

Published in final edited form as:

Science. 2013 April 5; 340(6128): . doi:10.1126/science.1231965.

Transposition driven genomic heterogeneity in the *Drosophila* brain*

Paola N. Perrat¹, Shamik DasGupta^{1,5}, Jie Wang², William Theurkauf³, Zhiping Weng², Michael Rosbash⁴, and Scott Waddell^{1,5,†}

¹Department of Neurobiology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605, USA

²Program in Bioinformatics and Integrative Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605, USA

³Program in Molecular Medicine, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605, USA

⁴Howard Hughes Medical Institute and National Center for Behavioral Genomics, Department of Biology, Brandeis University, Waltham, MA 02454, USA

⁵Center for Neural Circuits and Behaviour, University of Oxford, Mansfield Road, Oxford, OX1 3TA, UK

Abstract

Recent studies in mammals have documented neural expression and mobility of retrotransposons and suggested that neural genomes are diverse mosaics. Here we report transposition in memory-relevant neurons in the *Drosophila* brain. Cell-type specific gene expression profiling revealed transposon expression is more abundant in mushroom body (MB) $\alpha\beta$ neurons than in neighboring MB neurons. The PIWI-RNA (piRNA) proteins Aubergine and Argonaute 3, known to suppress transposons in the fly germline, are expressed in the brain and appear less abundant in $\alpha\beta$ MB neurons. Loss of piRNA proteins correlates with elevated transposon expression in brain. Lastly, paired-end deep sequencing identified over 200 *de novo* transposon insertions in $\alpha\beta$ neurons, including events into memory-relevant loci. Our observations indicate that genomic heterogeneity is a conserved feature of the brain.

Transposons comprise nearly 45% and 15-20% of the human and fly genome, respectively (1-3). Mobilized transposons can act as insertional mutagens and create lesions where they once resided (2). Recombination between homologous transposons can also delete intervening loci. Specific regions of the mammalian brain, most evidently the hippocampus, might be particularly predisposed to transposition (4, 5). LINE-1 (L1) retrotransposons mobilized during differentiation appear to insert in the open chromatin of neurally expressed genes (4-6). One such insertion in neural progenitor cells altered the expression of the receiving gene, and subsequent maturation of these cells into neurons (6). The mosaic nature of transposition could therefore provide additional neural diversity that might contribute towards behavioral individuality in addition to contributing towards neurological disorders (7-9).

*This manuscript has been accepted for publication in *Science*. This version has not undergone final editing. Please refer to the complete version of record at <http://www.sciencemag.org/>. The manuscript may not be reproduced or used in any manner that does not fall within the fair use provisions of the Copyright Act without the prior, written permission of AAAS.

†Correspondence: scott.waddell@cncb.ox.ac.uk.

The *Drosophila melanogaster* mushroom bodies (MBs) are brain structures critical for olfactory memory. The approximately 2000 intrinsic MB neurons are divisible into $\alpha'\beta'$, γ and $\alpha\beta$ by morphology and roles in memory processing (10-13). In this study, we used cell-type specific gene expression profiling (14) to gain insight into cellular properties of MB neurons. Intersectional genetics (15) (Fig. 1A) allowed us to exclusively label MB $\alpha'\beta'$, γ , and $\alpha\beta$ neurons in the brain with green fluorescent protein, GFP (Fig. 1B-D). We also assayed for comparison a genotype where GFP labels most neurons in the brain, except MB neurons - the no MB group, (Fig. 1E). Sixty brains per genotype were dissected from the head capsule and dissociated by proteolysis and agitation before GFP-expressing single cell bodies were collected using Fluorescence Activated Cell Sorting (FACS). Total RNA was isolated from 10000 cells per genotype and polyA RNA was amplified and hybridized to Affymetrix *Drosophila* 2.0 genome expression arrays. Each genotype was processed in four independent replicates (Fig. 2A).

Routine statistical analysis for differentially expressed genes, including a multiple-testing correction across all sixteen datasets, did not reveal significant differences at a false discovery rate (FDR) <0.05 . We therefore used CARMAweb (16) to identify 146 mRNAs whose average signal was 7 in $\alpha\beta$ neurons and that were also 2-fold higher than in $\alpha'\beta'$. 29 of the top 60 transcripts from this list, that were significantly different from $\alpha'\beta'$ signals by raw *P*-value, represent transposons (Fig. 2A and Table S1). Alignment of the corresponding values from the γ and no MB profiles showed a similarly significant bias in transposon expression over these samples. We identified retrotransposons that transpose via a replicative mechanism involving an RNA intermediate and DNA elements that use non-replicative excision and repair. Retrotransposons can be subdivided into long-terminal repeat (LTR) and Long Interspersed Nuclear Elements (LINEs). We found eleven LTR; *Tabor*, *mdg1*, *roo*, *qbert*, *gypsy*, *invader3*, *gypsy2*, *microcopia*, *412*, *accord* and *blood*, twelve LINE-like; *G6*, *Rt1b*, *HeT-A*, *Ivk*, *Cr1a*, *F* element, *Doc2*, *baggins*, *R2*, *Doc3* and *Doc* and four DNA elements; *Baril*, *pogo*, *Tc3* and *transib3*.

We further analyzed fourteen transposons, representing the most abundant in each class. Quantitative RT-PCR (QRT-PCR) of RNA from independently purified cell samples confirmed transposon expression was significantly higher in $\alpha\beta$ than other MB neurons (Fig. 2B). All transposons, other than R2, were also significantly higher in $\alpha\beta$ than the rest of the brain. R2 is unique, because it exclusively inserts in the highly repeated 28S ribosomal RNA locus and heterochromatin (17).

Transposition is ordinarily regulated by chromatin structure and post-transcriptional degradation of transposon mRNA using complementary RNAs (18-25). The small interfering RNA (siRNA) pathway has been implicated in somatic cells (26). In contrast, the PIWI-associated RNA (piRNA) pathway has a more established role in the germline (19-21). The microarray analysis skewed our attention towards piRNA because the translocated *Stellate* locus, *Stellate12D orphon* (*Ste12DOR*) mRNA, was >20 -fold higher in $\alpha\beta$ than other MB neurons and the rest of the brain (Fig. 2A and Table S1). *Stellate* repeat transcripts are usually curtailed by piRNA, not siRNA (27, 28). *Stellate* repeats encode a Casein Kinase II regulatory subunit and piRNA mutant flies form *Stellate* protein crystals in testis (29). Immunostaining *Stellate* in the brain labeled puncta within $\alpha\beta$ dendrites in the MB calyx, consistent with high *Ste12DOR* expression in wild-type $\alpha\beta$ neurons (Fig. S1).

piRNAs are loaded into PIWI clade argonaute proteins Piwi, Aubergine (Aub) and Argonaute 3 (Ago3) (21). Piwi and Aub can amplify piRNA pools with Ago3 (30, 31). To investigate piRNA involvement in differential transposon expression we immunolocalized PIWI proteins and co-localized CD8::GFP to assign signals to MB neuron type (Fig. 3). Aub and Ago3 differentially labeled MB subdivision in addition to structures throughout the

brain (Figs. 3, S2 and S3) but we did not detect Piwi (32). The ellipsoid body of the central complex, stained strongly for Aub but not at all for Ago3 (Figs. 3A, 3E, S2D and S3D) suggesting possible functional exclusivity of PIWI proteins in the brain.

Differential Aub and Ago3 label is most evident within axon bundles in the peduncle and lobes, where MB neuron types are anatomically discrete (Figs. 3A-H and S2). Aub protein co-localized with γ and $\alpha'\beta'$ in the peduncle and lobes but was reduced in $\alpha\beta$ in both locations (Figs. 3B-D and S2). Ago3 did not label MB lobes (Fig. S3) but co-localized with γ in the peduncle (Fig. 3H). Ago3 labeled core $\alpha\beta$ ($\alpha\beta_c$) neurons but did not label outer $\alpha\beta$ (Fig. 3F). Therefore outer $\alpha\beta$ neurons do not abundantly express Aub or Ago3 suggesting transposon suppression is relaxed. In contrast, γ express Aub and Ago3, providing potential for piRNA amplification, and $\alpha'\beta'$ express Aub. These patterns of Aub and Ago3 in the MB peduncle appear conserved in brains from *D. erecta*, *D. sechellia* and the more distantly related *D. pseudoobscura* species (Fig. S4).

Loss of siRNA function elevates transposon expression in the head (22). We replicated these findings using *ago2*⁴¹⁴ and *dcr-2*^{L811fsX} mutant flies (Fig. S5). In parallel we tested whether piRNA suppressed transposon expression using transheterozygous *aub* (*aub*^{HN2}/*aub*^{QC42}) and *ago3* (*ago3*^{t2}/*ago3*^{t3}) (19, 33) and transheterozygous *armitage* (*armi*¹/*armi*^{72.1}) heads (34) (Fig. 3I). Levels of the 14 LTR, LINE-like and TIR group transposons verified to be expressed in $\alpha\beta$ neurons (Fig. 2B) were assayed by QRT-PCR. 13 of the 14 transposons were significantly elevated in siRNA defective *ago2* and *dcr-2* mutants (Fig. S5). The piRNA defective *aub*, *ago3* and *armi* mutants also exhibited significantly elevated levels of 9 of the 14 elements (Fig. 3I). LTR elements *gypsy*, *Tabor* and *qbert*, LINE-like *HeT-A*, *RT1B* and *R2* and the TIR element *pogo* were higher in *ago3* mutants. In addition *blood*, *Tabor* and *R2* were elevated in *aub* flies and *blood*, *gypsy*, *Tabor*, *invader3*, *qbert*, *HeT-A* and *R2* in *armi* mutants. Therefore, the piRNA pathway contributes to transposon silencing in the brain and low level Aub and Ago3 may permit expression in $\alpha\beta$ neurons.

To determine whether transposons are mobile we mapped new insertions by deep sequencing (35) $\alpha\beta$ DNA (Fig. 4). We purified $\alpha\beta$ neurons by FACS, as for transcriptome analysis, but isolated genomic DNA. Insertions were defined by paired-end reads in which one end mapped to the annotated genome and the other to transposon sequence. To identify *de novo* transposition events in $\alpha\beta$ neurons we compared the genomic position of transposons within $\alpha\beta$ sequence to those located by sequencing DNA from genetically identical embryos. In addition we sequenced DNA from the remainder of the brain tissue from the FACS separation of $\alpha\beta$ neurons (Fig. 4A and Table S4).

These studies identified 3890 transposon insertions in embryo DNA that differed from the published *Drosophila* genome sequence (Fig. 4A). In comparison $\alpha\beta$ neuron DNA revealed 215 additional sites. The remaining brain tissue uncovered 200 new insertions, including 19 identical to those in $\alpha\beta$ neurons. The $\alpha\beta$ and other brain insertions are likely *de novo* because the sequencing depth was ten times higher for embryos (34.1 fold) than neurons (3.1 fold), due to the ease of collecting material. By randomly sampling reads to yield 1-fold genome coverage, we calculated 129 new transposon insertions per $\alpha\beta$ neuron genome (Table S2). Sequencing single neurons (36) would reveal the exact cellular frequency and heterogeneity of transposition events.

New $\alpha\beta$ insertions occurred across all chromosomes without obvious regional bias (Fig. 4B). In addition, insertions resulted from 49 different transposons representing LTR, LINE-like, TIR and *Foldback* (FB) classes (Fig. 4C). They included 11 of the 27 transposons in the $\alpha\beta$ transcriptome (Fig. 2A) and the number of insertions per class is consistent with their prevalence in the genome (Fig. 4D). Therefore, many transposons mobilize in $\alpha\beta$ neurons.

108 of the 215 *de novo* $\alpha\beta$ insertions mapped close to identified genes (Fig. 4E and S6). Of these, 35 disrupted exons, 68 introns and 5 fell in promoter regions (<1kb from transcription start site). The remaining 107 insertions located to piRNA clusters or intergenic regions and were not assigned to a particular gene. A similar distribution was observed for the 200 new insertions in the rest of the brain (Fig. 4E and S6). The reference fly genome has 258, 11110, 502 and 33008 transposon insertions in exonic, intronic, promoter and intergenic regions respectively. Therefore both groups of brain cells had a significantly larger fraction of *de novo* insertions within exons and fewer in intergenic regions than transposons annotated in the genome (Fig. 4E and Table S8). To test whether such a distribution was unique to neurons we analyzed *de novo* insertions in ovary, again using embryo sequence as the comparison. New insertions in ovary DNA revealed a similar skew towards exons (Fig. 4E and Table S8).

In mammals, active L1 elements appear to disrupt neurally expressed genes (4-6). New $\alpha\beta$ neuron insertions, but not those in other tissue (Tables S7 and S9), were significantly enriched in 12 Gene Ontology (GO) terms (Benjamini FDR <5%; Tables S4 and S5) all of which are related to neural functions. Moreover, promoter regions from 18 of 20 of the targeted genes drive expression in $\alpha\beta$ neurons (Table S10). We found exonic insertions in *gilgamesh*, *derailed* and *mushroom body defect* and intronic insertions in *dunce* and *rutabaga* (Table S3), all of which have established roles in MB development and function (37-40). In addition, MB neurons are principally driven by cholinergic olfactory projection neurons (41) and receive broad GABA-ergic inhibition (42) and dopaminergic modulation through G-protein coupled receptors (43). We identified intronic insertions in *nicotinic Acetylcholine Receptor α 80B*, *G protein-coupled receptor kinase 1* and *cyclic nucleotide gated channel-like* and an exonic insertion in *GABA-B-receptor subtype 1* (Table S3). Transposon-induced mosaicism could therefore alter integrative and plastic properties of individual MB $\alpha\beta$ neurons.

In conclusion, our data establish that transposon-mediated genomic heterogeneity is a feature of the fly brain and possibly other tissues. Taken with prior work in rodents and humans (4-6), it suggests genetic mosaicism may be a conserved characteristic of certain neurons. Work in mammals indicates L1 expression occurs because the L1-promoter is released during neurogenesis (6, 21). Our data are consistent with such a model and also that transposons avoid post-transcriptional piRNA silencing in adult $\alpha\beta$ neurons.

A recent study described a role for piRNA in epigenetic control of memory-related gene expression in *Aplysia* neurons (44). It is therefore possible that MB neurons differentially utilize piRNA to control memory-relevant gene expression, and transposon mobilization is an associated cost. Since we found transposon expression in $\alpha\beta$ neurons of adult flies, it is conceivable that disruptive insertions accumulate throughout life leading to neural decline and cognitive dysfunction.

Alternatively, permitting transposition may confer unique properties across the 1000 neurons in the $\alpha\beta$ ensemble and potentially produce behavioral variability between individual flies in the population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Data are available through ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) and NCBI Short Read (www.ncbi.nlm.nih.gov/sites/sra) archives as E-MEXP-3798 and SRP017718. We thank Anand Vodala, Jerome

Menet, Kate Abruzzi and Noreen Francis for advice with techniques. S.W. is funded by a Wellcome Trust Senior Research Fellowship in Basic Biomedical Sciences, The Gatsby Charitable Foundation, Oxford Martin School and grants MH069883 and MH081982 from NIH. M.R. is funded by NS044232 and NS045713 from NIH and the Ellison Medical Foundation. Z.W. and W.T. are funded by R01HD049116 from NIH.

References and Notes

1. Cordaux R, Batzer MA. *Nat Rev Genet.* 2009; 10:691. [PubMed: 19763152]
2. Kazazian HHJ. *Science.* 2004; 303:1626. [PubMed: 15016989]
3. Kaminker JS, et al. *Genome Biol.* 2002; 3 RESEARCH0084.
4. Coufal NG, et al. *Nature.* 2009; 460:1127. [PubMed: 19657334]
5. Baillie JK, et al. *Nature.* 2011; 479:534. [PubMed: 22037309]
6. Muotri AR, et al. *Nature.* 2005; 435:903. [PubMed: 15959507]
7. Muotri AR, Gage FH. *Nature.* 2006; 441:1087. [PubMed: 16810244]
8. Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. *Trends Neurosci.* 2010; 33:345. [PubMed: 20471112]
9. Gage FH, Muotri AR. *Sci Am.* 2012; 306:26. [PubMed: 22375319]
10. Blum AL, Li W, Cressy M, Dubnau J. *Curr Biol.* 2009; 19:1341. [PubMed: 19646879]
11. Trannoy S, Redt-Clouet C, Dura JM, Preat T. *Curr Biol.* 2011; 21:1647. [PubMed: 21962716]
12. Krashes MJ, Keene AC, Leung B, Armstrong JD, Waddell S. *Neuron.* 2007; 53:103. [PubMed: 17196534]
13. Yu D, Akalal DB, Davis RL. *Neuron.* 2006; 52:845. [PubMed: 17145505]
14. Nagoshi E, et al. *Nat Neurosci.* 2010; 13:60. [PubMed: 19966839]
15. Shang Y, Griffith LC, Rosbash M. *Proc Natl Acad Sci U S A.* 2008; 105:19587. [PubMed: 19060186]
16. Rainer J, Sanchez-Cabo F, Stocker G, Sturn A, Trajanoski Z. *Nucleic Acids Res.* 2006; 34:W498. [PubMed: 16845058]
17. Jakubczak JL, Burke WD, Eickbush TH. *Proc Natl Acad Sci U S A.* 1991; 88:3295. [PubMed: 1849649]
18. Muotri AR, et al. *Nature.* 2010; 468:443. [PubMed: 21085180]
19. Li C, et al. *Cell.* 2009; 137:509. [PubMed: 19395009]
20. Vagin VV, et al. *Science.* 2006; 313:320. [PubMed: 16809489]
21. Siomi MC, Sato K, Pezic D, Aravin AA. *Nat Rev Mol Cell Biol.* 2011; 12:246. [PubMed: 21427766]
22. Ghildiyal M, et al. *Science.* 2008; 320:1077. [PubMed: 18403677]
23. Saito K, Siomi MC. *Dev Cell.* 2010; 19:687. [PubMed: 21074719]
24. Brennecke J, et al. *Science.* 2008; 322:1387. [PubMed: 19039138]
25. Yang N, Kazazian HHJ. *Nat Struct Mol Biol.* 2006; 13:763. [PubMed: 16936727]
26. Ghildiyal M, Zamore PD. *Nat Rev Genet.* 2009; 10:94. [PubMed: 19148191]
27. Aravin AA, et al. *Curr Biol.* 2001; 11:1017. [PubMed: 11470406]
28. Aravin AA, et al. *Mol Cell Biol.* 2004; 24:6742. [PubMed: 15254241]
29. Bozzetti MP, et al. *Proc Natl Acad Sci U S A.* 1995; 92:6067. [PubMed: 7597082]
30. Brennecke J, et al. *Cell.* 2007; 128:1089. [PubMed: 17346786]
31. Gunawardane LS, et al. *Science.* 2007; 315:1587. [PubMed: 17322028]
32. Cox DN, et al. *Genes Dev.* 1998; 12:3715. [PubMed: 9851978]
33. Schupbach T, Wieschaus E. *Genetics.* 1991; 129:1119. [PubMed: 1783295]
34. Tomari Y, et al. *Cell.* 2004; 116:831. [PubMed: 15035985]
35. Khurana JS, et al. *Cell.* 2011; 147:1551. [PubMed: 22196730]
36. Evrony GD, et al. *Cell.* 2012; 151:483. [PubMed: 23101622]
37. Tomchik SM, Davis RL. *Neuron.* 2009; 64:510. [PubMed: 19945393]
38. Tan Y, Yu D, Pletting J, Davis RL. *Neuron.* 2010; 67:810. [PubMed: 20826312]

39. Moreau-Fauvarque C, Taillebourg E, Boissoneau E, Mesnard J, Dura JM. *Mech Dev.* 1998; 78:47. [PubMed: 9858681]
40. Hovhanyan A, Raabe T. *J Neurogenet.* 2009; 23:42. [PubMed: 19107630]
41. Yasuyama K, Meinertzhagen IA, Schurmann FW. *J Comp Neurol.* 2002; 445:211. [PubMed: 11920702]
42. Liu X, Davis RL. *Nat Neurosci.* 2009; 12:53. [PubMed: 19043409]
43. Waddell S. *Trends Neurosci.* 2010; 33:457. [PubMed: 20701984]
44. Rajasethupathy P, et al. *Cell.* 2012; 149:693. [PubMed: 22541438]

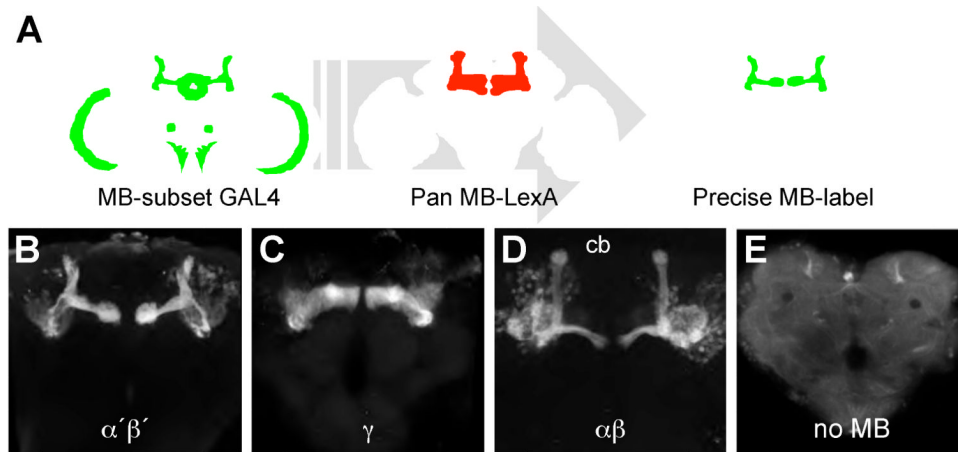


Fig. 1. Exclusive labeling of MB neuron subsets in the fly brain.

(A) MB-expressing GAL4s with expression elsewhere in the brain were intersected with an almost pan MB-expressing LexA. See Supporting Online Material for detail of genetic approach. Confocal projection of individual brains from flies labeling (B) $\alpha'\beta'$ (C) γ and (D) $\alpha\beta$ neurons. (E) all neurons in the brain, except MB neurons, no MB. Scale bar 40 μm .

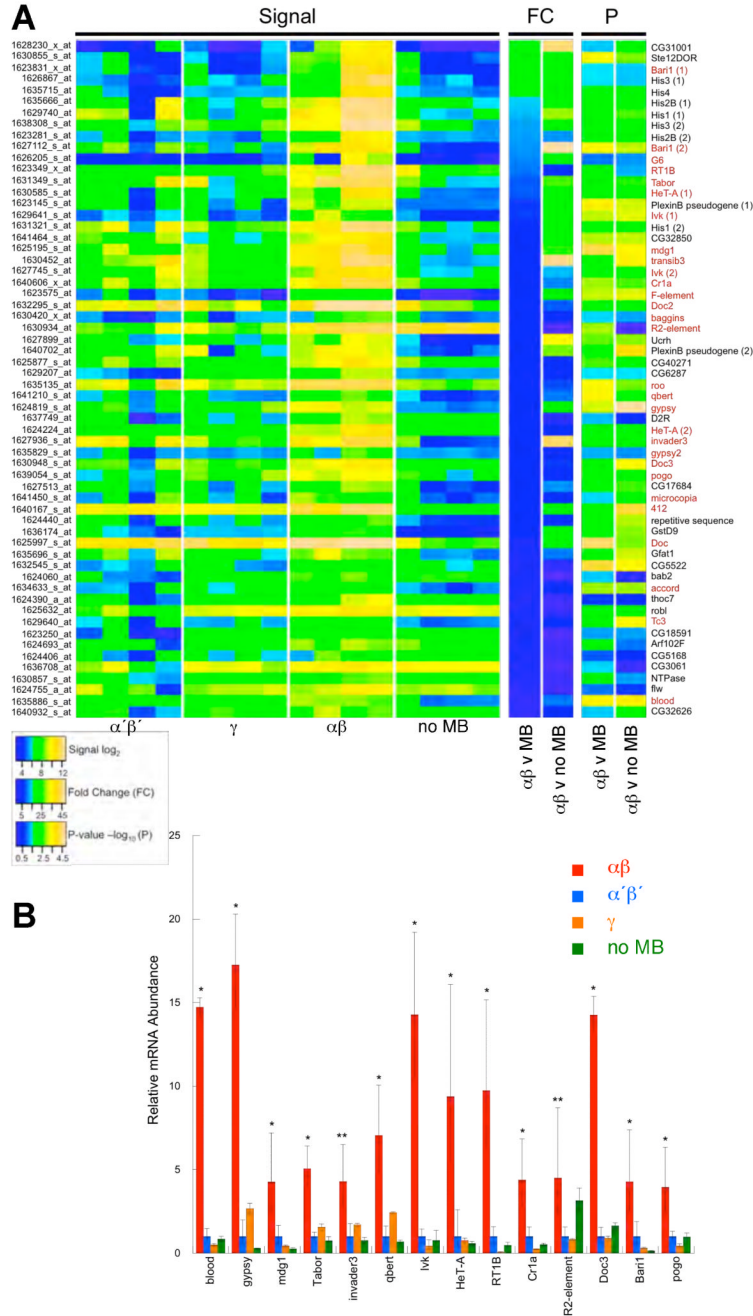


Fig. 2. Gene expression profiling MB neurons.

(A) Microarray data emphasizing elevated expression in $\alpha\beta$ neurons. Each signal column per category represents an independent replicate. Fold change (FC) and *t*-test *P* values are shown for $\alpha\beta$ versus other MB neurons (average value for $\alpha'\beta'$ and γ) and $\alpha\beta$ versus the no MB group. Transposons denoted in red. Scale bars linear for fold change, $-\log_{10}$ for *P*-value and \log_2 of the signal level. (B) QRT-PCR validation of increased transposon mRNA levels in $\alpha\beta$ neurons. * denotes $\alpha\beta$ signal significantly different to all other populations, all *P* 0.05, T-Test. ** denotes $\alpha\beta$ signal significantly different to MB but not no MB group for R2, and $\alpha\beta$ significantly different to all except γ for *invader3*.

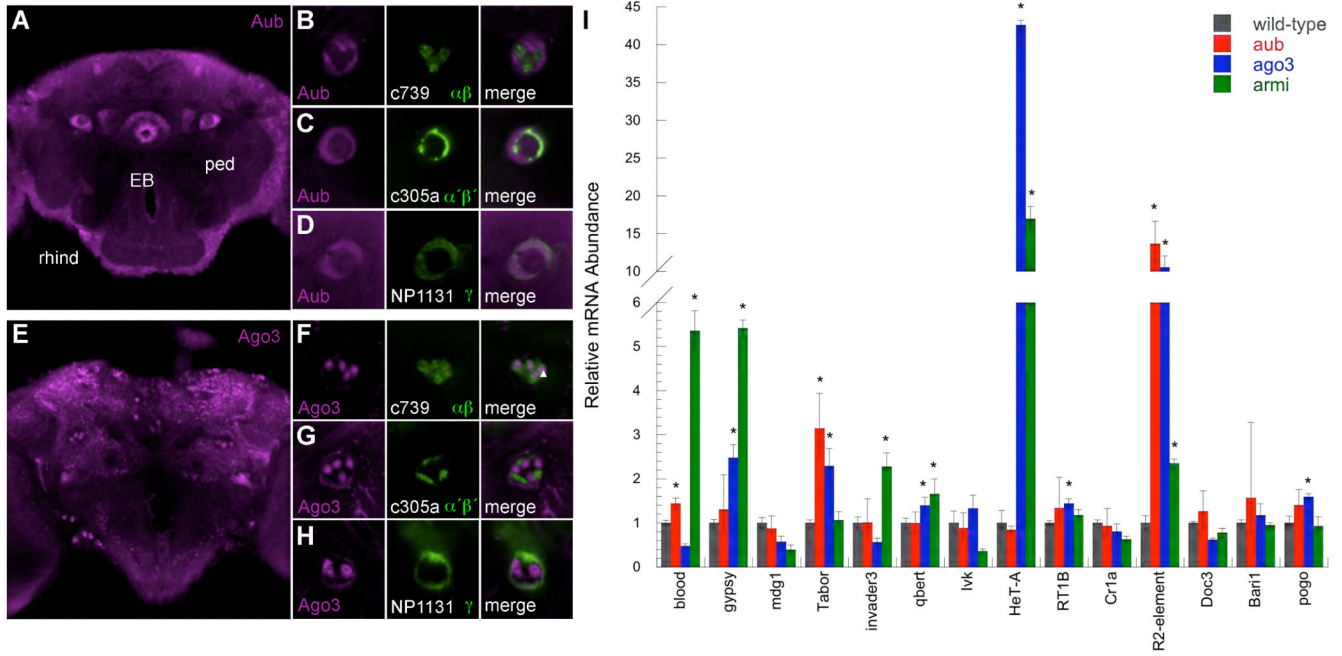


Fig. 3. Aub and Ago3 are not abundant in $\alpha\beta$ neurons.

(A) Aub immunostaining (magenta) labels the ellipsoid body (EB) and MB subdivision in the peduncle (ped). Single confocal section at the level of the MB peduncle. Dotted box denotes area in panels (B-D). (B) Aub labels $\alpha'\beta'$ and γ neurons in the peduncle more than $\alpha\beta$. Aub staining is mutually exclusive to GFP (green) expressing $\alpha\beta$ neurons but overlaps with (C) $\alpha'\beta'$ and (D) γ neurons. (B-D) left panels Aub staining; middle GFP; right panel merge. (E) Ago3 immunostaining (magenta) labels neurons throughout the brain and MB subdivision in the peduncle. Single confocal section at the level of the MB peduncle. Dotted box denotes area in panels (F-H). (F) Ago3 staining is prominent in the $\alpha\beta$ core (closed triangle) and does not overlap with outer GFP labeled $\alpha\beta$ neurons (open triangle) nor (G) $\alpha'\beta'$ but overlaps with (H) γ neurons. Left panels Ago3 staining; middle GFP; right panel merge. Scale bar 10 μ m. (I) Several transposon transcripts are elevated in *ago3*, *aub* and *armi* mutant fly brains. QRT-PCR analysis from wild-type, *aub*^{HN2}/*aub*^{QC42}, *ago3*^{t2}/*ago3*^{t3} and *armi*¹/*armi*^{27.1} mutant heads. Values normalized to wild-type heads. * denotes significant increase, *P*<0.05 (T-test).

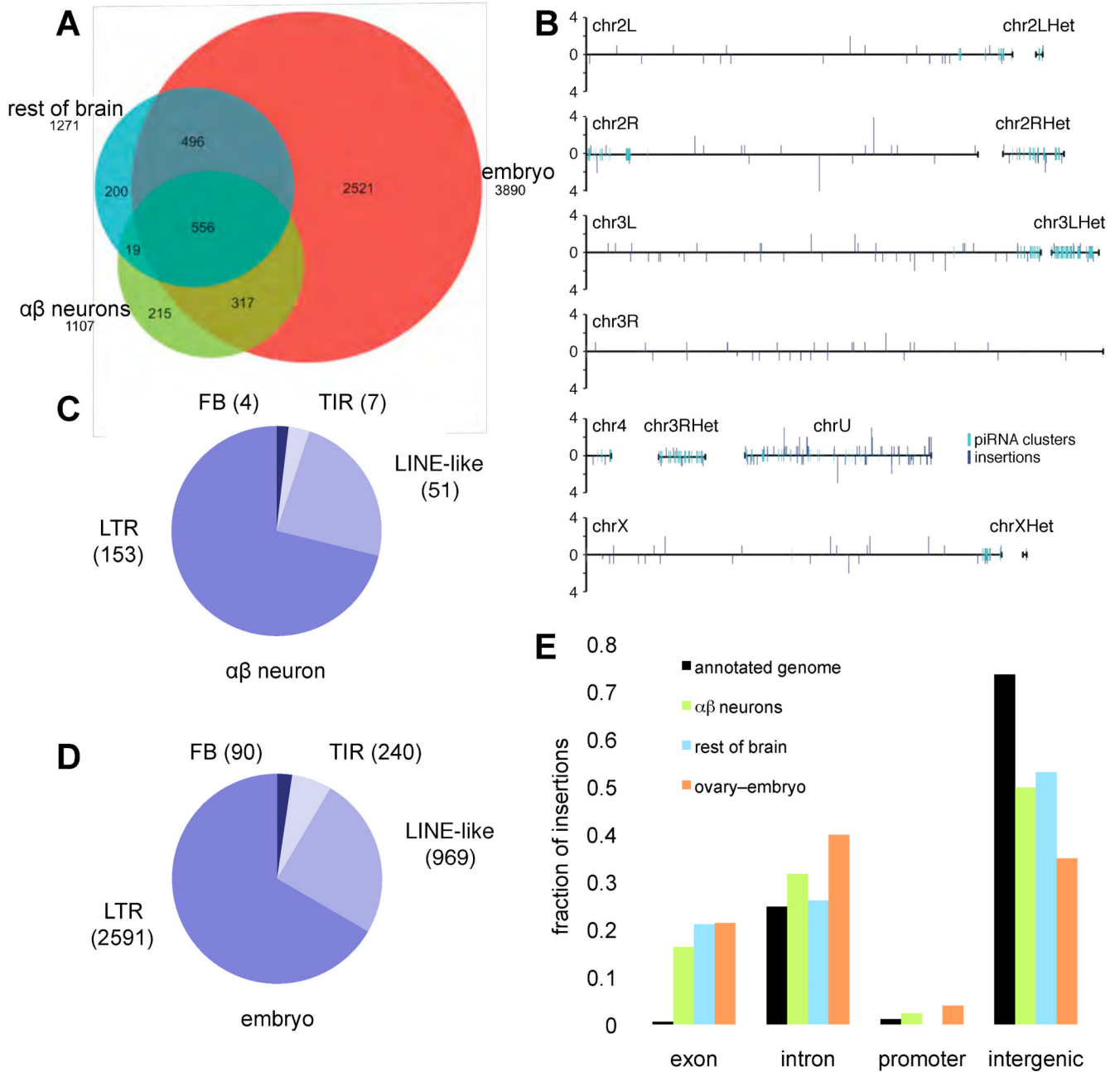


Fig. 4. Identification of *de novo* transposition events in $\alpha\beta$ neurons.

(A) Venn diagram of transposon insertions identified in this study. Direct genome sequencing identified 215 insertion sites in $\alpha\beta$ neuron DNA and 200 in DNA from other brain cells that were not present in embryo DNA from genetically identical sibling flies. (B) Chromosomal distribution of new transposon insertions in $\alpha\beta$ neurons. Bar height indicates number of insertions in each location. piRNA clusters are shown. (C) The proportion of new insertions found for each transposon class in $\alpha\beta$ neurons. (D) Proportion of insertions per transposon class in the inherited fly genome from our sequencing of embryo DNA. (E) Distribution of transposons in the annotated genome significantly differs from *de novo* insertions in $\alpha\beta$ neurons, the rest of brain and ovary DNA with respect to neighboring genes (p -value $< 2.2E-16$ for $\alpha\beta$ neurons, the rest of brain and ovary; χ^2 test).