nature research

Corresponding author(s): Bernd Bodenmiller

Last updated by author(s): Oct 24, 2019

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	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	All custom code used for this study is available in the linked github repository. The repository contains code written in Matlab version 2017b and R version 3.5. For segmentation and single-cell feature extraction CellProfiler version 2.1.1, Ilastik version 1.1.9 and histoCAT version 1.74 were used. CATALYST version 1.5.6 was used for spillover compensation. The PhenoGraph 2.0 implementation in histoCAT was used for single-cell clustering. The R package pvclust 2.0 was used for multiscale bootstrap resampling of the hierarchy subtrees during metaclustering, in order to assess the uncertainty of each subtree. For clustering on patients the cytofkit R implementation of PhenoGraph (version 1.10.0) was used. The R package entropy (version 1.2.1) was used for the calculation of both Shannon entropy and the Kullback-Leibler divergence. The Louvain community detection algorithm (C implementation by Lefebvre and Guillaume, version 0.2, wrapped by Matlab as used by the implementation of PhenoGraph 2.0 used by histoCAT/Cyt) was applied to identify highly interconnected spatial subunits in the tissue graph. Kaplan-Meier survival curves and coxph survival regression models were generated using the R package survival (version 2.42-4).

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 image, single-cell and patient data supporting the findings of this study were uploaded to figshare during resubmission (via the Manuscript Tracking System). All code that led to the results of this study will be available on https://github.com/BodenmillerGroup/SCPathology_publication upon publication.

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	In this investigative, first of its kind study, 281 breast cancer patients were analyzed. A second cohort containing 344 images from 71 patients was also studied and matched to the original cohort. As it was not possible to estimate the number of cellular phenotypes present before analysis we were not able to specify sample size.
Data exclusions	No data was excluded from this study. All samples that were successfully stained and imaged were used in our data analysis.
Replication	All cell types, patient groups and interactions were identified in multiple patient samples including multiple samples from the same patient. Further, a second patient cohort, acquired and processed at an independent institute was subject to identical analysis for the validation of initial classifications.
Randomization	In the first cohort from the University Hospital Basel, patients were not selected for any clinical or histological features, therefore patients in the cohort are representative of the population. In the second cohort, from the University Hospital Zurich, patients were selected to be equally divided between the three pathology grades of breast cancer and within each grade half did and half did not have metastasis.
Blinding	All samples were stained simultaneously. Image acquisition order was distributed spatially and blinded to clinical 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	All Antibody information is available in Supplemental Information Table #2
Validation	All antibodies were validated by immunofluorescence imaging prior to isotope-polymer conjugation. Antibodies were tested for cell type and inter-cell location specificity within positive control tissues including lymph nodes, breast tumors, and healthy breast tissue.

Human research participants

Policy information about studies involving human research participants

Population characteristics	All patient data is available as a .csv raw data file and summarized in Supplemental Information Table 1 and 9.
Recruitment	The specimens derived from patients diagnosed with primary breast cancer between 1991 and 2013 at the Institute of Pathology at the University Hospital Basel and between 2004 and 2013 at the Institute of Pathology and Molecular Pathology at the University Hospital Zurich.
Ethics oversight	Project to use samples from the University Hospital Basel were approved by the Ethikkommission Nordwest- und Zentralschweiz (EKNZ-2014-397) and for samples to use from the Institute of Pathology and Molecular Pathology, University Hospital Zurich were approved by the Ethikkommission Kanton Zürich (KEK-2012-553).

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	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	Histopathological data was obtained from the individual pathology reports
Outcomes	Clinical and survival data were extracted from the hospital database or from the patients' attending physician