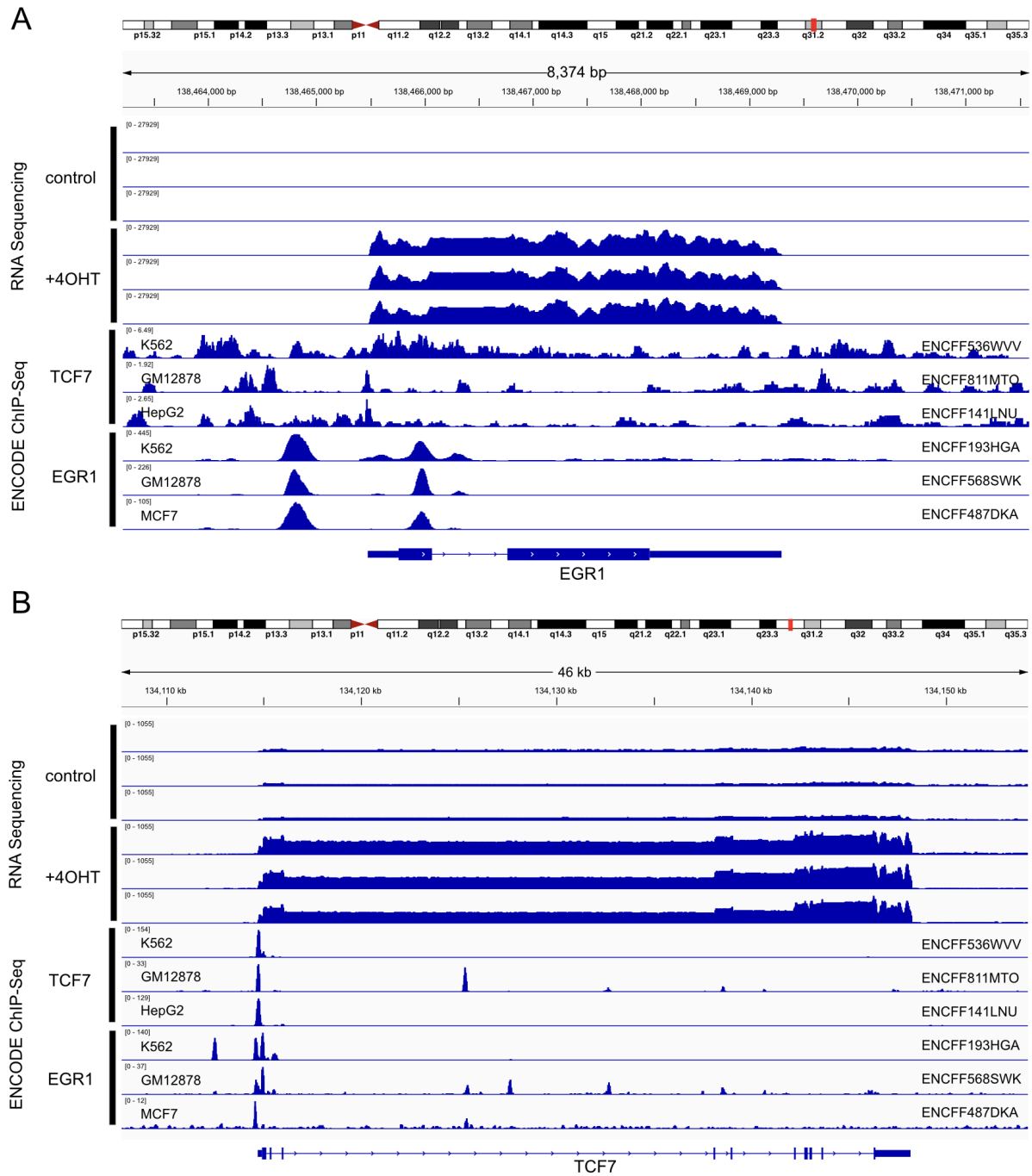
Supplementary Fig. 5 - Bulk RNA-Seq analysis of individual and combinatorial EGR1 and FOS perturbations. **(A)** Average $\log_2$ fold changes for genes that are identified as up- and down-regulated target genes of EGR1 and FOS in perturb-seq for bulk KO clones compared to parental clones. **(B)** Correlation between transcriptional changes detected via Perturb-seq and bulk RNA-Seq of EGR1- or FOS-perturbed cells from 2 different single knockout clonal lines of each gene. **(C)** Heatmap of transcriptional changes detected via bulk RNA-Seq from cells with individual or combinatorial EGR1 and FOS perturbations. Perturb-seq results from individual EGR1 and FOS perturbations are shown for comparison. **(D)** EGR2, EGR3, and EGR4 transcript levels determined via bulk RNA-Seq of EGR1/FOS single or double knockout clones, induced with 4OHT for 0h (utr), 6h, 12h, or 18h.



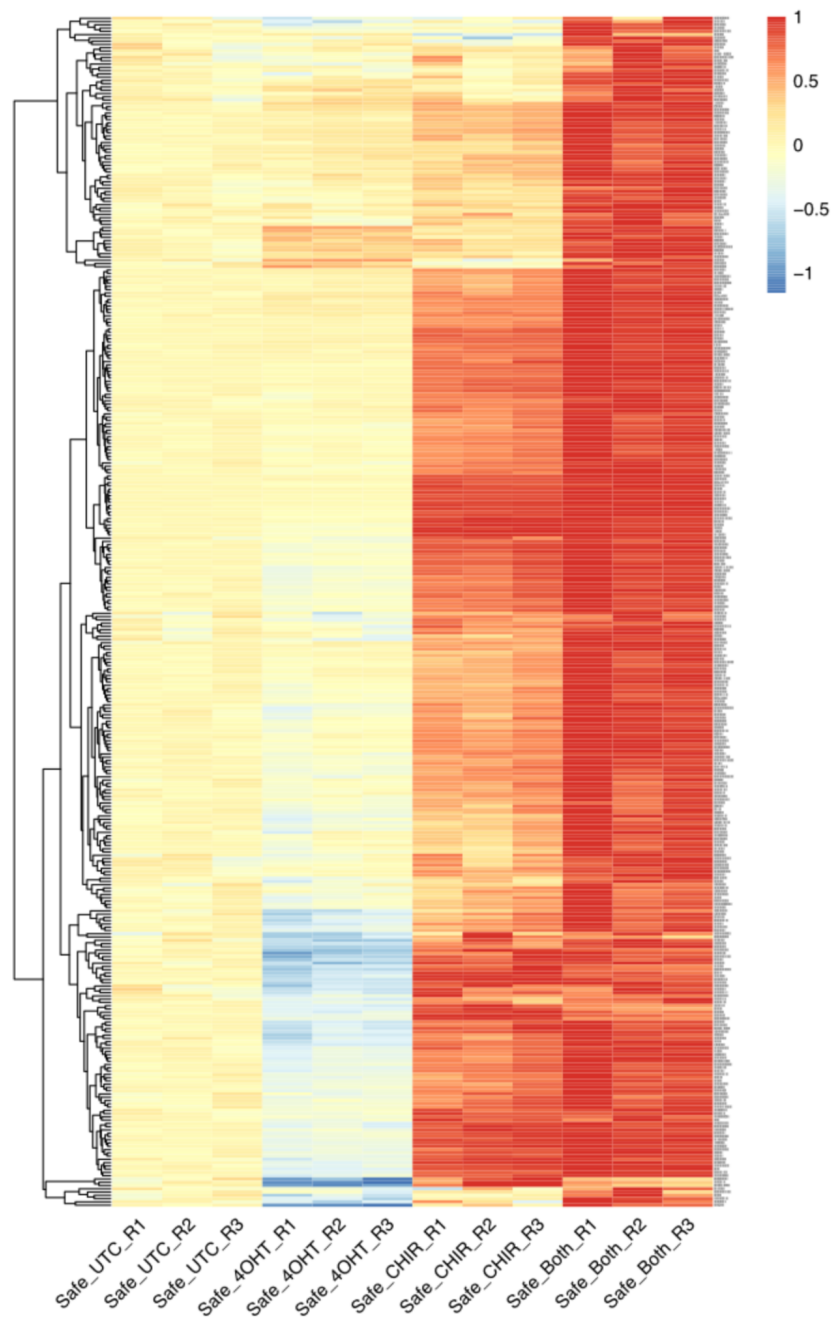
Supplementary Fig. 6 - Target gene expression changes following sgRNA perturbation. Log₂ fold change in target gene expression for individual sgRNAs at 6 h, 12 h, and 18 h after RAF1:ER activation, measured from Perturb-seq data. Red: significant changes compared to non-targeting controls ($p < 0.05$); Grey: non-significant changes.



Supplementary Fig. 7 - De novo model of the TF core network highlighting interactions represented in the OmniPath database. Directional transcriptional interactions between all perturbed TFs are shown. Blue edges indicate interactions listed in OmniPath without associated literature references; green edges indicate interactions in OmniPath supported by published experimental evidence; grey edges indicate interactions not found in the OmniPath database; dotted edges indicate interactions that were pruned in our de novo transcription model in Figure 3E.



Supplementary Fig. 8 - Public ENCODE ChIP-seq tracks showing promoter-proximal binding of EGR1 and TCF7 at the EGR1 and TCF7 loci. (A) Genome browser view of the EGR1 locus showing full-length RNA-seq signal from control and 4OHT-treated HEK293ΔRAF1:ER cells, together with ENCODE ChIP-seq tracks for TCF7 and EGR1 across the indicated cell lines. **(B)** Genome browser view of the TCF7 locus showing RNA-seq signal from control and 4OHT-treated HEK293ΔRAF1:ER cells and ENCODE ChIP-seq tracks for TCF7 and EGR1. RNA-seq coverage confirms expression of TCF7 transcripts in 4OHT-treated cells.



Supplementary Fig. 9 - Genes synergistically activated by combined MAPK and Wnt pathway activation. Heatmap of transcriptional changes detected via bulk RNA-Seq from cells treated with 0.5 μ M 4OHT for 12 h, 10 μ M CHIR99021 for 6 h, or both. Shown are the 364 Wnt signaling-associated genes identified as synergistically activated by the combined treatment. Color scale represents scaled expression values.