nature research

Corresponding author(s): Keisuke Kataoka

Last updated by author(s): Feb 5, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ImageQuant LAS 4000 control software version 1.2 Molecular Operating Environment (MOE) program version 2016.8 GROMACS version 2019.1 Compass software version 4.1.0 VICTOR Nivo Control Software version 3.0.2
Data analysis	LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver) Variant Effect Predictor version 95.3 vcf2maf version 1.6.16 SAMtools version 1.4.1 Minimap2 version 2.14 NanoQC version 0.9.0 Integrative Genomics Viewer (IGV) version 2.4.10 Python version 2.7.15 R version 3.6.0 Genomon version 2.6.2 Custom script for the permutation test (https://github.com/nccmo/Permutation-test) CisChecker (https://github.com/nccmo/CisChecker/) DNVChecker (https://github.com/nccmo/DNVChecker) pyMOL version 2.2.0 Gnuplot version 4.6.2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The findings of this study are supported by data that are available from public online repositories and data that are publicly available upon request of the data provider. See Methods for detail. Long-read WGS data of cell lines have been deposited in the European Genome-phenome Archive (EGA) under accession EGAS00001003763. Data generated in the current study are available as Source Data files that accompany Fig.4 and Extended Data Figs.10-11.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Sample size was determined by the availability of sequencing data. We enrolled all samples from public datasets.
Data exclusions	Samples with missing data, derived from xenografts, and duplicated in the same patient were excluded from the analysis. Hypermutator samples [with ≥500 and ≥20 coding mutations per exome (for the discovery cohort) and targeted region (for the additional cohort), respectively] were also removed from the analysis to minimize the effect of confounding passenger mutations. The specific exclusion criteria were not pre-established, although the cutoff values were determined based on previous publications (example: PMID: 29056346). No data were excluded for in vivo and in vitro experiments.
Replication	No experimental replication was performed for sequencing experiments. Molecular dynamics simulations were performed three times with different velocities. For in vivo and in vitro experiments, all attempts at replication were successful with biological replicates performed on separate cohorts of animals/cells.
Randomization	Samples were assigned to each group based on the number of mutations (samples with no mutation, single mutation, or multiple mutations in the gene of interest). Therefore, randomization was not required.
Blinding	Blinding was not relevant to the study, as there was no control and treatment arms involved.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Rabbit anti-PI3 Kinase p110α (C73F8) (Cell Signaling Technology, #4249, dilution 1:1000 [IB] or 1:50 [CI]) Rabbit anti-β-Actin (13E5) (Cell Signaling Technology, #4970, dilution 1:2000 [IB] or 1:50 [CI]) Rabbit anti-Akt (pan) (C67E7) (Cell Signaling Technology, #4691, dilution 1:1000 [IB] or 1:50 [CI]) Rabbit anti-phospho-Akt (Ser473) (D9E) (Cell Signaling Technology, #4060, dilution 1:2000 [IB] or 1:50 [CI]) Rabbit anti-mTOR (7C10) (Cell Signaling Technology, #2983, dilution 1:50 [CI])

Rabbit anti-phospho-mTOR (Ser2448) (D9C2) (Cell Signaling Technology, #5536, dilution 1:50 [CI])
Rabbit anti-p70 S6 Kinase (Cell Signaling Technology, #9202, dilution 1:50 [CI])
Rabbit anti-phospho-p70 S6 Kinase (Thr389) (Cell Signaling Technology, #9205, dilution 1:50 [CI])
Rabbit anti-PRAS40 (Cell Signaling Technology, #2610, dilution 1:50 [CI])
Rabbit anti-phospho-PRAS40 (Thr246) (C77D7) (Cell Signaling Technology, #2997, dilution 1:50 [CI])
Rabbit anti-GSK-3β (27C10) (Cell Signaling Technology, #9315, dilution 1:50 [CI])
Rabbit anti-Phospho-GSK-3β (Ser9) (5B3) (Cell Signaling Technology, #9323, dilution 1:50 [CI])
IB, immunoblot; CI, capillary-based immunoassay.

Validation

Antibodies with prior manufacturer validation were used. Validation statements at manufacturer's website were as follows:
Rabbit anti-PI3 Kinase p110α (C73F8) (Cell Signaling Technology, #4249): Western blot analysis of extracts from HeLa cells and neonatal mouse brain using PI3 Kinase p110α (C73F8) Rabbit mAb. Currently over 224 citations.
Rabbit anti-β-Actin (13E5) (Cell Signaling Technology, #4970): Western blot analysis of cell extracts from various cell lines (NIH/3T3, HeLa, PAE, A431) using beta-Actin (13E5) Rabbit mAb. Currently over 1,861 citations.
Rabbit anti-Akt (pan) (C67E7) (Cell Signaling Technology, #4691): Western blot analysis of recombinant Akt1, Akt2 and Akt3 proteins, and extracts from various cell lines (HeLa, NIH/3T3, C6, COS), using Akt (pan) (C67E7) Rabbit mAb. Currently over 1,851 citations.
Rabbit anti-phospho-Akt (Ser473) (D9E) (Cell Signaling Technology, #4060): Western blot analysis of extracts from PC-3 cells, untreated or LY294002/wortmannin-treated, and NIH/3T3 cells, serum-starved or PDGF-treated, using Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb (upper) or Akt (pan) (C67E7) Rabbit mAb #4691 (lower). Currently over 3,926 citations.
Rabbit anti-mTOR (7C10) (Cell Signaling Technology, #2983): Western blot analysis of extracts from 293, A431, COS, C6, and C2C12 cells, using mTOR (7C10) Rabbit mAb. Currently over 970 citations.
Rabbit anti-phospho-mTOR (Ser2448) (D9C2) (Cell Signaling Technology, #5536): Western blot analysis of extracts from serum-starved NIH/3T3 cells, untreated or insulin-treated (150 nM, 5 minutes), alone or in combination with λ-phosphatase, using Phospho-mTOR (Ser2448) (D9C2) XP® Rabbit mAb (upper) or mTOR (7C10) Rabbit mAb #2983. Currently over 650 citations.
Rabbit anti-p70 S6 Kinase (Cell Signaling Technology, #9202): Western blot analysis of extracts from HeLa, NIH-3T3, PC12 and COS-7 cells using p70 S6 Kinase Antibody. Currently over 939 citations.
Rabbit anti-phospho-p70 S6 Kinase (Thr389) (Cell Signaling Technology, #9205): Western blot analysis of HeLa, COS, C6 and 3T3 cells, serum-starved overnight, then treated with insulin, l-phosphatase or 20% serum as indicated. Upper panel probed with Phospho-p70 S6 Kinase (Thr389) Antibody #9205; lower panel probed with p70 S6 Kinase Antibody #9202. Currently over 904 citations.
Rabbit anti-PRAS40 (Cell Signaling Technology, #2610): Western blot analysis of extracts from various cell types (MCF-7, 293, HeLa, A204, RD, 3T3, RAW, KNRK, NBTII, and COS7) using PRAS40 Antibody. Currently over 40 citations.
Rabbit anti-phospho-PRAS40 (Thr246) (C77D7) (Cell Signaling Technology, #2997): Western blot analysis of extracts from serum starved H3255, Mkn45 and NIH/3T3 cells, untreated or treated with either Gefitinib (1 μM, 3 hours), Su11274 (1 μM, 3 hours) or insulin (150 nM, 15 minutes), using Phospho-PRAS40 (Thr246) (C77D7) Rabbit mAb (upper) or PRAS40 (D23C7) Rabbit mAb #2691 (lower). Currently over 119 citations.
Rabbit anti-GSK-3β (27C10) (Cell Signaling Technology, #9315): Western blot analysis of extracts from HeLa, NIH/3T3, COS, C6 and 293 cells using GSK-3β (27C10) Rabbit mAb. Currently over 666 citations.
Rabbit anti-Phospho-GSK-3β (Ser9) (5B3) (Cell Signaling Technology, #9323): Western blot analysis of extracts from NIH/3T3 cells, λ-phosphatase- or PDGF-treated, using Phospho-GSK-3β (Ser9) (5B3) Rabbit mAb (upper) or GSK-3β (27C10) Rabbit mAb #9315 (lower). Currently over 294 citations.
See https://www.citeab.com/ for details of relevant citations.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Ba/F3-CL1, WEHI-3, and 293T cell lines were obtained from the RIKEN Cell Bank, HEC-1 cell line from the JCRB Cell Bank, and MCF10A, BT-20, NCI-H1048, SUP-T1, and HRT-18 cell lines from ATCC. Lenti-X 293T cells were purchased from TaKaRa.

Authentication

All cell lines were authenticated by the providers using karyotype, isoenzmes, and/or microsatellite profiling (short tandem repeat or simple sequence length polymorphism).

Mycoplasma contamination

We confirmed that all cell lines were negative for mycoplasma contamination using MycoAlertTM Mycoplasma Detection Kit (Lonza, LT07-318).

Commonly misidentified lines
(See ICLAC register)

According to ICLAR register, cross-contamination was reported in BT-20 in 1976 (Reference PubMed ID: 6451928), although authentic stocks apparently do exist. We used the BT-20 cell line authenticated by ATCC using short tandem repeat (STR) profiling analysis. This cell line was selected because of the limited number of available cell lines harboring multiple mutations in PIK3CA.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Female BALB/c-nu/nu mice (6 weeks old)

Wild animals

The study did not involve wild animals.

Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight	All mouse experiments were approved by the Animal Ethics Committee of the National Cancer Center and strictly adhered to its guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.