nature portfolio

Corresponding author(s): SAMUEL APARICIO, JOAN BRUGGE, SOHRAB SHAH

Last updated by author(s): Oct 8, 2024

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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	No software used for data collection.
Data analysis	Single-cell pipeline for processing DLP+ data is available at https://github.com/shahcompbio/single_cell_pipeline. SIGNALS v.10.0 was used to call allele specific copy number Rv4.3 was used to generate all fugures Code to reproduce all the figures is available at https://github.com/marcjwilliams1/normal_brca_scdna

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw sequencing data is available from EGA under accession EGAS00001007716. Processed data including all single-cell copy number calls is available at 10.5281/zenodo.13645601.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Gender based analysis is not pertinent to this study. The samples were all obtained from biological female participants.
Reporting on race, ethnicity, or other socially relevant groupings	social groupings/demographics were not collected or analysed as part of this study.
Population characteristics	Subjects comprise biological females undergoing non-cancer surgery for cosmetic or risk reduction. Age (range 27-70), parity (Parous=22, Nulliparous=6) and menopausal status (Pre=15, Post surgical=8, Post non-surgical=2) are recorded. 4 donors had prior chemotherapy exposure due to previous cancer, 6 had prior history of cancer. 12 are BRCA1 carriers, 7 BRCA2 carriers.
Recruitment	Subjects undergoing reduction mammoplasty or non-cancer treatment risk reduction surgery were consented for participation in the study. Specimens were obtained from Brigham & Women's Hospital or Faulkner Hospital on the day of surgery. Inclusion was based on tissue availability and successful data generation, we do not believe these introduce any biases that would effect the results.
Ethics oversight	This study was reviewed by the Harvard Medical School Institutional Review Board (IRB) and deemed not human subjects research. Donors gave their informed consent to have their anonymized tissues used for scientific research purposes.

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

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 sample size calculation was performed for the number of donors included as this is an exploratory landscape study. Inclusion criteria were patients undergoing mastectomy for risk reduction or cosmetic reasons. Donors with current invasive cancer were not included. We ensured each sample was powered to detect CNAs at a baseline rate of above 1%. For this reason, samples with fewer than 300 cells were excluded from the study. This cutoff was based on requiring a 95% probability of sequencing at least 1 aneuploid cell if the baseline rate of aneuploidy was 1%.
Data exclusions	Filtering of low quality genomes was applied uniformly to all samples according to a procedure documented in Laks et al 2018 and this is described in the method s. S-phase cells were also identified and excluded from analysis as described in Laks et al 2019 and in the methods. This was applied uniformly to all samples.
Replication	Replication is not built into this survey sequencing study. Replication is not possible as tissue is scarce and only allows for running the scWGS assay once.
Randomization	Randomization was not applicable to this landscape survey. Randomization is not appropriate as this was an observational retrospective study from tissue collected over many years.
Blinding	All participants were de-identified. Blinding was not applied to knowledge of BRCA or WT genotype of the samples, however, no differences in tissue or sample processing were dependent on genotype. Data processing and analysis were applied uniformly without consideration of genotype.

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 and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Antibodies

Antibodies used

Alexa Fluor 647-conjugated anti-EpCAM (Biolegend 324212, Lot B347793)
PE-conjugated anti-CD49f (Biolegend 313612, Lot B346513)
FITC-conjugated anti-CD31 (Biolegend 303103, Lot B370631)
Alexa Fluor 488 anti-CD45 (Biolegend 304017, Lot B286002)

Validation

Links to biolegend product description pages provides technical details:
https://www.biolegend.com/en-ie/products/alexa-fluor-488-anti-human-cd45-antibody-2738
https://www.biolegend.com/en-ie/products/pe-anti-human-mouse-cd49f-antibody-4108
https://www.biolegend.com/en-ie/products/fitc-anti-human-cd31-antibody-881
https://www.biolegend.com/en-ie/products/alexa-fluor-488-anti-human-cd45-antibody-2738

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

184hTERT cell line was generated by us (SA) and is described in Burleigh et al.

Authentication

Identity of cells was confirmed by matching WGS

Mycoplasma contamination

Cell lines tested negative for mycoplasma.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used in the study

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.