nature portfolio

Corresponding author(s): Kiryl D. Piatkevich, Yi Zhu, Tiannan Guo

Last updated by author(s): Oct 2, 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	Images were acquired using Zeiss ZEN v3.1, NIS-Elements AR v5.42.01, and Leica LMD v8.4 software. MS data was acquired using Bruker OtofControl v6.2 and HyStar v5.1 for the timsTOF Pro, as well as Thermo Scientific Xcalibur v4.7.69.37 and Orbitrap Astral Tune Application v1.0.100.40 for the Orbitrap Astral MS.
Data analysis	MS data were analyzed using the FragPipe platform (version 18.0) with the MSFragger search engine (version 3.5) and the DIA-NN software (version 1.8.1).For all numerical comparisons, Welch's t-test was performed, and statistical visualizations were generated using PRISM (version 10.0.3) and R (version 4.4.0). MATLAB (version R2021a) was used for isotropic analysis, following a previously established protocol. Differentially expressed proteins (DEPs) were identified using the limma package in R, which fits linear models and incorporates the duplicateCorrelation function. "Region" was treated as the primary fixed effect without an intercept, and "Patient" was used as the blocking factor. Planned contrasts were defined for hypothesis testing between specific pairs of regions. An empirical Bayes moderated t-test, with FDR adjustment for multiple comparisons, was applied. For pairwise comparisons, contrast coefficients with an adjusted p-value < 0.05 and |logFC| > log2 were considered statistically significant, with the overall linear model also having an adjusted p-value < 0.05. The Pearson's correlation coefficient was calculated after excluding missing values. Hierarchical cluster analysis was conducted using Ward's method with Euclidean distance. The Uniform Manifold Approximation and Projection (UMAP) technique was applied for dimension reduction and visualization of high-dimensional data. The pvca R package (version 1.30.0) was used for Principal Variance Components Analysis (PVCA). The top-N principal components were selected based on a 90% explained variance threshold. "Patient," "slide," and "region" were included as random effects to account for variability. The magnitude of each effect was calculated as a proportion of total variance using a weighted average. Gene ID conversion was conducted using clusterProfiler (version 4.11.0). Gene Ontology (GO) term annotations were performed using org.Mm.eg.db (version 3.18.0), and redundant GO terms were manually simplified. Fast gene set enrichment analysis (GSEA) was conducted using clusterProfiler to identify Cellular Component Ontology terms ranging from 10 to 300. Reactome pathway-based analysis was conducted using the ReactomePA package (version 1.47.0) to identify gene sets ranging from 5 to 300. Gene sets with an adjusted p-value <

0.05, as reported by GSEA, were considered statistically significant.

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

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD041412 [https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD041412]. The raw images for Fig. 2a-d, Fig. 4c-e, and Fig. 5a, along with all isotropic analysis datasets used in this study, are available on Zenodo at https://doi.org/10.5281/zenodo.13848766. Source data are provided with this paper.

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	Tissue regions from both male and female patients were randomly selected from FFPE slides, with no analysis of sex or gender differences.
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable
Population characteristics	We included three colorectal tissue samples from patients with histologically confirmed cancer, representing various subtype regions, including normal, low-grade dysplasia, high-grade dysplasia, and carcinoma. No specific demographic factors, such as age, were analyzed.
Recruitment	Samples were retrospectively collected from archived FFPE material acquired at the Second Affiliated Hospital of Zhejiang University School of Medicine.
Ethics oversight	This study complies with all relevant regulations for human study participants and was conducted in accordance with the Declaration of Helsinki. The collection and use of human tissue were approved by the Institutional Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine and the Ethics Committee of Westlake University.

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 statistical methods were used to estimate sample size. Sample sizes for each experiment are provided in the figure legends, and they were selected to ensure they are sufficient for statistical comparisons between groups. The sample sizes were determined based on similar published studies (Nat Commun. https://doi.org/10.1038/s41467-022-34824-2).
Data exclusions	One file failed due to an MS bug during acquisition, resulting in a total of 131 files.
Replication	All attempts for replication were included in the manuscript.
Randomization	The experimental design for the colorectal cancer samples proteomics acquisition was randomized into six batches to reduce potential batch effects.
Blinding	Blinding was not applied in this study, as the experimental design did not necessitate it due to the specific objectives of the research.

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

1. Anti-ASGR1 (Proteintech, 11739-1-AP, 1:200 dilution for both pre- and post-expansion)
2. Goat anti-Rabbit IgG (Abcam, ab150086, 1:1000 dilution for pre-expansion and 1:200 for post-expansion)

Validation

All antibodies are commercially available and validated in the literature as cited on the manufacturer's websites as well as by the datasheets they provide. Antibody validation and validation criteria are available on the following websites:
1. Anti-ASGR1 (Proteintech, 11739-1-AP), https://www.ptglab.com/products/ASGR1-Antibody-11739-1-AP.htm
2. Goat anti-Rabbit IgG (Abcam, ab150086), https://www.abcam.com/en-us/products/secondary-antibodies/goat-rabbit-igg-h-l-alexa-fluor-555-preadsorbed-ab150086

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

Four male C57BL/6J wild-type mice (Mus musculus; four months old) and one female C57BL/6J wild-type mouse (Mus musculus; two months old) were used. All mice were obtained from the Laboratory Animal Resources Center of Westlake University and housed in controlled barrier facilities, with macroenvironmental temperatures maintained between 20 and 26°C and humidity levels between 40% and 70%. Food and water were provided ad libitum, and the mice were kept on a 12-hour light/dark cycle. Housing conditions were closely monitored and maintained.

Wild animals

N/A

Reporting on sex

Male and female mice were used in separate benchmarking experiments, without comparing sexes, so results are not disaggregated by sex.

Field-collected samples

N/A

Ethics oversight

All animal maintenance and experimental procedures were conducted according to the Westlake University Animal care guidelines, and all animal studies were approved by the Institutional Animal Care and Use Committee (IACUC) under protocol #22-083-GTN.

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

Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A