

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
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> 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)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input type="checkbox"/>	<input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), 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 was used for data collection.
Data analysis	Sequencing data was processed with CellRanger GEX and VDJ (v6.0.0) using the GRCh38-2020-A and vdj_GRCh38_alts_ensembl-5.0.0 human reference genomes, respectively. scRNA-seq data were analysed and integrated using the python packages scanpy (1.8.2) and scvi-tools (0.15.2). Cells were called as doublets by the python package scrublet (0.2.3). Minimum-Distortion Embedding were calculated using the Python package pymde (0.1.15). For visualisation of a large number of cells on either UMAP or MDE, the R package scattermore (1.0) was used. Differential expression between specified conditions was performed using the R package scran (1.26.2). Pathway analysis of differentially expressed genes was performed using the R package fgsea (1.24.0) and gene sets ccessed via the R package msigdb (7.5.1). Cell were scored with gene sets using the R package UCell (2.2.0). De novo label prediction tools were run with default parameters: Azimuth (https://azimuth.hubmapconsortium.org/) and Celltypist (https://www.celltypist.org/). For compositional analysis, the R package zCompositions (1.4.0-1) and compositions (2.0-6) were used. Mixed-effect models were implemented using the R package lmerTest (3.1-3). Clusters of TCRs with similar sequence features were identified using the Python package tcrdist3 (0.2.2), and networks were visualised using the R package igraph (1.4.2). HLA genotyping was performed using arcasHLA (3.24.0). Neoantigen prediction was performed using NetMHCpan (4.1). Survival analyses were performed using the R packages maxstat (0.7-25) and survival (3.5-5) with default parameters. For comparison of means in box plots P values were calculated by two-sided Wilcoxon test and for R and P values in correlations were calculated by Pearson correlation using the R package ggpubr (0.6.0).

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 unprocessed single-cell RNA and TCR data have been deposited in the Zenodo repository under the accession 13171648 (<https://zenodo.org/records/13171648>). The full integrated single-cell RNA and TCR datasets and cohort information have been deposited in the Zenodo repository under the accession 17418275 (<https://zenodo.org/records/17418275>). Specifically, data shared through the gene expression omnibus (GEO) can be accessed for Maura et al. under the accession GSE161195, Bailur et al. GSE163278, Oetjen et al. GSE120221, Granja et al. GSE139369, Zavidij et al. GSE124310, Kfoury et al. GSE143791, Zheng et al. GSE156728, Botta et al. GSE205393, and Friedrich et al. GSE216571. Data shared via dbGaP for Sklavenitis-Pistofidis et al. can be accessed under accession phs002476.v1.p1. Data shared online can be accessed for Stephenson et al. at covid19cellatlas.org, Conde et al. at tissueimmunecellatlas.org, and Liu et al. at humancellatlas.org/projects/2ad191cd-bd7a-409b-9bd1-e72b5e4cce81.

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	Biological sex and gender were not analysed in this study, but biological sex was collected and reported (female n = 76, male = 71)
Reporting on race, ethnicity, or other socially relevant groupings	Race, ethnicity and other socially relevant groupings were not collected, reported or analysed in this study.
Population characteristics	Human research participants included in this study were from a variety of clinical cohorts. Non-cancer controls in healthy individuals sampled for research (n = 87), non-cancer patients sampled during hip replacement treatment (n = 10), and deceased non-cancer patients (n = 10). Plasma cell dyscrasias patients included those diagnosed with myeloma precursor conditions monoclonal gammopathy of undetermined significance (MGUS, n = 62) and smouldering multiple myeloma (SMM, n = 236), alongside patients diagnosed with multiple myeloma (n = 94). These numbers quantify the number of patients across all assays. All patients were untreated. The median age of all participants was 58.5 years (range 21-87).
Recruitment	Bone marrow aspirates from individuals with myeloma or precursor conditions were obtained from patients included in one of four ongoing clinical trials: (1) Defining risk in smouldering myeloma (SMM) for early detection of multiple myeloma (COSMOS), a multicentre, observational UK study in smouldering myeloma (NCT05047107, COSMOS study UK Research Ethics Committee reference: 270077); (2) Risk-Adapted therapy Directed According to Response (RADAR), a randomised phase II/III trial in newly diagnosed patients with multiple myeloma eligible for transplant (UK Research Ethics Committee reference: 20/LO/0238) (42); (3) Carfilzomib/Cyclophosphamide/Dexamethasone with Maintenance Carfilzomib in Untreated Transplant-eligible Patients with Symptomatic MM to Evaluate the Benefit of Upfront ASCT (CARDAMON), a phase II trial (UK Research Ethics Committee reference: 148600) (43); (4) Biology of Myeloma, an observational study open to all plasma cell disorder patients treated at University College London Hospitals (Research ethics committee reference: 07/Q0502/17). Bone marrow aspirates from non-cancer controls were collected as a by-product of routine elective orthopaedic surgery (hip or knee replacements) via the UCL/ UCLH Biobank for Studying Health and Disease (UK Research Ethics Committee no: 272816). Material was obtained following written informed consent in accordance with the Declaration of Helsinki.
Ethics oversight	COSMOS: UK Research Ethics Committee reference: 270077; RADAR: UK Research Ethics Committee reference: 20/LO/0238; CARDAMON: UK Research Ethics Committee reference: 148600; Biology of Myeloma: Research ethics committee reference: 07/Q0502/17; UCL/ UCLH Biobank for Studying Health and Disease: UK Research Ethics Committee no: 272816. Material was obtained following written informed consent in accordance with the Declaration of Helsinki.

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 were performed. Sample sizes were chosen based on availability of data for cohorts of interest and deemed sufficient given comparisons all included a minimum of n = 5 individuals per group.
Data exclusions	No data were excluded from the analysis.
Replication	No replication was required as the study details descriptive findings of a fixed cohort.
Randomization	No randomization was performed as no interventions requiring randomisation (i.e. treatment arms) were given.
Blinding	No blinding was performed as no interventions (i.e. treatments) were given.

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
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	For flow cytometry, fluorochrome-conjugated antibody CD138 (PE, clone MI15, BioLegend, 1:100), CD38 (PE-CY7, Clone HB7, biolegend, 1:25), and a fixable viability dye (eFluor 780, eBioscience, 1:25). For cytometry by time-of-flight, CD3 (UCHT1, 89Y, Fluidigm, 1:200), CD45 (HI30, 106Cd, Fluidigm, 1:100), CD45 (HI30, 110Cd, Fluidigm, 1:100), CD45 (HI30, 111Cd, Fluidigm, 1:100), CD45 (HI30, 112Cd, Fluidigm, 1:100), CD45 (HI30, 113Cd, Fluidigm, 1:100), CD45 (HI30, 114Cd, Fluidigm, 1:100), CD16 (3G8, 116Cd, Fluidigm, 1:100), CD38 (HIT2, 141Pr, Fluidigm, 1:400), CD19 (HIB19, 142Nd, Fluidigm, 1:100), CD45RA (HI100, 143Nd, Fluidigm, 1:200), CD69 (FN50, 144Nd, Fluidigm, 1:200), CD4 (RPA-T4, 145Nd, Fluidigm, 1:200), CD8 (RPA-T8, 146Nd, Fluidigm, 1:200), CD127 (A019D5, 147Sm, BioLegend, 1:200), ICOS (C398.4A, 148Nd, Fluidigm, 1:200), CD56 (NCAM16.2, 149Sm, Fluidigm, 1:200), OX40 (ACT35, 150Nd, Fluidigm, 1:200), KLRG1 (MAFA (SA231A2), 151Eu, Fluidigm, 1:200), CD24 (ML5, 152Sm, BioLegend, 1:200), TIM3 (F38-2E2, 153Eu, Fluidigm, 1:200), TIGIT (MBSA43, 154Sm, Fluidigm, 1:200), PD1 (EH12.2H7, 155Gd, Fluidigm, 1:200), EOMES (WD1928, 156Gd, Fluidigm, 1:200), 41BB (4B4-1, 158Gd, Fluidigm, 1:200), CD357 (GITR) (621, 159Tb, Fluidigm, 1:200), CD28 (CD28.2, 160Gd, Fluidigm, 1:200), TBET (4B10, 161Dy, Fluidigm, 1:200), FOXP3 (259D/C7, 162Dy, Fluidigm, 1:200), TOX (REA473, 163Dy, Miltenyi, 1:100), CD45 RO (UCHL1, 164Dy, Fluidigm, 1:200), TCF7 (7F11A10, 165Ho, Biolegend, 1:100), NKG2D (ON72, 166Er, Fluidigm, 1:200), LAG3 (3DS223H, 167Er, Fluidigm, 1:100), CD138 (DL-101, 168Er, Fluidigm, 1:100), CD25 (2A3, 169Tm, Fluidigm, 1:200), CTLA4 (143D, 170Er, Fluidigm, 1:200), CD101 (BB27, 171Yb, Fluidigm, 1:200), KI67 (B56, 172Yb, Fluidigm, 1:200), GZMB (GB11, 173Yb, Fluidigm, 1:400), HLA-DR (L243, 174Yb, Fluidigm, 1:400), CD14 (M5E2, 175Lu, Fluidigm, 1:200), CD57 (HCD57, 176Yb, Fluidigm, 1:200), DNA (191Ir, Fluidigm, 1:1000), DNA (193Ir, Fluidigm, 1:1000), Cisplatin (195Pt, Fluidigm, 1:1000), CD11b (ICRF44, 209Bi, Fluidigm, 1:200).
Validation	No antibody validation was performed beyond that listed on manufacturers website.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.
Authentication	Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.
Mycoplasma contamination	Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See [ICLAC](#) register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

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

For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.

Wild animals

Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Reporting on sex

Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

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.

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

Individuals with myeloma or precursor conditions were obtained from patients included in one of three ongoing clinical trials: Defining risk in smouldering myeloma (SMM) for early detection of multiple myeloma (COSMOS; NCT05047107); Risk-Adapted therapy Directed According to Response (RADAR; CRUK Reference Number: RV-CL-MM-PI-13264), Carfilzomib/Cyclophosphamide/Dexamethasone with Maintenance Carfilzomib in Untreated Transplant-eligible Patients with Symptomatic MM to Evaluate the Benefit of Upfront ASCT (CARDAMON; NCT02315716).

Study protocol

Study protocols are published or available online for COSMOS (NCT05047107), RADAR (DOI: 10.1136/bmjopen-2022-063037), and CARDAMON (DOI: 10.1016/S2352-3026(22)00350-7).

Data collection

Bone marrow aspirates from individuals with myeloma or precursor conditions were obtained from patients included in one of the aforementioned trials. Patients were recruited from across the United Kingdom during the period of 2015-2022.

Outcomes

Time-to-progression data from the COSMOS study (NCT05047107) are reported.

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 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Public health |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> National security |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Crops and/or livestock |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Ecosystems |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Any other significant area |

Experiments of concern

Does the work involve any of these experiments of concern:

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

Plants

- | | |
|-----------------------|--|
| Seed stocks | <i>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.</i> |
| Novel plant genotypes | <i>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.</i> |
| Authentication | <i>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, mosaicism, off-target gene editing) were examined.</i> |

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	<i>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.</i>
Antibodies	<i>Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.</i>
Peak calling parameters	<i>Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.</i>
Data quality	<i>Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.</i>
Software	<i>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.</i>

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	Bone marrow (BM) mononuclear cells (MNCs) were isolated by Ficoll Paque density gradient centrifugation, using SepMate tubes (StemCell Technologies). Freshly isolated BM MNCs were stained with the fluorochrome-conjugated antibody CD138 (PE, clone MI15, BioLegend), CD38 (PE-CY7, Clone HB7, biolegend), and a fixable viability dye (eFluor 780, eBioscience).
Instrument	LSRFortessa, 4 laser 16 color
Software	FlowJo v10, BD Biosciences
Cell population abundance	The abundance of live tumour cells was determined as the frequency of live BM MNCs cells co-expressing CD38 and CD138 and the abundance of late different T-cell subsets by the expression of CD8+ in absence of IL7R as determined via manual gating.
Gating strategy	Mononuclear cells were first gated from all cells as those with greater than roughly 50,000 FSC-A with no cut-off of SSA-A values. Next, singlets were identified as the high density of cells fitting the linear correlation between FSC-A and FSC-H (excluding the FSC-A high population conventionally assumed to be doublets). Live cells were identified through low expression (roughly < 20,000) of the viability dye (eFluor 780, eBioscience). Live cells were then gated to CD138+CD38+ manually, with a clear double positive population emerging.

- ☒ 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	<i>Indicate task or resting state; event-related or block design.</i>
Design specifications	<i>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.</i>
Behavioral performance measures	<i>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).</i>

Acquisition

Imaging type(s)	<i>Specify: functional, structural, diffusion, perfusion.</i>
Field strength	<i>Specify in Tesla</i>
Sequence & imaging parameters	<i>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.</i>
Area of acquisition	<i>State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.</i>

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

A latent representation generated from a trained neural network model was used to create a k-nearest neighbors graph used for graph-based clustering using the Leiden algorithm.

A cell-cell interaction graph was created using differentially-expressed genes and the ligand-receptor interaction database OmniPath, where each node represented a differentially-expressed signaling molecule and each edge the presence of a reported interaction between a pair of molecules.

Multivariate modeling and predictive analysis

Cell type abundance data was used as input for a combination of intercept-only and additive regression models exploring the relationship between abundance and different covariates (for example, patient group, or patient group and age). Selected comparisons were also performed using a mixed-effect model with an additional random effect term (for example, study of origin)