nature portfolio

Corresponding author(s): Erwei Zuo, Chunlong Xu, Jinhai Huang, Yingsi Zhou

Last updated by author(s): Aug 16, 2024

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	TadA protein sequences were downloaded from NCBI database
Data analysis	Bioinformatics codes were deposited in GitHub repository (https://github.com/yszhou2016/TadA)

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 the sequencing data have been deposited in the NCBI SRA under project accession number PRJNA1061823. All materials are available upon reasonable request.

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	N/A

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	Sample sizes were determined based on reported literature for genome editing experiments.
Data exclusions	For high-throughput library experiments, the target sites were excluded with a coverage depth of less than 100 in each sample.
Replication	All main text figure experiments were repeated at least once and attempts at replication were successful.
Randomization	Human cells were grown at identical condition, after seeding cells into 24-wells plates, we randomly selected cells for test group and control group. DMD mice used for gene editing therapy were allocated to control or AAV9 treated group randomly.
Blinding	No blinding was applied, due to no subjective assessments were required.

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 for western blot: primary antibodies against dystrophin (1:1000 dilution, Sigma, D8168) and vinculin (1:1000 dilution, CST, 13901S). Secondary antibody (1:1000 dilution, Beyotime. A0216). Antibodies for immunofluorescence: primary antibodies against dystrophin (1:100 dilution, Abcam, ab15277) and spectrin (1500 dilution, Millipore, MAB1622). secondary antibodies secondary

	antibodies Alexa Fluor® 488 AffiniPure donkey anti-rabbit IgG (1:1000 dilution, Jackson ImmunoResearch labs, 711-545-152) and Alexa Fluor 647 AffiniPure donkey anti-mouse IgG (1:1000 dilution, Jackson ImmunoResearch labs, 715-605-151)
Validation	https://www.sigmaaldrich.cn/CN/en/product/sigma/D8168; https://www.cellsignal.com/products/primary-antibodies/vinculin-e1e9v-xp-rabbit-mab/13901; https://www.abcam.com/products/primary-antibodies/dystrophin-antibody-ab15277.html; https://www.merckmillipore.com/CN/en/product/msds/MM_NF-MAB1622

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HEK293T
Authentication	HEK293T (ATCC) cells were validated by supplier.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	None misidentified lines were used.

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	C57 BL/6J, Humanized DMD mice, 3weeks, 4weeks or 9weeks old. Mice were housed in a barrier facility with a 12-hour light/dark cycle and 18-23°C with 40-60% humidity. Diet and water were be accessible at all times. All of these were maintained in accordance with the Instructive Notions with Respect to Caring for Laboratory Animals issued by the Ministry of Science and Technology of China.
Wild animals	No wild animals were used in this study. Humanized DMDΔE54 mice were generated in STOCK Tg (DMD) 72Thoen/J mice (#018900) background using the CRISPR/Cas9 system. Specifically, two sgRNAs targeting the flanking intron of human DMD exon54 were designed, and the T7 promoter was added to the sgRNA template. The PCR product was then purified directly using the Omega gel extraction kit, and the templates were used for in vitro transcription with the MEGAshortscript T7 Kit. The sgRNAs were purified using a MEGAclear Kit and eluted with nuclease-free water. The concentration of target sgRNA was measured using a NanoDrop instrument. For cytoplasmic injection, spCas9 mRNA (100 ng/μl), sgRNA-L (50 ng/μl) and sgRNA-R (50 ng/μl) were mixed and then injected into fertilized eggs using a FemtoJet microinjector (Eppendorf) with constant flow settings. The injected zygotes were cultured in KSOM medium for 12 hours and surgically transferred to the oviduct of recipient mice 24 hours after estrus was observed. The DMDΔE54 mdx mice were derived by mating the humanized DMDΔE54 mice with mdx mice carrying stop mutation in mouse exon 23. After AAV9 intramuscular injection 6weeks, mice were anesthetized, euthanized and TA (tibialis anterior) muscle was collection.
Reporting on sex	Duchenne muscular dystrophy (DMD) is the most common sex linked lethal disease in man, thus male mice were selected for this study.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	All animal experiments were performed and approved by the Institutional Animal Care and Use Committee (IACUC) of HuidaGene Therapeutics Inc., Shanghai, China and Lingang Laboratory, Shanghai, China.

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

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	Post-transfected cells were culture medium, and analyzed by flow cytometry.
Instrument	For EGFP activation assay, Beckman CytoFlex was used, and BD FACSAria III was used for cell sorting.
Software	FACS data were analyzed with FlowJo X (v10.0.7)
Cell population abundance	EGFP activation rates = EGFP positive cell numbers / mCherry and BFP double positive cell numbers.
Gating strategy	For protein variants screening experiments, mCherry, BFP and EGFP positive cells were analyzed with fluorescence values over 1x10^4. For assay of genome editing in endogenous sites, all mecherry positive cells were sorted with fluorescence value over 1x10^3.

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