PNAS

Supporting Information for

Using single-cell RNA sequencing to generate predictive cell-type-specific split-GAL4 reagents throughout development

Yu-Chieh David Chen, Yen-Chung Chen, Raghuvanshi Rajesh, Nathalie Shoji, Maisha Jacy, Haluk Lacin, Ted Erclik, Claude Desplan

Corresponding authors: Dr. Claude Desplan or Dr. Yu-Chieh David Chen
Email: cd38@nyu.edu or ycc4@nyu.edu

This PDF file includes:

Supporting text
Figures S1 to S6
Tables S1 to S2
Legend for Dataset S1
SI References

Other supporting materials for this manuscript include the following:

Dataset S1

Supporting Information Text

MATERIALS AND METHODS

Fly strains.

Flies were reared on molasses-cornmeal-agar food at 25°C. The fly lines used in the paper are described in **SI Appendix, Table S1**. CRISPR-mediated T2A-split-GAL4 knock-in was performed by WellGenetics Inc. (Taipei, Taiwan). In brief, the gRNA sequences (specified in **SI Appendix, Table S2**) were cloned into U6 promoter plasmid(s). Cassette T2A-GAL4DBD (or AD)-GMR-RFP, which contains T2A, Zip-GAL4DBD (or AD), SV40 3'UTR, a floxed GMR-RFP, and two homology arms were cloned into pUC57-Kan as donor template for repair. Targeting gRNAs and hs-Cas9 were supplied in DNA plasmids, together with donor plasmid for microinjection into embryos of control strain w^{1118}. F1 flies carrying selection marker of GMR-RFP were further validated by genomic PCR and sequencing.

Split-GAL4 lines generated by injecting donor DNA plasmids into F1 embryos from crosses of coding intronic MiMIC and CRIMIC lines with ΦC31 integrase source. Embryo injections for transgenesis and transformant recovery was completed by BestGene (Chino Hills, CA). The split-GAL4 donor plasmids with appropriate splicing phases were obtained from Addgene, including *pBS-KS-attB2-SA(0)-T2A-GAL4DBD-Hsp70* (Addgene #62902), *pBS-KS-attB2-SA(1)-T2A-GAL4DBD-Hsp70* (Addgene #62903), *pBS-KS-attB2-SA(2)-T2A-GAL4DBD-Hsp70* (Addgene #62904), *pBS-KS-attB2-SA(0)-T2A-AD-Hsp70* (Addgene #62905), *pBS-KS-attB2-SA(1)-T2A-AD-Hsp70* (Addgene #62908), and *pBS-KS-attB2-SA(2)-T2A-AD-Hsp70* (Addgene #62915). The integration orientation of $yellow^-$ progenies from MiMIC injection and $3xP3\text{-}GFP^-$ progenies from CRIMIC injection were confirmed by PCR genotyping. See also **SI Appendix, Table S2** for further information.

Triple and single split-GAL4 donor transgenic lines.

For generating triple donor GAL4DBD plasmid pC(lox2-attB2-SA-T2A-GAL4DBD-Hsp70)3, we synthesized SphI_T2A-GAL4DBD_NotI, MluI-T2A-GAL4DBD_FesI, and BsiWI-T2A-GAL4DBD_AscI by GenScript (Piscataway, USA). The three fragments were subcloned into pC-(lox2-attB2-SA-T2A-GAL4-Hsp70)3 (Addgene #62957) by replacing T2A-GAL4 with T2A-GAL4DBD. For generating triple donor AD plasmid pC(lox2-attB2-SA-T2A-AD-Hsp70)3, we synthesized SphI_T2A-AD_NotI, MluI-T2A-VP16_FesI, and BsiWI-T2A-AD_AscI by GenScript (Piscataway, USA). The three fragments were subcloned into pC-(lox2-attB2-SA-T2A-GAL4-Hsp70)3 (Addgene #62957) by replacing T2A-GAL4 with T2A-AD. Both triple donor plasmids were sent for standard P-element-mediated transformation performed by WellGenetics Inc. (Taipei, Taiwan).

For single donor split-GAL4 plasmids, we synthesized BamHI_T2A-GAL4DBD_BamHI and BamHI_T2A-AD_BamHI by GenScript (Piscataway, USA). The fragments were subcloned into either pC-(loxP2-attB2-SA(0)-T2A-GAL4-Hsp70), Addgene #62954, pC-(loxP2-attB2-SA(1)-T2A-GAL4-Hsp70), Addgene #62955, or pC-(loxP2-attB2-SA(2)-T2A-GAL4-Hsp70); Addgene #62956 by replacing T2A-GAL4 with T2A-GAL4DBD or T2A-AD. The single donor plasmids were sent for standard P-element-mediated transformation performed by BestGene (Chino Hills, CA).

Immunohistochemistry.

Flies were anesthetized on ice, and the optic lobes were dissected in ice-cold Schneider's *Drosophila* Medium (Thermo Fisher, #21720024) for less than 30 min. Tissues were fixed for 20 min with 4% paraformaldehyde in 1X Dulbecco's phosphate-buffered saline (DPBS, Corning, #21031CV) at room temperature. After three washes with 1X PBST (DPBS with 0.3% Triton X-100), tissues were incubated in primary antibody solutions for 2 days at 4°C. Samples were washed with 1X PBST for at least 6 times (15 min per wash) followed by incubating in secondary antibodies 1-2 days at 4°C. Samples were washed again with 1X PBST for at least 6 times (15 min per wash) followed by the last wash in 1X DPBS. Both primary and secondary antibodies were prepared in 1X PBST with 5% goat serum (Thermo Fisher, #16210064). The antibodies used in the paper are described in **SI Appendix, Table S1**. Samples were mounted in

VECTASHIELD antifade mounting medium (Vector Laboratories, #H-1000) and stored at 4°C. Fluorescent images were acquired using a Leica SP8 confocal microscope with 400 Hz scan speed in 1024x1024 pixel formats. Image stacks were acquired at 0.5-1 μm optical sections. Unless otherwise noted, all images were presented as maximum projections of the *z* stack generated using Leica LAS AF software.

Identification of marker gene pairs with mixture modeling-inferred binarized expression. To identify marker gene pairs (two genes) that are specific to a cluster, we implemented a greedy search algorithm to minimize the number of clusters that express a given gene pair. Briefly, for each cluster, we begin with a gene that (1) is expressed in the cluster of interest (target cluster) and (2) is expressed in fewest other clusters (off-target cluster). The same steps are repeated once only among the clusters positive for the first gene selected to identify split-GAL4 candidates but can potentially be extended to identify combinations consisting of more than two genes.

To determine whether a gene is expressed in a cluster, we assign probability of whether a gene is expressed ($P(ON)$) to each cluster at each stage as previously described (1, 2). Briefly, we define (1) a baseline unimodal model that cluster average expression of a given gene follows a Gaussian distribution and (2) an alternative bimodal model that cluster average expression follows a mixture of two Gaussian distributions, representing ON and OFF respectively. Parameters of the two models were estimated, and the wellness of fit were compared with expected log predictive density (elpd, a measure of how well a model explains observed data) in 8-fold cross-validation in Stan (Stan Development Team. 2023. Stan Modeling Language Users Guide and Reference Manual, v2.29.2. https://mc-stan.org), a software package that implements Bayesian probabilistic model fitting with Markov chain Monte Carlo algorithm. Genes that fit better with the bimodal model ($\Delta\widehat{elpd} > 2\ SE_{\widehat{elpd}}$) were considered as bimodally distributed while other genes were considered as unimodal. Probability of whether a gene is expressed [*P(ON)*] in a cluster were then estimated with the models fitted above with R 4.0.4 (R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.) and mclust (Scrucca L, Fop M, Murphy TB, Raftery AE (2016). "mclust 5: clustering, classification and density estimation using Gaussian finite mixture models." The R Journal, 8(1), 289–317. https://doi.org/10.32614/RJ-2016-021.)

Gene × cluster probability matrices were generated for datasets corresponding to each stage, and we consider a gene with *P(ON)* > 0.5 as expressed. This process is equivalent to adaptively determining gene-specific binarization thresholds with Gaussian mixture models. Binary expression states not only allow efficient greedy combination search described above, but also enabled versatile integration across time series via Boolean algebra. Reduced computation resource requirement with binary data also allowed comprehensive marker combination search.

scMarco

scMarco (source code: https://github.com/chenyenchung/scMarco; an example site with optic lobe developmental atlas (2): https://apps.ycdavidchen.com/scMarco; documentations: https://docs.ycdavidchen.com/scMarco) is a Shiny application that provides interactive selection of **mar**ker **co**mbinations with modeled expression probability as an SQLite database. scMarko workflow requires the users to define (1) a probability cut-off to binarize expression status to ON and OFF for each stage/condition and (2) a criterion to define the expressed status in the whole dataset (e.g., A gene must be expressed in more than 3 stages including adult to be considered as expressed). scMarko uses the two rules to scrutinize each gene and generate a ON/OFF matrix for each gene. Selection of marker combinations starts from either a cluster of interest or a gene of interest.

In a cluster-centric workflow, scMarco selects and ranks specific marker combos by maximizing conditional mutual information (3). Alternatively, the user can manually select from all genes that are expressed in the cluster of interest. In this workflow, a list of other clusters (off-targets) that also express the selected genes will be generated to find second markers that is specific to each cluster if possible. In a gene-centric workflow, scMarco uses the gene as the first marker, and

lists all the clusters that express this gene, and find second markers that is specific to each of these clusters if possible.

scMarco provides visualization of expression status any gene combinations as a barcode plot to assess the quality of predicted marker combinations. To address potential information loss during binarization of continuous gene expression data, normalized cluster average expression is visualized as line plots for all clusters across all stages/conditions with the clusters of interest highlighted so the users can assess whether a selected gene is apparently bimodal.

To aid selection of convenient markers with available resources or genetic tools, scMarco takes gene lists so marker combination discovery can be done only for genes that has transgenetic reporters available, antibodies or hybridization probes available, or specific gene families that are of particular interest (e.g., transcription factors, EGFR downstream genes, and etc.).

Supporting Figures

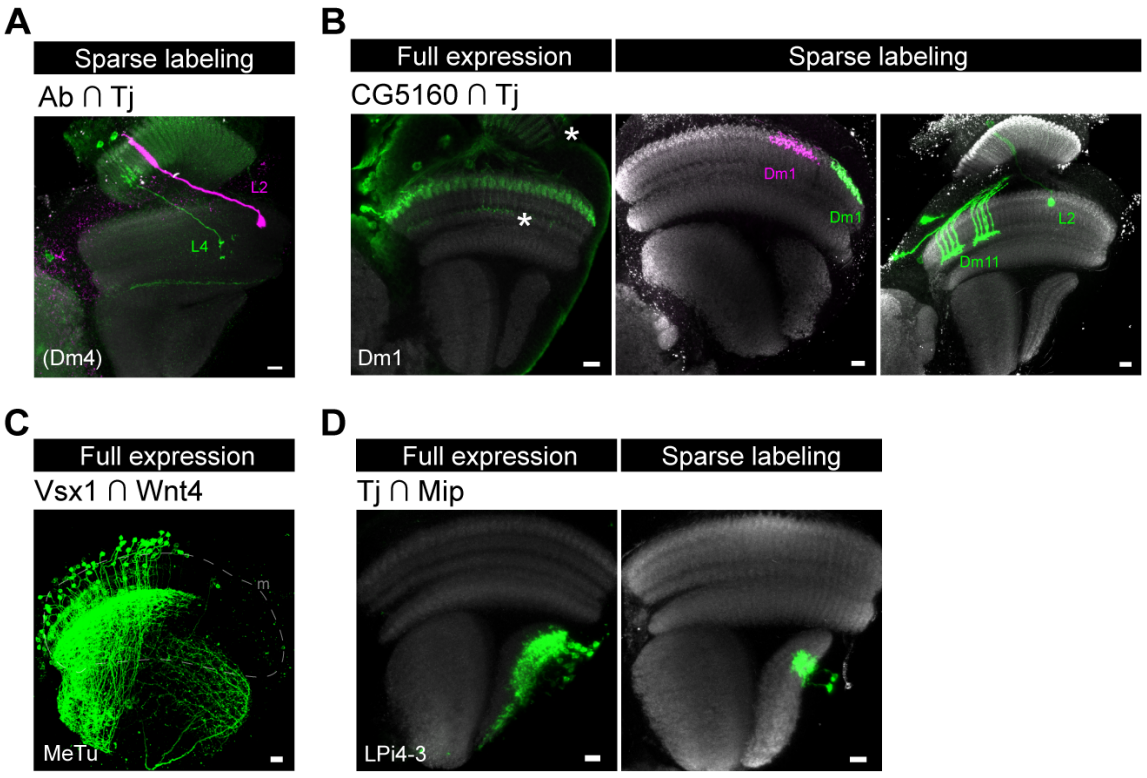

Fig. S1. Characterization of selected gene-specific split-GAL4 lines targeting different cell types/clusters. Targeted cell types predicted by scRNAseq expression are shown in the lower left corner for each split-GAL4 line. The expression pattern of each split-GAL4 line is shown either with UAS-myr-GFP reporter for full expression (right) or with UAS-MCFO lines for sparse labeling (right). (**A**) Sparse labeling of the split-GAL4 line targeting Dm4 showed additional L2 and L4 neuronal labeling. (**B**) A different split-GAL4 combination targeting Dm1. We also observed Dm11 and L2 neurons in this line. (**C**) The cell bodies of MeTu neurons are located in the dorsal half of medulla cortex in adult. (**D**) A different split-GAL4 combination targeting LPi4-3. Anti-NCad staining (gray) is used for visualizing neuropils. Images are substack projections from full expression labeling or segmented single cells from sparse labeling to show distinct morphological features of distinct cell types. Asterisks indicate expression in the cell types not predicted by mixture modeling. Scale bar: 10 μm.

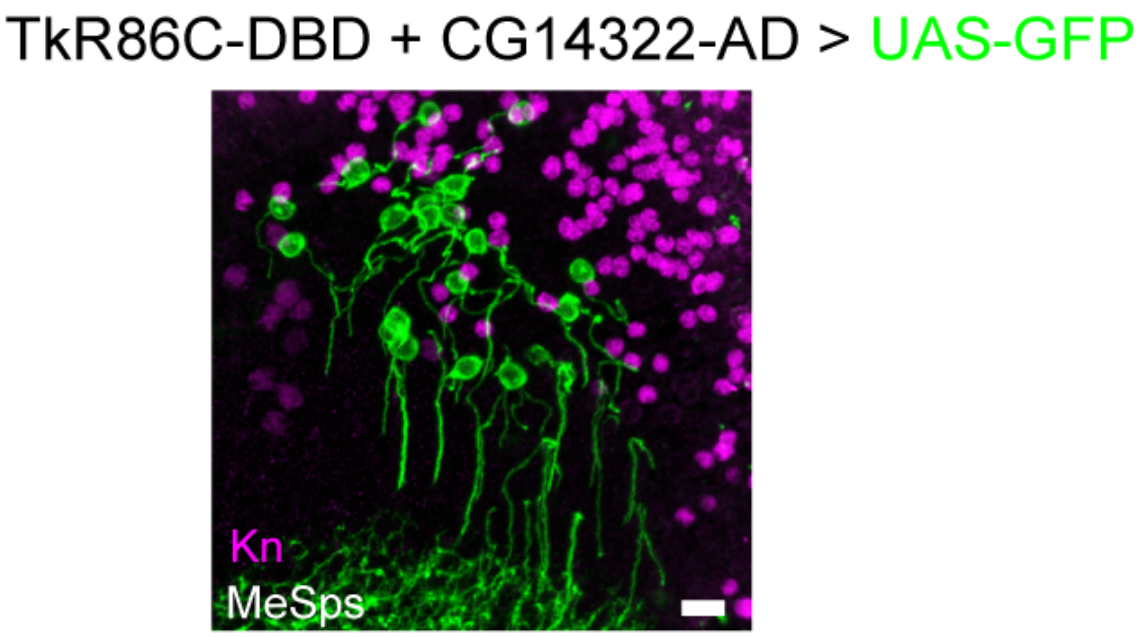

Fig. S2. Co-staining of MeSps neurons labeled by TkR86C ∩ CG14322 with anti-Kn (Magenta). MeSps neurons expressing GFP reporter are Kn negative. Scale bar: 10 μm

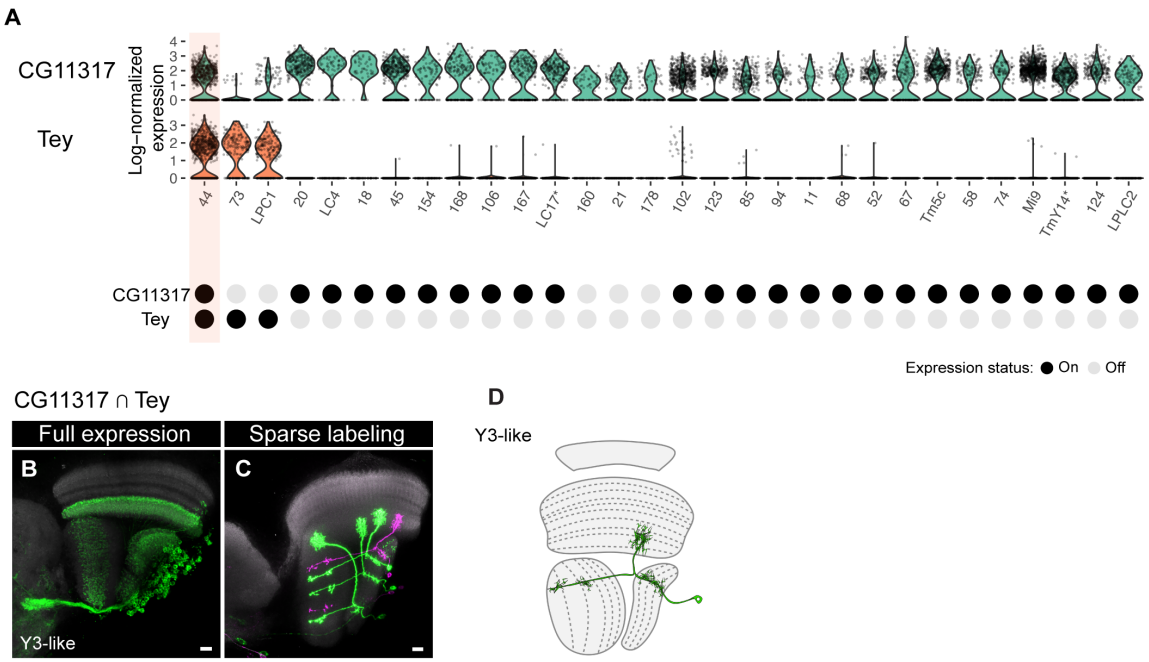

Fig. S3. (**A**) Log-normalized expression of CG11317 and Tey for different clutters (Top). Mixture modeling binarization of expression status for both genes is shown at the bottom. Note that cluster 44 is predicted to be the only cluster intersected by CG11317 and Tey. Top 30 clusters expressing CG11317 or Tey are shown. (**B**) The full expression pattern of CG11317 ∩ Tey line is shown with UAS-myr-GFP reporter. (**C**) Sparse labeling of Y3-like neurons using MCFO. (**D**) Schematic diagram of Y3-like neurons. Scale bar: 10 μm.

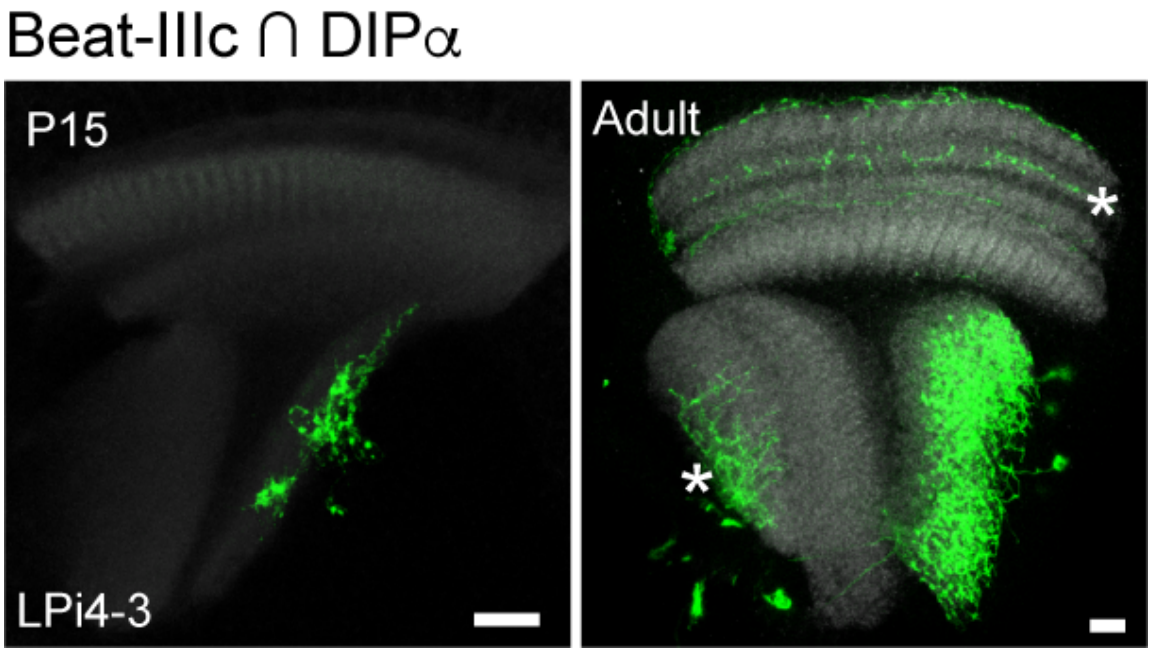

Fig. S4. The full expression pattern of Beat-IIIc ∩ DIPα line is shown with UAS-myr-GFP reporter at P15 (left) and adult (right) stages. Note that the targeted cell type (LPi4-3) is always observed at multiple developmental stages, although other cell types might be observed in addition to it (marked by asterisks). Anti-NCad staining (gray) is used for visualizing neuropils. Images are substack projections from full expression labeling to show distinct morphological features of distinct cell types. Scale bar: 10 μm.

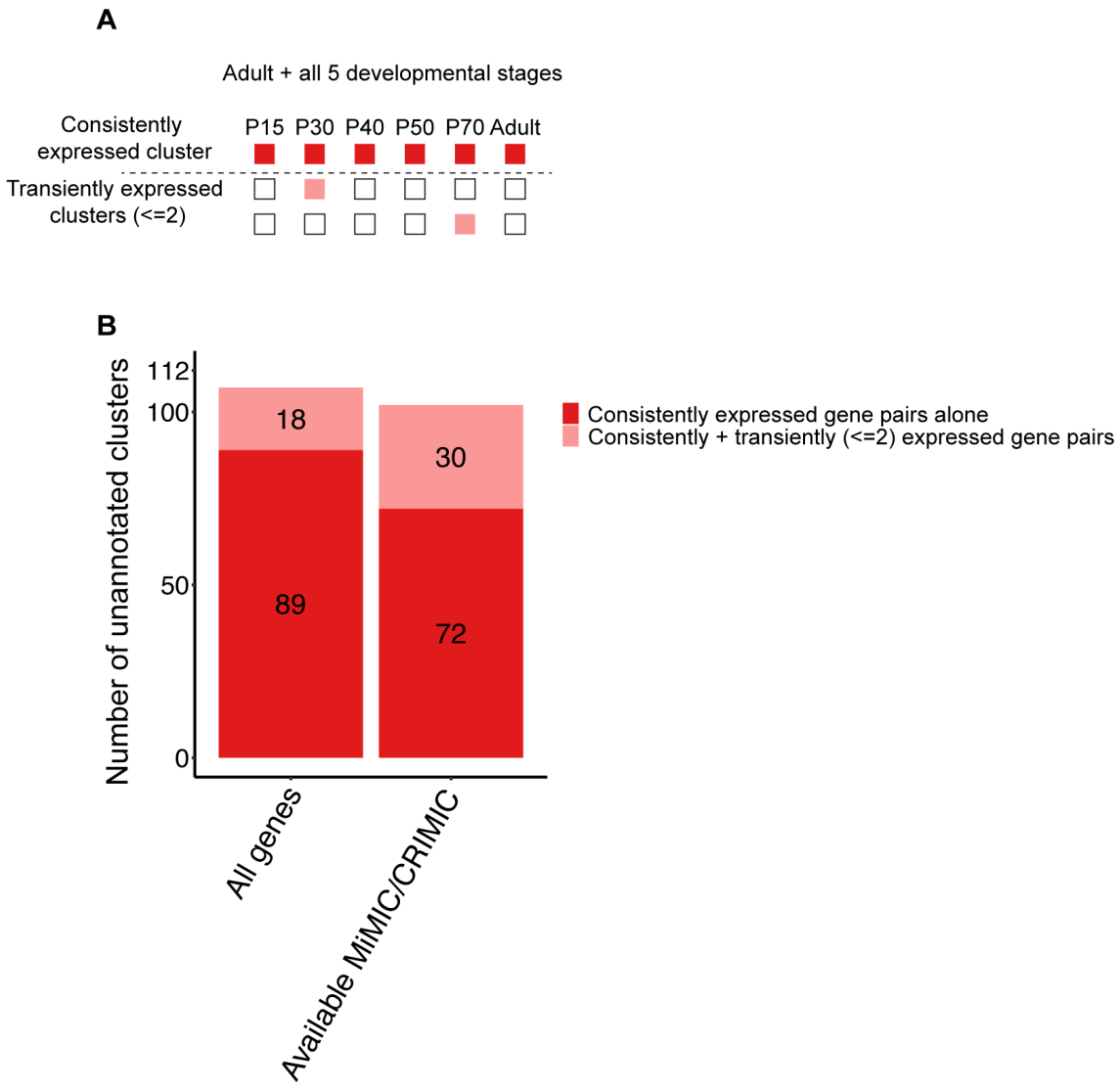

Fig. S5. (**A**) Illustration of the strategy used to identify gene pairs that mark a previously unannotated cluster of interest. Red boxes indicate the stages when a gene pair is predicted to be on in the cluster of interest while pink boxes indicate when the gene pair is predicted to be transiently active in other clusters. (**B**) Number of previously unannotated clusters that are predicted to be identified with gene pairs when all genes detected in the atlas are considered (left) or when only genes with coding intronic MiMIC or CRIMIC lines available are considered (right).

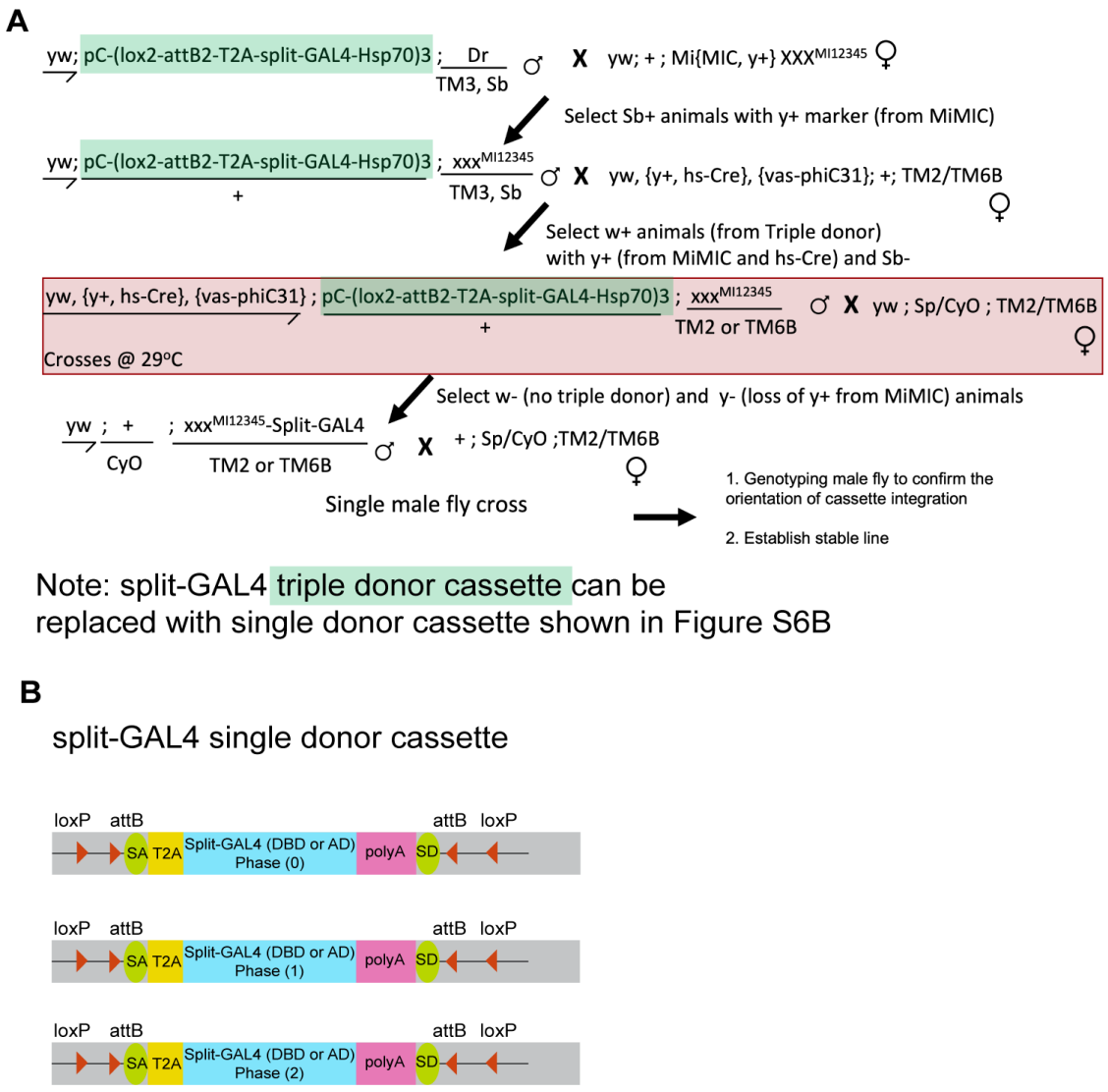

Fig. S6. (A) *In vivo* genetic crossing scheme for swapping T2A-split-GAL4 in coding intronic MiMIC lines. Flies with hs-Cre, vas-FC31, split-GAL4 donor and targeting MiMIC cross should be incubated at 29°C to induce the expression of Cre recombination. (**B**) Schematics of split-GAL4 single donor cassettes.

Table S1. Fly strains and antibodies

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fly strains		
Drosophila, yw;;	Desplan lab	N/A
Drosophila, 10XUAS-myr-GFP	BDSC	BDSC #32197, #32198
Drosophila, R57C10-Flp2::PEST; 10xUAS(FRT.stop)myr::smGdP-V5-THS-10xUAS(FRT.stop)myr::smGdP-FLAG} ; GMR-myr::RFP, Rh4-lacZ	Desplan lab and BDSC	BDSC #62124
Drosophila, 5-HT1B[MI05213]-T2A-GAL4DBD	This study	N/A
Drosophila, 5-HT2B[MI06500]-T2A-GAL4DBD	This study	N/A
Drosophila, ab-T2A-GAL4DBD	This study	N/A
Drosophila, AdoR[MI01202]-T2A-GAL4DBD	This study	N/A
Drosophila, ara-T2A-GAL4DBD	This study	N/A
Drosophila, Acj6[MI07818]-T2A-GAL4DBD	This study	N/A
Drosophila, BBS4[MI15481]-T2A-GAL4DBD	This study	N/A
Drosophila, Beat-IIIc[MI03726]-T2A-GAL4DBD	This study	N/A
Drosophila, bi-T2A-GAL4DBD	This study	N/A
Drosophila, bt[MI06578]-T2A-GAL4DBD	This study	N/A
Drosophila, CG11317[MI04510]-T2A-GAL4DBD	This study	N/A
Drosophila, CG11537[MI04840]-T2A-GAL4DBD	This study	N/A
Drosophila, CG13375[MI15214]-T2A-GAL4DBD	This study	N/A
Drosophila, CG13408[CR02202]-T2A-GAL4DBD	This study	N/A
Drosophila, CG14322[MI11688]-T2A-GAL4DBD	This study	N/A
Drosophila, CG14521(DIPγ)[MI03222]-T2A-GAL4DBD	(4)	N/A
Drosophila, CG15353-T2A-GAL4DBD	This study	N/A
Drosophila, CG1688[CR01570]-T2A-GAL4DBD	This study	N/A
Drosophila, CG2772[CR02306]-T2A-GAL4DBD	This study	N/A
Drosophila, CG31183[MI02001]-T2A-GAL4DBD	This study	N/A
Drosophila, CG31689[MI03245]-T2A-GAL4DBD	This study	N/A
Drosophila, CG32432[MI03670]-T2A-GAL4DBD	This study	N/A
Drosophila, CG34411[MI06794]-T2A-GAL4DBD	This study	N/A
Drosophila, CG42368[MI09901]-T2A-GAL4DBD	This study	N/A
Drosophila, CG43737[MI10158]-T2A-GAL4DBD	This study	N/A
Drosophila, CG44325[MI09901]-T2A-GAL4DBD	This study	N/A

Drosophila, CG5160-T2A-GAL4DBD	This study	N/A
Drosophila, CG5910[MI09800]-T2A-GAL4DBD	This study	N/A
Drosophila, CG7191[MI03423]-T2A-GAL4DBD	This study	N/A
Drosophila, CG9896-T2A-GAL4DBD	This study	N/A
Drosophila, CngA[MI05524]-T2A-GAL4DBD	This study	N/A
Drosophila, CNMaR[MI01128]-T2A-GAL4DB	This study	N/A
Drosophila, Fer2[MI09483]-T2A-GAL4DBD	This study	N/A
Drosophila, foxo[MI00493]-T2A-GAL4DBD	This study	N/A
Drosophila, glob1[MI14203]-T2A-GAL4DBD	This study	N/A
Drosophila, heph[MI04030]-T2A-GAL4DBD	This study	N/A
Drosophila, hig[MI05774]-T2A-GAL4DBD	This study	N/A
Drosophila, Kek3[CR00930]-T2A-GAL4DBD	This study	N/A
Drosophila, Lkr[MI08640]-T2A-GAL4DBD	This study	N/A
Drosophila, Meltrin[MI09878]-T2A-GAL4DBD	This study	N/A
Drosophila, mld[MI07175]-T2A-GAL4DBD	This study	N/A
Drosophila, ort-T2A-GAL4DBD	This study	N/A
Drosophila, PHDP[CR01516]-T2A-GAL4DBD	This study	N/A
Drosophila, Pka-R1[MI06935]-T2A-GAL4DBD	This study	N/A
Drosophila, rdo[MI14128]-T2A-GAL4DBD	This study	N/A
Drosophila, Reck-T2A-GAL4DBD	This study	N/A
Drosophila, rst[MI04842]-T2A-GAL4DBD	This study	N/A
Drosophila, RunxB[MI10372]-T2A-GAL4DBD	This study	N/A
Drosophila, Rya-R[MI01843]-T2A-GAL4DBD	This study	N/A
Drosophila, salm[MI13018]-T2A-GAL4DBD	This study	N/A
Drosophila, Tj-T2A-GAL4DBD	This study	N/A
Drosophila, TkR86C[MI05788]-T2A-GAL4DBD	This study	N/A
Drosophila, TrissinR[MI05488]-T2A-GAL4DBD	This study	N/A
Drosophila, tsh-T2A-GAL4DBD	This study	N/A
Drosophila, Tusp[MI04698] -T2A-GAL4DBD	This study	N/A
Drosophila, Vsx1-T2A-GAL4DBD	This study	N/A
Drosophila, Wnt10-T2A-GAL4DBD	(5)	N/A
Drosophila, Beat-IIIc[MI03726]-T2A-VP16	This study	N/A
Drosophila, bru1[MI00135]-T2A-VP16	This study	N/A
Drosophila, CCAP-R[MI05804]-T2A-VP16	This study	N/A
Drosophila, CG14010 (DIPη)[MI07948]-T2A-VP16	This study	N/A
Drosophila, CG14322[MI11688]-T2A-VP16	This study	N/A

Drosophila, CG14431[MI10741]-T2A-VP16	This study	N/A
Drosophila, CG14521(DIPγ)[MI03222]-T2A-VP16	(4)	N/A
Drosophila, CG14947-T2A-VP16	This study	N/A
Drosophila, CG2016[MI04995]-T2A-p65	This study	N/A
Drosophila, CG31191[MI07868]-T2A-VP16	This study	N/A
Drosophila, CG31689[MI03245]-T2A-VP16	This study	N/A
Drosophila, CG32105-T2A-VP16	This study	N/A
Drosophila, CG32791 (DIPα)[MI02031]-T2A-p65	This study	N/A
Drosophila, CG4238[MI13092]-T2A-VP16	This study	N/A
Drosophila, CG43737[MI10158]-T2A-VP16	This study	N/A
Drosophila, CG5397[MI15629]-T2A-VP16	This study	N/A
Drosophila, CG9109[MI01660]-T2A-VP16	This study	N/A
Drosophila, CG9743[MI09362]-T2A-VP16	This study	N/A
Drosophila, dac-T2A-GAL4AD	This study	N/A
Drosophila, Dop1R2[CR02529]-T2A-VP16	This study	N/A
Drosophila, erm-T2A-VP16	This study	N/A
Drosophila, eya-T2A-VP16	This study	N/A
Drosophila, Fife[MI08414]-T2A-VP16	This study	N/A
Drosophila, fs[MI04308]-T2A-VP16	This study	N/A
Drosophila, Gad1[MI09277]-T2A-VP16	(6)	BDSC #60322
Drosophila, Kn[MI15480]-T2A-p65	(7)	N/A
Drosophila, Lim1-T2A-VP16	This study	N/A
Drosophila, mbl[MI00139]-T2A-VP16	This study	N/A
Drosophila, Mip-T2A-VP16	This study	N/A
Drosophila, mirr-T2A-VP16	This study	N/A
Drosophila, MsR2[CR01840]-T2A-p65	This study	N/A
Drosophila, NetA[MI04563]-T2A-VP16	This study	N/A
Drosophila, orb[MI04761]-T2A-VP16	This study	N/A
Drosophila, ort-T2A-VP16	This study	N/A
Drosophila, Ple-T2A-p65	(8)	BDSC #84712
Drosophila, rn[MI09946]-T2A-VP16	This study	N/A
Drosophila, run-T2A-VP16	This study	N/A
Drosophila, Rx[CR00377]-T2A-p65	This study	N/A
Drosophila, sfl[MI02467]-T2A-p65	This study	N/A
Drosophila, slou-T2A-VP16	This study	N/A
Drosophila, SoxN-T2A-VP16	This study	N/A

Drosophila, svp[MI01102]-T2A-VP16	This study	N/A
Drosophila, tey-T2A-VP16	This study	N/A
Drosophila, tj-T2A-VP16	This study	N/A
Drosophila, tsh-T2A-VP16	This study	N/A
Drosophila, VGlut[MI04979]-T2A-p65	(9)	BDSC #82986
Drosophila, Wnt10-T2A-p65	(5)	N/A
Drosophila, P{lox(Trojan-GAL4DBD)x3}	This study	N/A
Drosophila, P{lox(Trojan-VP16)x3}	This study	N/A
Drosophila, P{loxP(Trojan-GAL4DBD.0)}	This study	N/A
Drosophila, P{loxP(Trojan-GAL4DBD.1)}	This study	N/A
Drosophila, P{loxP(Trojan-GAL4DBD.2)}	This study	N/A
Drosophila, P{loxP(Trojan-VP16.0)}	This study	N/A
Drosophila, P{loxP(Trojan- VP16.1)}	This study	N/A
Drosophila, P{loxP(Trojan- VP16.2)}	This study	N/A
Antibodies		
Sheep anti-GFP (1:200)	BioRad	4745-1051 (RRID:AB_619712)
Chicken anti-V5 (1:5000)	Abcam	ab9113 (RRID:AB_307022)
Rat anti-FLAG (1:200)	Novus Biologicals	NBP1-06712 (RRID:AB_1625981)
Rat anti-NCad (1:20)	DSHB	DN-ex#8 (RRID:AB_528121)
Mouse anti-Chaoptin (1:20)	DSHB	24B10 (RRID:AB_528161)
Mouse anti-Brp (1:20)	DSHB	nc82 (RRID:AB_2314866)
Mouse anti-pros (1:33)	DSHB	MR1A (RRID:AB_528440)
Guinea pig anti-Tj (1:250)	(10), Gift from Dorothea Godt	N/A
Rabbit anti-Toy (1:500)	(2), Desplan lab	N/A
Guinea pig anti-Fd59A (1:400)	(11), Gift from James Skeath	N/A
Guinea pig anti-Kn (1:200)	(12), Desplan lab	N/A
Donkey anti-sheep Alexa Fluor 488 (1:500)	Jackson ImmunoResearch	713-545-147 (RRID:AB_2340745)

Donkey anti-chicken Alexa Fluor 488 (1:500)	Jackson ImmunoResearch	703-545-155 (RRID:AB_2340375)
Donkey anti-rat Cy3 (1:500)	Jackson ImmunoResearch	712-165-153 (RRID:AB_2340667)
Donkey anti-mouse Alexa Fluor 555 (1:500)	Invitrogen	A31570 (RRID:AB_2536180)
Donkey anti-rabbit Alexa Fluor 555 (1:500)	Invitrogen	A31572 (RRID:AB_162543)
Donkey anti-rat Alexa Fluor 647 (1:200)	Jackson ImmunoResearch	712-605-153 (RRID:AB_2340694)
Donkey anti-guinea pig Alexa Fluor 647 (1:200)	Jackson ImmunoResearch	706-605-148 (RRID:AB_2340476)

Table S2. Gene-specific split-GAL4 informaion

DBD	BDSC #	MI # or CR #	SA (?) DBD	Addgene #	CRISPR knock in	gRNA sequence	Reference
5-HT1B	41063	MI05213	SA (0) DBD	62902			
5-HT2B	60251	MI06500	SA (1) DBD	62903			
ab			GAL4DBD		C-terminal knock in	GGGAGTGCACAACACATAGG[AGG]	
Acj6	51212	MI07818	SA (0) DBD	62902			
AdoR	35940	MI01202	SA (1) DBD	62903			
ara			GAL4DBD		C-terminal knock in	CTTGAAGAGCGGCTTGAACA[TGG]	
BBS4	61065	MI15481	SA (2) DBD	62904			
Beat-IIIc	41041	MI03726	SA (1) DBD	62903			
bi			GAL4DBD		C-terminal knock in	GGTGCGAGTGTAGATGGGTC[TGG]	
bt	60797	MI06578	SA (1) DBD	62903			
CG11317	38565	MI04510	SA (1) DBD	62903			
CG11537	38057	MI04840	SA (0) DBD	62902			
CG13375	60995	MI15214	SA (0) DBD	62902			
CG13408	91471	CR02202	SA (2) DBD	62904			
CG14322	56686	MI11688	SA (0) DBD	62902			
CG14521	35928	MI03222	VP16				(4)
CG15353			GAL4DBD		C-terminal knock in	TGTCTACTCCGCCGGCTATC[CGG]	
CG1688	91261	CR01570	SA (1) DBD	62903			
CG2772	91508	CR02306	SA (0) DBD	62902			
CG31183	33160	MI02001	SA (1) DBD	62903			
CG31689	44287	MI03245	SA (1) DBD	62903			
CG32432	37070	MI03670	SA (0) DBD	62902			
CG34411	42179	MI06794	SA (1) DBD	62903			
CG42368	56383	MI11827	SA (1) DBD	62903			
CG43737	53823	MI10158	SA (1) DBD	62903			
CG44325	53795	MI09901	SA (0) DBD	62902			
CG5160			GAL4DBD		C-terminal knock in	ACGGATGAGCATTTCCACCA[GGG]	
CG5910	53784	MI09800	SA (1) DBD	62903			
CG7191	41388	MI03423	SA (2) DBD	62904			
CG9896			GAL4DBD		N-terminal knock in	CTGGCCGCATTCCTCCGAAT[TGG]	
CngA	42329	MI05524	SA (0) DBD	62902			
CNMaR	34150	MI01128	SA (0) DBD	62902			
Fer2	53114	MI09483	SA (2) DBD	62904			
foxo	31027	MI00493	SA (0) DBD	62902			
glob1	59679	MI14203	SA (0) DBD	62902			
heph	38561	MI04030	SA (0) DBD	62902			
hig	42114	MI05774	SA (1) DBD	62903			
Kek3	79364	CR00930	SA (2) DBD	62904			
Lkr	51094	MI08640	SA (1) DBD	62903			
Meltrin	53790	MI09878	SA (0) DBD	62902			
mld	44324	MI07175	SA (0) DBD	62902			
ort			GAL4DBD		N-terminal knock in	TGCATGAAGCACTACGCCAA[AGG]	
PHDP	86430	CR01516	SA (0) DBD	62902			
Pka-R1	43600	MI06935	SA (0) DBD	62902			

rdo	59263	MI14128	SA (1) DBD	62903			
Reck			GAL4DBD		N-terminal knock in	AAAAATGCGCCTGAGTGGCT[GGG]	
rst	38589	MI04842	SA (1) DBD	62903			
RunxB	55474	MI10372	SA (1) DBD	62903			
Rya-R	35074	MI01843	SA (0) DBD	62902			
salm	58017	MI13018	SA (2) DBD	62904			
tj			GAL4DBD		N-terminal knock in	GATCTGCACCACCTGGAGGT[GGG]	
TkR86C	42116	MI05788	SA (2) DBD	62904			
TrissinR	60783	MI05488	SA (2) DBD	62904			
tsh			GAL4DBD		C-terminal knock in	CTTCACGCAGAATTGTCGCA[AGG]	
Tusp	37835	MI04698	SA (0) DBD	62902			
Vsx1			GAL4DBD		C-terminal knock in	GAAGCAGGAGCAGGCGCATC[TGG]	
						/ ATCGATGTACATGTTTCATC[TGG]	
wnt10			GAL4DBD				(5)
AD	**BDSC #**	**MI # or CR #**	**SA (?) AD**	**Addgene #**	**CRISPR knock in**	**gRNA sequence**	**Refernece**
Beat-IIIc	41041	MI03726	SA (1) VP16	62905			
bru1	30631	MI00135	SA (1) VP16				
CCAP-R	40788	MI05804	SA (1) VP16	62905			
CG14010	44718	MI07948	SA (1) VP16	62905			
CG14322	56686	MI11688	SA (0) VP16	62905			
CG14431	56233	MI10741	SA (0) VP16	62905			
CG14521	35928	MI03222	VP16				(4)
CG14947			VP16		C-terminal knock in	TGGATTCGGCGTGGACACGG[TGG]	
CG2016	37997	MI04995	SA (2) p65	62915			
CG31191	44890	MI07868	SA (1) VP16	62908			
CG31689	44287	MI03245	SA (1) VP16	62908			
CG32105			VP16		C-terminal knock in	CTTCACGCAGAATTGTCGCA[AGG]	
CG32791	34458	MI02031	SA (1) p65	62914			
CG4238	58583	MI13092	SA (1) VP16	62908			
CG43737	53823	MI10158	SA (1) VP16	62908			
CG5397	61112	MI15629	SA (0) VP16	62905			
CG9109	34232	MI01660	SA (0) VP16	62905			
CG9743	52108	MI09362	SA (1) VP16	62908			
dac			GAL4AD		C-terminal knock in	TCCGCTGTCGGAGCGACCGA[CGG]	
Dop1R2	92685	CR02529	SA (0) VP16	62905			
erm			VP16		N-terminal knock in	ATCACGGGGCTCTTCGAAGG[CGG]	
eya			VP16		C-terminal knock in	CGACATGGGCTTCTTATGAA[AGG]	
Fife	51081	MI08414	SA (1) VP16	62908			
fs	42987	MI04308	SA (1) VP16	62908			
Gad1	52090	MI09277					(6)
kn	61064	MI15480	p65				(7)
Lim1			VP16		C-terminal knock in	GAACTCCGGCGAATTGGTGC[GGG]	
mbl	30632	MI00139	SA (0) VP16	62905			
Mip			VP16		C-terminal knock in	GCAGAGAATCGCCGTAGCAC[AGG]	
mirr			VP16		C-terminal knock in	CAGCGAGGTCACATTCACGC[CGG]	

MsR2	91323	CR01840	SA (2) p65	62915			
NetA	38035	MI04563	SA (1) VP16	62908			
orb	37978	MI04761	SA (0) VP16	62905			
ort			VP16		N-terminal knock in	TGCATGAAGCACTACGCCAA[AGG]	
ple	84712		p65				(8)
rn	53200	MI09946	SA (0) VP16	62905			
run			VP16		C-terminal knock in	ACCGTGTGGCGGCCCTACTA[GGG]	
Rx	79247	CR00377	SA (2) p65	62915			
sfl	36447	MI02467	SA (2) p65	62915			
slou			VP16		C-terminal knock in	GTAGGGGCCGTATGGCGGAT[AGG]	
SoxN			VP16		N-terminal knock in	CCACATGAGCGCCCATATCG[CGG]	
svp	32734	MI01102	SA (1) VP16	62908			
tey			VP16		C-terminal knock in	TCCGCAGCAGAGCGTTTCGC[AGG]	
tj			VP16		N-terminal knock in	GATCTGCACCACCTGGAGGT[GGG]	
tsh			VP16		C-terminal knock in	CTTCACGCAGAATTGTCGCA[AGG]	
VGlut	82986	MI04979	p65				(9)
wnt4			p65				(5)

Dataset S1 (separate file). All possible split-GAL4 combinations.

SI References

1. F. P. Davis, *et al.*, A genetic, genomic, and computational resource for exploring neural circuit function. *eLife* **9**, e50901 (2020).

2. M. N. Özel, *et al.*, Neuronal diversity and convergence in a visual system developmental atlas. *Nature* **589**, 88–95 (2021).

3. F. Fleuret, Fast binary feature selection with conditional mutual information. *J. Mach. Learn. Res.* **5**, 1531–1555 (2004).

4. M. Courgeon, C. Desplan, Coordination between stochastic and deterministic specification in the Drosophila visual system. *Science* **366**, eaay6727 (2019).

5. B. Ewen-Campen, T. Comyn, E. Vogt, N. Perrimon, No Evidence that Wnt Ligands Are Required for Planar Cell Polarity in Drosophila. *Cell Rep.* **32**, 108121 (2020).

6. F. Diao, *et al.*, Plug-and-play genetic access to drosophila cell types using exchangeable exon cassettes. *Cell Rep.* **10**, 1410–1421 (2015).

7. Q. Xie, *et al.*, Temporal evolution of single-cell transcriptomes of Drosophila olfactory projection neurons. *eLife* **10**, e63450 (2021).

8. B. Deng, *et al.*, Chemoconnectomics: Mapping Chemical Transmission in Drosophila. *Neuron* **101**, 876-893.e4 (2019).

9. H. Lacin, *et al.*, Neurotransmitter identity is acquired in a lineage-restricted manner in the Drosophila CNS. *eLife* **8**, e43701 (2019).

10. F. Gunawan, M. Arandjelovic, D. Godt, The Maf factor Traffic jam both enables and inhibits collective cell migration in Drosophila oogenesis. *Dev. Camb. Engl.* **140**, 2808–2817 (2013).

11. H. Lacin, *et al.*, Genome-wide identification of Drosophila Hb9 targets reveals a pivotal role in directing the transcriptome within eight neuronal lineages, including activation of nitric oxide synthase and Fd59a/Fox-D. *Dev. Biol.* **388**, 117–133 (2014).

12. N. Konstantinides, *et al.*, A complete temporal transcription factor series in the fly visual system. *Nature* **604**, 316–322 (2022).