

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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ 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	n/a
Data analysis	(1) python 3.12 - data exploration, statistical analysis, visualization (2) R 4.5.1 - data exploration, statistical analysis, visualization (3) Rosace 1.1 - function score modeling (4) Cutadapt 5.1 - sequence data processing (5) PEAR 0.9.6 - sequence data processing (6) BWA 0.7.17-r1188 - sequence data processing (7) samtools 1.22.1 - sequence data processing (8) snpEff 5.1 - variant annotation

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](#)

Variant function scores and associated annotations are supplied in Supplementary Tables 4 and 5. Processed variant counts are available at NCBI (accession GEO: GSE312720), and raw sequencing reads are available at NCBI (accession: PRJNA1375770).

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	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	University of Michigan IRB (declared exempt)

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](https://www.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	Not applicable as all variants were analyzed regardless of presence in any cohorts
Data exclusions	Variants at low library abundance (<100 counts per million counts) were filtered out and not considered further.
Replication	MAVE experiments performed in biological triplicate per standard practice. Cell culture seeded with at least 3 million cells to avoid bottlenecks. 16 replicate PCR reactions were pooled for each gDNA sample to maintain library complexity. Transient gene expression in cell lines and in vitro biochemical experiments were performed as multiple independent biological replicates.
Randomization	N/A
Blinding	Variants are effectively blinded within pooled screens as genotypes are not evident until sequencing

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	Antibodies against RAD51D (Abcam, 202063, 1:1000), KAR2 (Santa Cruz, sc-33630, 1:1000), and α -tubulin (Cell Signaling Technologies, 2144, 1:1000) (Jackson anti-rabbit #111-165-003, 1:10,000)
Validation	All antibodies used are commercially available (https://www.abcam.com/en-us/products/primary-antibodies/rad51d-antibody-epr16205-ab202063 , https://datasheets.scbt.com/sc-33630.pdf , https://www.cellsignal.com/products/primary-antibodies/alpha-tubulin-antibody/2144 , https://www.jacksonimmuno.com/catalog/products/111-165-003)

Eukaryotic cell lines

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

Cell line source(s)	U2OS-SCR18-WT (via Mauro Modesti, https://doi.org/10.1371/journal.pgen.1008355), U2OS-SCR18-RAD51D-KO (DSMZ, #ACC835)
Authentication	Shallow pass WGS
Mycoplasma contamination	Tested monthly by PCR
Commonly misidentified lines (See ICLAC register)	N/A

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

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	Following a 72 hr incubation, cells were collected by trypsinization and fixed in an aqueous 2% paraformaldehyde solution. Fixed cells were re-suspended in phosphate-buffered saline (PBS) and analyzed on the BD Accuri C6 Plus flow cytometer. For
--------------------	---

	each sample, 30,000 events were recorded, and results were normalized to the complemented RAD51D WT control. Each variant was analyzed in at least three biological replicates.
Instrument	BD Accuri C6 Plus
Software	BD-Accuri
Cell population abundance	Normalized to WT complemented control
Gating strategy	<i>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.</i>

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