Supplementary Materials for

Transposition driven genomic heterogeneity in the *Drosophila* brain

Paola N. Perrat, Shamik DasGupta, Jie Wang, William Theurkauf, Zhiping Weng,

Michael Rosbash and Scott Waddell

correspondence to: scott.waddell@cncb.ox.ac.uk

This PDF file includes:

Materials and Methods
Author contributions
Figs. S1 to S6
Tables S1 to S10
References

Materials and Methods

Fly strains

Fly stocks were raised on standard cornmeal food at 25°C and 40-50% relative humidity. The MB-LexA (aka 247-LexA) flies are described (*45*). MB-LexA labels most MB neurons but misses a significant population in the αβ core. We combined c305a-GAL4, c739-GAL4 and NP1131-GAL4 with MB-LexA to generate the driver lines c305a-GAL4; MB-LexA/TM3Sb, c739-GAL4; MB-LexA/TM3Sb and NP1131-GAL4; MB-LexA/TM3Sb, which label the MB α′β′, αβ and γ neurons respectively. We combined lexAop-FLP (*15*) with UAS>STOP>CD8::GFP (*46*) to generate lexAop-FLP; UAS>STOP>CD8::GFP flies. We crossed lexAop-FLP; UAS>STOP>CD8::GFP females with males of the driver lines for the immunohistochemistry and gene expression profiling experiments. elavGAL4/MBGAL80;uas-CD8::GFP flies that label most neurons other than MB neurons were generated by crossing elavGAL4/CyO with MBGAL80;uas-CD8::GFP flies. MBGAL80 flies are described (*12*). Transheterozygous aub^{HN2}/ aub^{QC42} flies were generated by crossing aub^{HN2}/CyO and aub^{QC42}/CyO flies (*33*). Transheterozygous $ago3^{t2}$/$ago3^{t3}$ flies were generated by crossing $ago3^{t2}$/TM6Tb and $ago3^{t3}$/TM6Tb flies (*19*). Transheterozygous $armi^{1}$/$armi^{72.1}$ flies were generated by crossing $armi^{1}$/TM3 to $armi^{72.1}$/TM3 flies (*34*). The $ago2^{414}$ and $dcr\text{-}2^{L811fsX}$ flies are described (*22*). The wild-type *Drosophila* control strain used in the QRT-PCR experiments is w1 (*19*). *D. sechellia* and *D. erecta* were obtained from Joel Levine (University of Toronto) and *D. pseudoobscura* from Aki Ejima (Brandeis University). Oregon R flies used for ovary sequencing are those in (*35*).

Immunohistochemistry

Fly brains were dissected in chilled 1X Phosphate Buffered Saline (PBS) (1.86 mM NaH_2PO_4, 8.41 mM Na_2HPO_4, and 175 mM NaCl), 0.1% Triton X-100 (PBS-T) and fixed for 30 minutes in chilled 4% paraformaldehyde in 1XPBS. After fixation, brains were rinsed four times in PBS-T for 15 minutes at room temperature. The following primary antibodies were added to 1:200 final concentration in PBS-T and brains were incubated overnight at 4°C with gentle rocking: rabbit anti-Aub (*19*), rabbit anti-Ago3 (*19*), and mAb anti-GFP (Invitrogen). Brains were rinsed in PBS-T and incubated overnight with the appropriate secondary antibodies (Jackson Laboratories). Confocal analysis was performed on a Zeiss LSM 5 Pascal confocal microscope. Confocal stacks were processed using ImageJ and Adobe Photoshop.

Cell sorting

RNA isolation and preparation protocol was performed as described previously (*18*). 50-60 brains were dissected from ~5 day old flies into ice-cold modified dissecting saline (9.9 mM HEPES-KOH buffer, 137 mM NaCl, 5.4 mM KCl, 0.17 mM NaH_2PO_4, 0.22 mM KH_2PO_4, 3.3 mM glucose, 43.8 mM sucrose, pH 7.4) containing 50 μM d(–)-2-amino-5-phosphono-valeric acid (AP5), 20 μM 6,7-dinitroquinoxaline-2,3-dione (DNQX), 0.1 μM tetrodotoxin (TTX), and immediately transferred them into modified SM^{active} medium (SM^{active} medium containing 5 mM Bis-Tris, 50 μM AP5, 20 μM

DNQX, 0.1 μM TTX) on ice. Brains were digested with l-cysteine-activated papain (50 units/ml in dissecting saline; Worthington) for 30 minutes at 25°C. Digestion was stopped with five volumes of medium, and brains were washed twice with the chilled medium. Brains were triturated with a flame-rounded 1,000 μl pipetter tip with filter until most of the tissues were dissociated to single cells. The resulting cell suspension was filtered with a 40 μm Nylon Cell Strainer (BD Falcon). GFP-positive cells were sorted using a BD FACSAria Flow Cytometer with FACSDiVa 6.1.1 software. The quality of purified GFP-labeled material was verified by microscopy.

Gene expression analysis

Total RNA was extracted from approximately 10,000 GFP-positive cells and poly(A) RNA was amplified by two-cycle linear amplification as previously described (*14*) and hybridized to a GeneChip *Drosophila* Genome 2.0 array (Affymetrix). Four biological replicates were profiled for each cell type. Scanned Affymetrix image data were processed with the Affymetrix GCOS software to convert to probe level signals. The probe signal files were then processed in CARMAweb (*16*) (https://carmaweb.genome.tugraz.at/carma/index.jsp) using the GCRMA algorithm to normalize and calculate summary values for each probe set. CARMAweb tools were used to identify differentially expressed genes. Genes from each population of cells with average value of replicates ≥7.0 were selected and compared between populations. Genes with a ≥2-fold change difference between populations are reported.

Real-Time PCR

Total RNA from adult fly heads was isolated with Trizol (Invitrogen) and cleaned with RNeasy Micro Kit (Qiagen) with DNAse I treatment. RNA (1 μg) was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and oligo$(dT)_{12-18}$. The cDNA was used for quantitative real-time PCR with ABI PRISM® 7000 Sequence Detection System (Applied Biosystems) with standard cycling parameters (2 min at 50 °C, 10 min at 95 °C, and 40 alternate cycles of 15 s at 95 °C and 60 s at 60 °C). The PCR mixture contained TaqMan® Gene Expression Master Mix and the appropriate Gene Expression Assay (Applied Biosystems). TaqMan custom-made qPCR assays were ordered for all transposons assayed. GAPDH (AB: Dm01841185_m1) was used as endogenous control for normalization of each gene (ΔC_T value). The increase in expression ($\Delta\Delta C_T$ value) was calculated and transformed to the exponential scale.

Deep-sequencing

DNA was prepared for whole genome sequencing by shearing material prepared from approximately 5000 sorted cells (GFP-positive and GFP-negative samples) or 10 μg DNA from embryos. Shearing was performed using 5 cycles of 20 pulses spaced by 30 seconds using a Sonifier 450 (Branson) with the following settings: Output 1, Duty Cycle 50%. The sequencing library was prepared as previously described (*47*) involving DNA end repair, addition of A bases to the 3′ ends and adapter ligation (Illumina PE Adapter Oligo Mix). Adapter-ligated products were purified using the DNA Clean&Concentrator-5 kit (Zymo Research) and amplified by PCR using Phusion DNA polymerase (New England Biolabs) for 30 cycles using the inPE PCR primers 1.0 and 2.0, and individual index primers for each sample (Illumina). PCR fragments of the desired range (200-400 base pairs) were purified using a 2% agarose gel. DNA quality was assessed and quantified using a High Sensitivity DNA Assay on a 2100 Bioanalyzer (Agilent Technologies). Computational identification of transposon insertion sites was performed as previously described (*35*). In total, 12.8M, 8.4M and 58.8M 100-nt paired-end reads were obtained for αβ neurons (GFP-positive), other cells in the brain (GFP-negative), and embryos. Reads were aligned against the *Drosophila* reference genome (not masked for repeats) using the BWA algorithm (*48*), allowing insertions, deletions, and up to three mismatches per read (100 nt). For the three datasets, 2.1M, 2.5M, and 23.3M pairs of reads map to the reference genome properly, i.e., both reads in the pair map and their mapped genomic locations are within 500 nt. For the detection of transposon insertions that were in the experimental genome but not in the reference genome, discordant read pairs were identified with one read mapping to a location in the reference genome and the other mapping to a transposon sequence. Oregon R strain ovarian and embryo DNA was sequenced as described previously (*35*) and new insertions in the ovary sample were determined by comparison to embryo sequence, essentially as above for neurons. 33.1M and 59.2M pairs of reads properly mapped to the reference genome for ovary and embryo sequence respectively.

Author contributions

S.W., P.N.P. and M.R. conceived the idea of this project. S.D. and Y.S. contributed to the development of the intersectional strategy and S.D. provided the intersectional images. W.T. suggested deep-sequencing, provided the ovary data and aided J. W and Z.W. in the computational process required to map transposon insertions. P.N.P designed and performed all other experiments; fly genetics, cell type–specific purification and expression profiling, immunostaining and molecular biology with advice from S.W. and M.R. throughout. S.W. and P.N.P wrote the manuscript with assistance from the other authors.

Fig. S1.

Punctate Stellate labeling in the mushroom body calyx**.** Immunostaining for Ste (magenta) in the brain labels discrete and variable patterns of puncta (white arrow heads) within the MB calyces of flies. The puncta colocalize with the dendrites of MB αβ neurons (green). Scale bar 10μm.

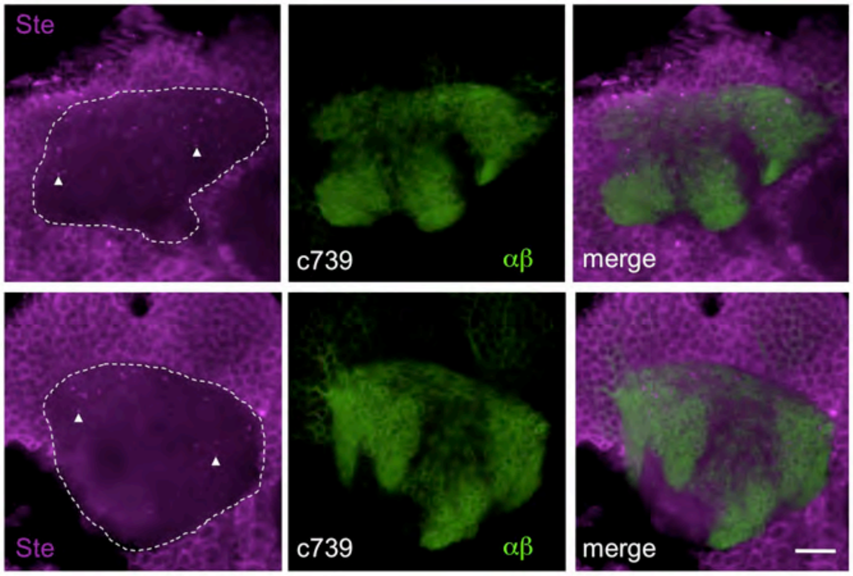

Fig. S2.

Aub labels the α′β′ and γ lobes of the mushroom body and the ellipsoid body of the central complex. Aub immunostaining labels the processes of α′β′ and γ neurons more strongly than those of the αβ neurons in the MB lobes. (**A**) Aub immunostaining (magenta) does not co-localize with the αβ neurons labeled by c739 (green). (**B**) Aub immunostaining (magenta) co-localizes with the α′β′ neurons labeled by c305a (green) and (**C**) the γ neurons labeled by NP1131 (green). (**D**) Aub also strongly labels the ellipsoid body of the central complex (dashed circle). The MB neurons are labeled with FLP-out GFP as in Fig. 1. Scale bar 20μm.

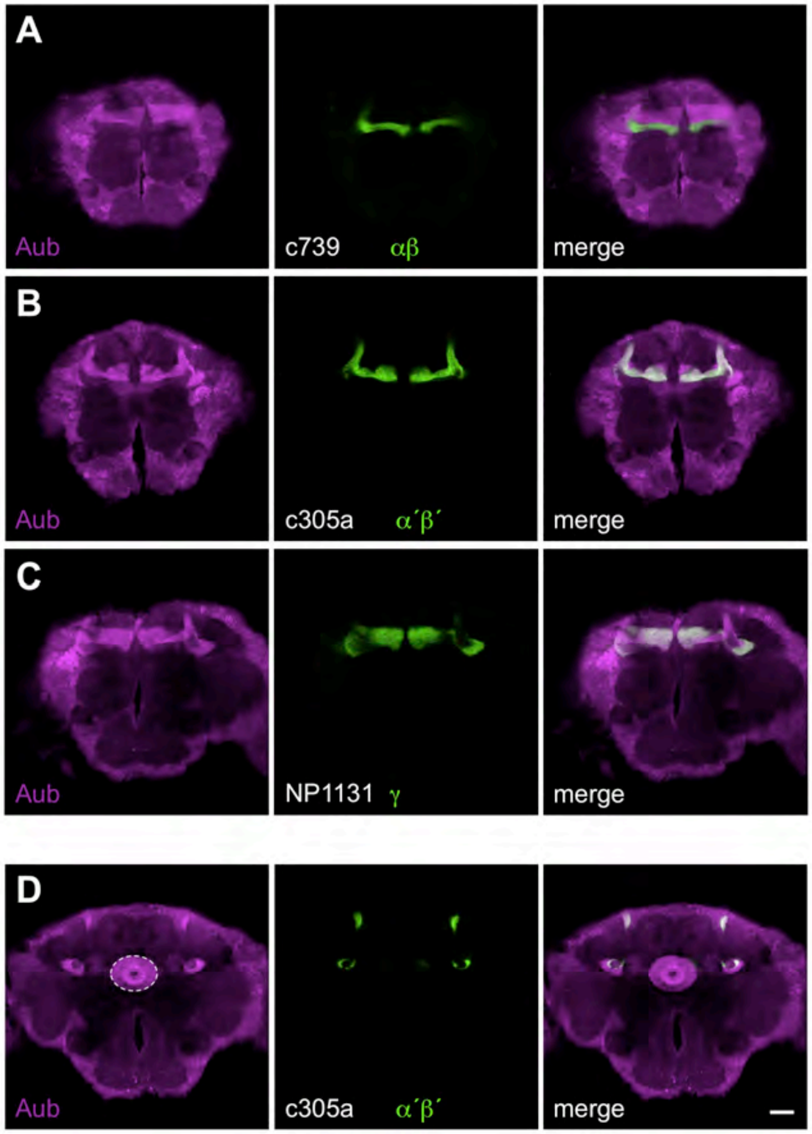

Fig. S3.

Ago3 does not label the mushroom body lobes or the ellipsoid body. Ago3 immunostaining does not colocalise in the MB lobes with (**A**) the αβ neurons labeled by c739, (**B**) the α´β´ neurons labeled by c305a or (**C**) the γ neurons labeled by NP1131 (green). (**D**) Ago3 immonolabeling is also notably absent from the ellipsoid body of the central complex (dashed circle). The MB neurons are labeled with FLP-out GFP as in Fig. 1. Scale bar 20μm.

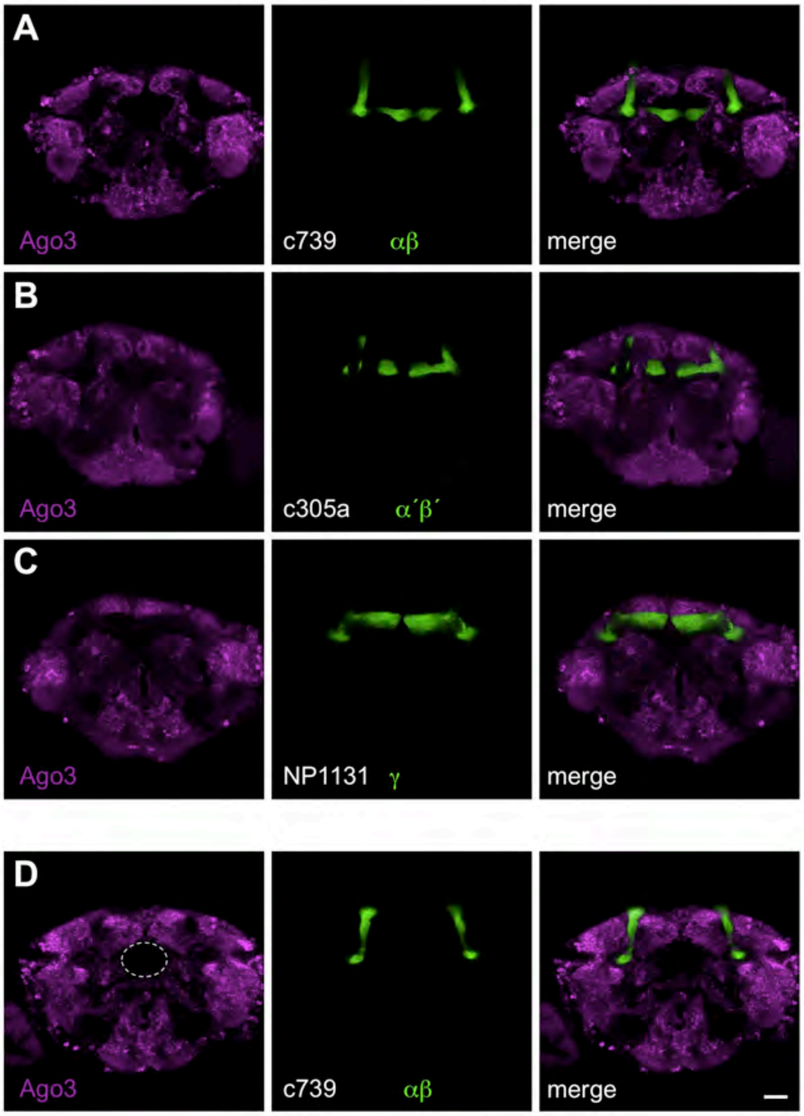

Fig. S4.

Aub and Ago3 indicate similarly organized neurons in the MB peduncle of other Drosophilids. Immunostaining of adult brains from *D. erecta*, *D. sechellia* and *D. pseudoobscura*. Cross sections through the peduncle reveal weaker Aub labeling in the centre of the peduncle than the periphery. Ago3 also labels the periphery of the peduncle and discrete bundles in the centre. Both of these patterns are consistent with those seen in *D.melanogaster* (Fig. 3) where Aub preferentially labels α´β´ and γ neurons and Ago3 labels γ and αβ core neurons. Scale bar 20μm.

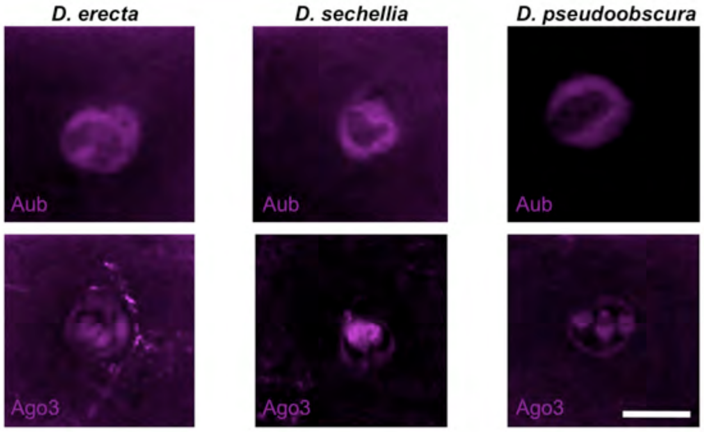

Fig. S5.

Transposon expression is elevated in heads from siRNA mutant flies. Quantitative RT-PCR of transposon transcripts from wild-type, *ago2*^{414} and *dcr-2*^{L811fsX} mutant fly heads. Values normalized to those from wild-type flies. Asterisks denote significant increase from wild-type samples, $P<0.05$ (T-test).

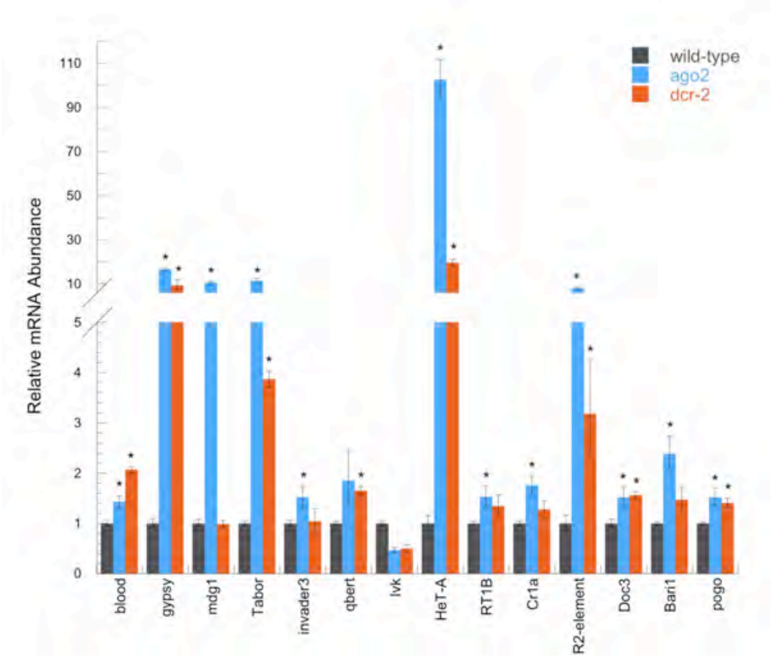

Fig. S6.

Distribution of *de novo* transposon insertions in αβ neurons and in the rest of the brain sample with respect to neighboring genes. Numbers identified in each category are listed above each column.

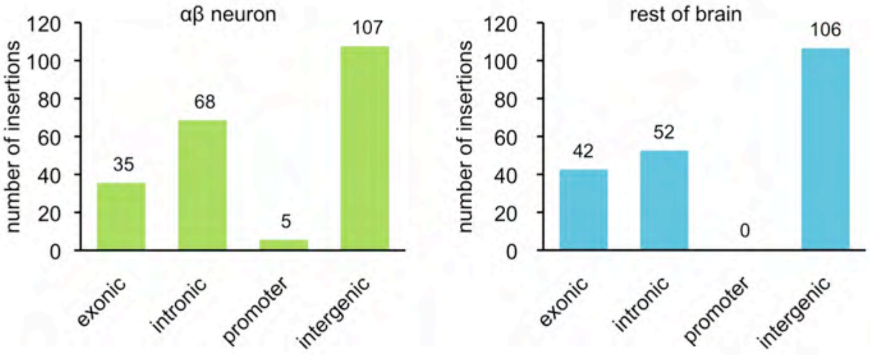

Table S1.

Raw microarray data and analysis illustrated in Fig. 2.

		a'b' replicates				gamma replicates				ab replicates				No MBs replicates				Fold Change		Ttest		Ttest (-log10 scale)	
probeset_id	symbol	305 A.CEL	305 B.CEL	305 C.CEL	305 D.CEL	1131 B.CEL	1131 C.CEL	1131 D.CEL	1131 E.CEL	739 A.CEL	739 B.CEL	739 D.CEL	739 E.CEL	No MBs A.CEL	No MBs B.CEL	No MBs D.CEL	No MBs E.CEL	FC ab/MB	FC ab/NoMB	Ttest ab/MB	Ttest ab/No MB	Ttest ab/MB	Ttest ab/No MB
1628230_x_at	CG31001	2.750484217	3.108032696	2.94586156	5.971881753	4.80837	2.480156	2.604539	5.565506	4.439910633	8.508112912	9.605416587	10.07394845	3.919307	2.498804	2.512647	2.4963	20.78532273	39.39887817	0.034036683	0.02166577	1.468052772	1.664225866
1630855_s_at	Ste12DOR	5.416472124	6.564625968	3.954350984	3.234210561	5.691621	3.948884	4.138842	5.327353	8.340499898	8.092136657	10.07085077	10.06321708	6.340589	3.933091	3.867659	4.614921	20.49402068	21.89623777	0.000481364	0.001333169	3.317526032	2.875114757
1623831_x_at	Bari1 (1)	5.423630917	6.411578849	4.26546081	3.801390114	6.060234	3.275831	3.989595	5.074648	8.008802173	6.415959355	11.04043127	10.90530747	7.970912	3.796465	3.432922	3.641764	19.76435515	20.85193433	0.025600924	0.031740163	1.591744364	1.498390854
1626867_at	His3 (1)	5.327610847	4.739536426	2.946715384	4.342519859	6.81095	5.737295	4.619836	6.934964	7.593281615	7.345697694	11.16509587	11.64166902	7.720578	4.661354	5.689034	3.113973	19.0802426	17.63294201	0.025595534	0.033365619	1.591835804	1.476700812
1635715_at	His4	7.054680505	5.232079254	3.048681054	3.827472135	5.451306	5.057373	5.40309	5.151196	7.730341208	7.21392797	10.52541215	10.67375245	5.598684	4.894795	5.005832	4.515849	16.08477332	16.35963259	0.013782272	0.018232237	1.860679191	1.739160041
1635666_at	His2B (1)	6.92908543	7.653992132	3.512701039	9.687268506	7.115474	5.706132	6.243682	8.429464	8.816270623	9.914799688	12.03452738	11.98302971	6.950037	5.993195	6.389337	5.677715	13.71261639	21.62436479	0.007892312	0.007670332	2.102795733	2.115185812
1629740_at	His1 (1)	8.932577888	8.602257623	3.947352787	9.886514427	7.83409	5.005036	7.326277	6.169507	10.36214912	9.243435998	12.02003848	11.97326236	7.571815	5.759393	5.163635	6.269989	12.87740417	26.14591202	0.004785117	0.001788442	2.32010742	2.747525172
1638308_s_at	His3 (2)	7.17031798	7.267084295	5.524718172	8.310408644	8.430632	7.062304	6.794978	8.425312	9.540564502	9.334820332	12.36127283	12.51162805	8.914674	5.96791	6.625275	4.967673	11.82568659	19.94822436	0.018733268	0.011648114	1.727386465	1.933744379
1623281_s_at	His2B (2)	5.388329924	6.089971672	4.021206207	5.099143562	5.331999	4.923526	4.444916	5.94268	7.044917602	7.753855303	10.26436696	9.692351258	5.378637	5.07682	5.356065	4.932511	11.58070545	11.33619464	0.014589977	0.01860206	1.835945403	1.730438965
1627112_s_at	Bari1 (2)	6.552343274	8.517532155	4.47885604	8.453778261	7.94547	5.0448	6.104266	4.912055	8.965121815	10.28979956	10.5053388	10.08444575	6.365792	4.097193	3.962808	3.773072	11.00463119	42.5610026	0.000453782	0.000736686	3.343153203	3.132717683
1626205_s_at	G6	3.848224973	3.781624455	2.862366312	3.88307434	4.147621	3.014934	2.956827	4.519435	6.495785611	3.412588581	9.707362693	8.53868128	3.284131	2.853763	3.546646	2.773656	10.64306044	15.17953342	0.08810256	0.064056802	1.055011473	1.193434748
1623349_x_at	RT1B	7.269100507	7.630704237	5.977891307	6.595116876	7.404572	7.292809	7.502578	8.327297	9.629103981	9.043961754	11.87843651	11.82529343	9.748104	7.713699	7.542569	7.882655	10.15550659	5.17816942	0.01489393	0.042735307	1.826990703	1.36921317
1631349_s_at	Tabor	7.727650104	7.524563707	6.554575776	10.64199832	10.255897	5.433161	7.80441	6.04311	9.315967723	11.46353977	11.42508003	11.47184771	8.390501	5.899294	5.499959	5.612837	9.006322068	23.72705367	0.00419104	0.00227333	2.37767818	2.643337479
1630585_s_at	HeT-A (1)	6.782952399	6.105841958	4.166968286	6.626571833	6.051374	5.593808	5.934395	6.9269	8.648087827	7.263955551	10.66719436	10.19324827	6.168085	4.610694	4.856809	4.475001	8.997474251	17.94455514	0.019114876	0.006672862	1.718628521	2.175687856
1623145_s_at	PlexinB pseudogene (1)	6.662679974	6.004905544	3.877281351	6.147596001	7.805013	4.891025	6.273921	4.249171	9.043219687	7.855353022	8.975315887	8.977144481	4.626327	4.688658	4.623283	4.816064	7.856077493	16.2703713	0.000281342	0.000625728	3.550765843	3.203614655
1629641_s_at	Ivk (1)	4.488626904	5.807030988	4.333164895	5.684732975	6.67619	4.197349	4.967805	4.242637	7.325095978	9.128581266	7.954323142	7.554195201	5.204239	3.291164	3.321195	3.341419	7.678671884	18.39248675	0.000761912	0.000558084	3.118095256	3.253300474
1631321_s_at	His1 (2)	8.591447007	9.597592975	5.008287084	9.76135233	8.911817	7.05676	7.973798	8.174473	9.764517787	10.34263301	11.9566701	11.85145412	8.452103	6.403026	5.466868	6.991583	7.18196105	17.75832682	0.005717756	0.002598515	2.242774392	2.585274819
1641464_s_at	CG32850 (1)	8.231118092	6.824724366	7.383606634	8.896891445	5.803326	5.945558	5.743512	6.232602	8.656747259	9.699385394	9.996674933	10.52064925	7.214473	5.40108	5.414946	5.179481	7.138875752	15.09364361	0.000855537	0.000806199	3.067761011	3.093557796
1625195_s_at	mdg1	7.723370742	9.125102534	7.131230515	7.474198432	8.603922	7.691796	7.493817	7.632168	10.00668145	10.58229863	11.24039611	10.89469547	7.305291	6.001374	5.06755	6.109862	7.069299479	23.58828562	5.56585E-05	0.000436432	4.254468889	3.360083534
1630452_at	transib3	7.791952141	8.354744376	8.738096496	10.77079408	8.917823	6.38563	7.83314	6.059947	9.898904532	10.14771856	11.80747945	11.7000788	6.448621	5.383949	4.97508	5.390919	6.878192278	40.47342636	0.004340211	0.000270541	2.362489162	3.567767514
1627745_s_at	Ivk (2)	6.24905598	7.690930128	6.752211137	9.310272631	8.667096	5.912318	6.621723	6.110378	8.329007058	10.63986831	10.49971966	10.20937125	7.837039	5.778275	5.496041	5.837945	6.751665644	12.83638063	0.00540783	0.002883554	2.266977009	2.540071903
1640606_x_at	Cr1a	7.807733392	8.962528129	6.767736892	8.969036558	8.623506	7.263444	7.737445	6.896079	9.514741456	10.4941312	11.24463047	11.10423724	8.61642	6.759861	6.382143	6.935796	6.547737428	10.67289824	0.001102166	0.001938489	2.957753165	2.712536671
1623575_at	F-element	4.812765264	6.910016196	4.010662913	5.322563214	4.869178	2.833016	4.496545	2.603482	6.963013185	7.577246672	6.942273395	7.067611128	4.063233	2.426925	2.533144	2.895808	6.299588189	17.84884227	0.000755511	0.000549692	3.121759207	3.259880683
1632295_s_at	Doc2	9.126150011	9.937592768	8.927501147	11.34700806	10.549423	8.455724	9.325692	7.606635	10.68713079	12.31017848	12.57351387	12.40889395	10.573059	8.393798	8.490272	7.792436	6.002084058	9.079029075	0.002704172	0.006645681	2.567965716	2.177460492
1630420_x_at	baggins	5.191184816	4.14709923	2.810984986	5.056727291	6.760737	3.216531	5.363317	5.409498	5.219247488	7.361102365	8.097289872	8.37613421	6.460582	4.710974	4.427994	4.727143	5.731582674	4.537099295	0.027340231	0.049217145	1.563197825	1.30788358
1630934_at	R2-element	8.270138491	7.853319485	7.772221878	9.520945556	9.400255	8.325928	7.94057	8.767086	9.964753399	11.1904391	11.3071834	11.53197956	11.202932	10.8212	11.066428	10.863226	5.725020329	1.007054923	0.001100336	0.979276133	2.958474669	0.00909483
1627899_at	Ucrh	7.515951643	6.251747126	5.036763949	5.781445111	6.774356	7.149992	5.377248	6.324225	8.705295768	7.680291305	9.460798453	9.121575033	5.695259	2.909047	3.259708	3.701523	5.523275291	28.85212931	0.001708145	0.001175321	2.767475371	2.929843411
1640702_at	PlexinB pseudogene (2)	5.838303911	7.533703674	5.841388617	9.071992912	9.302044	3.895943	7.615844	5.644073	8.651655906	9.764568714	9.473226322	9.12829323	5.351091	4.282998	3.732864	4.992703	5.320361973	25.36068912	0.007810761	9.45866E-05	2.107306678	4.024170472
1625877_s_at	CG40271	7.427269237	8.685916539	6.628961437	8.498625751	7.380271	6.47914	7.098056	6.48639	8.805901209	9.528178143	10.61374852	9.832700449	8.564625	6.330894	6.193006	6.201037	5.13211442	7.324557382	0.001791008	0.008434846	2.746902574	2.073922855
1629207_at	CG6287	5.756287358	5.263531819	4.137824116	3.431772525	6.384746	6.740503	5.966742	6.470525	6.450650023	6.949808947	8.818415256	8.908307057	5.146616	4.927105	5.229884	5.108479	4.799232956	6.403116774	0.025996435	0.023508051	1.585086208	1.628783377
1635135_at	roo	8.983301869	9.97689871	7.587740349	10.97001886	9.347114	8.279764	9.036076	8.782942	11.24616651	11.36942572	11.38844381	11.2355401	8.706187	8.371763	7.435548	8.481967	4.561195676	8.345669486	0.000502155	0.001404768	3.299162373	2.852395369
1641210_s_at	qbert	5.517911161	7.65889294	4.729016373	5.966258828	6.267848	5.165433	4.918646	4.756537	7.577123593	7.456490767	8.251013779	7.783352068	5.75106	4.140657	3.609597	4.347481	4.42116672	9.881953735	0.000331088	0.002841722	3.480056459	2.546418418
1624819_s_at	gypsy	7.367753392	7.979557299	4.863867011	7.031827555	7.203636	7.417215	7.253301	8.61297	9.501090451	9.90430194	8.920566055	9.090265771	5.087993	5.028359	4.495533	4.953252	4.400874245	22.0509943	0.000695318	1.24546E-05	3.157816347	4.904670192
1637749_at	D2R	6.084513814	6.945988665	2.597440356	4.527469446	6.303451	5.513498	6.022514	5.25239	6.27461253	6.57621533	8.822935093	8.42969964	6.543676	6.47436	5.909626	6.712122	4.346811361	2.167331219	0.036024678	0.180949066	1.443399897	0.742443655
1624224_at	HeT-A (2)	7.373690343	7.326084999	7.72082736	7.369439349	7.378971	7.303665	7.37369	7.37369	8.791972155	8.657532741	10.00842868	10.60590238	7.369439	6.516674	6.859901	6.409939	4.327253777	6.620640243	0.020387946	0.005554018	1.690626532	2.255392681
1627936_s_at	invader3	9.199035578	10.14244377	6.6216489	10.22406138	8.922932	7.214795	8.02305	7.035854	10.39353002	9.226559805	11.21509866	11.19608951	8.449743	4.043619	4.088926	4.395904	4.242286093	38.40631111	0.01409213	0.010176673	1.85102336	1.992394198
1635829_s_at	gypsy2	5.237727552	6.288969813	3.388325045	5.021198452	5.245953	4.849798	5.305538	4.870252	7.233742773	5.14911949	8.214933035	7.764289856	5.146619	5.122248	5.285284	4.860567	4.18303786	3.963683556	0.047058446	0.059614844	1.327362417	1.224645585
1630948_s_at	Doc3	8.031929864	8.134976298	4.820305663	8.80364412	8.483896	7.517098	7.98342	6.989797	9.011115854	9.360522009	10.02352325	10.21256038	7.302352	6.557231	6.994852	5.964012	4.159173895	7.713141744	0.002852529	0.000340299	2.544769954	3.468139773
1639054_s_at	pogo	5.538545701	7.668476971	5.274432006	7.802789238	8.29978	6.853604	7.651876	8.723612	9.206096376	8.91855194	9.331611896	9.426809624	8.025577	6.366962	5.293716	6.423339	3.983752362	6.46822094	0.002455952	0.015460987	2.609780212	1.810762778
1627513_at	CG17684 (1)	6.491688438	6.021671505	4.961061902	7.777076637	6.690588	4.694914	6.157491	5.37416	7.588903084	6.916679589	8.766451196	8.457665195	5.881943	3.771789	4.395124	3.750898	3.761591882	11.17719494	0.009850856	0.001948541	2.006526049	2.710290346
1641450_s_at	microcopia	5.040311617	5.870867667	3.921589995	8.34854305	6.601501	5.395948	5.92558	4.252171	8.775287325	7.917627936	6.819584906	6.642011089	6.396961	4.551791	3.86753	4.704009	3.652954471	6.314003515	0.027831436	0.011149515	1.555464383	1.952744008
1640167_s_at	412	9.6660505	9.126946536	10.06334039	9.601451085	9.531107	9.638427	9.126487	9.134134	11.16177384	10.35745084	11.70242556	11.63924149	7.681359	6.941426	6.915912	6.602506	3.315508133	18.12516449	0.006942661	6.30012E-05	2.158474013	4.200651038
1624440_at	repetitive sequence	6.050024254	6.328289878	4.098797048	6.054589255	5.82632	5.329089	5.984577	5.113031	6.814119154	6.625288632	7.698504266	8.13637238	5.335648	4.23922	4.08478	4.074305	3.295463584	7.387482757	0.00760864	0.000949488	2.118692965	3.022510729
1636174_at	GstD9	6.485044803	5.755333282	4.810223003	3.848126843	5.505552	5.521123	5.396317	5.422899	6.027545732	7.222251038	7.890894704	7.022669563	3.304268	3.307386	2.974095	2.944227	3.243975431	15.01514248	0.011128727	0.001296162	1.953554507	2.887340631
1625997_s_at	Doc	10.271787	10.36836397	10.86392787	11.22264933	11.529171	10.175759	10.833691	9.800066	12.25258586	12.01527739	12.55824569	12.40839964	9.605896	8.029389	7.831525	8.103897	3.194190259	15.094493	3.48835E-05	0.001444993	4.45737988	2.840134267
1635696_s_at	Gfat1	6.31359999	8.384019155	5.36483354	8.858774641	7.696898	6.16085	6.491182	5.693123	8.233247831	9.344926973	8.348875216	8.080024253	6.298019	4.870257	5.136285	4.477243	3.098045918	9.892379374	0.012486807	0.000707747	1.903548592	3.150122103
1632545_s_at	CG5522	4.882850506	5.696051668	4.878497066	5.974935729	5.909682	6.322687	5.106054	4.784862	7.386459561	6.996821587	6.820092177	7.074276013	5.911414	6.011927	5.684698	5.819772	3.084335788	2.317323701	5.91788E-05	0.00035937	4.22783405	3.444458254
1624060_at	bab2	6.655533174	4.986385869	4.846206158	4.269932112	7.236555	6.715308	6.357587	6.855924	8.106580005	6.20773646	8.312205641	7.823561932	6.385078	7.887677	7.872642	7.301585	3.078210934	1.189846536	0.034379342	0.689280293	1.463702441	0.161604138
1634633_s_at	accord	5.246641692	5.446843215	3.79134762	6.126816502	6.292941	5.714985	6.026766	5.900168	7.627692548	7.43210687	6.928397669	6.720759632	5.43261	4.724705	3.959759	4.961272	3.050245898	5.306102996	0.001086473	0.001050794	2.963981033	2.978482432
1624390_a_at	thoc7	7.419064978	7.034249887	3.659974287	4.7912996	6.914326	7.6175	6.814607	6.528603	5.885767499	7.570668055	9.000049071	9.339861634	7.415589	7.207194	6.944379	7.233254	3.034867373	1.680607194	0.140140327	0.412895621	0.853436873	0.384159723
1625632_at	robl	7.929735353	8.934394316	6.00574409	6.711881426	8.973552	9.1534	9.504997	9.714068	9.844726995	9.726014201	10.21583197	10.01877482	9.172431	9.258362	9.478841	9.165007	3.000838079	1.605115093	0.012930807	0.002718335	1.888374375	2.565697065
1629640_at	Tc3	5.777900883	6.883147742	3.517523777	6.253157028	5.364615	5.045136	6.014311	5.140666	7.307510104	6.98903137	7.533324635	6.465856423	5.332057	4.37531	4.613116	4.918326	2.978061308	4.803974134	0.004067133	0.000357221	2.390711665	3.447062472
1623250_at	CG18591	4.436426217	6.815750307	2.946611414	2.740754559	6.888383	7.16096	6.492469	6.91467	6.973418094	7.573870783	7.113329113	6.692155289	6.557116	6.503421	6.806118	6.348947	2.905306271	1.448232086	0.055839532	0.054786393	1.253058231	1.26132729
1624693_at	Arf102F	6.6736374	7.000694345	7.548952295	4.221351924	6.507424	7.666205	7.584631	7.710369	8.623725688	7.67248756	8.836539175	8.309285565	8.407907	7.358358	7.153912	7.287693	2.821282976	1.751440524	0.011446948	0.080670656	1.941310285	1.09328441
1624406_at	CG5168	6.266359371	5.011069152	3.39268845	6.857362421	5.984599	5.389048	6.22491	5.554712	5.214313893	8.363570964	7.833216257	6.902311825	6.077197	5.926901	6.27753	5.989459	2.81524345	2.01472298	0.117671554	0.239852893	0.929328513	0.620055039
1636708_at	CG3061	8.469922004	9.273040131	6.359133936	5.477315387	9.254413	9.131039	9.420459	8.427184	9.221016453	9.860504946	10.20707802	9.542715995	9.845349	9.561151	9.577911	9.576724	2.791935371	1.047931999	0.028553544	0.777896651	1.544339982	0.109078098
1630857_s_at	NTPase	6.149907347	6.474828747	5.629114339	5.089230175	5.905095	6.969255	6.191752	6.984074	6.989673758	7.083077211	8.037252811	8.341144741	6.890192	6.487153	6.340033	6.730685	2.710633485	2.001069646	0.013152389	0.053443518	1.880995342	1.272104958
1624755_a_at	flw	9.071842419	6.907994378	8.33376232	4.063701058	8.12132	8.654358	8.785397	9.154784	8.415046084	8.986345766	9.902203436	9.683254636	8.514204	8.661093	8.927645	8.201686	2.566972043	1.591685696	0.077542093	0.14265404	1.110462481	0.845715926
1635886_s_at	blood	7.117483123	6.908364241	5.841644636	7.061758097	7.503893	7.182727	7.278087	7.703209	8.710140387	8.483239499	8.380681251	8.015387942	5.818415	4.548246	4.772964	4.925895	2.501366589	10.41782549	0.000312218	0.00021738	3.505541598	3.662781016
1640932_s_at	CG32626	7.205855925	7.945741735	4.313664508	5.427602467	7.327457	8.01552	7.708027	7.215494	7.80480275	8.986516718	7.830494475	8.235528497	6.383722	6.290532	6.5282	6.421194	2.495649415	3.502593614	0.035498768	0.006224873	1.44978672	2.205869522

Table S2.

Summary table of deep-sequencing data analyzed in Fig. 4.

	MB neuron	non-MB neuron	embryo	Oregon R ovary	Oregon R embryo
read pairs	6,413,050	4,226,518	29,399,946	34,363,229	62,590,728
properly paired pairs	2,078,839	2,511,774	23,309,307	33,103,300	59,244,707
non properly paired pairs	4,334,211	1,714,744	6,090,639	1,259,929	3,346,021
N of insertions	1,107	1,271	3,890	7,683	8,400
genome coverage (fold)	3	4	34	48	87
N of insertions not in embryo	234	219	-	3,860	-
N of insertions not in embryo at 1-fold genome coverage	129				

Table S3.

Genomic location of exonic, intronic and promoter insertions in αβ neurons. TSS, transcription start site.

exonic insertions

chr	start	end	transposonName	strand	transposonStrand	gene	full name
chr2L	2972995	2973095	FBgn0000004_17.6	-	antisense	FBgn0004584,FBgn0260639	Recombination repair protein 1; gamma-Tubulin at 23C
chr2L	4682874	4682974	FBgn0005384_3S18	-	sense	FBgn0031620	
chr2L	7233939	7234039	FBgn0044355_Quasimodo	+	antisense	FBgn0259111	Na[+]-driven anion exchanger 1
chr2L	15034699	15034799	FBgn0000481_Doc	-	sense	FBgn0260446	metabotropic GABA-B receptor subtype 1
chr2L	18051917	18052017	FBgn0000638_FB	-	sense	FBgn0243486	reduced ocelli
chr2L	19208084	19208185	FBgn0002698_mdg3	+	antisense	FBgn0015380	derailed
chr2L	19606191	19606291	FBgn0000005_297	+	antisense	FBgn0000464	Leukocyte-antigen-related-like
chr2R	5850464	5850564	FBgn0000155_roo	+	sense	FBgn0033466	Peptidyl-alpha-hydroxyglycine alpha-amidating lyase 1
chr2R	6323927	6324027	FBgn0061485_rover	+	sense	FBgn0033523,FBgn0262114	
chr2R	8368693	8368793	FBgn0000155_roo	-	sense	FBgn0033744	
chr2R	8530163	8530263	FBgn0000004_17.6	-	sense	FBgn0033760	
chr2R	8771081	8771181	FBgn0000155_roo	-	antisense	FBgn0002069	Aspartyl-tRNA synthetase
chr2R	10502619	10502719	FBgn0067380_invader6	+	sense	FBgn0026427	
chr3L	1559891	1559991	FBgn0000349_copia	-	sense	FBgn0035233,FBgn0035235	
chr3L	3902821	3902921	FBgn0001207_HMS-Beagle	-	antisense	FBgn0003943	Ubiquitin-63E
chr3L	6320214	6320397	FBgn0026065_Idefix	-	sense	FBgn0041624	Odorant receptor 65b
chr3L	9399329	9399429	FBgn0000155_roo	+	antisense	FBgn0035986	
chr3L	9444697	9444797	FBgn0044355_Quasimodo	-	sense	FBgn0035996	
chr3L	12133342	12133442	FBgn0000349_copia	-	sense	FBgn0036257	Rho GTPase-activating protein 68F
chr3L	14660017	14660117	FBgn0000481_Doc	+	sense	FBgn0041604	dally-like
chr3L	17474151	17474251	FBgn0003007_opus	+	antisense	FBgn0036731	
chr3L	20420598	20420698	FBgn0000155_roo	+	sense	FBgn0036987,FBgn0036988	Transmembrane protein 104 homolog
chr3R	3576205	3576305	FBgn0000155_roo	-	antisense	FBgn0037513	
chr3R	5408606	5408706	FBgn0005384_3S18	+	sense	FBgn0037717,FBgn0053654	
chr3R	7617657	7617757	FBgn0000155_roo	+	sense	FBgn0025802	SET domain binding factor
chr3R	12101043	12101143	FBgn0000005_297	-	antisense	FBgn0250823	gilgamesh
chr3R	12294543	12294643	FBgn0001249_I-element	-	sense	FBgn0250754	
chr3R	15026475	15026575	FBgn0043055_Ivk	+	sense	FBgn0051475	
chr3R	17001743	17001870	FBgn0000638_FB	-	antisense	FBgn0025803	SNF4/AMP-activated protein kinase gamma subunit
chr3R	21677552	21677652	FBgn0000155_roo	-	sense	FBgn0039379	
chr3R	25874141	25874241	FBgn0000155_roo	-	antisense	FBgn0039741	
chr4	355258	355358	FBgn0000155_roo	+	antisense	FBgn0259214	
chrU	2929947	2930046	FBgn0000481_Doc	-	sense	FBgn0085520	
chrX	1244548	1244648	FBgn0000006_412	-	antisense	FBgn0024365	
chrX	3979121	3979221	FBgn0000481_Doc	+	antisense	FBgn0263772	
chrX	14147521	14147621	FBgn0000199_blood	+	antisense	FBgn0002873	mushroom body defect
chrX	14152429	14152530	FBgn0040267_Transpac	+	antisense	FBgn0002873	mushroom body defect
chrX	18466687	18466787	FBgn0001167_gypsy	-	sense	FBgn0000242	Beadex

intronic insertions

chr	start	end	transposonName	strand	transposonStrand	gene	gene fullname
chr2L	167031	167131	FBgn0000199_blood	-	antisense	FBgn0016977	split ends
chr2L	6311078	6311178	FBgn0005773_Bari1	+	antisense	FBgn0053531	Discoidin domain receptor
chr2L	7551610	7551710	FBgn0000155_roo	-	sense	FBgn0085403	Rapgap1
chr2L	7788913	7789013	FBgn0000005_297	+	antisense	FBgn0031948	
chr2L	8911266	8911366	FBgn0063897_Stalker4	+	sense	FBgn0085427	
chr2L	8911266	8911366	FBgn0003519_Stalker	+	sense	FBgn0085427	
chr2L	14227207	14227307	FBgn0000007_1731	+	sense	FBgn0016930	smell impaired 35A
chr2L	17819966	17820065	FBgn0000349_copia	-	sense	FBgn0262018	Cadherin-N2
chr2L	18521463	18521563	FBgn0063897_Stalker4	+	antisense	FBgn0032685	
chr2L	18521463	18521563	FBgn0003519_Stalker	+	antisense	FBgn0032685	
chr2L	19357124	19357224	FBgn0000652_F-element	+	antisense	FBgn0024245	doughnut on 2
chr2L	21764851	21764951	FBgn0002697_mdg1	-	sense	FBgn0051612	
chr2R	236865	236965	FBgn0043055_Ivk	+	sense	FBgn0260798	G protein-coupled receptor kinase 1
chr2R	7060624	7060777	FBgn0002697_mdg1	-	sense	FBgn0003396	schnurri
chr2R	12571449	12571569	FBgn0002697_mdg1	+	antisense	FBgn0050463	
chr2R	14360560	14360660	FBgn0000652_F-element	-	antisense	FBgn0028496	
chr2R	15506829	15507018	FBgn0000638_FB	+	sense	FBgn0003435	smooth
chr2R	16464761	16464861	FBgn0010302_Burdock	-	sense	FBgn0028622	quasimodo
chr2R	16780579	16780679	FBgn0026065_Idefix	-	sense	FBgn0015524	orthopedia
chr2R	18787072	18787172	FBgn0000349_copia	+	antisense	FBgn0034797	nahoda
chr2R	20977572	20977672	FBgn0000638_FB	-	sense	FBgn0003715	
chr2RHet	1738882	1738982	FBgn0000349_copia	+	antisense	FBgn0020908	Sarcoplasmic calcium-binding protein 1
chr2RHet	1883038	1883138	FBgn0063897_Stalker4	+	antisense	FBgn0260995	
chr2RHet	1883038	1883138	FBgn0003519_Stalker	+	antisense	FBgn0260995	
chr2RHet	3151985	3152085	FBgn0000005_297	+	antisense	FBgn0262124	unextended
chr3L	1169329	1169429	FBgn0000224_BS	-	sense	FBgn0025525	bric a brac 2
chr3L	2382353	2382453	FBgn0000004_17.6	-	antisense	FBgn0035338,FBgn0261551	
chr3L	2438252	2438352	FBgn0000481_Doc	-	sense	FBgn0261551	
chr3L	3343281	3343381	FBgn0000155_roo	-	antisense	FBgn0004167	karst
chr3L	5003744	5003844	FBgn0043969_diver	-	antisense	FBgn0005775	Connectin
chr3L	5269184	5269284	FBgn0000005_297	-	antisense	FBgn0052423	alan shepard
chr3L	6470991	6471091	FBgn0001167_gypsy	+	antisense	FBgn0261801	
chr3L	8521510	8521610	FBgn0000155_roo	-	antisense	FBgn0024236	fear-of-intimacy
chr3L	10646082	10646182	FBgn0001249_I-element	-	antisense	FBgn0052066	
chr3L	14487533	14487633	FBgn0002698_mdg3	+	sense	FBgn0087007	big bang
chr3L	15596598	15596698	FBgn0063429_invader2	+	antisense	FBgn0036518	RhoGAP71E
chr3L	15617853	15617953	FBgn0063917_McClintock	-	sense	FBgn0259236	comm3
chr3L	15991804	15991904	FBgn0000638_FB	-	antisense	FBgn0036556	Limpet
chr3L	16730090	16730190	FBgn0001167_gypsy	-	antisense	FBgn0036659	
chr3L	16868582	16868746	FBgn0001283_jockey	+	antisense	FBgn0261565	
chr3L	17200666	17200919	FBgn0000005_297	+-	antisense	FBgn0260943	RNA-binding protein 6
chr3L	17708204	17708342	FBgn0043055_Ivk	+	antisense	FBgn0052183	Ccn
chr3L	19369673	19369773	FBgn0000005_297	+	antisense	FBgn0052206	
chr3L	23276051	23276151	FBgn0000155_roo	-	sense	FBgn0037212	nicotinic Acetylcholine Receptor alpha 80B
chr3LHet	1539109	1539209	FBgn0040267_Transpac	+	sense	FBgn0085612	
chr3R	5421583	5421683	FBgn0000638_FB	-	antisense	FBgn0037720	
chr3R	7598601	7598701	FBgn0000155_roo	-	sense	FBgn0086910	lethal (3) neo38
chr3R	8965310	8965410	FBgn0000155_roo	+	antisense	FBgn0038118,FBgn0260793	timeout
chr3R	10293337	10293437	FBgn0000004_17.6	+	sense	FBgn0086901	crossveinless c
chr3R	10401492	10401592	FBgn0000006_412	-	antisense	FBgn0038244	
chr3R	14333327	14333427	FBgn0000006_412	+	sense	FBgn0004652	fruitless
chr3R	14421607	14421707	FBgn0000155_roo	+	sense	FBgn0038629,FBgn0263974	qin
chr3R	17476805	17476905	FBgn0046110_Juan	+	sense	FBgn0011766	E2F transcription factor
chr3R	19895533	19895633	FBgn0000005_297	-	sense	FBgn0053100	4EHP
chr3R	25612177	25612277	FBgn0010302_Burdock	-	antisense	FBgn0001297	kayak
chr3RHet	388880	388979	FBgn0040267_Transpac	+	sense	FBgn0040034	
chr3RHet	1200593	1200693	FBgnnnnnnnn_HMS-Beagle2	+	antisense	FBgn0058232	
chr3RHet	1261015	1261115	FBgn0000481_Doc	+	antisense	FBgn0058232	
chrU	125084	125184	FBgn0001249_I-element	-	antisense	FBgn0058378	
chrU	455427	455527	FBgn0063428_invader3	-	antisense	FBgn0261338	
chrX	1461735	1461835	FBgn0000155_roo	-	antisense	FBgn0025390	Mucin related 2B
chrX	1461738	1461838	FBgn0000155_roo	-	sense	FBgn0025390	Mucin related 2B
chrX	2803098	2803198	FBgn0000155_roo	-	antisense	FBgn0028369	kin of irre
chrX	3170814	3170914	FBgn0043055_Ivk	-	antisense	FBgn0000479	dunce
chrX	3647494	3647592	FBgn0046110_Juan	-	sense	FBgn0086899	Tousled-like kinase
chrX	4238801	4238901	FBgn0000155_roo	+	sense	FBgn0262738	no receptor potential A
chrX	8388138	8388238	FBgn0000652_F-element	-	sense	FBgn0053181	
chrX	12081980	12082080	FBgn0000199_blood	+	antisense	FBgn0259240	Tenascin accessory
chrX	13169418	13169518	FBgn0000005_297	-	sense	FBgn0041210	HDAC4
chrX	14699692	14699792	FBgn0043055_Ivk	-	antisense	FBgn0003301	rutabaga
chrX	15296435	15296536	FBgn0000004_17.6	+	sense	FBgn0263257	CNG channel-like
chrX	19327236	19327336	FBgn0005773_Bari1	+	sense	FBgn0031016	kekkon5
chrX	21125471	21125571	FBgn0026065_Idefix	+	antisense	FBgn0052521	

promoter insertions

chr	start	end	transposonName	strand	transposonStrand	gene	distance to TSS
chr2R	15051397	15051497	FBgn0000155_roo	-	sense	FBgn0034420	-254
chr3R	485015	485115	FBgn0000224_BS	-	sense	FBgn0263346	-289
chr3R	11570346	11570446	FBgn0063897_Stalker4	-	antisense	FBgn0038355	-29
chr3R	11570346	11570446	FBgn0003519_Stalker	-	antisense	FBgn0038355	-29
chrU	8541	8641	FBgn0000005_297	-	sense	FBgn0261404	-796

Table S4.

GO-term analysis of genes disrupted by αβ neuron insertions (from -1k to the end of the gene). The Count column lists the number of genes with insertions in αβ neurons for each GO term. The Benjamini FDR column indicates the false discovery rate after multiple-testing correction. GO terms with FDR less than 5% are bold.

Category	Term	Count	%	P-Value	Benjamini FDR
GOTERM_BP_FAT	**cell morphogenesis involved in neuron differentiation**	**10**	**11.2**	**1.90E-04**	**2.70E-02**
GOTERM_BP_FAT	**cell motion**	**10**	**11.2**	**2.90E-04**	**2.80E-02**
GOTERM_BP_FAT	**cell morphogenesis involved in differentiation**	**10**	**11.2**	**2.70E-04**	**3.10E-02**
GOTERM_BP_FAT	**cell projection morphogenesis**	**10**	**11.2**	**4.10E-04**	**3.40E-02**
GOTERM_BP_FAT	**neuron projection development**	**10**	**11.2**	**1.80E-04**	**3.50E-02**
GOTERM_BP_FAT	**neuron differentiation**	**11**	**12.4**	**5.80E-04**	**3.70E-02**
GOTERM_BP_FAT	**cell part morphogenesis**	**10**	**11.2**	**5.20E-04**	**3.80E-02**
GOTERM_BP_FAT	**neuron development**	**10**	**11.2**	**7.40E-04**	**4.20E-02**
GOTERM_BP_FAT	**axonogenesis**	**9**	**10.1**	**7.50E-05**	**4.30E-02**
GOTERM_CC_FAT	**plasma membrane**	**16**	**18**	**5.20E-04**	**4.70E-02**
GOTERM_BP_FAT	**cell projection organization**	**10**	**11.2**	**1.10E-03**	**5.00E-02**
GOTERM_BP_FAT	**neuron projection morphogenesis**	**10**	**11.2**	**1.80E-04**	**5.00E-02**
GOTERM_BP_FAT	cell morphogenesis	11	12.4	1.10E-03	5.40E-02
GOTERM_BP_FAT	Wnt receptor signaling pathway	5	5.6	1.60E-03	7.00E-02

Table S5.

Genes represented in each significant GO-term in αβ neuron insertions.

cell morphogenesis involved in neuron differentiation; cell morphogenesis involved in differentiation; cell projection morphogenesis; neuron projection development; cell part morphogenesis; neuron development; cell projection organization; neuron projection morphogenesis		**transposon**
FBgn0005775	Connectin	FBgn0043969_diver
FBgn0011766	E2F transcription factor	FBgn0046110_Juan
FBgn0000464	Leukocyte-antigen-related-like	FBgn0000005_297
FBgn0041604	dally-like	FBgn0000481_Doc
FBgn0015380	derailed	FBgn0002698_mdg3
FBgn0024245	doughnut on 2	FBgn0000652_F-element
FBgn0000479	dunce	FBgn0043055_Ivk
FBgn0003301	rutabaga	FBgn0043055_Ivk
FBgn0003435	smooth	FBgn0000638_FB
FBgn0016977	split ends	FBgn0000199_blood

neuron differentiation		
FBgn0005775	Connectin	
FBgn0011766	E2F transcription factor	
FBgn0000464	Leukocyte-antigen-related-like	
FBgn0041604	dally-like	
FBgn0015380	derailed	
FBgn0024245	doughnut on 2	
FBgn0000479	dunce	
FBgn0001297	kayak	FBgn0010302_Burdock
FBgn0003301	rutabaga	
FBgn0003435	smooth	
FBgn0016977	split ends	

axonogenesis		
FBgn0005775	Connectin	
FBgn0000464	Leukocyte-antigen-related-like	
FBgn0041604	dally-like	
FBgn0015380	derailed	
FBgn0024245	doughnut on 2	
FBgn0000479	dunce	
FBgn0003301	rutabaga	
FBgn0003435	smooth	
FBgn0016977	split ends	

cell motion		
FBgn0000464	Leukocyte-antigen-related-like	
FBgn0086901	crossveinless c	FBgn0000004_17.6
FBgn0041604	dally-like	
FBgn0015380	derailed	
FBgn0024245	doughnut on 2	
FBgn0024236	fear-of-intimacy	FBgn0000155_roo
FBgn0250823	gilgamesh	FBgn0000005_297
FBgn0001297	kayak	
FBgn0003435	smooth	
FBgn0016977	split ends	

plasma membrane		
FBgn0005775	Connectin	
FBgn0033744	Dmel_CG12370	FBgn0000155_roo
FBgn0259214	Dmel_CG42314	FBgn0000155_roo
FBgn0000464	Leukocyte-antigen-related-like	
FBgn0259111	Na[+]-driven anion exchanger 1	FBgn0044355_Quasimodo
FBgn0033466	Peptidyl-alpha-hydroxyglycine alpha-amidating	FBgn0000155_roo
FBgn0259240	Tenascin accessory	FBgn0000199_blood
FBgn0087007	big bang	FBgn0002698_mdg3
FBgn0015380	derailed	
FBgn0024245	doughnut on 2	
FBgn0024236	fear-of-intimacy	
FBgn0250823	gilgamesh	
FBgn0004167	karst	FBgn0000155_roo
FBgn0028369	kin of irre	FBgn0000155_roo
FBgn0037212	nicotinic Acetylcholine Receptor alpha 80B	FBgn0000155_roo
FBgn0003301	rutabaga	

Table S6.
Genomic locations of new insertions in other brain cells.

FLYBASE_GENE_ID	Gene Name
FBgn0000479	dunce
FBgn0000578	enabled
FBgn0001269	invected
FBgn0003295	roughoid
FBgn0003415	skuld
FBgn0004168	Serotonin receptor 1A
FBgn0004198	cut
FBgn0005410	strawberry notch
FBgn0005633	flightin
FBgn0010247	Poly-(ADP-ribose) polymerase
FBgn0011206	boule
FBgn0011259	Semaphorin-1A
FBgn0011288	Synapse protein 25
FBgn0011829	Ret oncogene
FBgn0013467	igloo
FBgn0013997	Neurexin IV
FBgn0014343	mirror
FBgn0014455	Adenosylhomocysteinase at 13
FBgn0015561	unplugged
FBgn0015600	toucan
FBgn0020305	dribble
FBgn0023171	ribonuclease H1
FBgn0023511	Dmel_CG3810
FBgn0025726	Dmel_CG2999
FBgn0027539	Dmel_CG5807
FBgn0028866	Dmel_CG18420
FBgn0030073	Dmel_CG10962
FBgn0030898	Dmel_CG12611
FBgn0031195	Dmel_CG17600
FBgn0031730	Dmel_CG7236
FBgn0031902	Protein Wnt
FBgn0032120	Dmel_CG33298
FBgn0032211	Dmel_CG13138
FBgn0032362	Dmel_CG14928
FBgn0032471	Dmel_CG5122
FBgn0032627	Dmel_CG12621
FBgn0033524	Probable cytochrome P450 49a1
FBgn0035173	Dmel_CG13907
FBgn0035266	Dmel_CG7995
FBgn0035888	Dmel_CG7120
FBgn0036433	Dmel_CG9628
FBgn0036511	Dmel_CG6498
FBgn0037057	Dmel_CG10512
FBgn0037130	Syntrophin-like 1
FBgn0037215	Dmel_CG12582
FBgn0037427	Osiris 17
FBgn0037635	Dmel_CG9837
FBgn0038680	Probable cytochrome P450 12a5, mitochondrial
FBgn0038878	Dmel_CG3301
FBgn0038930	Dmel_CG5778
FBgn0039016	Dicer-1
FBgn0039494	Gram-positive Specific Serine protease
FBgn0039862	Dmel_CG1804
FBgn0039900	Dmel_CG2381
FBgn0050087	Dmel_CG30087
FBgn0050296	Dmel_CG30296
FBgn0050468	Dmel_CG30468
FBgn0051524	Dmel_CG31524
FBgn0052138	Formin-like protein CG32138
FBgn0052676	Dmel_CG32676
FBgn0052758	Dmel_CG32758
FBgn0053202	Dmel_CG33202
FBgn0053337	Dmel_CG33337
FBgn0053543	Dmel_CG33543
FBgn0053554	Transcription-associated protein 1
FBgn0053555	bitesize
FBgn0058006	Dmel_CG40006
FBgn0058178	Dmel_CG40178
FBgn0083950	Dmel_CG34114
FBgn0083975	Dmel_CG34139
FBgn0085399	Dmel_CG34370
FBgn0086604	Dmel_CG12484
FBgn0086675	found in neurons
FBgn0086758	Chronologically inappropriate morphogenesis
FBgn0259239	Dmel_CG42337
FBgn0259244	Dmel_CG42342
FBgn0260642	Antennapedia
FBgn0260941	Dmel_CG42318
FBgn0260943	RNA-binding protein Musashi homolog Rbp6
FBgn0260959	Microcephalin
FBgn0260987	Dmel_CG17436
FBgn0261549	
FBgn0262103	
FBgn0263102	
FBgn0263873	
FBgn0261509	
FBgn0261649	
FBgn0261698	
FBgn0262362	
FBgn0262738	
FBgn0261642	
FBgn0261914	
FBgn0263391	
FBgn0058232	

Table S7.

GO-term analysis of genes disrupted in other brain cells. No term reaches statistical significance.

Category	Term	Count	%	P-Value	Benjamini FDR
GOTERM_BP_FAT	synaptic vesicle transport	5	6.2	1.80E-03	6.40E-01
GOTERM_BP_FAT	synaptic vesicle exocytosis	4	4.9	3.70E-03	6.60E-01
GOTERM_CC_FAT	plasma membrane part	6	7.4	6.30E-02	7.00E-01
GOTERM_BP_FAT	exocytosis	4	4.9	6.70E-03	7.30E-01
GOTERM_CC_FAT	plasma membrane	9	11.1	4.90E-02	7.50E-01
GOTERM_BP_FAT	neurotransmitter transport	4	4.9	3.60E-02	7.60E-01
GOTERM_CC_FAT	coated vesicle	3	3.7	6.30E-02	7.70E-01
GOTERM_BP_FAT	compound eye cone cell fate commitment	2	2.5	5.10E-02	7.70E-01
GOTERM_BP_FAT	regulation of transcription, DNA-dependent	8	9.9	6.10E-02	7.70E-01
GOTERM_BP_FAT	wing disc pattern formation	3	3.7	5.90E-02	7.70E-01
GOTERM_BP_FAT	cyclic-nucleotide-mediated signaling	2	2.5	9.40E-02	7.70E-01
GOTERM_BP_FAT	cell morphogenesis involved in differentiation	5	6.2	9.70E-02	7.80E-01
GOTERM_BP_FAT	secretion	4	4.9	3.50E-02	7.80E-01
GOTERM_BP_FAT	regulation of RNA metabolic process	8	9.9	9.30E-02	7.80E-01
GOTERM_CC_FAT	clathrin-coated vesicle	3	3.7	4.00E-02	7.80E-01
GOTERM_BP_FAT	actin filament-based process	4	4.9	5.00E-02	7.80E-01
GOTERM_BP_FAT	secretion by cell	4	4.9	3.10E-02	7.80E-01
GOTERM_BP_FAT	regionalization	7	8.6	4.60E-02	7.80E-01
GOTERM_BP_FAT	cell-cell signaling	5	6.2	3.30E-02	7.80E-01
GOTERM_BP_FAT	cell morphogenesis	7	8.6	4.10E-02	7.80E-01
GOTERM_BP_FAT	open tracheal system development	4	4.9	5.80E-02	7.80E-01
GOTERM_BP_FAT	respiratory system development	4	4.9	5.80E-02	7.80E-01
GOTERM_BP_FAT	neurological system process	7	8.6	9.30E-02	7.90E-01
GOTERM_BP_FAT	cAMP-mediated signaling	2	2.5	8.90E-02	7.90E-01
GOTERM_BP_FAT	pattern specification process	7	8.6	5.70E-02	7.90E-01
GOTERM_BP_FAT	actin cytoskeleton organization	4	4.9	5.00E-02	7.90E-01
GOTERM_BP_FAT	peripheral nervous system development	3	3.7	9.20E-02	7.90E-01
GOTERM_BP_FAT	cell fate commitment	5	6.2	4.50E-02	7.90E-01

Table S8.

$\chi 2$ test reveals the distribution of de novo insertions in αβ neurons, other brain cells and ovary is statistically different from the positions of transposons annotated in the published genome.

	Transposons in the reference genome (measured)	αβ neurons	rest of brain	ovary–embryo
exon	258	35	42	823
intron	11110	68	52	1538
promoter	502	5	0	152
intergenic	33008	107	106	1347
total	44878	215	200	3860
CHISQ.TEST against measured		8.75E-181	2.63E-273	<2E-16

Table S9.

GO-term analysis of genes disrupted in ovary cells. 102 genes were randomly sampled from the ovary data to normalize numbers to those in the brain samples. No term reaches statistical significance.

Category	Term	Count	%	P-Value	Benjamini FDR
GOTERM_BP_FAT	intrinsic to membrane	15	16.5	2.90E-02	7.20E-01
GOTERM_BP_FAT	plasma membrane	10	11	4.40E-02	7.30E-01
GOTERM_CC_FAT	integral to membrane	15	16.5	2.50E-02	8.90E-01

Table S10.
Genes receiving de novo transposon insertions in αβ neurons are likely expressed in αβ neurons. Promoter fragments from 20 of the genes receiving transposon insertions in αβ have been analyzed for their ability to direct neuron specific gene expression as transgenic fly lines (49). 18 of the 20 confer expression in MB αβ neurons (black text). One does not label anything in the brain and the other labels α′β′ neurons (red text).

Gene	*FlyLight GAL4 transgenic line(s)*
dnc	R12C07
rut	R14H06, R15E09
mud	R29B12
drl	R26E12 α′β′ labeled
nAChRalpha 80B	R18G03, R18G05
gprk	
cngl	R89D07
GABA B-R1	R74G01, R73G03
E2F	R47B08, R47B10, R48C03
kayak	R42A04
spen	
smooth	R15G02 no brain expression
gilgamesh	R27E01, R27E03, R28C09, R28C10
fear of intimacy	
doughnut on 2	
dally-Like (dlp)	R53H05, R52H12, R54B01
cv-c	
Lar	
Ndae1	
kin of irre	
karst	
big bang	
Ten-A	
Pal1	
connectin	
CG12370	R19A04
CG42314	
Non-GOterm clustered hits	
beadex	R41C08, R42C07
SNF4gamma	R30H03, R33A06
Sbf	
RhoGAP68F	
limpet	R27C02
ccn	R53C04, R53C09
norpA	R12B05, R14A07, R14D06, R14H01
tlk	
mur2B	
HDAC4	
kek5	
Timeout	
shepard	
nahoda	
bab2	R36E08
uex	
orthopedia	R83E01, R83E02
quasimodo	
Smi35A	
scp1	

References

1. R. Cordaux, M. A. Batzer, *Nat Rev Genet* **10**, 691 (2009).
2. H. H. J. Kazazian, *Science* **303**, 1626 (2004).
3. J. S. Kaminker *et al.*, *Genome Biol* **3**, RESEARCH0084 (2002).
4. N. G. Coufal *et al.*, *Nature* **460**, 1127 (2009).
5. J. K. Baillie *et al.*, *Nature* **479**, 534 (2011).
6. A. R. Muotri *et al.*, *Nature* **435**, 903 (2005).
7. A. R. Muotri, F. H. Gage, *Nature* **441**, 1087 (2006).
8. T. Singer, M. J. McConnell, M. C. Marchetto, N. G. Coufal, F. H. Gage, *TrendsNeurosci* **33**, 345 (2010).
9. F. H. Gage, A. R. Muotri, *Sci Am* **306**, 26 (2012).
10. A. L. Blum, W. Li, M. Cressy, J. Dubnau, *Curr Biol* **19**, 1341 (2009).
11. S. Trannoy, C. Redt-Clouet, J. M. Dura, T. Preat, *Curr Biol* **21**, 1647 (2011).
12. M. J. Krashes, A. C. Keene, B. Leung, J. D. Armstrong, S. Waddell, *Neuron* **53**, 103 (2007).
13. D. Yu, D. B. Akalal, R. L. Davis, *Neuron* **52**, 845 (2006).
14. E. Nagoshi *et al.*, *Nat Neurosci* **13**, 60 (2010).
15. Y. Shang, L. C. Griffith, M. Rosbash, *Proc Natl Acad Sci U S A* **105**, 19587 (2008).
16. J. Rainer, F. Sanchez-Cabo, G. Stocker, A. Sturn, Z. Trajanoski, *Nucleic Acids Res* **34**, W498 (2006).
17. J. L. Jakubczak, W. D. Burke, T. H. Eickbush, *Proc Natl Acad Sci U S A* **88**, 3295 (1991).
18. A. R. Muotri *et al.*, *Nature* **468**, 443 (2010).
19. C. Li *et al.*, *Cell* **137**, 509 (2009).
20. V. V. Vagin *et al.*, *Science* **313**, 320 (2006).
21. M. C. Siomi, K. Sato, D. Pezic, A. A. Aravin, *Nat Rev Mol Cell Biol* **12**, 246 (2011).
22. M. Ghildiyal *et al.*, *Science* **320**, 1077 (2008).
23. K. Saito, M. C. Siomi, *Dev Cell* **19**, 687 (2010).
24. J. Brennecke *et al.*, *Science* **322**, 1387 (2008).
25. N. Yang, H. H. J. Kazazian, *Nat Struct Mol Biol* **13**, 763 (2006).
26. M. Ghildiyal, P. D. Zamore, *Nat Rev Genet* **10**, 94 (2009).
27. A. A. Aravin *et al.*, *Curr Biol* **11**, 1017 (2001).
28. A. A. Aravin *et al.*, *Mol Cell Biol* **24**, 6742 (2004).
29. M. P. Bozzetti *et al.*, *Proc Natl Acad Sci U S A* **92**, 6067 (1995).
30. J. Brennecke *et al.*, *Cell* **128**, 1089 (2007).
31. L. S. Gunawardane *et al.*, *Science* **315**, 1587 (2007).
32. D. N. Cox *et al.*, *Genes Dev* **12**, 3715 (1998).
33. T. Schupbach, E. Wieschaus, *Genetics* **129**, 1119 (1991).
34. Y. Tomari *et al.*, *Cell* **116**, 831 (2004).
35. J. S. Khurana *et al.*, *Cell* **147**, 1551 (2011).
36. G. D. Evrony et al., *Cell* **151**, 483 (2012).
37. S. M. Tomchik, R. L. Davis, *Neuron* **64**, 510 (2009).
38. Y. Tan, D. Yu, J. Pletting, R. L. Davis, *Neuron* **67**, 810 (2010).

39. C. Moreau-Fauvarque, E. Taillebourg, E. Boissoneau, J. Mesnard, J. M. Dura, *Mech Dev* **78**, 47 (1998).
40. A. Hovhanyan, T. Raabe, *J Neurogenet* **23**, 42 (2009).
41. K. Yasuyama, I. A. Meinertzhagen, F. W. Schurmann, *J Comp Neurol* **445**, 211 (2002).
42. X. Liu, R. L. Davis, *Nat Neurosci* **12**, 53 (2009).
43. S. Waddell, *Trends Neurosci* **33**, 457 (2010).
44. P. Rajasethupathy *et al.*, *Cell* **149**, 693 (2012).
45. J. L. Pitman *et al.*, *Curr Biol* **21**, 855 (2011).
46. P. Stockinger, D. Kvitsiani, S. Rotkopf, L. Tirian, B. J. Dickson, *Cell* **121**, 795 (2005).
47. D. Schmidt *et al.*, *Methods* **48**, 240 (2009).
48. H. Li, R. Durbin, *Bioinformatics* **25**, 1754 (2009).
49. A. Jenett *et al.*, *Cell Rep* 2, 991 (2012).