nature research

Corresponding author(s): Sarah-Maria Fendt

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Mass spectrometry data were collected using MSD Chemstation Data Analysis (vE.02.0.2.1431) or Agilent Mass Hunter (vB.0802 Build 8.2.8260.0) followed by an in house developed Matlab script.
Data analysis	Mass spectrometry data were analyzed using MSD Chemstation Data Analysis (vE.02.0.2.1431) or Agilent Mass Hunter (vB.0802 Build 8.2.8260.0) followed by an in house developed Matlab script. Metastatic area and metastases number were quantified by Zen Blue software (2011). Collagen intensity was quantified with Imaris Image Analysis Software 8 (Bitplane). GNU Imagin Manipulation Program (GIMP 2.10.8) was used to manually stitch all picro-sirius images belonging to the same metastasis together. The composite images were analyzed with Image J 1.45. Microsoft excel 2013 was used for data output. Statistical data analysis was performed using GraphPad Prism 7 software.

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 upon request. 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 authors declare that all data supporting the findings of this study are available within the article, its extended data files, or from the corresponding author upon reasonable request

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	In vitro sample sizes were based on previous similar studies that have given statistical results. For in vivo experiments, sample size was determined using power calculations with B=0.8 and P<0.05, based on preliminary data and respects the limited use of animals in line with the 3 R system: Replacement, Reduction, Refinement.
Data exclusions	Detection of outliers was performed using Grubb's test in Graphpad.
Replication	All experiments were performed at least in triplicates. All attempts at replications were successful.
Randomization	Mice were randomized into control and treatment groups.
Blinding	Mice were given a number prior to data collection and analysis. Data was collected and analyzed, and subsequently grouped in the corresponding cohorts for statistical analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	Collagen staining was performed with either anti-Collagen I (Abcam, Ab34710) or anti-Collagen III (Abcam, Ab7778). Alexa Fluor 555 (Life technologies, A31272) was used as conjugated secondary antibody. Western blot analysis was performed with the following antibodies: MCT2 (LabNed 0315312), GPT2 (Santa Cruz, 398383), P5CS (Santa Cruz, 515443), GDH (Abcam, 153973), P4HA1 (Abcam, 59497), B-Actin (Sigma, A5441) and ERK1/2 (Cell signaling, 4695S).
Validation	Antibody were used as recommended in the respective data-sheets. MCT2 antibody was validated based on positive and negative control cell lines.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	MCF10A, MCF7, HCC70 and 4T1 cell lines were purchased from ATCC. The EMT6.5 cell line was kindly provided by professor Robin Anderson (Peter MacCallum cancer center). Fibroblasts were kindly provided by Prof. Ludo Van Den Bosch (VIB). Myofibroblasts were kindly provided by Prof. Akira Orimo (Juntendo University).
Authentication	MCF10A, MCF7 and HCC70 cell lines were validated by DNA fingerprinting.
Mycoplasma contamination	All cel llines were confirmed to be mycoplasma free by Mycoalert detection kit (Lonza).
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

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

Laboratory animals	Six weeks old female BALB/C mice were injected with mouse breast cancer cells through either mammary fat pad injection or intravenously. Mice were sacrificed after 2-4 weeks.
Wild animals	No wild animals were used.
Field-collected samples	No field-collected samples were used.