nature portfolio

Corresponding author(s): Gozde S Demirer

Last updated by author(s): May 19, 2026

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	No code was used for data collection.
Data analysis	GraphPad Prism Version 10.3.1 and ImageJ were used to analyze all data. Other code used to analyze data is available at https://github.com/DemirerLab. Specifically, 1. R2 5' Junction Classification 2. TTISS On- and Off-Target Insertion Analysis 3. R2 25S integration site mapping 4.3' junction analysis 5. rDNA and 3' junction methylation analysis 6. Flow cytometry data analysis All are publicly available at https://github.com/DemirerLab.

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 authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files.
All sequencing results are accessible at NCBI SRA with accession number PRJNA1463879.

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	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

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 method was used to predetermine the sample size. At minimum 3 biological replicates have been performed, which is accepted to be sufficient for these types of assays in the literature.
Data exclusions	No data was excluded.
Replication	3 researchers independently repeated the experiments at least 3 times. We were able to replicate the data in presented in the manuscript.
Randomization	Fully random allocation of plants into treatment groups were performed.
Blinding	Blinding was not possible because there was only one researcher performing data collection and analysis at a given time.

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

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	Col-0 A. thaliana and Nicotiana benthamiana seeds were obtained from the Meyerowitz Lab at Caltech. Solanum lycopersicum seeds were obtained from commercial sources.
Novel plant genotypes	Plant genotypes with mCherry and RUBY reporter DNA were generated.
Authentication	NA

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	Mesophyll protoplasts were isolated from fully expanded leaves of 4-week-old N. benthamiana or A. thaliana (Col-0) by enzymatic digestion following Yoo et al. (2007) with minor modifications. Protoplasts were adjusted to 5×10^5 cells mL^{-1} in MMG solution and transfected by PEG 4000-mediated delivery of 20 µg plasmid DNA per 100 µL protoplast suspension. Following two W5 washes, transfected protoplasts were resuspended in WI solution and incubated at room temperature in the dark for 24–72 h prior to flow cytometry analysis. In a separate protocol, protoplasts were extracted from 5-week-old N. benthamiana leaves 7 days post-Agrobacterium infiltration and analyzed directly.
Instrument	Flow cytometry data were acquired on a Beckman Coulter CytoFlex S analyzer equipped with four excitation lasers (405 nm, 488 nm, 561 nm, and 638 nm). Data were collected using CytExpert onboard software.
Software	Raw FCS files were analyzed using custom Python scripts (Python 3.x) employing the fcsparser, NumPy, and Matplotlib libraries. Sequential polygon gates were applied computationally using matplotlib.path.Path.contains_points(). Summary statistics were exported as tab-separated files.
Cell population abundance	From the flow cytometry analyzer, the abundance of fluorescent reporter-positive cell populations was determined within the viable singlet protoplast gate. Transfection efficiency varied across experiments and constructs, and was quantified as the percentage of gated protoplasts exceeding the fluorescence threshold established by the wild-type negative control for each channel. In an double negative control and mCherry positive control sample, FDA has been used to check the viability of cells and confirm that P1 and P2 gates enrich live protoplasts.
Gating strategy	Protoplasts were gated sequentially as follows. A morphology gate (P1) was applied in FSC-A vs. SSC-A space to exclude debris and non-cellular events. P1 gate is selected to locate in a region validated to contain live protoplasts based on FDA staining in the double negative control and Cherry positive control protoplasts. Singlets were selected within P1 using a FSC-A vs. FSC-Width gate (P2) to exclude doublets and aggregates. Where sequential reporter gating was applied, each downstream population was gated within its parent positive population. Fluorescent transformation reporter-positive populations were identified within P2 using polygon gates in FSC-A vs. fluorescence channel space. Gate boundaries were set above the autofluorescence ceiling of wild-type untransfected protoplasts processed identically, serving as the double negative control for both transformation and R2Tg mediated insertion for each experiment. Within the transformed cell gate, mCherry positive cells were gated accordingly above the maximum signal observed in the transformation marker only negative control as well as the double negative control.

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