AKT1^{low} quiescent cancer cells promote solid tumor growth

Cleidson P. Alves[1,2,*], Ipsita Dey-Guha[1,2,*], Sheheryar Kabraji[1,2], Albert C. Yeh[1,2], Nilesh P. Talele[1,2], Xavier Solé[1,2], Joeeta Chowdhury[1,2], Mari Mino-Kenudson[1,2], Massimo Loda[1,3], Dennis Sgroi[1,2], Anne-Lise Borresen-Dale[4,5], Hege Russnes[4,5], Kenneth N. Ross[1,2], Sridhar Ramaswamy[1,2,6,7,8,**]

AFFILIATIONS

[1] Massachusetts General Hospital Cancer Center, Boston, Massachusetts.

[2] Harvard Medical School, Boston, Massachusetts.

[3] Dana-Farber Cancer Institute, Boston, Massachusetts.

[4] Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway.

[5] Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway.

[6] Broad Institute of Harvard & MIT, Cambridge, Massachusetts.

[7] Harvard Stem Cell Institute, Cambridge, Massachusetts.

[8] Harvard-Ludwig Center for Cancer Research, Boston, Massachusetts.

*Correspondence to: Sridhar Ramaswamy, Massachusetts General Hospital Cancer Center, 185 Cambridge Street, Boston, MA 02114. E-mail: sridhar@mgh.harvard.edu

This PDF file includes:

Materials and Methods

Materials and Methods

Cell lines

HCT116 colon, MCF7 breast, MDA-MB-231 breast, A375 melanoma, and PC9 lung were purchased from ATCC, were they were validated. HCT116 AKT1/2-/- was purchased from Horizon Discovery, where it was validated. MCF7, MDA-MB-231 and A375 were maintained in DMEM, 10% FCS, 40 mMol glutamine, 100 U/mL penicillin, and 100 mg/mL streptomycin; HCT116 and HCT116 AKT1/2-/- in McCoy's 5α medium supplemented with 10% FCS, 100 U/mL penicillin, and 100 mg/mL streptomycin; PC9 in RPMI, 25% glucose, 1% sodium pyruvate, 100 U/mL penicillin, and 100 mg/mL streptomycin. All the cells were grown at 37°C and 5% CO_2.

DNA constructs and viral infection

The double-strand DNA sequence of AKT1 (NM_001014431.1) were synthesized and cloned into pLVX-One by GenScrip. The AKT1 sequence was then amplified by PCR and cloned into plasmid pLenti-GFP-Puro or pTRIPZ between the AgeI and MluI sites:

5'-ACCGGTTTTTGACCTCCATAGAAG;

5'-ACGCGTTAAACAGAATTAATTCCAAACTC).

Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent) to produce the E17K mutation:

5'-CCAGGTCTTGATGTATTTCCCTCGTTTGTGCAGCC;

5-GGCTGCACAAACGAGGGAAATACATCAAGACCTGG).

The MGH DNA Core Facility performed sequencing verification of the constructs.

Virus carrying the desired fusion gene was produced by transfecting HEK-293T cells with target vector pTRIPZ-AKT1-WT or pTRIPZ-AKT1-E17K and packaging vectors using the Mirus TransIT-293 transfection reagent and established protocols. Virus was collected 72 hours after transfection and filtered in 0.44μm filter. Infection was performed in 6 well plates by adding 1 mL of pooled virus, and 1μL 1,000× polybrene per well. A media change was performed the following day and cells were allowed to grow to confluence before splitting into a 10 cm dish and selection with 2μmol/L puromycin. After selection, cells were kept in 1μmol/L puromycin. The expression was induced using 2 μg/mL doxycycline for 72 hours.

Western blots

We used standard protocols for SDS-PAGE electrophoresis and used the following primary antibody: AKT1, phosphor-AKT1 (S473), S6, phosphor-S6 (S235) from Cell Signaling and GAPDH (Sigma).

Flow cytometry

Cells were fixed with cold methanol for 30 minutes at -20°C followed by PBS wash. AKT1Antibody incubation was performed in PBS containing 10% FBS for blocking. After 3 hours, cells were washed 3x with PBS and incubated with NucBlue Fixed Cell ReadyProbes Reagent (Invitrogen) for DNA content. Flow cytometry analysis was performed in a Becton Dickinson FACSAria II. Akt1 Alexa Fluor647 Conjugate was used (Cell Signaling).

In vivo xenograft tumors

A total of $5x10^5$ cells were injected subcutaneously into the flanks of 5- to 6-week-old, female immunocompromised NU/NU mice (Charles River Laboratories), and the growing tumors were measured by calliper. Mice were treated with TS2/16 antibody (18 mg/kg IP, every week for 5 weeks) or Paclitaxel (20 mg/kg IP, every week for five weeks). The TS2/16 hybridoma was culture and antibody production was made by the Monoclonal Antibody Core at Dana-Farber Cancer institute. For induction of AKT1-WT and AKT1-E17K, mice were given water containing 2ug/mL of doxycycline two days after cell injection.

Cells and tumor immunohistochemistry

Cells were grown directly on collagen IV–coated coverslips (Sigma). Cells were fixed in 3.7% formaldehyde, permeabilized using 0.1% Triton X-100, and treated with 0.1% SDS. They were blocked in 1% BSA and then incubated with primary antibody diluted in blocking solution, washed, and incubated with the respective secondary antibody. Cells were mounted using mounting media containing DAPI (Vector Laboratories). All secondary antibodies were Alexa Fluor conjugates (488, 555, 568, 633, and 647; Invitrogen). For immunohistochemistry, 5 mm sections of formalin-fixed paraffin-embedded (FFPE) tissues were dewaxed with xylene and rehydrated. Antigen retrieval was achieved by microwaving in unmasking solution (Vector Laboratories). For immunohistochemistry, 5μm sections of formalin-fixed paraffin-embedded (FFPE) tissues were dewaxed with xylene and rehydrated.

Antigen retrieval was achieved by microwaving in unmasking solution (Vector Laboratories). After washing, sections were blocked in 5% FCS and immunoperoxidase analysis was performed on tumors for MKI67 and cleaved Caspase3 (Abcam). The slides were counterstained lightly with hematoxylin for viewing negatively stained cells. H&E slides were used to assess the morphological integrity and geographical variation in morphology of the tissue samples. The slides were analyzed under the LEICA DC500 microscope.

AKT^{low} cells were identified using the previously validated markers H3K9me2, Hes1 (Abcam), MCM2 and AKT1 (Cell Signaling), as described in Dey-Guha *et. al.*, 2015. Cells were considered AKT^{low} when presented the $H3K9me2^{low}$, $MCM2^{low}$, $AKT1^{low}$ and $HES1^{high}$ profile.

For labeling $AKT1^{low}$ cancer cells in human primary tumor samples, we used a multi-step tyramide-amplified immunofluorescence (TSA-IF) protocol as previously described (Feng et al., 2015). Briefly, four micron-thick formalin-fixed paraffin-embedded tumor sections were de-paraffinized and rehydrated, followed by target antigen retrieval via a single microwave step and blocking with serum-free protein block (DAKO X0909). Each labeling cycle consisted of application of a primary antibody, followed by a secondary antibody conjugated to horse radish peroxidase (HRP), and tyramide-conjugated fluorophore (FITC, CY3, CY5, TSA Plus System, Perkin Elmer, Inc.) with intervening wash steps. The slides were sequentially incubated (and TSA-IF-labelled) with the antibody against H3K9me2 (ABCAM AB1220), pan-AKT (Cell Signaling 4961), HES1 (Millipore AB5702), and cytokeratin (Dako M0821) (where needed for additional tumor cell confirmation). Primary antibodies were omitted to generate a negative control for each staining cycle. TSA reagents were obtained from Perkin Elmer, Inc. Images were acquired on Nikon Ti confocal microscope at x60. $AKT1^{low}$ cancer cells were identified based on the following immunofluorescence pattern: DAPI+ / $AKT1^{low}$ / $H3K9me2^{low}$ / $HES1^{high}$ in a known area of tumor tissue (confirmed by comparison with a sequential hematoxylin and eosin-stained tumor section). 10 fields were randomly imaged and the number of $AKT1^{low}$ cancer cells per field was counted by a single observer using ImageJ. Mean $AKT1^{low}$ cancer cell count per tumor section was determined as the percentage of $AKT1^{low}$ cancer cells per total cancer cells in each image, averaged across the 10 images acquired per slide.

Feng, Z., et al., Multispectral imaging of formalin-fixed tissue predicts ability to generate tumor-infiltrating lymphocytes from melanoma. J Immunother Cancer, 2015. 3: p. 47.

Senescence detection

Tumors from three each wild type and E17K mice of 5 different cell lines were frozen in liquid nitrogen, and mounted in OCT. Thin sections (5 μm) were cut, mounted onto glass slides, fixed in 1% formalin in PBS for 1 min at room temperature, washed in PBS, immersed overnight in SA-3-Gal staining solution then viewed under bright field at 10-20X.

In vitro assays

Proliferation: After three days of doxycycline treatment, cells were plated in 96 well plates and the number of cells was estimated in different time points using the CellTiter 96 AQ_{ueous} One Solution Cell Proliferation Assay (Promega). Cells were analyzed for growth under different stress conditions including growth in 1% FBS supplemented media, hypoxia (4% oxygen), and paclitaxel treatment.

Migration and invasion: Migration was assayed using Boyden chamber (Corning Costar) containing a polycarbonate membrane filter (6.5 mm diameter, 8 μm pore size). Invasion was performed in matrigel-coated chamber. Cells were starved for 8 hours and then plate at the upper chamber in media contained 1% FBS, and the lower chamber contained 10% FBS (chemoattractant). Cells were incubated for 28 hours at 37°C in 5% CO_2. Non-migrated cells were scraped off the upper surface of the membrane with a cotton swab. Migrated cells remaining on the bottom surface were counted after staining with DAPI.

Soft agar: Colony formation in agar was performed as described by Borowicz *et. al*. 2014.

Colony formation: cells were seeded at a density of 500 cells per well in six-well plates and incubated for 15 days. Colonies were fixed and stained with Coomassie blue and counted.

Spheroid generation

25,000 cells were seeded per well in ultra low attachment clear bottom 96 well (cat.#4520 Corning, NY). The plates were incubated at standard cell culture conditions of 37^{o}C and 5% CO_2. Uniformity of spheroid size and compactness was assessed microscopically after 24 hours. Integrin β1 activation antibody TS2/16 (final concentration: 10 μg/ml) and Paclitaxel

(final concentration: 0.5 μM) were added to wells at day 8, and incubated for 3 days before imaging. Drug/antibody response analyses were carried out using stably transfected AKT1-GFP cell lines. Similarly, tumor spheroid growth curve analysis was carried out using stably transfected AKT-WT inducible vector and AKT-E17K inducible vector for 12 days. Growth curves were characterized using standard AlamarBlue® Cell Viability Reagent (cat# DAL1025, ThermoFisher, NY).

Spheroid staining and imaging

Prior to imaging, spheroids were stained overnight with counter stain Hoechst 33342 (final concentration: 5 μg/ml) and dead cell marker-propidium iodide (final concentration: 1 μg/ml). Spheroids were imaged using Eclipse Ti A1R-A1 confocal microscope (Nikon, USA).

*** Computational methods**

TCGA survival analysis

TCGA data for somatic mutations from whole exome sequencing data and the associated clinical information for those samples was downloaded from the TCGA data matrix access portal (http://cancergenome.nih.gov/). TCGA somatic mutation data was downloaded during April 2015 and TCGA clinical data was downloaded during December 2015. Follow-up clinical data files were merged with the original clinical data file to ensure that the most up-to-date patient follow-up information was used for survival analysis. The publicly available level 2 maf files for the TCGA whole exome sequencing data were reformatted into a matrix of Hugo gene symbols by sample barcode with each element of the matrix containing information about any mutations found in that gene and sample. When TCGA whole exome sequencing data was available from multiple platforms for a cancer disease type, the data from the maf files was merged into a single matrix. Using the TCGA whole exome sequencing data, mutation status for AKT1 in each tumor sample was determined as AKT1 WT, AKT1 E17K mutation, other AKT1 missense mutation, silent AKT1 mutation, deletion within AKT1, or AKT1 nonsense mutation.

The Kaplan-Meier plots and the log-rank p-value for TCGA BRCA of Supplemental Figure XX were generated using R with the 916 TCGA BRCA primary tumors that had whole

exome sequencing data, clinical information with at least 6 months of follow-up information or a death event, and were either AKT1 E17K mutant or AKT1 WT. AKT1 WT samples were sub-categorized according to ER+ status using the 'er_status_by_ihc' information in the clinical data tables. Kaplan-Meier plots used overall survival with death from any cause as the end point and patients still alive at last follow-up were censored at last follow-up time. The sheet labeled 'TCGA BRCA AKT1 Mut. w Clinical' of the supplemental excel file contains the clinical information for the 1002 TCGA BRCA samples with whole exome sequencing data. Patient data from three additional cases of primary breast cancer with AKT1 E17K mutations was available and these samples were combined with the 916 TCGA BRCA samples to produce the Kaplan-Meier plots and log-rank p-values of Figure 4b. The sheet labeled 'Combined TCGA MGH BRCA' of the supplemental excel file contains the clinical information for the 919 samples in the combined TCGA MGH breast cancer sample set made up of samples with at least 6 months of follow-up or a death event and AKT1 status as either WT or E17K mutant.

The Kaplan-Meier plots and the log-rank p-value for Combined TCGA data of Figure 4c were generated using clinical information and AKT1 somatic mutation information from the 7898 samples with the 31 cancer disease types in TCGA that have whole exome sequencing data. Summaries of the sample numbers and AKT1 mutation distributions for the cancer types are listed in the sheet labeled 'TCGA Combined Summary' of the supplemental excel file. Detailed information about the AKT1 mutations and sample names for the combined TCGA Kaplan-Meier plots are contained in the sheet labeled 'TCGA Combined AKT1 Mutations' of the supplemental excel file. Detailed clinical information for each of cancer disease types in the TCGA combined dataset can be found in the sheets labeled 'TCGA XXXX AKT1 Mut. w Clinical' of the supplemental excel file where XXXX is the TCGA dataset abbreviations for one of the 31 cancer disease types.

Association of AKT1 E17K mutations with clinical parameters in TCGA ER+ BRCA

R was used to test the significance of associations between clinical parameters and AKT1 E17K mutations in TCGA ER+ breast cancer (BRCA) data. The 726 ER+ TCGA BRCA samples (labeled as 'Positive' in the 'er_status_by_ihc' in the clinical data tables) with AKT1 status as either AKT1 wild-type (WT) or AKT1 E17K mutant were compared to clinical parameters by producing confusion matrices and performing significance tests on the

confusion matrices. Significance was evaluated using a Fisher exact test for binary clinical parameters (such as whether the patient received neoadjuvant treatment), a Chi-Square test for multi-way comparisons (such as menopause status), and a proportional trend test for trends (such as age at diagnosis). Raw p-values from the significance tests were corrected for multiple hypothesis testing using the Benjamini-Hochberg (BH) correction (in R using the multtest package) to find the BH false discovery rates (FDR). Results of testing for associations between clinical parameters and AKT1 E17K mutations in TCGA ER+ BRCA data are shown in the sheet labeled 'ER+ TCGA BRCA Clinical Assoc.' of the supplemental excel file. Only two clinical parameters had an association with AKT1 E17K mutations with BH FDR values less than 0.25: testing 'ajcc_nodes_pathologic_pn' found AKT1 E17K mutant samples more likely to be N0 and testing 'pr_status_by_ihc' found AKT1 E17K mutant samples slightly more likely to be PR positive.

Association of AKT1 E17K mutations with other somatic mutations in TCGA ER+ BRCA

The significance of associations between somatic mutations from whole exome sequencing data in genes other than AKT1 and AKT1 E17K mutations in TCGA ER+ breast cancer (BRCA) data was tested using a Chi-square test in R. The 726 ER+ TCGA BRCA samples (labeled as 'Positive' in the 'er_status_by_ihc' in the clinical data tables) with AKT1 status as either AKT1 wild-type (WT) or AKT1 E17K mutant were compared to all other genes that had at least 10 samples with non-silent mutations. There were 913 genes with at least 10 ER+ BRCA samples with non-silent mutations. Raw p-values from the significance tests were corrected for multiple hypothesis testing using the Benjamini-Hochberg (BH) correction (in R using the multtest package) to find the BH false discovery rates (FDR). Results of testing for associations between somatic mutations and AKT1 E17K mutations in TCGA ER+ BRCA data are shown in the sheet labeled 'ER+ TCGA BRCA Mutation Assoc.' of the supplemental excel file. Twelve genes with somatic mutations had significant p-values but none were significant after using BH FDR to correct for multiple hypothesis testing.

Association of AKT1 E17K mutations with RPPA protein expression in TCGA ER+ BRCA

R was used to test the significance of associations between protein expression measured by reverse phase protein arrays (RPPA) and AKT1 E17K mutations in TCGA ER+ breast cancer (BRCA) data. There were 577 TCGA ER+ BRCA samples that were in both the set of 937 samples with RPPA data and in the set of 726 ER+ BRCA samples (labeled as 'Positive' in the 'er_status_by_ihc' in the clinical data tables) with somatic mutation information (14 are AKT1 E17K mutant and 563 are AKT1 WT). Associations between AKT1 E17K mutations and RPPA protein expression were tested using a t-test and raw p-values were corrected for multiple hypothesis testing using the Benjamini-Hochberg (BH) correction (in R using the multtest package) to find the BH false discovery rates (FDR). Results of testing for associations between RPPA protein expression and AKT1 E17K mutations in TCGA ER+ BRCA data are shown in the sheet labeled 'ER+ TCGA BRCA RPPA Assoc.' of the supplemental excel file. Fifteen RPPA probes had significant p-values and 10 had BH FDR values less than 0.25. Among those probes most highly associated AKT1 E17K mutations are probes for c-Myc, Akt pS473, and Akt pT308.

Association of AKT1 E17K mutations with RNASeqV2 gene expression in TCGA ER+ BRCA

R was used to test the significance of associations between gene mRNA expression measured by RNASeqV2 and AKT1 E17K mutations in TCGA ER+ breast cancer (BRCA) data. TCGA BRCA had RNASeqV2 data from 1093 primary tumor samples and 718 of those samples were ER+ BRCA samples (labeled as 'Positive' in the 'er_status_by_ihc' in the clinical data tables) and either AKT1 wild type (WT) or E17K mutant (18 AKT1 E17K mutant vs. 700 AKT1 WT). Significance of association of AKT1 E17K mutations with gene expression was evaluated using log2 RNASeqV2 data after filtering out low expression and low fold change genes (prior to taking the log2 the data was thresholded to a minimum value of 1 and genes were filtered to leave only those with a minimum maximum fold change across all the samples of 3 fold, a minimum absolute difference of 5 across all the samples, and minimum standard deviation of 0.5) which reduced the number of genes tested to 18809 from 20531. Because of the severe imbalance between the number of AKT1 E17K mutant and AKT1 WT samples, a permutation-based 'bagging' approach was used to calculate the p-value for a t-test between AKT1 E17K mutant and AKT1 WT samples. This 'bagging' approach randomly selected 36 AKT1 WT samples to compare to the 18 AKT1 E17K

mutant samples for 2500 iterations and used the mean t-test p-value from those 2500 iterations as the p-value for the gene. Raw p-values from the significance tests were corrected for multiple hypothesis testing using the Benjamini-Hochberg (BH) correction (in R using the multtest package) to find the BH false discovery rates (FDR). Results from testing for associations between RNASeqV2 gene expression and AKT1 E17K mutations in TCGA ER+ BRCA data are shown in the sheet labeled 'ER+ TCGA BRCA Bagged T-test' of the supplemental excel file. After the bagged t-test, zero genes were significant with FDR BH<0.25 and 1236 genes had p<0.05 (252 genes had p<0.01 and 19 genes had p<0.001). The sheet in the excel file has a plot of fold change versus the bagged t-test p-value and a heat map of the genes with a p-value less than 0.01.

Association of AKT1 E17K mutations with pathways using RNASeqV2 gene expression in GSEA in TCGA ER+ BRCA

Gene Set Enrichment Analysis (GSEA) (http://www.broadinstitute.org/gsea/index.jsp) was used to assess association of gene expression to molecular pathways and other published signatures with AKT1 E17K mutations in TCGA ER+ breast cancer (BRCA) data. TCGA BRCA had 1093 samples from primary tumors with RNASeqV2 data and 718 of these samples were ER+ BRCA samples (labeled as 'Positive' in the 'er_status_by_ihc' in the clinical data tables) and AKT1 wild type (WT) or E17K mutant (18 AKT1 E17K mutant vs. 700 AKT1 WT). GSEA used signal-to-noise of AKT1 wild type (WT) vs. E17K mutant to evaluate enrichment of the c2 gene sets (curated gene sets) from v4 of molecular signature database (MSigDB, www.broadinstitute.org/msigdb) with p-values determined by 2500 permutations of the phenotype. A summary of the GSEA results from evaluating AKT1 E17K mutations in TCGA ER+ BRCA data are shown in the sheet labeled 'ER+ TCGA BRCA GSEA Summary' of the supplemental excel file. After permutation testing, 129/3782 gene sets had p<0.05 but none had a BH FDR p-value of FWER p-value less than 0.25.

Association of AKT1 E17K mutations with gene copy number in TCGA ER+ BRCA

The significance of associations between gene copy numbers from GISTIC2 and AKT1 E17K mutations in TCGA ER+ breast cancer (BRCA) data was tested using R. TCGA BRCA had 1080 samples from primary tumors with copy number (from Affymetrix genome-wide SNP array 6.0) data and 717 of these samples were ER+ BRCA samples (labeled as

'Positive' in the 'er_status_by_ihc' in the clinical data tables) and either AKT1 wild type (WT) or E17K mutant (18 AKT1 E17K mutant vs. 699 AKT1 WT). GISTIC2 processed copy number data was downloaded from firebrowse.org (downloaded April 2015 as gdac.broadinstitute.org_BRCA-TP.CopyNumber_Gistic2.Level_4.2015040200.0.0.tar).

Associations between AKT1 E17K mutations and copy numbers from GISTIC2 all_data_by_genes.txt were tested using a t-test and raw p-values were corrected for multiple hypothesis testing using the Benjamini-Hochberg (BH) correction (in R using the multtest package) to find the BH false discovery rates (FDR). Results of testing for associations between GISTIC2 gene copy numbers and AKT1 E17K mutations in TCGA ER+ BRCA data are shown in the sheet labeled 'ER+ TCGA BRCA GISTIC2 CN Assoc.' of the supplemental excel file. 5993 genes had p-values less than 0.05 and 2678 had BH FDR values less than 0.05. Many of these significant genes were a result of small average copy number changes driven by a few samples across large regions. Those cases with the difference being driven by focal changes were almost exclusively driven by focal changes in a relatively small number of AKT1 WT samples.