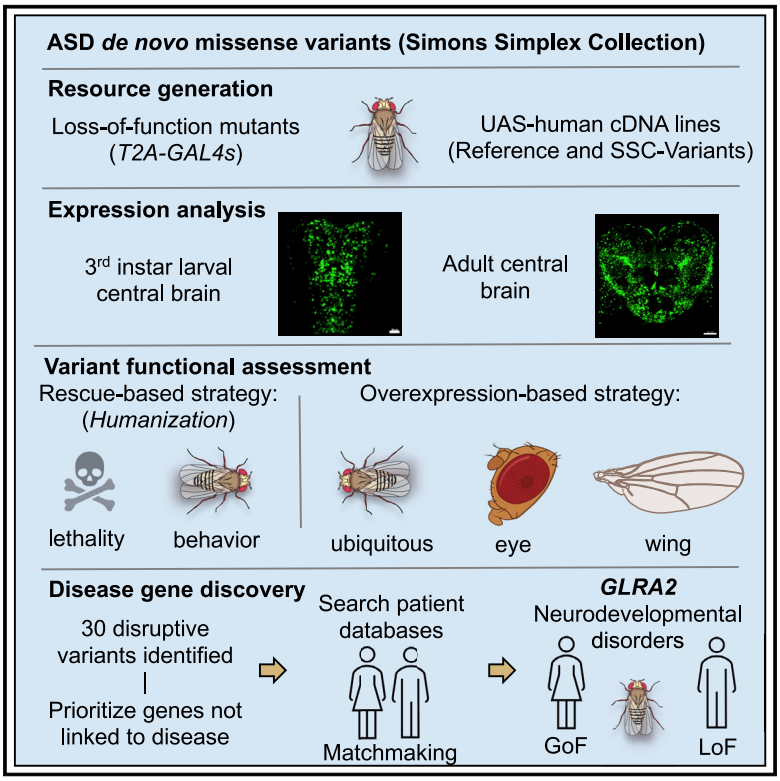
Cell Reports

Resource

Drosophila functional screening of *de novo* variants in autism uncovers damaging variants and facilitates discovery of rare neurodevelopmental diseases

Graphical abstract



Highlights

- We generate and characterize >300 (TG4 and cDNA) *Drosophila* mutants and transgenics
- Humanization and overexpression strategies to functionally assess ASD variants *in vivo*
- ASD variant data in flies help identify *GLRA2*-related neurodevelopmental disorders
- Basic and clinical collaboration facilitates variant testing and disease gene discovery

Authors

Paul C. Marcogliese, Samantha L. Deal, Jonathan Andrews, ..., Ronit Marom, Michael F. Wangler, Shinya Yamamoto

Correspondence

mw147467@bcm.edu (M.F.W.), yamamoto@bcm.edu (S.Y.)

In brief

Marcogliese et al. generate >300 *Drosophila* mutants and use complementary rescue-based and overexpression approaches to study the function of *de novo* missense variants found in autism. They find that 38% of missense changes have functional consequences and identify variants in *GLRA2* that cause a variable neurological disorder.

Marcogliese et al., 2022, Cell Reports *38*, 110517
March 15, 2022
https://doi.org/10.1016/j.celrep.2022.110517

CellPress

Cell Reports

CellPress
OPEN ACCESS

Resource

Drosophila functional screening of *de novo* variants in autism uncovers damaging variants and facilitates discovery of rare neurodevelopmental diseases

Paul C. Marcogliese,[1,2,37] Samantha L. Deal,[2,3,37] Jonathan Andrews,[1,2,37] J. Michael Harnish,[1,2,37] V. Hemanjani Bhavana,[1,2] Hillary K. Graves,[1,2] Sharayu Jangam,[1,2] Xi Luo,[1,2,4] Ning Liu,[1,2,36] Danqing Bei,[1,2] Yu-Hsin Chao,[1,2] Brooke Hull,[1,2] Pei-Tseng Lee,[1,2] Hongling Pan,[1,2] Pradnya Bhadane,[1,2] Mei-Chu Huang,[1,2] Colleen M. Longley,[2,3] Hsiao-Tuan Chao,[1,2,5,6,7,8,9] Hyung-lok Chung,[1,2,10] Nele A. Haelterman,[1,2] Oguz Kanca,[1,2] Sathiya N. Manivannan,[1,2] Linda Z. Rossetti,[1] Ryan J. German,[1,2] Amanda Gerard,[1,8] Eva Maria Christina Schwaibold,[11] Sarah Fehr,[12] Renzo Guerrini,[13] Annalisa Vetro,[13] Eleina England,[14] Chaya N. Murali,[1,8] Tahsin Stefan Barakat,[15]

(Author list continued on next page)

[1]Department of Molecular and Human Genetics, Baylor College of Medicine (BCM), Houston, TX 77030, USA
[2]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital (TCH), Houston, TX 77030, USA
[3]Program in Developmental Biology, BCM, Houston, TX 77030, USA
[4]Department of Pediatrics, Division of Hematology/Oncology, BCM, Houston, TX 77030, USA
[5]Department of Pediatrics, Division of Neurology and Developmental Neuroscience, BCM, Houston, TX 77030, USA
[6]Department of Neuroscience, BCM, Houston, TX 77030, USA
[7]McNair Medical Institute, The Robert and Janice McNair Foundation, Houston, TX 77030, USA
[8]TCH, Houston, TX 77030, USA
[9]Development, Disease Models & Therapeutics Graduate Program, BCM, Houston, TX 77030, USA
[10]Howard Hughes Medical Institute, Houston, TX 77030, USA
[11]Institute of Human Genetics, Heidelberg University, Heidelberg, Germany
[12]Praxis für Humangenetik Tübingen, Tübingen, Germany
[13]Neuroscience Department, Children's Hospital Meyer-University of Florence, Florence, Italy
[14]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[15]Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, the Netherlands
[16]Department of Medical Genetics, Lyon University Hospital, Université Claude Bernard Lyon 1, Lyon, France
[17]Institut NeuroMyoGène, CNRS UMR 5310 - INSERM U1217, Université Claude Bernard Lyon 1, Lyon, France
[18]Department of Pediatric Neurology, Lyon University Hospitals, Lyon, France
[19]Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA 02115, USA
[20]The Manton Center for Orphan Disease Research, Boston Children's Hospital, Boston, MA 02115, USA

(Affiliations continued on next page)

SUMMARY

Individuals with autism spectrum disorder (ASD) exhibit an increased burden of *de novo* mutations (DNMs) in a broadening range of genes. While these studies have implicated hundreds of genes in ASD pathogenesis, which DNMs cause functional consequences *in vivo* remains unclear. We functionally test the effects of ASD missense DNMs using *Drosophila* through "humanization" rescue and overexpression-based strategies. We examine 79 ASD variants in 74 genes identified in the Simons Simplex Collection and find 38% of them to cause functional alterations. Moreover, we identify *GLRA2* as the cause of a spectrum of neurodevelopmental phenotypes beyond ASD in 13 previously undiagnosed subjects. Functional characterization of variants in ASD candidate genes points to conserved neurobiological mechanisms and facilitates gene discovery for rare neurodevelopmental diseases.

INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition with impairments in social interaction, communication, and restricted interests or repetitive behaviors (APA, 2013). Individuals affected by ASD, particularly in severe cases, exhibit a higher burden of *de novo* mutations (DNMs) in an expanding list of genes (Coe et al., 2019; Fischbach and Lord, 2010; Iossifov et al., 2014). The genetic burden of DNMs in ASD subjects has been estimated to account for ~30% of

Marieke F. van Dooren,[15] Martina Wilke,[15] Marjon van Slegtenhorst,[15] Gaetan Lesca,[16,17] Isabelle Sabatier,[18] Nicolas Chatron,[16,17] Catherine A. Brownstein,[19,20,21] Jill A. Madden,[19,20] Pankaj B. Agrawal,[19,20,21,22] Boris Keren,[23] Thomas Courtin,[23] Laurence Perrin,[24] Melanie Brugger,[25] Timo Roser,[26] Steffen Leiz,[27] Frederic Tran Mau-Them,[28,29] Julian Delanne,[28] Elena Sukarova-Angelovska,[30] Slavica Trajkova,[31] Erik Rosenhahn,[32] Vincent Strehlow,[32] Konrad Platzer,[32] Roberto Keller,[33] Lisa Pavinato,[31,34] Alfredo Brusco,[31,35] Jill A. Rosenfeld,[1,36] Ronit Marom,[1,8] Michael F. Wangler,[1,2,8,9,*] and Shinya Yamamoto[1,2,3,6,9,38,*]

[21]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA
[22]Division of Newborn Medicine, Boston Children's Hospital, Boston, MA 02115, USA
[23]Genetic Department, Pitié-Salpêtrière Hospital, APHP.Sorbonne Université, Paris 75013, France
[24]Genetic Department, Robert Debré Hospital, APHP.Nord-Université de Paris, Paris 75019, France
[25]Institute of Human Genetics, Technical University Munich, Munich, Germany
[26]Division of Pediatric Neurology, Developmental Medicine and Social Pediatrics, Department of Pediatrics, Dr. von Hauner Children's Hospital, Ludwig-Maximilians-University, Lindwurmstraße 4, 80337 Munich, Germany
[27]Department of Pediatrics and Adolescent Medicine, Hospital Dritter Orden, Munich, Germany
[28]INSERM U1231, LNC UMR1231 GAD, Burgundy University, 21000 Dijon, France
[29]Laboratoire de Génétique, Innovation en Diagnostic Génomique des Maladies Rares UF6254, Plateau Technique de Biologie, CHU Dijon, 14 Rue Paul Gaffarel, BP 77908, 21079 Dijon, France
[30]Department of Endocrinology and Genetics, University Clinic for Children's Diseases, Medical Faculty, University Sv. Kiril i Metodij, Skopje, Republic of Macedonia
[31]Department of Medical Sciences, University of Torino, Turin, Italy
[32]Institute of Human Genetics, University of Leipzig Medical Center, Leipzig, Germany
[33]Adult Autism Center, Mental Health Department, Health Unit ASL Città di Torino, Turin, Italy
[34]Institute of Human Genetics and Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany
[35]Medical Genetics Unit, Città della Salute e della Scienza, University Hospital, Turin, Italy
[36]Baylor Genetics Laboratories, Houston, TX 77021, USA
[37]These authors contributed equally
[38]Lead contact
*Correspondence: mw147467@bcm.edu (M.F.W.), yamamoto@bcm.edu (S.Y.)
https://doi.org/10.1016/j.celrep.2022.110517

disease causation (Iossifov et al., 2014; Rubeis et al., 2014; Sanders et al., 2012; Satterstrom et al., 2020; Takata et al., 2018; Yuen et al., 2017). While these studies have implicated hundreds of genes in ASD pathogenesis, which of these genes and variants causally contribute to this disease remains unknown. Missense DNMs in particular present a unique challenge because most genes lack established functional assays. *Drosophila melanogaster* is a genetically tractable system that is widely used to study human diseases (Bellen et al., 2019; Link and Bellen, 2020; Marcogliese and Wangler, 2001). In addition to studying disease mechanisms by establishing preclinical models, flies can be used as a "living test tube" to study functional consequences of variants of unknown significance found in subjects. Here, we integrate a number of state-of-the-art technologies in the fly field to establish an *in vivo* pipeline to effectively study the functional impact of DNMs identified in a large cohort of ASD subjects.

RESULTS

Prioritization of ASD variants to study in *Drosophila*

We prioritized genes with coding DNMs identified in ASD probands from the Simons Simplex Collection (SSC) (Iossifov et al., 2014) that are conserved in *Drosophila*. We primarily focused on missense variants and in-frame indels because functional consequences of these variants are more difficult to predict compared with nonsense and frameshift alleles. However, we also tested a few truncating variants in single-exon genes because these transcripts escape nonsense-mediated decay. In this cohort, 1,708 ASD proband-specific *de novo* missense or in-frame indels were identified through whole-exome sequencing (WES) (Figure 1A), corresponding to 1,519 unique human genes. Of these, 920 fly genes corresponding to 1,032 human genes were identified. Based on multiple ortholog prediction algorithms scores, 487 human genes had no or weak ortholog candidates in *Drosophila* (cut off: DIOPT <4/16; Hu et al., 2011; Table S1). By overlapping these 920 *Drosophila* genes with available fly lines containing *Minos*-mediated integration cassette (MiMIC) transposons within coding introns that permit targeting of all annotated protein isoforms ("gold"; 1,732 insertions; Nagarkar-Jaiswal et al., 2015; Venken et al., 2011), we identified reagents for 122 fly genes corresponding to 143 human genes and 179 ASD proband variants from the SSC. Compared with the entire genome, the SCC subset and our study subset showed enrichment for constrained genes by assessing gene-level metrics, such as probability of loss of function intolerance (pLI) (Figure 1B; Lek et al., 2016), loss of function observed or expected (o/e) upper bound fraction (LOEUF) (Figure 1C; Karczewski et al., 2020), or missense o/e (Figure 1D; Karczewski et al., 2020).

Of the 122 fly genes that met our selection criterion, we were able to successfully generate 108 T2A-GAL4 (TG4) lines via recombinase-mediated cassette exchange (Diao et al., 2015; Gnerer et al., 2015; Lee et al., 2018). These 108 TG4 lines correspond to 128 SSC genes (some fly genes correspond to multiple human genes with variants in SSC; Figure S1). These fly lines behave as loss-of-function (LoF) alleles that simultaneously produce a GAL4 transactivator in the same temporal and spatial pattern as the gene of interest (Figure 1E; Tang et al., 2009).

CellPress
OPEN ACCESS

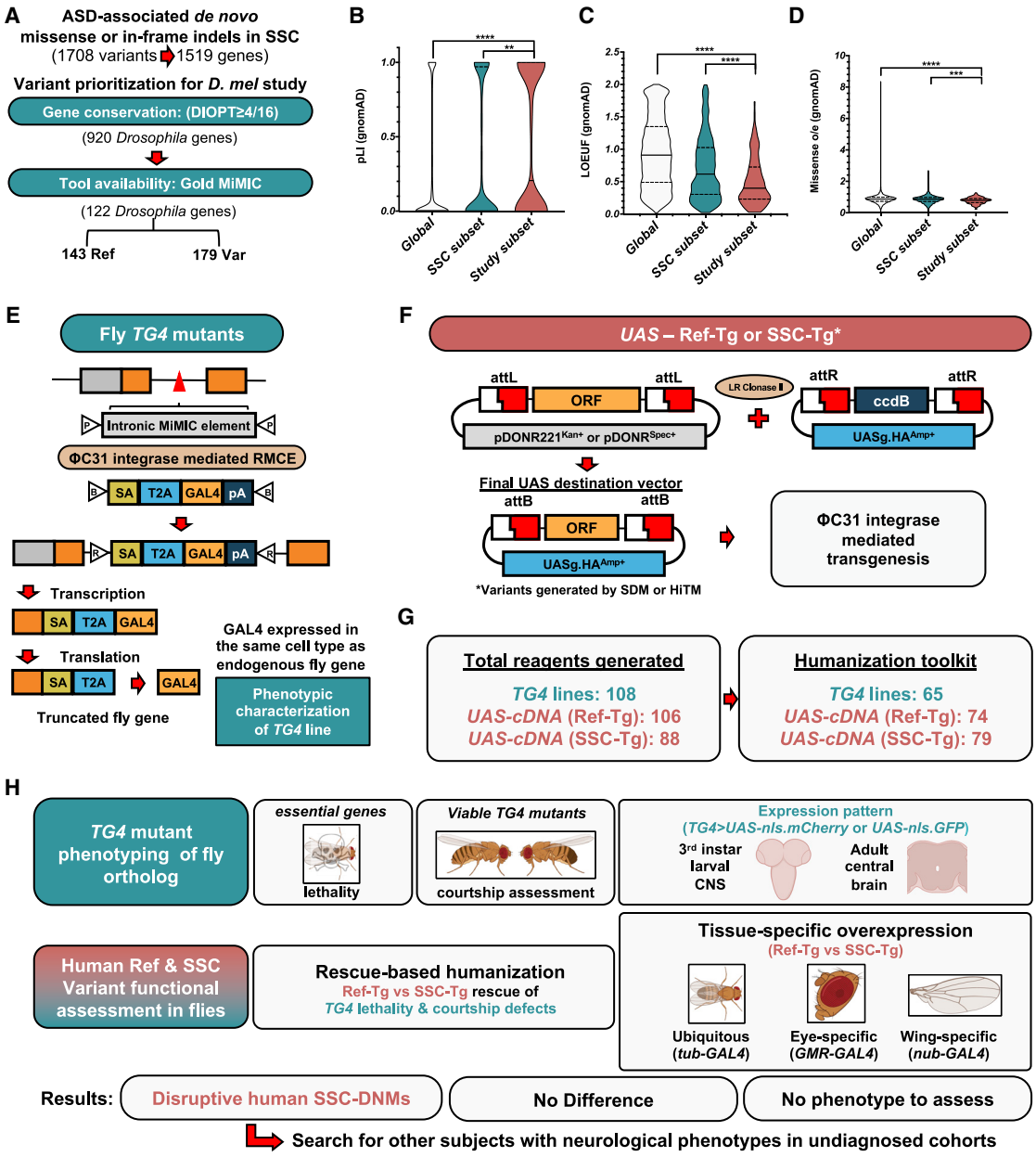


Figure 1. Gene and variant prioritization, resource generation, and screening outline

(A) Criteria to prioritize ASD candidate genes and variants for this study.

(B–D) Gene level constraints from control individuals (gnomAD) for (B) probability of loss of function intolerance (pLI), (C) loss of function observed or expected (o/e) upper bound fraction (LOEUF), and (D) missense o/e. **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

(E) Schematic depicting generation and effect of *TG4* lines on gene function.

(F) Schematic illustrating generation of UAS-human cDNA constructs.

(G) Total number of *Drosophila* reagents generated for this study.

(H) Screening paradigms using both humanization and overexpression strategies to assess SSC-DNM function.

To generate upstream activating sequence (UAS) human reference transgenic (Ref-Tg) and human SSC candidate variant transgenic (SSC-Tg) fly lines, we obtained human open reading frame (ORF) collections from the Mammalian Gene Collection (MGC Project Team et al., 2009) or commercial sources (Figure 1F). Out of the 143 human genes and 179 SSC-DNMs of interest that we attempted to generate, we were successful in generating 194 *UAS-cDNA* (106 Ref-Tgs; 88 SSC-Tgs) flies (Figures 1G and S1; Tables S2 and S3). The *UAS-Ref-Tg* and *UAS-SSC-Tg* were inserted into the same genomic docking site in the fly genome and are generated with the same construct, only differing in the variant point

mutation, allowing for direct comparison of function by controlling for positional effects.

We were able to make a complete set of *TG4*, *UAS-Ref-Tgs*, and *UAS-SSC-Tgs* lines for 65 fly genes corresponding to 74 human genes and 79 variants (again, some fly genes correspond to multiple human genes and multiple SSC variants are found for a small subset of human genes), which were critical to test variant function using a rescue-based humanization strategy. In summary, 302 *Drosophila* stocks were generated for this project as a resource for the community, and these stocks are available from the Bloomington *Drosophila* Stock Center (BDSC) or in the process of being transferred and registered at BDSC (please see Figure S1 for detailed screen pipeline data).

In order to determine whether there are functional differences between the Ref-Tg and SSC-Tg, we used them in combination with TG4 lines to "humanize" *Drosophila* genes or crossed these lines to ubiquitous and tissue-specific drivers to ectopically overexpress reference or variant human proteins and assessed them for phenotypic differences (Bellen et al., 2019; Figure 1H). If the SSC variant lacked a function a reference allele possessed, we classified it as a LoF allele (e.g., amorph and hypomorph). If the variant had some function that the reference allele did not possess, we classified it as a gain-of-function (GoF) allele (e.g., hypermorph, antimorph, and neomorph) in this study.

Humanization of essential *Drosophila* genes reveal loss-of-function ASD variants

We identified 47 of 65 lethal TG4 mutants that remained lethal when placed in *trans* with a corresponding deficiency line. These 47 TG4 lines correspond to 60 human ASD candidate genes for both reference and variant human cDNA transgenic fly lines (Figure S1; Table S4). To assess whether the human homolog can replace the corresponding fly genes, we determined whether *UAS-Ref-Tg* can rescue the lethality of lethal TG4 mutants. We assessed rescue at four temperatures (18°C, 22°C, 25°C, and 29°C) as the GAL4/UAS system is temperature dependent (Nagarkar-Jaiswal et al., 2015). We found that lethality was suppressed in 17 of 37 genes tested (46%; Figure 2A; Table S5). We next tested whether SSC-DNMs have functional consequences by comparing the rescue efficiency of *UAS-Ref-Tg* and *UAS-SSC-Tg*. We observed significant functional differences in the ability for SSC variants to rescue lethality for *ABL2*, *CAT*, *CHST2*, *TRPM6* (two variants), and *TRIP12* (Figures 2B–2D). For *ABL2* and *CAT*, we further found that humanized flies carrying the SSC-DNM ($ABL2^{A1099T}$ or CAT^{G204E}) had significantly decreased lifespan compared with reference animals (Figures 2E and 2F). Overall, we found that 32% (6/19) of the tested SSC-DNMs functionally differed from the reference *in vivo*, all behaving as LoF alleles.

To assess whether the fly homologs of human ASD candidate genes from SSC are expressed in the central nervous system (CNS), we crossed each TG4 line to UAS-nls.mCherry (red fluorescent protein with a nuclear localization signal) and performed co-staining with neuronal (Elav) and glial (Repo) nuclear markers. We chose the anterior central brain of the fly CNS to image as it is enriched for neuronal nuclei. All five genes associated with deleterious LoF DNMs were expressed in the adult central brain (Figures 2G and S2A). Examining neuronal and glia expression throughout the central brain revealed that all five genes were found in subsets of neurons and some glia, where *Abl* (corresponding to *ABL2*) and *ctrip* (*TRIP12*) are expressed in neurons and a wider set of glia (Figure S2B). All five genes are also expressed in the third-instar larval CNS (Figure S2C).

Humanization of viable *Drosophila TG4s* reveals ASD variants with altered function

While we were able to test the function of 37 human genes based on rescue of lethality as mentioned above, 61 TG4 lines corresponding to 68 SSC-ASD candidate genes were viable and did not exhibit any obvious morphological phenotypes that can be used for variant functional studies. Note that 18 of 61 TG4 lines were homozygous lethal but were viable as compound heterozygotes over a molecularly defined deficiency that covers the locus, indicating that these lines carry second site lethal mutations (Figure S1), which has been previously reported in a subset of MiMIC strains (Nagarkar-Jaiswal et al., 2015). For the 43 of 61 remaining viable TG4 lines, we attempted to generate humanized *TG4*; *UAS-Ref-Tg* and *TG4; UAS-SSC-Tg* lines in an appropriate genetic background. Before carrying out behavioral studies, we had to replace the X chromosome of the *TG4* lines with the X chromosome from a *Canton-S* strain to eliminate the effect of *yellow* (*y*) and *white* (*w*) alleles that were present in the original stocks (Figure S3). We were successful in generating all the necessary strains for eight genes. Using these humanized flies in a $y^+ w^+$ background, we performed courtship assays to assess social interactions in mutant and humanized flies considering that ASD patients exhibit social deficits (Figure 3A; Liu, 2013). Fly courtship involves a complex set of neurological components involving sensory input, processing, and motor output (Guo et al., 2019). We measured the amount of time a *TG4* male fly spent performing wing extensions as a proxy for courtship, as well as the amount of time spent copulating with a wild-type (*Canton-S*) female. In addition to quantifying these two parameters, we also measured the time flies spent moving within the test chamber to assess their locomotion and tracked grooming, a stereotypic behavior in flies that involves a complex neurocircuit (Seeds et al., 2014).

Of the eight SSC-DNMs tested, we found five variants that showed functional alterations from the reference allele in at least one of four behavioral paradigms (Figure 3B). Humanized $KCND3^{R86P}$ flies displayed increased movement and decreased grooming behavior when compared with the humanized reference flies (Figures 3C, S4A, and S4B). The humanized $KDM2A^{R449K}$ flies showed decreased time copulating compared with the humanized reference (Figures 3D and S4C–S4E). Humanized $USP30^{P200S}$ flies displayed decreased grooming behavior when compared with humanized reference flies (Figures 3E and S4F–S4H). While these variants have clear functional differences from the reference allele, it was difficult to classify them as clean LoF or GoF alleles. Humanized $ALDH1L1^{N900H}$ flies displayed a significant reduction in courtship and an increase in grooming behavior when compared with the humanized reference fly or the *TG4* mutant alone, potentially indicating the variant acts as some kind of GoF allele (Figures 3F, S4I, and S4J). Finally, humanized reference *GLRA2* flies failed to copulate at all but still exhibited normal movement while

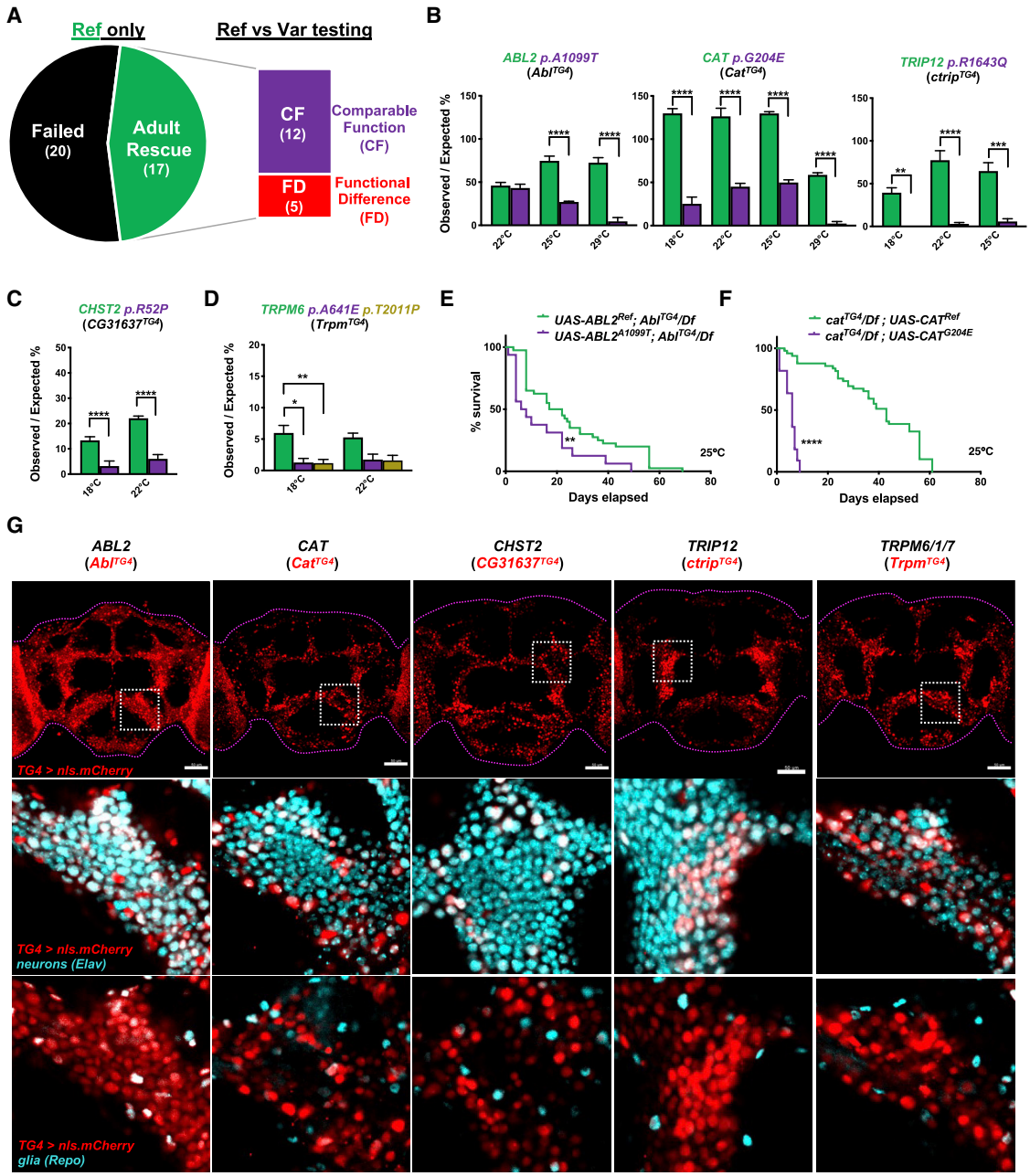


Figure 2. Assessment of SSC-DNM function through humanization of essential fly genes

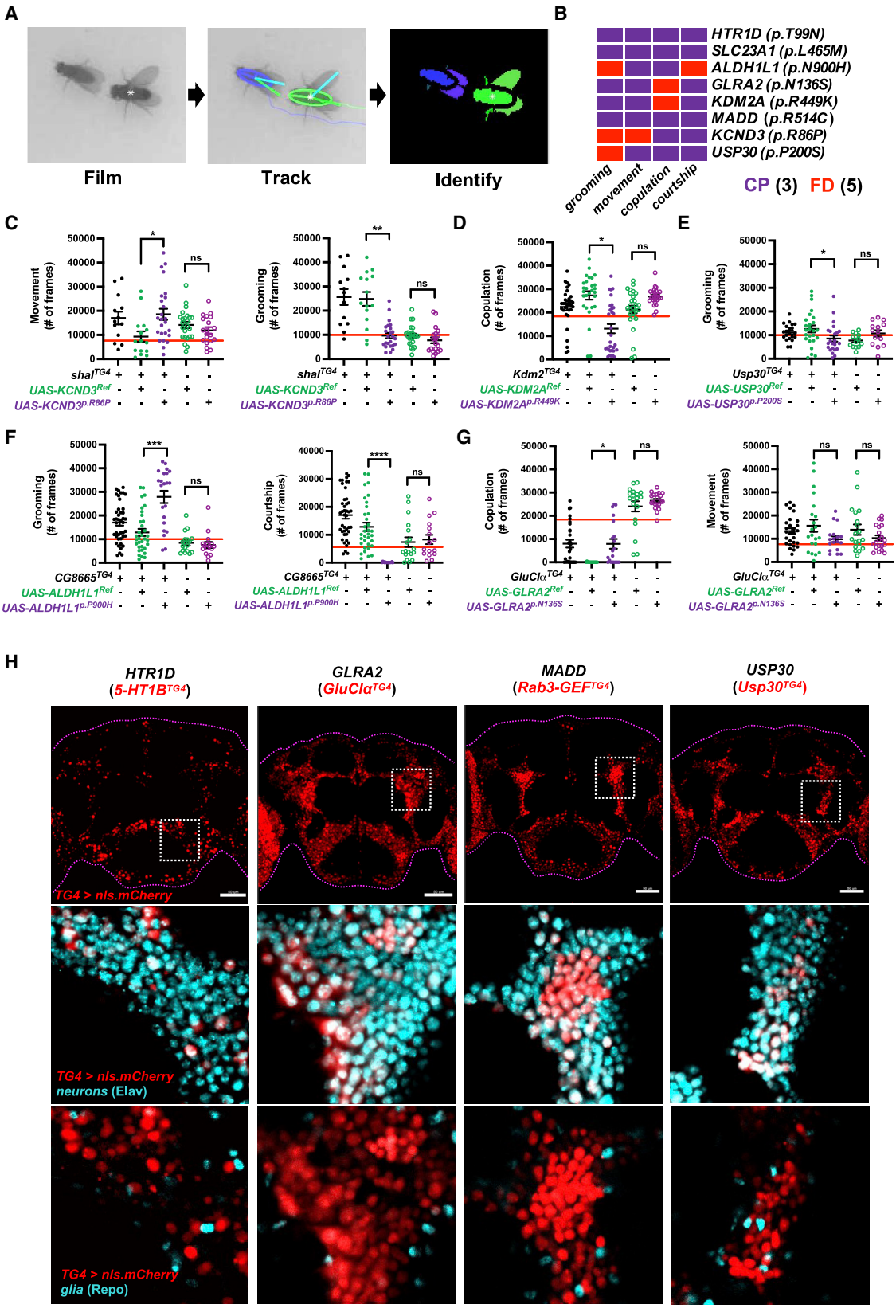
(A) Rescue of lethality to adult stage by *TG4*-driven UAS-reference human cDNA and subsequent comparison of reference and variant cDNA.

(B–D) Observed/expected Mendelian ratios for rescue of humanized *TG4* mutants across different temperatures. Three independent crosses were set per genotype, and n > 50 flies were quantified for each cross. Statistical analyses were performed by ANOVA followed by Sidak's multiple comparisons test. Error bars are +SEM (standard error of mean).

(E and F) Lifespan analysis of humanized *TG4* lines at 25°C. Survival comparisons obtained by log rank (Mantel-Cox) test with a minimum of 11–49 flies for each genotype from three independent crosses.

(G) Single focal confocal images showing expression pattern of *UAS-nls.mCherry* driven by *TG4* (red) in the anterior of the central adult brain. Bottom two rows depict the 5× zoom from dotted white box of co-localization between TG4 reporter and Elav (neurons) or Repo (glia) in cyan. Co-localization is depicted in white. Scale bars represent 50 μm. Dotted magenta lines outline of the brain. *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.

the humanized *GLRA2*^{N136S} flies were capable of copulating within the trial period similar to the *TG4* mutant alone, suggesting it behaves as a LoF allele (Figures 3G, S4K, and S4L). Courtship assessment of humanized SSC-DNMs in *HTR1D*, *SLC23A1*, and *MADD* did not show altered function to reference (Figures S4M–S4P). In summary, 63% (5/8) SSC-DNMs act functionally



(legend on next page)

different from reference alleles using quantitative behavioral measurements in flies.

Finally, we determined the CNS expression of TG4 lines corresponding to all eight lines we were able to humanize. Surprisingly, we only detected expression of 4/8 genes in the adult (Figures 3H, S5A, and S5B) and larval CNS (Figure S5C). All are expressed in a subset of neurons, whereas *GluClα* (*GLRA2*) is also expressed in some glia (Figure S5B).

Overexpression assays reveal ASD variants with diverse functional consequences

We complemented our rescue-based assays by overexpressing Ref-Tg and SSC-Tg in a wild-type background using ubiquitous (*tub-GAL4*), eye-specific (*GMR-GAL4*), and wing-specific (*nub-GAL4*) drivers (Figure 4A). This approach has routinely been employed to discern functional differences between reference and disease-associated variant proteins *in vivo*, regardless of whether the phenotypic readout has similarities to the patient's conditions (Ansar et al., 2019; Chung et al., 2020; Goodman et al., 2021; Harel et al., 2016; Huang et al., 2020; Kanca et al., 2019; Liu et al., 2018; Marcogliese et al., 2018; Post et al., 2020; Ravenscroft et al., 2021; Splinter et al., 2018). Critically, *UAS-Ref-Tg* and *UAS-SSC-Tg* are inserted into the same genomic landing site in the fly genome, and the construct only differs by the point mutation, allowing for direct functional comparison. Across the three drivers and testing 66 human genes (73 SSC variants), we found 21/73 SSC-DNMs (in 19 fly genes) showed functional alteration in phenotypic assays, 17/73 displayed phenotypes comparable to reference, and 35/73 did not produce a scorable phenotype (Table S5).

Twelve variants in eleven human genes ($ATP2B2^{T818M}$, $EPHA1^{V567I}$, $GLRA2^{N136S}$, $GRK4^{P385A}$, $ITGA8^{R748G}$, $IRF2BPL^{F30L}$, $IRF2BPL^{N701fs66*}$, $KCND3^{R86P}$, $MINK1^{C269R}$, $NPFFR2^{M163I}$, $PDK2^{R120Q}$, and $TSC2^{R1557W}$) behaved as LoF alleles. $GLRA2^{N136S}$, $GRK4^{P385A}$, $ITGA8^{R748G}$, $KCND3^{R86P}$, $MINK1^{C269R}$, $NPFFR2^{M163I}$, $PDK2^{R120Q}$, and $TSC2^{R1557W}$ were annotated as LoF alleles using a ubiquitous driver because they failed to reduce the expected viability to the extent of the corresponding reference alleles upon overexpression (Figure 4B). Notably, $GLRA2^{N136S}$ and $KCND3^{R86P}$ were also annotated as LoF alleles in our rescued based assay, showing consistency (Figures 3C and 3G). Moreover, $KCND3^{R86P}$ and $IRF2BPL^{N701fs66*}$ variants behaved as LoF alleles when assessed with multiple drivers. $KCND3^{R86P}$ abolished the activity of the reference transgene, which caused lethality when overexpressed with a ubiquitous driver (Figure 4B). In the wing, $KCND3^{R86P}$ failed to produce a severe notching phenotype that is observed by expression of the reference transgene (Figure 4D).

Ubiquitous or wing-specific overexpression of reference *IRF2BPL* caused lethality, whereas the $IRF2BPL^{N701Tfs66*}$ frameshift allele (note that *IRF2BPL* is a single-exon gene) does not cause any phenotype (Figures 4B and 4D). Interestingly, the missense variant, $IRF2BPL^{F30L}$, behaved as a LoF using the wing driver but was indistinguishable using the ubiquitous driver, indicating it is likely to be a partial LoF allele (Figures 4B and 4D). In addition to *IRF2BPL* DNMs, variants in two genes (*ATP2B2* and *EPHA1*) were identified as LoF alleles based on the wing-specific driver and assay. Wing-specific overexpression of reference *ATP2B2* in the developing wing disc causes a curled wing phenotype while the $ATP2B2^{T818M}$ variant fails to do so (Figure 4D). Expression of reference *EPHA1* caused a wing-size reduction and wing-margin serration, whereas the $EPHA1^{V567I}$ variant caused serrated wings of normal size (Figure 4D), indicating partial LoF.

Seven variants (ACE^{Y818C}, $GPC5^{M133T}$, $MYH9^{R1571Q}$, PC^{P1024R}, $SLC23A1^{L465M}$, $HTR1D^{T99N}$, and $BAIAP2L1^{A481V}$) behaved as GoF alleles. Flies overexpressing variant forms of *ACE*, *GPC5*, *MYH9*, *PC*, and *SLC23A1* exhibited enhanced lethality when compared with reference protein (Figure 4B). The GoF nature of $MYH9^{R1571Q}$ was also observed in the wing-size-based assay (Figure 4D). $HTR1D^{T99N}$ displayed consistent stronger phenotypes compared with reference when expressed in the eye or the wing, resulting in eye size reduction and absent wing phenotype, respectively (Figures 4C and 4D). $BAIAP2L1^{A481V}$ caused a smaller, more crumpled wing phenotype compared with its reference allele (Figure 4D).

Intriguingly, $EPHB1^{V916M}$ and $MAP4K1^{M725T}$ exhibited conflicting results in the eye and wing; therefore, they could not be categorized as simple LoF or GoF variants. While overexpression of reference *EPHB1* or *MAP4K1* in the developing eye causes eye-size-reduction phenotype, SSC variant forms of either gene result in normal eyes, indicating they behave as LoF alleles in this tissue. However, the same variant transgenes for these two genes expressed in the wing result in blistered or crumpled wings, respectively, that are phenotypically stronger than the reference alleles (Figures 4C and 4D), indicating they behave as GoF alleles in this tissue. In summary, when a scorable phenotype was present, 48% (21/44) of the SSC-DNMs tested with an overexpression strategy impact function. Furthermore, we found diverse SSC variant consequences, including 12 LoF, 7 GoF, and 2 with complex phenotypes.

While this overexpression-based screening approach was not directly investigating the function of genes in the nervous system, expression analysis revealed that most (15/19) fly genes that correspond to SSC-DNMs with a functional difference identified through our overexpression assay are expressed in the

Figure 3. Assessment of SSC-DNM function through humanization of viable TG4 lines and behavioral analysis

(A) Analysis pipeline used to evaluate *Drosophila* behavior.

(B) SSC-DNMs in which variants display significant differences in time spent performing a specific behavior (courtship, copulation, movement, or grooming) when compared with reference humanized flies.

(C–G) The number of frames male flies spent performing courtship (single-wing extensions), copulating, moving within the chamber, or grooming during a 30-min test period. The red line represents the average number of frames a *Canton-S* (control) male spends performing the same task. n = 10–40 flies were used per genotype. Statistical analysis was performed by non-parametric Kruskal-Wallis one-way ANOVA and the Dunn's multiple comparison test.

(H) Single focal confocal images showing expression pattern of *UAS-nls.mCherry* driven by *TG4* (red) in the anterior of the central adult brain. Bottom two rows depict the 5× zoom from dotted white box of co-localization between TG4 reporter and Elav (neurons) or Repo (glia) in cyan. Co-localization is depicted in white. Scale bars represent 50 μm. *p < 0.05, ***p < 0.001, ****p < 0.0001, and ns, not significant. Error bars are ± SEM.

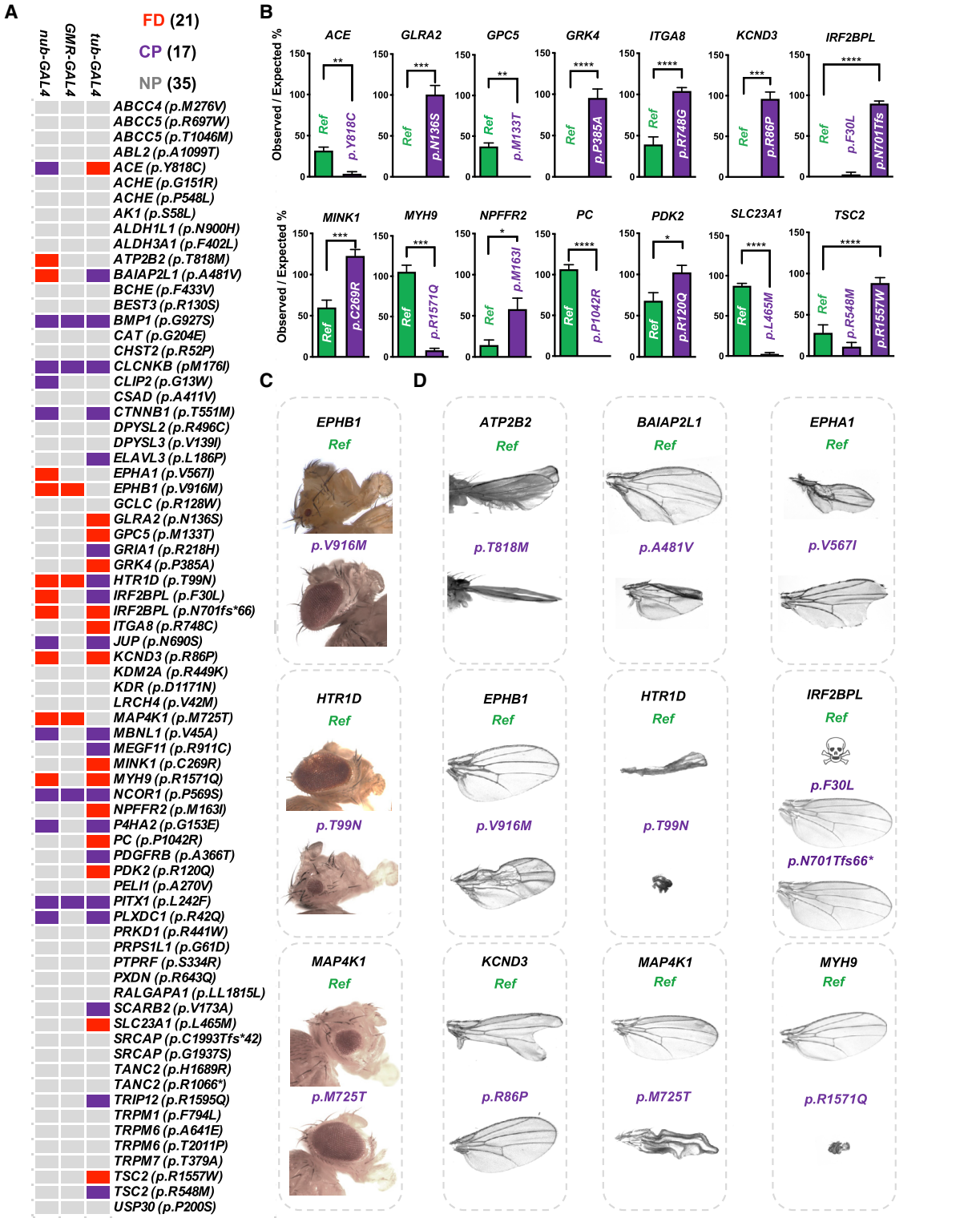


Figure 4. Variant assessment by overexpression of reference and SSC-DNMs

(A) Phenotypes observed upon overexpressing the reference and variant cDNAs using a ubiquitous driver (*tub-GAL4*) at 25°C, an eye-specific driver (*GMR-GAL4*) at 29°C, or a wing-specific driver (*nub-GAL4*) at 25°C. Black denotes there was no phenotype (NP), purple there was a comparable phenotype (CP), or red there was a functional difference (FD).

(legend continued on next page)

adult (Figures 3H, S6, S7, and S8A) and larval (Figure S8B) CNS (see Marcogliese et al., 2018 for *Pits* [corresponding to human *IRF2BPL*]). The deleterious nature of variants in three of these genes was identified in our behavioral screen (*GluClalpha*, *5-HT1B*, and *Usp30*; Figure 3H). Four genes (*CG6293*, *Mhc*, *Shal*, and *Ance*) were not detected in the brain in our analysis, which could be because they are expressed at very low levels or may primarily function in non-neural tissues. While most genes are enriched in neuronal subpopulations, *Pdk* (*PDK2*) and *hppy* (*MAP4K1*) are enriched in glia as well. Interestingly, *if ITGA8* is not detected in either neurons or glia but revealed a unique expression pattern, that may reflect its expression in tracheal cells. In addition, based on imaging with an nls.GFP that leaks into the cytoplasm, it may also be present within cells that wrap around neurons reminiscent of cells in *pars intercerebralis*, a neuroendocrine organ analogous to the mammalian hypothalamus (de Velasco et al., 2007; Figure S8A). Taken together, most of the fly genes corresponding to SSC-DNMs in which we found *in vivo* alterations using overexpression-based assays are expressed in the fly CNS, similar to hits from rescue-based studies.

Identification of 30 deleterious SSC variants by merging all functional data

In total, we found 29 missense and 1 frameshift SSC-DNMs that displayed functional differences when compared with their respective reference allele (for a total of 28 genes: one variant for 26 genes and two variants for two genes; Tables 1 and S5). Approximately 53% (30/57) of the SSC-DNMs exhibited functional differences compared with the reference. Intriguingly, in our study, we only found GoF variants for genes corresponding to viable TG4 fly mutants based on both rescue-based and overexpression-based assays (Figure S9A). Interestingly, while we were able to classify the variants into LoF or GoF based for most genes, we found in two cases where different assays gave different results.

When we informatically surveyed the genes and variants with functional consequences identified through our screen in comparison to other genes included in our study (variants with comparable function or those lacking a phenotype to assess) using the MARRVEL tool (Wang et al., 2017), we did not find any significant differences in gene-level metrics, such as pLI (Lek et al., 2016); LOEUF (Karczewski et al., 2020); missense o/e (Karczewski et al., 2020); pathogenicity prediction scores based on several *in silico* algorithms, including sorting intolerant from tolerant (SIFT) (Vaser et al., 2016), PolyPhen-2 (Adzhubei et al., 2010), and combined annotation-dependent depletion (CADD) (Kircher et al., 2014); or absence or presence of identical variant in gnomAD (Lek et al., 2016) (Tables S5 and S6; Figures S9B–S9Z). By analyzing Gene Ontology (GO) by PantherDB (Ashburner et al., 2000) and visualizing with reduce visualize gene ontology (REVIGO) (Supek et al., 2011), ASD candidate genes from SSC with deleterious variants *in vivo* were compared with all protein-coding genes. We found most significant enrichment for genes with GO terms for "synapse (cellular component)" (Figure S10A) and "ion binding (function)" (Figure S10B). Finally, we systematically imaged the expression pattern of 41 additional TG4 lines generated through our study to document their expression in the adult (Figure S11A) and larval (Figure S11B) CNS as a resource for the community.

Loss- or gain-of-function alleles in *GLRA2* cause X-linked neurodevelopmental disorders

Many genes implicated in ASD are also associated with neurodevelopmental disorders beyond autism (Levitt and Campbell, 2009; Sullivan and Geschwind, 2019). Therefore, we asked whether additional variants in genes with disruptive SSC-DNMs could also be responsible for neurological diseases beyond ASD by identifying human subjects with rare, potentially deleterious variants that have not previously been associated with neurological disease (Chong et al., 2015; Gahl et al., 2016; Sobreira et al., 2015). Out of 28 genes in which we identified damaging SSC-DNMs, eight are associated with Mendelian diseases that have neurological presentations documented in OMIM (Amberger et al., 2019; Table 1). For one of these genes (*IRF2BPL*), we recently reported *de novo* truncating variants as the cause of a severe neurodevelopmental disorder that presents with abnormal movements, loss of speech, and seizures. This work was done in collaboration with the Undiagnosed Diseases Network (Gahl et al., 2016; Marcogliese et al., 2018). Here, aided by using the online matchmaking software, GeneMatcher (Sobreira et al., 2015), internal human genetics databases, and re-analysis of clinical exome sequencing data, we report variants in *GLRA2* as a cause of a neurodevelopmental syndrome with developmental delay (DD), intellectual disability (ID), ASD, and epilepsy (Figure 1H; Table 2).

We identified rare *GLRA2* variants in 13 unrelated subjects with or without autistic features. In addition to developmental and cognitive delay of variable severity, 4/13 subjects have microcephaly, 6/13 subjects have a history of epilepsy, and 10/13 subjects have ocular manifestations, including congenital nystagmus that improved with age in three of the cases (Table 2; see GLRA2 subject case histories). *Glycine receptor alpha 2* (*GLRA2*) is an X-linked gene that encodes a subunit of a glycine-gated chloride channel (Zeilhofer et al., 2018). All female subjects harbored DNMs, including a recurring $GLRA2^{T296M}$ *de novo* variant found in 6/8 female subjects. The $GLRA2^{T296M}$ variant was also identified in a female subject in previous large-scale developmental disorder study (Deciphering Developmental Disorders Study, 2017). The five male subjects had inherited *GLRA2* variants from unaffected mothers. The mother of

(B) Quantification of viability upon overexpression of reference or variant human cDNAs using a ubiquitous driver for genes where the variants showed a functional difference. Minimum of three independent crosses were set with two independent UAS-transgenic lines. We scored 50–100 flies (a minimum of 10 if overexpression caused survival defects). Statistical analyses were performed by unpaired t test.

(C and D) Representative optical sections of eyes and wings for variants with a functional difference using eye-specific (*GMR-GAL4*) and wing-specific (*nub-GAL4*) drivers, respectively, at 25°C.

*p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001. Error bars are +SEM.

Table 1. Identification of 30 SSC-DNMs with functional consequences

H. sap gene	pLI (LOEUF)	Missense O/E	OMIM disease	SSC variant	CADD	*D. mel gene*	TG4 lethality	Functional assay	SSC-CV consequence
ABL2	0 (0.58)	0.81	–	p.A1099T	28.2	*Abl*	yes	RB	LoF
ACE	0 (1.08)	1.07	267430 (AR)	p.Y818C	7.5	*Ance*	no	OE	GoF
ALDH1L1	0 (0.78)	0.93	–	p.N900H	9.8	*CG8665*	no	RB	GoF?
ATP2B2	1 (0.15)	0.54	601386 (AR)[a]	p.T818M	33.0	*PMCA*	yes	OE	LoF
BAIAP2L1	0 (0.65)	0.90	–	p.A481V	17.8	*IRSp53*	no	OE	GoF
CAT	0 (1.05)	1.01	614097 (AR)	p.G204E	28.1	*Cat*	yes	RB	LoF
CHST2	0.02 (0.81)	0.66	–	p.R52P	12.8	*CG31637*	yes	RB	LoF
EPHA1	0 (1.04)	0.97	–	p.V567I	1.3	*Eph*	no	OE	LoF
EPHB1	1 (0.26)	0.73	–	p.V916M	34.0	*Eph*	no	OE	complex
GLRA2	0.97 (0.30)	0.43	–	p.N136S	25.1	*GluClα*	no	RB, OE	LoF
GPC5	0 (1.08)	1.09	–	p.M133T	24.3	*dally*	no[b]	OE	GoF
GRK4	0 (1.09)	1.09	–	p.P385A	26.0	*Gprk2*	yes	OE	LoF
HTR1D	0 (1.30)	0.98	–	p.T99N	19.4	*5-HT1B*	no	OE	GoF
IRF2BPL	0.84 (0.41)	0.90	618088 (AD)[a]	p.F30L	24.8	*Pits*	yes	OE	LoF
				p.N701fs	–			OE	LoF
ITGA8	0 (0.66)	1.03	191830 (AR)	p.R748C	35.0	*if*	yes	OE	LoF
KCND3	0.99 (0.28)	0.48	607346 (AD)[a]	p.R86P	32.0	*Shal*	no	RB, OE	LoF?
KDM2A	1 (0.04)	0.43	–	p.R449K	5.7	*Kdm2*	no	RB	?
MAP4K1	0.99 (0.29)	0.59	–	p.M725T	21.3	*hppy*	no	OE	complex
MINK1	1 (0.13)	0.60	–	p.C269R	26.8	*msn*	yes	OE	LoF
MYH9	1 (0.09)	0.71	603622 (AD)[a]	p.R1571Q	35.0	*Mhc*	no[b]	OE	GoF
NPFFR2	0 (1.13)	1.20	–	p.M163I	13.3	*SIFaR*	yes	OE	LoF
PC	0.01 (0.43)	0.69	266150 (AR)[a]	p.P1042R	24.6	*PCB*	no	OE	GoF
PDK2	0 (0.92)	0.63	–	p.R120Q	25.3	*Pdk*	yes	OE	LoF
SLC23A1	0.02 (0.54)	0.71	–	p.L465M	17.9	*CG6293*	no	OE	GoF
TRIP12	1 (0.06)	0.60	617752 (AD)[a]	p.R1643Q	36.0	*ctrip*	yes	RB	LoF
TRPM6	0 (0.45)	0.86	602014 (AR)[a]	p.T2011P	12.2	*Trpm*	yes	RB	LoF
				p.A641E	17.8			RB	LoF
TSC2	1 (0.07)	1.03	613254 (AD)[a]	p.R1557W	16.0	*gig*	yes	OE	LoF
USP30	0 (0.66)	0.76	–	p.P200S	14.7	*Usp30*	no	RB	?

List of all human genes and corresponding SSC variants determined to have a functional difference across all assays in this study. GoF, gain of function; LoF, loss of function; OE, overexpression; RB, rescue-based.
[a]Nervous system disorder
[b]Known lethal mutants

subject 12 has a history of learning problems. The CADD scores for all five male subjects are predicted to be damaging (Table 2). Four of the five male subjects had diagnosed or suspected ASD. A maternally inherited microdeletion of *GLRA2* was previously reported in a single male patient with ASD (Pinto et al., 2010), indicating that hemizygous LoF allele of this gene in males may cause ASD. Indeed, the *de novo* $GLRA2^{N136S}$ variant present in the SSC in a male subject acts as a LoF allele based on overexpression studies (Figure 4B), which is supported by our behavioral assay on humanized flies, as the $GLRA2^{N136S}$ variant loses the toxic effect on copulation caused by expression of the humanized reference protein (Figure 3G).

To better understand the functional consequences of variants found in our *GLRA2* cohort, we generated additional transgenic flies to assay the function of p.T296M (found in six female subjects) and p.R252C (found in a male subject) variants (Figures 5A and S12A). By overexpressing reference and variant *GLRA2* using a ubiquitous driver, we found that $GLRA2^{R252C}$ behaves as a LoF allele (Figure 5B), similar to $GLRA2^{N136S}$ (Figure 4B). In contrast, this assay did not distinguish $GLRA2^{T296M}$ from the reference (Figure 5B). Given the recurrent nature of this variant, as well as structural prediction that the residue has a potential role in obstruction of the ion pore in the closed conformation (Figure S12G; Du et al., 2015; Moraga-Cid et al., 2015), we further tested $GLRA2^{T296M}$ and other alleles using additional GAL4 drivers. Using a *pnr-GAL4* that is expressed in the dorsocentral stripe in the notum, we found that $GLRA2^{T296M}$, but not the reference or any other *GLRA2* variant tested, causes lethality when expressed at high levels (Figure S12B). When we expressed $GLRA2^{T296M}$ at lower levels by manipulating the

Table 2. Salient features of subjects with *GLRA2* variants

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13
GLRA2 variant (hg19, NM_001118886.1)	c.887C>T, p.Thr296Met	c.887C>T, p.Thr296Met	c.887C>T, p.Thr296Met	c.887C>T, p.Thr296Met	c.887C>T, p.Thr296Met	c.887C>T, p.Thr296Met	c.140T>C, p.Phe47Ser	c.777C>G, p.Ile259Met	c.754C>T, p.Arg252Cys	c.862G>A, p.Ala288Thr	c.1186C>A, p.Pro396Thr	c.1199C>T, p.Pro400Leu	c.1334G>A, p.Arg445Gln
Inheritance	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	maternal	maternal	maternal	maternal	maternal
CADD score	27	27	27	27	27	27	27.8	23.4	31	27.2	20.9	22.3	33
Gender	female	female	female	female	female	female	female	female	male	male	male	male	male
Age at most recent evaluation (years)	6.7	6.5	5.5	0.5	5.4	0.8	6.7	5	0.9	7	34	15	5.9
Developmental delay/ intellectual disability	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes, with regression	yes	yes	yes
Hypotonia/ incoordination	no	no	yes, ataxic gait	yes	no	no	no	yes, incoordination	yes	yes, ataxia	yes	no	yes, impaired fine motor coordination
Autism spectrum disorder	no	yes	no	NA	no	NA	no	no, diagnosis of child psychosis	NA	yes	yes	yes	suspected
Inattention/ hyperactivity	yes	yes	no	NA	yes	NA	yes	no	NA	no	no	yes	yes
Sleep disturbance	no	yes	no	no	no	NA	yes	yes	yes	no	no	no	yes
Microcephaly	no	yes	yes	yes	yes	borderline	no	no	no	no	no	no	no
Ocular features	myopia, astigmatism, and nystagmus (improved with age)	nystagmus (improved with age)	alternating exotropia, borderline opsoclonus	none	oculomotor apraxia, ptosis	upbeat nystagmus (starting 6 weeks after birth)	strabismus, nystagmus (improved with age)	none	strabismus	myopia	myopia, astigmatism	none	reduced visual acuity
Epilepsy	no	yes	no	yes	no	yes	yes	no	no	yes	no	yes	yes
EEG findings	slow background suggestive of mild encephalopathy	bilateral synchronized high-amplitude spikes, epileptic potentials	normal	slow background, infantile spasms, multifocal spikes during sleep	not performed	intermittent hypsarrythmic pattern, infantile spasms	infantile spasms and then normal interictal EEG	left fronto-temporal spike waves focus, which diffuses in the right frontal region, activated by sleep	not performed	generalized slowing and generalized epileptiform discharges associated with myoclonic jerks	not performed	right temporal focus of high and polymorphic alfa spike-wave complexes, with ipsilateral propagation	left and right posterior and right frontal intermittent slowing, bilateral polyspikes during sleep, and excessive beta-activity (with medications)

(*Continued on next page*)

Table 2. *Continued*

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13
Brain MRI findings	normal	delayed myelination, a small arachnoid cyst	mild cortical atrophy, thinning of corpus callosum	normal	normal	normal	cortical and white-matter atrophy, including vermian atrophy	increased signal intensity in FLAIR of the subcortical white matter of the frontal region	not performed	minimally increased T2 signal intensity on the occipital lobes	normal	increased signal intensity in FLAIR of the cortical matter of the parietal region	normal

CADD, combined annotation-dependent depletion; EEG, electroencephalography; MRI, magnetic resonance imaging.

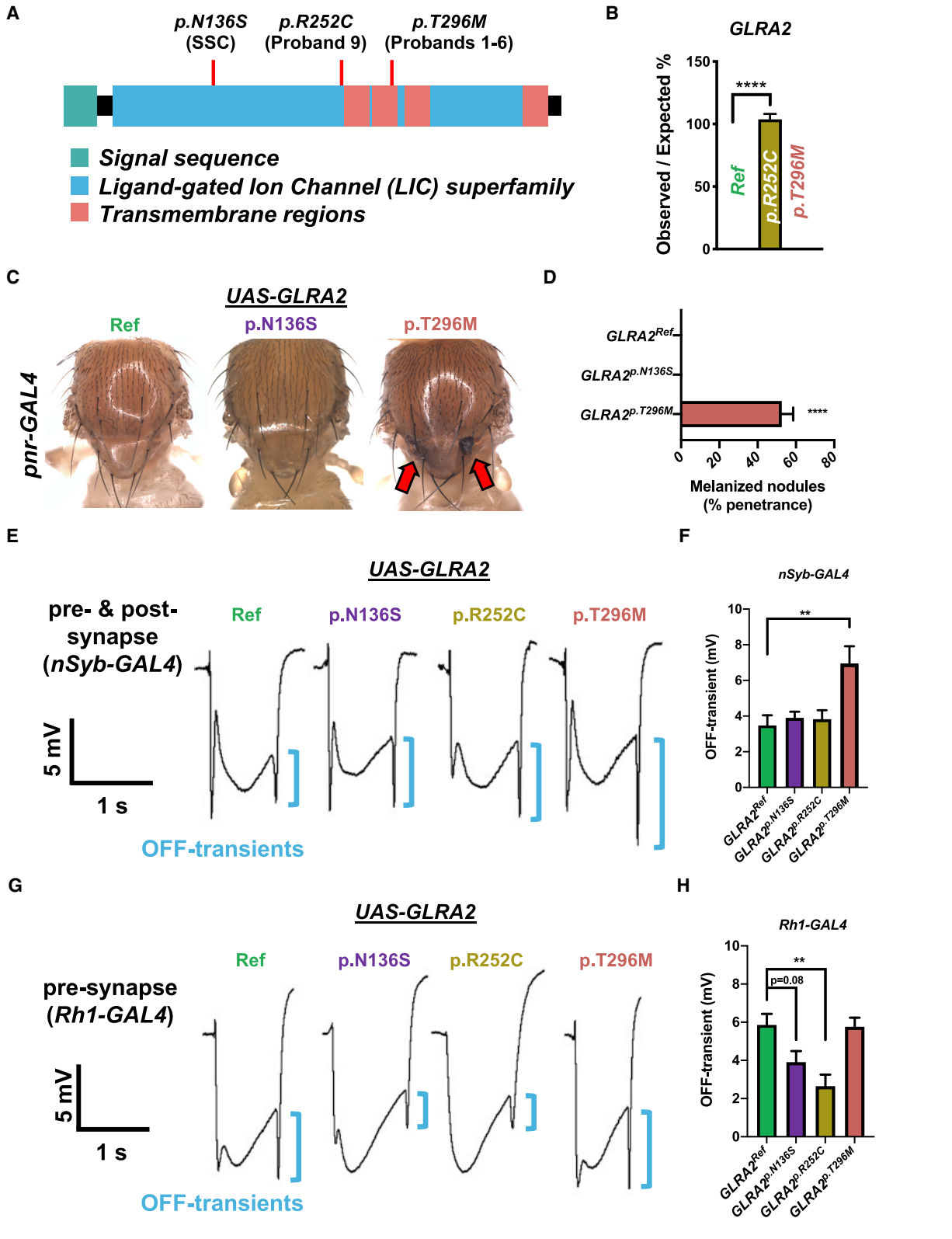
temperature, we found that this variant induces the formation of melanized nodules in the thorax, a phenotype that we never observe when the reference or other *GLRA2* variants are overexpressed (Figures 5C, 5D, and S12B).

To further examine the functional consequences of overexpression of reference and variant *GLRA2* in the nervous system, we performed electroretinogram (ERG) recordings on the fly eye expressing human *GLRA2* using two distinct drivers. Pan-neuronal driver (*nSyb-GAL4*; Pauli et al., 2008) allows one to express *GLRA2* in both pre-synaptic photoreceptors and post-synaptic neurons in the nervous system. Using this driver, we found a significant increase in amplitudes of "*OFF*" transients with $GLRA2^{T296M}$ (Figures 5E, 5F, S12C, and S12D). This indicates an increase in synaptic transmission (Deal and Yamamoto, 2018), supporting the finding in the notum that p.T296M behaves as a GoF allele. Interestingly, when we limited the expression of *GLRA2* to pre-synaptic photoreceptors using *Rh1-GAL4* (Xiong et al., 2012), we did not observe any functional difference between $GLRA2^{T296M}$ and the reference allele. However, with this driver, we were able to discern that both $GLRA2^{R252C}$ seen in subject 9 and $GLRA2^{N136S}$ found in an SSC subject behave as LoF alleles based on observing a decrease in amplitude of "*OFF*" transients, indicating a decrease in synaptic transmission (Figures 5G, 5H, S12E, and S12F). Hence, we have identified a cohort of subjects with deleterious variants in an X-linked gene *GLRA2* and have shown that a recurrent missense DNM in females acts as a GoF allele, whereas rare variants found in male subjects behave as LoF alleles.

DISCUSSION

In this study, we generated >300 fly strains that allow functional studies of human variants and homologous fly genes *in vivo*. These reagents can be used to study many coding variants that are being identified through next-generation sequencing efforts in the human genomics field in diverse disease cohorts beyond ASD. Our screen elucidated 30 SSC variants with functional differences compared with reference, which was over half (~53%) of the genes in which we were able to perform a comparative functional assay.

Our screen was part of a larger effort to characterize the functional consequences of missense *de novo* changes from the SSC dataset using different strategies. One approach was based on proteomics by performing yeast two-hybrid assays on 109 SSC-DNMs found in subjects, showing 20% of protein-protein interactions that are found in reference proteins are disrupted by variants (Chen et al., 2018). Another study reported that ~70% of 37 SSC-DNMs knocked into homologous *C. elegans* genes caused scorable phenotypes (Wong et al., 2019). These studies are complementary to each other because, while some variants have been identified as deleterious by more than one approach (e.g., $GLRA2^{N136S}$ identified in both worm and fly screens), others are uniquely identified in one study, which could be due to technical limitations. For example, our approach utilizing human cDNA transgenes allowed us to test variant function, regardless of residue conservation in *Drosophila*. Of the 29 disruptive missense SSC-DNMs



(legend on next page)

identified through our study, 14 affect residues that are conserved in flies and 10 in worms.

To take an unbiased approach, our gene prioritization was only based on gene-level conservation and tool availability (e.g., intronic MiMIC lines and full-length human cDNA) rather than based on gene level constraints and variant-level pathogenicity prediction scores. Hence, our study subset, although somewhat limited, can be considered a random sample of ASD-implicated genes and variants. Interestingly, we could not find any significant difference in pathogenicity prediction for disruptive variants *in vivo*. It should also be noted that we were limited by the availability of full-length human cDNA, which could select against genes encoding larger transcripts, for which reagents are often harder to obtain. We were able to generate 13 additional *TG4* lines that were homozygous viable and successfully crossed back to a *Canton-S* X chromosome. We assessed their behavior phenotypes in comparison to the reference *Canton-S* files, which is presented in Figures S4Q–S4T as an additional resource for the community.

Of the 29 missense SSC-DNMs that had functional consequences in our assays, four were not predicted as damaging variants (CADD < 10), nine had moderate scores (CADD: 10–20), and 16 were predicted to be disruptive (CADD > 20). Understanding how variants that are not predicted to be damaging based on state-of-the-art informatics programs impact protein function may provide guidance to improve the accuracy of *in silico* tools.

To study the functional consequences of SSC-DNMs, we took two conceptually different approaches (rescue-based humanization strategy and overexpression-based strategy). Indeed, the two approaches were complementary, as only two variants (*GLRA2*^{N136S} and *KCND3*^{R86P}) were detected in both screens, showing consistent LoF effects using both approaches. However, it should be noted that variant interpretation is not always straightforward for behavioral analysis. Interestingly, *GRK4*, *NPFFR2*, and *PDK2* SSC-DNMs were found to be LoF variants when overexpressed ubiquitously, yet these variants were able to rescue lethality in a similar manner to their respective reference alleles (Figure 4; Table S5). This suggests that these variants are partial LoF alleles and different drivers and assays have different sensitivity. Moreover, two disruptive SSC-DNMs, *EPHB1*^{V916M} and *MAP4K1*^{M725T}, behaved as complex alleles, displaying discordant phenotypes in the eye and wing. This suggests that these variants may behave in a context-dependent manner, acting as a GoF allele in one tissue while behaving as a LoF allele in another. Thus, one functional assay may not be enough to reveal the full nature of pathogenic mechanisms, and some disease-associated variants may act differently in different tissues or cell types. It is also important to note that a variety of factors may explain why functional differences are not observed across assays, as different *GAL4* lines may have variable developmental timing, strength, and context dependency. In addition, the presence of the endogenous protein in cells and the cell-type-specific effect of exogenous protein may also contribute to our functional readout. Therefore, variant annotation in *Drosophila* should be supplemented by deeper characterization of the loss-of-function mutant and gene-expression studies to provide the clearest supportive evidence to a molecular diagnosis.

Starting from a single *de novo* hemizygous missense variant that we identified as a LoF allele in *GLRA2* (p.N136S), we identified a cohort of subjects with overlapping neurodevelopmental phenotypes carrying LoF or GoF variants in this gene. Interestingly, female subjects harbored DNMs and male subjects carried maternally inherited variants in this X-linked gene, which undergoes random X inactivation in females, but not in males (Barakat and Gribnau, 2012). The X-linked status of *GLRA2* may mean that variants causing reduced *GLRA2* activity lead to disease in males but can be tolerated in heterozygous females. This is supported by asymptomatic mothers of male probands who had maternally inherited alleles (subjects 9–11 and 13). In contrast, GoF variants in this channel could be overrepresented in females since hyperactivation of this channel may cause neurological defects (Zhang et al., 2017). Of the eight female subjects, six carried the identical recurring DNM, p.T296M. None of the variants observed in females are present in control databases, arguing strongly that these variants are pathogenic. While the exact mechanism of how the p.T296M variant affects GLRA2 function remains unclear, the presence of melanized nodules in flies expressing this variant are indicative of an innate immune response (Dudzic et al., 2019), potentially as a result of leaky ion channel function (Feske et al., 2015). Fittingly, our structural analysis revealed that the p.T296 residue is adjacent to a critical amino acid that is likely important for keeping the ion pore in a conformationally closed state (Figure S12G).

In summary, we utilized a model organism-based *in vivo* functional genomics approach to study the functional consequence of rare genetic events in a common neurological disorder, ASD. In addition to garnering variant functional data for ASD subjects in the SSC, we leveraged this information to identify and document a rare neurological condition through

Figure 5. *GLRA2*^{T296M} found in female probands acts as a GoF allele while *GLRA2*^{R252C} and *GLRA2*^{N136S} found in male probands behave as LoF alleles

(A) Schematic diagram of domain structure of GLRA2 and the relative positions of subject variants functionally assessed in *Drosophila*.

(B) Mendelian ratios upon overexpression of the *GLRA2* reference or variant human cDNAs using a ubiquitous driver (*tub-GAL4*). A minimum of three independent crosses were set.

(C and D) Representative images and quantification of melanized nodules formed on the notum of flies expressing *GLRA2*^{T296M} driven by a dorsocentral thorax-specific (*pnr-GAL4*) driver at 25°C.

(E–H) Representative traces of ERG and quantification of "*OFF*"-transient amplitude (blue bracket) in animals expressing GLRA2 pan-neuronally (both pre-synaptic photoreceptors and post-synaptic laminar neurons; *nSyb-GAL4*) or only in the pre-synaptic photoreceptors (*Rh1-GAL4*). ERG was analyzed by ANOVA followed by Bonferroni's multiple comparison test. Four to ten flies were examined for each genotype. Recording was repeated at least three times per fly. **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$. Error bars are +SEM.

matchmaking and collaboration. Such bidirectional communication and collaboration between bench scientists and clinicians greatly facilitates functional studies of human variants found in common diseases, such as ASD, and can also lead to novel discoveries that have an impact on rare-disease research.

Limitations of the study

Although this study revealed a number of rare variants found in ASD patients that have functional consequences, there are several caveats to recognize based on the design and assays used in this screen. First, it is not clear for the majority of the 30 hits found by the screen whether the disruptive variants are truly pathogenic and directly contribute to ASD. Identification of additional patients with similar genotypes and phenotypes will be necessary to establish a causal relationship between these variants and ASD pathogenesis. Second, we do not argue that this screen was able to identify all variants that had functional consequences, considering most variants were shown to be deleterious based on one phenotypic assay. Hence, if one performs additional assays in different biological contexts, more variants with altered function may be discovered. Third, it is not clear how many phenotypic assays are required to determine whether a variant has functional consequences and which phenotypic assays in flies are more relevant to complex human phenotypes, such as ASD. Additional studies of variants identified as having functional consequences in an invertebrate model organism should be followed up using mammalian models or human cells, tissues, and organoids to assess whether the deleterious variants affect biological processes that relate to ASD. Fourth, although the majority of the hits were alleles of genes in which the fly homolog is expressed in the nervous system, some were in genes in which we failed to detect any expression in this organ system in *Drosophila*. This could be because some of these genes are expressed at low levels that are beyond the detection limit of our assay systems; only expressed at a specific time point during neural development; expressed in the nervous system of humans, but not flies; or they may be contributing to neurological phenotypes through their function in non-neural organ systems. Additional gene expression and functional studies will be required to fully understand their mechanistic contributions to ASD. Fifth, although we were able to identify some deleterious variants that affect fly behavior, which is the most relevant phenotype to ASD out of all assays we performed in this screen, we have not been able to assess this for all variants due to technical limitations. In order to perform clean behavioral experiments, one must control the genetic background since this could be a significant confounding factor. Although we tried to eliminate some of the genetic variability by inserting the reference and variant transgenic constructs into the same genomic location and swapping out mutant chromosomes (*y w*) that are known to affect behavioral outcomes, we did not isogenize all chromosomes through multiple rounds of back crossing to facilitate the speed of our screen. Hence, additional behavioral studies performed on a more standardized genetic background (e.g., cantonized flies) will likely provide additional information regarding the role of these genes in fly behavior, which could provide additional insights into their links to ASD. Finally, we would like to emphasize that this work should be considered as a pilot screen rather than a comprehensive screen of *de novo* missense variants identified in the SSC. The genetic variants that we were able to study in our screen were limited by the availability of intronic MiMIC elements in fly genes because recombinase-mediated cassette exchange (RMCE) was the only efficient way to generate TG4 lines at the time of project initiation. With the advent of CRISPR-mediated integration cassette (CRIMIC) and KozakGAL4 insertion technologies (Lee et al., 2018; Kanca et al., 2021), virtually any fly gene is now targetable to generate a strong LoF allele with GAL4 expression that allows for subsequent humanization experiments. Therefore, future screening strategies could employ these and other emerging techniques to assess functional consequences of many rare missense variants found in ASD or other disease patients.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Generation of TG4 lines
  - Generation of UAS-human cDNA lines
  - Fly husbandry
  - Fly stocks used that were not generated here
  - Patient recruitment and consent
  - GLRA2 subject case histories
- METHOD DETAILS
  - Ortholog candidate identification
  - Electroretinograms (ERG)
  - Complementation test of lethality in TG4 lines
  - Rescue of lethality in TG4 lines by UAS-human cDNA transgenes
  - Lifespan assays
  - Behavioral assays
  - Overexpression assays to assess lethality and morphological phenotypes
  - Imaging of adult fly morphology
  - Expression analysis of TG4 lines in larval and adult brains
  - GeneOntology (GO) analysis
  - Exome sequencing and identification of GLRA2 variants
  - SDS-PAGE/Western blot
  - Structural biological analysis of GLRA2 patient variants
  - Image generation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Gene and variant level statistics
  - Rescue-based and overexpression-based survival
  - Lifespan analysis
  - Behavior
  - Assessment of melanized nodules
  - Electroretinograms

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110517.

ACKNOWLEDGMENTS

We express our appreciation to the subjects and families for their participation in this study. We thank Nora Duran, Mark Durham, Shelley Gibson, Yuchun He, Matthew Lagarde, Wen-Wen Lin, and Dr. Karen Schulze for technical or administrative assistance. We thank Dr. Hugo Bellen for insightful scientific discussions and valuable comments on this manuscript. We sincerely thank the late Dr. Kenneth Scott for access to many of the human cDNAs used in this study. This work was primarily supported by a Simons Foundation Autism Research Initiative (SFARI) Functional Screen Award (368479) to M.F.W. and S.Y. Generation of human cDNA transgenic lines were in part supported by NIH/ORIP (R24OD022005). Confocal microscopy is supported in part by NIH/NICHD (U54HD083092) to the Intellectual and Developmental Disabilities Research Center (IDDRC) Neurovisualization Core at BCM. P.C.M. is supported by CIHR (MFE-164712) and the Stand by Eli Foundation. J.A. is supported by NIH/NINDS (F32NS110174). H.C. is supported by the Warren Alpert Foundation. R.M. is supported by NIH/NIGMS (T32GM07526-43) and through BCM Chao physician-scientist award. C.M.L. is part of the BCM Medical Scientist Training Program and McNair MD/PhD Student Scholars, supported by the McNair Medical Institute at the Robert and Janice McNair Foundation and NIH F30 Award (F30MH118804). B.H. is supported in part by the Postbaccalaureate Research Education Program NIGMS (R25 GM069234). H.T.C. received support by the American Academy of Neurology and CNCDP-K12. T.S.B. is supported by the Netherlands Organisation for Scientific Research (ZonMW Veni, grant 91617021), an Erasmus MC Fellowship 2017, and Erasmus MC Human Disease Model Award 2018. R.G. and A.V. received support from The DECODE-EE project (Health Research Call 2018; Tuscany Region) and are members of the European Reference Network EpiCARE. A.B., L. Pavinato, and S.T. received funding specifically appointed to Department of Medical Sciences from the Italian Ministry for Education, University and Research (Ministero dell'Istruzione, dell'Università e della Ricerca—MIUR) under the program "Dipartimenti di Eccellenza 2018–2022" project code D15D18000410001. For subject 10, sequencing and analysis were provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG) and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung and Blood Institute grant UM1 HG008900 and in part by National Human Genome Research Institute grant R01 HG009141.

AUTHOR CONTRIBUTIONS

M.F.W. and S.Y. conceived and designed the project. P.C.M., J.A., S.L.D., and J.M.H. designed and conducted most fly experiments and analyzed the data. V.H.B. and Y.-H.C. performed cloning and mutagenesis. H.K.G. performed cloning and coordination. S.J. performed immunostaining and confocal microscopy. X.L. performed structural analysis and generated reagents. N.L., D.B., Y.-H.C., P.L., B.H., H.P., P.B., M.-C.H., C.M.L., H.-T.C., H.C., N.A.H., O.K., and S.N.M. contributed to reagent generation and some fly experiments. R.M., A.G., E.M., C.S., S.F., R.G., A.V., E.E., C.N.M., T.S.B., M.F.v.D., M.W., M.v.S., G.L., I.S., N.C., C.A.B., J.A.M., P.B.A., B.K., T.C., L. Perrin, M.B., T.R., S.L., F.T.M.-T., J.D., E.S.-A., S.T., E.R., V.S., K.P., R.K., L. Pavinato, and A.B. reported and described *GLRA2* subjects. L.Z.R., R.J.G., and J.A.R. aided in subject matchmaking and collection and organization of patient data. P.C.M., J.A., S.L.D., J.M.H., R.M., M.F.W., and S.Y. wrote and revised the manuscript.

DECLARATION OF INTERESTS

The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing completed at Baylor Genetics Laboratories.

Received: February 2, 2021
Revised: September 23, 2021
Accepted: February 18, 2022
Published: March 15, 2022

REFERENCES

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.

Amberger, J.S., Bocchini, C.A., Scott, A.F., and Hamosh, A. (2019). OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Res. *47*, D1038–D1043.

Ansar, M., Chung, H., Al-Otaibi, A., Elagabani, M.N., Ravenscroft, T.A., Paracha, S.A., Scholz, R., Magid, T.A., Sarwar, M.T., Shah, S.F., et al. (2019). Biallelic variants in IQSEC1 cause intellectual disability, developmental delay, and short stature. Am. J. Hum. Genet. *105*, 907–920.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29.

American Psychiatric Association (APA) (2013). Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Pub).

Barakat, T.S., and Gribnau, J. (2012). X chromosome inactivation in the cycle of life. Development *139*, 2085–2089.

Bellen, H.J., Wangler, M.F., and Yamamoto, S. (2019). The fruit fly at the interface of diagnosis and pathogenic mechanisms of rare and common human diseases. Hum. Mol. Genet. *28*, R207–R214.

Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. U S A *104*, 3312–3317.

Brunet, T., Jech, R., Brugger, M., Kovacs, R., Alhaddad, B., Leszinski, G., Riedhammer, K.M., Westphal, D.S., Mahle, I., Mayerhanser, K., et al. (2021). De novo variants in neurodevelopmental disorders—experiences from a tertiary care center. Clin. Genet. *100*, 14–28.

Burley, S.K., Berman, H.M., Bhikadiya, C., Bi, C., Chen, L., Costanzo, L.D., Christie, C., Dalenberg, K., Duarte, J.M., Dutta, S., et al. (2019). RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res. *47*, D464–D474.

Calleja, M., Herranz, H., Estella, C., Casal, J., Lawrence, P., Simpson, P., and Morata, G. (2000). Generation of medial and lateral dorsal body domains by the pannier gene of Drosophila. Development *127*, 3971–3980.

Chen, S., Fragoza, R., Klei, L., Liu, Y., Wang, J., Roeder, K., Devlin, B., and Yu, H. (2018). An interactome perturbation framework prioritizes damaging missense mutations for developmental disorders. Nat. Genet. *50*, 1032–1040.

Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al. (2015). The genetic basis of mendelian phenotypes: discoveries, challenges, and opportunities. Am. J. Hum. Genet. *97*, 199–215.

Chung, H.-L., Wangler, M.F., Marcogliese, P.C., Jo, J., Ravenscroft, T.A., Zuo, Z., Duraine, L., Sadeghzadeh, S., Li-Kroeger, D., Schmidt, R.E., et al. (2020). Loss- or gain-of-function mutations in ACOX1 cause axonal loss via different mechanisms. Neuron *106*, 589–606.e6.

Coe, B.P., Stessman, H.A.F., Sulovari, A., Geisheker, M.R., Bakken, T.E., Lake, A.M., Dougherty, J.D., Lein, E.S., Hormozdiari, F., Bernier, R.A., et al. (2019). Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat. Genet. *51*, 106–116.

Deal, S.L., and Yamamoto, S. (2018). Unraveling novel mechanisms of neurodegeneration through a large-scale forward genetic screen in Drosophila. Front. Genet. *9*, 700.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A

framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.

Diao, F., Ironfield, H., Luan, H., Diao, F., Shropshire, W.C., Ewer, J., Marr, E., Potter, C.J., Landgraf, M., and White, B.H. (2015). Plug-and-play genetic access to drosophila cell types using exchangeable exon cassettes. Cell Rep. *10*, 1410–1421.

Du, J., Lü, W., Wu, S., Cheng, Y., and Gouaux, E. (2015). Glycine receptor mechanism elucidated by electron cryo-microscopy. Nature *526*, 224–229.

Dudzic, J.P., Hanson, M.A., Iatsenko, I., Kondo, S., and Lemaitre, B. (2019). More than black or white: melanization and toll share regulatory serine proteases in Drosophila. Cell Rep. *27*, 1050–1061.e3.

Eyjolfsdottir, E., Branson, S., Burgos-Artizzu, X.P., Hoopfer, E.D., Schor, J., Anderson, D.J., and Perona, P. (2014). Detecting Social Actions of Fruit Flies. In Computer Vision – ECCV 2014, Lecture Notes in Computer Science (Springer), pp. 772–787.

Feske, S., Wulff, H., and Skolnik, E.Y. (2015). Ion channels in innate and adaptive immunity. Annu. Rev. Immunol. *33*, 291–353.

Fischbach, G.D., and Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron *68*, 192–195.

Gahl, W.A., Mulvihill, J.J., Toro, C., Markello, T.C., Wise, A.L., Ramoni, R.B., Adams, D.R., and Tifft, C.J.; UDN (2016). The NIH undiagnosed diseases program and Network: applications to modern medicine. Mol. Genet. Metab. *117*, 393–400.

Gnerer, J.P., Venken, K.J.T., and Dierick, H.A. (2015). Gene-specific cell labeling using MiMIC transposons. Nucleic Acids Res. *43*, e56.

Goodman, L.D., Cope, H., Nil, Z., Ravenscroft, T.A., Charng, W.-L., Lu, S., Tien, A.-C., Pfundt, R., Koolen, D.A., Haaxma, C.A., et al. (2021). TNPO2 variants associate with human developmental delays, neurologic deficits, and dysmorphic features and alter TNPO2 activity in Drosophila. Am. J. Hum. Genet. *108*, 1669–1691.

Guelman, S., Suganuma, T., Florens, L., Weake, V., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2006). The essential gene wda encodes a WD40 repeat subunit of Drosophila SAGA required for histone H3 acetylation. Mol. Cell Biol. *26*, 7178–7189.

Guo, H., Bettella, E., Marcogliese, P.C., Zhao, R., Andrews, J.C., Nowakowski, T.J., Gillentine, M.A., Hoekzema, K., Wang, T., Wu, H., et al. (2019). Disruptive mutations in TANC2 define a neurodevelopmental syndrome associated with psychiatric disorders. Nat. Commun. *10*, 4679–4717.

Harel, T., Yoon, W.H., Garone, C., Gu, S., Coban-Akdemir, Z., Eldomery, M.K., Posey, J.E., Jhangiani, S.N., Rosenfeld, J.A., Cho, M.T., et al. (2016). Recurrent de novo and biallelic variation of ATAD3A, encoding a mitochondrial membrane protein, results in distinct neurological syndromes. Am. J. Hum. Genet. *99*, 831–845.

Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., and Mohr, S.E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics *12*, 357–416.

Huang, Y., Mao, X., van Jaarsveld, R.H., Shu, L., Terhal, P.A., Jia, Z., Xi, H., Peng, Y., Yan, H., Yuan, S., et al. (2020). Variants in CAPZA2, a member of an F-actin capping complex, cause intellectual disability and developmental delay. Hum. Mol. Genet. *29*, 1537–1546.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216–221.

Jagadeesh, K.A., Wenger, A.M., Berger, M.J., Guturu, H., Stenson, P.D., Cooper, D.N., Bernstein, J.A., and Bejerano, G. (2016). M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat. Genet. *48*, 1581–1586.

Kabra, M., Robie, A.A., Rivera-Alba, M., Branson, S., and Branson, K. (2013). JAABA: interactive machine learning for automatic annotation of animal behavior. Nat. Methods *10*, 64–67.

Kanca, O., Andrews, J.C., Lee, P.-T., Patel, C., Braddock, S.R., Slavotinek, A.M., Cohen, J.S., Gubbels, C.S., Aldinger, K.A., Williams, J., et al. (2019). De novo variants in WDR37 are associated with epilepsy, colobomas, dysmorphism, developmental delay, intellectual disability, and cerebellar hypoplasia. Am. J. Hum. Genet. *105*, 413–424.

Kanca, O., Zirin, J., Hu, Y., Tepe, B., Dutta, D., Lin, W.W., Ma, L., Ge, M., Zuo, Z., Liu, L.P., and Levis, R.W. (2021). An expanded toolkit for Drosophila gene tagging using synthesized homology donor constructs for CRISPR mediated homologous recombination. bioRxiv.

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310–315.

Krstic, D., Boll, W., and Noll, M. (2013). Influence of the white locus on the courtship behavior of Drosophila males. PLoS One *8*, e77904.

Lee, P.-T., Zirin, J., Kanca, O., Lin, W.-W., Schulze, K.L., Li-Kroeger, D., Tao, R., Devereaux, C., Hu, Y., Chung, V., et al. (2018). A gene-specific T2A-GAL4 library for Drosophila. ELife *7*, 1377.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.

Levitt, P., and Campbell, D.B. (2009). The genetic and neurobiologic compass points toward common signaling dysfunctions in autism spectrum disorders. J. Clin. Invest. *119*, 747–754.

Li, Q., and Wang, K. (2017). InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. Am. J. Hum. Genet. *100*, 267–280.

Link, N., and Bellen, H.J. (2020). Using Drosophila to drive the diagnosis and understand the mechanisms of rare human diseases. Development *147*, dev191411.

Liu, T. (2013). Sensory processing and motor skill performance in elementary school children with autism spectrum disorder. Percept Mot. Skill *116*, 197–209.

Liu, N., Schoch, K., Luo, X., Pena, L.D.M., Bhavana, V.H., Kukolich, M.K., Stringer, S., Powis, Z., Radtke, K., Mroske, C., et al. (2018). Functional variants in TBX2 are associated with a syndromic cardiovascular and skeletal developmental disorder. Hum. Mol. Genet. *27*, 2454–2465.

Liu, X., Jian, X., and Boerwinkle, E. (2011). dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum. Mutat. *32*, 894–899.

Marcogliese, P.C., and Wangler, M.F. (2001). Drosophila as a Model for Human Diseases (Encyclopedia of Life Sciences).

Marcogliese, P.C., Shashi, V., Spillmann, R.C., Stong, N., Rosenfeld, J.A., Koenig, M.K., Martínez-Agosto, J.A., Herzog, M., Chen, A.H., Dickson, P.I., et al. (2018). IRF2BPL is associated with neurological phenotypes. Am. J. Hum. Genet. *103*, 245–260.

Massey, J.H., Chung, D., Siwanowicz, I., Stern, D.L., and Wittkopp, P.J. (2019). The yellow gene influences Drosophila male mating success through sex comb melanization. Elife *8*, e49388.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

Mehta, S.Q., Hiesinger, P.R., Beronja, S., Zhai, R.G., Schulze, K.L., Verstreken, P., Cao, Y., Zhou, Y., Tepass, U., Crair, M.C., et al. (2005). Mutations in Drosophila sec15 reveal a function in neuronal targeting for a subset of exocyst components. Neuron *46*, 219–232.

Moraga-Cid, G., Sauguet, L., Huon, C., Malherbe, L., Girard-Blanc, C., Petres, S., Murail, S., Taly, A., Baaden, M., Delarue, M., et al. (2015). Allosteric and hyperekplexic mutant phenotypes investigated on an α1 glycine receptor transmembrane structure. Proc. Natl. Acad. Sci. U S A *112*, 2865–2870.

Nagarkar-Jaiswal, S., Lee, P.-T., Campbell, M.E., Chen, K., Anguiano-Zarate, S., Gutierrez, M.C., Busby, T., Lin, W.-W., He, Y., Schulze, K.L., et al. (2015). A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila. ELife *4*, 2743.

Pauli, A., Althoff, F., Oliveira, R.A., Heidmann, S., Schuldiner, O., Lehner, C.F., Dickson, B.J., and Nasmyth, K. (2008). Cell-type-specific TEV protease cleavage reveals cohesin functions in Drosophila neurons. Developmental Cell *14*, 239–251.

Perenthaler, E., Nikoncuk, A., Yousefi, S., Berdowski, W.M., Alsagob, M., Capo, I., van der Linde, H.C., van den Berg, P., Jacobs, E.H., Putar, D., et al. (2020). Loss of UGP2 in brain leads to a severe epileptic encephalopathy, emphasizing that bi-allelic isoform-specific start-loss mutations of essential genes can cause genetic diseases. Acta Neuropathol. *139*, 415–442.

Pinto, D., Pagnamenta, A.T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T.R., Correia, C., Abrahams, B.S., et al. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. Nature *466*, 368–372.

Post, K.L., Belmadani, M., Ganguly, P., Meili, F., Dingwall, R., McDiarmid, T.A., Meyers, W.M., Herrington, C., Young, B.P., Callaghan, D.B., et al. (2020). Multi-model functionalization of disease-associated PTEN missense mutations identifies multiple molecular mechanisms underlying protein dysfunction. Nat. Commun. *11*, 2073.

Ravenscroft, T.A., Phillips, J.B., Fieg, E., Bajikar, S.S., Peirce, J., Wegner, J., Luna, A.A., Fox, E.J., Yan, Y.-L., Rosenfeld, J.A., et al. (2021). Heterozygous loss-of-function variants significantly expand the phenotypes associated with loss of GDF11. Genet. Med. *23*, 1–12.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of medical genetics and genomics and the association for molecular Pathology. Genet. Med. *17*, 405–423.

Rubeis, S.D., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209–215.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature *485*, 237–241.

Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., Rubeis, S.D., An, J.-Y., Peng, M., Collins, R., Grove, J., Klei, L., et al. (2020). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. Cell *180*, 568–584.e23.

Seeds, A.M., Ravbar, P., Chung, P., Hampel, S., Midgley, F.M., Mensh, B.D., and Simpson, J.H. (2014). A suppression hierarchy among competing motor programs drives sequential grooming in Drosophila. ELife *3*, e02951.

Simon, J.C., and Dickinson, M.H. (2010). A new chamber for studying the behavior of Drosophila. PLoS One *5*, e8793.

Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.

Splinter, K., Adams, D.R., Bacino, C.A., Bellen, H.J., Bernstein, J.A., Cheatle-Jarvela, A.M., Eng, C.M., Esteves, C., Gahl, W.A., Hamid, R., et al. (2018). Effect of genetic diagnosis on patients with previously undiagnosed disease. New Engl. J. Med. *379*, 2131–2139.

Deciphering Developmental Disorders Study (2017). Prevalence and architecture of de novo mutations in developmental disorders. Nature *542*, 433–438.

Sullivan, P.F., and Geschwind, D.H. (2019). Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. Cell *177*, 162–183.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. Plos One *6*, e21800.

Takata, A., Miyake, N., Tsurusaki, Y., Fukai, R., Miyatake, S., Koshimizu, E., Kushima, I., Okada, T., Morikawa, M., Uno, Y., et al. (2018). Integrative analyses of de novo mutations provide deeper biological insights into autism spectrum disorder. Cell Rep. *22*, 734–747.

Tang, W., Ehrlich, I., Wolff, S.B.E., Michalski, A.-M., Wölfl, S., Hasan, M.T., Lüthi, A., and Sprengel, R. (2009). Faithful expression of multiple proteins via 2A-peptide self-processing: a versatile and reliable method for manipulating brain circuits. J. Neurosci. : Official J. Soc. Neurosci. *29*, 8621–8629.

MGC Project Team, Temple, G., Gerhard, D.S., Rasooly, R., Feingold, E.A., Good, P.J., Robinson, C., Mandich, A., Derge, J.G., Lewis, J., et al. (2009). The completion of the mammalian gene collection (MGC). Genome Res. *19*, 2324–2333.

Tsang, Y.H., Dogruluk, T., Tedeschi, P.M., Wardwell-Ozgo, J., Lu, H., Espitia, M., Nair, N., Minelli, R., Chong, Z., Chen, F., et al. (2016). Functional annotation of rare gene aberration drivers of pancreatic cancer. Nat. Commun. *7*, 10500–10511.

Vaser, R., Adusumalli, S., Leng, S.N., Sikic, M., and Ng, P.C. (2016). SIFT missense predictions for genomes. Nat. Protoc. *11*, 1–9.

de Velasco, B., Erclik, T., Shy, D., Sclafani, J., Lipshitz, H., McInnes, R., and Hartenstein, V. (2007). Specification and development of the pars intercerebralis and pars lateralis, neuroendocrine command centers in the Drosophila brain. Developmental Biol. *302*, 309–323.

Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747–1751.

Venken, K.J.T., Schulze, K.L., Haelterman, N.A., Pan, H., He, Y., Evans-Holm, M., Carlson, J.W., Levis, R.W., Spradling, A.C., Hoskins, R.A., et al. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nat. Methods *8*, 737–743.

Verstreken, P., Koh, T.-W., Schulze, K.L., Zhai, R.G., Hiesinger, P.R., Zhou, Y., Mehta, S.Q., Cao, Y., Roos, J., and Bellen, H.J. (2003). Synaptojanin is recruited by endophilin to promote synaptic vesicle uncoating. Neuron *40*, 733–748.

Vetro, A., Pisano, T., Chiaro, S., Procopio, E., Guerra, A., Parrini, E., Mei, D., Virdò, S., Mangone, G., Azzari, C., et al. (2020). Early infantile epileptic-dyskinetic encephalopathy due to biallelic PIGP mutations. Neurol. Genet. *6*, e387.

Wang, J., Al-Ouran, R., Hu, Y., Kim, S.-Y., Wan, Y.-W., Wangler, M.F., Yamamoto, S., Chao, H.-T., Comjean, A., Mohr, S.E., et al. (2017). MARRVEL: integration of human and model organism genetic resources to facilitate functional annotation of the human genome. Am. J. Hum. Genet. *100*, 843–853.

Wong, W.-R., Brugman, K.I., Maher, S., Oh, J.Y., Howe, K., Kato, M., and Sternberg, P.W. (2019). Autism-associated missense genetic variants impact locomotion and neurodevelopment in Caenorhabditis elegans. Hum. Mol. Genet. *28*, 2271–2281.

Xiong, B., Bayat, V., Jaiswal, M., Zhang, K., Sandoval, H., Charng, W.-L., Li, T., David, G., Duraine, L., Lin, Y.-Q., et al. (2012). Crag is a GEF for Rab11 required for rhodopsin trafficking and maintenance of adult photoreceptor cells. PLoS Biol. *10*, e1001438.

Yuan, S., Chan, H.C.S., Filipek, S., and Vogel, H. (2016). PyMOL and inkscape bridge the data and the data visualization. Structure *24*, 2041–2042.

Yuen, R.K.C., Merico, D., Bookman, M., Howe, J.L., Thiruvahindrapuram, B., Patel, R.V., Whitney, J., Deflaux, N., Bingham, J., Wang, Z., et al. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat. Neurosci. *20*, 602–611.

Zeilhofer, H.U., Acuña, M.A., Gingras, J., and Yévenes, G.E. (2018). Glycine receptors and glycine transporters: targets for novel analgesics? Cell Mol. Life Sci. : CMLS *75*, 447–465.

Zhang, Y., Ho, T.N.T., Harvey, R.J., Lynch, J.W., and Keramidas, A. (2017). Structure-function analysis of the GlyR $\alpha 2$ subunit autism mutation p.R323L reveals a gain-of-function. Front. Mol. Neurosci. *10*, 158.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-elav	DSHB Cat# 7E8A10	RRID:AB_528218
anti-Repo	DSHB Cat# 8D12	RRID:AB_528448
anti-mouse-647	Jackson ImmunoResearch Labs Cat# 715-605-151	RRID:AB_2340863
anti-rat-Cy3	Jackson ImmunoResearch Labs Cat# 712-165-153	RRID:AB_2340667
anti-GFP	Abcam Cat# ab6662	RRID:AB_305635
anti-HA	BioLegend Cat# 902301	RRID:AB_2565018
anti-actin	Millipore Cat# MAB1501	RRID:AB_2223041
Experimental models: Organisms/strains		
UAS-2xEGFP, hs-Cre, vas-dφC31	Diao et al., 2015	N/A
Tub-Gal4	BDSC	RRID: BDSC_5138
nubbin-GAL4	BDSC	RRID: BDSC_51635
GMR-GAL4	BDSC	RRID: BDSC_1104
UAS-nls.GFP	BDSC	RRID: BDSC_4775
UAS-nls.mCherry	BDSC	RRID: BDSC_38424
nSyb-GAL4	BDSC	RRID: BDSC_51635
Rh1-GAL4	BDSC	RRID: BDSC_8691
pnr-GAL4	BDSC	RRID: BDSC_3039
$CG4562^{TG4}$	BDSC	RRID:BDSC_76740
Abl^{TG4}	BDSC	RRID:BDSC_67429
$Adk1^{TG4}$	Lee et al., 2018	N/A
$CG7470^{TG4}$	BDSC	RRID:BDSC_76749
$CG8665^{TG4}$	BDSC	RRID:BDSC_66811
$RhoGAP19D^{TG4}$	BDSC	RRID:BDSC_76687
osa^{TG4}	Lee et al., 2018	N/A
$PMCA^{TG4}$	BDSC	RRID:BDSC_76741
Cat^{TG4}	BDSC	RRID:BDSC_76660
$Cep135^{TG4}$	BDSC	RRID:BDSC_66853
$CG31637^{TG4}$	BDSC	RRID:BDSC_76647
$ClC\text{-}a^{TG4}$	BDSC	RRID:BDSC_66801
$CLIP\text{-}190^{TG4}$	BDSC	RRID:BDSC_66834
$Dh44\text{-}R2^{TG4}$	BDSC	RRID:BDSC_66865
b^{TG4}	BDSC	RRID:BDSC_76724
arm^{TG4}	BDSC	RRID:BDSC_66903
mbc^{TG4}	BDSC	RRID:BDSC_66840
spg^{TG4}	BDSC	RRID:BDSC_76205
$CG17684^{TG4}$	Lee et al., 2018	N/A
$shot^{TG4}$	BDSC	RRID:BDSC_76760
fne^{TG4}	BDSC	RRID:BDSC_77796
dom^{TG4}	BDSC	RRID:BDSC_76192
Eph^{TG4}	BDSC	RRID:BDSC_66800
fry^{TG4}	BDSC	RRID:BDSC_76736

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Gclc^{TG4}	BDSC	RRID:BDSC_76654
dally^{TG4}	BDSC	RRID:BDSC_66830
GluRIB^{TG4}	BDSC	RRID:BDSC_76135
Nmdar2^{TG4}	BDSC	RRID:BDSC_76705
Gprk2^{TG4}	BDSC	RRID:BDSC_66828
Lerp^{TG4}	BDSC	RRID:BDSC_77798
Pits^{TG4}	BDSC	RRID:BDSC_77731
if^{TG4}	BDSC	RRID:BDSC_66867
Lpt^{TG4}	BDSC	RRID:BDSC_76714
wb^{TG4}	BDSC	RRID:BDSC_76189
Lrch^{TG4}	BDSC	RRID:BDSC_77756
LRP1^{TG4}	BDSC	RRID:BDSC_76640
Rab3-GEF^{TG4}	BDSC	RRID:BDSC_76623
hppy^{TG4}	BDSC	RRID:BDSC_67447
mbl^{TG4}	BDSC	RRID:BDSC_66779
msn^{TG4}	BDSC	RRID:BDSC_76204
Mhc^{TG4}	BDSC	RRID:BDSC_76653
ck^{TG4}	BDSC	RRID:BDSC_76720
Nlg3^{TG4}	BDSC	RRID:BDSC_76134
Cad99C^{TG4}	BDSC	RRID:BDSC_67483
a^{TG4}	BDSC	RRID:BDSC_76725
Pli^{TG4}	BDSC	RRID:BDSC_77693
Piezo^{TG4}	BDSC	RRID:BDSC_76658
Ptx1^{TG4}	BDSC	RRID:BDSC_67497
l(1)G0289^{TG4}	BDSC	RRID:BDSC_67467
CG31211^{TG4}	BDSC	RRID:BDSC_76718
Spn^{TG4}	This study	N/A
CG6767^{TG4}	BDSC	RRID:BDSC_77797
otk^{TG4}	BDSC	RRID:BDSC_76759
Lar^{TG4}	BDSC	RRID:BDSC_67451
Pxn^{TG4}	BDSC	RRID:BDSC_66850
CG5521^{TG4}	BDSC	RRID:BDSC_76180
emp^{TG4}	BDSC	RRID:BDSC_66904
sdk^{TG4}	BDSC	RRID:BDSC_76628
retm^{TG4}	BDSC	RRID:BDSC_66816
Sema5c^{TG4}	BDSC	RRID:BDSC_77809
CG6293^{TG4}	BDSC	RRID:BDSC_76761
CG18304^{TG4}	BDSC	RRID:BDSC_76128
rols^{TG4}	BDSC	RRID:BDSC_76150
CG7744^{TG4}	BDSC	RRID:BDSC_76662
Trpm^{TG4}	BDSC	RRID:BDSC_77748
Nipped-A^{TG4}	BDSC	RRID:BDSC_76723
gig^{TG4}	BDSC	RRID:BDSC_67515
Tusp^{TG4}	BDSC	RRID:BDSC_66798
unc80^{TG4}	BDSC	RRID:BDSC_76686
Usp30^{TG4}	BDSC	RRID:BDSC_76704
bchs^{TG4}	BDSC	RRID:BDSC_76762
CG6225^{TG4}	BDSC	RRID:BDSC_76769

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Yip1d1^{TG4}	BDSC	RRID:BDSC_67492
Ance^{TG4}	BDSC	RRID:BDSC_76676
Ace^{TG4}	BDSC	RRID:BDSC_76688
Aldh-III^{TG4}	BDSC	RRID:BDSC_77692
CG33298^{TG4}	BDSC	RRID:BDSC_76700
IRSp53^{TG4}	BDSC	RRID:BDSC_67637
Best1^{TG4}	BDSC	RRID:BDSC_76671
tok^{TG4}	BDSC	RRID:BDSC_76679
CASK^{TG4}	BDSC	RRID:BDSC_76631
Ddr^{TG4}	This study	N/A
cv-c^{TG4}	This study	N/A
CRMP^{TG4}	This study	N/A
CG11594^{TG4}	This study	N/A
GluClalpha^{TG4}	BDSC	RRID:BDSC_77841
Galphao^{TG4}	This study	N/A
5-HT1B^{TG4}	BDSC	RRID:BDSC_76668
Itp-r83A^{TG4}	This study	N/A
Shal^{TG4}	This study	N/A
Kdm2^{TG4}	This study	N/A
Pvr^{TG4}	BDSC	RRID:BDSC_76657
beta-Man^{TG4}	This study	N/A
drpr^{TG4}	This study	N/A
Smr^{TG4}	BDSC	RRID:BDSC_76743
Ndg^{TG4}	BDSC	RRID:BDSC_76768
Nos^{TG4}	BDSC	RRID:BDSC_76766
SIFaR^{TG4}	BDSC	RRID:BDSC_76670
NetB^{TG4}	BDSC	RRID:BDSC_76730
PH4alphaEFB^{TG4}	BDSC	RRID:BDSC_76678
PCB^{TG4}	BDSC	RRID:BDSC_66832
Pdk^{TG4}	BDSC	RRID:BDSC_77785
PKD^{TG4}	BDSC	RRID:BDSC_76706
Rim^{TG4}	This study	N/A
CG9098^{TG4}	BDSC	RRID:BDSC_76763
ctrip^{TG4}	BDSC	RRID:BDSC_76764
CG1815^{TG4}	This study	N/A
UAS-ABCC4	BDSC	RRID:BDSC_78481
UAS-ABCC4.M276V	BDSC	RRID:BDSC_92726
UAS-ABCC5	BDSC	RRID:BDSC_78508
UAS-ABCC5.R697W	BDSC	RRID:BDSC_92728
UAS-ABCC5.T1046M	BDSC	RRID:BDSC_92727
UAS-ABL2.A1084T	BDSC	RRID:BDSC_92729
UAS-ABL2	BDSC	RRID:BDSC_78453
UAS-ACE.Y818C	BDSC	RRID:BDSC_92730
UAS-ACE	This study	N/A
UAS-ACHE	BDSC	RRID:BDSC_78466
UAS-ACHE.G151R	BDSC	RRID:BDSC_92732
UAS-ACHE.P548L	BDSC	RRID:BDSC_92731
UAS-AK1	BDSC	RRID:BDSC_78462

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UAS-AK1.S58L	BDSC	RRID:BDSC_92733
UAS-ALDH18A1.D703H.HA	BDSC	RRID:BDSC_92734
UAS-ALDH18A1.HA	BDSC	RRID:BDSC_78488
UAS-ALDH1L1.N900H	BDSC	RRID:BDSC_92735
UAS-ALDH1L1	This study	N/A
UAS-ALDH3A1.F402L.HA	BDSC	RRID:BDSC_92926
UAS-ALDH3A1.HA	BDSC	RRID:BDSC_78497
UAS-ATP10A	This study	N/A
UAS-ATP2B2.T818M	BDSC	RRID:BDSC_92952
UAS-ATP2B2	This study	N/A
UAS-ATP2B4	BDSC	RRID:BDSC_78455
UAS-BAIAP2L1.A481V	BDSC	RRID:BDSC_92922
UAS-BAIAP2L1	BDSC	RRID:BDSC_78446
UAS-BCHE.HA	This study	N/A
UAS-BCHE.F433V.HA	This study	N/A
UAS-BEST3	This study	N/A
UAS-BEST3.R130S	This study	N/A
UAS-BMP1.G927S.HA	BDSC	RRID:BDSC_92931
UAS-BMP1.HA	BDSC	RRID:BDSC_77944
UAS-CAMK2A	This study	N/A
UAS-CAMK2A.E183V	This study	N/A
UAS-CARS.HA	BDSC	RRID:BDSC_79001
UAS-CARS.N348S.HA	BDSC	RRID:BDSC_92918
UAS-CASK	This study	N/A
UAS-CAT.HA	BDSC	RRID:BDSC_78471
UAS-CAT.G204E.HA	This study	N/A
UAS-CEP135.HA	BDSC	RRID:BDSC_78458
UAS-CEP135.S947P.HA	This study	N/A
UAS-CHST2.HA	BDSC	RRID:BDSC_78472
UAS-CHST2.R52P.HA	This study	N/A
UAS-CLCNKB.HA	BDSC	RRID:BDSC_77934
UAS-CLCNKB.M176I.HA	BDSC	RRID:BDSC_92910
UAS-CLIP2.G13W.HA	BDSC	RRID:BDSC_92934
UAS-CLIP2.HA	BDSC	RRID:BDSC_78461
UAS-CSAD.HA	BDSC	RRID:BDSC_78498
UAS-CSAD.A411V.HA	This study	N/A
UAS-CTNNB1	BDSC	RRID:BDSC_78496
UAS-CTNNB1.T551M	This study	N/A
UAS-DDR2.B	BDSC	RRID:BDSC_78483
UAS-DLC1.HA	BDSC	RRID:BDSC_78089
UAS-DPP6	This study	N/A
UAS-DPYSL2.HA	BDSC	RRID:BDSC_77938
UAS-DPYSL2.R496C.HA	BDSC	RRID:BDSC_92946
UAS-DPYSL3.HA	BDSC	RRID:BDSC_77939
UAS-DPYSL3.V139I.HA	BDSC	RRID:BDSC_92945
UAS-ELAVL3	BDSC	RRID:BDSC_78505
UAS-ELAVL3.L186P	BDSC	RRID:BDSC_92912
UAS-EP400.B	BDSC	RRID:BDSC_78493

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UAS-EPHA1.HA	BDSC	RRID:BDSC_77931
UAS-EPHA1.V567I.HA	BDSC	RRID:BDSC_92963
UAS-EPHB1	BDSC	RRID:BDSC_78473
UAS-EPHB1.V916M	BDSC	RRID:BDSC_92932
UAS-EPT1	BDSC	RRID:BDSC_78485
UAS-EPT1.H82N	This study	N/A
UAS-EXD2.E513D.HA	BDSC	RRID:BDSC_92942
UAS-EXD2.HA	BDSC	RRID:BDSC_77933
UAS-FGGY	BDSC	RRID:BDSC_78091
UAS-GCLC.HA	This study	N/A
UAS-GCLC.R128W.HA	This study	N/A
UAS-GLRA2.HA	BDSC	RRID:BDSC_77954
UAS-GLRA2.N136S.HA	BDSC	RRID:BDSC_92915
UAS-GNAO1.HA	BDSC	RRID:BDSC_79003
UAS-GPC5.HA	BDSC	RRID:BDSC_77936
UAS-GPC5.M133T.HA	BDSC	RRID:BDSC_92913
UAS-GRIA1.HA	BDSC	RRID:BDSC_78474
UAS-GRIA1.R218H.HA	BDSC	RRID:BDSC_92930
UAS-GRIK5	BDSC	RRID:BDSC_77957
UAS-GRK4.HA	BDSC	RRID:BDSC_78503
UAS-GRK4.P385A.HA	BDSC	RRID:BDSC_92916
UAS-HTR1D.HA	This study	N/A
UAS-HTR1D.T99N.HA	This study	N/A
UAS-IGF2R.HA	BDSC	RRID:BDSC_78454
UAS-IRF2BPL	BDSC	RRID:BDSC_78509
UAS-IRF2BPL.F30L	This study	N/A
UAS-IRF2BPL.N701X	This study	N/A
UAS-ITGA8.HA	BDSC	RRID:BDSC_78501
UAS-ITGA8.R748C.HA	BDSC	RRID:BDSC_92924
UAS-JUP.HA	BDSC	RRID:BDSC_78500
UAS-JUP.N690S.HA	This study	N/A
UAS-KCND3.HA	BDSC	RRID:BDSC_78475
UAS-KCND3.R86P.HA	BDSC	RRID:BDSC_92921
UAS-KCNH8.HA	BDSC	RRID:BDSC_78470
UAS-KDM2A	BDSC	RRID:BDSC_78476
UAS-KDM2A.R449K	BDSC	RRID:BDSC_92927
UAS-KDR	BDSC	RRID:BDSC_78451
UAS-KDR.D1171N	BDSC	RRID:BDSC_92955
UAS-LAMA2	BDSC	RRID:BDSC_79000
UAS-LRCH4	BDSC	RRID:BDSC_78477
UAS-LRCH4.V42M	BDSC	RRID:BDSC_92928
UAS-MADD	BDSC	RRID:BDSC_78457
UAS-MADD.R514C	BDSC	RRID:BDSC_92933
UAS-MANBA	BDSC	RRID:BDSC_77960
UAS-MAP4K1.HA	BDSC	RRID:BDSC_77946
UAS-MAP4K1.M725T.HA	This study	N/A
UAS-MBNL1.HA	BDSC	RRID:BDSC_78467
UAS-MBNL1.V45A.HA	BDSC	RRID:BDSC_92908

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UAS-MEGF11	BDSC	RRID:BDSC_78460
UAS-MEGF11.R911C	This study	N/A
UAS-MINK1.HA	BDSC	RRID:BDSC_78489
UAS-MINK1.C269R.HA	This study	N/A
UAS-MYH9.HA	BDSC	RRID:BDSC_79002
UAS-MYH9.R1571Q.HA	This study	N/A
UAS-NCOR1	BDSC	RRID:BDSC_78486
UAS-NCOR1.P569S	BDSC	RRID:BDSC_92953
UAS-NID2	This study	N/A
UAS-NLGN1	This study	N/A
UAS-NLGN1.H795Y	BDSC	RRID:BDSC_92936
UAS-NLGN3.R195W	This study	N/A
UAS-NOS3	This study	N/A
UAS-NPFFR2	BDSC	RRID:BDSC_78478
UAS-NPFFR2.M163I	BDSC	RRID:BDSC_92917
UAS-NR2F1	BDSC	RRID:BDSC_77959
UAS-NR2F1.R404H	This study	N/A
UAS-NTN1.A449D	BDSC	RRID:BDSC_92958
UAS-NTN1	BDSC	RRID:BDSC_78495
UAS-NTN5	BDSC	RRID:BDSC_78507
UAS-P4HA2.G153E.HA	BDSC	RRID:BDSC_92938
UAS-P4HA2.HA	BDSC	RRID:BDSC_77935
UAS-PC.HA	BDSC	RRID:BDSC_77928
UAS-PC.P1042R.HA	This study	N/A
UAS-PDGFRB.A366T.HA	BDSC	RRID:BDSC_92941
UAS-PDGFRB	This study	N/A
UAS-PDK2.HA	BDSC	RRID:BDSC_77949
UAS-PDK2.R120Q	BDSC	RRID:BDSC_92962
UAS-PEAR1	BDSC	RRID:BDSC_77956
UAS-PEAR1.T824I	BDSC	RRID:BDSC_92911
UAS-PELI1.A270V.HA	BDSC	RRID:BDSC_92919
UAS-PELI1.HA	BDSC	RRID:BDSC_78506
UAS-PITX1.HA	BDSC	RRID:BDSC_92043
UAS-PITX1.L242F.HA	BDSC	RRID:BDSC_92914
UAS-PLXDC1	BDSC	RRID:BDSC_78504
UAS-PLXDC1.R42Q	BDSC	RRID:BDSC_92909
UAS-PPP1R9A	BDSC	RRID:BDSC_78499
UAS-PRKD1	BDSC	RRID:BDSC_78494
UAS-PRKD1.R441W	BDSC	RRID:BDSC_92950
UAS-PRPS1L1	This study	N/A
UAS-PRPS1L1.G61D	BDSC	RRID:BDSC_92929
UAS-PTK7.R570Q	This study	N/A
UAS-PTPRF	BDSC	RRID:BDSC_92404
UAS-PTPRF.S334R	BDSC	RRID:BDSC_92937
UAS-PXDN.HA	BDSC	RRID:BDSC_77955
UAS-PXDN.R643Q	BDSC	RRID:BDSC_92951
UAS-RALGAPA1	BDSC	RRID:BDSC_78449
UAS-RALGAPA1.LL1769L	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UAS-RIMS2.HA	BDSC	RRID:BDSC_77937
UAS-SCARB2.HA	BDSC	RRID:BDSC_77929
UAS-SCARB2.V173A.HA	This study	N/A
UAS-SDK2.HA	BDSC	RRID:BDSC_78487
UAS-SEC14L5	BDSC	RRID:BDSC_78092
UAS-SH2D3C.HA	BDSC	RRID:BDSC_78492
UAS-SH2D3C.R227Q.HA	This study	N/A
UAS-SLC23A1	BDSC	RRID:BDSC_78459
UAS-SLC23A1.L465M	BDSC	RRID:BDSC_92943
UAS-SLC8A2	BDSC	RRID:BDSC_78490
UAS-SLC8A2.G792R	This study	N/A
UAS-SLCO4A1	BDSC	RRID:BDSC_78479
UAS-SLCO4A1.V679I	BDSC	RRID:BDSC_92920
UAS-SOGA3.HA	BDSC	RRID:BDSC_78469
UAS-SOGA3..R873P.HA	This study	N/A
UAS-SRCAP	BDSC	RRID:BDSC_78450
UAS-SRCAP.2137del	This study	N/A
UAS-SRCAP.G1937S	This study	N/A
UAS-TANC2	BDSC	RRID:BDSC_78452
UAS-TANC2.H1689R	BDSC	RRID:BDSC_92925
UAS-TRIP12	BDSC	RRID:BDSC_78518
UAS-TRIP12.R1643Q	BDSC	RRID:BDSC_92956
UAS-TRPM1	BDSC	RRID:BDSC_78517
UAS-TRPM1.F794L	BDSC	RRID:BDSC_92954
UAS-TRPM6.HA	BDSC	RRID:BDSC_77958
UAS-TRPM6.A641E.HA	BDSC	RRID:BDSC_92939
UAS-TRPM6.T2011P.HA	BDSC	RRID:BDSC_92940
UAS-TRPM7.HA	BDSC	RRID:BDSC_78447
UAS-TRPM7.T379A.HA	BDSC	RRID:BDSC_92923
UAS-TSC2.HA	BDSC	RRID:BDSC_78465
UAS-TSC2.R1557W.HA	BDSC	RRID:BDSC_92949
UAS-TSC2.R548M.HA	BDSC	RRID:BDSC_92947
UAS-TULP4	BDSC	RRID:BDSC_92408
UAS-USP30.HA	BDSC	RRID:BDSC_78480
UAS-USP30.P200S.HA	BDSC	RRID:BDSC_92957
UAS-YIPF5.HA	BDSC	RRID:BDSC_82197
UAS-ZMYND8.HA	BDSC	RRID:BDSC_77945
Deposited data		
TG4 imaging videos	This paper: Mendeley Data	https://doi.org/10.1016/10.17632/64jrz799sb.1
Human GLRA2 variant p.Thr296Met	ClinVar	SCV002055997
Human GLRA2 variant p.Phe47Ser	ClinVar	SCV002056017
Human GLRA2 variant p.Ile259Met	ClinVar	SCV002056018
Human GLRA2 variant p.Arg252Cys	ClinVar	SCV002056022
Human GLRA2 variant p.Ala288Thr	ClinVar	SCV002056021
Human GLRA2 variant p.Pro396Thr	ClinVar	SCV002056020
Human GLRA2 variant p.Pro400Leu	ClinVar	SCV002056019
Human GLRA2 variant p.Arg445Gln	ClinVar	SCV002056023

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
pUASg-HA.attB	Drosophila Genomics Resource Center,	DGRC_ 1423
Software and algorithms		
Prism8	Graph Pad	https://www.graphpad.com
Zen Blue	Zeiss	https://www.zeiss.com/microscopy/us/products/microscope-software/zen-lite.html
Zen Black	Zeiss	https://www.zeiss.com/microscopy/us/products/microscope-software/zen-lite.html
Snapgene	Snapgene	https://www.snapgene.com
Imaris	Imaris	https://imaris.oxinst.com/
Lab Chart	Ad instruments	https://www.adinstruments.com/products/labchart

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Shinya Yamamoto (yamamoto@bcm.edu).

Materials availability

The fly lines generated in this study have been deposited to the Bloomington *Drosophila* Stock Center (BDSC). Please see key resources table for unique identifiers. Please contact the Lead Contact for further information.

Data and code availability

- Confocal imaging movies of neuron/glia colocalization with *TG4>UAS-nls.mCherry* have been deposited at Mendeley Data and are publicly available as of the date of publication. The DOI is listed in the key resources table. De-identified GLRA2 variant data have been deposited at ClinVar. They are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This study didn't generate any code.
- Any additional information required to reanalyze the data reported in this work paper is available from the Lead Contact upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Generation of TG4 lines

All *TG4* alleles in this study were generated by ϕC31-mediated recombination-mediated cassette exchange of MiMIC (Minos mediated integration cassette) insertion lines (Gnerer et al., 2015; Nagarkar-Jaiswal et al., 2015; Venken et al., 2011). Conversion of the original MiMIC element was performed via genetic by crossing *UAS-2xEGFP, hs-Cre,vas-dϕC31, Trojan T2A-GAL4* triplet flies to each MiMIC strain and following a crossing scheme (Diao et al., 2015). 73 TG4 lines were described previously but not extensively characterized (Lee et al., 2018), while 35 lines were generated specifically for this study.

Generation of UAS-human cDNA lines

The majority of reference human cDNA clones were obtained in either pDONR221 or pDONR223 donor vectors. The LR clonase II (ThermoFisher) enzyme was used to shuttle ORFs into the p.UASg-HA.attB destination vector via Gateway™ cloning. Some ORFs that were not Gateway compatible were obtained from additional sources (Table S2), amplified with flanking *attB* sites and cloned into pDONR223 plasmid using BP clonase II (ThermoFisher). Sequence-verified variants were generated in the DONR vectors by either site-directed mutagenesis (SDM) or High-Throughput Mutagenesis (HiTM) as previously described (Tsang et al., 2016). SDM was performed with primers generated using NEBaseChanger (Table S3) with the Q5® mutagenesis kit (NEB). Sequence-verified reference and variant ORFs in the pUASg-HA.attB destination plasmid were microinjected into ~200 embryos in one three *attP* docking sites (*attP86Fb*, *VK00037* or *VK00033*) docking sites by ϕC31 mediated transgenesis (Bischof et al., 2007; Venken et al., 2006). The docking site of choice were selected based on the genomic locus of the corresponding fly gene. In principal, VK00037 docking site on the 2nd chromosome was used for human genes that correspond to fly genes on the X, 3rd or 4th chromosome,

whereas VK00033 or attP86Fb docking site on the 3rd chromosome was used for human genes that correspond to fly genes on the 2nd chromosome.

Fly husbandry

Unless otherwise noted, all flies used in experiments were grown in a temperature and humidity-controlled incubator at 25°C and 50% humidity on a 12-hour light/dark cycle. Some experiments were conducted at different temperatures that are specifically indicated in the text and figures. Stocks were reared on standard fly food (water, yeast, soy flour, cornmeal, agar, corn syrup, and propionic acid) at room temperature (~22°C) and routinely maintained.

Fly stocks used that were not generated here

tub-GAL4 (*y*^{1} *w***; P{w[+mC]=tubP-GAL4}LL7/TM3, Sb*^{1} *Ser*^{1}) BDSC_5138, *GMR-GAL4* (*w***; P{w[+mC]=GAL4-ninaE.GMR}12*) BDSC_1104, *nub-GAL4* (*P{GawB}nubbin-AC-62*) (Calleja et al., 2000), *nSyb-GAL4* (*y*^{1} *w***; P{nSyb-GAL4.S}3*) BDSC_51635, *Rh1-GAL4* (*P{ry[+t7.2]=rh1-GAL4}3, ry[506])* BDSC_8691, *pnr-GAL4* (*y*^{1} *w*^{1118}*; P{w[+mW.hs]= GawB}pnr[MD237]/TM3, P{w[+mC]=UAS-y.C}MC2, Ser*^{1}) BDSC_3039, UAS-nlsGFP (*w*^{1118}*; P{w[+mC]=UAS-GFP.nls}14*) BDSC_4775, and UAS-nls.mCherry (*w***; P{w[+mC]=UAS-mCherry.NLS}3*) BDSC_38424.

Patient recruitment and consent

Affected individuals were investigated by their referring physicians at local sites. Prior to research studies, informed written consent for testing and publication was obtained according to the institutional review boards (IRB) and ethnics committees of each institution. Individuals who were ascertained in diagnostic testing procedures (and/or their legal guardians) gave clinical written informed consent for testing, and their permission for inclusion of their anonymized data in this cohort series. This was obtained using standard forms at each local site by the responsible referring physicians.

GLRA2 subject case histories

Subject 1 is an 8 years old female with global developmental and cognitive delay. Pregnancy was naturally conceived and uncomplicated, other than decreased fetal movements noted by the mother. She was delivered at term (39 weeks gestational age) via c/section due to breech presentation. Birth weight was 3,600 grams. Neonatal period was uneventful. There were no feeding difficulties and her growth remained within the normal limits. She was delayed with all her milestones but most significantly for speech (walked at 18 months, scribbled with a crayon at 3.5 years, first words at 24 months and combined words to sentences at 4-5 years of age). She was diagnosed with mixed expressive-receptive speech delay and received speech therapy, occupational therapy and physical therapy interventions. In school she exhibits learning problems, inattention and is below her grade level. She has a modified curriculum and is receiving resources in reading and math. There is no history of developmental regression or seizures. The medical history is otherwise significant for nystagmus that was first noted in infancy and improved with age, as well as myopia and astigmatism requiring corrective glasses. Family ethnicity is Hispanic and the family history is non-contributory. The patient had a normal brain MRI (magnetic resonance imaging) at 6 months of age. EEG (electroencephalography) at 4 years of age showed a slow and poorly formed background, indicative of mild encephalopathy, but did not detect epileptiform activity. Genetic testing included: mitochondrial DNA sequencing which detected a pathogenic variant m.13042 G>A though at heteroplasmy level of 1.9%, which was felt unlikely to explain the phenotype. CMA (chromosomal microarray) was negative. Trio whole exome sequencing (WES) detected a *de novo*, heterozygous variant of unknown clinical significance in *GLRA2*, c.887C>T, p.Thr296Met (NC_000023.10, chrX: g.14627284C>T). This variant is absent in gnomAD. More recent clinical reanalysis of exome data did not detect any other candidates that may explain the phenotype.

Subject 2 is a 6 years old female with epilepsy, developmental delay (DD), mild intellectual disability (ID) and autism spectrum disorder (ASD). Pregnancy was uncomplicated and she was delivered at term (41 weeks gestational age) via vaginal delivery with vacuum extraction. The neonatal period was uneventful. At the age of 6 months, she developed a severe epileptic encephalopathy with myoclonic seizures. Seizure control was achieved with medications, and she has been seizure-free without medications since the age of about two years old. Delayed psychomotor development was noted, most significantly for her speech with a mixed expressive-receptive speech delay (non-verbal). Her ability to concentrate is poor and she displays mood swings. The medical history is otherwise significant for nystagmus that was first noted in infancy (6 weeks old) and improved with age, and sleep disturbance. She has mild microcephaly [< 1st centile: -2.84 standard deviation (SD)] and mild bilateral cutaneous 3rd-4th syndactyly, with no other congenital anomalies. Family ethnicity is European (German/Italian) and the family history is significant for a maternal aunt that had epilepsy in adulthood but her cognitive development was normal. Brain MRI showed delayed myelination at 7 months old and a small arachnoid cyst. EEG was abnormal for bilateral synchronized, sometimes high amplitude spike/polyspike-waves-complexes, and bitemporo-occipital hints for severe functional defects with epileptic potentials. Chromosomal analysis, Angelman syndrome methylation study, epilepsy next generation sequencing (NGS) gene panel and *MECP2* sequencing were negative. Trio WES detected a *de novo*, heterozygous variant of unknown clinical significance in *GLRA2*, c.887C>T, p.Thr296Met (NC_000023.10, chrX: g.14627284C>T). This variant is absent in gnomAD.

Subject 3 is a 5 year and 6 months old female with DD, microcephaly, abnormal eye movements and ataxic gait. Pregnancy was uncomplicated and she was born at term via c/section. Abnormal eye movements were noticed two weeks after birth, during

hospitalization due to a lower respiratory tract infection. At the age of 6 months, clinical examination revealed mildly delayed developmental milestones and erratic conjugate eye movements akin to opsoclonus. At age 4 years OFC (occipitofrontal circumference) was 43 cm (< 1st centile: -4.28 SD) and ophthalmological evaluation revealed alternating exotropia, for which patching therapy was initiated. Language was limited to a few words and neuropsychological evaluation documented moderate developmental delay (Bayley-III). The patient could walk unsupported with ataxic gait. At age 5 years 6 months, erratic eye movements were considerably reduced and she could walk independently but her expressive language was still limited to a few words, with delayed receptive speech and nonverbal communicative skills. Family ethnicity is European and the family history is unremarkable. Brain MRI at 6 months of age showed mild cortical atrophy with thinning of the corpus callosum. EEG, while awake and asleep, laboratory and metabolic investigations were unremarkable. Array-CGH (comparative genomic hybridization) highlighted a maternally inherited 3q25.32 duplication (chr3:157746089-158324659, hg19) that was interpreted as likely benign. Trio WES detected a *de novo* heterozygous variant in *GLRA2*, c.887C>T, p.Thr296Met (NC_000023.10, chrX: g.14627284C>T). This variant is absent in gnomAD. In addition, it detected a *de novo* variant in *CACNA1B*, c.5381C>T, p.Thr1794Met (NC_000009.11:g.141000212C>T), which is a variant of unknown significance in a gene that is linked to an autosomal recessive condition (Neurodevelopmental disorder with seizures and nonepileptic hyperkinetic movements, MIM #618497). Failure to identify a second allele in this gene reduces the likelihood that this variant is responsible for this patient's phenotype.

Subject 4 was a female infant with seizures and severe developmental delay who passed away at 7 months of age secondary to complications of COVID-19 infection. Pregnancy was uneventful and she was born at term (40 weeks gestational age). She was noted to have focal seizures at 2-3 weeks of age, and was diagnosed with infantile spasms when she was 5 months old. At 6 months of age she was not reaching for objects, not sitting up and only making high-pitched sounds. She had borderline microcephaly with dysmorphic features including midface retrusion, apparent hypotelorism, deep set eyes, thick eyebrows, downturned corners of the mouth, and wide-spaced nipples. Family ethnicity is Hispanic and the family history is non-contributory. She had normal plasma and CSF (cerebrospinal fluid) lactate, pipecolic acid and piperideine-6-carboxylate, ammonia, urine organic acids, plasma amino acids, acylcarnitine profile, and CSF amino acids. An Epilepsy gene panel was non-diagnostic. Trio WES detected a *de novo* heterozygous variant in *GLRA2*, c.887C>T, p.Thr296Met (NC_000023.10, chrX: g.14627284C>T). This variant is absent in gnomAD.

Subject 5 is a 5 years and 4 months old female with global developmental delay. She was born in Afghanistan to consanguineous parents, and there is limited information available regarding her birth history. Pregnancy was uneventful. She walked at 2 years and is currently non-verbal. She has ptosis and oculomotor apraxia. On physical exam she is noted to have broad halluces. Brain MRI at 5 years old was normal. Metabolic testing was non-revealing. Array-CGH, molecular genetic testing of *FMR1* and screening for congenital disorders of glycosylation resulted negative. WES detected a *de novo*, heterozygous variant of unknown clinical significance in *GLRA2*, c.887C>T, p.Thr296Met (NC_000023.10, chrX: g.14627284C>T). This variant is absent in gnomAD.

Subject 6 is a 9 months old female with developmental delay, West syndrome, microcephaly, up-beat nystagmus and myopia. She was born at term (38 gestational weeks, birth weight 2770 g (10^{th} percentile), OFC 34 cm (33^{rd} percentile)) via c/section due to breech presentation. She had postnatal respiratory distress. Abnormal eye movements were noticed at 6 weeks of age. Ophthalmologic assessment was normal with normal fixation apart from up-beat nystagmus. Clinical evaluation at 4 months of age revealed developmental delay and an OFC of 38.5 cm (3rd percentile). At approximately 7 months of age, she developed infantile spasms with an intermittent hypsarrhythmic pattern in EEG (West syndrome). Therapy with Sulthiame resulted in resolution of seizure activity and EEG normalisation. She has made developmental progress since starting treatment with antiepileptic medications. At 9 months of age she is showing mild gross motor delay. Family ethnicity is European, and the family history is non-contributory. Brain MRI, abdominal ultrasound, and metabolic screening labs (including plasma amino acids, acylcarnitine profile, CSF analysis and CSF neurotransmitters) were normal. Trio exome sequencing detected a heterozygous *de novo* missense variant in *GLRA2* (NM_002063.4), c.887C>T, p. Thr296Met (NC_000023.10, chrX: g.14627284C>T, hg19). This variant is absent in gnomAD.

Subject 7 is a 6 years and 7 months old female with a history of infantile spasms, epilepsy and intellectual disability. She was born at term and first presented with infantile spasms at 3 months of age. This evolved to atonic and tonic-clonic seizures as she grew up. She was delayed with all milestones (walked at 4.5 years old and remains non-verbal). She had nystagmus that improved with age and strabismus. The medical history is otherwise significant for hyperactivity, inattention and sleep disturbance. Her ethnicity is African (Senegal). Brain MRI at 3 years of age showed cortical and white matter atrophy, including vermian atrophy. EEG showed hypsarrythmia at onset and she had normal interictal EEG afterwards. She had normal SNP (single nucleotide polymorphism)-array, negative targeted epilepsy panel and negative metabolic lab results. Trio WES identified a *de novo* heterozygous variant in *GLRA2*, c.140T>C, p.Phe47Ser (NC_000023.10, chrX: g.14550432T>C). This variant is absent in gnomAD.

Subject 8 is a 5 years and 3 months old female with mild developmental delay and learning disabilities. Pregnancy was naturally conceived and uncomplicated. She was born near term following premature rupture of membrane. Birth weight was 3,280g, birth length was 50cm and OFC was 33.5cm. Apgar scores was 8/9 at $1^{st}/5^{th}$ minutes of life. She had feeding difficulties during the neonatal period. With regards to her milestones, she was able to sit unsupported at 9 months and walked independently at 17 months. Speech was not delayed but she has difficulties in pronunciation and articulation. There is no history of developmental regression. She has a School aide for learning difficulties but is interacting well with other children. She is receiving therapy for fine motor difficulties and is followed by a psychiatrist with a diagnosis of infantile psychosis. She also has a history of sleep disturbance (short sleep duration). The family history is non-contributory, she is the only child to her parents as a couple and has two paternal half-siblings that are healthy. Her growth is normal, with weight at 22 kg (+1.1 SD), height at 117.5 cm (+1.4 SD) and OFC 51 cm (+0.2 SD). On

exam she has mild dysmorphic features, including long face, pointed chin, and overlapping 1st and 2nd toes. Although she does not have clinical seizures, EEG was abnormal for left fronto-temporal spike-waves focus which diffuses in the right frontal region, activated by sleep but not meeting criteria for Epilepsy with continuous spike-wave during sleep (CSWS). Brain MRI showed nonspecific dot-like hypersignal in FLAIR of the subcortical white matter of the frontal region. Fragile X testing was negative. Exome sequencing detected a *de novo* heterozygous variant in *GLRA2*, c.777C>G, p.Ile259Met (NC_000023.10, chrX: g.14627174C>G). This variant is absent in gnomAD.

Subject 9 is an 11 months old male with hypotonia, DD and dysmorphic craniofacial features. Pregnancy was uncomplicated, he was delivered at term (38 and 3/7 weeks gestational age) and the neonatal period was uneventful. Soon after birth dysmorphic features were noted, including an elongated face, high anterior hairline, epicanthal folds, downslanting palpebral fissures and a bulbous nose. Growth remains within the normal limits. His medical history is otherwise significant for obstructive sleep apnea and strabismus. Family ethnicity is European (Dutch) and the family history is significant for the maternal grandfather who has not further specified unexplained neurological complaints, and which could not be further investigated. Investigations for metabolic disorders, Fragile X syndrome and a SNP-array were normal. Trio WES identified a rare variant in *GLRA2*, c.754C>T, p.Arg252Cys (NC_000023.10, chrX: g.14627151C>T), which was inherited from mother. No other possible disease explaining variant was identified. The mother displayed skewed X chromosome inactivation (82% on two measurements). The variant was absent in the maternal uncle and the maternal grandmother, but was inherited from the maternal grandfather, who was not available for clinical investigations. His level of functioning remains unknown. This variant is present in one heterozygous female in gnomAD.

Subject 10 is a 7 years old male with epilepsy, DD with regression, and ASD. Pregnancy was uncomplicated. He was born full term via uncomplicated delivery, and was meeting his early developmental milestones. He was speaking in sentences at 2.5 years old when he started having generalized tonic-clonic seizures. He developed staring spells, ataxia, and an increased frequency of myoclonic jerks, which around the age of 6 years old were occurring 20 times per day on average, with 5-6 atonic seizures per day each lasting less than 30 seconds. Following seizure onset, he experienced developmental regression. At 3 years of age he was diagnosed with ASD. At 6 years of age his vocabulary was about 20 words, with gains in development lost following significant seizures. His ethnicity is European, and the family history is significant for a younger brother with ASD, although he has not presented with seizures. Neither mother nor father have a history of seizures or delays. At age 3, EEG depicted generalized slowing and generalized epileptiform discharges associated with myoclonic jerks. MRI showed minimal increased T2 signal intensity on the occipital lobes that was thought to be within normal limits. Genetic testing for Fragile X syndrome, Prader-Willi/Angelman syndromes, and congenital disorders of glycosylation was normal. Additional tests, including plasma amino acids, lysosomal enzymes, and cerebral creatine deficiency were also normal. Array-CGH reported a maternally inherited 1p33 deletion of unknown significance (chr1: 48,688,391-49,922,153). The patient was enrolled to The Manton Center for Orphan Disease Gene Discovery Core protocol. Trio WES identified a maternally inherited variant in *GLRA2*, c.862G>A, p.Ala288Thr (NC_000023.10, chrX: g.14627259G>A). This variant is absent in gnomAD.

Subject 11 is a 35 years old male with a history of DD, learning disabilities and ASD. Pregnancy was uncomplicated and he was born at term (40 weeks gestational age). Since early childhood he showed slow movement and difficulties in motor coordination. He walked and said his first words at 24 months, and first sentences at 3 years of age. In school learning disabilities were noted, including difficulties in writing, reading, praxias, temporal orientation, calculation, drawing, and visuo-spatial organization. He graduated high school and continued to higher education, though he did not complete a degree. Neuropsychiatric assessment in adulthood was consistent with ASD and social and cognitive deficits. There is no history of seizures. The medical history is otherwise significant for environmental allergies, myopia, and astigmatism. The ethnicity is European, and the family history is non-contributory. The patient had a normal brain MRI at 29 years old. Trio WES identified a maternally inherited variant in *GLRA2*, c.1186C>A, p.Pro396Thr (NC_000023.10, chrX: g.14748434C>A). This variant is present in 3 heterozygous females and 1 hemizygous male in gnomAD.

Subject 12 is a 15 years old male with a history of epilepsy, DD, expressive language disorder, and ASD. Pregnancy was uncomplicated and he was born at term. He walked at 17 months and said his first words at 13 months. Following febrile convulsions at 18 months of age, more pronounced developmental regression was noted. At 7 years of age, he had focal (partial) motor seizures for which he still receives antiepileptic medications. In school he is enrolled in mainstream classes, but he has learning problems and inattention, and is behind his age-peers. Speech remains delayed, he has limited vocabulary and receives speech therapy. He is frequently agitated and has social anxiety. The medical history is otherwise significant for obesity. The ethnicity is European, and the family history is significant for learning problems in his mother. On exam, he is noted to have mild dysmorphic features, including broad face, Widow's peak, horizontal and broad eyebrows, long prominent eyelashes, and a broad nasal tip. Brain MRI at 7 years of age showed increased signal intensity in FLAIR of the cortical matter of the parietal region. EEG at that age showed abnormal alpha waves with high polymorphic spikes with a focus in the right temporal region (consistent with the clinical presentation of partial simple seizures). Genetic testing for Fragile X syndrome, and metabolic labs were normal. Trio WES identified a maternally inherited variant in *GLRA2*, c.1199C>T, p.Pro400Leu (NC_000023.10, chrX: g.14748447C>T). This variant is absent in gnomAD. X chromosome inactivation study in the mother, who is mildly affected, showed moderately unbalanced X inactivation (60:40) in two independent experiments.

Subject 13 is a 6 years old male with intractable epilepsy, developmental delay, and suspected autism spectrum disorder. Pregnancy was uncomplicated and he was delivered at term (39 and 2/7 weeks gestational age, weight: 3440 g (-0.19 SD), length: 49 cm (-1.34 SD), OFC: 35 cm (-0,3 SD)). The neonatal period was uneventful. He learned to walk at the age of 21 months. He spoke

his first words at 18 months, but his speech development was delayed. At the age of 4 years and 3 months he first presented with generalized tonic-clonic and tonic seizures. At most recent evaluation he was reported to have 4-5 generalized tonic-clonic or tonic seizures, and absence seizures daily (while being treated with five anti-epileptic medications simultaneously). EEG at age 5 years and 6 months showed excessive beta-activity most likely due to medication; left and right posterior and right frontal Intermittent slowing; and bilateral polyspikes during sleep with a frequency of < 1/minute three times associated with tonic-clonic seizures. Brain MRI at the same age showed no abnormalities. Concomitant to the epilepsy, he was diagnosed with DD with a focus on speech development and cognitive impairment. He is presenting with poor expressive speech (partly able to speak in sentences). The behavioral abnormalities include a short attention span, poor impulse control, sleep disturbances and recently also some compulsive traits have been reported by the mother. An early-childhood autism spectrum disorder is suspected, and an evaluation has been initiated. He visits an integrative Kindergarten with permanent one-on-one care and receives speech therapy, occupational therapy, and physiotherapy. Growth parameters at the last assessment (5 years and 11 months) were normal (weight at 0.51 SD, height at 1.12 SD and OFC at -0,72 SD). On clinical exam, he was noted to have downslanting palpebral fissures and a high forehead. There is no family history of developmental delay and epilepsy. Neonatal metabolic screening was unremarkable. Genetic testing for Fragile X syndrome as well as array-CGH, karyotype and proband exome sequencing were non-revealing. Trio exome sequencing on a research basis identified a maternally inherited hemizygous missense variant in *GLRA2* c.1334G>A, p.Arg445Gln (NC_000023.10, chrX: g.14748582G>A). This variant is present in three heterozygous females in gnomAD and has not been observed in a hemizygous state. With Sanger sequencing, it was shown that the variant is absent in the unaffected maternal half-brother of subject 13.

METHOD DETAILS

Ortholog candidate identification

We utilized DIOPT (Hu et al., 2011) to determine the fly ortholog of the 1519 human genes from the SSC. DIOPT scores lower than 4/16 were excluded. If there were multiple fly paralogs with equal DIOPT scores we referred to the weighted score. If the weighted scores were equal, we chose the ortholog in which a Gold MiMIC present in the other paralogs of equal strength.

Electroretinograms (ERG)

ERG recordings on adult flies were performed on *nSyb-GAL4* (Pauli et al., 2008) and *Rh1-GAL4* (Xiong et al., 2012) driven *UAS-GLRA2* at 5 days post-eclosion raised at 25°C in 12h light/12h dark cycle as previously described (Verstreken et al., 2003) using LabChart software (AD instruments). 4-10 flies were examined for each genotype. Recording was repeated at least 3 times per fly. Quantification and statistical analysis was performed using ANOVA followed by Bonferroni's multiple comparison test using Prism 8.0.

Complementation test of lethality in TG4 lines

Out of the 108 *TG4* mutants generated, 65 TG4 mutants were homozygous lethal. Because lethality can be caused by disruption of the gene of interest or due to second site lethal mutations carried on the same chromosome, we performed complementation test using standard methodology. For genes on the 2nd and 3rd chromosome, female heterozygous TG4 lines balanced with either *SM6a* or *TM3, Sb, Ser*, respectively, were crossed with male flies carrying a corresponding deficiency (Df) that covers the gene of interest (see Table S4). Three independent crosses were set at 25°C for each TG4 line and we determined if any *TG4* flies survived to the adult stage *in trans* with their corresponding Df (*TG4/Df*). If viable, a second Df line covering the same gene was used to validate this finding to make sure the complementation is not due to some problematic Df lines. If TG4 was viable over two independent Df lines, we ascribed the lethality to a second site mutation on the *TG4* chromosome. *TG4* that remained lethal in trans with a Df line are be considered to be disrupting an essential gene in flies. For five genes on the X-chromosome of the fly, complementation was performed by first rescuing hemizygous TG4 males with a duplication (Dp) line obtained from BDSC (Table S4), and crossing these rescued flies to female TG4/FM7 flies. If TG4/Y; Dp/+ lines were viable, we ascribed the lethality of TG4 to the gene of interest. All Df and Dp lines were obtained from BDSC, and the specific stock used in our analysis are listed in Table S4.

Through this experiment, we found 65 *TG4* mutant lines that were homozygous lethal, and 47 of 65 remained lethal when *in trans* with a corresponding deficiency line (Figure S1; Table S4). The 47 essential genes in *D. mel* corresponded to 60 SSC related human genes (Figure S1). The lethality of 18 *TG4* lines corresponding to 19 human genes were due to a second site lethal mutation, potentially present in the original MiMIC line, or introduced during RMCE which has been reported previously (Nagarkar-Jaiswal et al., 2015). These TG4 lines together with viable TG4 lines are likely associated with non-essential genes in *Drosophila*.

Rescue of lethality in TG4 lines by UAS-human cDNA transgenes

In order to assess the ability of human reference or SSC variant cDNAs to rescue lethality observed in *TG4* mutants in essential genes, we first double balanced all Df lines that fail to complement a lethal TG4 line with UAS-reference or variant cDNA lines. For genes on the 2nd chromosome, we generated *Df/CyO; UAS-cDNA/(TM3, Sb, Ser)* stocks. For genes on the 3rd chromosome, we generated *UAS-cDNA/(CyO); Df/TM3, Sb, Ser.* Heterozygous *TG4/Balancer* females were crossed to double balanced Df/Balancer, UAS-human cDNA males at multiple temperatures (18°C, 22°C, 25°C, 29°C) to determine rescue of lethality to adult stage. A minimum of two independent crosses were conducted at each temperature. For the five genes on the X-chromosome of the fly, we attempted

rescue by crossing female *TG4/FM7* flies to UAS-cDNA/(SM6a) males to generate hemizygous TG4 males that expresses human cDNA (TG4/Y; UAS-cDNA/+) to test their viability. Statistical analysis was performed using Two-way ANOVA followed by Sidak's multiple comparison test across temperature and genotype.

Lifespan assays

Lifespan analysis was performed as previously described (Chung et al., 2020). Briefly, newly eclosed flies were separated by genotype and sex and incubated at 25°C. Flies were transferred into a fresh vial every two days and survival was determined once a day. 11-49 flies were tested per group. Statistical analysis was performed using Log-rank (Mantel-Cox) test.

Behavioral assays

Of 61 *TG4* mutants that were viable when *in trans* with a corresponding deficiency, 18 lines exhibited lethality in homozygous states, indicating the presence of a second site lethal mutation. Out of 43 *TG4* mutants that were homozygous viable, we prioritized to study 21 *TG4* mutants based on reagent availability. Courtship assay was performed as previously described (Guo et al., 2019). Due to the reported effects of the $y^1\ w^*$ mutations on behavior (Krstic et al., 2013; Massey et al., 2019), we replaced the X-chromosome containg $y^1\ w^*$ with the $y^+\ w^+$ X-chromosome from a *Canton-S* strain. This *Canton-S* strain was provided by Dr. Hugo Bellen and has been in his stock collection since the 1980s. Please see Figure S3 for the crossing strategy. Collection of socially naive adults was performed by isolating pupae in 16 x 100 polystyrene vials containing approximately 1 mL of fly food. After eclosion, flies were anesthetized briefly with CO_2 to ensure they were healthy and lacking wing damage. Anesthetized flies were returned to their vials and allowed 24 hours to recover before testing. Courtship assays were performed in a 6 well acrylic plate with 40mm circular wells, with a depth of 3 mm and a slope of 11 degrees, as per the chamber design (Simon and Dickinson, 2010). One Canton-S virgin female (6-10 days post-eclosion), and one *TG4* mutant male fly (3-5 days post-eclosion) with or without UAS-human cDNAs were simultaneously introduced into the chamber via aspiration. Recordings were taken using a Basler 1920UM, 1.9MP, 165FPS, USB3 Monochromatic camera using the BASLER Pylon module, with an adjusted capturer rate of 33 fps (frames per second). Conversion of captured images into a movie file was performed via a custom MatLab script, and tracking of flies in the movie was performed using the Caltech Flytracker (Eyjolfsdottir et al., 2014). Machine learning assessment of courtship was performed using JAABA (Kabra et al., 2013) using classifiers that scored at 95% or higher accuracy during ground-truthing trials. At least 10 animals were tested per genotype. Analysis of data was performed using Excel (Microsoft) and Prism (GraphPad). A ROUT (Q=1%) test was performed in Prism to identify outliers. Determination of significance in behavior tests was performed using the Kruskal-Wallis one-way analysis of variance and the Dunn's multiple comparison test. P-values of 0.05 or less were considered significant.

Overexpression assays to assess lethality and morphological phenotypes

To detect any differences in the phenotypes induced by overexpression of reference and variant human cDNA in order to assess variant function, we crossed UAS-human cDNAs with reference or variant alleles to ubiquitous (*tub-GAL4*) (Guelman et al., 2006), wing (*nub-GAL4*) (Calleja et al., 2000) or eye (*GMR-GAL4*) (Mehta et al., 2005) specific drivers. In the ubiquitous expression screen, 3-4 virgin females of *tub-GAL4/TM3 Sb* flies were crossed to 24 males of the UAS-cDNA reference and variant at 25°C. After 3-4 days, the parents were transferred into new vials, and the new vial was placed at 29°C while the old vial was kept at 25°C, allowing us to test two temperatures simultaneously. The parents were discarded after 3-5 days. Flies were collected after most of the pupae eclosed. The total number of flies were counted and scored with the genotype of interest (i.e. *tub-GAL4>UAS-cDNA*) as well as all other genotypes, (i.e. genotypes with balancers). A minimum of 10 flies were scored per experiment, though for the majority of crosses 50-100 flies were scored in this primary analysis. Viability was calculated by taking the % of observed/expected based on Mendelian ratio, and any UAS-cDNA with survival less than 70% was recorded as having scorable phenotype (lethal or semi-lethal). All of lines showing a phenotype at 29°C also showed phenotypes at 25°C, so subsequent experiments were performed at 25°C. To validate our hits, we performed the same viability assay, except each UAS-cDNA was tested at least three times to statistically validate that there is a difference between reference and variant. In addition, two independent UAS-cDNA transgenic lines established from the same construct were tested for each reference and variant. A variant was considered to have functional consequence (true hit) if both transgenic lines showed the same phenotype. In the cases where the difference is rather minor (e.g. <20% difference between survival), this was considered within the variation of the experiment paradigm, and the variant phenotype was documented. Functional study using wing or eye drivers were performed using similar strategies, but morphological phenotypes were scored instead of lethality.

Imaging of adult fly morphology

Drosophila eyes, wings and nota (dorsal thorax) were imaged after flies were frozen at -20°C for at least 24 hours. Wings for some flies were dissected in 70% EtOH and mounted onto slides for imaging. Images were obtained with the Leica MZ16 stereomicroscope equipped with Optronics MicroFire Camera and Image Pro Plus 7.0 software to extend the depth-of-field for Z-stack images.

Expression analysis of TG4 lines in larval and adult brains

All *TG4* lines are crossed with *UAS-nls.mCherry* (3rd chromosome) or *UAS-nls.GFP* (3rd chromosome) at room temperature. Note that the nls.GFP that is being used here is prone to leak outside the nucleus while nls.mCherry is retained in the nucleus. The brains of

mCherry/GFP positive third instar larvae and 3-5 days old adult flies were dissected in 1X phosphate-buffered saline (PBS). Adult brains were fixed immediately in 4% paraformaldehyde (PFA) and incubated at 4°C overnight (o/n) on a shaker. Next day these brains were post-fixed with 4% PFA with 2% Triton-X in PBS (PBST), kept in a vacuum container for an hour to get rid of the air from the tracheal tissue also make the tissue more permissive. Fixative was replaced every 10 minutes during this post-fixation step. Larval brains were fixed for 50 minutes on a rotator at room temperature. After thorough washing with PBS with 0.2% Triton (PBTX) both adult and larval brains were incubated with primary antibodies overnight (o/n) at 4°C on a shaker. The sample were extensively washed with 0.2% PBTX before secondary antibodies were applied at room temperature for 2 hours. Samples were thoroughly washed with PBST and mounted on a glass slide using Vectasheild (Vector Labs, H-1000-10). Primary antibodies used: Mouse anti-repo (DSHB: 8D12) 1:50, Rat anti-elav (DSHB: 7E8A10) 1:100, Goat anti-GFP (abcam: ab6662) 1:500. Secondary antibodies used: Anti-mouse-647 (Jackson ImmunoResearch: 715-605-151) 1:250, Anti-rat-Cy3 (Jackson ImmunoResearch: 712-165-153) 1:500. The samples were scanned using a laser confocal microscope (Zeiss LSM 880) with a 20X objective, and images were processed using ZEN (Zeiss) and Imaris (Oxford Instruments) software. Co-localization between mCherry and Elav or mCherry and Repo was performed with default thresholds in Imaris.

GeneOntology (GO) analysis

GO analysis was performed based on the PANTHER (Protein Analysis Through Evolutionary Relationships) system (http://www.pantherdb.org; date last accessed October 31, 2020 (Ashburner et al., 2000). Statistical analysis was performed by using the default PANTHER Overrepresentation Test (Released 20200728), Annotation Version and Release Date: GO Ontology database https://doi.org/10.1016/10.5281/zenodo.4033054 Released 2020-09-10 which used the Fisher's Exact test with a false discovery rate $p < 0.05$.

Exome sequencing and identification of GLRA2 variants

Subjects 1, 2, 4, 5 and 8 had clinical exome sequencing at GeneDx (Gaithersburg, MD, United States), at the Praxis für Humangenetik Tubingen (Tubingen, Germany), at Baylor Genetics (Houston, TX, United States), at Pitié-Salpêtrière Hospital Genetics lab (Paris, France) and at Integragen (Evry, France), respectively. Subject 3 WES was performed at the Meyer Children's Hospital, University of Florence, in the context of the DESIRE program and as previously described (Vetro et al., 2020). Briefly, the SureSelectXT Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA) was used for library preparation and target enrichment, and paired-end sequencing was performed using Illumina sequencer (NextSeq550, Illumina, San Diego, CA, USA) to obtain an average coverage of above 80x, with 97.6% of target bases covered at least 10x. Reads were aligned to the GRCh37/hg19 human genome reference assembly by the BWA software package, and the GATK suite was used for base quality score recalibration, realignment of insertion/deletions (InDels), and variant calling (DePristo et al., 2011; McKenna et al., 2010). Variant annotation and filtering pipeline included available software (VarSeq, Golden Helix, Inc v1.4.6), focusing on non-synonymous/splice site variants with minor allele frequency (MAF) lower than 0.01 in the GnomAD database (Karczewski et al., 2020) (http://gnomad.broadinstitute.org/), an internal healthy control database and pre-computed genomic variants score from dbNSFP (Liu et al., 2011). Subject 6 WES was performed at the Institute of Human Genetics, Technical University of Munich, Germany as described previously (Brunet et al., 2021). DNA was extracted from blood samples by the Gentra Puregene Blood Kit (Qiagen, Hilden, Germany). WES was performed using the Sure Select Human All Exon 60 Mb V6 Kit (Agilent, Santa Clara, CA, USA) for exomic enrichent according to the manufacturer's protocol. Sequencing was performed on a NovaSeq 6000 sequencer (Illumina, San Diego, CA, USA) to an average coverage of >94X. Reads were aligned to the UCSC human reference assembly (hg19) with the BWA algorithm v.0.5.9. For detection of single-nucleotide variants (SNVs) as well as small insertions and deletions SAMtools v.0.1.19 was applied. Copy number variations (CNVs) were called with the software ExomeDepth. In-house custom Perl scripts were used for variant annotation. Variant analysis was performed using I) a recessive filter for homozygous or compound-heterozygous variants with a minor allele frequency (MAF) of <1% in our in-house database of 22,000 exome datasets, II) a filter for X-chromosomal variants with a MAF<0,1%, III) a filter for de novo variants (MAF <0.01%) and IV) a phenotype-based filter using search terms of characteristic phenotypic traits and a MAF <0,1%. CNVs were assessed using a MAF <0,01%. Current criteria for variant classification according to the American College of Human Genetics (ACMG) were used for variant interpretation (Richards et al., 2015). Subject 7 had exome sequencing at Lyon Universiy Hospital (Lyon, France). The SeqCap EZ Medexome kit (Roche, Pleasanton, CA, USA) was used for library preparation and target enrichment before paired-end sequencing using an Illumina instrument (NextSeq500, Illulina, San Diego, CA, USA). A mean depth of coverage of 133x was obtained with 99.0% of target bases covered at least 10x. Reads were aligned to the GRCh37/hg19 human genome reference assembly by the BWA software package, and the GATK suite was used for base quality score recalibration, realignment of insertion/deletions (InDels), and variant calling (DePristo et al., 2011; McKenna et al., 2010). Variant annotation was performed with SnpEFF and filtering pipeline focused on non-synonymous/splice site variants with minor allele frequency (MAF) lower than 0.01 in the GnomAD database (Karczewski et al., 2020) (http://gnomad.broadinstitute.org/). Subject 9 WES was performed at the Erasmus MC as previously described (Perenthaler et al., 2020). In brief, exome-coding DNA was captured with the Agilent SureSelect Clinical Research Exome (CRE) kit (v2). Sequencing was performed on an Illumina HiSeq 4000 platform with 150-bp paired-end reads. Reads were aligned to hg19 using BWA (BWA-MEM v0.7.13) and variants were called using the GATK Haplotype Caller (McKenna et al., 2010) v3.7 (https://www.broadinstitute.org/gatk/). Detected variants were annotated, filtered and prioritized using the Bench lab NGS v5.0.2 platform (Agilent technologies). Subject 10 WES and data processing were performed by the Genomics Platorm at the Broad Institute of MIT and Harvard with an Illumina Nextera or Twist exome capture (~38 Mb target), and sequenced (150 bp

paired reads) to cover >80% of targets at 20x and a mean target coverage of >100x. WES data was processed through a pipeline based on Picard and mapping done using the BWA aligner to the human genome build 38. Variants were called using Genome Analysis Toolkit (GATK) HaplotypeCaller package version 3.5 (McKenna et al., 2010) (https://www.broadinstitute.org/gatk/). WES for subjects 11 and 12 was performed in collaboration with the Autism Sequencing Consortium (ASC) at the Broad Institute on Illumina HiSeq sequencers using the Illumina Nextera exome capture kit. Exome sequencing data was processed through a pipeline based on Picard and mapping done using the BWA aligner to the human genome build 37 (hg19). Variants were called using Genome Analysis Toolkit (GATK) HaplotypeCaller package version 3.4 (McKenna et al., 2010) (https://www.broadinstitute.org/gatk/). Variant call accuracy was estimated using the GATK Variant Quality Score Recalibration (VQSR) approach. High-quality variants with an effect on the coding sequence or affecting splice site regions were filtered against public databases (dbSNP150 and gnomAD V.2.0) to retain (i) private and clinically associated variants; and (ii) annotated variants with an unknown frequency or having minor allele frequency <0.1%, and occurring with a frequency <2% in an in-house database including frequency data from > 1,500 population-matched WES. The functional impact of variants was analyzed by CADD V.1.3, Mendelian Clinically Applicable Pathogenicity V.1.0 (Jagadeesh et al., 2016; Kircher et al., 2014), and using InterVar V.0.1.6 to obtain clinical interpretation according to American College of Medical Genetics and Genomics/Association for Molecular Pathology 2015 guidelines (Li and Wang, 2017).

Trio exome sequencing of subject 13 and his parents was performed at the Institute of Human Genetics, University of Leipzig Medical Center, in the context of the research project ''Genetics of rare disorders''. Testing in a research setting was approved by the ethics committee of the University of Leipzig (224/16-ek and 402/16-ek). Library preparation was done using the Nextera DNA Flex Pre-Enrichment LibraryPrep with Illumina Nextera DNA UD Indexes by Illumina. Target enrichment was achieved by using the Human Core Exome hybridization probes from Twist Bioscience. Paired-end Next-Generation-Sequencing was then performed on a NovaSeq 6000 Instrument using an S1 Reagent Kit (300 cycles) by Illumina. Analysis of the raw data was performed using the software Varfeed (Limbus, Rostock). Variants were annotated and prioritized using the software Varvis (Limbus, Rostock). Rare variants (minor allele frequency below 1 % in the general population) were prioritized based on inheritance mode, impact on protein, clinical relevance in variant databases and in silico prediction. GeneMatcher (Sobreira et al., 2015) (https://genematcher.org/) assisted in the recruitment of Subjects 2, 3 and 5-13.

SDS-PAGE/Western blot

Five heads of *nSyb-GAL4 UAS-GLRA2* reference and variant flies aged for 5 days post eclosion were lysed in 30μL NETN buffer (50mM Tris pH 7.5, 150mM NaCl, 0.5% NP-40, 1 mM EDTA) with an electric douncer for 10 seconds for three times on ice. 30μL of 2x Laemmli Sample Buffer (Bio-Rad) with 10% 2-mercaptoethanol was added to the lysis and incubated on ice for 10 min. Samples were boiled at 95°C and spun at 14,000 RPM for 5 minutes at 4 °C. The soluble fraction was loaded onto a standard SDS-PAGE gel. PVDF (polyvinylidene difluoride) membrane activated for 1 minute with 100% methanol. After running and wet transfer, the membrane was blocked in 5% skim milk for 1 hour. The membrane was incubated (overnight, shaking, at 4°C) with mouse anti-HA (HA.11, 1:1,000, 901501, BioLegend) and mouse anti-Actin (C4) (1:50,000, MAB1501, EMD Millipore) primary antibodies in 3% BSA (bovine serum albumin), followed by 10 minute washes (3 times) with 1% Triton-X in Tris-buffered saline (TBST). We incubated this with goat anti-mouse HRP-conjugated (1:15000, 115-035-146, Jackson ImmunoResearch) secondary antibody in skim milk. The membrane was washed three times with 1% TBST and detected with Western Lightning™ Chemiluminescence Reagent Plus (perkinelmerNEL104001EA) ECL solution using the Bio-Rad ChemiDoc MP imaging system.

Structural biological analysis of GLRA2 patient variants

Protein residues that corresponds to GLRA2 patient variants were mapped onto the crystal protein structure of GLRA1 protein in Protein Data Bank (PBD, ID: 4X5T) (Burley et al., 2019) using the PyMOL (https://pymol.org/) (Yuan et al., 2016) because GLRA1 and GLRA2 are highly homologous proteins (85% similarity, 78% identity and 3% gaps) based on DIOPT (Hu et al., 2011).

Image generation

Cartoon images in Figure 1H were generated with BioRender.com.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was performed with GraphPad Prism version 8. Significance was defined as *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

Gene and variant level statistics

Gene level constraints for Figures 1B–1D (pLI, LOEUF, missense O/E) for all genes (Global, n=19704), the SSC subset (n=1493), and the study subset (n=143) were based on metric availability from gnomAD and analyzed by ANOVA followed by Dunn's multiple comparison test. The same test was used in Figure S9B–S9Y for gene and variant analysis. The contingency graph for variant consequences (Figure S9A) grouped by variants corresponding to lethal (essential genes) or viable (non-essential genes) TG4 mutants was analyzed by Chi square, df (20, 3, p<0.0002).

Rescue-based and overexpression-based survival

For humanized rescue of TG4 lethality (Figures 2B–2D), 3 independent crosses were set per genotype and a minimum of n>50 flies were quantified for each cross. Statistical analyses were performed by ANOVA followed by Sidak's multiple comparisons test. For lethality caused by overexpression with *Tub-GAL4* (Figures 4B, 5B, and S12B), a minimum of 3 independent crosses were set with two independent UAS-transgenic lines. 50–100 flies (a minimum of 10 if overexpression caused survival defects) were scored. Statistical analyses were performed by unpaired t test.

Lifespan analysis

Lifespan analysis of humanized *TG4* lines at 25°C (Figures 2E and 2F) were analyzed by Log rank (Mantel-Cox) test with a minimum of 11-49 flies for each genotype from three independent crosses.

Behavior

Fore courtship analysis (Figures 3C–3G and S4A–S4T), at least 10 animals were tested per genotype. Analysis of data was performed using Excel (Microsoft) and Prism (GraphPad). A ROUT (Q=1%) test was performed in Prism to identify outliers. Data was analyzed with Kruskal-Wallis ANOVA followed by Dunn's multiple comparison test.

Assessment of melanized nodules

Melanized nodules formed on the notum of flies expressing *GLRA2*^{T296M} driven by *pnr-GAL4* (Figure 5D) were analyzed by ANOVA followed by Tukey's multiple comparison test. 3 independent crosses were set and minimum of 50 flies examined for presence of melanized nodules per genotype.

Electroretinograms

Electrophysiological field potential recordings (ERG) of *nSyb-GAL4* or *Rh1-GAL4* flies expressing *UAS-GLRA2* constructs (Figures 5E–5H and S12C–S12F) were analyzed by ANOVA followed by Bonferroni's multiple comparison test. 4-10 flies were examined for each genotype. Recording was repeated at least 3 times per fly.