nature research

Corresponding author(s): Bernd Bodenmiller and Carlos Caldas

Last updated by author(s): 19Dec2019

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	Fluidigm CyTOF Software
Data analysis	In house image preprocessing scripts: https://github.com/BodenmillerGroup/imctools Ilastik 1.2.0 Cellprofiler version 2.2 Stata SE Version 14.2 R Version 3.4.3 R packages: clusterprofiler 3.10.1 flowcore 1.48.1 flowsom 1.14.1 reactomepa 1.26.0 glmnet 2.0_16 rphenograph 0.99.1 rtsne 0.15

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

Imaging mass cytometry data, including cell masks and processed single cell data, have been deposited to the Image Data Resource (https://idr.openmicroscopy.org/) under the accession code idr0076.

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	Sample size was determined by a combination of sample availability and feasibility of assay throughput. Sample size was considered sufficient to represent the breadth of disease (all subtypes) and for survival analyses since standard prognostic factors showed expected effects.
Data exclusions	Samples comprising fewer than 100 segmented cells were excluded from tumour-level analyses owing to insufficient representation of tumour characteristics, as detailed in the manuscript. Criteria were not pre-established.
Replication	Analysis of observational population data; further replication analyses not conducted.
Randomization	Randomisation not relevant to the analysis of observational data.
Blinding	All data acquisition was conducted blinded to associated clinical/molecular data.

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

Antibodies are detailed in Supplementary Table 1 which includes the following information:
Antigen Antibody Clone Concentration Host Species Clonality Catalogue Number Lot number Vendor
Estrogen Receptor Alpha SP1 8 ug/mL Rabbit Monoclonal M3011* 131011 Spring Bioscience currently Thermo Fisher
Estrogen Receptor Alpha EP1 8 ug/mL Rabbit Monoclonal IR084* AC-0015EU Epitomics, currently Agilent
Histone H3 D1H2 8 ug/mL Rabbit Monoclonal 4499 12 Cell Signaling Technology
Tri-Methyl-Histone H3 (Lys27) C36B11 8 ug/mL Rabbit Monoclonal 9733 11 Cell Signaling Technology
Cytokeratin 5 EP1601Y 6 ug/mL Rabbit Monoclonal ab52635 GR299320-I AbCam
Fibronectin 10/Fibronectin 4 ug/mL Mouse Monoclonal 610078 6251888 BD Biosciences
Cytokeratin 19 Troma-III 6 ug/mL Mouse Monoclonal Troma-III 3/7/13 Dev Studies Hybridoma Bank
Cytokeratin 8/18 C51 4 ug/mL Mouse Monoclonal 4546 2 Cell Signaling Technology
Twist1 Polyclonal 8 ug/mL Rabbit Polyclonal ABD29 2754704 Merck Millipore
CD68 KP1 8 ug/mL Mouse Monoclonal 14-0688-80 E15987-105 eBioscience (ThermoFisher)

Cytokeratin 14 Polyclonal 1 ug/mL Rabbit Polyclonal PA5-16722 SE2391461H eBioscience (ThermoFisher)
SMA 1A4 2 ug/mL Mouse Monoclonal ab7817 033M4768 AbCam
Vimentin D21H3 1 ug/mL Rabbit Monoclonal 5741 3 Cell Signaling Technology
c-Myc 9E10 8 ug/mL Mouse Monoclonal 626802 B201394 Biolegend
c-erbB-2 3B5 8 ug/mL Mouse Monoclonal 554299 6182799 BD Biosciences
CD3ε D7A6E 8 ug/mL Rabbit Monoclonal 85061 2 Cell Signaling Technology
Histone H3 Phospho (Ser28) HTA28 6 ug/mL Rat Monoclonal 641002 B200946 Biolegend
ERK1/2 (pT202/pY204) 20A 8 ug/mL Rabbit Monoclonal 612359 2153932 BD Biosciences
Slug 666633 8 ug/mL Rabbit Monoclonal MAB7408 CFZB021603A R&D Systems
Unconjugated Goat Anti-Rabbit IgG Antibody Polyclonal 4 ug/mL Goat Polyclonal AI-1000 ZA0911 Vector laboratories
Progesterone Receptor A/B EP2 8 ug/mL Rabbit Monoclonal AC-0028EU* N/A Epitomics, currently Abcam
Progesterone Receptor A/B SP2 8 ug/mL Rabbit Monoclonal M3024 131017 Spring Bioscience, currently Abcam
p53 7F5 8 ug/mL Rabbit Monoclonal 2527 Custom purification Cell Signaling Technology
CD44 Polyclonal 6 ug/mL Sheep Polyclonal AF3660 CFOE0216011 R&D Systems
EpCAM (CD326) 9C4 8 ug/mL Mouse Monoclonal 324208 B190626 Biolegend
CD45 2B11 8 ug/mL Mouse Monoclonal 14-9457-80 4298126 eBioscience (ThermoFisher)
GATA3 L50-823 8 ug/mL Mouse Monoclonal 558686 6132744 BD Biosciences
CD20 L26 8 ug/mL Mouse Monoclonal 14-0202-80 4285442 eBioscience (ThermoFisher)
Non-phospho (Active) β-Catenin (Ser33/37/Thr41) (D13A1) D13A1 8 ug/mL Rabbit Monoclonal 8814 2 Cell Signaling Technology
Carbonic Anhydrase IX Polyclonal 8 ug/mL Goat Polyclonal AF2188 VNQ0215032 R&D Systems
E-Cadherin / P-Cadherin 36/E-Cadherin 8 ug/mL Mouse Monoclonal 610181 5274693 BD Biosciences
Ki-67 8D5 3 ug/mL Mouse Monoclonal 9449 2 Cell Signaling Technology
EGF Receptor D38B1 8 ug/mL Rabbit Monoclonal 4267 13 Cell Signaling Technology
Phospho-S6 Ribosomal Protein (Ser235/236) (D57.2.2E) D57.2.2E 6 ug/mL Rabbit Monoclonal 4858 10 Cell Signaling Technology
Sox9 Polyclonal 6 ug/mL Rabbit Polyclonal AB5535 2894178 Merck Millipore
von Willebrand Factor Polyclonal 6 ug/mL Rabbit Polyclonal AB7356 2700933 Merck Millipore
CD31 HC1/6 8 ug/mL Mouse Monoclonal M0823 2557591 Merck Millipore
mTOR, phospho (Ser2448) 49F9 8 ug/mL Rabbit Monoclonal 2976 5536BF Cell Signaling Technology
Cytokeratin 7 RCK105 8 ug/mL Mouse Monoclonal 550507 6083676 BD Biosciences
pan Cytokeratin AE1 0.5 ug/mL Mouse Monoclonal MAB1612 2341224 Merck Millipore
Keratin Epithelial AE3 0.5 ug/mL Mouse Monoclonal MAB1611 2607604 Merck Millipore
Cleaved PARP (Asp214) F21-852 8 ug/mL Mouse Monoclonal 558576 2150663 BD Biosciences
Cleaved Caspase3 C92-605 8 ug/mL Rabbit Monoclonal 559565 559565 BD Biosciences

Validation

Expected antibody staining patterns were confirmed in a variety of tissues including invasive breast carcinoma. Details of vendor validation and use in previous publications are provided as hyperlinks to https://antibodyregistry.org pages in Supplementary Table 1.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Women diagnosed with primary breast cancer between 1985 and 2005.
Recruitment	Selected as a retrospective representative case series; no specific exclusion criteria or patient self-selection were applied.
Ethics oversight	Addenbrooke's Hospital; Cambridge University Hospitals

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	The study is not a clinical trial.
Study protocol	The study did not involve a clinical intervention.
Data collection	Tertiary hospital setting; ongoing data collection/updates conducted from approximately 2007 to date.
Outcomes	No specific outcome included in study design.