Supporting Information

An Estrogen Regulated Feedback Loop Limits the Efficacy of Estrogen Receptor Targeted Breast Cancer Therapy

www.pnas.org/cgi/doi/10.1073/pnas.1722617115

Supplementary Figure

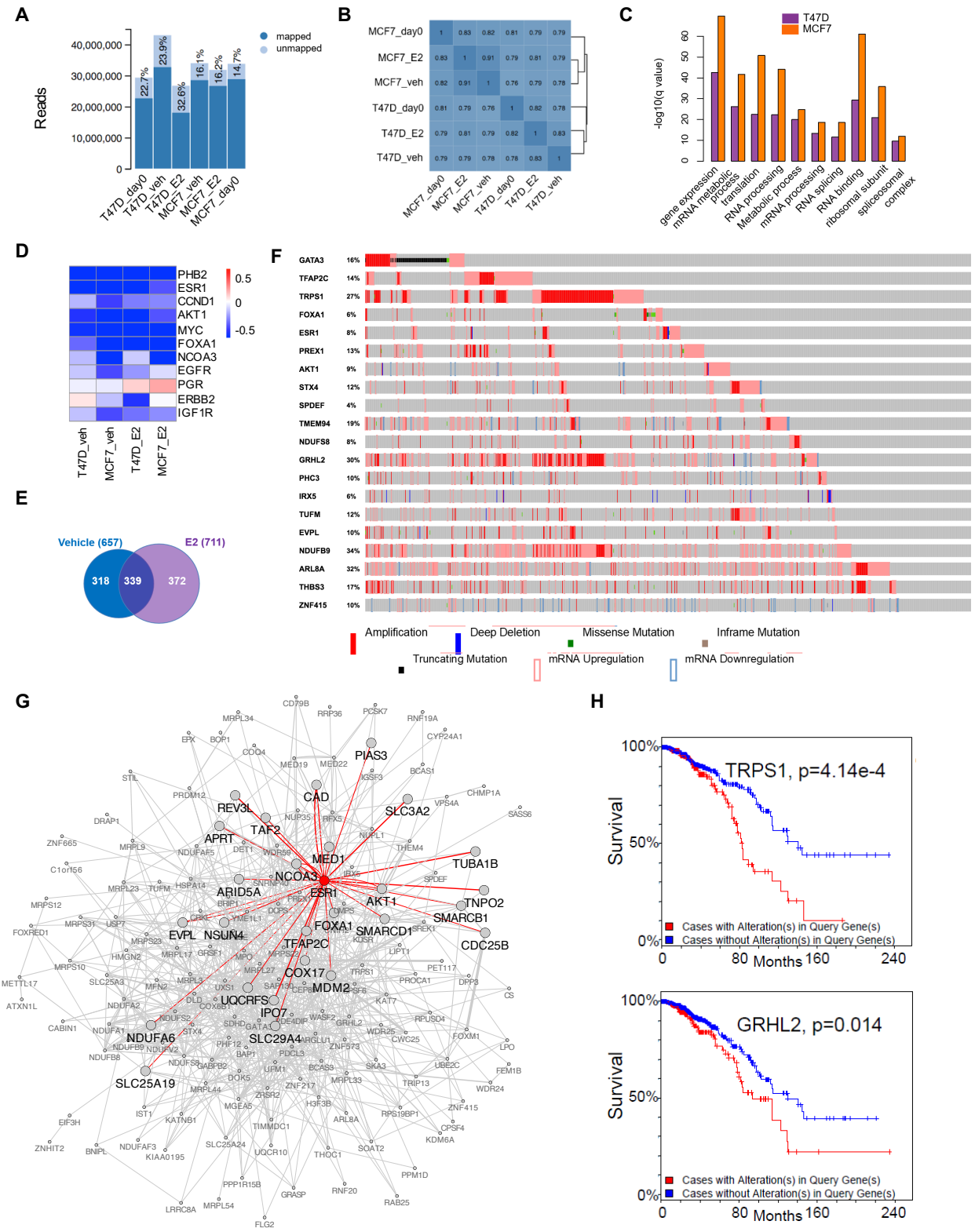

Fig. S1. CRISPR functional screens on two breast cancer cell lines, T47D and MCF7
(A-B) The quality control measurements of T47D and MCF7 CRISPR screens, including total reads and the percentage of unmapped reads (A), and sample correlation and clustering results (B).
(C) Enriched Gene Ontology (GO) terms in negatively selected genes. The functional enrichment is analyzed using GOrilla, the Gene Ontology enrichment tool.
(D) The overlap of top essential genes in vehicle and E2 (10nM) conditions. The top essential genes are genes that are consistently ranked top 1k in each condition in both T47D and MCF7 cells. There is a significant overlap between two sets of genes ($p<2.2e-16$, Fisher's exact test).
(E) The β scores of genes in ER downstream pathways.
(F**)** Genetic alterations of top 20 breast cancer specific essential genes in TCGA breast cancer dataset. Data is downloaded and visualized from cBioPortal (www.cbioportal.org).
(G) A network view of breast cancer specific essential genes. Dots represent essential genes, and edges indicate two genes have genetic, physical interactions, are co-localized, or are in the same pathway. The network is extracted from GeneMANIA (www.genemania.org). Genes connected with ER (large nodes) are highlighted.
(H) Alterations of TRPS1 and GRHL2 predicts worse clinical outcome.

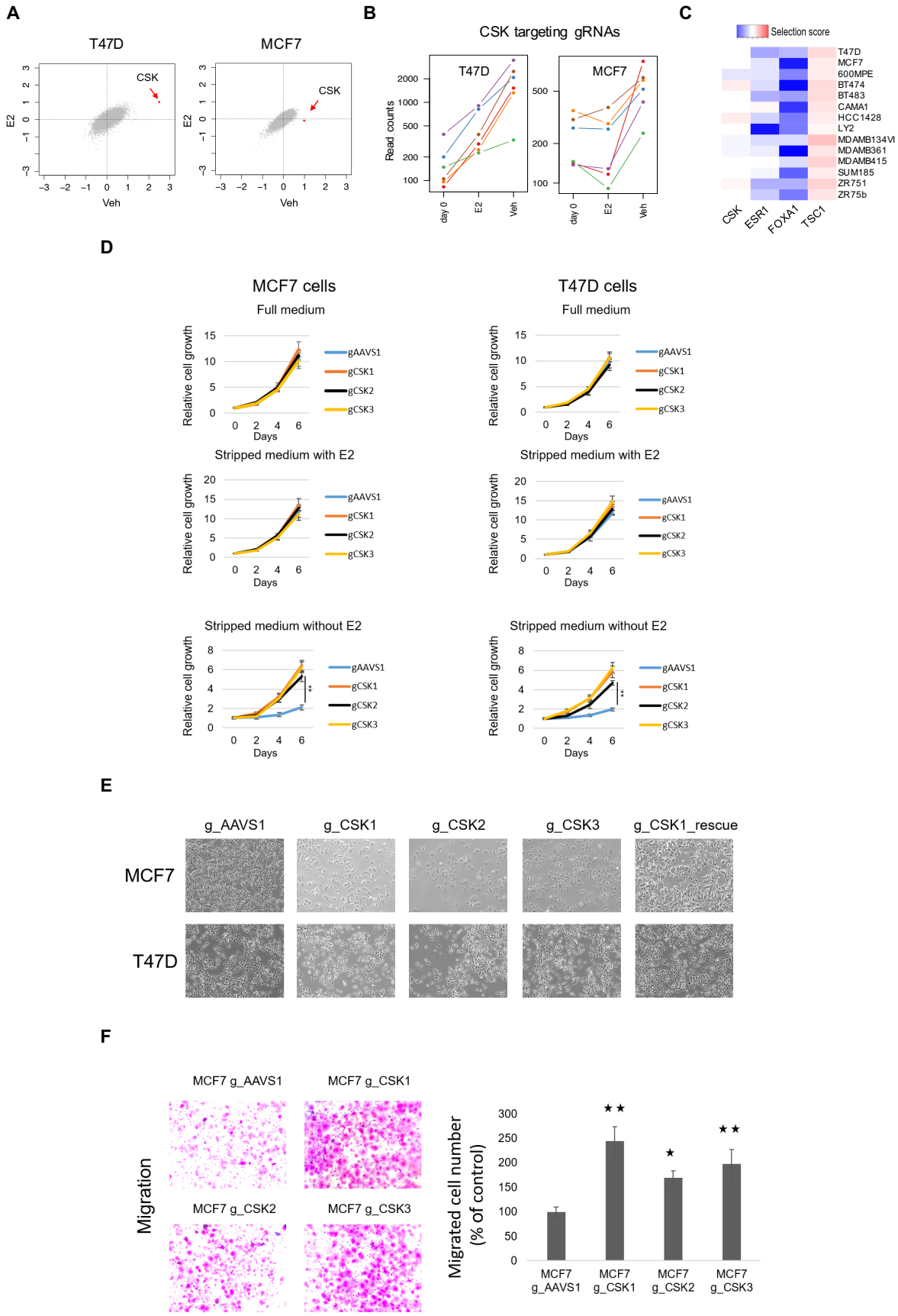
A
T47D
MCF7
CSK
E2
Veh
B
CSK targeting gRNAs
T47D
MCF7
Read counts
day 0
E2
Veh
C
Selection score
T47D
MCF7
600MPE
BT474
BT483
CAMA1
HCC1428
LY2
MDAMB134VI
MDAMB361
MDAMB415
SUM185
ZR751
ZR75b
CSK
ESR1
FOXA1
TSC1
D
MCF7 cells
T47D cells
Full medium
Stripped medium with E2
Stripped medium without E2
Relative cell growth
Days
gAAVS1
gCSK1
gCSK2
gCSK3
E
g_AAVS1
g_CSK1
g_CSK2
g_CSK3
g_CSK1_rescue
MCF7
T47D
F
Migration
MCF7 g_AAVS1
MCF7 g_CSK1
MCF7 g_CSK2
MCF7 g_CSK3
Migrated cell number (% of control)
MCF7 g_AAVS1
MCF7 g_CSK1
MCF7 g_CSK2
MCF7 g_CSK3

Fig. S2. CSK mediates hormone independent breast cancer cell growth
(A) CSK shows positively selected in vehicle treated conditions compared with E2 (10nM) treated conditions in both T47D and MCF7 cell lines. The β scores of all genes in two conditions (vehicle and E2) are shown.
(B) The normalized read counts of gRNAs targeting CSK in two cell lines.
(C) The selection score of CSK, as well as ESR1, FOXA1 and TSC1 in a public RNAi screen on several ER+ breast cancer cell lines (Marcotte et al. Cell 2016).
(D) Cell proliferation of MCF7 (left) and T47D (right) cells harboring three different gRNA against CSK, and one gAAVS1 as control. The cells were cultured in full medium with 10%FBS, or hormone depleted medium (10% DCC-FBS) plus/minus E2 (10nM). The data were represented as mean ± SD (for n = 3, *P < 0.05, **P < 0.005, two-tailed student's t-test).
(E) The morphology change of cell shapes after knocking out CSK in T47D and MCF7 cells. Rescuing CSK expression recovers the original cell shapes in both cell lines.
(F) Migration of MCF7 cells upon CSK loss. MCF7 cells were infected with lentiviral gRNAs directed against either AAVS1 (control) or CSK plated onto transwell chambers for 24 h in DMEM medium. After 24 h, the migrated cells were fixed, stained and quantified. Results shown are normalized to control cell migration and represent the means ± SD of data from 3 independent experiments. *P<0.05; **P<0.005; two-tailed student's t-test.

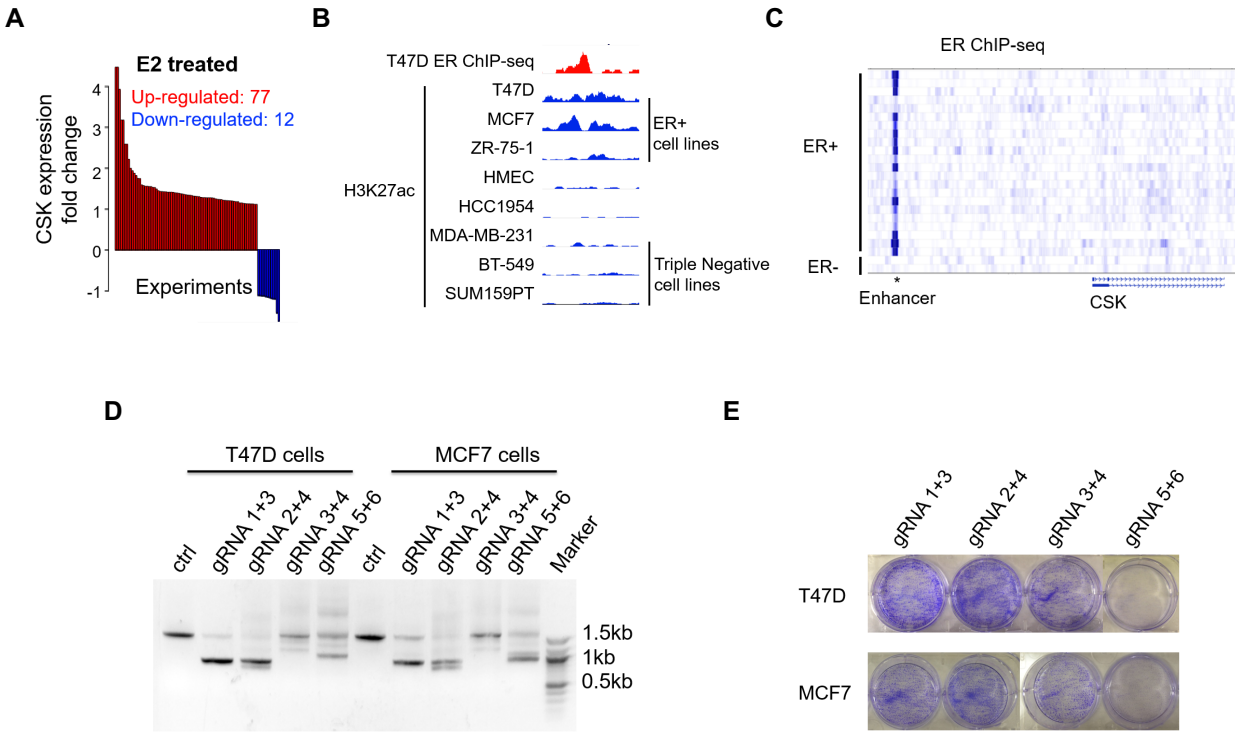

Fig. S3. ER binds to CSK enhancer and regulate CSK expression

(A) The CSK expression change in E2 treated samples in cell lines or mouse models from Nuclear Receptor Signaling Atlas (NURSA) database (https://www.nursa.org). Only samples with CSK expression fold change >1.1 or <-1.1 are considered.

(B) The H3K27ac signals of the enhancer in different breast cancer cell lines. Data is extracted from Cistrome database (http://cistrome.org/db).

(C) ER binds to the enhancer of CSK in 86% (19/22) ER+ breast cancer patients in a public ER ChIP-seq dataset (Ross-Inns et al. Nature 2012).

(D) The knockout efficiency of CSK enhancer deletions using different pairs of gRNAs.

(E) The effects of deleting enhancers and flanking regions on cell growth. Cell growth measured by crystal violet staining assays is shown. All of the cells were cultured in hormone-depleted medium.

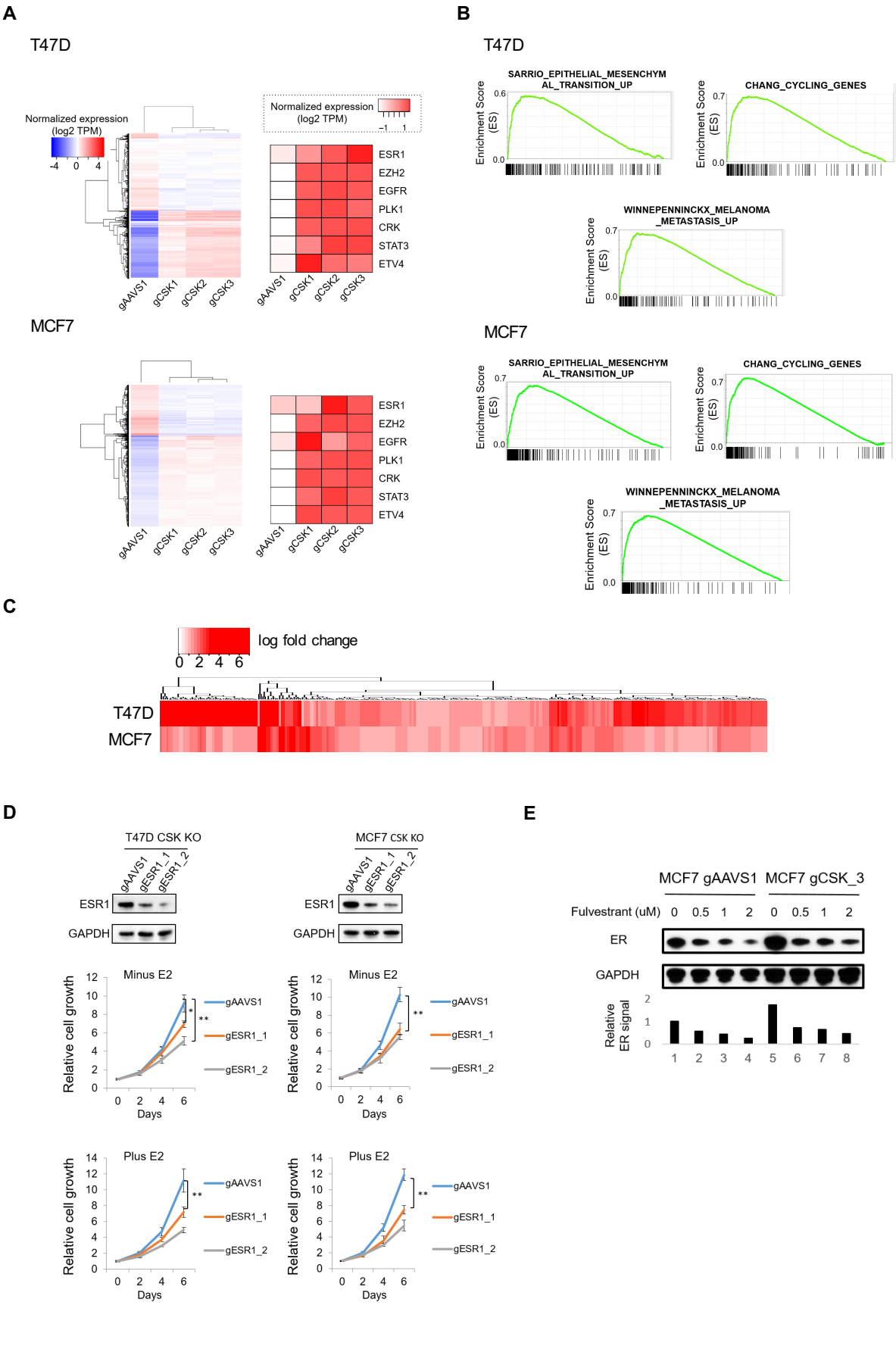
A
T47D
Normalized expression (log2 TPM)
-4 0 4
Normalized expression (log2 TPM)
-1 1
ESR1
EZH2
EGFR
PLK1
CRK
STAT3
ETV4
gAAVS1
gCSK1
gCSK2
gCSK3
MCF7
B
T47D
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP
CHANG_CYCLING_GENES
WINNEPENNINCKX_MELANOMA_METASTASIS_UP
Enrichment Score (ES)
MCF7
C
log fold change
0 2 4 6
T47D
MCF7
D
T47D CSK KO
MCF7 CSK KO
gAAVS1
gESR1_1
gESR1_2
ESR1
GAPDH
Minus E2
Plus E2
Relative cell growth
Days
E
MCF7 gAAVS1
MCF7 gCSK_3
Fulvestrant (uM) 0 0.5 1 2 0 0.5 1 2
ER
GAPDH
Relative ER signal

Fig. S4. Growth factor and ER signaling changes induced by CSK loss
(A) The expression patterns of differentially expressed genes (FDR=1e-5) between CSK null and wild-type cells (left), and the normalized expression of selected genes in control and CSK null cells. The expressions of genes are measured in Transcripts Per Million (TPM) from RNA-seq. All the cells were cultured in hormone depleted medium for 48 hr before RNA isolation.
(B) Enriched pathways in up- and down-regulated genes in CSK null cells using Gene Set Enrichment Analysis (GSEA) in T47D and MCF7 cells. A complete list of GSEA results can be found in Table S7.
(C) The log-fold change of CSK-signature genes (or up-regulated genes upon CSK loss) in both cell lines.
(D) Growth curve of T47D and MCF7 CSK null cells. Cell proliferation of T47D (left) and MCF7 (right) CSK nulls cells (gCSK3) harboring two independent gRNAs against ESR1, and one gAAVS1 as control. All the cells were cultured in hormone depleted medium plus 10% DCC-FBS with or without E2 (10nM). The data were represented as mean ± SD (for n = 3, *P < 0.05, **P < 0.005, two-tailed student's t-test).
(E) The expressions of ER in the MCF7 control (gAAVS1) and MCF7 CSK null cells (gCSK3) upon treatments of fulvestrant at 0, 0.5, 1, or 2 uM for 24hr. Immunoblot analysis for indicated proteins of control (gAAVS1) and CSK null cells is shown. GAPDH was used as a loading control.

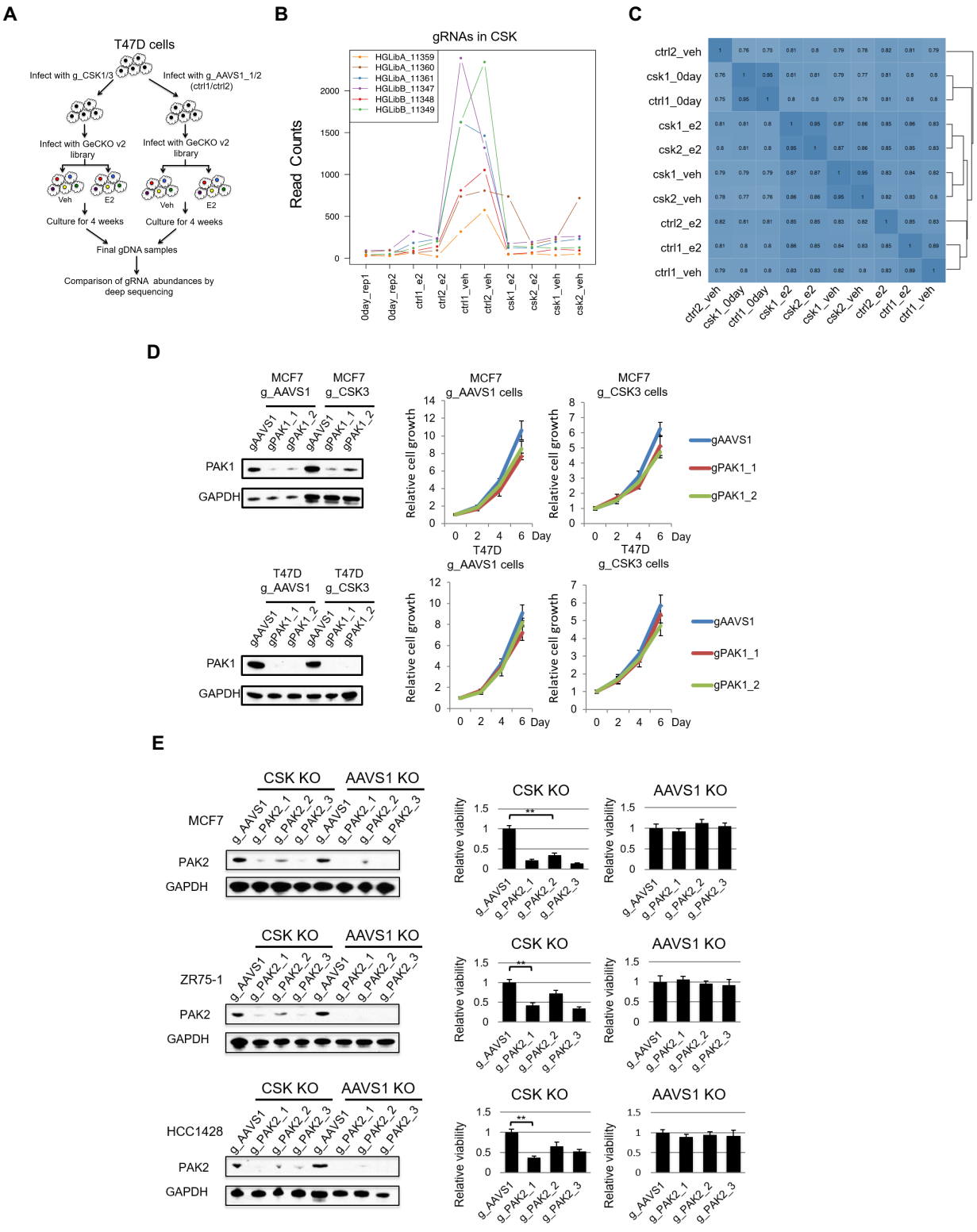
A
T47D cells
Infect with g_CSK1/3
Infect with g_AAVS1_1/2 (ctrl1/ctrl2)
Infect with GeCKO v2 library
Infect with GeCKO v2 library
Veh
E2
Veh
E2
Culture for 4 weeks
Culture for 4 weeks
Final gDNA samples
Comparison of gRNA abundances by deep sequencing
B
gRNAs in CSK
HGLibA_11359
HGLibA_11360
HGLibA_11361
HGLibB_11347
HGLibB_11348
HGLibB_11349
Read Counts
0day_rep1
0day_rep2
ctrl1_e2
ctrl2_e2
ctrl1_veh
ctrl2_veh
csk1_e2
csk2_e2
csk1_veh
csk2_veh
C
ctrl2_veh
csk1_0day
ctrl1_0day
csk1_e2
csk2_e2
csk1_veh
csk2_veh
ctrl2_e2
ctrl1_e2
ctrl1_veh
D
MCF7 g_AAVS1
MCF7 g_CSK3
gAAVS1
gPAK1_1
gPAK1_2
PAK1
GAPDH
MCF7 g_AAVS1 cells
MCF7 g_CSK3 cells
Relative cell growth
Day
gAAVS1
gPAK1_1
gPAK1_2
T47D g_AAVS1
T47D g_CSK3
T47D g_AAVS1 cells
T47D g_CSK3 cells
E
CSK KO
AAVS1 KO
MCF7
ZR75-1
HCC1428
g_AAVS1
g_PAK2_1
g_PAK2_2
g_PAK2_3
PAK2
GAPDH
Relative viability

Fig. S5. PAK2 is synthetic lethal to CSK loss

(A) The screening strategy. For the second round of genome-wide CRISPR screens, we first infected the T47D cells with lentiviral gCSK_1, gCSK_3, gAAVS1_1, gAAVS_2 cloned by pLX-gRNA vector. After blasticidin selection, we generated these four types of T47D cells with stable expression of gCSK_1, gCSK_3, gAAVS1_1, gAAVS_2 respectively. Then the Genome-wide CRISPR screens were performed in these four cell types.

(B) The normalized counts of CSK-targeting gRNAs in each screen. The six CSK-targeting gRNA IDs in the GeCKO2 library are shown in the legend.

(C) The sample correlation and clustering results of all samples involved.

(D) The relative viability of MCF7&T47D CSK null cells (gCSK3) and control (gAAVS1) cells harboring two gRNAs against PAK1, and one gAAVS1 as control cells. The control cells were cultured in hormone-depleted medium (10% DCC-FBS) plus E2 (10nM) and the CSK null cells were cultured in hormone-depleted medium (10% DCC-FBS) plus vehicle. The data were represented as mean ± SD (for n = 3, *P < 0.05, **P < 0.005, two-tailed student's t-test). Immunoblot analysis for indicated proteins of control (gAAVS1) and CSK null cells is shown on the left panel. GAPDH was used as a loading control.

(E) Validation of a synthetic lethal interaction between PAK2 and CSK in MCF7, ZR75-1 and HCC1428 cells. Relative viability of MCF7, ZR75-1 and HCC1428 CSK null (gCSK3) as well as control (gAAVS1) cells harboring three gRNAs against PAK2, and one gAAVS1 as control (mean ± SD, for n = 3) is shown. The control cells were cultured in hormone-depleted medium (10% DCC-FBS) plus E2 (10nM) and the CSK null cells were cultured in hormone-depleted medium (10% DCC-FBS) plus vehicle. The immunoblot analysis indicated proteins of PAK2 upon control (AAVS1) and CSK–null cells is displayed on the left panel. GAPDH was used as a loading control.


A
Tamoxifen Fulvestrant
CSK KO
Days 0 2 4 6 2 4 6
p-Y419 SFK
p-S141 PAK2
CSK
PAK2
SFK
GAPDH
B
Relative gene expression
MCF7
MCF7-LTED
MCF7-TAMR
MCF7-FULR
T47D
T47D-LTED
T47D-TAMR
T47D-FULR
CSK
C
Median(36.4)
3xM
10xM
D
CSK-supp. signature gene score
Pearson correlation=-0.54
p=3.63e-5
CSK expression in log2(RPKM)
E
Lymphocytes
CSK expression
Tissues
CSK-supp. signature gene expression
F
Tamoxifen treated patients (GSE6532)
Tamoxifen treated patients (GSE9195)
Fraction
8 (62%)
4 (30%)
1 (8%)
19 (19%)
59 (58%)
24 (24%)
12 (63%)
5 (26%)
2 (11%)
12 (31%)
15 (38%)
12 (31%)
Grade 3
Grade 2
Grade 1
High CSK-supp. sig. exp(n=13)
Others (n=102)
High CSK-supp. sig. exp (n=19)
Others (n=39)

Fig. S6. Clinical relevance of an ER regulated feedback loop
(A) Activating phosphorylation of SFK and PAK2, and gene expression of CSK, SFK and PAK2 upon tamoxifen and fulvestrant treatment in MCF7 cells for indicated days. GAPDH was used as a loading control.
(B) The relative expressions of CSK in T47D and MCF7, as well as the long-term estrogen deprivation (LTED) cells and tamoxifen/fulvestrant-resistent (TAMR/FULR) cells. The relative gene expression was measured by qRT-PCR after normalizing to the amount of GAPDH signal (mean ± SD, for n = 3) ** $p<0.005$, student's t test.
(C) The expression of CSK across different tissue types in BioGPS (http://biogps.org).
(D-E) The correlation of CSK gene expression and CSK-suppressed signature gene expression in different tissues in GTEx Consortium (1).
(F) Patients with higher expression of CSK-suppressed signature genes ("High CSK-supp. sig. exp") correspond to higher grade tumors in tamoxifen-treated patients from public datasets (indicated by the cohort name or GSE accession number).

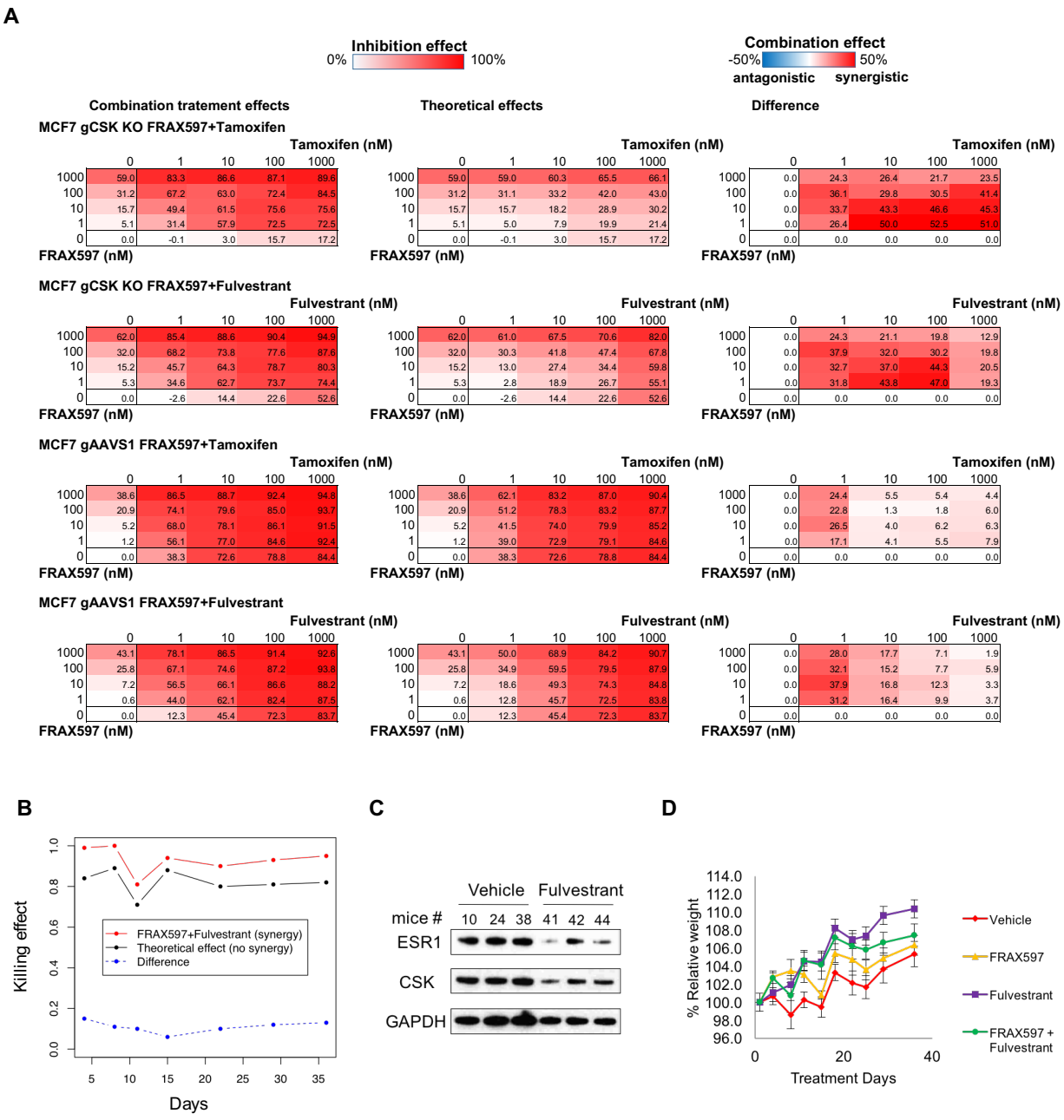
A
Inhibition effect
0%
100%
Combination effect
-50%
50%
antagonistic
synergistic
Combination tratement effects
Theoretical effects
Difference
MCF7 gCSK KO FRAX597+Tamoxifen
Tamoxifen (nM)
0 1 10 100 1000
1000 59.0 83.3 86.6 87.1 89.6
100 31.2 67.2 63.0 72.4 84.5
10 15.7 49.4 61.5 75.6 75.6
1 5.1 31.4 57.9 72.5 72.5
0 0.0 -0.1 3.0 15.7 17.2
FRAX597 (nM)
Tamoxifen (nM)
0 1 10 100 1000
1000 59.0 59.0 60.3 65.5 66.1
100 31.2 31.1 33.2 42.0 43.0
10 15.7 15.7 18.2 28.9 30.2
1 5.1 5.0 7.9 19.9 21.4
0 0.0 -0.1 3.0 15.7 17.2
FRAX597 (nM)
Tamoxifen (nM)
0 1 10 100 1000
1000 0.0 24.3 26.4 21.7 23.5
100 0.0 36.1 29.8 30.5 41.4
10 0.0 33.7 43.3 46.6 45.3
1 0.0 26.4 50.0 52.5 51.0
0 0.0 0.0 0.0 0.0 0.0
FRAX597 (nM)
MCF7 gCSK KO FRAX597+Fulvestrant
Fulvestrant (nM)
0 1 10 100 1000
1000 62.0 85.4 88.6 90.4 94.9
100 32.0 68.2 73.8 77.6 87.6
10 15.2 45.7 64.3 78.7 80.3
1 5.3 34.6 62.7 73.7 74.4
0 0.0 -2.6 14.4 22.6 52.6
FRAX597 (nM)
Fulvestrant (nM)
0 1 10 100 1000
1000 62.0 61.0 67.5 70.6 82.0
100 32.0 30.3 41.8 47.4 67.8
10 15.2 13.0 27.4 34.4 59.8
1 5.3 2.8 18.9 26.7 55.1
0 0.0 -2.6 14.4 22.6 52.6
FRAX597 (nM)
Fulvestrant (nM)
0 1 10 100 1000
1000 0.0 24.3 21.1 19.8 12.9
100 0.0 37.9 32.0 30.2 19.8
10 0.0 32.7 37.0 44.3 20.5
1 0.0 31.8 43.8 47.0 19.3
0 0.0 0.0 0.0 0.0 0.0
FRAX597 (nM)
MCF7 gAAVS1 FRAX597+Tamoxifen
Tamoxifen (nM)
0 1 10 100 1000
1000 38.6 86.5 88.7 92.4 94.8
100 20.9 74.1 79.6 85.0 93.7
10 5.2 68.0 78.1 86.1 91.5
1 1.2 56.1 77.0 84.6 92.4
0 0.0 38.3 72.6 78.8 84.4
FRAX597 (nM)
Tamoxifen (nM)
0 1 10 100 1000
1000 38.6 62.1 83.2 87.0 90.4
100 20.9 51.2 78.3 83.2 87.7
10 5.2 41.5 74.0 79.9 85.2
1 1.2 39.0 72.9 79.1 84.6
0 0.0 38.3 72.6 78.8 84.4
FRAX597 (nM)
Tamoxifen (nM)
0 1 10 100 1000
1000 0.0 24.4 5.5 5.4 4.4
100 0.0 22.8 1.3 1.8 6.0
10 0.0 26.5 4.0 6.2 6.3
1 0.0 17.1 4.1 5.5 7.9
0 0.0 0.0 0.0 0.0 0.0
FRAX597 (nM)
MCF7 gAAVS1 FRAX597+Fulvestrant
Fulvestrant (nM)
0 1 10 100 1000
1000 43.1 78.1 86.5 91.4 92.6
100 25.8 67.1 74.6 87.2 93.8
10 7.2 56.5 66.1 86.6 88.2
1 0.6 44.0 62.1 82.4 87.5
0 0.0 12.3 45.4 72.3 83.7
FRAX597 (nM)
Fulvestrant (nM)
0 1 10 100 1000
1000 43.1 50.0 68.9 84.2 90.7
100 25.8 34.9 59.5 79.5 87.9
10 7.2 18.6 49.3 74.3 84.8
1 0.6 12.8 45.7 72.5 83.8
0 0.0 12.3 45.4 72.3 83.7
FRAX597 (nM)
Fulvestrant (nM)
0 1 10 100 1000
1000 0.0 28.0 17.7 7.1 1.9
100 0.0 32.1 15.2 7.7 5.9
10 0.0 37.9 16.8 12.3 3.3
1 0.0 31.2 16.4 9.9 3.7
0 0.0 0.0 0.0 0.0 0.0
FRAX597 (nM)
B
Killing effect
0.0 0.2 0.4 0.6 0.8 1.0
FRAX597+Fulvestrant (synergy)
Theoretical effect (no synergy)
Difference
5 10 15 20 25 30 35
Days
C
Vehicle
Fulvestrant
mice # 10 24 38 41 42 44
ESR1
CSK
GAPDH
D
% Relative weight
114.0 112.0 110.0 108.0 106.0 104.0 102.0 100.0 98.0 96.0
0 20 40
Treatment Days
Vehicle
FRAX597
Fulvestrant
FRAX597 + Fulvestrant

Fig. S7. Potential therapeutic strategies to overcome endocrine resistance
(A) Combination of fulvestrant and tamoxifen with PAK2 inhibitor (FRAX597) for treatment of the MCF7 CSK null cells (gCSK3, top two rows) and the control cells (gAAVS1, bottom two rows). The cells were cultured in full medium treated with indicated drugs (three replicates) for five days. The combination treatment effects (left) are measured in terms of normalized percentage of inhibition, while the theoretical effects of drug combination (middle) are calculated using the BLISS model (Bliss et al. Annals of Applied Biology 2008). The differences between the two (right) indicate whether the effect is synergistic (positive value) or antagonistic (negative value).
(B) In PDX model, the killing effect of FRAX597 and fulvestrant (red) compared with the theoretical combination effect of both drugs using the BLISS model (black), assuming there is no synergy between them.
(C) Protein expression of ESR1 and CSK in the PDX tumors treated with vehicle or fulvestrant. Tumor lysates from three mice treated with vehicle or fulvestrant were analyzed by western blot using antibodies targeting the ESR1 or CSK. GAPDH was used as a loading control.
(D) Single or combination treatments of vehicle, FRAX597 and fulvestant in TM00386 PDX (Jackson Labs) tumors for 35 days in each respective group (n=8). Relative mouse weights were plotted as an average of % of the first measurement (% relative weight) for each mouse in each respective group. The data were represented as mean ± SD.

SUPPLEMENTAL TABLES

Table S1. The β scores of all the coding genes from E2 and vehicle conditions in T47D and MCF7 cells, calculated using MAGeCK-VISPR.

Table S2. The top essential genes in E2 and vehicle conditions. Top essential genes are genes that are consistently ranked top 1k in each condition in both T47D and MCF7 cells. The overlap of two genes is also marked.

Table S3. Breast cancer specific essential genes, calculated from the β scores of 2 breast cancer cell lines and 10 non-breast cancer cell lines.

Table S4. The breast cancer specific essential genes are enriched in breast cancer related gene sets. Breast cancer specific essential genes are compared with genes in MSigDB (http://software.broadinstitute.org/gsea/msigdb/), and the enrichment test is performed using the overlap test in the MSigDB website.

Table S5. Top genes that have stronger positive selection in vehicle over E2 conditions in T47D and MCF7 cells.

Table S6. Differentially expressed genes between CSK knockout cells (vehicle treated) and CSK wild-type cells (vehicle treated) in both T47D and MCF7 cells. DESeq2 was used to perform the differential expression analysis of RNA-seq data.

Table S7. The Gene Set Enrichment Analysis (GSEA) results of up-regulated genes in CSK null T47D cells compared with control cells.

Table S8. The β scores of all the coding genes from E2 and vehicle conditions in T47D CSK null cells, calculated using MAGeCK-VISPR.

Table S9. Top genes that are specifically essential in CSK null cells (vehicle treated) compared with other non-breast cancer cell lines.

Table S10. Top genes that are specifically essential in CSK null cells (vehicle treated) compared with CSK wild-type cells (E2 treated).

Supplementary Methods

Plasmids and inhibitors

The lentiviral gCSK, gAAVS1, gPAK-2 and gCSK_enhancer vectors were generated by ligation of hybridized oligos (below) into LentiCRISPR-v2 vector (Addgene) linearized with BsmBI using quick ligase (NEB).

CACCGTACAAAGCCAAAAACAAGG	gCSK_1_F
AAACCCTTGTTTTTGGCTTTGTAC	gCSK_1_R
CACCGGAGCGGCTTCTGTACCCGC	gCSK_2_F
AAACGCGGGTACAGAAGCCGCTCC	gCSK_2_R
CACCGGCAACTGCGGCATAGCAACC	gCSK_3_F
AAACGGTTGCTATGCCGCAGTTGCC	gCSK_3_R
CACCGGTGACCTGCCCGGTTCTCAG	gAAVS1_1_F
AAACCTGAGAACCGGGCAGGTCACC	gAAVS1_1_R
CACCGCGGGGACACAGGATCCCTGG	gAAVS1_2_F
AAACCCAGGGATCCTGTGTCCCCGC	gAAVS1_2_R
CACCGGTTAGGGCATGCCAGAACAG	gPAK2_1_F
AAACCTGTTCTGGCATGCCCTAACC	gPAK2_1_R
CACCGATGGTGTGCTCAAAATCAGA	gPAK2_2_F
AAACTCTGATTTTGAGCACACCATC	gPAK2_2_R
CACCGGGACATCCAGCACAGCCTG	gPAK2_3_F
AAACCAGGCTGTGCTGGATGTCCC	gPAK2_3_R
CACCGGTTTGAGAAGATTGGACAA	gPAK1_1_F
AAACTTGTCCAATCTTCTCAAACC	gPAK1_1_R
CACCGCCTCTGCCTCCAAACCCAG	gPAK1_2_F
AAACCTGGGTTTGGAGGCAGAGGC	gPAK1_2_R
CACCGCGCCGCAGCCTCAGACCCG	gESR1_1_F
AAACCGGGTCTGAGGCTGCGGCGC	gESR1_1_R
CACCGCACCATTGATAAAAACAGG	gESR1_2_F
AAACCCTGTTTTTATCAATGGTGC	gESR1_2_R
CACCGTGGTGGTGGCCTTCAAATCA	CSK_eh_gRNA1_F
AAACTGATTTGAAGGCCACCACCAC	CSK_eh_gRNA1_R
CACCGCAGGGAGCAGCCCACGGTAG	CSK_eh_gRNA2_F
AAACCTACCGTGGGCTGCTCCCTGC	CSK_eh_gRNA2_R
CACCGAGCGCCACCAGAGACCAGAC	CSK_eh_gRNA3_F
AAACGTCTGGTCTCTGGTGGCGCTC	CSK_eh_gRNA3_R
CACCGCTAGAATCCAGTCTGGTCTC	CSK_eh_gRNA4_F
AAACGAGACCAGACTGGATTCTAGC	CSK_eh_gRNA4_R
CACCGAGTAATCACCCAGAGTGCAA	CSK_eh_gRNA5_F
AAACTTGCACTCTGGGTGATTACTC	CSK_eh_gRNA5_R
CACCGAGAGGACTTGGAGTCGCTGA	CSK_eh_gRNA6_F
AAACTCAGCGACTCCAAGTCCTCTC	CSK_eh_gRNA6_R
CCTTGAAGGAAGATGATCAAATGAGAGC	CSK-eh-PCR_F
CCAGCCTGGGGCCAGTTCTTATC	CSK-eh_PCR_R

For enhancer deletion by pairs of gRNA, we first modified the LentiCRIPSR V2 vector by substituting blasticidin resistant gene for puromycin resistant gene. Then we cloned the

CSK_eh_gRNA1, CSK_eh_gRNA2, CSK_eh_gRNA3, and CSK_eh_gRNA5 into LentiCRISPR_puro vector, and CSK_eh_gRNA3, CSK_eh_gRNA4, CSK_eh_gRNA3, and CSK_eh_gRNA6 into LentiCRISPR_blast vector. After a pair of gRNA (gRNA1+gRNA3) was delivered into cells by lentivirus, we selected the cell by both puromycin and blasticidin.

The pLX-gRNA vector (Addgene) was used to generate lentiviral gCSK_1, gCSK_3, gAAVS1_1, gAAVS_2 vectors for the secondary CRISPR screens by the protocol from Addgene.

The vectors of inducible overexpression of CSK and PAK2 were generated by cloning the ORFs of CSK and PAK2 genes into the pCW-Cas9 vectors. We substituted the CSK or PAK2 genes for Cas9 by double restriction enzyme digestion (NheI and BamHI). The primers were used as follows:

AGTCAGCTAGCATGTCAGCAATACAGGCCGCCTG	CSK_nheI_F
AGTCAGGATCCTCAAGCGTAATCTGGAACATCGTATGGGTACAGGTGCAGCTCGTG GGTTTTG	CSK_BamHI_R
AGTCAGCTAGCATGTCTGATAACGGAGAACTGGAAGATAAGCC	PAK2_nheI_F
AGTCAGGATCCTTACTTGTCGTCATCGTCTTTGTAGTCACGGTTACTCTTCATTGCTT CTTTAGCTGCC	PAK2_BamHI_R

The gCSK resistant CSK cDNAs were generated by introducing a mutation (NGG->NTG) at PAM without changing the amino acid. And the Q5® Site-Directed Mutagenesis Kit (NEB) was used with the primers:

GTTGGCCGTGAGGGCATCATC	CSK1_mut_R_F
CTTGTTTTTGGCTTTGTACCAGTTGG	CSK1_mut_R_R
CGCCTGAGACAGGCCTGTTCCTG	CSK2_mut_R_F
GGTACAGAAGCCGCTCAGCCTGCTC	CSK2_mut_R_R
CTTGTGCAGCTCCTGGGCGTGA	CSK3_mut_R_F
GTTGCTATGCCGCAGTTGCGTCA	CSK3_mut_R_R

The gPAK2_3 targets the intron-exon boundary of PAK2 in human genome, thus it will not affect the PAK2 cDNA.

Amino-acid substitution mutants of PAK2 (Y130F, Y139F, Y194F) were generated by the Q5® Site-Directed Mutagenesis Kit (NEB) with the following primers:

AAGTTCTTCGACTCCAACACAGTGAAGCAGA	PAK2_mut_Y130_F
TAGGACATCCAGCACAGCCTGAGG	PAK2_mut_Y130_R
CAGAAATTTCTGAGCTTTACTCCTCCTGAGAAAGATG	PAK2_mut_Y139_F
CTTCACTGTGTTGGAGTCGTAGAACTTTAGGAC	PAK2_mut_Y139_R
TCAATTTTCACACGGTCTGTAATTGACCCTG	PAK2_mut_Y194_F
TTTCGTATGATCCGGTCGCGG	PAK2_mut_Y194_R

Inhibitors used in this work include: dasatinib, saracatinib, and FRAX597 were purchased from Selleck Chemicals. E2, Tamoxifen and fulvestrant were purchased from Sigma.

CRISPR screens

GeCKO v2 library(2) from Addgene was used for the genome-wide CRISPR screens. Cells of interest were infected at a low MOI (0.3～0.5) to ensure that most cells receive only 1 viral construct with high probability. To find optimal virus volumes for achieving an MOI of 0.3–0.5, each new cell type and new virus lots were tested by spinfecting $3x10^6$ cells with several different volumes of virus. Briefly, $3x10^6$ cells per well were plated into a 12 well plate in the appropriate standard media for the cell type (see below) supplemented with 8 ug/ml polybrene. For T47D cells, standard media was RPMI 1640 supplemented with 10 % FBS. Each well received a different titrated virus amount (usually between 5 and 50 ul) along with a no-transduction control. The 12-well plate is centrifuged at 2,000 rpm for 2 h at 37°C. After the spin, media was aspirated and fresh media (without polybrene) was added. Cells were incubated overnight and then enzymatically detached using trypsin. Cells were counted and each well was split into duplicate wells. One replicate receives 2 ug/mL puromycin for MCF7 cells or 3.5 ug/ml puromycin for T47D cells. After 3 days (or as soon as no surviving cells remained in the no-transduction control under puromycin selection), cells were counted to calculate a percent transduction. Percent transduction was calculated as cell count from the replicate with puromycin divided by cell count from the replicate without puromycin multiplied by 100. The virus volume yielding a MOI closest to 0.4 was chosen for large-scale screening.

For each cell lines, large-scale spin-infection of $2x10^8$ cells was carried out using four of 12-well plates with $4x10^6$ cells per well. Wells were pooled together into larger flasks on the day after spin-infection. For most cell types, 0.5-4 ug/ml puromycin works well, although the minimum dose that kills all cells without any viral transduction was determined in advance and the minimum concentration was used for selection. After three days of puromycin selection, the surviving cells (T47D and MCF7) were divided into three groups (0 day control, vehicle, and E2 (10nM)) and cultured in hormone depleted medium for four weeks before genomic DNA extraction and analysis. Two rounds of PCR were performed after gDNA had been extracted, and 300 ug DNA per sample was used for library construction. Each library was sequenced at 30~40 million reads to achieve ~300X average coverage over the CRISPR library. The 0 day sample library of each screen could serve as controls to identify positively or negatively selected genes or pathways.

For the second round of Genome-wide CRISPR screens, we first infected the T47D cells with lentiviral gCSK_1, gCSK_3, gAAVS1_1, gAAVS_2 cloned by pLX-gRNA vector. After blasticidin selection, we generated these four types of T47D cells with stable expression of gCSK_1, gCSK_3, gAAVS1_1, gAAVS_2 respectively. Then the Genome-wide CRISPR screens were performed in these four cell types by the above method.

PCR primers for library construction:

The first round of PCR:

AATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG	lentiCRISPR_F1
TCTACTATTCTTTCCCCTGCACTGTACCTGTGGGCGATGTGCGCTCTG	lentiCRISPR_R1

The second round of PCR:

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCG ATCTTCTTGTGGAAAGGACGAAACACCG	Cri_library_F
CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTXXXXXXTCTACTATTCTTTCCCCTGCACTGTACC	Cri_library_R

(XXXXXX denotes the sample barcode)

Sequencing primer (read1): GCTCTTCCGATCTTCTTGTGGAAAGGACGAAACACCG

Indexing primer: CATCGCCCACAGGTACAGTGCAGGGGAAAGAATAGTAGA

Lentivirus production and purification

T-225 flasks of 293FT cells were cultured at 40%~50% confluence the day before transfection. Transfection was performed using Lipofectamine 2000 (Life Technologies). For each flask, 20 ug of lentivectors, 5 ug of pMD2.G, and 15 ug of psPAX2 (Addgene) were added into 4 ml OptiMEM (Life Technologies). 100 ul of Lipofectamine 2000 was diluted in 4 ml OptiMEM and, after 5 min, it was added to the plasmid mixture. The complete mixture was incubated for 20 min before being added to cells. After 6 h, the media was changed to 30 ml DMEM + 10% FBS. After 60 h, the media was removed and centrifuged at 3,000 rpm at 4 °C for 10 min to pellet cell debris. The supernatant was filtered through a 0.45 um low protein binding membrane. The virus was ultracentrifuged at 24,000 rpm for 2 h at 4 °C and then resuspended overnight at 4°C in DMEM + 10% FBS. Aliquots were stored at –80°C.

Real-time RT-PCR

Real-time RT-PCR was performed as described before (3). Data are presented as mean ± standard deviation (SD). Primers used for RT-PCR are listed as follows:

GGTGTGAACCATGAGAAGTATGA	GAPDH_qF1
GAGTCCTTCCACGATACCAAAG	GAPDH_qR1
CGGAATCCTTCTCTGGGAAATC	CSK_qF1
CATCCATCTTGTAGCCCTTCTC	CSK_qR1

Immunoblot

The western blotting was performed as described before (4). Specific antibodies used in our work include: anti-CSK (sc-286), anti-c-Src (sc-18), anti-p-c-Src Tyr530 (sc-101803), anti-p-c-Src Tyr 419 (sc-101802), anti-GAPDH (sc-25778), and anti-ESR1 (sc-543) from Santa Cruz Biotechnology, anti-PAK2 (A301-264A) from Bethyl Lab, anti-p-PAK2 Ser141 (2606) and anti-PAK1 (2602) from Cell Signaling technology.

Copy number, gene expression and epigenetics profiling analysis

The copy number variation (CNV) data from both T47D and MCF7 cells are downloaded from the Cancer Cell Line Encyclopedia (CCLE) (5) project. The gene expressions of CSK null and AAVS1 knockout T47D cells are quantified and analyzed from RNA-seq reads using Kallisto (6) and DESeq2 (7). We use the expression profiles in Cancer Cell Line Encyclopedia (CCLE) (5) to compare between breast cancer and non-breast cancer cell lines (Fig. 1F). The processed gene expression values are downloaded directly from the CCLE website.

We make use of several public epigenetics profiles in T47D cells in Fig. 3A and 3B, including genomic DNase-I footprints (8), ER ChIP-seq (9), FOXA1 ChIP-seq (10) and GATA3 ChIP-seq (11). The data from FOXA1 and GATA3 ChIP-seq is not shown (Fig. 3B) since there are no FOXA1/GATA3 bindings in the putative CSK enhancer. The raw reads of these studies (together with reads from H3K27ac ChIP-seq experiments) are first mapped to human hg38 reference genome using Bowtie2 (12), and the peaks are identified using MACS2 (13).

Survival analysis

The gene expressions of breast cancer patients from different cohorts are downloaded from NCBI Gene Expression Omnibus (GEO) with accession number GSE12093, GSE6532, GSE9195, GSE17705. The R “survival” package is used for the survival analysis.

The CSK-suppressed gene signatures and the gene expression data from clinical trials are used to evaluate the effect of CSK loss. The 292 CSK-suppressed signature genes are those that are up-regulated consistently upon CSK knockout from both MCF7 and T47D (log2 fold change >1 and adjusted p-value <0.01, Fig. S4C). For each clinical trial expression dataset, patients are separated into two groups according to the expressions of CSK gene signatures: Higher expression of CSK-suppressed signature genes as a surrogate of CSK inhibition after treatment and Others. The overall expression of these signature genes, or the "signature gene score", is calculated by weight-averaging the expressions of all genes, using the method described in (14-16).

Network analysis

GeneMania (17) is used to construct the network of primary screens (Fig. 1C) and CSK synthetic lethal gene network (Fig. 5A and S5C). In GeneMania, different networks collected from public datasets, including co-localization, genetic interaction, pathway, physical interaction, and shared protein domain networks are used to connect genes. Network construction is performed through GeneMania CytoScape plugin (18), while the networks are visualized using Cytoscape (19).

Accession numbers

RNA-seq data is deposited in NCBI Short Read Archive (SRA) under the accession number SRX1653056 (T47D) and SRX2188492 (MCF7). H3K27ac ChIP-seq data is deposited in SRA under the accession number SRX1653117. CRISPR screening data is deposited in SRA under the accession number SRX1653118.

SRR3278937 RNA-seq of T47D cells, AAVS1 knock-out
SRR3278984 RNA-seq of T47D cells, CSK knock-out using g_CSK1, vehicle treatment
SRR3278985 RNA-seq of T47D cells, CSK knock-out using g_CSK2, vehicle treatment
SRR3278992 RNA-seq of T47D cells, CSK knock-out using g_CSK3, vehicle treatment
SRR4293006 RNA-seq of MCF7 cells, CSK knock-out using g_CSK1, vehicle treatment
SRR4293010 RNA-seq of MCF7 cells, CSK knock-out using g_CSK2, vehicle treatment
SRR4293016 RNA-seq of MCF7 cells, CSK knock-out using g_CSK3, vehicle treatment
SRR4293022 RNA-seq of MCF7 cells, AAVS1 knock-out
SRR3278995 H3K27ac ChIP-seq of T47D cells, replicate 1
SRR3278996 H3K27ac ChIP-seq of T47D cells, replicate 2
SRR3278998 CRISPR screens, T47D, day0
SRR3278999 CRISPR screens, T47D, vehicle treatment, week 4
SRR3279002 CRISPR screens, T47D, E2 treatment, week 4
SRR3279335 CRISPR screens, MCF7, day0
SRR3279336 CRISPR screens, MCF7, vehicle treatment, week 4
SRR3279337 CRISPR screens, MCF7, E2 treatment, week 4
SRR3279406 CRISPR secondary screens, T47D, AAVS1 knockout, day0
SRR3279408 CRISPR secondary screens, T47D, CSK knockout, day0
SRR3279409 CRISPR secondary screens, T47D, AAVS1 knockout, E2 treatment
SRR3279410 CRISPR secondary screens, T47D, AAVS1 knockout, E2 treatment
SRR3279413 CRISPR secondary screens, T47D, CSK knockout, vehicle treatment
SRR3279414 CRISPR secondary screens, T47D, CSK knockout, vehicle treatment
SRR3284761 CRISPR secondary screens, T47D, AAVS1 knockout, vehicle treatment
SRR3284767 CRISPR secondary screens, T47D, AAVS1 knockout, vehicle treatment
SRR3284781 CRISPR secondary screens, T47D, CSK knockout, E2 treatment
SRR3284806 CRISPR secondary screens, T47D, CSK knockout, E2 treatment

References

1. GTEx Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348(6235):648–660.

2. Sanjana NE, Shalem O, Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11(8):783–784.

3. Xiao T, et al. (2012) A differential sequencing-based analysis of the C. elegans noncoding transcriptome. *RNA* 18(4):626–639.

4. Xiao T, et al. (2015) Long Noncoding RNA ADINR Regulates Adipogenesis by Transcriptionally Activating C/EBPα. *Stem Cell Reports* 5(5):856–865.

5. Barretina J, et al. (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483(7391):603–607.

6. Bray N, Pimentel H, Melsted P, Pachter L (2015) Near-optimal RNA-Seq quantification.

7. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 15(12):550.

8. Neph S, et al. (2012) An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489(7414):83–90.

9. Ross-Innes CS, et al. (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481(7381):389–393.

10. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS (2011) FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet* 43(1):27–33.

11. Gertz J, et al. (2013) Distinct properties of cell-type-specific and shared transcription factor binding sites. *Mol Cell* 52(1):25–36.

12. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.

13. Zhang Y, et al. (2008) Model-based analysis of ChIP-Seq (MACS). 9(9):R137.

14. Lim E, et al. (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 15(8):907–913.

15. Hanker AB, et al. (2013) Mutant PIK3CA accelerates HER2-driven transgenic mammary tumors and induces resistance to combinations of anti-HER2 therapies.

110(35):14372–14377.

16. Zhao X, et al. (2015) RAS/MAPK Activation Drives Resistance to Smo Inhibition, Metastasis, and Tumor Evolution in Shh Pathway-Dependent Tumors. *Cancer Res* 75(17):3623–3635.

17. Warde-Farley D, et al. (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 38(Web Server issue):W214–20.

18. Montojo J, et al. (2010) GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* 26(22):2927–2928.

19. Shannon P, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504.