nature portfolio

Corresponding author(s): Carlos Bais

Last updated by author(s): 04/24/2025

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	For IMblaze370 and Guardant 360 data collection, no software was used. For Foundation Medicine core database, Data were collected by examining the Foundation Medicine database comprising patients who received comprehensive genomic profiling during routine clinical care.
Data analysis	1. For IMblaze370 data: Nonnegative Matrix Factorization was performed using the the package NMF. R package version 0.27, https://cran.r-project.org/package=NMF. The xCell enrichment was performed using the package xCell. R package version 1.1.0 (devtools::install_github ('dviraran/xCell')). GSVA scores were determined using the package GSVA. R package version 1.52.2(https://www.bioconductor.org/packages/release/bioc/html/GSVA.html). CMS subtypes were annotated 591 using the CMScaller version 0.99.2 (https://github.com/Lothelab/CMScaller). Statistical comparison between groups were performed using the package matrixTests. R package version 0.2.3 (https://cran.r-project.org/web/packages/matrixTests/index.html). Heatmaps were generated using the package ComplexHeatmap. R package version 2.20.0 (https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html). Volcano plots were generated using the package EnhancedVolcano. R package version 1.22.0. (https://bioconductor.org/packages/release/bioc/html/EnhancedVolcano.html). For developing a machine learning classifier, the randomforest algorithm was used using RandomForest R package version 4.6.14 (https://cran.r-project.org/web/packages/randomForest/index.html). For predictions, the NMF subtypes the caret R package was used version 6.0.94. 2. For Guardant 360 data, SAS 9.4 was used to perform the statistical analyses. R was used to generate visualizations (ComplexHeatMap). 3. For FoundationCore database, Statistics, and computation were carried out using fisher_exact functions within the SciPy package (version 1.2.1) in Python 2.7.18 (Python Software Foundation). Plotting was carried out using the ggpubr (version 0.6.0) and ggplot2 (version 3.4.4) packages in R 4.3.1 (R Foundation for Statistical Computing).

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

1. For data generated from IMbalze370 study: IMBLAZE 370 RNA-seq data will be made available indefinitely to qualified researchers at the European Genome-Phenome Archive under an accession number EGAS00001005952. To request access to such data, researchers can submit a data access request via the standard EGA process. Following publication, the data will be released upon application with necessary agreements to enforce terms such as security, patient privacy and consent of specified data use, consistent with evolving, applicable data protection laws. The general turnaround right now is 2-3 months to make data available to requesters and the data availability on the platform is 'indefinite' per our statement. The user, per the Data Use Agreement only has a year to use the data, after that it must be deleted.

2. For Guardnant 360 data: The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request.

3. For FoundationCore database: The sequencing data utilized for this study are derived from clinical samples. All the relevant data supporting the findings of this study are provided within the main manuscript and its supplementary files. Due to HIPAA requirements, we are not authorized to share individualized patient genomic data, which contains potentially identifying or sensitive patient information. Foundation Medicine is committed to collaborative data analysis, with well-established, and widely utilized mechanisms by which investigators can query our core genomic database of >700,000 de-identified sequenced cancers to obtain aggregated datasets. For more information and mechanisms of access, please contact the corresponding author(s) or the Foundation Medicine, Inc. Data Governance Council at data.governance.

4. For the PDX analysis, we used publicaly available data from GEO accession number GSE76402.

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	These variables were not considered in this study.
Reporting on race, ethnicity, or other socially relevant groupings	None of these socially relevant groupings were performed in this study.
Population characteristics	1. For IMblaze370 study:Please refer to Lancet Oncol, 2019; 20: 849-861 2. For Guardant360 data: All patients with colorectal cancer entered as cancer type on Guardant360 Test Requisition Form, with linked administrative claims data were included in the study. 3. For FoundationCore database: The study included a total of 36966 tumor DNA and 1367 cfDNA samples from patients with CRC who received comprehensive genomic profiling using during routine clinical care were included in this study.
Recruitment	1. For IMblaze 370 study: Please refer to Lancet Oncol, 2019; 20: 849-861 2. For Guardant 360 data: All patients with at least one Guardant360 are included in this HIPAA-compliant, fully de-identified dataset. Recruitment was not required. 3. For FoundationCore database: Patients with a diagnosis of CRC who underwent testing as part of routine clinical care and consented for research were included in this study. Genomic profiling is typically performed later in clinical care, so the data may be biased for later-stage cases.
Ethics oversight	1. For IMblaze370 study: The study followed Good Clinical Practice guidelines and the Declaration of Helsinki. Ethical approval was obtained from independent ethic review boards at each study sites. All participants gave informed consent for biomarker analyses described in the protocol and for publication of the results. T 2. For Guardant 360 data: Participants' consent was not necessary because deidentified research data sets generated by Guardant Health are approved by the Advarra Institutional Review Board with a waiver of consent. The GuardantINFORM database is a fully deidentified database that complies with sections 164.514(a)-(n)1ii of the U.S. Health Insurance Portability and Accountability Act regarding the determination and documentation of statistically deidentified data. 3. For FoundationCore database: Approval for this study, including a waiver of informed consent and a Health Insurance Portability and Accountability Act waiver of authorization, was obtained from the WCG Institutional Review Board (Protocol No. 20152817). The Institutional Review Board granted a waiver of informed consent under 45 CFR § 46.116 based on review and determination that this research meets the following requirements: (i) the research involves no more than minimal risk to the subjects; (ii) the research could not practicably be carried out without the requested waiver; (iii) the waiver will not adversely affect the rights and welfare of the subjects. No compensation was provided as samples were obtained during routine clinical care.

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

Field-specific reporting

Life sciences study design

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

Sample size	1. For IMblaze370 study: Out of 363 enrolled patients, 254 patients, who received either prior anti-EGFR, or anti-VEGF agents and with both tumor NGS data at diagnosis, and ctDNA data at baseline post prior lines of treatment were included in this study 2. For Guardant 360 data: There is no a priori hypothesis therefore formal sample size and power calculation was not required. 3. For FoundationCore Database: All available samples were included and no sample size calculation was performed. This study included a total of 36966 patients with CRC who received tissue biopsy-based comprehensive genomic profiling (CGP) and 1367 patients with CRC who recieved liquid biopsy during routine clinical care.
Data exclusions	1. For IMblaze370 study: Samples without prior treatment history or without NGS data at diagnosis or ctDNA data at baseline were excluded from this analysis 2. For Guardant 360 data:There were no data exclusion criteria, however, due to the nature of administrative claims data used as clinical data, uninsured patients were not included in the parent dataset. 3. For FoundationCore Database:Samples with a failed clinical report were excluded.
Replication	1. For IMblaze 370, Guardant 360 and Foundation Core Database: Because the data are from large cohort of patient samples, the analysis can not be replicated 2. For CRC organoid and drug sensitivity data: Data are represented of three independent experiments
Randomization	1. For IMblaze370 study: It was a randomized trial, however, randomization was not relevant to this analysis 2. For Guardant 360 data: Randomization was not relevant for this study since this is an observational descriptive analysis of cancer biomarker dynamics among colorectal cancer patients receiving two different types of therapy. 3. For FoundationCore Database:This is an observational study, so no randomization was performed.
Blinding	1. For IMblaze370 study: This study was open label and investigators and participants were not masked to treatment allocation. 2. For Guardant 360 data: Blinding was not relevant for this study since this is a retrospective observational descriptive study. 3. For FoundationCore Database: No participants were allocated to groups.

Behavioural & social sciences study design

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

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

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

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

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

Phospho-Erk (Thr202/Tyr204), Cell Signaling Technology, Cat. #9101, 1:1000
Erk1, Cell Signaling Technology, Cat. #9102, 1:1000
Phospho-S6 Ribosomal Protein (Ser235/236), Cell Signaling Technology, Cat. #2211, 1:1000
S6 Ribosomal Protein, Cell Signaling Technology, Cat. #2217, 1:1000
Phospho-Akt (Ser473), Cell Signaling Technology, Cat. #9271, 1:1000
Akt (pan), Cell Signaling Technology, Cat. #9272, 1:1000
Tubulin, Cell Signaling Technology, Cat. #2146, 1:2500

Validation

Data provided in this manuscript: phospho-ERK (Thr202/Tyr204), total ERK, phospho-S6 (Ser 235/236), total S6, Phospho-AKT (Ser 473), AKT, Tubulin: Western blots of KRAS mutant organoid. All listed antibodies have validation provided in manufacture website

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

CellLine: LS-174T
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CL-188
Gender: Female

CellLine: SW 480
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CCL-228
Gender: Male

CellLine: SW 620
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CCL-227
Gender: Male

CellLine: SK-CO-1
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: HTB-39
Gender: Male

CellLine: ATRFLOX
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CRL-2780
Gender: Male

CellLine: HCT 116
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CCL-247
Gender: Male

CellLine: LoVo
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CCL-229
Gender: Male

CellLine: DLD-1
Species: Homo sapiens
Tissue: Colon
Vendor: ATCC
Cat#: CCL-221
Gender: Male

CellLine: CL-11
Species: Homo sapiens
Tissue: Colon
Vendor: DSMZ
Cat#: ACC467
Gender: Male

Authentication

Short Tandem Repeat (STR) Profiling: STR profiles are determined for each line using the Promega PowerPlex 16 System. This is performed once and compared to external STR profiles of cell lines (when available) to determine cell line ancestry. Loci analyzed: Detection of sixteen loci (fifteen STR loci and Amelogenin for gender identification), including D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, AMEL, vWA, D8S1179 and TPOX.

SNP fingerprinting: SNP profiles are performed each time new stocks are expanded for cryopreservation. Cell line identity is verified by high-throughput SNP profiling using Fluidigm multiplexed assays. SNPs were selected based on minor allele frequency and presence on commercial genotyping platforms. SNP profiles are compared to SNP calls from available internal and external data (when available) to determine or confirm ancestry. In cases where data is unavailable or cell line ancestry is questionable, DNA or cell lines are re-purchased to perform profiling to confirm cell line ancestry. SNPs analyzed: rs11746396, rs16928965, rs2172614, rs10050093, rs10828176, rs16888998, rs16999576, rs1912640, rs2355988, rs3125842, rs10018359, rs10410468, rs10834627, rs11083145, rs11100847, rs11638893, rs12537, rs1956898, rs2069492, rs10740186, rs12486048, rs13032222, rs1635191, rs17174920, rs2590442, rs2714679, rs2928432, rs2999156, rs10461909, rs11180435, rs1784232, rs3783412, rs10885378, rs1726254, rs2391691, rs3739422, rs10108245, rs1425916, rs1325922, rs1709795, rs1934395, rs2280916, rs2563263, rs10755578, rs1529192, rs2927899, rs2848745, rs10977980.

Mycoplasma contamination

All stocks are tested for mycoplasma prior to and after cells are cryopreserved. Method used is the Lonza MycoAlert Mycoplasma Detection Kit (LT07-318).

Commonly misidentified lines
(See ICLAC register)

none

Palaeontology and Archaeology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

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

Mouse strain used in this manuscript: C57BL/6 (Jackson lab, Cat# 000664)

Wild animals

No wild animals were used in in this study

Reporting on sex

Female mice were used exclusively for this experiment due to their availability at the time of the experiment. The biological sex of the donor animals is not expected to influence the outcomes or interpretations of our findings, as the experimental design and endpoints were not sex dependent.

Field-collected samples	No field -collected samples were used in this study
Ethics oversight	Animal studies were approved by Genentech's Institutional Animal Care and Use Committee and adhere to the NRC Guidelines for the Care and Use of Laboratory Animals.

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	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

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.*

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication. — *For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*

Files in database submission — *Provide a list of all files available in the database submission.*

Genome browser session
(e.g. UCSC) — *Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates — *Describe the experimental replicates, specifying number, type and replicate agreement.*

Sequencing depth — *Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*

Antibodies — *Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*

Peak calling parameters — *Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*

Data quality — *Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*

Software — *Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation — *Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*

Instrument — *Identify the instrument used for data collection, specifying make and model number.*

Software — *Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)

Specify: functional, structural, diffusion, perfusion.

Field strength

Specify in Tesla

Sequence & imaging parameters

Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition

State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization

If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template

Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal

Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See Eklund et al. 2016)

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity	*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*
Graph analysis	*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*
Multivariate modeling and predictive analysis	*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*