

SUPPLEMENTAL INFORMATION

A TFAP2C Gene Signature is Predictive of Outcome in HER2-positive Breast Cancer

Vincent T. Wu^{1,8}, Boris Kiriazov^{1,8}, Kelsey E. Koch¹, Vivian W. Gu², Anna C. Beck¹, Nicholas Borcharding³, Tiandao Li¹, Peter Addo¹, Zachary J. Wehrspan⁴, Weizhou Zhang⁵, Terry A. Braun⁶, Bartley J. Brown⁶, Vimla Band⁷, Hamid Band⁷, Mikhail V. Kulak¹, Ronald J. Weigel^{1,2,4*}

¹Department of Surgery, University of Iowa, Iowa City, IA, 52242 USA

²Department of Molecular Physiology and Biophysics, University of Iowa, Iowa City, IA, 52242 USA

³Department of Pathology, University of Iowa, Iowa City, IA, 52242 USA

⁴Department of Biochemistry, University of Iowa, Iowa City, IA, 52242 USA

⁵Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, FL 32610 USA

⁶Department of Biomedical Engineering, University of Iowa, Iowa City, IA, 52242 USA

⁷Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, NE, 68198 USA

⁸These authors contributed equally

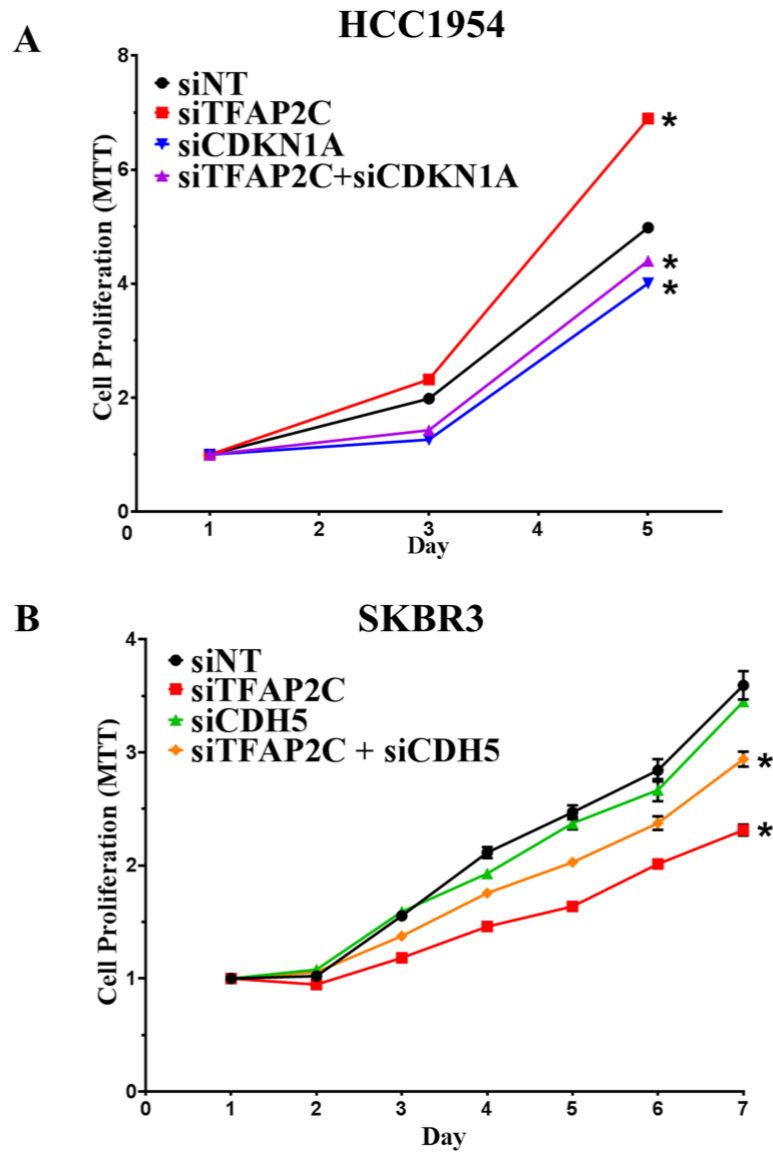


Figure S1. Proliferative Response to CDKN1A and CDH5 with Knockdown of TFAP2C. **A.** HCC1954 cells with knockdown of *TFAP2C*, *CDKN1A* or both showing proliferation determined by cell viability measured by MTT assay; these data were performed in parallel with data presented in Figure 5. **B.** Parallel experiments in SKBR3 cells with knockdown of *TFAP2C*, *CDH5* or both showing proliferation determined by cell viability measured by MTT assay; these data were performed in parallel with data presented in Figure 5.

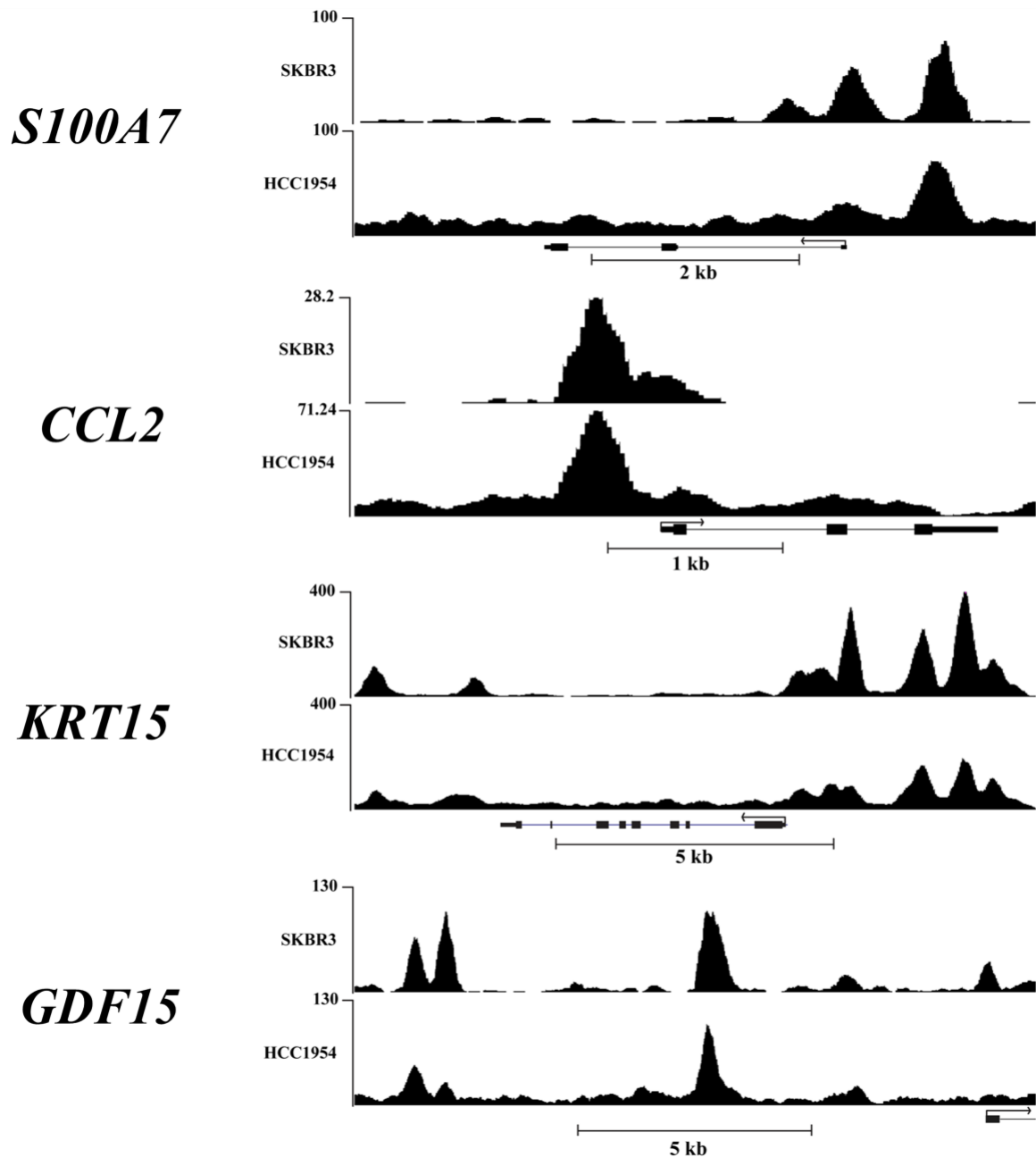


Figure S2. ChIP-seq for Similarly Regulated Genes. Examples of the regulatory regions of four genes of showing AP-2 γ occupancy for SKBR3 and HCC1954 cells.

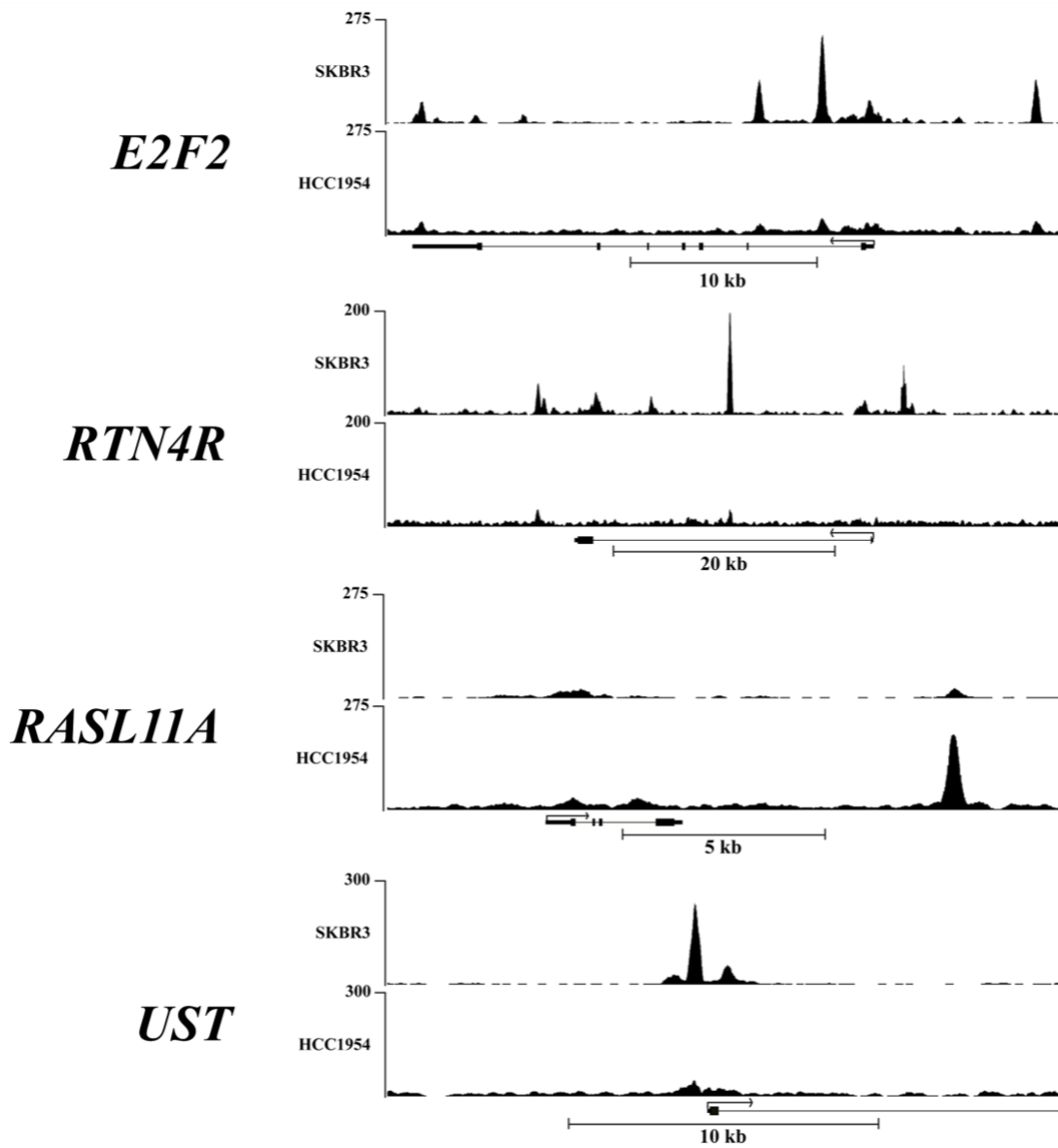


Figure S3. ChIP-seq for Differentially Regulated Genes Demonstrating Differential AP-2 γ Occupancy. Shown are examples of ChIP-seq data for four genes demonstrating differences in AP-2 γ occupancy comparing SKBR3 and HCC1954.

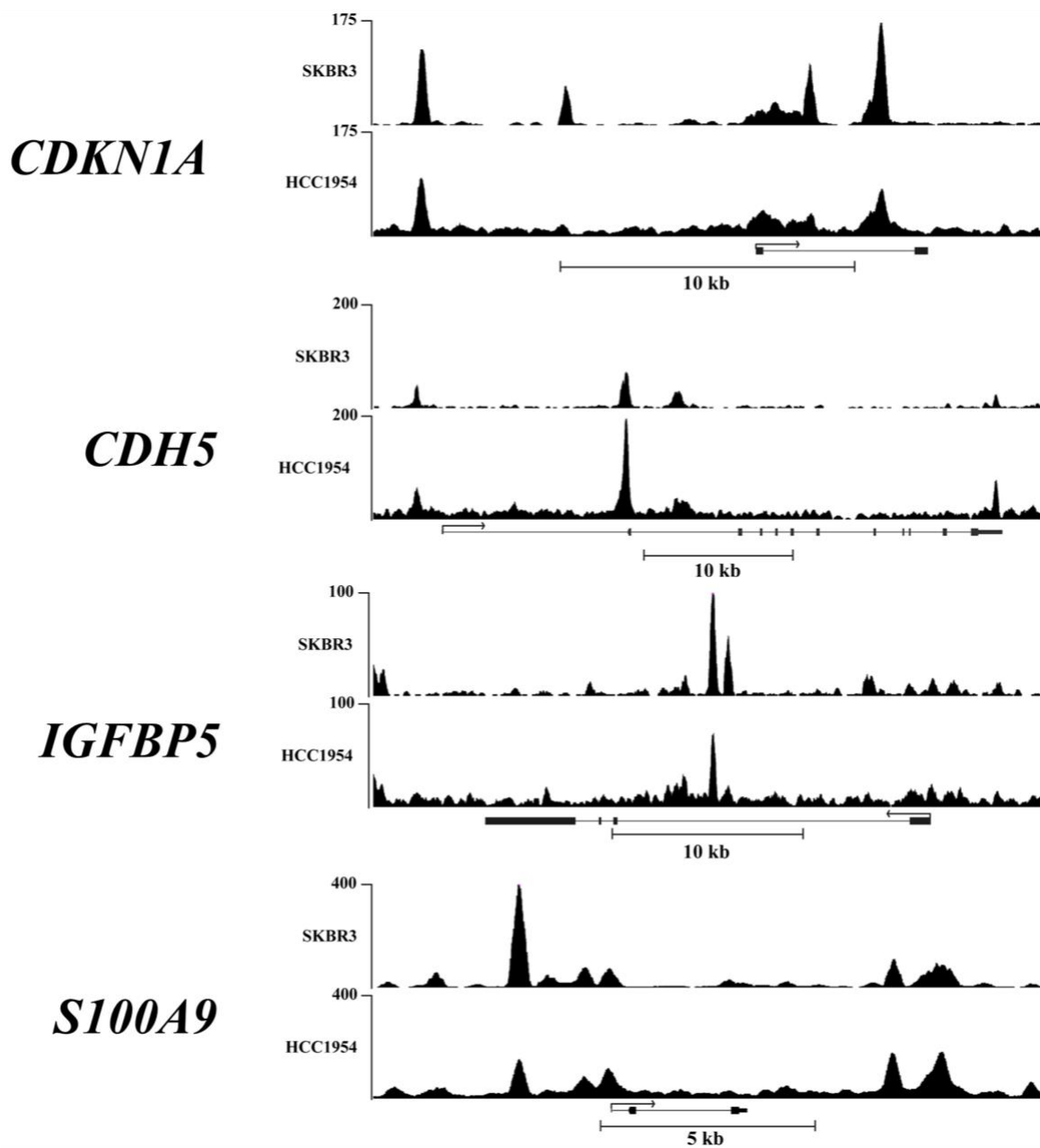


Figure S4. ChIP-seq for Differentially Regulated Genes Demonstrating Similar AP-2 γ Occupancy. Shown are examples of ChIP-seq data for four genes demonstrating similar patterns of AP-2 γ occupancy.

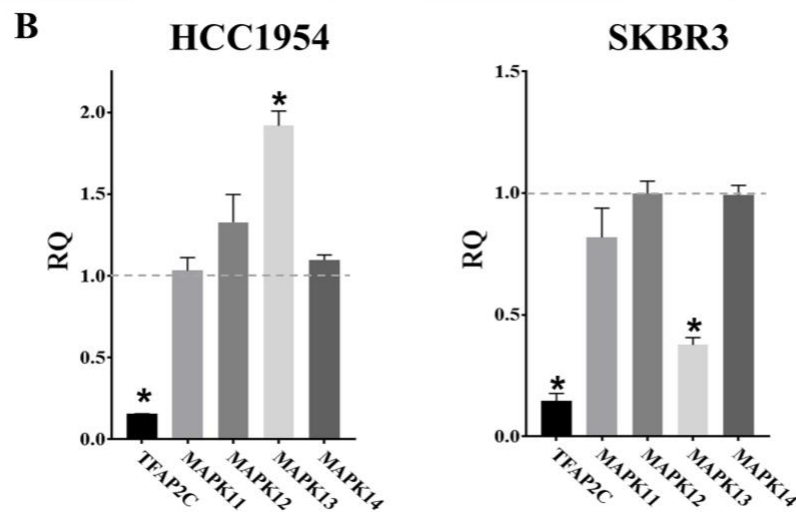
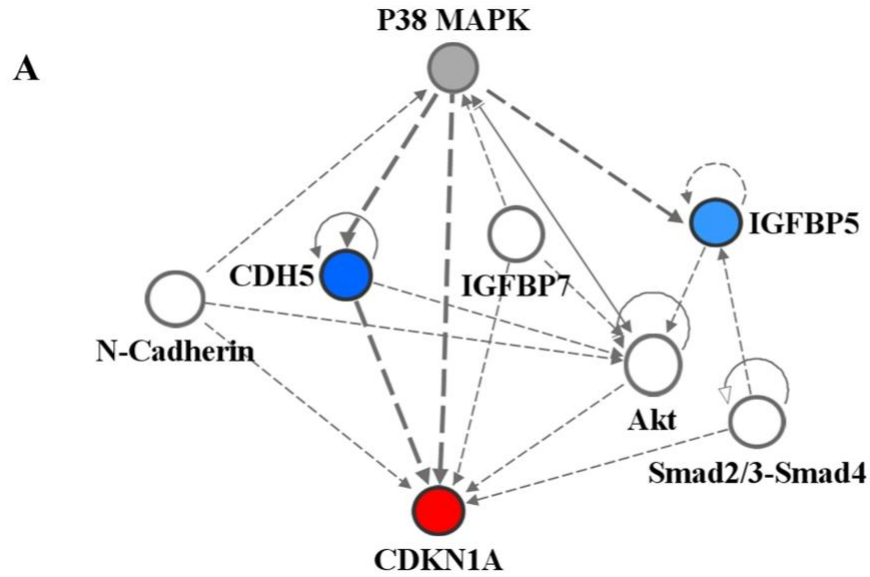


Figure S5. Pathway Analysis Implicates MAPK13/p38.

A. Pathway analysis shows the potential role of MAPK in regulation of differentially regulated genes in HER2 cancer lines. Relative color corresponds to change in relative expression and matches data in Figure 3. Data were analyzed through the use of Ingenuity Pathways Analysis. **B.** HCC1954 and SKBR3 cells were transfected with siRNA against *TFAP2C* vs. *NT* and relative RNA was measured using RT-PCR; data was normalized to transfection with *NT*; expression with *NT* was 1.0 (gray dotted line). Data demonstrate successful knockdown of *TFAP2C* and also significant changes in *MAPK13*; * $p < 0.05$.

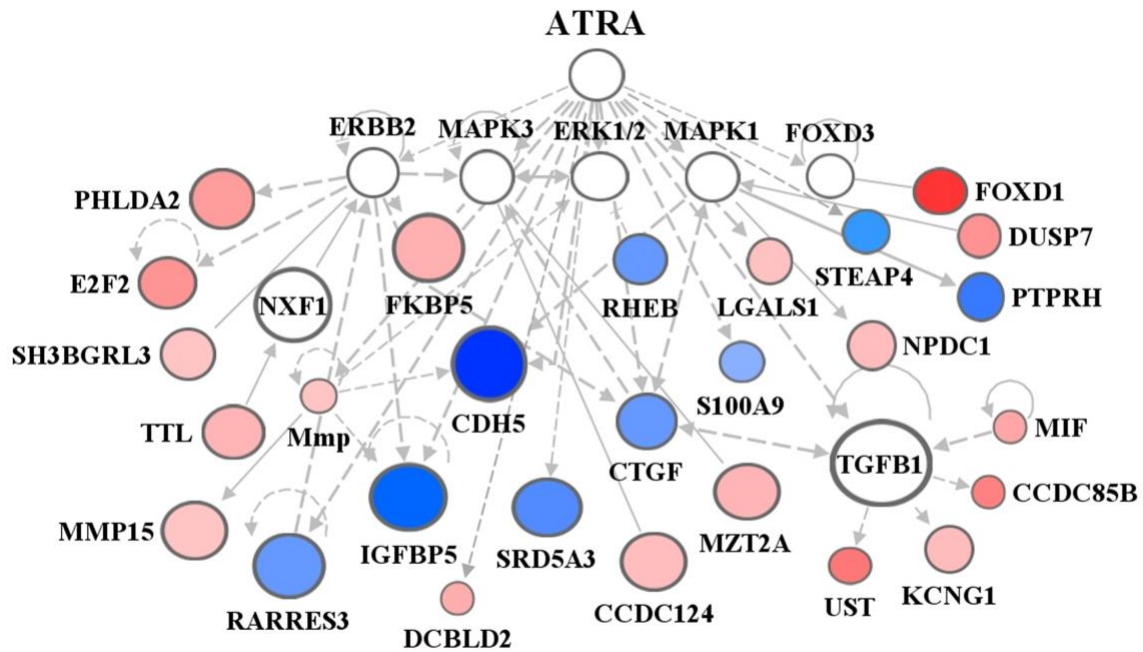


Figure S6. Pathway Analysis Implicates All-trans Retinoic Acid.

Pathway analysis shows the potential role of the retinoic acid pathway in regulating expression of differentially regulated gene signature. Relative color corresponds to change in relative expression and matches data in Figure 3. Data were analyzed through the use of Ingenuity Pathways Analysis.

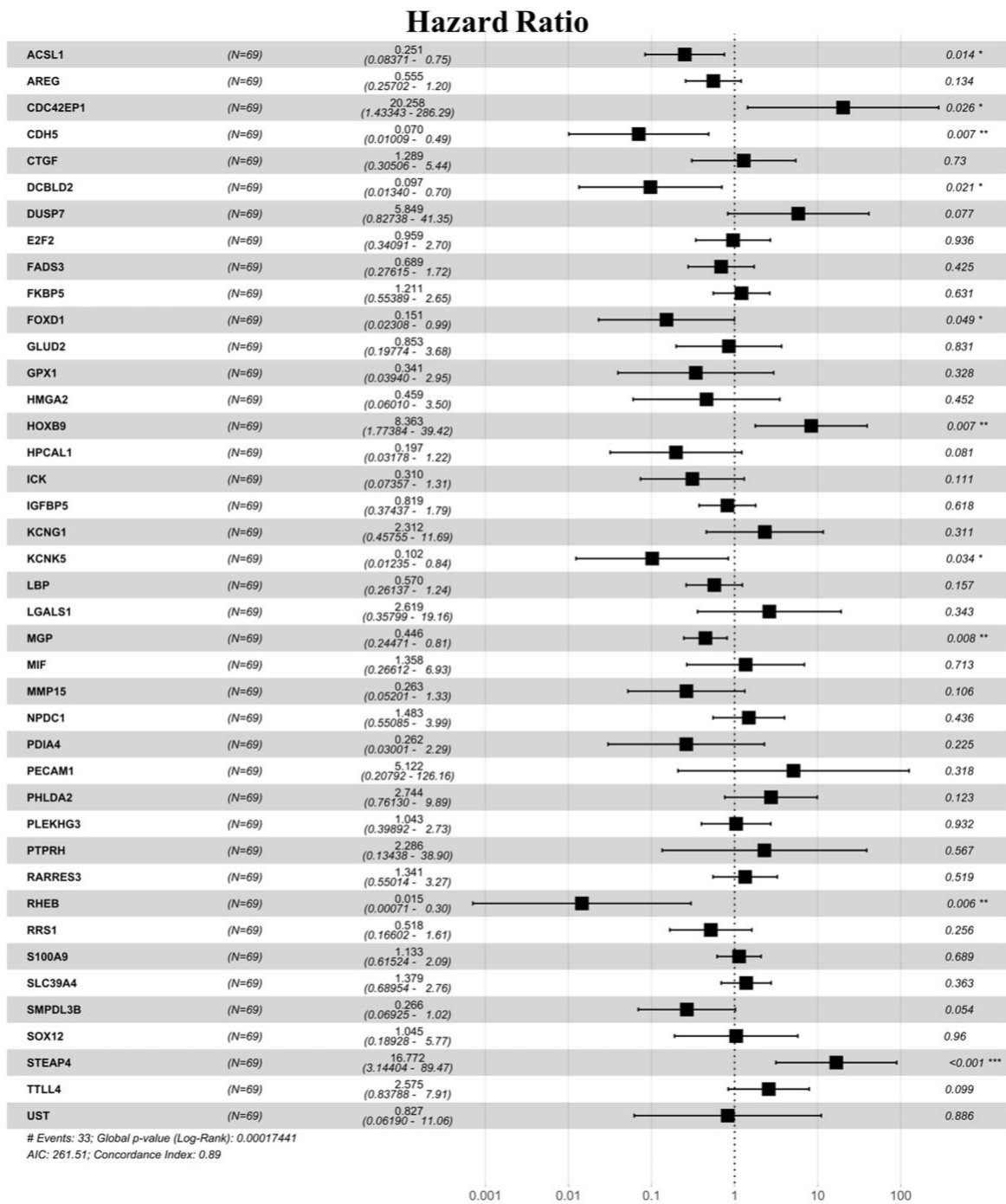


Figure S7. Hazard Ratio for Expression of Differentially Regulated Genes. Hazard Ratio is shown with confidence interval for 41 genes in Yau et al. (Yau et al., 2010) dataset. The p-values are shown in last column with significant differences noted by an asterisk.

Table S1: Calculation of Recurrence Score (41-gene panel).

Gene	coef	exp(coef)	se(coef)	z	Pr(> z)	
ACSL1	-1.38176	0.25114	0.56053	-2.465	0.013698	*
AREG	-0.58843	0.5552	0.39295	-1.497	0.134274	
CDC42EP1	3.00853	20.25762	1.35128	2.226	0.025985	*
CDH5	-2.65749	0.07012	0.98926	-2.686	0.007224	**
CTGF	0.25367	1.28875	0.73518	0.345	0.730059	
DCBLD2	-2.33267	0.09704	1.01021	-2.309	0.020939	*
DUSP7	1.76634	5.84939	0.99789	1.77	0.076715	.
E2F2	-0.04216	0.95872	0.52756	-0.08	0.936312	
FADS3	-0.37209	0.68929	0.46669	-0.797	0.42528	
FKBP5	0.19173	1.21134	0.39925	0.48	0.631078	
FOXD1	-1.8873	0.15148	0.96002	-1.966	0.04931	*
GLUD2	-0.15893	0.85306	0.74587	-0.213	0.831266	
GPX1	-1.07624	0.34087	1.10091	-0.978	0.328273	
HMGA2	-0.77915	0.4588	1.03707	-0.751	0.452473	
HOXB9	2.12377	8.36257	0.79115	2.684	0.007266	**
HPCAL1	-1.62459	0.19699	0.93088	-1.745	0.080945	.
ICK	-1.17035	0.31026	0.7343	-1.594	0.110974	
IGFBP5	-0.19944	0.81919	0.39954	-0.499	0.617659	
KCNQ1	0.83829	2.31242	0.82663	1.014	0.310528	
KCNK5	-2.28431	0.10184	1.07652	-2.122	0.033843	*
LBP	-0.56253	0.56977	0.39761	-1.415	0.157138	
LGALS1	0.96282	2.61908	1.01536	0.948	0.343001	
MGP	-0.80773	0.44587	0.30611	-2.639	0.008322	**
MIF	0.30578	1.35769	0.83144	0.368	0.713039	
MMP15	-1.33639	0.26279	0.82652	-1.617	0.105902	
NPDC1	0.39397	1.48285	0.50524	0.78	0.435533	
PDIA4	-1.33988	0.26188	1.10537	-1.212	0.225453	
PECAM1	1.63346	5.12156	1.63476	0.999	0.317696	
PHLDA2	1.00954	2.74434	0.65423	1.543	0.122807	
PLEKHG3	0.04211	1.04301	0.49037	0.086	0.931567	
PTPRH	0.82698	2.28641	1.44599	0.572	0.56738	
RARRES3	0.29317	1.34067	0.45447	0.645	0.518881	
RHEB	-4.23282	0.01451	1.5428	-2.744	0.006077	**
RRS1	-0.65848	0.51764	0.58019	-1.135	0.256401	
S100A9	0.12475	1.13286	0.31148	0.4	0.688791	
SLC39A4	0.32158	1.37931	0.35374	0.909	0.363294	
SMPDL3B	-1.32281	0.26639	0.68734	-1.925	0.054286	.
SOX12	0.04393	1.04491	0.87168	0.05	0.959805	
STEAP4	2.81971	16.77191	0.8542	3.301	0.000963	***
TTLL4	0.94583	2.57494	0.57282	1.651	0.098702	.
UST	-0.1894	0.82746	1.32286	-0.143	0.886155	

coef - the beta coefficient

exp(coef) - the hazard ratio

se(coef) - the standard error for beta coefficient

z - the wald statistic, calculated by dividing beta with its standard error

Pr - Pvalue corresponding to z statistics. If lower than 5% then null hypothesis of beta=0 can be rejected for 95% confidence level

*p<0.05, **p<0.01, ***p<0.001

$$RS = \sum_{k=1}^n E_k C_k$$

RS = recurrence score, where n is the number of genes used, E_k is the observed expression value for gene k , and C_k is the Cox-regression coefficient for gene k

Table S2: Calculation of Recurrence Score (16-gene panel).

Gene	coef	exp(coef)	se(coef)	z	Pr(> z)	
ACSL1	-0.98495	0.37346	0.73434	-1.341	0.17983	
CDC42EP1	1.12295	3.07392	0.90493	1.241	0.21463	
DCBLD2	-0.23381	0.79152	0.60608	-0.386	0.69967	
DUSP7	0.17424	1.19034	0.33437	0.521	0.60231	
FADS3	0.43487	1.54476	0.55792	0.779	0.43572	
GLUD2	-0.21373	0.80756	0.46138	-0.463	0.64319	
ICK	-2.12797	0.11908	0.76997	-2.764	0.00571	**
KCNQ1	-1.9555	0.14149	0.66812	-2.927	0.00342	**
MMP15	0.49272	1.63676	0.34161	1.442	0.1492	
NPDC1	0.19501	1.21533	0.30042	0.649	0.51625	
PHLDA2	-0.585	0.55711	0.47832	-1.223	0.22132	
RARRES3	-0.45798	0.63256	0.32929	-1.391	0.16428	
S100A9	0.03405	1.03464	0.09974	0.341	0.73279	
SLC39A4	-0.77984	0.45848	0.4964	-1.571	0.11619	
SMPDL3B	-0.31273	0.73144	0.29331	-1.066	0.28633	
STEAP4	0.82779	2.28826	0.36423	2.273	0.02304	*

coef - the beta coefficient

exp(coef) - the hazard ratio

se(coef) - the standard error for beta coefficient

z - the wald statistic, calculated by dividing beta with its standard error

Pr - Pvalue corresponding to z statistics. If lower than 5% then null hypothesis of beta=0 can be rejected for 95% confidence level

*p<0.05, **p<0.01