nature portfolio

Corresponding author(s): Brian P. Brooks
Bin Guan
Rodrigo Young

Last updated by author(s): 04/07/2025

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	d27a79a1-8a5f-404d-8063-52e19122ef49.h5ad (https://cellxgene.cziscience.com/collections/4c6eaf5c-6d57-4c76-b1e9-60df8c655f1e) 88444d73-7f55-4a62-bcfe-e929878c6c78.h5ad (https://cellxgene.cziscience.com/collections/4c6eaf5c-6d57-4c76-b1e9-60df8c655f1e) https://hpc.nih.gov/~mcgaugheyd/eyeIntegration/2023/gene_counts.csv.gz https://hpc.nih.gov/~mcgaugheyd/eyeIntegration/2023/eyeIntegration23_meta_2023_09_01.built.csv.gz
Data analysis	http://www.yasara.org/ Code for single cell and bulk RNA seq and multiple sequence alignment (MSA) figures are accessioned in Zenodo (10.5281/zenodo.14164786) NGS analyses: https://github.com/NIH-NEI/NGS_genotype_calling&https://github.com/NIH-NEI/variant_prioritization

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 genome sequencing data from NEI participants generated in this study have been deposited in the dbGaP database under accession code phs003996.v1. The genome data are available under restricted access for participant consent and confidentiality, access can be obtained by following dbGaP data access policy. Data for all bioinformatic analyses have been deposited in 10.5281/zenodo.14757568. All other data are provided in the main text or in the Supplementary Information. Source data files are provided with this paper. The controlled access to the data is governed by NIH policies. The UK100K genome data can be accessed through Genomics England Limited.

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	Sex and/or gender of the participants involved in the study is provided in Table 1 of the manuscript. Sex information was collected based on patient reports of sex at birth for possible differential manifestation of coloboma in males and females. The number of affected individuals is small and therefore not amenable to statistical analysis. OVR presented as an autosomal dominant trait, sex and gender were not considered as independent variables, also note that affected individuals were both male and females (Please refer the pedigree charts).
Reporting on race, ethnicity, or other socially relevant groupings	Not reported.
Population characteristics	Not reported.
Recruitment	NEI patients were recruited by physician or self-referral. It is possible that the NEI and UK100KGP cohorts may be enriched with patients without pathogenic variants in known disease genes.
Ethics oversight	National Institutes of Health-Internal Review Board takes care of the ethics oversight. UK data ethics oversight is by Moorfields Eye Hospital.

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 Statistical methods were used to determine the sample size. Sample sizes were chosen based on accepted standards in the field. https://doi.org/10.1242/dev.187047 https://doi.org/10.1242/dev.181420 https://doi.org/10.1016/j.ydbio.2010.12.004 https://doi.org/10.1242/dev.162453
Data exclusions	No data was exculded
Replication	All the experiments were repeated twice and the results were consistent.
Randomization	All zebrafish experiments were carried out by random selection of males an females for breeding, which were maintained at standard conditions. Allocation of zebrafish embryos while injecting into controls and test group were carried at random.
Blinding	Blinding was not applicable due to quantitative nature of the experiments.

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
[ ]	[x] Antibodies
[ ]	[x] Eukaryotic cell lines
[x]	[ ] Palaeontology and archaeology
[ ]	[x] Animals and other organisms
[ ]	[x] Clinical data
[x]	[ ] Dual use research of concern
[x]	[ ] Plants

Methods

n/a	Involved in the study
[x]	[ ] ChIP-seq
[ ]	[x] Flow cytometry
[x]	[ ] MRI-based neuroimaging

Antibodies

Antibodies used	FIBRILLARIN from Invitrogen with Cat#- MA1-91878, Lot#- ZA389553, Clone #- DYKDDDDK Tag Monoclonal Antibody (FG4R) raised in mouse was used at 1:1000 dilution
Validation	https://www.fishersci.com/shop/products/dykddddk-tag-monoclonal-antibody-fg4r-invitrogen/MA191878

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	In house lot of HEK293 cells.
Authentication	Not authenticated
Mycoplasma contamination	Cell line not tested for mycoplasma
Commonly misidentified lines (See ICLAC register)	Not on the list

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	Danio rerio/AB TL and C57Bl/6 mice (The Jackson laboratory, Strain #:000664)
Wild animals	We did not use any wild animals in the study.
Reporting on sex	We did not used any sex based study, we have analyzed the phenotype of the zebrafish morphants at 72hpf and our findings does not apply to a specific sex. OVR syndrome is autosomal dominant, so we did not consider sex as a variable in any of our animal studies.
Field-collected samples	We did not use any field collected samples in our study
Ethics oversight	National Eye Institute, Animal Care and Use Committee and NIH Animal Research Advisory Committee

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJEguidelines for publication of clinical research and a completedCONSORT checklist must be included with all submissions.

Clinical trial registration	www.clinicaltrials.gov
Study protocol	NCT01778543, NCT01087320, NCT02077894
Data collection	Data is collected at NIH clinical Center, Bethesda, Maryland, USA and Moorfields Eye Hospital, London, UK.

Outcomes	Does not apply

Plants

Seed stocks	Does not apply to our study
Novel plant genotypes	Does not apply to our study
Authentication	Does not apply to our study

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	Sample preparation and source has been described in the methods section
Instrument	Mentioned in the methods section
Software	Mentioned in the methods section
Cell population abundance	No sorting was performed. Cell samples were 80% viable. Only single cells were used in the analysis, about 100, 00 cells were analyzed per sample.
Gating strategy	Gating Strategy is also shown in Supplementary Fig 5

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