nature portfolio

Corresponding author(s): Joshua N. Leonard

Last updated by author(s): 6/22/2025

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	Flow cytometry data were collected using BD FACS DIVA software (v8.0.3).
Data analysis	Flow cytometry data was analyzed using FlowJo software (v10.8.1). Where indicated, western blots were analyzed using ImageJ2 version: 2.9.0/1.53t. Graphs were generated using Microsoft Excel 365. Statistical analyses were performed using Microsoft Excel 365, RStudio, and GraphPad. Figures were compiled using Adobe Illustrator 2022. Microscopy images were analyzed using MATLAB R2020b (9.9.0.1467703).

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

Plasmids used in this study and the corresponding figure panels in which they were used are provided in Supplementary Data 1. Plasmid maps for all plasmids used in this study are provided as annotated GenBank files in Supplementary Data 2. Notable receptor amino acid sequences can also be found in Supplementary Note

10. Plasmids used in this study will be deposited with and distributed by Addgene, including complete and annotated GenBank files, at https://www.addgene.org/Joshua_Leonard/. The exact doses of DNA used in each experiment are listed in Supplementary Data 3. All reported experimental data for main figures are provided in Source Data 1, all reported experimental data for supplementary figures are provided in Supplementary Data 4, unprocessed western blot images are available in Source Data 2, and unprocessed microscopy images are provided in Source Data 3. The raw datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Not applicable
Population characteristics	Not applicable
Recruitment	Not applicable
Ethics oversight	Not applicable

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	Nearly all experiments were conducted in biological triplicate (i.e. N = 3), which is a standard size for such experiments given prior observations with similar systems (e.g., Schwarz, K. et al. Rewiring human cellular input-output using modular extracellular sensors. Nat Chem Biol 13, 202–209 (2017). https://doi.org/10.1038/nchembio.2253; Dolberg, T. et al. Computation-guided optimization of split protein systems. Nat Chem Biol 17, 531–539 (2021). https://doi.org/10.1038/s41589-020-00729-8)
Data exclusions	In one experiment, CAR surface expression was quantified using N=2 (Fig. 3d). In one experiment to validate split intein function, cells were transfected with N=1 (Supplementary Fig. 9f). The number of replicates used in each figure panel is indicated in each respective figure legend and source data files, so these exceptions are clearly marked.
Replication	For the majority of the experiments in this manuscript, the published data is an experimental repeat of a similar, but not identical, initial pilot experiment. Similar but not identical means that the pilot experiments may have included different controls, comparison samples, or plasmid doses, and the design choices for the published experiments represent iterations on the designs from the pilot experiments. In all cases, the initial pilot and final published experiments gave highly similar, repeatable results. Generally each experiment was conducted independently two times (a pilot experiment and a repeat).
Randomization	Randomization is not relevant to this study, as the same cell lines (uniform biological material) were grown up and divided between wells prior to transfection or treatment.
Blinding	Blinding is not relevant to this study, as flow cytometry data acquisition and analysis is not subject to investigator bias that could be mitigated through blinding. Western blots are presented as images alone or images with post-hoc quantification, such that no subjective analysis was used other than discussing the data as presented. Sample preparation and analysis across all experiments followed uniform protocols as described in the Methods section.

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

Methods

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

Antibodies

Antibodies used

Monoclonal ANTI-FLAG clone M2 antibody produced in mouse (Sigma Cat# F1804; RRID: AB_262044)
Anti-mouse IgG, HRP-linked antibody (Cell Signaling Technology Cat# 7076; RRID: AB_330924)
Monoclonal anti-myc clone 9E10 antibody produced in mouse (Abcam Cat# ab32, RRID: AB_303599
Monoclonal Anti-FLAG clone M2, APC-linked antibody produced in mouse (Abcam Cat# ab72569, RRID:AB_1310127)
Polyclonal Anti-human/primate EGF, biotinylated antibody produced in goat (R&D Systems Cat# BAF236, RRID:AB_356307)
Streptavidin, APC-linked antibody (Abcam Cat# 134362)

Validation

From Sigma: "Our standard antibody validation processes include verification for each recommended immunodetection application. Each of the thousands of antibodies in or portfolio are certified through our standard validation process to ensure quality and reproducibility."
From Cell Signaling Technology: "Every CST antibody undergoes rigorous application-specific validation testing customized according to the target and needs of the individual antibody."
From Abcam: "Antibody validation must be application-specific to be effective and information on which applications an antibody has been validated in can be found in the Tested Applications section on any antibody datasheet."
From R&D Systems: "R&D Systems® takes rigorous steps towards antibody validation and reproducibility. We have been since the beginning. For 30 years, we have used our industry-leading production standards and quality control specifications to develop antibodies that can be relied on for specificity and reproducibility. By developing and testing our products in-house, we can ensure a validated and specific antibody. We are confident in our antibodies and provide 100% guarantee for our products."
Whenever possible, negative controls are included in the antibody experiments. For Western blots, this includes analogous samples that do not contain the tag being detected. For flow cytometry surface stains, this includes analogous cell lines or cell samples lacking the antibody's target receptor or an appropriate isotype control.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

HEK293FT cells, Thermo Fisher R70007
Jurkat T cells, ATCC TIB-152
SKOV3 cells, ATCC HTB-77
Other cell lines derived from those above using lentiviral or transposon-mediated integration
HEK293FT-LP cells, gift from Ron Weiss lab
Other cell lines derived from HEK293FT-LP cells using landing pad Bxb1-mediated integration

Authentication

The cell line vendors perform routine authentication, so no further authentication was performed. The HEK293FT-LP line was authenticated only by flow cytometric analysis of EYFP expression, which was shown to be highly homogeneous and stable over time, consistent with the original description of this cell line in the literature.

Mycoplasma contamination

HEK293FT, HEK293FT-LP, Jurkat T cells, and SKOV3 cells tested negative for mycoplasma. Cell lines derived from these parent cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

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	For fluorescent reporter assays with HEK293FT-based cells, cells were harvested for flow cytometry by incubating with Tryspin-EDTA for 3 minutes at 37 degrees Celsius and quenching with completed media, and the resulting cell solution was added to at least 2 volumes of FACS buffer (PBS pH 7.4 with 2-5mM EDTA and 0.1% BSA). For fluorescent reporter assays with Jurkat cells, cell suspensions were pipetted directly into at least 2 volumes of FACS buffer (PBS pH 7.4 with 2-5mM EDTA and 0.1% BSA). Cells were spun at 150 x g for 5 min, FACS buffer was decanted, and 1 drop of fresh FACS buffer was added. For surface staining, cells were harvested by incubating with FACS buffer for 5 minutes at room temperature. For anti-FLAG staining, cells were spun at 150 x g for 5 min, incubated with human IgG isotype control for 5 minutes at 4 degrees Celsius, incubated with staining antibody for 30 minutes at 4 degrees Celsius, and washed three times post-staining with 1 mL of FACS buffer, centrifuging at 150 x g for 5 min and decanting the supernatant after each wash. For anti-CAR staining, cells were spun at 150 x g for 5 min, incubated with human IgG isotype control for 5 minutes at 4 degrees Celsius, incubated with primary staining antibody for 20 minutes at 4 degrees Celsius, washed once with 1 mL of FACS buffer (spin at 150 x g for 5 min and decant), incubated with secondary staining antibody for 20 minutes at 4 degrees Celsius, and washed two times post-staining with 1 mL of FACS buffer, centrifuging at 150 x g for 5 min and decanting the supernatant after each wash. Cells were resuspended in two drops of FACS buffer prior to flow cytometry. For experiments with Jurkat cells, a viability dye was included in FACS buffer.
Instrument	Analytical flow cytometry was run on a BD LSR Fortessa Special Order Research Product (Robert H. Lurie Cancer Center Flow Cytometry Core). FACS was conducted on a BC FACS Aria SORP (Robert H. Lurie Cancer Center Flow Cytometry Core). Instrument configurations can be found in the Methods section.
Software	Samples were analyzed using FlowJo v10 software (FlowJo, LLC).
Cell population abundance	For Bxb1-recombinase-engineered HEK293FT-LP reporter cell lines: HEK293FT-LP cells were transfected with recombinase and landing pad integration vector and selected for genomic integration using puromycin. To gate for cell sorting, single cells were first identified as above. Then, a quadrant gate in FITC (EYFP) and APC-Cy7 (miRFP720) channels was drawn to include only miRFP720+ and EYFP- cells. The quadrant gate was set in the FITC channel such that 99.5% of unmodified HEK293FT cells were considered EYFP-. The quadrant gate was set in the APC-Cy7 channel such that 99.5% of unmodified HEK293FT-LP cells were considered miRFP720-. Only EYFP- and miRFP720+ cells were included in further gating. Of the EYFP-/miRFP720+ cells, the cells in the 88-98th percentile of miRFP720 expression were collected. For transposon-engineered HEK293FT and HEK293FT-LP cell lines: Cells were transfected with transposase and transposon vector and selected for genomic integration using the appropriate antibiotic (puromycin or hygromycin). To gate for cell sorting, single cells were first identified as above. Then, a gate on FITC (mNeonGreen) was set to include only mNeonGreen+ cells, defined by a sample of non-modified parent HEK293FT cells (non-fluorescent) such that less than 1% of parent cells are considered mNeonGreen+. From here, two different sorting strategies were employed. For some experiments, the mNeonGreen+ cells were divided into octiles that each contained 12.5% of the mNeonGreen+ population and the top (brightest mNeonGreen) four octiles were collected separately. For other experiments, an additional gate to isolate low background reporter (PE-TexasRed, DsRedExpress2) expressing cells, defined by a sample of non-modified parent HEK293FT cells such that 99% of parent cells are considered low reporter expressers. Then, the same octile approach described previously was used to collect each of the top 4 mNeonGreen+ octiles separately. For transposon-engineered Jurkat cell lines: Cells were transfected with transposase and transposon vector and selected for genomic integration using puromycin. To gate for cell sorting, single cells were first identified as above. Then, a gate on FITC (mNeonGreen) was set to include only mNeonGreen+ cells, defined by a sample of non-modified parent Jurkat cells (non-fluorescent) such that less than 1% of parent cells are considered mNeonGreen+. The top 12.5% of mNeonGreen+ cells (the brightest mNeonGreen-expressing cells, the top octile) was collected.
Gating strategy	Live HEK293FT cells were identified by FSC-A vs. SSC-A gating, and single cells were then identified by FSC-A vs. FSC-H gating. Jurkat cells were identified by FSC-A vs. SSC-A gating, single cells were then identified by FSC-A vs. FSC-H gating, and live cells were identified by gating for DAPI- cells. An unstained Jurkat sample (no DAPI) was used to draw a gate around DAPI-negative cells such that less than 1% of unstained cells are considered DAPI+. Beads were identified by FSC-A vs. SSC-A gating. A fluorescent protein was used to identify transfected or genomically modified cells in relevant experiments. A sample transfected with an empty vector or an unmodified parent cell sample was used to draw a gate identifying transfected/ engineered cells such that less than 1% of non-fluorescent cells were included in the analysis.

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