nature portfolio

Corresponding author(s): Ed Boyden, Kiryl D. Piatkevich

Last updated by author(s): Aug 1, 2021

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	P.27 Data collection for fluorescence spectra of NIR FPs were done using https://www.fpbase.org.
Data analysis	P.27 Data were analyzed offline using NIS-Elements Advance Research software, Excel (Microsoft), OriginPro, ImageJ, the Microscope online application (https://www.fpbase.org/microscope), and Arivis Vision4D.

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

All other data generated or analyzed during this study are available from the corresponding author on reasonable request. The gene sequences of the new proteins will be deposited to GenBank.

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	P. 27 We did not perform a power analysis, since our goal was to create a new technology; and recommended by the NIH, "In experiments based on the success or failure of a desired goal, the number of animals required is difficult to estimate…" As noted in the aforementioned paper, "The number of animals required is usually estimated by experience instead of by any formal statistical calculation, although the procedures will be terminated [when the goal is achieved]." These numbers reflect our past experience in developing neurotechnologies.
Data exclusions	P. 26 Fish larvae exhibiting no fluorescence in the corresponding channels were excluded from further imaging. P. 27 No larvae or adult flies, carrying the genes of target proteins, were excluded from the study. P. 27 worms without green fluorescence were excluded from further imaging...
Replication	P.27 All attempts at replication of the experiments were successful.
Randomization	P.27 No randomization was used in the study.
Blinding	P. 27 No blinding was used in the study.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293FT (Invitrogen) HeLa (ATCC CCL-2), NIH3T3 (DSMZ).
Authentication	Cells were authenticated by the manufacturer and tested for mycoplasma contamination to their standard levels of stringency and were here used because they are common cell lines for testing new tools.
Mycoplasma contamination	Cells were authenticated by the manufacturer and tested for mycoplasma contamination to their standard levels of stringency and were here used because they are common cell lines for testing new tools.
Commonly misidentified lines (See ICLAC register)	N/A

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Swiss Webster (Taconic) mice were used for primary neuronal culture preparation disregarding the gender. Embryonic day (E) 15.5 timed-pregnant female Swiss Webster mice (Taconic) were used for in utero electroporation. Acute brain slices were obtained from Swiss Webster mice (Taconic) at P20–30 without regard for sex. The pigmentation-compromised zebrafish brass strain and

	homozygous nacre embryos of the pan-neuronal expressing Gal4 line, tg(elavl3:GAL4-VP16)nns6 were used for proteins expression in zebrafish.
Wild animals	No wild type animals were used.
Field-collected samples	N/A
Ethics oversight	All procedures involving mice at MIT and Westlake University were conducted in accordance with the US National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Massachusetts Institute of Technology or Westlake University Committee on Animal Care. All experiments involving zebrafish at MIT and Technische Universität Braunschweig were conducted in accordance with protocols approved by Massachusetts Institute of Technology Committee on Animal Care following guidelines described in the US National Institutes of Health Guide for the Care and Use of Laboratory Animals or by German legislation following European Union guidelines (EU Directive 2010_63) according to location of the respective experiments.

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

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	P.19-20 To sort the gene library-transfected HEK293FT cells using flow cytometry, cells were harvested from a culture dish ~ 48h after gene library transfection by applying trypsin for 5–10min (Cellgro) and then washed twice by centrifuging the cell suspension for 5min at 500rpm and re-suspending cells in PBS (Cellgro). The washed cells were then re-suspended in PBS supplemented with 4% FBS (Corning) and 10mM EDTA at a density of 1–2·106 cells/ml and filtered through a 30-µm filter (Falcon) to prevent clogging on the FACS machine. The filtered cells were sorted by FACSAria (BD Biosciences) running BD FACS Diva 8.0 software and equipped with standard 488- and 640-nm solid-state lasers. Debris, dead cells, and cell aggregates were gated out using forward and side scatter before desired fluorescence signals were detected. For each library several hundred cells exhibiting highest fluorescent intensity in the corresponding channel were collected
Instrument	FACSAria (BD Biosciences)
Software	BD FACS Diva 8.0
Cell population abundance	Positive cells population was about 1-2%.
Gating strategy	Top ~0.0001-0.0005% of cell population

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