## SUPPLEMENTAL INFORMATION

## A TFAP2C Gene Signature is Predictive of Outcome in HER2positive Breast Cancer

Vincent T. Wu<sup>1, 8</sup>, Boris Kiriazov<sup>1, 8</sup>, Kelsey E. Koch<sup>1</sup>, Vivian W. Gu<sup>2</sup>, Anna C. Beck<sup>1</sup>, Nicholas Borcherding<sup>3</sup>, Tiandao Li<sup>1</sup>, Peter Addo<sup>1</sup>, Zachary J. Wehrspan<sup>4</sup>, Weizhou Zhang<sup>5</sup>, Terry A. Braun<sup>6</sup>, Bartley J. Brown<sup>6</sup>, Vimla Band<sup>7</sup>, Hamid Band<sup>7</sup>, Mikhail V. Kulak<sup>1</sup>, Ronald J. Weigel<sup>1, 2, 4 \*</sup>

<sup>1</sup>Department of Surgery, University of Iowa, Iowa City, IA, 52242 USA
<sup>2</sup>Department of Molecular Physiology and Biophysics, University of Iowa, Iowa City, IA, 52242 USA
<sup>3</sup>Department of Pathology, University of Iowa, Iowa City, IA, 52242 USA
<sup>4</sup>Department of Biochemistry, University of Iowa, Iowa City, IA, 52242 USA
<sup>5</sup>Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, FL 32610 USA
<sup>6</sup>Department of Biomedical Engineering, University of Iowa, Iowa City, IA, 52242 USA
<sup>7</sup>Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, NE, 68198 USA

<sup>8</sup>These authors contributed equally

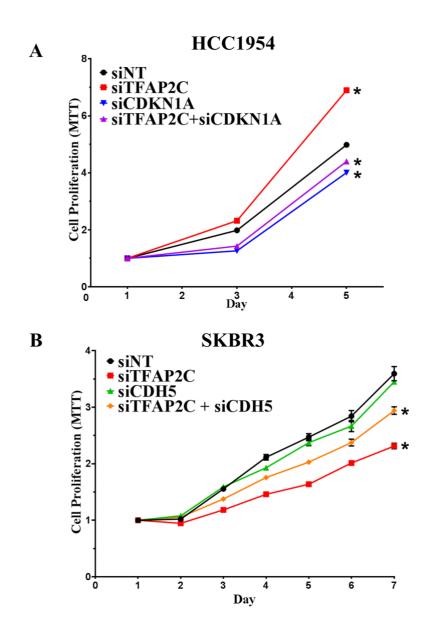
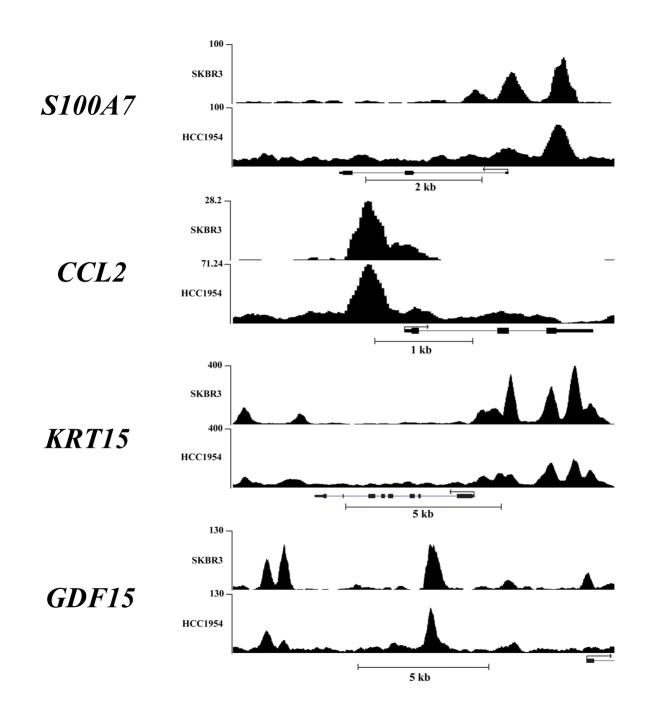


Figure S1. Proliferative Response to CDKN1A and CDH5 with Knockdown of

**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-2γ occupancy for SKBR3 and HCC1954 cells.

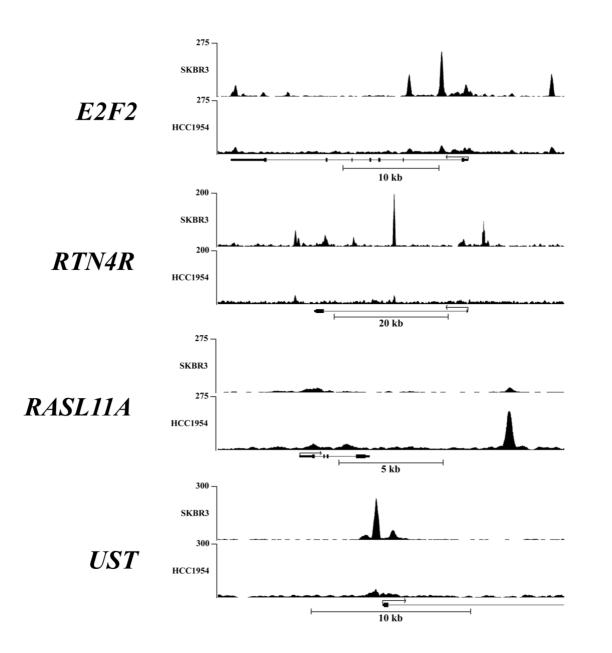


Figure S3. ChIP-seq for Differentially Regulated Genes Demonstrating Differential AP-2 $\gamma$  Occupancy. Shown are examples of ChIP-seq data for four genes demonstrating differences in AP-2 $\gamma$  occupancy comparing SKBR3 and HCC1954.

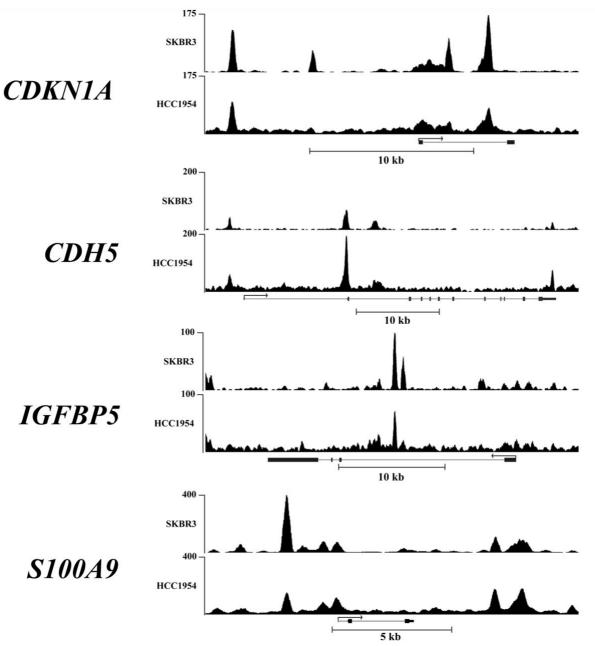


Figure S4. ChIP-seq for Differentially Regulated Genes Demonstrating Similar AP-

 $2\gamma$  Occupancy. Shown are examples of ChIP-seq data for four genes demonstrating similar patters of AP- $2\gamma$  occupancy.

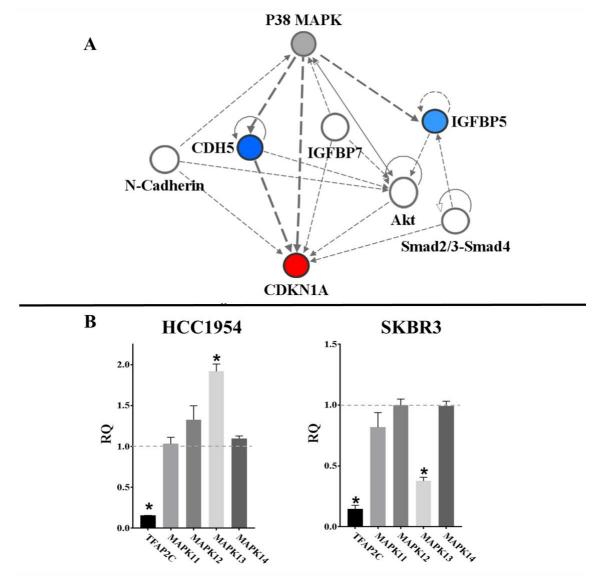


Figure S5. Pathway Analysis Implicates MAPK13/p38.

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.

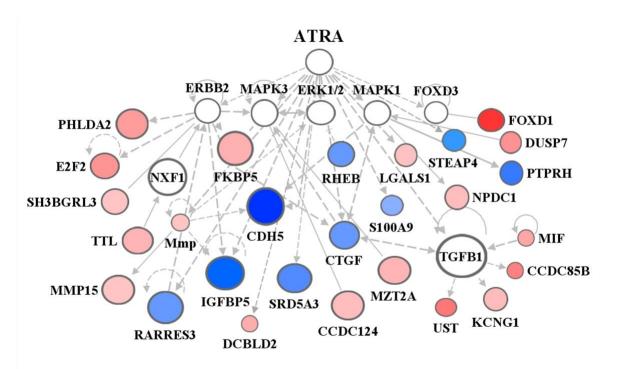


Figure S6. Pathway Analysis Implicates All-trans Retinoic Acid.

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.

		Hazard Ratio	· · · ·
ACSL1	(N=69)	0.251 (0.08371 - 0.75)	0.014 *
AREG	(N=69)	0.555 (0.25702 - 1.20)	0.134
CDC42EP1	(N=69)	20.258 (1.43343 - 286.29)	• 0.026 *
CDH5	(N=69)	0.070 (0.01009 - 0.49)	0.007 **
CTGF	(N=69)	(0.30506 - 5.44)	0.73
DCBLD2	(N=69)	0.097 (0.01340 - 0.70)	- 0.021 *
DUSP7	(N=69)	5.849 (0.82738 - 41.35)	0.077
E2F2	(N=69)	0.959 (0.34091 - 2.70)	0.936
FADS3	(N=69)	0.689 (0.27615 - 1.72)	0.425
FKBP5	(N=69)	1.211 (0.55389 - 2.65)	0.631
FOXD1	(N=69)	0.151 (0.02308 - 0.99)	0.049 *
GLUD2	(N=69)	0.853 (0.19774 - 3.68)	0.831
GPX1	(N=69)	0.341 (0.03940 - 2.95)	0.328
HMGA2	(N=69)	0.459 (0.06010 - 3.50)	0.452
HOXB9	(N=69)	8.363 (1.77384 - 39.42)	0.007 **
HPCAL1	(N=69)	0.197 (0.03178 - 1.22)	0.081
іск	(N=69)	0.310 (0.07357 - 1.31)	0.111
IGFBP5	(N=69)	0.819 (0.37437 - 1.79)	0.618
KCNG1	(N=69)	2.312 (0.45755 - 11.69)	0.311
KCNK5	(N=69)	0.102 (0.01235 - 0.84)	0.034 *
LBP	(N=69)	0.570 (0.26137 - 1.24)	0.157
LGALS1	(N=69)	2.619 (0.35799 - 19.16)	0.343
MGP	(N=69)	0.446 (0.24471 - 0.81)	0.008 **
MIF	(N=69)	1.358 (0.26612 - 6.93)	0.713
MMP15	(N=69)	0.263 (0.05201 - 1.33)	0.106
NPDC1	(N=69)	1.483 (0.55085 - 3.99)	0.436
PDIA4	(N=69)	0.262 (0.03001 - 2.29)	0.225
PECAM1	(N=69)	5.122 (0.20792 - 126.16)	0.318
PHLDA2	(N=69)	(0.76130 - 9.89)	0.123
PLEKHG3	(N=69)	1.043 (0.39692 - 2.73)	0.932
PTPRH	(N=69)	2.286 (0.13438 - 38.90)	0.567
RARRES3	(N=69)	(0.1340 - 30.27) 1.341 (0.55014 - 3.27)	0.519
RHEB	(N=69)	0.015 (0.00071 - 0.30)	0.006 **
RRS1	(N=69)	0.518 (0.16602 - 1.61)	0.256
S100A9	(N=69)	(0.7602 - 7.07) 1.133 (0.61524 - 2.09)	0.689
SLC39A4	(N=69)	(0.61524 - 2.09) 1.379 (0.68554 - 2.76)	0.363
SMPDL3B	(N=69)	(0.68954 - 2.76) 0.266 (0.06925 - 1.02)	0.054
SOX12	(N=69)	(0.09925 - 1.02) 1.045 (0.18928 - 5.77)	0.96
STEAP4	(N=69)	(0.18928 - 5.77) 16.772 (3.14404 - 89.47)	<0.001
TTLL4	(N=69)	(3.14404 - 89.47) 2.575 (0.83788 - 7.91)	0.099
UST	(N=69)	(0.83788 - 7.91) 0.827 (0.06190 - 11.06)	0.886
	(Log-Rank): 0.00017441	(0.06190 - 11.06)	

**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.

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	
KCNG1	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	

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

coef - the beta coefficient

exp(coef) - the harzard ratio

se(coef) - the standard error for beta coefficient

z - the wald statistict, 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* 

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	**
KCNG1	-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	*

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

coef - the beta coefficient

exp(coef) - the harzard ratio

se(coef) - the standard error for beta coefficient

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

 $\ensuremath{\mathsf{Pr}}$  -  $\ensuremath{\mathsf{Pvalue}}$  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