

## SUPPLEMENTARY MATERIALS

### **m<sup>6</sup>A-mRNA reader *YTHDF2* identified as a potential risk gene in autism with disproportionate megalencephaly**

Sierra S. Nishizaki<sup>1,2,3,4</sup>, Nicholas K. Haghani<sup>1,5</sup>, Gabriela N. La<sup>1,5</sup>, Natasha Ann F. Mariano<sup>1,6</sup>, José M. Uribe-Salazar<sup>1,5</sup>, Gulhan Kaya<sup>1</sup>, Melissa Regester<sup>3,4</sup>, Derek Sayre Andrews<sup>3,4</sup>, Christine Wu Nordahl<sup>3,4</sup>, David G. Amaral<sup>3,4</sup>, Megan Y. Dennis<sup>1,3,4,5†</sup>

<sup>1</sup>Genome Center, <sup>2</sup>Autism Research Training Program, <sup>3</sup>Department of Psychiatry and Behavioral Sciences, <sup>4</sup>MIND Institute, <sup>5</sup>Department of Biochemistry & Molecular Medicine, <sup>6</sup>Postbaccalaureate Research Education Program, University of California Davis, CA, USA

†Corresponding author:

Megan Y. Dennis, Ph.D.

University of California, Davis

School of Medicine

One Shields Avenue

Genome Center, 4303 GBSF

Davis, CA 95616

Email: [mydennis@ucdavis.edu](mailto:mydennis@ucdavis.edu)

## SUPPLEMENTARY TABLES

**Table S1.** CRISPR zebrafish knockout gRNA and primer sequences.

**Table S2.** ASD-DM and ASD-M probands from the SSC and APP cohorts. DM, disproportionate macrocephaly; RM, relative macrocephaly; SO, somatic overgrowth.

**Table S3.** Loss-of-function *de novo* variants from ASD-DM and ASD-M probands.

**Table S4.** Enriched gene ontologies from network analysis of ASD-DM candidate genes v a human background from DAVID.

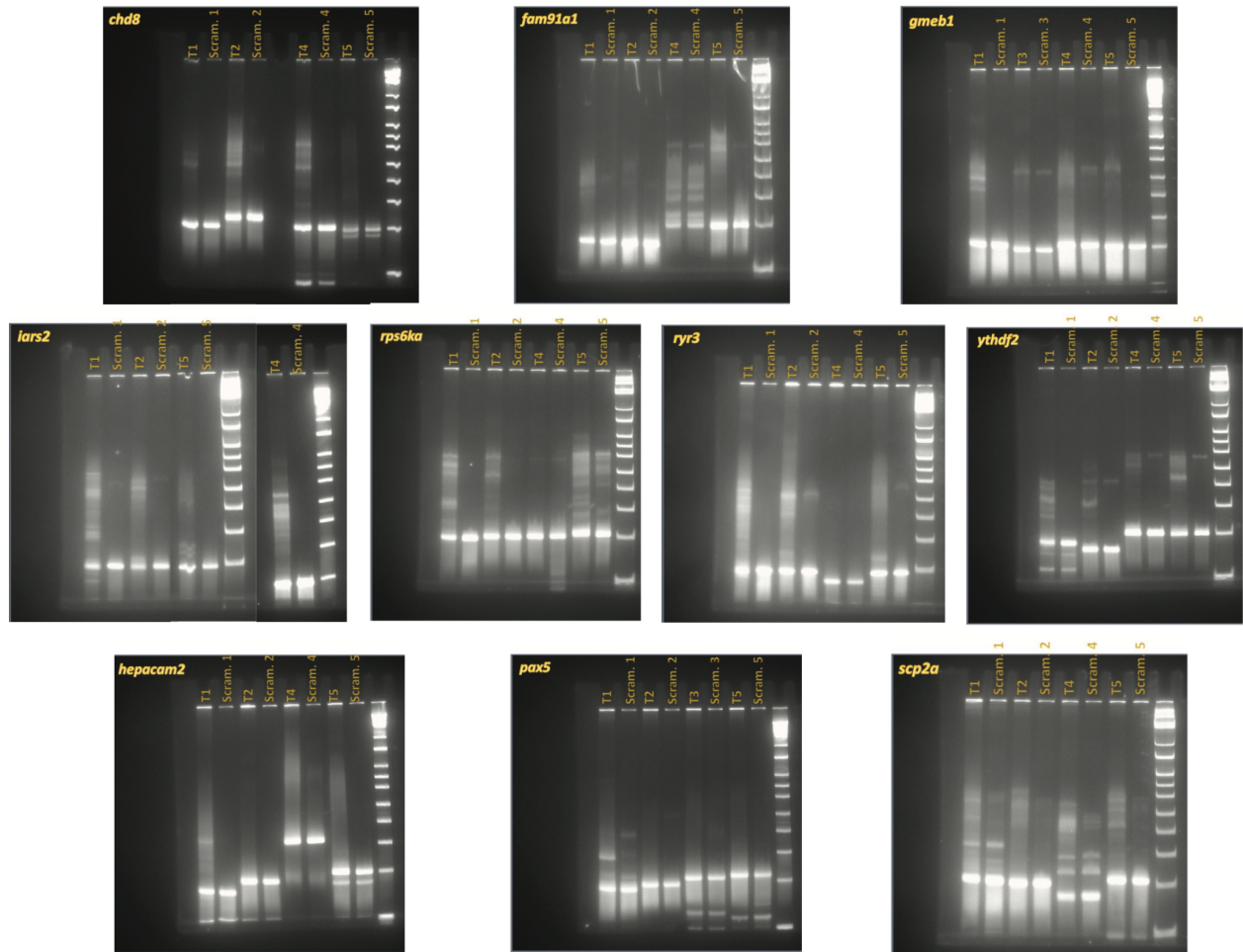
**Table S5.** Network analysis seed ASD-DM candidate genes and top 10 interactors.

**Table S6.** Enriched gene ontologies from network analysis of ASD-DM candidate genes and top 10 interactors.

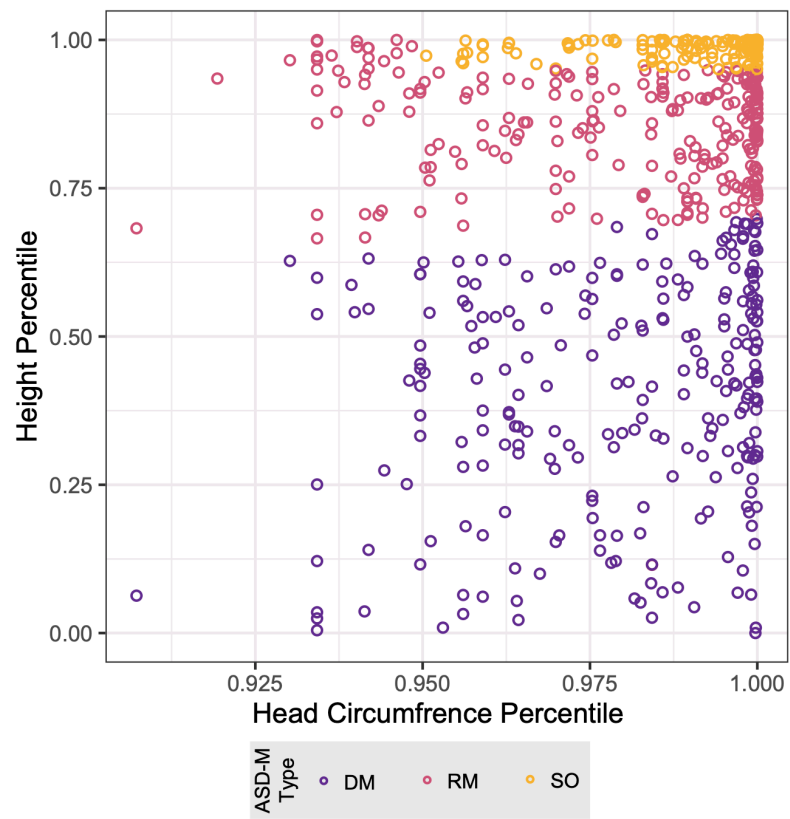
**Table S7.** Technical statistics for scRNA-seq replicates.

**Table S8.** Differential gene expression results from scRNA-seq and FMRP-gene assignment.

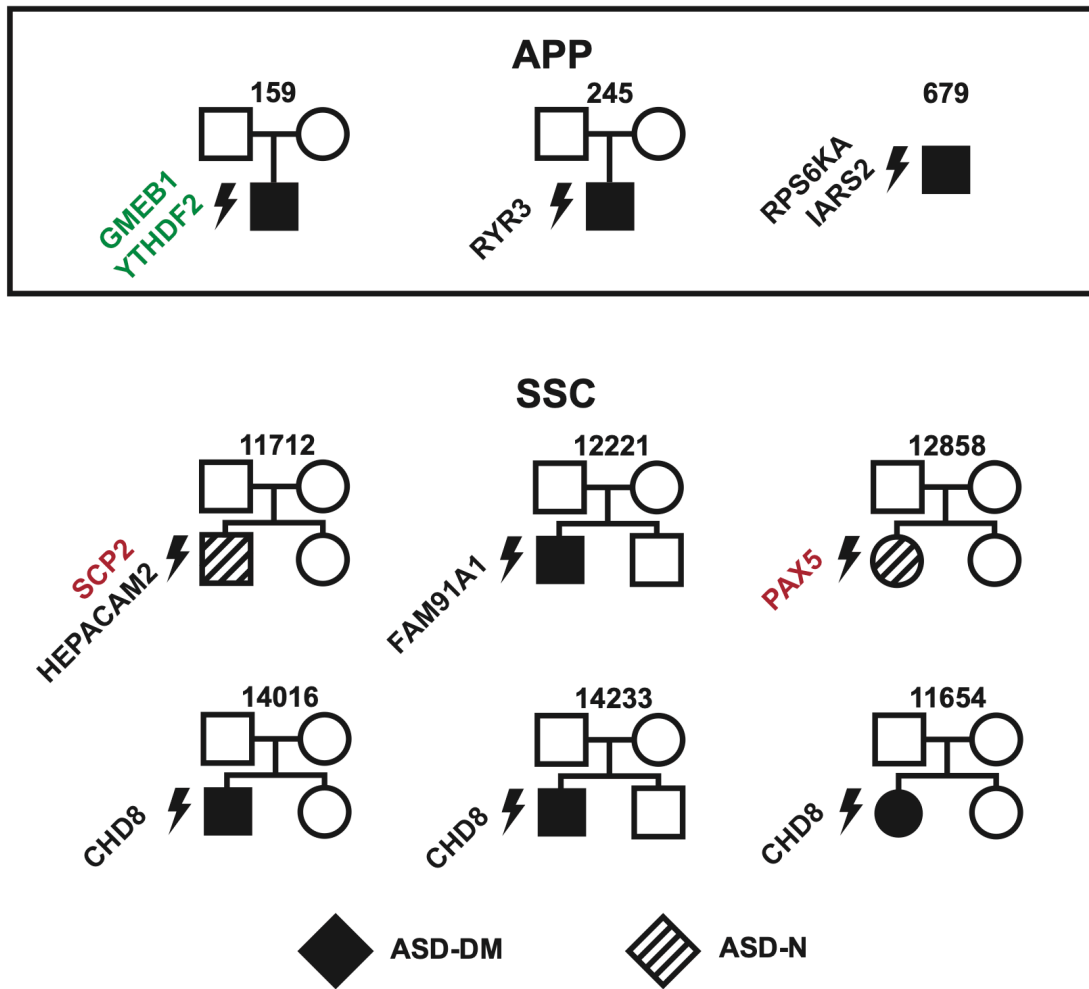
## SUPPLEMENTARY FIGURES



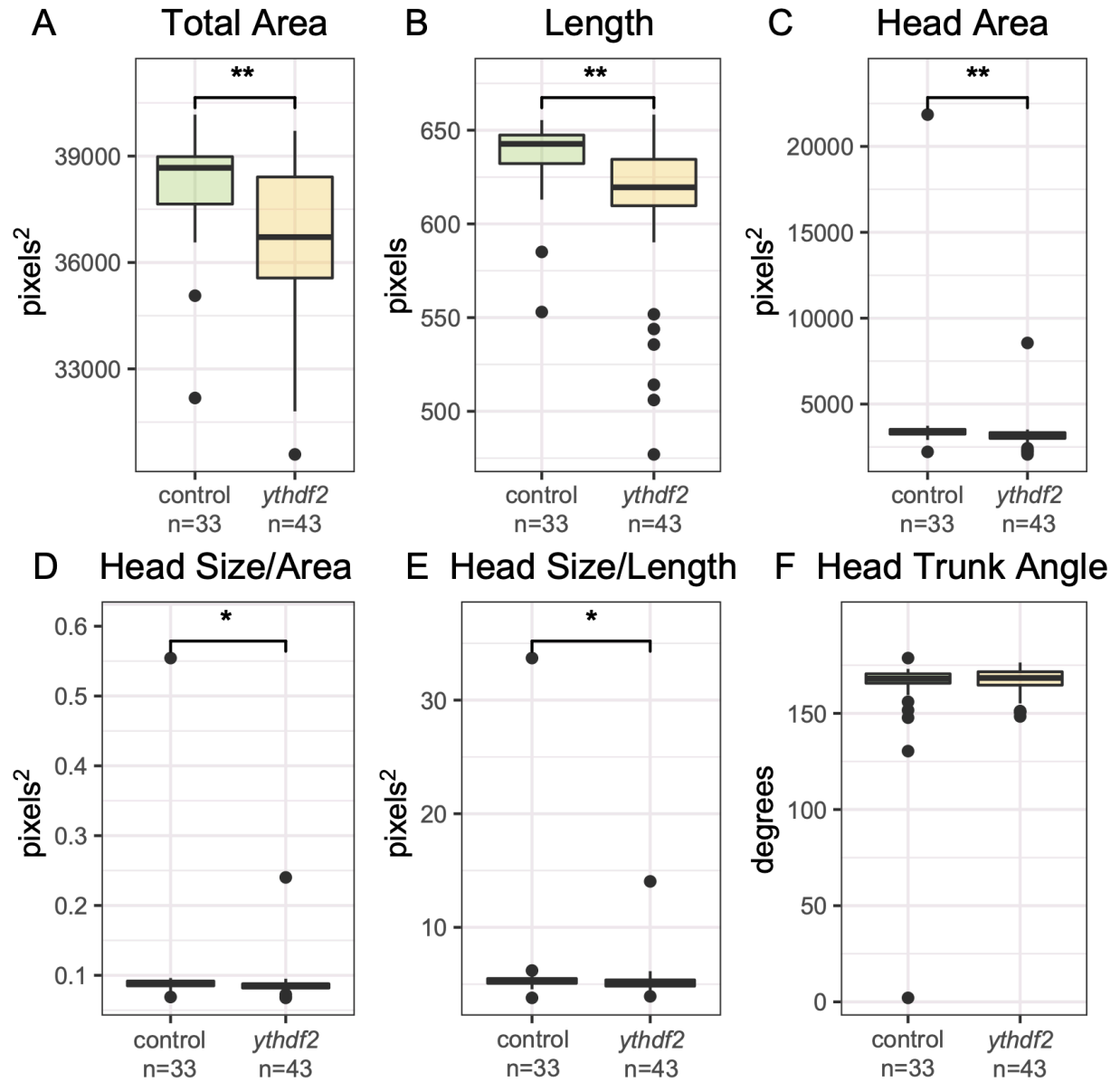
**Figure S1. Mutational efficiencies of CRISPR targets in zebrafish embryos.** Polyacrylamide gels showing the efficiency of each of the four CRISPR gRNAs chosen for the 10 genes tested in zebrafish.



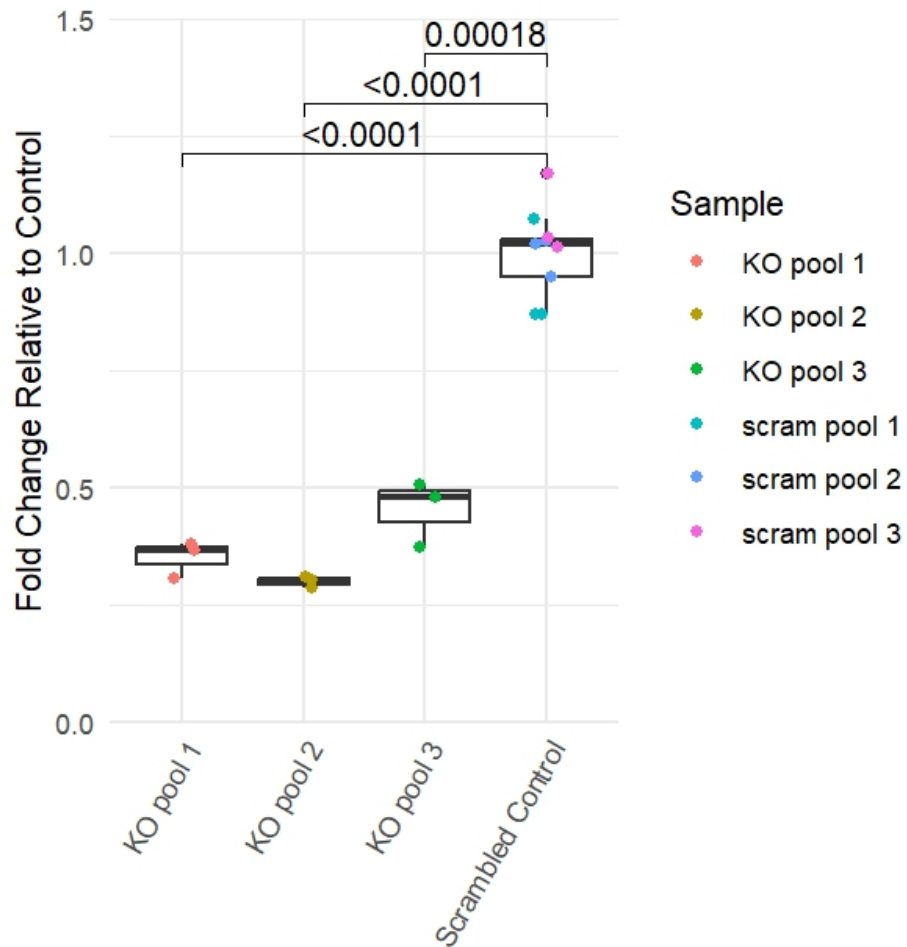
**Figure S2. Head circumference v. height percentile of ASD-M SSC probands.**



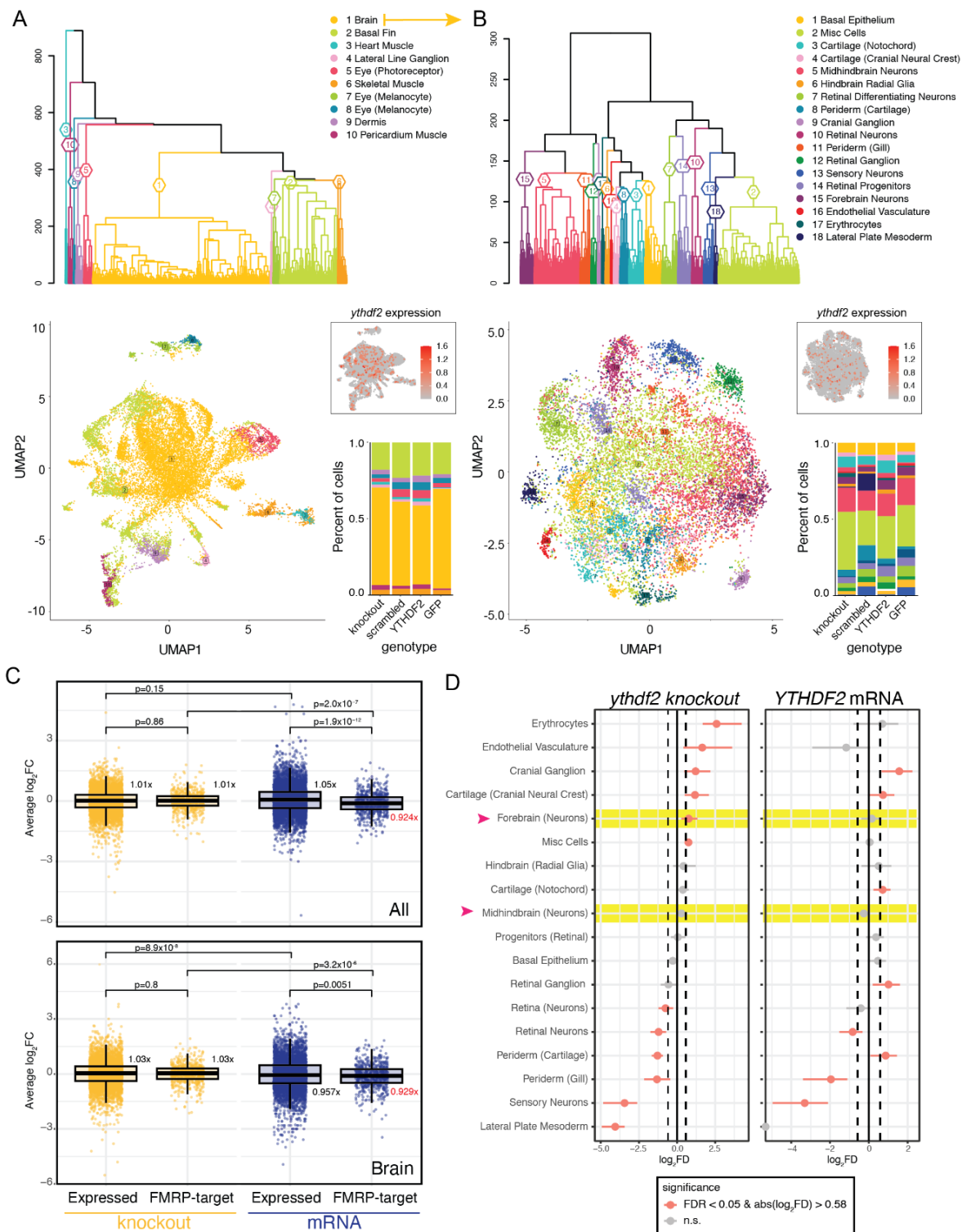
**Figure S3. ASD-DM and ASD-M families with *de novo* variants.** Families were from the Autism Phenome Project (APP) and the Simons Simplex Cohort (SSC) with impacted genes that were screened using zebrafish in this study. Green font signifies a duplication and red font signifies a deletion.



**Figure S4. Lateral VAST measurements of *ythdf2* knockout embryos at 3 dpf.** Measurements of total area (A), length (B), and head area (C) were significantly reduced (Wilcoxon T-test  $p$ -value  $< 0.00041$ ), with a significantly reduced head size maintained after normalization by area (D) (Wilcoxon T-test  $p$ -value = 0.032) and length (E) (Wilcoxon T-test  $p$ -value = 0.044). No significant change noted in head trunk angle (F) (Wilcoxon T-test  $p$ -value = 0.56). These results are consistent with overall smaller embryos and microcephaly, but not developmental delay.



**Figure S5. Gene expression analysis of *ythdf2* in crispant mutants.** RT-qPCR results for across *ythdf2* knockout (KO) and scrambled control biological replicates. All samples demonstrate consistent *act1b* expression, while the pooled *ythdf2* KO samples demonstrate lower *ythdf2* expression than the scrambled control. *p*-values were calculated using a Student's T-test.



**Figure S6. Single-cell transcriptomic analysis of *YTHDF2* zebrafish models.** Dendrograms were created to cluster cells with similar expression profiles and visualized via a UMAP plot, with colors indicating assigned cluster IDs for (A) all cells and (B) brain cells assigned in A. The proportion of cells assigned to each cluster per model is a percentage of total cells indicated as a barplot. Cells with *ythdf2* transcripts are colored red based on a continuous scale of natural log normalized expression [1]. (C) The average log<sub>2</sub>FC per expressed gene (0.01% of cells) was plotted across all cells and brain cells (see Methods) for all and FMRP-target genes for the *ythdf2* knockout and *YTHDF2* mRNA zebrafish models. Comparisons were made using T-tests either paired (between models) or unpaired (within models). Median fold changes versus respective controls (scrambled for knockout and GFP for mRNA) are indicated next to plots. (D) Abundances of cells per brain subcluster were plotted as log<sub>2</sub> fold difference (FD) with error bars representing 95% confidence intervals and an FDR threshold of 0.05 (scProportionTest).





## SUPPLEMENTARY REFERENCES

1. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol.* 2024;42:293–304.
2. White RJ, Collins JE, Sealy IM, Wali N, Dooley CM, Digby Z, et al. A high-resolution mRNA expression time course of embryonic development in zebrafish [Internet]. Available from: <http://dx.doi.org/10.1101/107631>
3. Yang H, Luan Y, Liu T, Lee HJ, Fang L, Wang Y, et al. A map of cis-regulatory elements and 3D genome structures in zebrafish. *Nature.* 2020;588:337–43.
4. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369:1318–30.