Maurer et al. Supplementary Information

Additional Acknowledgements.

We thank Dario Alessi, Yoshihiro Nakatani, and Nancy Hynes for providing constructs, Victor Velculescu for *PDPK1* sequencing primers, Gary Nolan for Phoenix-ampho cells, Devin McQuaid for assistance with array CGH, and Chi-Ming Li, Martha Posada, Benjamin Tycko and Thomas Ludwig for their technical assistance with Southern blotting. We also thank Barry Fine, Amy Gerstein, Megan Keniry, Ita Horan, and Wei Xia. We thank Drs. Hale, Davies, Gonzalez, Yu, Carey, Agarwal and Sahin for providing RPPA and sequencing data. We thank Drs. Gabor Balazsi and Katrina Podsypanina for discussion. We thank Donald Landry for helping with the synthesis of BX-795.

Author Contributions

T.S. performed the PDPK1 RT-PCR as well as assisted with the breast cancer tissue bank; L.H.S. was the major contributor to the *PIK3CA* and *PTEN* sequencing, and the PTEN IHC; S.K. made the PDK1 shRNA cell lines, assisted in generating expression array data, and assisted with the T47D proliferation assay; B.D.H. assisted in all aspects of the mouse experiments; C.R.B., J.W., and J.J.Z. performed the HMEC/hTERT/p53DD/H1047R colony formation and wortmannin proliferation assays; S.N. and V.V.V.S.M. performed the *PDPK1* FISH; B.D. and G.B.M. performed the analysis of the protein lysate array; Y.X. synthesized BX-795; Y.R.C. and A.T. performed the AKT siRNA migration assay; D.K. confirmed the PTEN RNAi signaling experiment; J.S.F. and M.B.T. performed the *PDPK1* ICN patient survival analysis and assisted with Table 1; S.K.G. generated the SNP array data and assisted in generating the expression array data; M.L. and J.I. performed the *HER2* CISH; X.W. assisted with the PDPK1 RT-PCR;

L.M. constructed the tissue microarray; A.R. maintained the breast cancer clinical data; T.M. and C.C. performed the cell block PDK1 IHC; J.S.Y. provided technical assistance; H.H. founded the directs the Columbia Breast Macromolecule bank which was the source of the tissue, DNA, and RNA used to study the Columbia cohort of breast cancers, provided technical assistance and scored the PTEN and PDK1 IHC; M.M. performed all other experiments and analyses; M.M. and R.P. conceived the work and wrote the manuscript; All authors discussed the results and commented on the manuscript.

Supplementary Table S1. Clinical and pathologic tumor characteristics and their association with increased *PDPK1* copy number (Columbia cohort).

	total	(%)	≤4 copies	(%)	≥5 copies	(%)	p value†
Aneuploid	53	(50)	40	(48)	13	(62)	
Diploid	30	(29)	28	(33)	2	(10)	
Tetraploid	22	(21)	16	(19)	6	(29)	
							0.09
ER positive	90	(70)	70	(69)	20	(74)	
ER negative	39	(30)	32	(31)	7	(26)	
							0.58
PR positive	80	(62)	61	(60)	19	(70)	
PR negative	49	(38)	41	(40)	8	(30)	
							0.31
Stage 1	28	(24)	22	(24)	6	(24)	
Stage 2	59	(50)	48	(52)	11	(44)	
Stage 3	30	(26)	22	(24)	8	(32)	
							0.68
≤50 years	52	(40)	42	(41)	10	(37)	
>50 years	77	(60)	60	(59)	17	(63)	
							0.70
White	82	(64)	63	(62)	19	(70)	
Non-White	47	(36)	39	(38)	8	(30)	
							0.41

*p-value < 0.05 denoting significant enrichment in the number of breast cancers with *PDPK1* increased copy number (ICN, ≥ 5 copies) among those with clinical and pathologic tumor characteristics (as noted) and the remainder of cases.

† p-value from Fisher's Exact test, all other p-values from Chi square test.

Supplementary Materials and Methods

PDK1 siRNA and Cell Line IHC. siRNA target sequences: PDK1-1: 5′-AACTGGCAACCTCCAGAGAAT-3′(1), PDK1-2: 5′-AAGTCCGCCTGTAAGAGTTCA-3′(2), ctrl: siControl Nontargeting siRNA #1 (Dharmacon). $1x10^8$ cells were transfected with HiPerFect using standard protocols with 100 pmols siRNA. Cells were harvested at 24 hours and first embedded in 1% PBS agarose before fixation and paraffin embedding. PDK1 IHC used antibody PKB Kinase (E-3) Santa Cruz, 1:5000.

***PDPK1* Map.** The BAC employed in the FISH analysis (RP11-67B18) contains genetic sequence from genes *ATP6V0C*, *CEMP1*, and *AMDHD2*, as well as the *PDPK1* pseudogene (*PDPK2*) that lies directly centromeric of *PDPK1* (Supplementary Fig. S1*C*). *PDPK1* sequence within the BAC was confirmed by colony PCR of exon 11, sequence that is unique to PDPK1 and not contained within the pseudogene. The pseudogene, orientated in the opposite direction, contains 100% exon and intron sequence homology beginning within exon 2 and extends to the end of exon 10, sparing exons 1 and 11-14.

Southern blotting. Primers to amplify *PDPK1* exon 11 were 5'-GTCCCATGGAGGAGAATCAG and 5'-GCTGAGGGAGGCAGTGAC. Primers for MX1 probe were obtained from Chi-Ming Li (unpublished). Standard protocols were used as previously described(3).

PDPK1 quantitative RT-PCR: **RNA extraction:** Total RNA was isolated using Qiazol (Qiagen,CA), and purified with miRNeasy Kit (Qiagen, CA). During purification, DNase1 treatment was performed on kit column according to the manufacturer's instruction. RNA integrity was confirmed by agarose gel electrophoresis and the concentration was quantified by

measuring OD260nm in BioPhotometer (Eppendorf, NY). **Reverse transcription:** Reverse transcription of cDNA first strand was carried out as following: 3ug of total RNA, 400ng of random primer (Invitrogen, CA) and 150ng oligo dT24 were incubated together in a total volume of 12ul in 70°C for 10min and chilled on ice. Then it was mixed with 8ul of reverse transcription reaction mixture to get final 20ul volume (final concentration: 1x first-strand buffer, 10mM DTT, 500nM dNTPs and 200 unit SuperScript III reverse transcriptase, all from Invitrogen, CA). The reaction was incubated at 25°C for 10min, and then 50°C for 60min, and followed by 85°C 5min. **PCR Primer design:** mRNA sequence were obtained from public GeneBank database (www.ncbi.nlm.nih.gov), and primers were designed using Primer3 software from MIT university (frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). All primers were synthesized by Invitrogen Company, and quality confirmed with no primer dimer and only one peak in dissociation curve in pre-PCR, and only one band in agarose gel electrophoresis. All PCR products were within 100-160bp range. **Quantitative PCR:** Comparative quantitative PCR was performed in Mx3005p real-time PCR machine (Stratagene, CA), using QuantiTect SYBR Green PCR kit (Qiagen, CA). PCR was carried out in total 20ul volume containing final concentration of 1x reaction buffer, 300nM each forward and reverse primer and 10ng DNA or cDNA (based on amount of input RNA). The PCR was hot-started at 95 °C for 15 min, and then 45 cycles of 95 °C for 20sec, 54°C for 30sec, 72°C for 30sec, and then followed by dissociation curve measurement from 54 to 95°C. No template control (NTC) and no reverse transcription control (NRT) were run together with every assay and all samples were run at least in duplicates. PCR data was saved and analyzed with software Mxpro (Stratagene, CA). Comparative quantitative analysis was performed based on delta-delta Ct method using housekeeping gene GAPDH as internal control.

Primer sequence used in qPCR.

PrimerName	EnsemblTranscriptId	GBId	GeneSymbol	Length	Sequence	Location	Start
PDPK1-BF	ENST00000342085	AF017995	PDPK1	1891	gatctggactcgaactcctttg	exon11	1332
PDPK1-BR	ENST00000342085	AF017995	PDPK1	1891	aaacccttccgcttatccac	exon12-13	1486
GAPDH-BF	ENST00000229239	NM_002046	GAPDH	1310	ggcctccaaggagtaagacc	exon9	1095
GAPDH-BR	ENST00000229239	NM_002046	GAPDH	1310	aggggtctacatggcaactg	exon9	1241

Single nucleotide polymorphism microarrays. SNPs were genotyped using Affymetrix Gene Chip Human Mapping 250K StyI Arrays (Affymetrix, Santa Clara, CA) according to the manufacturer's instructions. 250 ng of genomic tumor DNA was used and the labeled products were hybridized and scanned in the Microarray Facility at the Herbert Irving Comprehensive Cancer Center, Columbia University. The dChip software 2007 was used to analyze the array data as described previously(4). Data was normalized to a baseline array with median signal intensity at the probe intensity level using the invariant set normalization method. A model-based (PM/MM) method was used to obtain the signal values for each SNP in each array. Signal values for each SNP were compared with the average intensities from 22 normal DNA samples from the same patient set. To infer the DNA copy number from the raw signal data, we used median smoothing with a window size of 11 SNPs based on the assumption of diploidy for normal samples. Mapping information of SNP locations and cytogenetic band were based on curation of Affymetrix and University of California Santa Cruz hg17.

Expression array analysis: RNA was isolated from breast tumors as described under RT-PCR methods. 500ng total RNA was labeled using Agilent Low RNA Input Linear Amplification Kit, PLUS, Two-color (catalog number 5188-5340) as per manufacturer's instructions. Universal Human Reference RNA (Stratagene) was used as a reference RNA. Microarray chips (Agilent

#G4112F [4 x 44K]) were hybridized at the Herbert Irving Comprehensive Cancer Center microarray facility and data was extracted using Agilent Feature Extraction version 9.5.3.1.

CGH: DNA sent to Roswell Park Cancer Institute employing their human 19K RPCI-11 BAC array with an average resolution of 0.286 megabases (range 0.234-0.330). ArrayExpress accession: E-MEXP-2141.

Statistics. We estimated survival curves and multivariable models using Kaplan-Meir and Cox Proportional Hazard modeling, respectively. We estimated hazard ratios adjusting for age and stage. We tested whether adjustment by hormone receptor status, tumor ploidy, and race altered the hazard ratio for increased PDPK1 copy number by more than 10%.

PTEN RNAi: Cells transfected (HiPerFect, Qiagen) as per protocol with 10nmol final concentration of PTEN RNAi oligo (Dharmacon Individual siGenome Duplex, D-003023-05) or control oligo (Dharmacon, siCONTROL Non-Targeting siRNA #1). Cells were harvested at 48 hours.

Matrigel tissue culture. Cell lines were purchased from the ATCC with MCF10A cells grown as described (5).

IHC of paraffin-embedded matrigel. antigen retrieval via microwave in citrate, PDK1 (PKB Kinase (E-3) Santa Cruz, 1:100), E-cadherin (BD Biosciences, 1:500), anti-laminin 5 (Chemicon MAB19562, 1:200), N-cadherin (clone 3B9, Chemicon IHC2129-6, as prepared), cleaved caspase 3 (Asp 175, Cell Signaling, 1:100).

HMEC methods. tHMEC-DD-H1047R cells were generated as described previously (6). For the wortmannin dose response curve, cells were cultured for 3 days. For the growth curve, cells were cultured for 4 days with 15 nM of wortmannin. Wild type and myristolated human PDPK1

sequences were cloned into the retroviral vector pWZL-neo. The shRNA (sh-luc and sh-PDPK1s) were from the RNAi consortium at the Broad Institute. Standard protocols were used.

Supplementary Figure Legends

Figure S1. Increased copy number (ICN) of *PDPK1* and increased expression of PDK1 are found in a significant portion of human breast cancers. (*A*) immunoblot from MCF10A cells showing knockdown of PDK1 with two separate siRNAs. (*B*) IHC staining for PDK1 in MCF10A cells over-expressing PDK1 with control or PDK1 siRNA. (*C*) *PDPK1* genomic region schematic showing BAC RP11-67B18 encompassing both *PDPK1* and its pseudogene *AJ785968*. (*D*) *PDPK1* Southern blot yielding a single band. The signal is compared with probe to the gene *MX1* located on chromosome 21. Tumors 355 and 462 had *PDPK1* ICN by FISH (both with 6 copies, tumor 23 was not tested by FISH). (*E*) Immunoblot of PDK1 protein isolated from breast cancers compared with quantitative RT-PCR for expression of mRNA, FISH copy number, and IHC score.

Figure S2. *PDPK1* ICN cases clustered within two separate groups and *PDPK1* located within peak of the uniquely narrow 16p13.3 amplicon of case #432. (*A*) SNP array analysis of chromosome 16 (each block is average of 100 SNPs, red=increased, green=decreased) with unsupervised clustering listed by tumor number-*PDPK1* copy number (column) against chromosome 16 location, corresponding PI3K lesions identified above (PTEN = loss (IHC), *ERBB2* = amplification (CISH), *PIK3CA* = mutation). (*B*) expression array ($\log_{10}$ single [dotted line] and dual channel [solid line]) with criteria for gene labeling of at least a three-fold increase in expression compared with control (dual channel) and at least a 10 fold increase in expression compared to the median of all genes in the sample (single channel); all aligned against the SNP amplicon map (below), both plotted against 16p13.3 chromosomal location.

Figure S3. CGH data of chromosome 16 showing the different patterns of *PDPK1* gain. Tumor 69: broad amplicon. 432: steep peak directly at *PDPK1* locus. 582: gain of the entire 16p arm.

591: multiple regions of gain throughout the p arm. 614 and 619: likely increase of 16p. 210 and 73: no obvious gain at *PDPK1* locus. *PDPK1* locus indicated by arrow at top or page and pink square at the top of each plot. X axis (megabases), Y axis (log_2 ratio (test/control)).

Figure S4. Increased copy number of PDPK1 is associated with shortened patient survival. Kaplan-Meier survival curves of breast cancer patients with (lower curve) and without *PDPK1* ICN. Hazard ratio (HR) = 3.14 [95% CI = 1.3-7.6] for *PDPK1* ICN cases compared to those without ICN adjusted for age at diagnosis and stage of disease (p=0.04, Log-Rank Test).

Figure S5. Over-expression of PDK1 enhances growth and signaling in the setting of upstream PI3K activation. (*A*) Proliferation assay of MCF10A cells as indicated in absence of growth factors (s.d., n=3). (*B*) Immunoblots showing RNAi knockdown of PTEN protein expression vs. control in MCF10A cells with and without over-expression of PDK1. (*C*) Immunoblots from MCF7 cells in normal growth media or starved of growth factor for 24 hours, both with and without over-expression of PDK1.

Figure S6. Over-expression of PDK1 enhances transformation and migration in the setting of NeuT. (*A*) Immunohistochemical analysis of cells pictured in Fig. 3*A* with the indicated antibodies. (*B*) Scratch test comparing the ability of MCF10A cells over expressing NeuT vs. cells over-expressing both NeuT and PDK1 to migrate in the setting of serum starvation over 36 hours.

Figure S7. In cells with PI3K activation, PDK1 levels are a determinant of signaling, proliferation, transformation, and pathway inhibition. (*A*) Proliferation assay of T47D cells with stable knockdown of PDK1 with two separate PDK1 shRNA constructs compared with control (s.d., n=3). (*B*) Immunoblots from MCF10A cells starved of growth factor for 24h then stimulated with EGF either alone or with a 30 minute pre-incubation with BX-795 in μM

concentrations as indicated. (*C*) Colony quantification from colony formation assay of HMEC/hTERT/p53DD cells over-expressing mutant p110α (H1047R) or control, or p110α (H1047R) in the setting of stable PDK1 shRNA compared with control shRNA (s.d., n=3). (*D*) Growth proliferation assay in 15 nM wortmannin of HMEC/hTERT/p53DD/H1047R cells over-expressing wild type PDK1 or myristolated PDK1 compared to control (s.d., n=3).

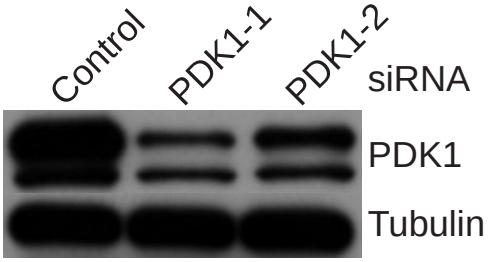
(*E*) Growth proliferation assay of HMEC/hTERT/p53DD/H1047R cells over-expressing wild type PDK1 or myristolated PDK1 compared to control over a dose range of wortmannin [nM] (s.d., n=3). (*F*) Immunoblot of PDK1 in HMEC/hTERT/p53DD/H1047R cells compared with cells over-expressing wild type PDK1 or myristolated PDK1.

Supplementary References.

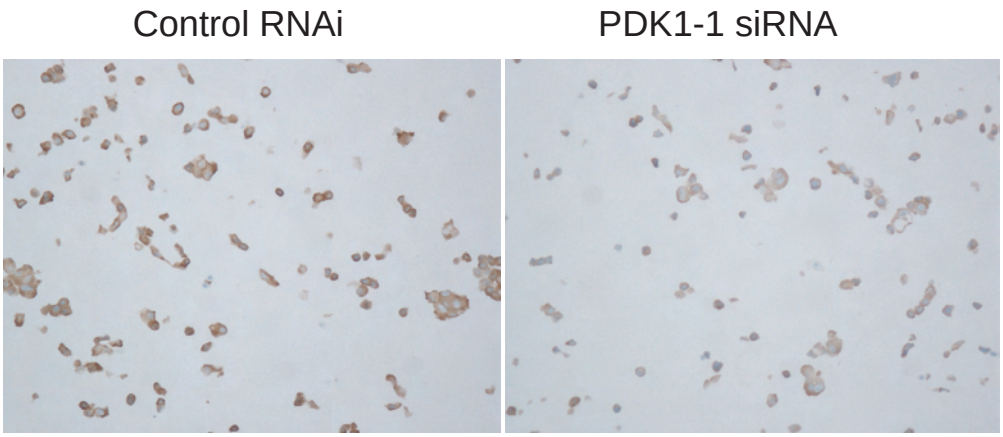
1. Bilanges B and Stokoe D Direct comparison of the specificity of gene silencing using antisense oligonucleotides and RNAi. Biochem J 2005;388:573-583.
2. Xie Z, Zeng X, Waldman T, and Glazer RI Transformation of mammary epithelial cells by 3-phosphoinositide- dependent protein kinase-1 activates beta-catenin and c-Myc, and down-regulates caveolin-1. Cancer Res 2003;63:5370-5375.
3. Li CM, Margolin AA, Salas M, *et al.* PEG10 is a c-MYC target gene in cancer cells. Cancer Res 2006;66:665-672.
4. Zhao X, Li C, Paez JG, *et al.* An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 2004;64:3060-3071.
5. Debnath J, Muthuswamy SK, and Brugge JS Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. Methods 2003;30:256-268.
6. Zhao JJ, Liu Z, Wang L, Shin E, Loda MF, and Roberts TM The oncogenic properties of mutant p110alpha and p110beta phosphatidylinositol 3-kinases in human mammary epithelial cells. Proc Natl Acad Sci U S A 2005;102:18443-18448.

Fig. S1

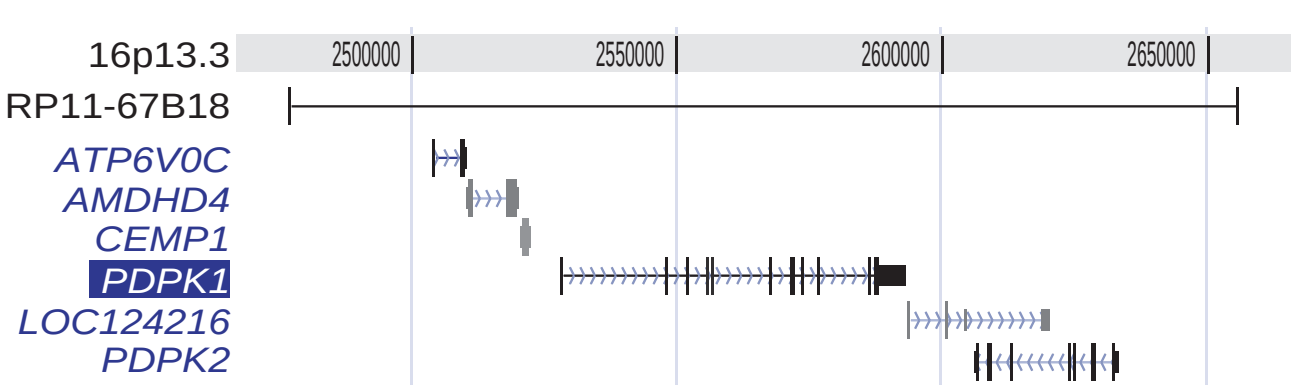
A


Control
PDK1-1
PDK1-2
siRNA
PDK1
Tubulin

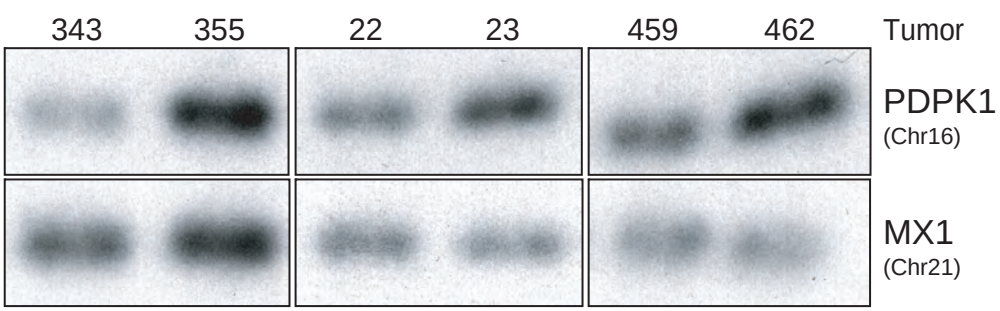
B


Control RNAi
PDK1-1 siRNA

C


16p13.3
2500000
2550000
2600000
2650000
RP11-67B18
ATP6V0C
AMDHD4
CEMP1
PDPK1
LOC124216
PDPK2

D


343
355
22
23
459
462
Tumor
PDPK1
(Chr16)
MX1
(Chr21)

E

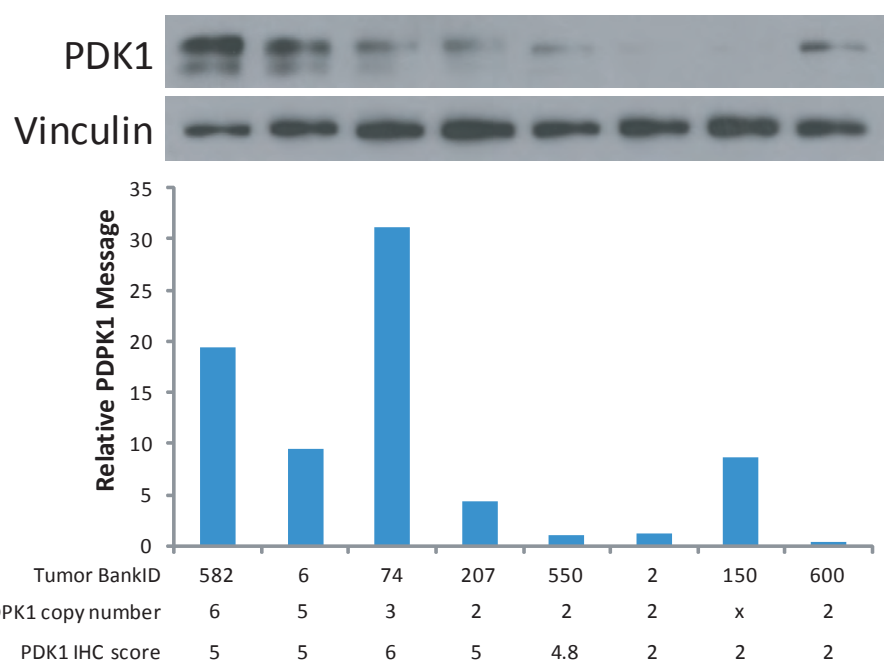
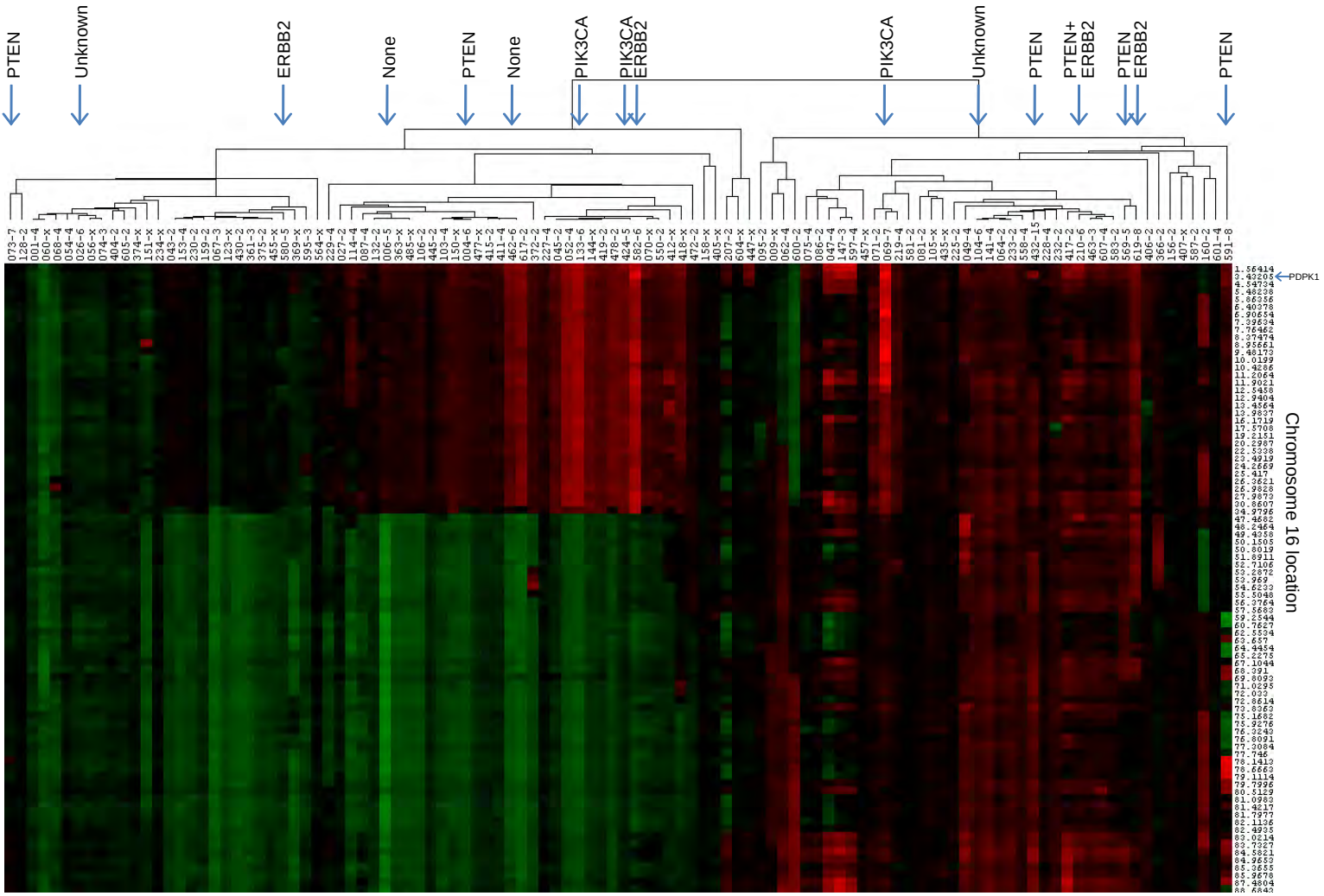
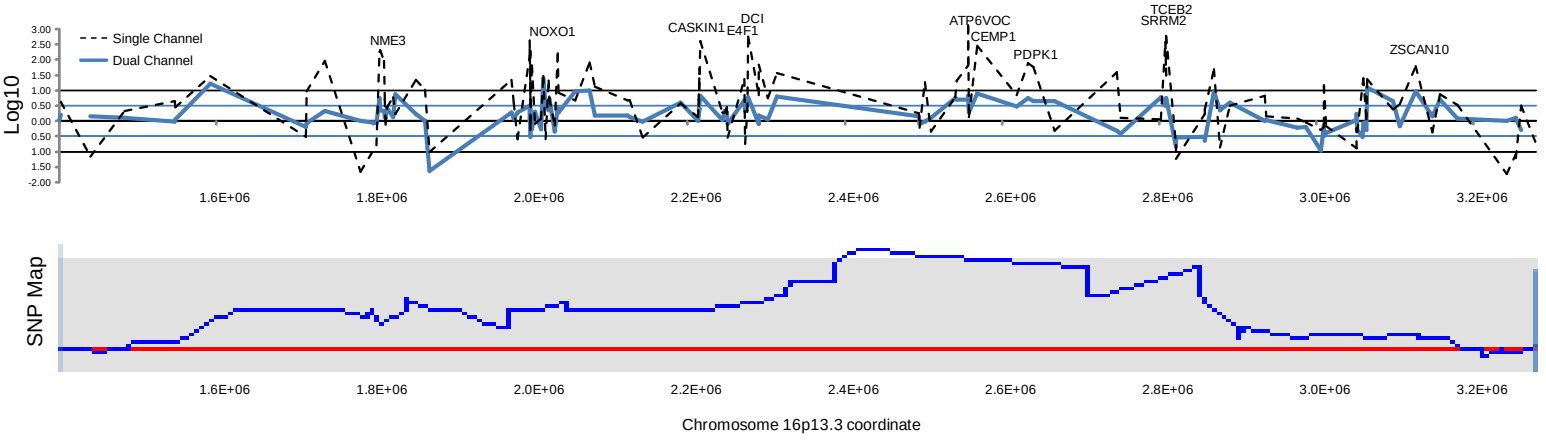

PDK1
Vinculin
Relative PDPK1 Message
35
30
25
20
15
10
5
0
Tumor BankID 582 6 74 207 550 2 150 600
PDPK1 copy number 6 5 3 2 2 2 x 2
PDK1 IHC score 5 5 6 5 4.8 2 2 2

Fig. S2

A


PTEN
Unknown
ERBB2
None
PTEN
None
PIK3CA
PIK3CA ERBB2
PIK3CA
Unknown
PTEN
PTEN+ ERBB2
PTEN ERBB2
PTEN
PDPK1
Chromosome 16 location

B


Single Channel
Dual Channel
Log10
NME3
NOXO1
CASKIN1
DCI
E4F1
ATP6VOC
CEMP1
PDPK1
TCEB2
SRRM2
ZSCAN10
1.6E+06
1.8E+06
2.0E+06
2.2E+06
2.4E+06
2.6E+06
2.8E+06
3.0E+06
3.2E+06

SNP Map
1.6E+06
1.8E+06
2.0E+06
2.2E+06
2.4E+06
2.6E+06
2.8E+06
3.0E+06
3.2E+06
Chromosome 16p13.3 coordinate

Fig. S3

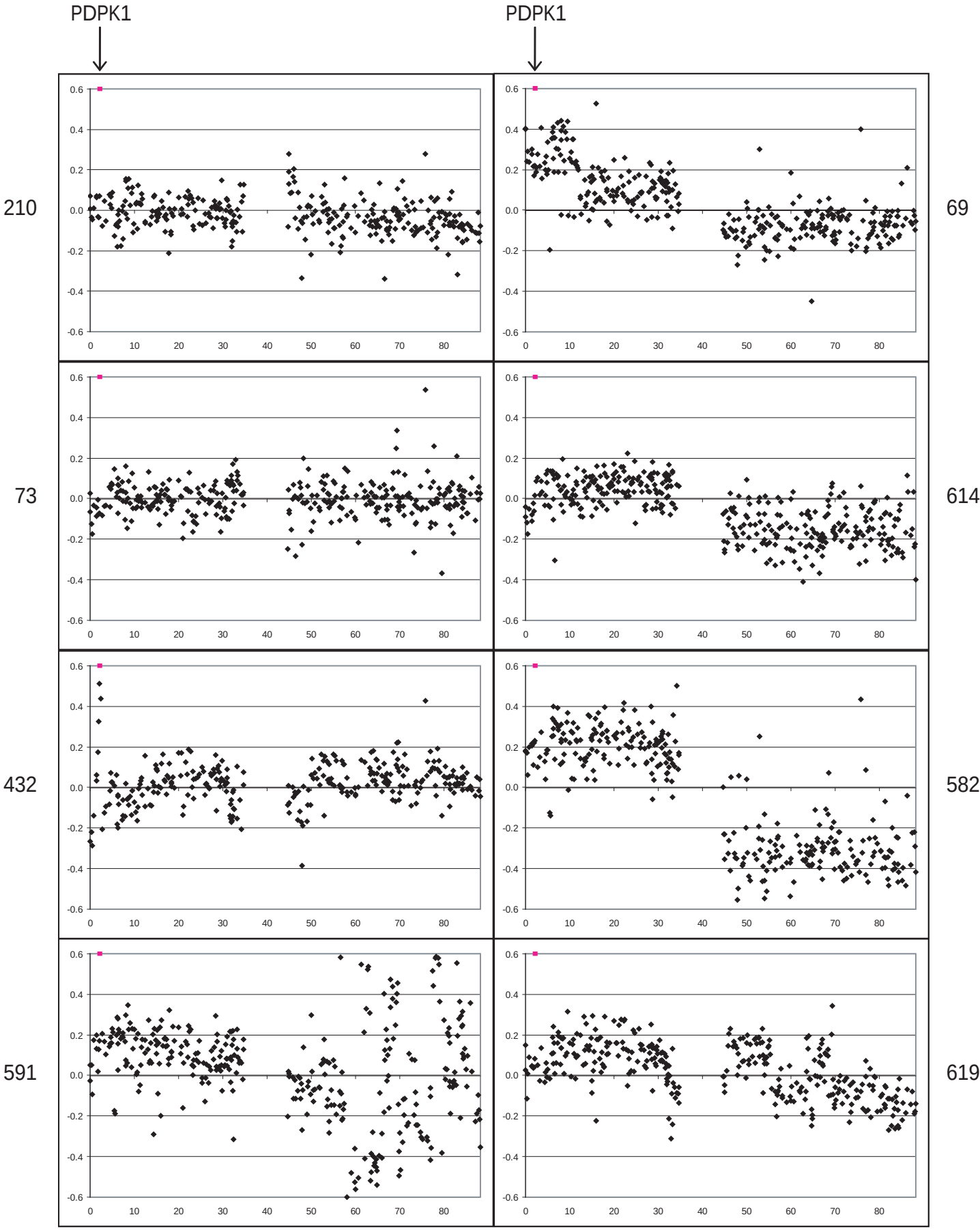


Fig. S4.

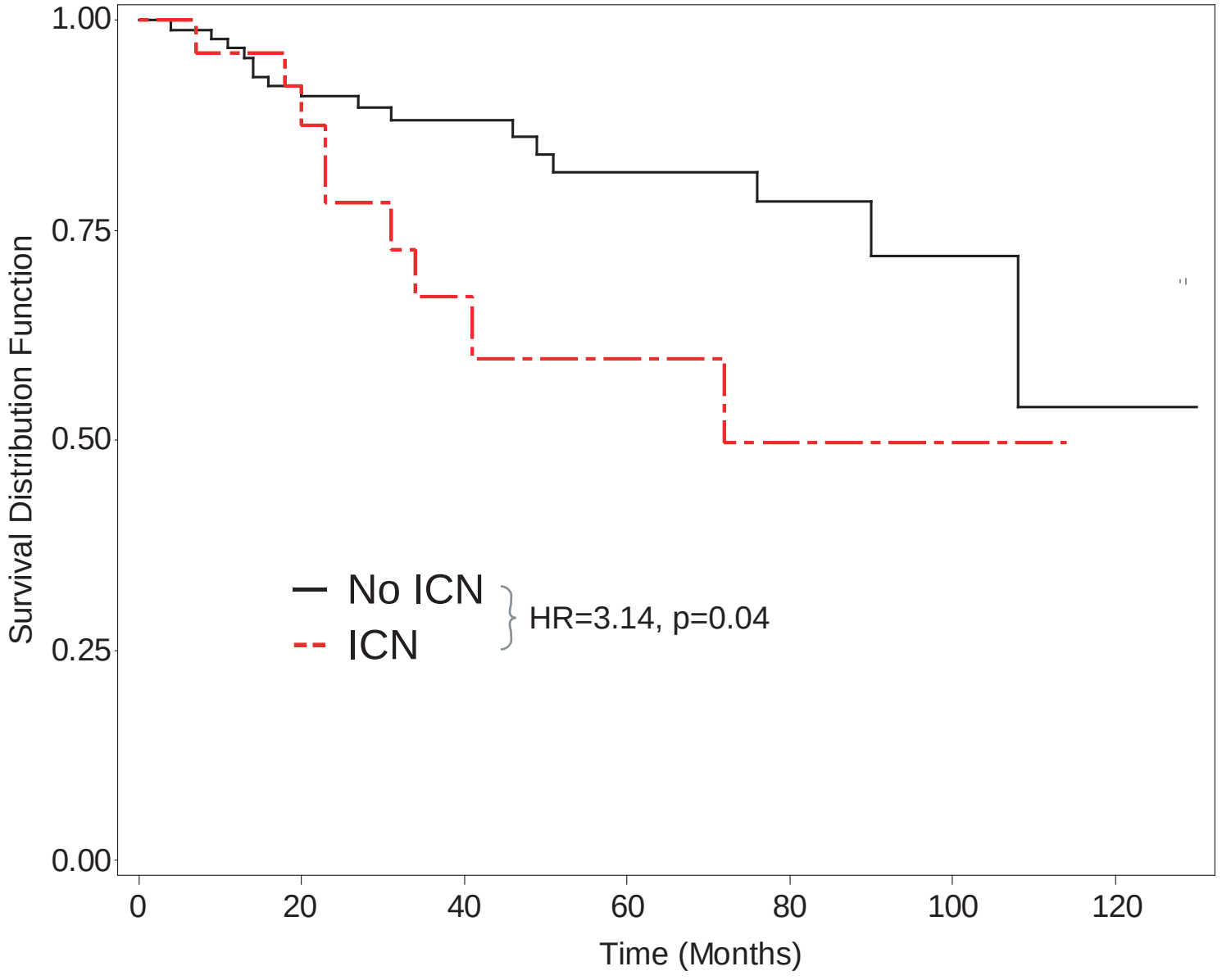
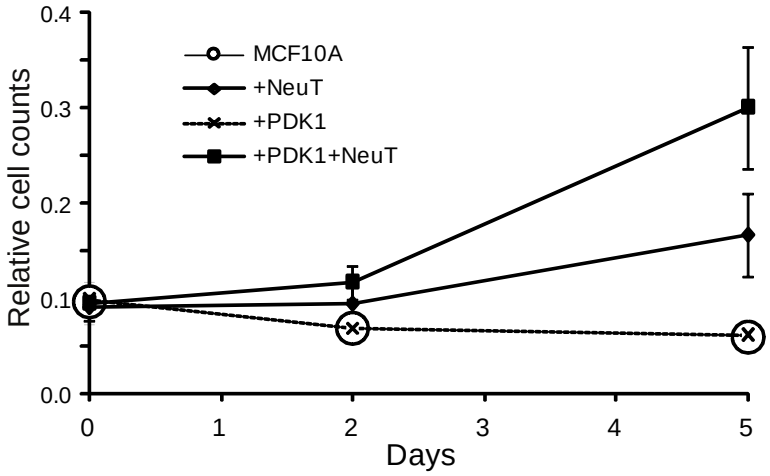
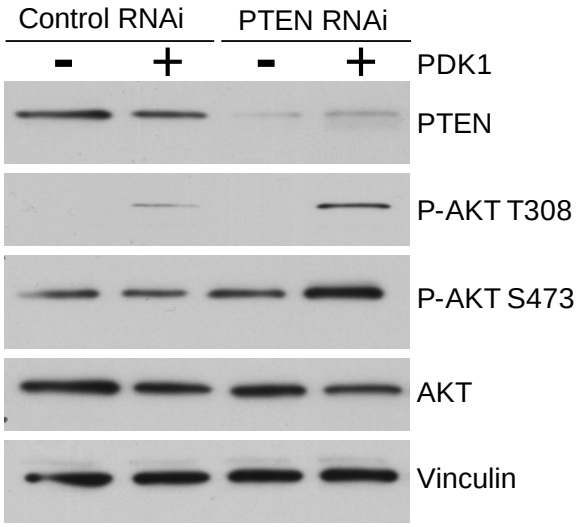

1.00
0.75
0.50
0.25
0.00
Survival Distribution Function
0
20
40
60
80
100
120
Time (Months)
No ICN
ICN
HR=3.14, p=0.04

Fig. S5

A


MCF10A
+NeuT
+PDK1
+PDK1+NeuT
Relative cell counts
0.0
0.1
0.2
0.3
0.4
0
1
2
3
4
5
Days

B


Control RNAi
PTEN RNAi
-
+
-
+
PDK1
PTEN
P-AKT T308
P-AKT S473
AKT
Vinculin

C

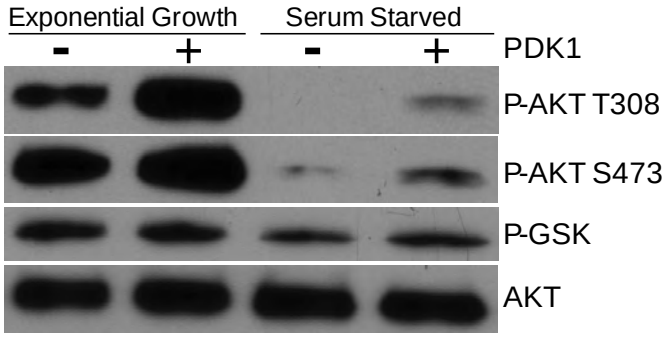
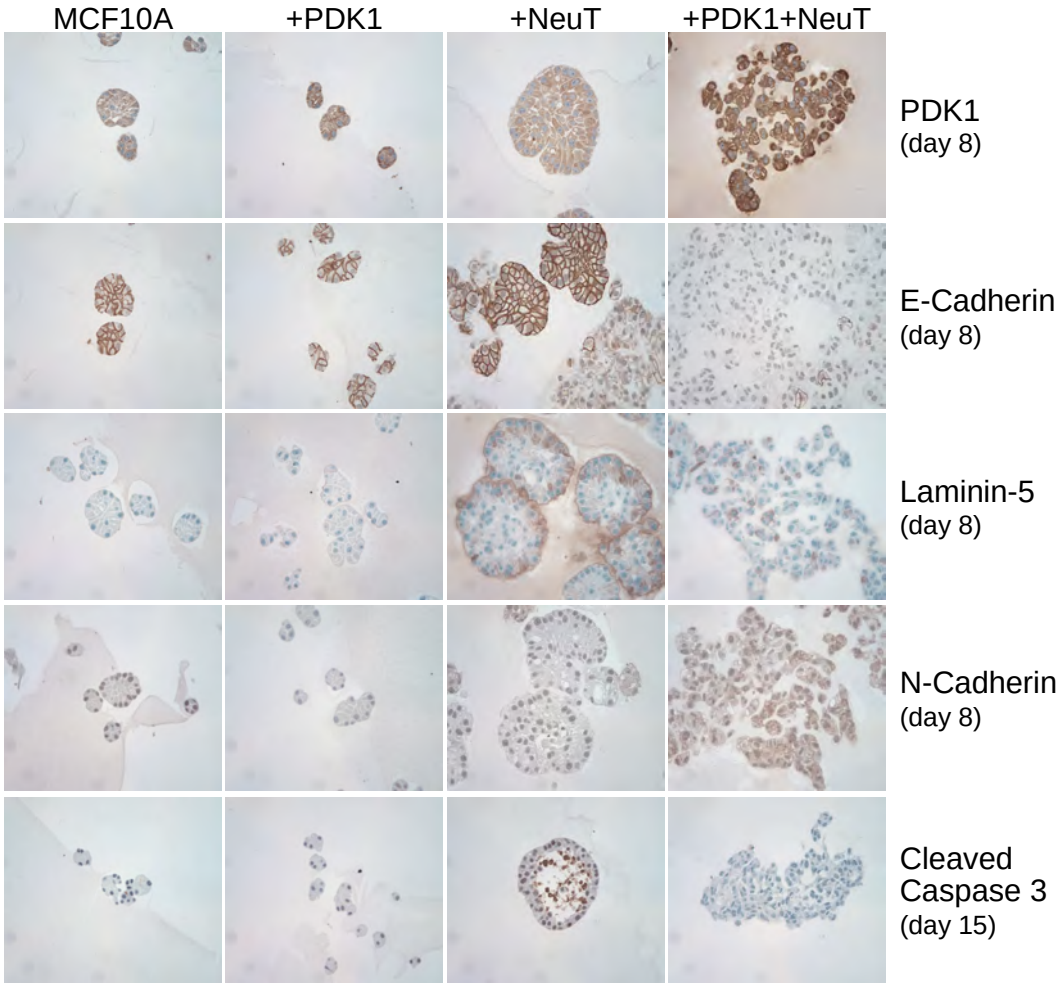

Exponential Growth
Serum Starved
-
+
-
+
PDK1
P-AKT T308
P-AKT S473
P-GSK
AKT

Fig. S6

A


MCF10A
+PDK1
+NeuT
+PDK1+NeuT
PDK1
(day 8)
E-Cadherin
(day 8)
Laminin-5
(day 8)
N-Cadherin
(day 8)
Cleaved
Caspase 3
(day 15)

B

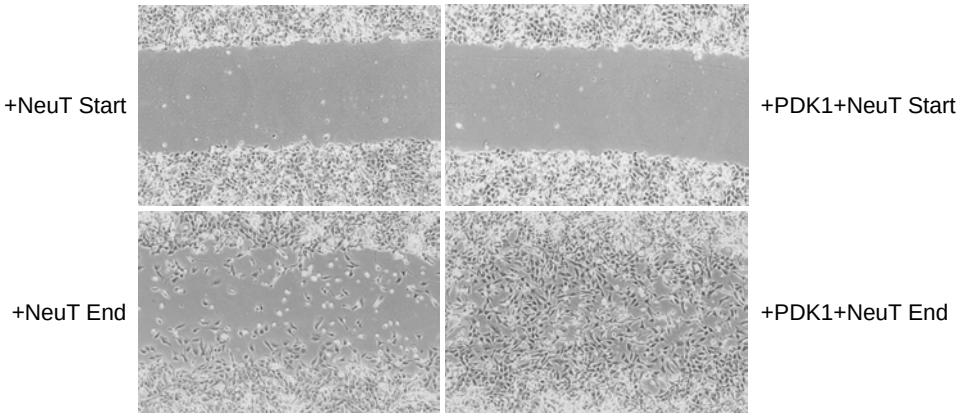
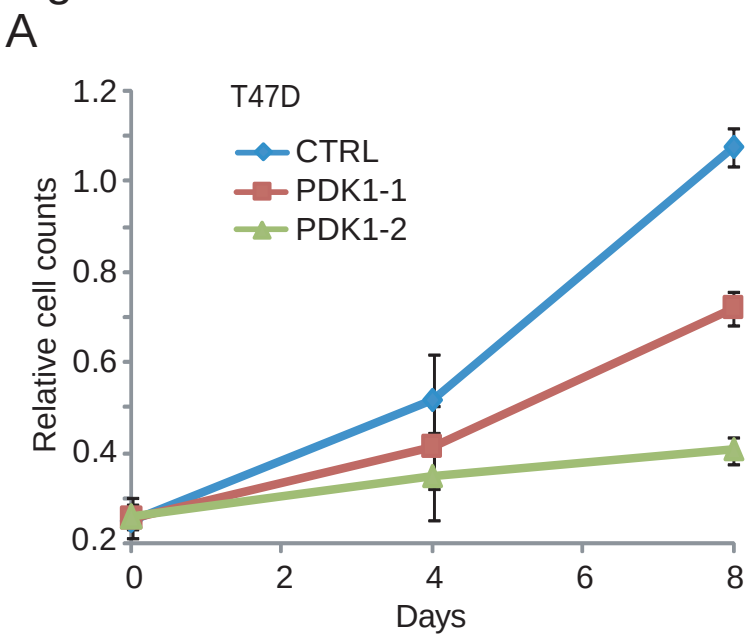
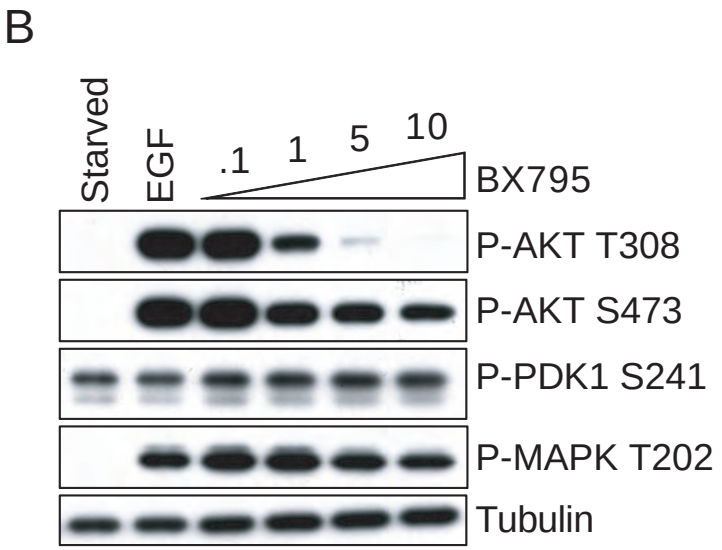
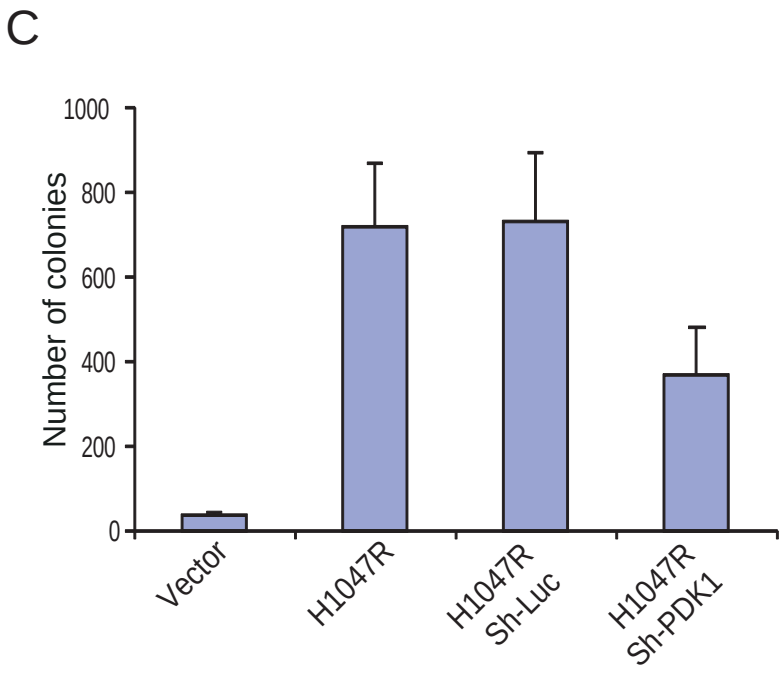
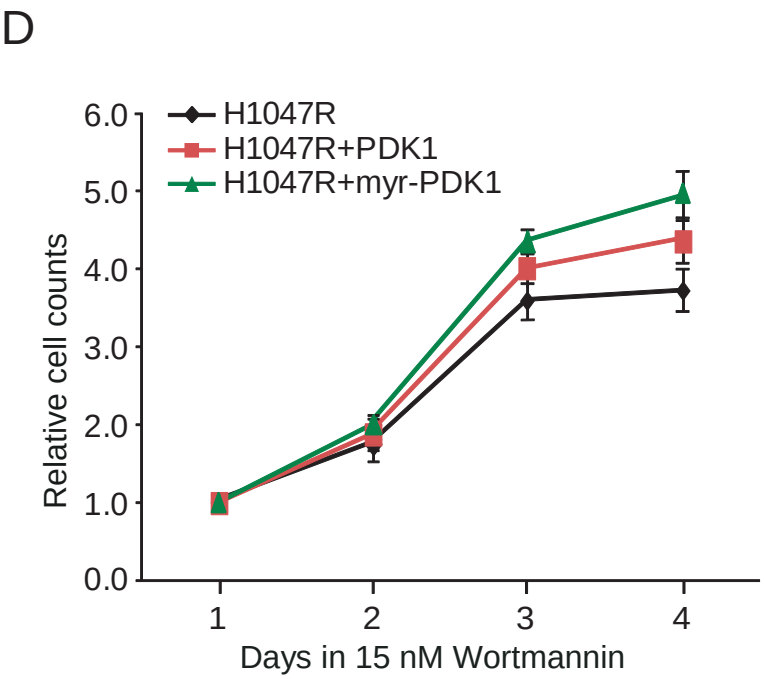

+NeuT Start
+PDK1+NeuT Start
+NeuT End
+PDK1+NeuT End

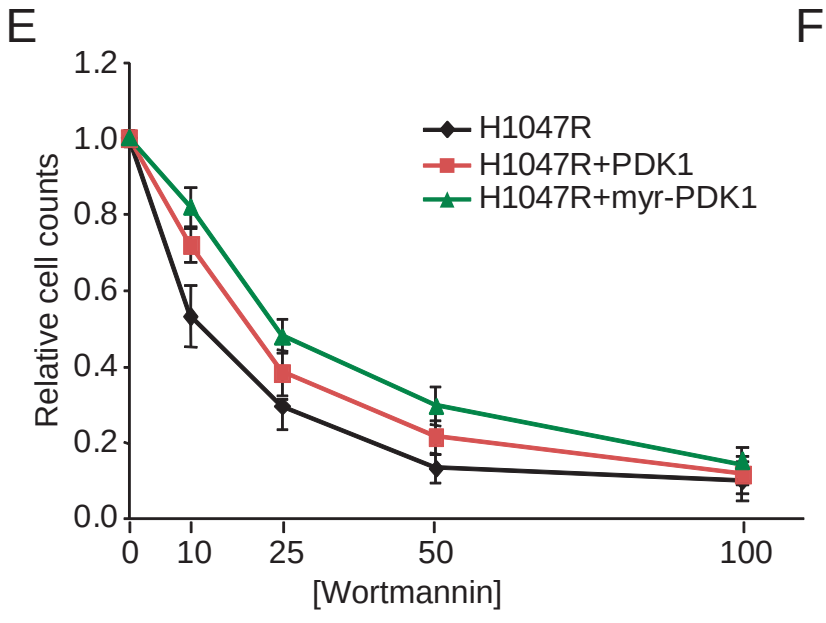
Fig. S7

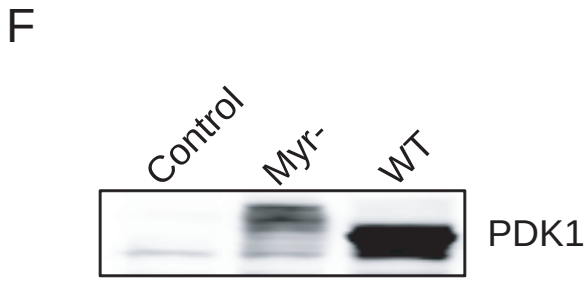

A
T47D
CTRL
PDK1-1
PDK1-2
Relative cell counts
Days


B
Starved
EGF
.1
1
5
10
BX795
P-AKT T308
P-AKT S473
P-PDK1 S241
P-MAPK T202
Tubulin


C
Number of colonies
Vector
H1047R
H1047R Sh-Luc
H1047R Sh-PDK1


D
H1047R
H1047R+PDK1
H1047R+myr-PDK1
Relative cell counts
Days in 15 nM Wortmannin


E
H1047R
H1047R+PDK1
H1047R+myr-PDK1
Relative cell counts
[Wortmannin]


F
Control
Myr-
WT
PDK1