iScience, Volume *24*

Supplemental information

Systems-level network modeling deciphers the master regulators of phenotypic plasticity and heterogeneity in melanoma

Maalavika Pillai and Mohit Kumar Jolly

Supplementary Figures

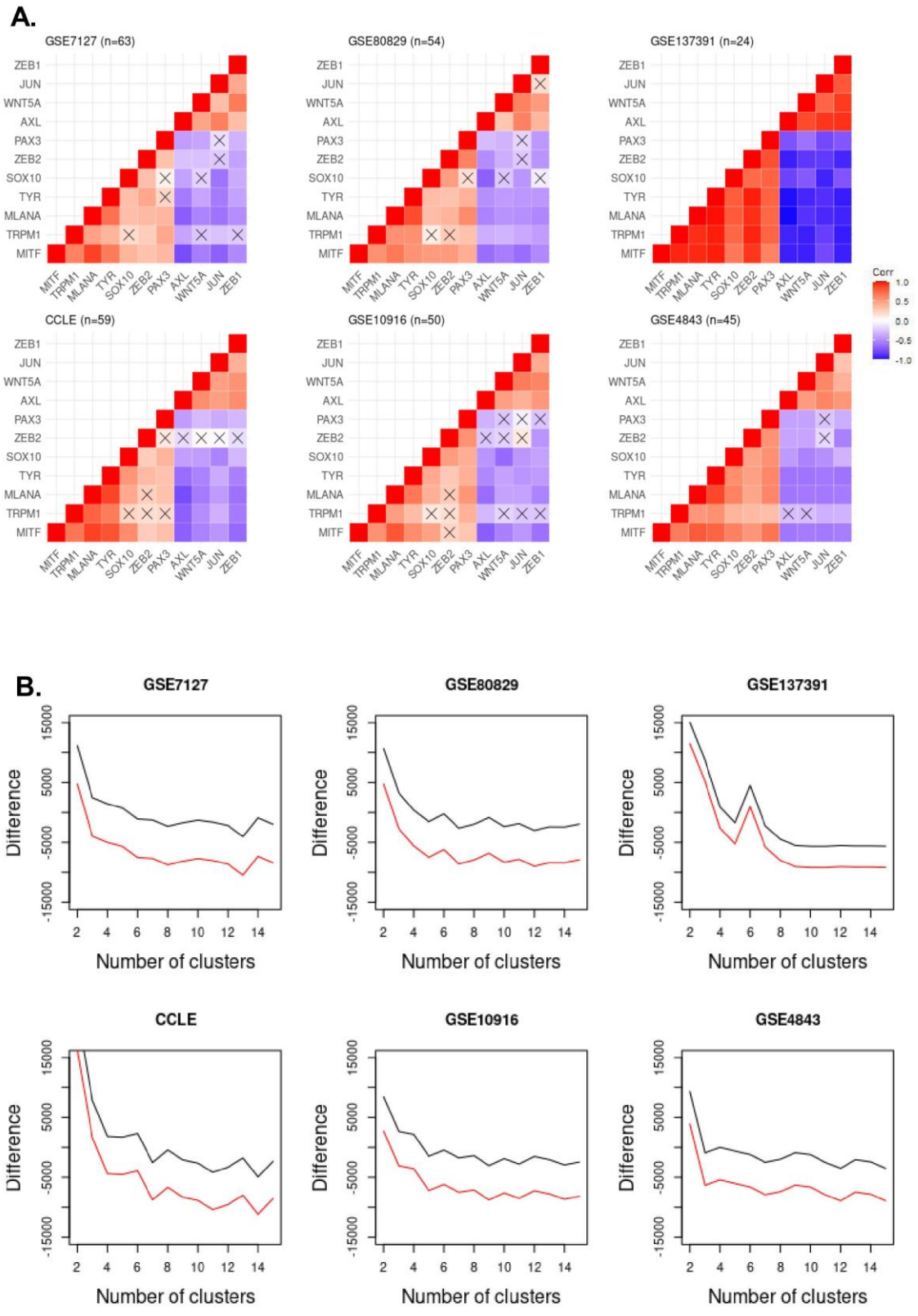

Fig S1. Two distinct groups of genes give rise to phenotypic heterogeneity in, related to Figure 1. **A.** Pearson's Correlation of regulators of phenotypic heterogeneity for GSE7127, GSE80829, GSE137391 (left to right, top panel) and CCLE, GSE10916, GSE4843 (left to right, bottom panel) . Crosses indicate $p > 0.05$. Colorbar denotes correlation coefficient. 'n' stands for number of samples in each dataset. **B.** Difference between n^{th} and $(n-1)^{th}$ AIC (red line) and BIC (black line) scores for 2-15 clusters for GSE7127, GSE80829, GSE137391 (left to right, top panel) and CCLE, GSE10916, GSE4843 (left to right, bottom panel).

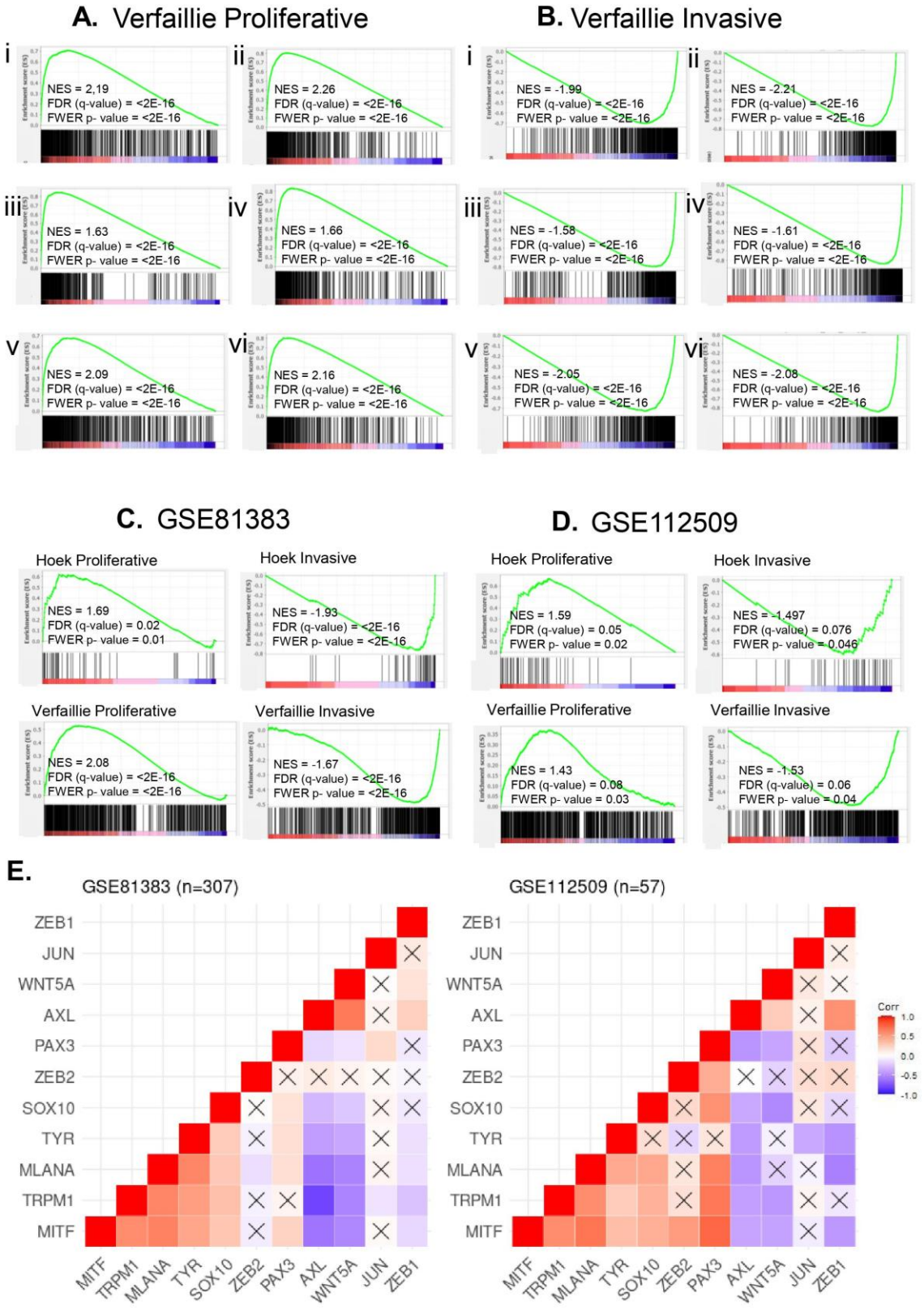

Fig S2. Two classes of samples (proliferative and invasive) exist across multiple datasets, related to Figure 2. A,B GSEA for Verfaillie proliferative and invasive geneset (respectively) for **i**. GSE7127 (n=63) **ii**. GSE80829 (n=53) **iii.** GSE137391 (n=24) **iv.** CCLE (n=59) **v.** GSE10916 (n=50) **vi.** GSE4843 (n=45). GSEA for Hoek Proliferative, Hoek invasive (left to right, top panel), Verfaillie proliferative and Verfaillie invasive geneset (left to right, bottom panel), in GSE81383 (**C.**) and GSE112509 (Primary tumour cells only) (**D.**). **E.** Spearman's Correlation Coefficient matrix for regulators of phenotypic heterogeneity for GSE81383 and GSE112509 (Primary tumour cells only). Crosses indicate $p > 0.05$. Colorbar denotes correlation coefficient. 'n' stands for number of samples in each dataset.

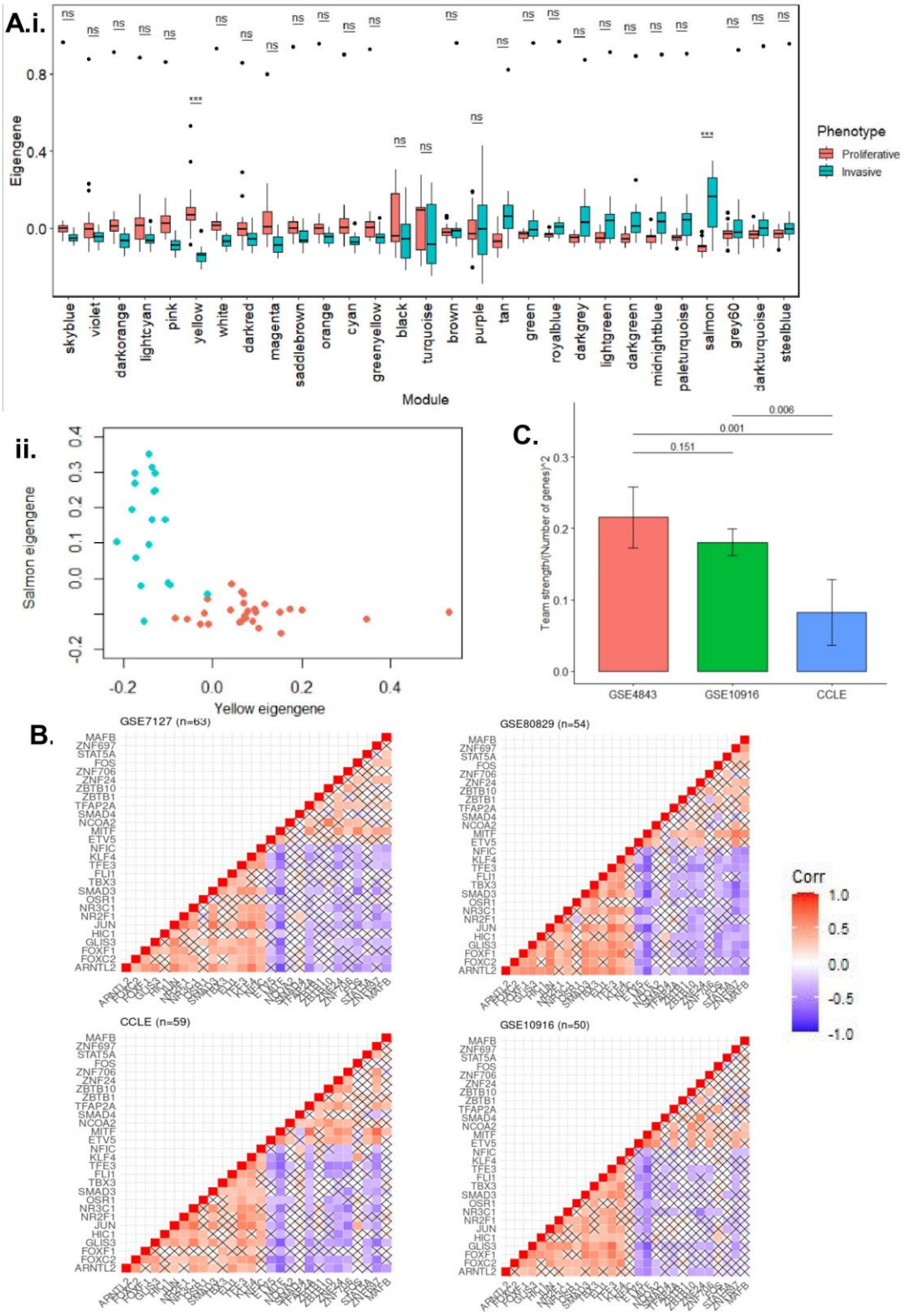

Fig S3 Identification of key regulators, related to Figure 3. A. Eigengene values for all modules in proliferative and invasive samples for GSE4843. Significance is represented for Bonferonni adjusted p-value using * for p< 0.001, ** for p < 0.0001, *** for p < 0.00001. **ii**. Module eigengene scatterplot for proliferative (Orange) and invasive (cyan) samples for Salmon (Invasive) and yellow (Proliferative) modules. **B.** Spearman's Correlation Coefficient for genes identified as regulators from GSE4843 in CCLE, GSE7127,10916 and GSE80829. Crosses indicate p > 0.05. Colorbars denote correlation coefficient. 'n' stands for number of samples in each dataset. **C.** Average team strength for candidate master regulators identified from GSE4843, GSE10916 and CCLE.

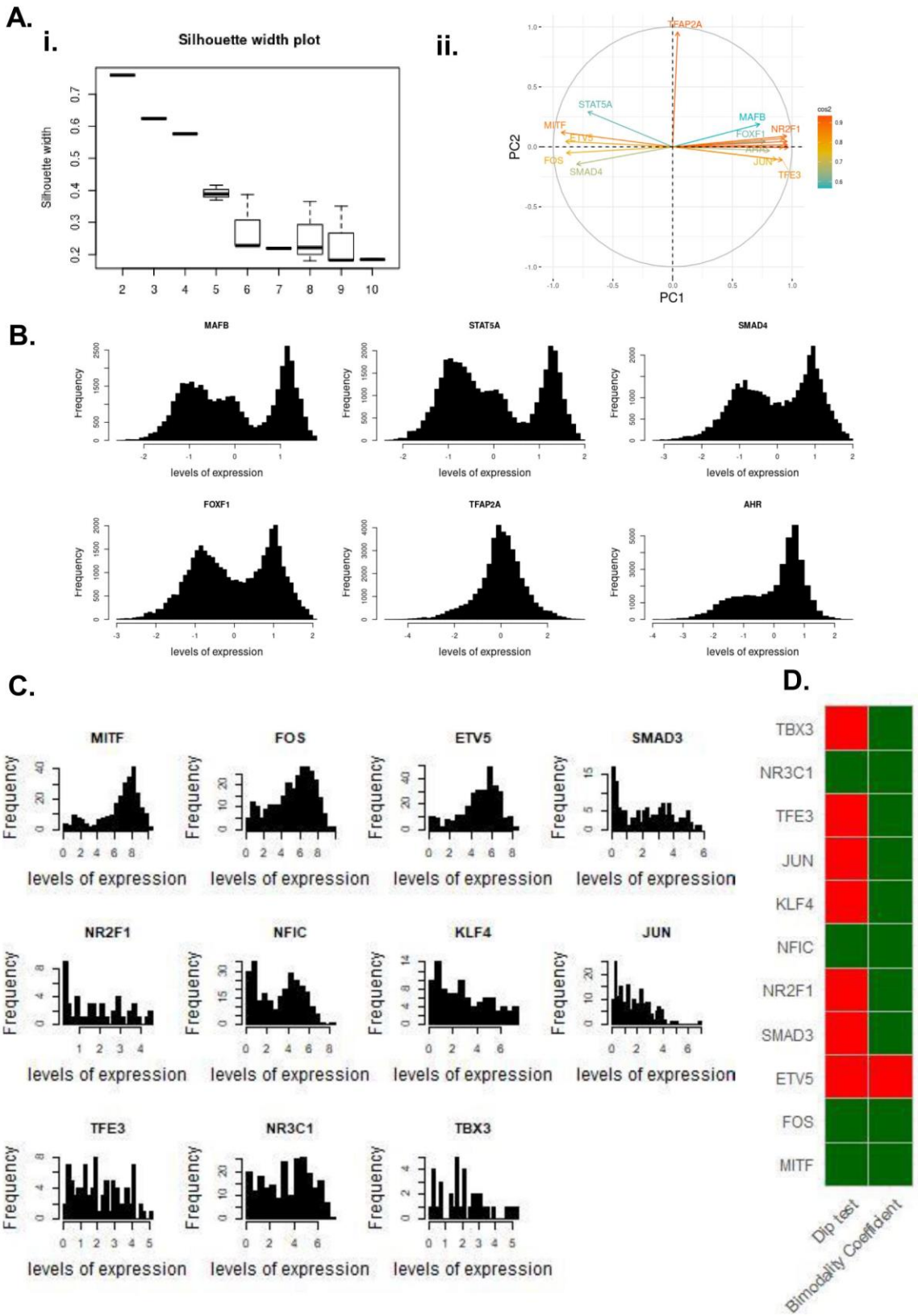

Fig S4. **Simulations identify 2 clusters of steady state solutions, related to Figure 4. A. i.** Average silhouette width for k-means clustering measuring for varying number of clusters (k = 2 to 10). Error bars indicate standard deviation. **ii.** Correlation circle with squared cosines of all genes in the network for PC1 and PC2. **B.** Histograms for gene expression simulated using RACIPE. **C.** Histograms for gene expression in GSE81383 **D.** Bimodality test using Bimodality coefficient (BC) (Green boxes indicate BC > 0.555 and red boxes indicate BC < 0.555) and Hartigan's dip test (Green boxes indicate $p < 0.05$ and red boxes indicate $p > 0.05$) for GSE81383.

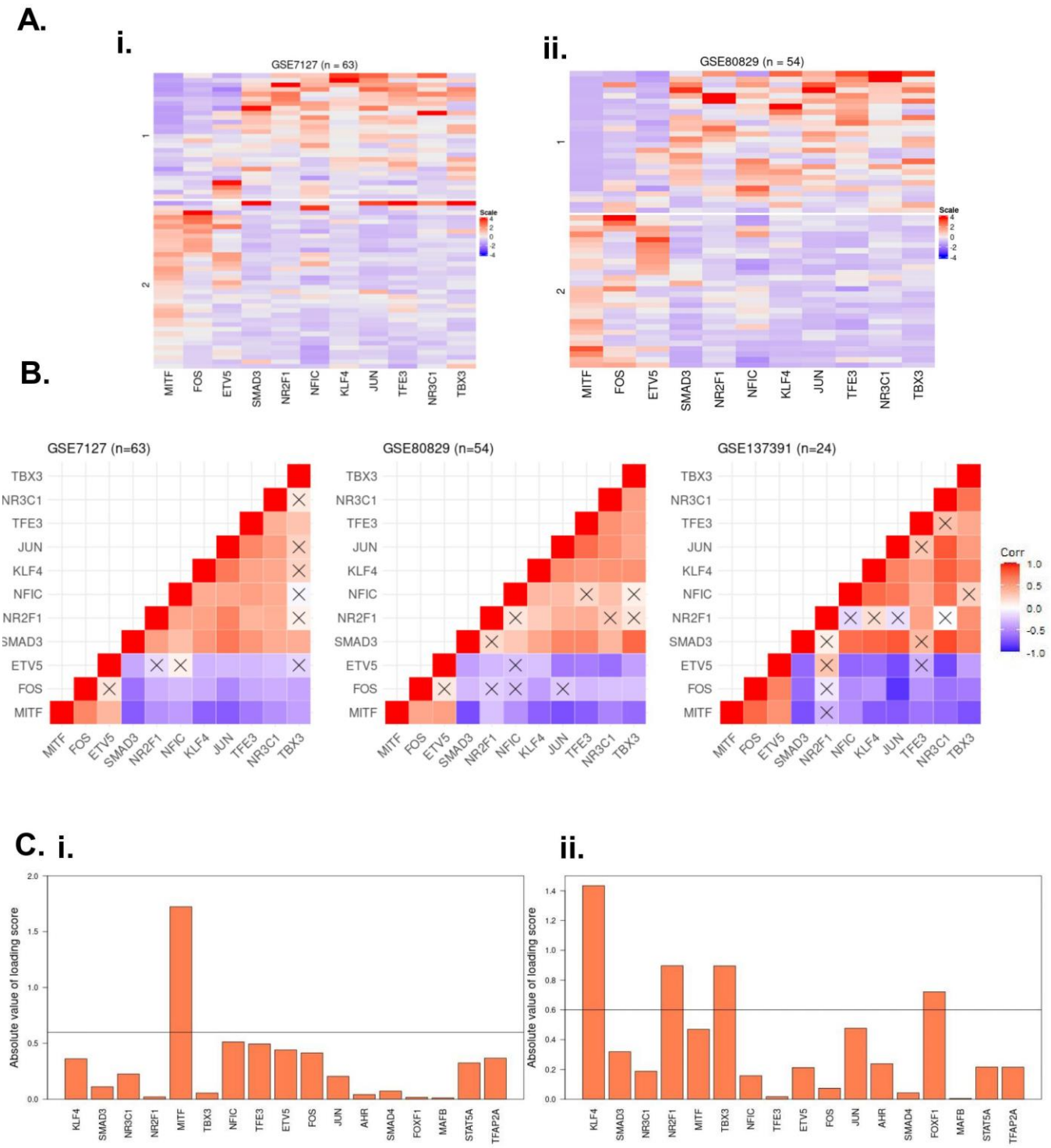

Fig S5. Identification of key master regulators that distinguish between phenotypes, related to Figure 5. A. Differential expression of master regulators in proliferative and invasive samples in GSE7127 and GSE80829 **B.** Spearman's Correlation Coefficient matrix for a subset of network genes in GSE7127 , GSE80829 and GSE137391. 'n' stands for number of samples in each dataset. **C.** LDA loading scores for master regulators in **i.** Proliferative subclusters and **ii.** Invasive subclusters. Cut off value of 0.6 was set to select genes.

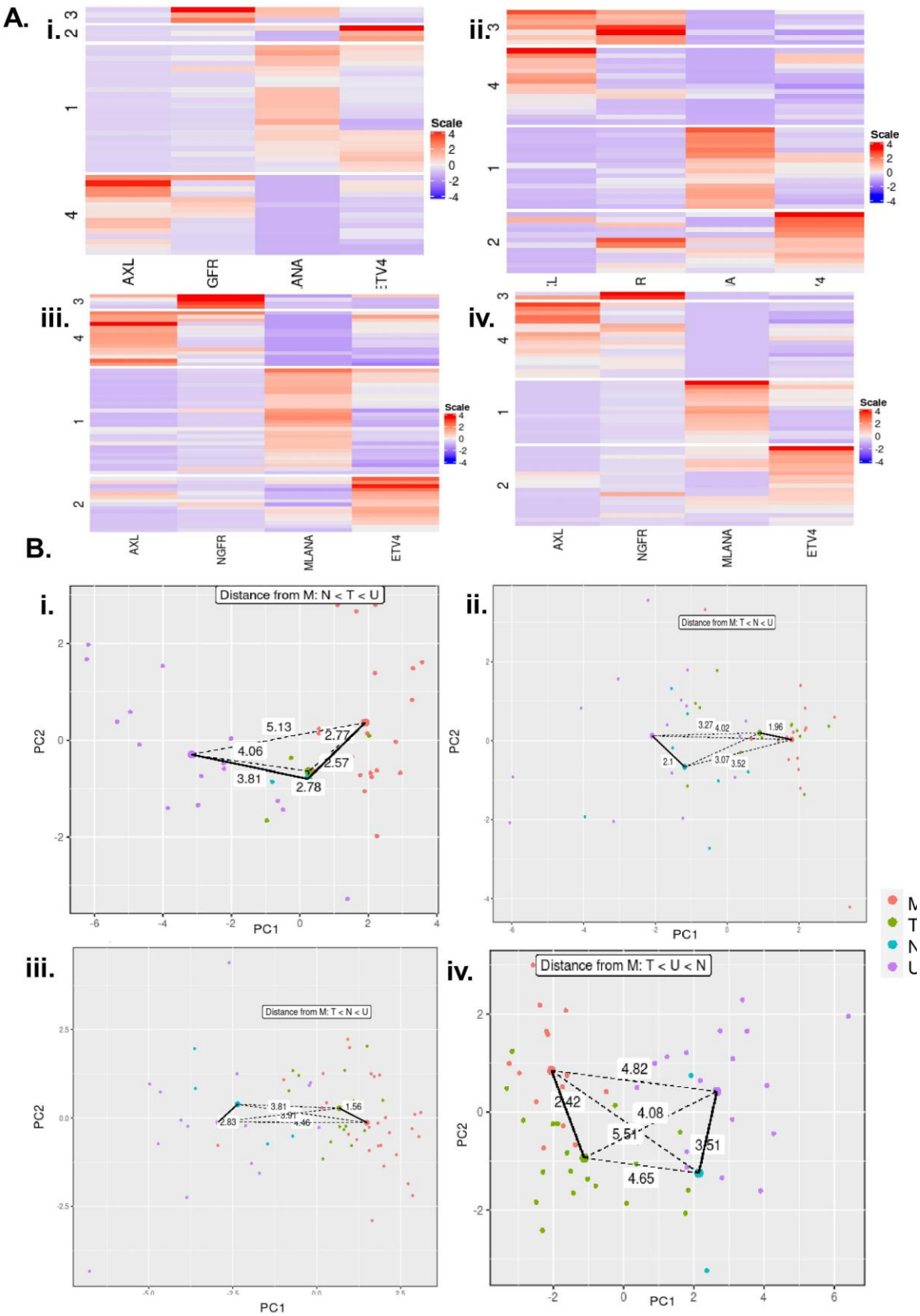

Fig S6. Network genes explain the distinct de-differentiation trajectory observed in melanoma cells, related to Figure 6. A. Heatmap of marker genes expression levels for the 4 phenotypes (AXL, NGFR, MLANA and ETV4) for **i.** GSE4843, **ii.** GSE10916 **iii**. GSE7127 **iv**. GSE80829. **C.** All possible trajectories (dotted line), corresponding Euclidean distance and shortest distance to a cluster centre from melanocytic and undifferentiated clusters (solid line) [right] for **i.** GSE4843 **ii**. GSE10916 **iii.** GSE7127 **iv.** GSE80829.

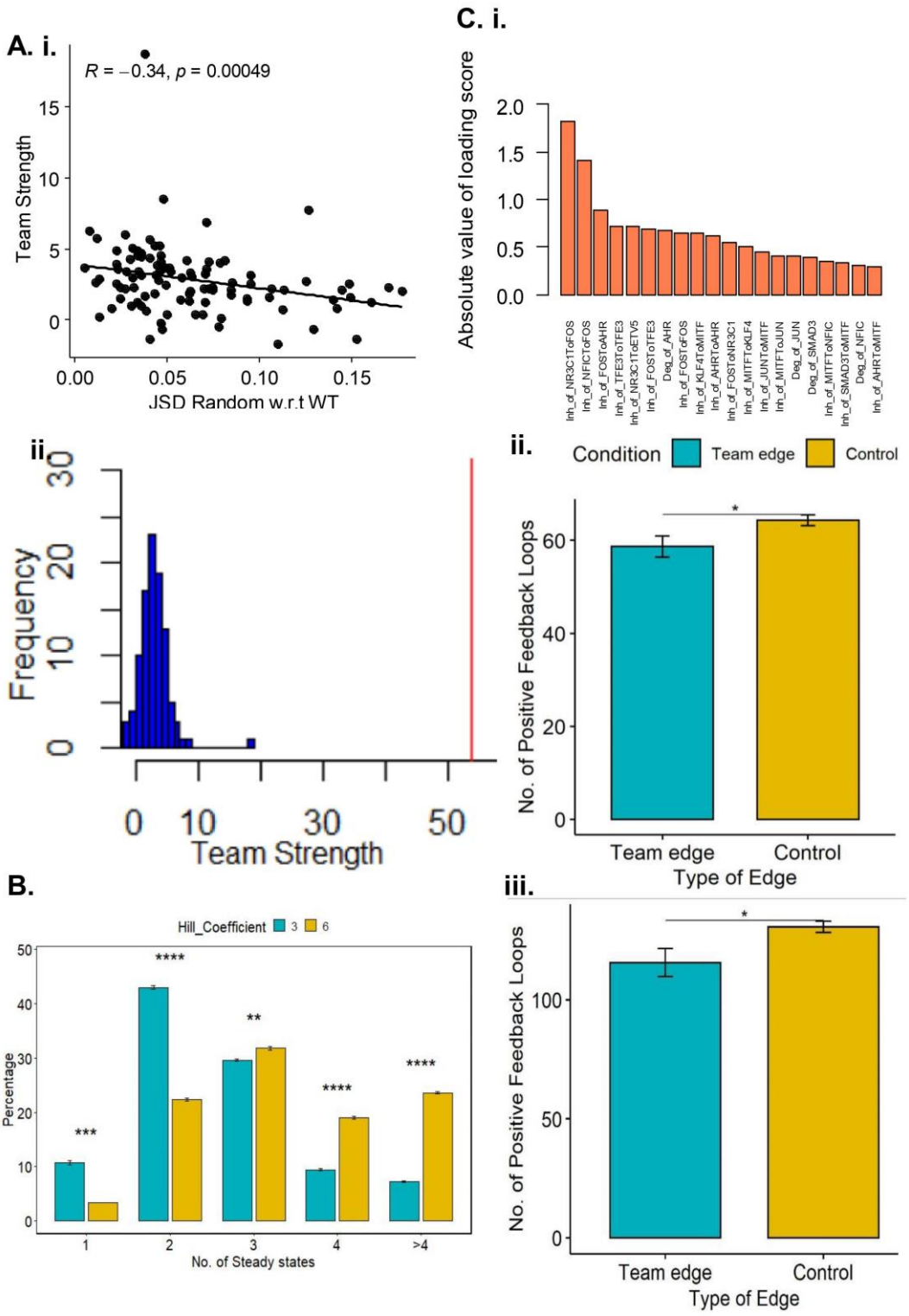

Fig S7. Analysis od design principles of the network, related to STAR Method Edge Removal. A. i. Spearman correlation between team strength for random networks and JSD between distribution of states in the WT and random network. **ii.** Histogram for team strengths of 100 random networks. Red line indicates the team strength metric of the WT network **B.** Distribution of models giving rise to 1,2,3,4 or >4 steady states with maximum Hill coefficient set to 3 and 6 in RACIPE simulations (error bars show standard deviation over three replicates). **C. i**. Absolute value of loading score for LD1 segregating parameters contributing to monostability v/s multistability. Average number of positive feedback loops having less than **ii.** 5 edges and **iii.** 6 edges in case of Team edge removal and controls. ***** indicates p-value <0.05.

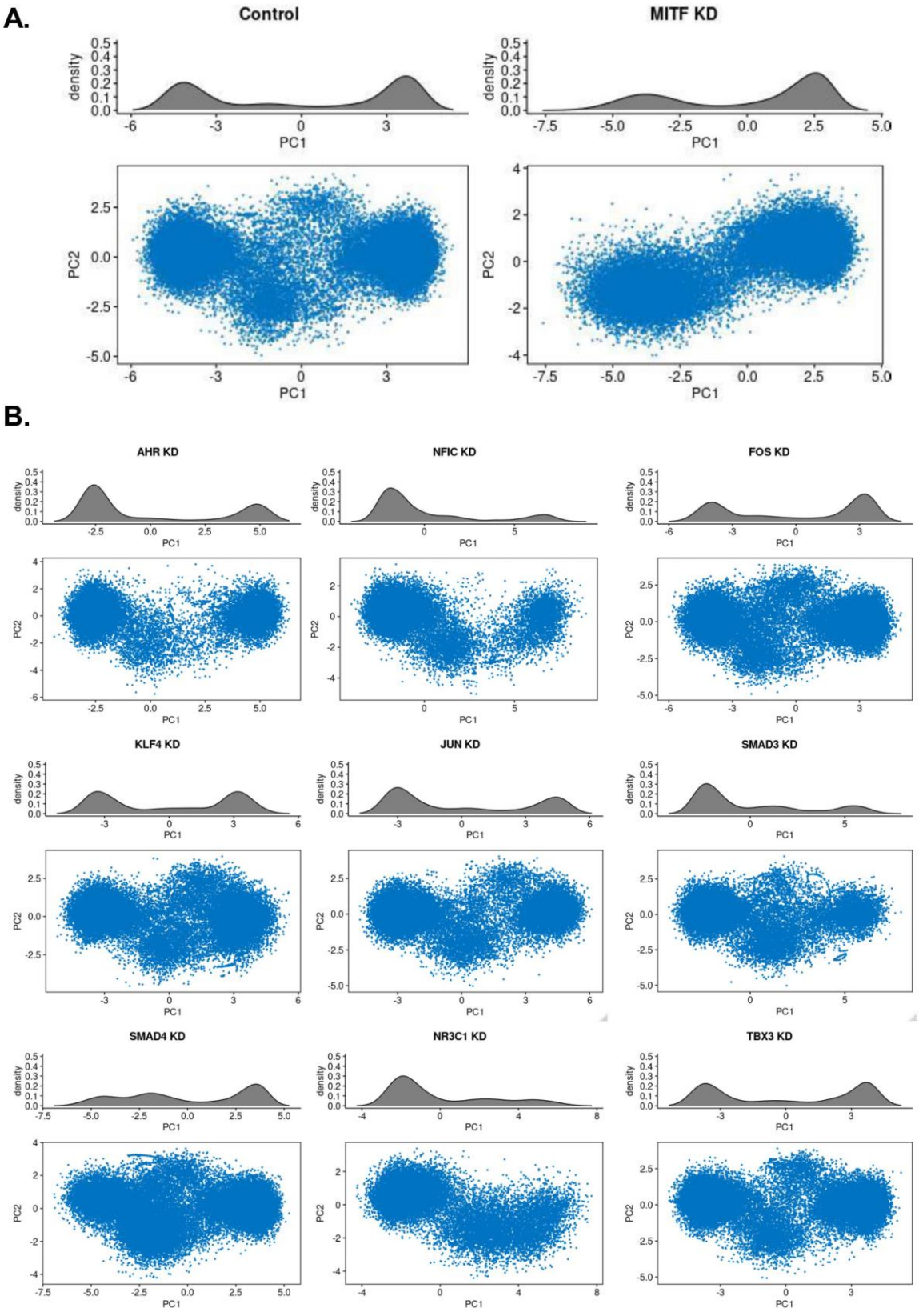

Fig S8. MITF Knockdown recapitulates phenotype switching seen during BRAFi/MAPKi, related to Figure 7. Density distribution of datapoints for **A.** complete network and MITF knock down **B.** Knock-down of rest of the genes in the network.