Supplemental Methods

Exome and genome analysis

Exome or genome sequencing was performed with a variety of standard capture kits and the general assertion criteria for variant classification referred to the Online Mendelian Inheritance in Man (OMIM) database and gnomAD genomic dataset. Other candidate variants in these patients that survived filtration and analysis by using either dominant or recessive models were thoroughly analyzed with their gene and domain functions, gene expression pattern, and genomic constraint feature. It was found that no other variants could better explain the phenotypes (Supplemental Table 8).

RNASeq and analysis

RNA was extracted using the RNeasy Mini Kit (Qiagen) for all the fly or hPSC samples each at the same time to minimize the batch effect. Extracted RNA samples underwent quality control assessment using Bioanalyzer (Agilent, Santa Clara, CA, USA) and were quantified using NanoDrop Technologies (Wilmington, DE, USA). RNA libraries were prepared using the polyA method and were sequenced using NovaSeq 6000 sequencer (Illumina, Inc., San Diego, CA, USA) at Center for Applied Genomics at the Children's Hospital of Philadelphia per standard protocols (paired-end 100 bp). The RNA-Seq data were aligned on the BDGP6 fly or GRCh38 human reference genome using HISAT2 v2.1.0 (1), transcripts were assembled using StringTie v2.0 (2), and features were counted by featureCounts (3). DESeq2 was then used to detect differentially expressed genes(4). DEXSeq (5), limma/diffSplice (6), and JunctionSeq (7) were applied to identify differentially used exons. Since *Prp19* RNAi with 77% reduction is more efficient than *U2af50* (49.0%), we applied different thresholds for differential gene expression and exon usage fold changes (FC), $|\log_2 FC| \geqslant 1$ and $P_{adj} \leqslant 0.01$ for *Prp19;* and $|\log_2 FC| > 0.7$ and $P_{adj} \leqslant 0.01$ for *U2af50*, in order to get comparable numbers of significant genes.

pyPY minigene assay

For transfection experiments, HEK293T cells were seeded into 24-well plates (125,000 cells per well) and grown as monolayers in MEM supplemented with 10% (v/v) of heat-inactivated fetal bovine serum. After one day, the cells were transiently transfected with either 50 ng of *pyPY* plasmid or a mixture of 250 ng of U2AF2 variant and 50 ng of *pyPY* plasmid per well using appropriately adjusted FugeneHD ratio according to the manufacturer's instructions (Promega). For reverse transcription PCR (RT-PCR), the total RNA was isolated two days post transfection using the RNeasy Plus kit (Qiagen). cDNA was synthesized using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosciences). The RT-PCR reaction comprised 30 cycles (98 °C for 10 s, 69 °C for 30 s, and 72 °C for 20 s) with forward (5'-TGAGGGGAGGTGAATGAGGAG-3') and reverse (5'-TCCACTGGAAAGACCGCGAAG-3') primers for the *pyPY* product. The RT-PCR products were separated by 2% agarose gel electrophoresis and stained with ethidium bromide.

Co-immunoprecipitation

500,000 HEK293T cultured in 6-well plates for 24 h were transfected with 2 μg plasmid encoding PRPF19 with an N-terminal FLAG tag using FugeneHD as per manufacturer's instructions. A mock transfection excluding the plasmid was included as a negative control. 48 h post-transfection, cells were lysed using an immunoprecipitation buffer (25 mM Tris-HCl, 150

mM NaCl, 1% (v/v) NP-40, 1% (w/v) sodium deoxycholate) supplemented with a protease inhibitor cocktail (Roche). Samples of cell extract were both reserved as "inputs" and incubated with DynaBeads Protein G (Invitrogen) having immobilized an anti-FLAG antibody for capturing PRPF19 complexes. Co-immunoprecipitation was carried out following the instructions provided by the manufacturer. Proteins were eluted from the beads using RIPA buffer and analyzed by Western blot for PRPF19 and CDC5L. The following antibodies were used: PRPF19 antibody (1:1000, Invitrogen, #PA5-88644) and CDC5L antibody (1:1000, #12974-1-AP, ProteinTech, #12974-1-AP).

Immunohistochemistry

3rd instar larvae were dissected, and their brains were subjected to immunostaining. The following antibodies were used: mouse anti-FasII antibody (1:100, DHSB, 1D4), rat anti-phospho Histone3 antibody (1:400, abcam, ab10543), and fluorescence-conjugated secondary antibodies (1:1000, Jackson ImmunoResearch).

Fly stocks

The *w^{1118}, UAS-U2af50 RNAiBL27542, UAS-U2af50 RNAiv24176, UAS-U2af50 RNAiBL55153, UAS-Prp19 RNAiBL32865, UAS-Prp19 RNAiBL41438, UAS-CG13287, UAS-RpS19a, elav-Gal4, D42-Gal4, UAS-Dcr2, OK107-Gal4* and *UAS-mCD8::GFP* stocks were obtained from Bloomington *Drosophila* Stock Center, FlyORF or VDRC. *Ppk-CD4tdGFP* and *ppk-Gal4* have been previously described (8). *UAS-Rbfox1* was a generous gift from Michael Buszczak. *UAS-Iswi* was a gift from Matthew Kayser. To generate the *UAS-$U2AF2^{WT}$, UAS- $U2AF2^{Arg149Trp}$, UAS- $U2AF2^{Arg150Cys}$, UAS-$PRPF19^{WT}$, UAS-$PRPF19^{Gly404Ser}$, UAS-$PPRF19^{His273Thrfs*37}$* and *UAS-$PPRF19^{Leu499Phe}$* stocks, the coding sequences were cloned from patients and construct into a pACU2 vector, then injected to fly eggs (Rainbow Transgenic Flies, Inc).

RT-qPCR

RT-qPCR was performed as previously described (9). For each sample, 30-40 WT or pan-neuronal knockdown larval brains were collected, with three biological replicates for each genotype. RNA was extracted with TRIZOL (Ambion, 15596026), and reverse transcription was done with kit. The mRNA level of *RpS19a* was normalized to *rp49*. For *Iswi, Brm,* and *Rbfox1*, since their overall mRNA level were slightly altered in *U2af50* or *Prp19* knockdown flies as shown by RNASeq data, the expression level of the area of interest were normalized to their own coding exons (ctrl primers). The mRNA level of the tested genes in *U2af50* or *Prp19* knockdown brains was normalized to that in WT. The primers are listed as follows:

rp49
F: CAGTCGGATCGATATGCTAAGCTG R: TAACCGATGTTGGGCATCAGATAC
Iswi ctrl
F: ACACCCGTAACACCTTCATTAG R: GAATACAGACTTCTCGCGGATAC
Iswi
F: ACTAACAGGTTGCACACTCG R: CGGCGTCTGTCTCTAAATGAA
Brm ctrl
F: CACGAGACGAGATACCACTTTG R: CCTCTAGCTTCTGTTGCTTCTC
Brm
F: GGTAATACTTTCGGTGCCAAAT R: CGGCTAAAGATACGCCAAATC
RpS19a

F: GTCTGTGAGAGAATCGTGGAAA R: GCGTGCTGGTCAATATCCT

Rbfox1 ctrl

F: CTACGCCAGTCAGCCAAATA R: CTACGCCAGTCAGCCAAATA

Rbfox1

F: ATCACGTCGGCATTTCGTAG R: TCCGTTAGGAAGCACAGTTATTT

Heat-induced seizure

Heat induced paralysis assays were performed as described (10). Briefly, 10 flies were transferred to a transparent fly vial and the whole vial was submerged in 39 °C water. The number of immobile flies was recorded in 30-s intervals and the assay continued for a total of 6 min.

Social space assay

Social space assay was performed as described (11). In brief, 5-7 days post-eclosion, ~40 virgin females were aspirated into the social space arena. The flies were given 20 min to freely investigate the arena and then images were captured. The distance between any of the two flies was measured by Image J and normalized to the average body length of flies in pixels.

Negative geotaxis assay

10-15 flies were gently transferred to an empty vial and allowed 2 min to habituate. The flies were then tapped to the bottom and were given 10 seconds to climb up. The percentage of flies climbing across the 8 cm line was recorded.

Larval locomotion assay

In order to track locomotion activity of fly larvae, we utilized a well-characterized program derived from *C. elegans* experiments named wrmTrck (12). Third-instar larvae were placed into apple juice plates and allowed to acclimate for approximately thirty seconds. Afterwards, the larvae were recorded for sixty seconds utilizing a video-capturing device. Ultimately, we were left with several quantifiable parameters, including distance, length traveled, and speed. In addition to quantification of locomotion parameters, we can also qualitatively analyze the larval movement over the experimental session.

Pluripotent stem cell maintenance

H9 human pluripotent stem cells (hPSCs) (WiCell WA09) and edited lines were cultured on Matrigel (Corning, 354277) coated plates in mTeSR1 media (STEMCELL Technologies, # 85850). The medium was changed every day and cells were passaged when they reached 70% confluence, approximately every 5-7 days, using ReleSR (STEMCELL Technologies, #100-0484) at a 1:30 ratio.

Generation of genetically modified H9 hPSCs by CRISPR/Cas9

The sgRNA targeting the mutation area was synthesized by Integrated DNA Technologies (IDT) and cloned into the PX458 vector (Addgene, #48138). Single-stranded DNA oligonucleotides (ssODNs) were also synthesized by IDT. The plasmid and ssODNs were introduced into H9 cells by electroporation with Human Stem Cell Nucleofector Kit 1 (Lonza, #VPH-5012) and Amaxa Nucleofector II device, following the manufacturer's protocol. Specifically, to generate heterozygous clones, 3.5 μg plasmid and 2 μg ssODN (1 μg each template) were diluted in 100

μL transfection solution and mixed with 1.2×10^6 singularized cells. B-016 program was selected on the device. Single GFP-positive cells were later sorted with BD FACSAria™ Fusion Flow Cytometer. Single colonies from single cells were genotyped by Sanger sequencing with the primers listed below.

sgRNA
5'GTTGCCCACGTAGAGACACC3'
Silent template
5'GCTGTGACCCCAACGCCGGTGCCCGTGGTCGGGAGCCAGATGACCAGACAAGCTC
GGCGTCTCTACGTGGGCAACATCCCCTTTGGCATCACTGAGGTACTGCCCTCC3'
Template_149 mutation
5'GCTGTGACCCCAACGCCGGTGCCCGTGGTCGGGAGCCAGATGACCAGACAAGCT[
T]GGCGTCTCTACGTGGGCAACATCCCCTTTGGCATCACTGAGGTACTGCCCTCC3'
Template_150 mutation
5'GCTGTGACCCCAACGCCGGTGCCCGTGGTCGGGAGCCAGATGACCAGACAAGCTC
GG[T]GTCTCTACGTGGGCAACATCCCCTTTGGCATCACTGAGGTACTGCCCTCC3'
Forward sequencing primer
5'ACCATGACCCCTGACGGTCTGG3'
Reverse sequencing primer
5'TGGCTCAACTCCATCGGTCCGT3'

Lentiviral infection and neural induction
UbC-rtTA-Ngn2:2A:EGFP (plasmid from Addgene, #127288) was packaged into lentiviral capsids and was collected in mTeSR1. The isogenic hESCs were singularized with accutase and infected with the viral media by mixing 0.2×10^6 hESCs with 100 μL virus-containing mTeSR1, 5 μL LentiBlast (OZ Biosciences, LBPX500), and 10 μM Rho-associated protein kinase (ROCK) inhibitor (Y-27632, Tocris, Cat. No.1254). Virus-containing media was removed after 24 h and replaced with fresh mTeSR1. Starting from 48 h after infection, 0.3-1 μg/mL puromycin was added to mTeSR1. Colonies that survived the selection were expanded.

Starting from 3 days prior to differentiation, stem cells were maintained in media containing the following: knockout DMEM/F12 (GIBCO, 12660012) as the base, supplemented with 1× MEM Non-Essential Amino Acids (GIBCO, 11140050), 1× N2 (GIBCO, 17502048), 10 ng/mL NT-3 (PeproTech, 450-03), 10 ng/mL BDNF (R&D Systems, 248-BD/CF), 0.2 μg/mL Mouse Laminin (GIBCO, 23017015), and 2 μg/mL doxycycline hydrochloride (Sigma-Aldrich, D3447) to induce the expression of NGN2. On Day 0 of differentiation, stem cells were counted and plated at 1×10^5 per coverslip in the same media, supplemented with 10 μM ROCK inhibitor. Change the media the second day to remove ROCK inhibitor. On Day 2, one set of the cells were fixed with 4% paraformaldehyde for 20 min, rinsed twice with PBS, and stored at 4˚C for further neurite growth immunostaining. For the rest of the cells, on Day 2, the media was fully changed to the same as the previous day. On Day 3, the media was fully changed to the one made with the following recipe: half DMEM/F12 (GIBCO, 11320033) and half Neurobasal-A (GIBCO, 10888022) as the base, supplemented with 1× MEM Non-Essential Amino Acids, 0.5× GlutaMAX (GIBCO, 35050061), 0.5× N2, 0.5× B27 (GIBCO, 17504044), 10 ng/mL NT-3, 10 ng/mL BDNF, 1 μg/mL Mouse Laminin, and 2 μg/mL doxycycline hydrochloride. In addition, 4 μM Arabinocytidine hydrochloride (Ara-C, Sigma-Aldrich, C6645) was added to the media to

reduce the proliferating non-neuronal cells. On Day 4, the media was fully changed to remove Ara-C. On Day 10, the media was fully changed to remove doxycycline. On Day 15, the cells were fixed with 4% paraformaldehyde for 20 min, rinsed twice with PBS, and stored at 4˚C for further immunostaining of U2AF2.

Immunostaining and neurite length quantification of the induced neurons
The coverslips with induced neurons were permeabilized and blocked in blocking buffer (5% goat serum (Sigma-Aldrich, G9023) and 0.1% Triton X-100 in PBS) for 1 h at room temperature, and incubated overnight at 4 °C with anti-MAP2 (1:1000, Abcam, ab5932) and anti-TUJ1 (1:1000, Abcam, ab18207) or anti-U2AF2 (1:50, Sigma-Aldrich, U4758) diluted in the blocking buffer. The next day, cells were incubated for 1h at room temperature with 1:500 of Alexa Fluor 488 and 555-conjugated secondary antibodies and 0.1 μg/mL Hoechst 33342, then mounted with Mowiol mounting medium (sigma #81381). Images were acquired with Leica SP8 confocal microscope with Leica LAS X software and analyzed with FIJI ImageJ (NIH).

TUJ1-positive neurites were quantified as followed. First, TUJ1-positive trace images were generated with ImageJ by selecting Process > Make Binary > Skeletonize. Then, the total length of each image was measured by selecting Analyze > Measure > Area. The numbers of the total area are equal to the numbers of the total length, as the width of the skeleton lines was 1px. The absolute total neurite length per image was then calculated using the embedded scale bar. The average neurite length per cell was calculated by dividing the absolute total neurite length per image by cell number (Hoechst positive dots) of the image.

Data from each genotype was collected from 2 independent differentiations of 2 separate clones (4 replicates). 200-750 cells from each replicate were quantified. Statistical analyses were performed in Graphpad Prism 9.

TrkB-BAC minigene-expressing cell line and immunoblot analysis for RBFOX1 variants
The mouse BAC clone RP23-424E11 (from Bacpac Genomics) containing 194 kb of the Ntrk2 (TrkB) genomic locus from exon 11 (encoding the transmembrane region) to exon 14 (corresponding to the second exon of the kinase domain), was modified using the BAC manipulation system as described in Sharan et al. to create a functioning minigene that can express both the truncated and full-length transcripts by alternative splicing (13). Briefly, the synthetic CAG promoter and the TrkB extracellular domain cDNA sequence from the ATG to exon 10 was fused in frame with the BAC exon 11. The TrkB kinase domain cDNA from exon 15 to the STOP codon was fused in frame with the downstream BAC exon 14 followed by a pGKneo cassette. The TrkB-BAC minigene was transfected into HEK293 cells using Targefect (Targeting Systems, Targefect-BAC) as instructed by the manufacturer to generate a minigene-expressing stable cell line by selection with G418 Sulfate at 300 μg/mL (ThermoFisher Scientific 10131035). Cells expressing the TrkB-BAC minigene were transiently transfected (48h) using expression vectors (pcDNA3; Genewiz) containing the following cDNA: mouse Rbfox1 (NM_001359723.1), murine mutant F158A, human RBFOX1 (NM_018723.4), human mutants R118Q, G154C and E187K. Transfections were done by using X-tremeGENE 9 DNA Transfection Reagent (6365779001, Millipore-Sigma). Cells were lysed using RIPA lysis buffer (20-188, Millipore-Sigma) and incubated 30 min at 4°C. After the incubation, lysates were centrifuged at top speed (13,000 rpm) using a table-top centrifuge at 4°C. Supernatants were

collected and transferred into new tubes. The total amount of protein was then quantified using BCA assay (23225, ThermoFisher Scientific) and samples were prepared in order to have the same amount of total protein before adding 2 × Laemmli sample buffer (S3401, Sigma-Aldrich). Samples were heated at 95°C for 5 min (protein denaturation step) before being loaded in 4–12% NuPAGE precast gels for Western analysis (ThermoFisher Scientific). After being transferred to PVDF membranes (LC2005, ThermoFisher Scientific), blots were blocked in 5% non-fat milk in TBS-Tween (0.1%) and incubated overnight at 4°C with specific antibodies. Primary antibodies were anti-panTrk (Abcam ab76291) used for detecting TrkB.FL - against the intracellular kinase-domain of Trk, anti-Rbfox1 (MABE985, clone 1D10, Millipore-Sigma) and anti GAPDH (MAB374; Millipore). After incubation with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibodies (Millipore), membranes were incubated with enhanced chemiluminescent substrate (34076, ThermoFisher Scientific) for detection of HRP enzyme activity and visualized/quantified in a Syngene gel documentation system (GeneSys).

GestaltMatcher analysis

We utilized the next-generation phenotyping approach, GestaltMatcher (14), to analyze the facial phenotypes of *U2AF2* patients. GestaltMatcher first used deep convolutional neural networks that trained on 6,617 frontal images with 204 different disorders to learn the facial dysmorphic features. Later, GestaltMatcher encoded each photo into a 320-dimensional facial phenotype descriptor (FDP), and the FDPs spanned the Clinical Face Phenotype Space. In this space, the similarity between the two photos was quantified by the cosine distance. We then ranked the patients by sorting the cosine distance in descending order. To analyze 19 *U2AF2* patients, we compared them to 4,306 photos with 257 different disorders to simulate the real-world scenario. We then performed the pairwise rank analysis on *U2AF2* patients and reported the results in Supplemental Figure 1A. We further performed tSNE to project the 19 *U2AF2* patients to two-dimensional space and tried to cluster the patient regarding their mutations in Supplemental Figure 1B.

We generated the average face to investigate the difference between the *U2AF2* face and the face of healthy individuals. For each *U2AF2* patient, we randomly selected ten faces that matched the age and sex from UTKFace (https://arxiv.org/abs/1702.08423) to simulate the healthy faces. For both *U2AF2* and healthy faces, we detected the 68 facial landmarks to align the faces and later performed the average over the faces in each cohort.

Methods References

1. Kim D, et al. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357-60.

2. Pertea M, et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 2015;33(3):290-5.

3. Liao Y, et al. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923-30.

4. Love MI, et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

5. Anders S, et al. Detecting differential usage of exons from RNA-seq data. *Genome Res.* 2012;22(10):2008-17.

6. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.

7. Hartley SW, and Mullikin JC. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. *Nucleic Acids Res.* 2016;44(15):e127.

8. Han C, et al. Enhancer-driven membrane markers for analysis of nonautonomous mechanisms reveal neuron-glia interactions in Drosophila. *Proc Natl Acad Sci U S A.* 2011;108(23):9673-8.

9. Li F, et al. The Atr-Chek1 pathway inhibits axon regeneration in response to Piezo-dependent mechanosensation. *Nat Commun.* 2021;12(1):3845.

10. Hope KA, et al. Glial overexpression of Dube3a causes seizures and synaptic impairments in Drosophila concomitant with down regulation of the Na(+)/K(+) pump ATPalpha. *Neurobiol Dis.* 2017;108:238-48.

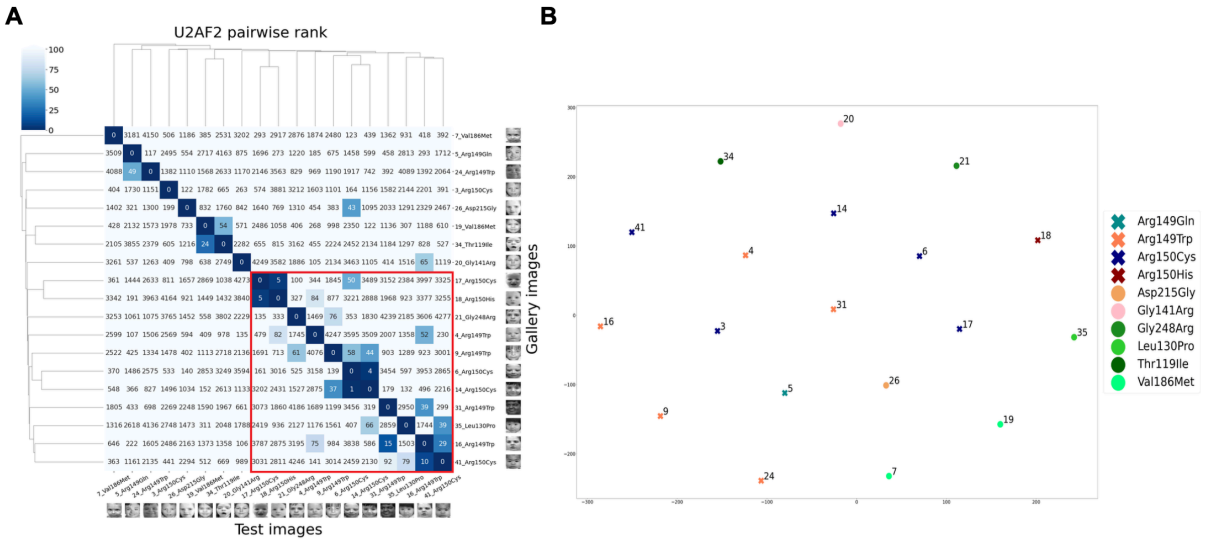
11. Gong NN, et al. The chromatin remodeler ISWI acts during Drosophila development to regulate adult sleep. *Sci Adv.* 2021;7(8).

12. Brooks DS, et al. Optimization of wrMTrck to monitor Drosophila larval locomotor activity. *J Insect Physiol.* 2016;93-94:11-7.

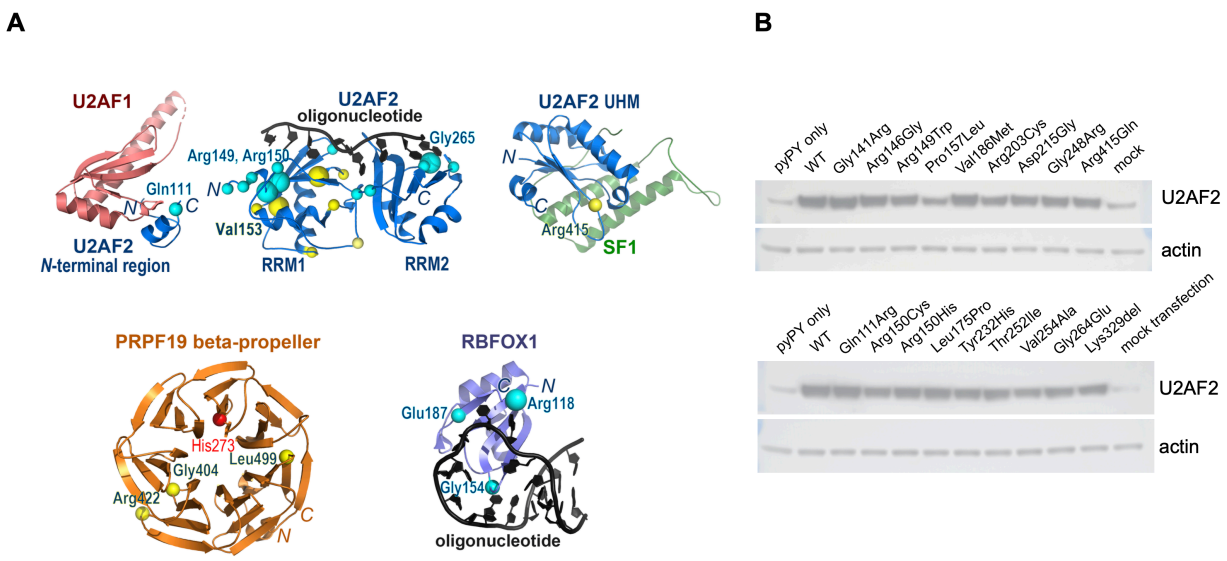
13. Sharan SK, et al. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc.* 2009;4(2):206-23.

14. Hsieh TC, et al. GestaltMatcher facilitates rare disease matching using facial phenotype descriptors. *Nat Genet.* 2022;54(3):349-57.

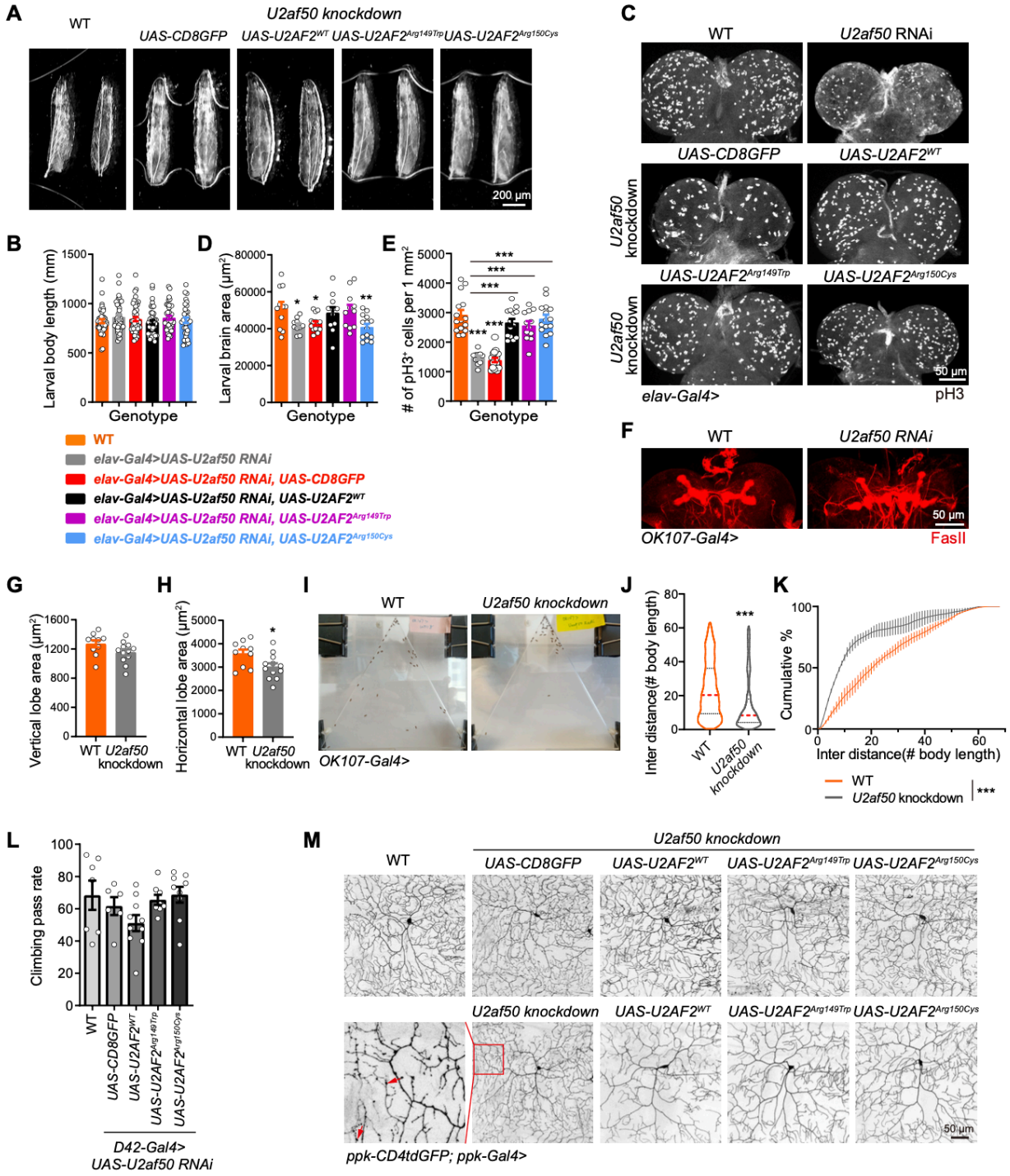
Supplemental Figures



Supplemental Figure 1. Clinical photographs of affected individuals with *U2AF2* variants demonstrating a pairwise rank comparison of 19 *U2AF2* patients to 4,306 facial photos. **(A)** Pairwise rank analysis on *U2AF2* patients. The labeling consists of individual numbering and variant correlate with those in Supplemental Table 1. For example, 41_Arg150Cys means individual 41 with Arg150Cys. Each column is a ranking result of testing the image in the column and trying to match the rest of patients in the rows. For example, column 6_Arg150Cys (the sixth column from the right) showed that individual 6 is at the first rank of individual 14. It means that individual 6 is the most similar to individual 14 compared to the other 4,306 photos with 257 different disorders in the database. The red box indicates a potential cluster for *U2AF2* patients. **(B)** tSNE visualization showing a trend of clustering of individuals with variants affected Arg149 or Arg150.

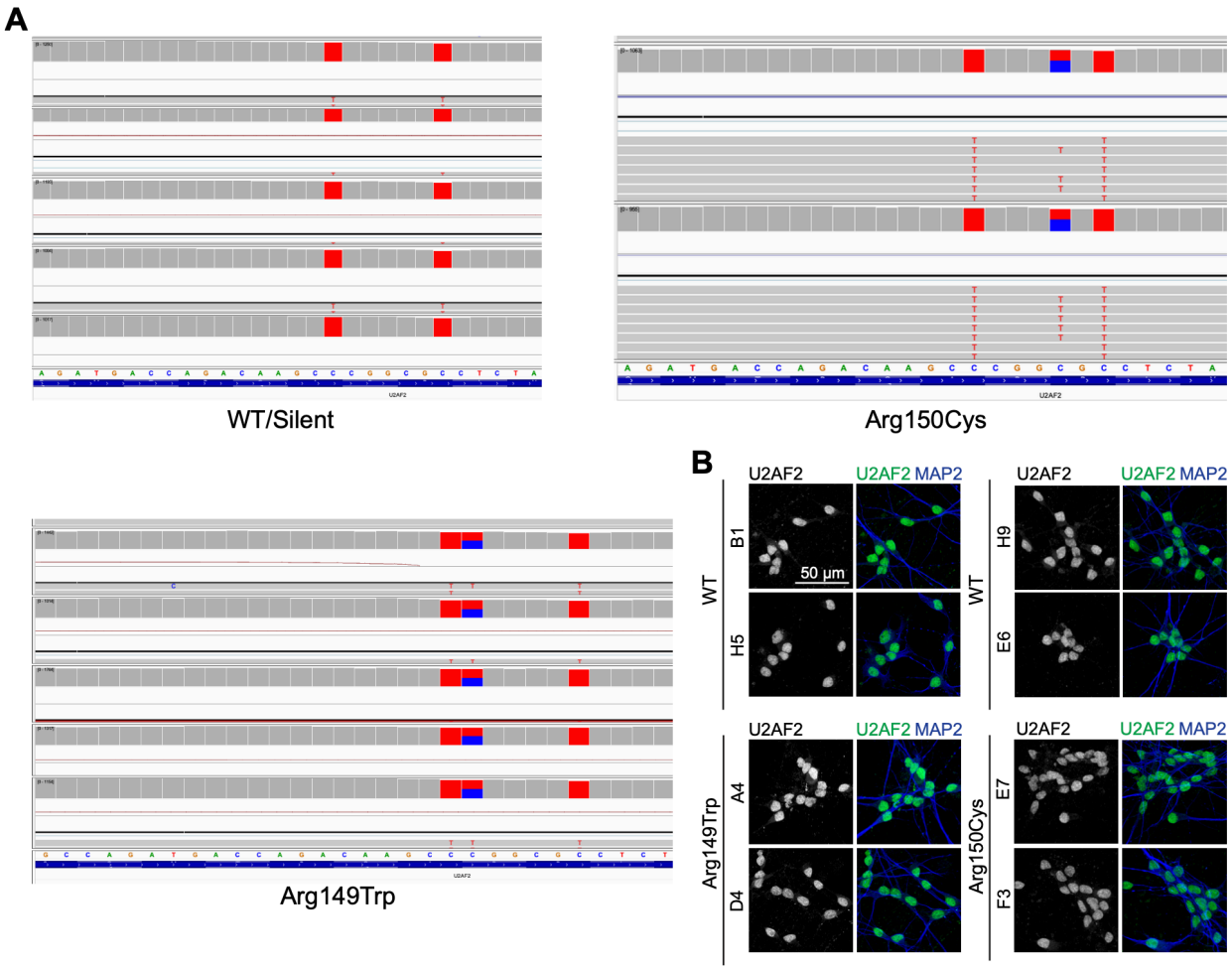


Supplemental Figure 2. Comparable relative abundance of U2AF2 WT and mutants and three-dimensional structure modeling of variants identified. (A) Representative western blot images from minigene splicing assay showed comparable U2AF2 expression levels. $N = 3$ independent experiments. **(B)** The neurodevelopmental disease-associated *U2AF2* variants mapped on available atomic resolution structures. Upper left panel: one mutated U2AF2 residue (Gln111Arg) in a U2AF2 N-terminal peptide (also called U2AF Ligand Motif or ULM, blue) bound to U2AF1 (salmon) (PDB: 1JMT). Upper middle panel: 19 mutated U2AF2 residues (Gly141Arg inclusive to Lys329del) in the U2AF2 RRM-containing region (also called U2AF2^{12L}, blue) bound to Py tract oligonucleotide (black) (PDB ID 6XLW). Upper right panel: one mutated U2AF2 residue (Arg415Gln for the predominant *U2AF2* isoform *b*, also called Arg419Gln using the isoform *a* numbering of the PDB coordinates) in the U2AF2 UHM (blue) bound to SF1 (green) (PDB: 4FXW). Two remaining mutated U2AF2 residues, which are located in the linker between the ULM and RRM1, are not present in available 3D structures. Lower left panel: three variants and one frameshift of PRPF19. Lower right panel: three variants of RBFOX1. The Cα-atoms of mutated residues are represented by spheres with enlarged radii marking the recurrently mutated residues (U2AF2 R149, R150, V153, V186 and G265; RBFOX1 R118). Mutants predicted to disrupt intermolecular interactions (with structurally-characterized RNA or protein interfaces) are colored cyan. Mutants predicted to disrupt local protein folding are colored yellow. The amino- (*N*) and carboxy- termini (*C*) of the U2AF2, PRPF19, and RBFOX1 domains are labeled in italicized font. The views of U2AF2 in upper panels are arranged to match the N-to-C-terminal orientation of the protein sequence.

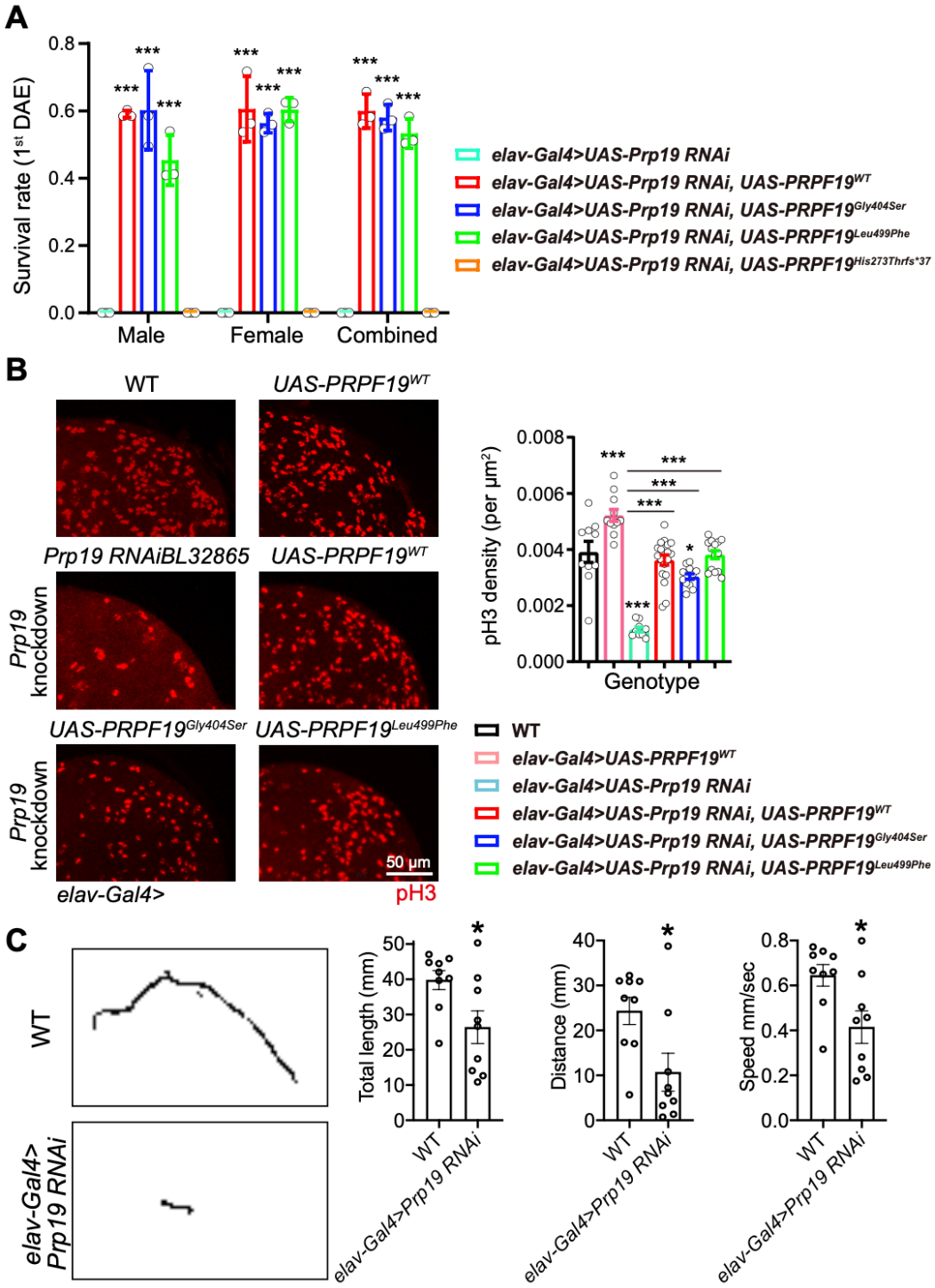


Supplemental Figure 3. Knocking down *U2af50* with different Gal4 drivers.
(A-E) Knocking down *U2af50* with *elav-Gal4* leads to smaller brain size and decreased proliferation without affecting the larval body length. (**A**) Representative images showing comparable 3rd instar larvae. Scale bar, 200 μm. (**B**) Quantification of the larval body length in (**A**). *n* = 33-48 larvae. (**C**) Representative images of larval brains. Scale bar, 50 μm. (**D**) Quantification of larval brain lobe area, *n* = 10-15. (**E**) Quantification of pH3 positive cells in the brain lobes, *n* = 10-17. Data are mean ± SEM, and analyzed by one-way ANOVA followed by

Dunnett's test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. **(F-K)** Mushroom body specific *U2af50* knockdown causes modest structural defects and impaired social performance. (**F**) Representative images showing mushroom body. Scale bar, 50 μm. (**G** and **H**) Quantification of the vertical and horizontal lobe area in (**F**). WT: $n = 10$, *U2af50:* $n = 12$. Data are mean ± SEM, and analyzed by two-tailed unpaired t-test. Vertical lobe: $P = 0.0805$, horizontal lobe: $P = 0.0225$. **(I)** Representative images showing *U2af50* knockdown virgin female exhibit reduced inter distance. $n = 3$ trials, with 40-41 flies for WT and 39-49 flies for *U2af50* knockdown. **(J** and **K)** Quantification of the distance between any of the two flies in one trial. Data are analyzed by two-tailed unpaired t-test (**J**) or two-way ANOVA followed by Sidak's test (**K**), ***$p < 0.001$. **(L)** The negative geotaxis assay shows that *U2af50* knockdown by the *D42-Gal4* driver has no gross locomotion deficits in flies. $n = 6$-10 groups, with 8-16 flies per group. ns: not significant. **(M)** Knocking down *U2af50* in dendritic arborization sensory neurons leads to degenerative-like morphology, which is reversed by the expression of human $U2AF2^{WT}$ or patient variants. Scale bar, 50 μm.



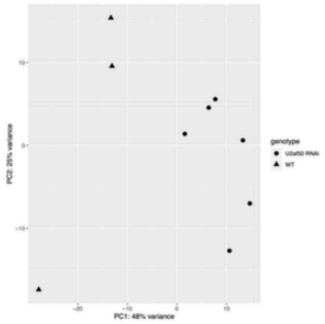
Supplemental Figure 4. RNA-Seq Integrative Genomics Viewer (IGV) view of knock-in silent and U2AF2 variants in RNA samples that were derived from hPSC colonies (A) and representative immunofluorescence images of induced neurons on Day 15 of U2AF2 and dendrite marker MAP2 demonstrated no localization change for two knock-in U2AF2 variants (B). Representative images. Scale bar, 50 µm.



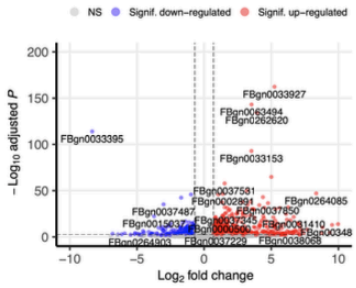
Supplemental Figure 5. *Prp19* knockdown and rescue study.
(A) Pan-neural *Prp19* knockdown is lethal to adult flies. While expressing of PRPF19^{WT}, PRPF19^{Gly404Ser}, or PPRF19^{Leu499Phe} increases the survival rate on the first day after eclosion (DAE), expressing PPRF19^{His273Thrfs*37} fails to rescue the lethality phenotype. n = 3 groups, with 44-91 flies per group for male and 40-101 flies for female. Data are mean ± SEM, and analyzed by two-way ANOVA followed by Turkey's test. **(B)** Pan-neural *Prp19* knockdown decreases

cell proliferation as revealed by pH3 immunostaining in the larval brain lobes with quantification on the right bar graph. Expression of $PRPF19^{WT}$, $PRPF19^{Gly404Ser}$, or $PPRF19^{Leu499Phe}$ fully or partially rescues the decreased proliferation. Notably, overexpressing $PRPF19^{WT}$ in the setting of WT significantly increases cell proliferation. n = 9-18. Data are mean ± SEM, and analyzed by one-way ANOVA followed by Dunnett's test. **(C),** Pan-neural *Prp19* knockdown larvae exhibit reduced motor activity. Left panel showed typical traces of WT and *Prp19* knockdown larvae. n = 9. Data are analyzed by two-tailed unpaired t-test with Welch's correction. $^{*}P < 0.05$, $^{***}P < 0.001$.

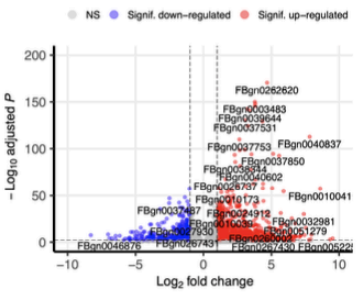
A



B



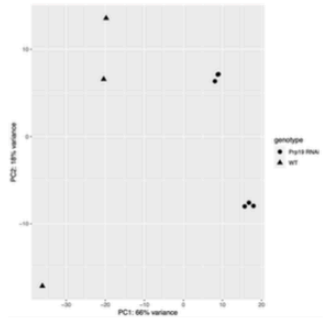
E



C

Term	Count	Fold Enrichment	P_Value	FDR
GO:0006357~regulation of transcription from RNA polymerase II promoter	62	4.0	1.9E-21	1.4E-18
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	21	4.1	1.7E-07	6.4E-05
GO:0045893~positive regulation of transcription, DNA-templated	14	4.3	2.1E-05	5.1E-03
GO:0007411~axon guidance	15	3.7	5.5E-05	9.0E-03
GO:0007406~negative regulation of neuroblast proliferation	6	13.4	6.0E-05	9.0E-03
GO:0031290~retinal ganglion cell axon guidance	5	17.7	1.3E-04	1.4E-02
GO:0030182~neuron differentiation	8	6.9	1.3E-04	1.4E-02
GO:0006355~regulation of transcription, DNA-templated	20	2.7	1.6E-04	1.5E-02
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	20	2.6	2.3E-04	1.9E-02

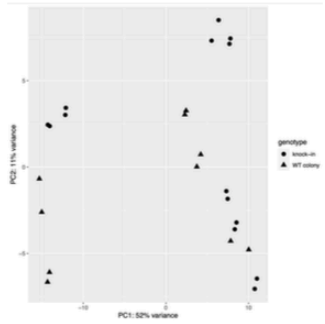
D



F

Term	Count	Fold Enrichment	P_Value	FDR
GO:0006357~regulation of transcription from RNA polymerase II promoter	54	3.6	9.3E-17	6.2E-14
GO:0030182~neuron differentiation	12	10.8	8.2E-09	2.7E-06
GO:0050767~regulation of neurogenesis	10	13.4	3.1E-08	6.8E-06
GO:0001708~cell fate specification	9	12.9	3.0E-07	4.9E-05
GO:0007219~Notch signaling pathway	9	8.3	1.0E-05	1.3E-03
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	16	3.3	1.1E-04	1.3E-02
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	20	2.7	1.4E-04	1.3E-02
GO:0001748~optic lobe placode development	4	29.5	2.1E-04	1.8E-02
GO:0045165~cell fate commitment	6	10.2	2.5E-04	1.8E-02
GO:1902692~regulation of neuroblast proliferation	5	14.8	2.8E-04	1.8E-02
GO:0009952~anterior/posterior pattern specification	6	9.8	3.0E-04	1.8E-02
GO:0007411~axon guidance	13	3.3	5.3E-04	2.9E-02
GO:0007155~cell adhesion	8	5.3	7.4E-04	3.7E-02

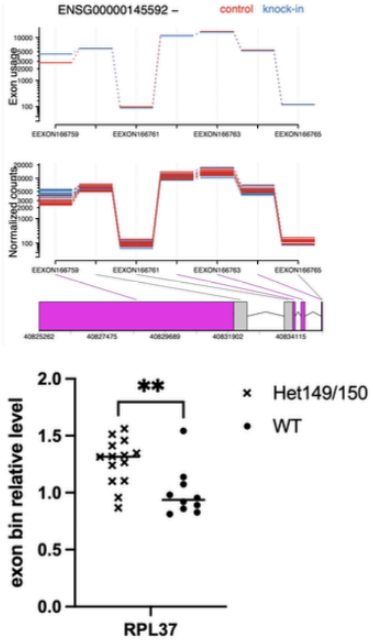
H



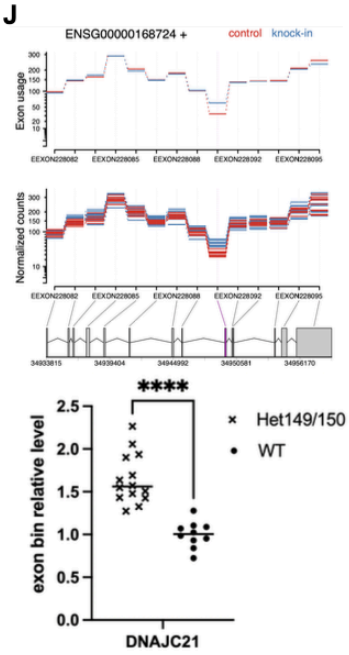
G

Term	Count	Fold Enrichment	P_Value	FDR
GO:0002181~cytoplasmic translation	32	5.2	1.9E-14	3.5E-11
GO:0006468~protein phosphorylation	43	3.5	9.1E-13	8.4E-10
GO:0007411~axon guidance	39	3.7	2.2E-12	1.4E-09
GO:0048813~dendrite morphogenesis	29	4.1	1.6E-10	7.6E-08
GO:0000381~regulation of alternative mRNA splicing, via spliceosome	24	4.5	1.3E-09	4.7E-07
GO:0016319~mushroom body development	21	4.2	5.3E-08	1.6E-05
GO:0035556~intracellular signal transduction	27	3.4	6.1E-08	1.6E-05
GO:0007409~axonogenesis	17	4.7	2.9E-07	6.6E-05
GO:0045879~negative regulation of smoothened signaling pathway	16	4.7	8.1E-07	1.7E-04
GO:0018105~peptidyl-serine phosphorylation	16	4.6	1.0E-06	1.9E-04
GO:0032436~positive regulation of proteasomal ubiquitin-dependent protein catabolic process	14	5.1	1.7E-06	2.9E-04
GO:1903688~positive regulation of border follicle cell migration	13	5.3	2.8E-06	4.3E-04
GO:0007391~dorsal closure	21	3.3	3.9E-06	5.5E-04

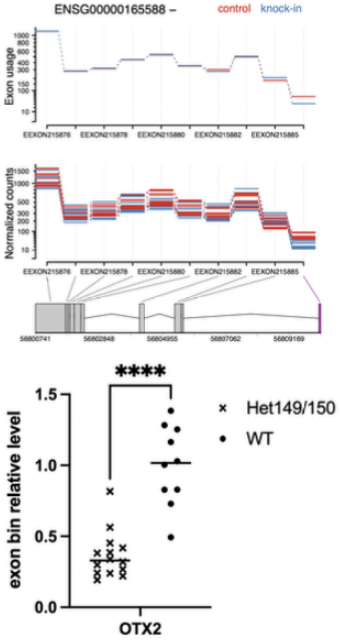
I



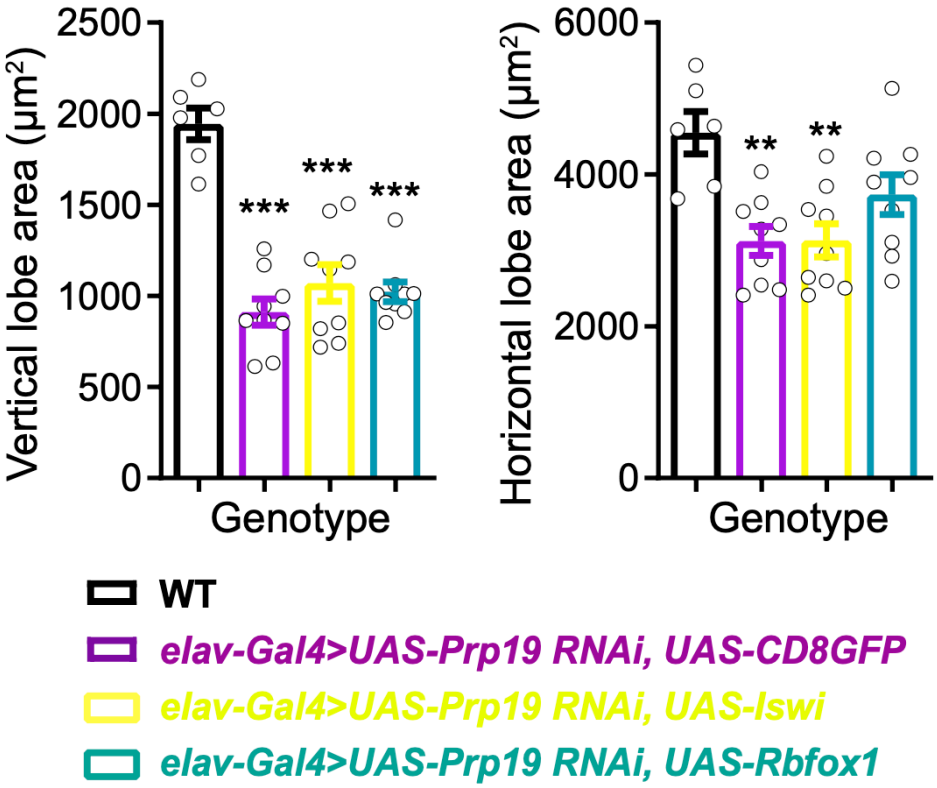
J



K



Supplemental Figure 6. Analysis of differentially expressed genes and differentially used exons from fly brain and hPSC RNASeq.
(A) Principal component analysis (PCA) of WT or RNAi of *U2af50* by using DESeq2 package. **(B)** Volcano plot of fly *U2af50* RNAi brain ($|\log_2 FC|>0.7$ and $P_{adj} \leq 0.01$). **(C)** GO enrichment analysis of downregulated genes in fly *U2af50* RNAi brain. **(D)** PCA of WT or RNAi of *Prp19* by using DESeq2 package. **(E)** Volcano plot of fly *Prp19* RNAi brain ($|\log_2 FC|\geq 1$ and $P_{adj} \leq 0.01$). **(F)** GO enrichment analysis of downregulated genes in fly *Prp19* RNAi brain. **(G)** GO enrichment analysis of genes with differentially used exon bin detected in fly *Prp19* RNAi brain. **(H)** PCA of knockin hPSC colonies. **(I-K)** Mis-splicing events identified in hPSC isogenic model and RT-qPCR confirmation. (**I**) A plot by plotDEXSeq showing differential exon usage for the first exon bin of *RPL37* (left panel) and RT-qPCR confirming the elevated exon usage in knock-in colonies. Purple triangle indicates the exon of interest. (**J**) A plot by plotDEXSeq showing differential exon usage for *DNAJC21* (left panel) and RT-qPCR confirming the elevated exon usage in knock-in colonies. (**K**) A plot by plotDEXSeq showing differential exon usage for *OTX2* (left panel) and RT-qPCR confirming the reduced exon usage in knock-in colonies. $^{**}P < 0.01$; $^{****}P < 0.0001$.



Supplemental Figure 7. Iswi and Rbfox1 fail to rescue the mushroom body structural defects in pan-neural *Prp19* knockdown. n = 6-9. Data are analyzed by one-way ANOVA followed by Dunnett's test, ***$P < 0.001$.