

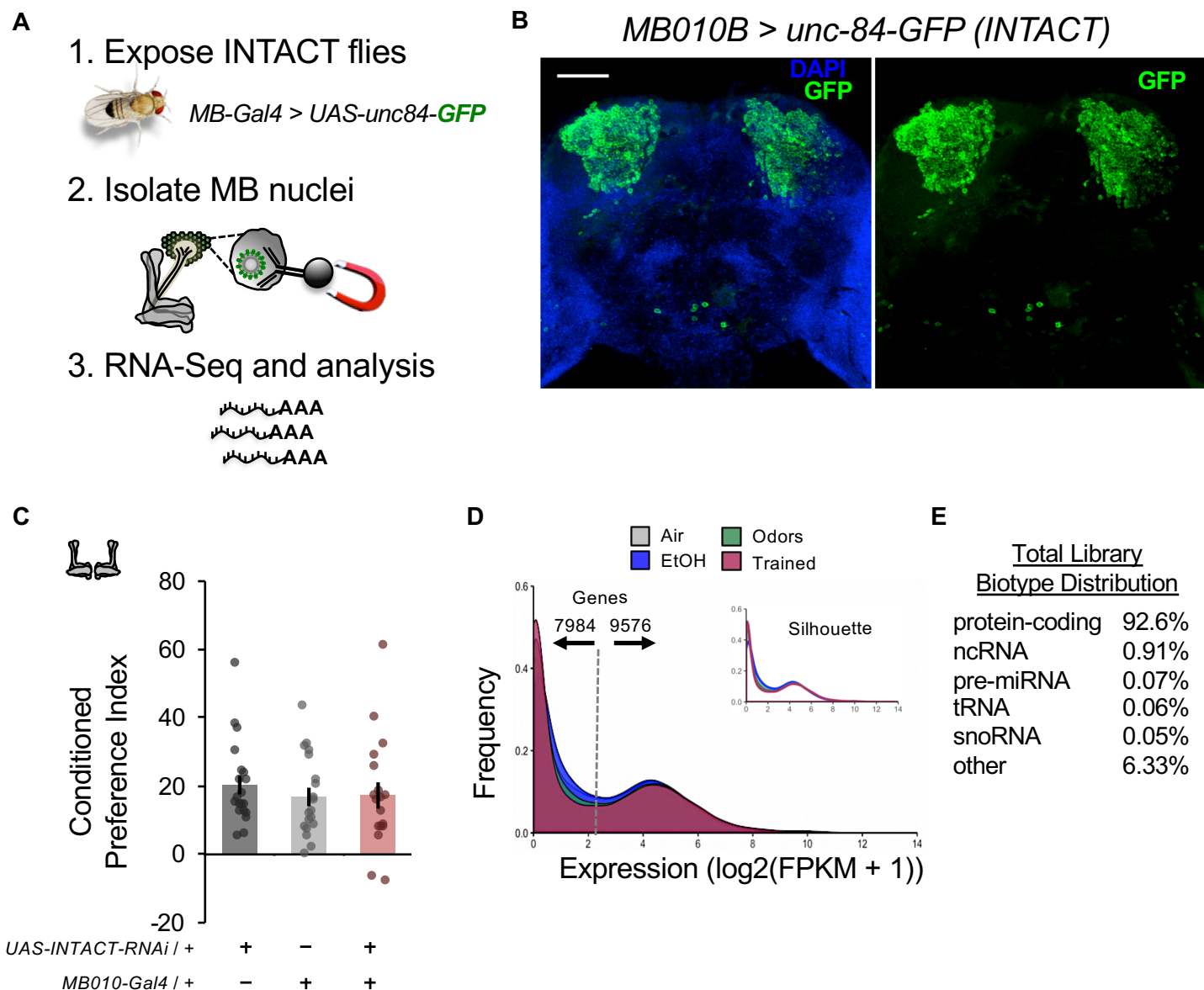
# Alcohol causes lasting differential transcription in *Drosophila* mushroom body neurons

## Supplemental Information

**Authors:** Emily Petruccelli\*†, Tariq Brown\*, Amanda Waterman\*, Nicolas Ledru\*, Karla R. Kaun\*

**Authors' Affiliations:** \*Department of Neuroscience, Brown University, Providence, Rhode Island, USA, 02912. †Department of Biological Sciences, Southern Illinois University Edwardsville, Edwardsville, Illinois, USA, 62026.

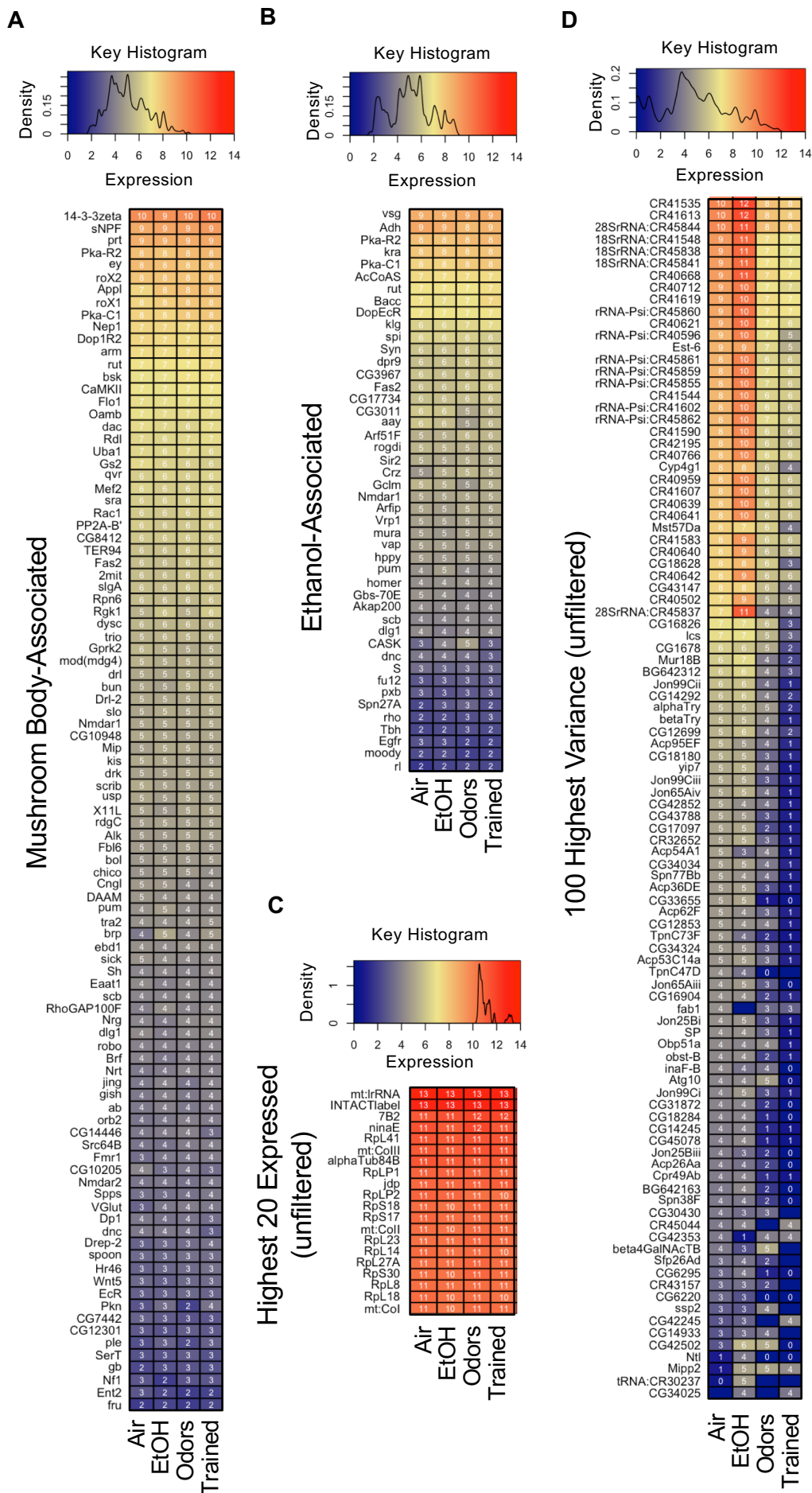
Table of Contents	page
Supplemental Title page.....	1
Supplemental Figure 1 .....	2
MB INTACT paradigm and initial analysis.	
Supplemental Figure 2 .....	3
QualiMap (BAM QC) analysis of the RNA-seq libraries.	
Supplemental Figure 3 .....	4
Expression of genes-of-interest, highest expressed, and most variable genes across behavioral conditions.	
Supplemental Figure 4 .....	5
Comparison of top differentially expressed transcripts across treatments.	
Supplemental Figure 5 .....	6
Comparison of top 200 folding-changing transcripts across treatments.	
Supplemental Figure 6 .....	7
DAVID analysis of gene ontology (GO) of top 200 p-value genes.	
Supplemental Figure 7 .....	8
Sashimi plots (IGV, Broad Institute) for alternatively expressed transcripts.	
Supplemental Figure 8 .....	9
Decreasing expression of <i>Cdc5</i> in adult neurons affects <i>Stat92E</i> expression.	
Supplemental Figure 9 .....	10
Knockdown of splice-associated targets in adult MB neurons does not affect odor responses.	
Supplemental Figure 10 .....	11
Efficacy of RNAis of spliceosome associated genes.	
Supplemental Figure 11 .....	12
RNA-seq comparison to previous studies.	



**Supplemental Figure 1. MB INTACT paradigm and initial analysis.** **A)** RNA-seq was performed on MB nuclei of flies expressing the INTACT transgene (*MB010B-Gal4 > UAS-unc-84-GFP*). Each fly brain contains ~4000 MB nuclei. **B)** A representative immunohistochemical staining of the restricted INTACT (*unc-84-GFP*) tag expressed in *MB010B-splitGal4* neurons in a 3-5-day-old adult male brain (DAPI staining DNA, Blue; anti-GFP staining GFP, Green). Confocal max stack image taken at 20x resolution (Scale bar 50  $\mu$ m). **C)** Expression of the INTACT transgene (*unc<sup>84</sup>-GFP*) in MB neurons did not interfere with odor cue-induced ethanol memory  $F(2,56)=71.36$ ,  $p=0.68$ . **D)** Histogram of the mean gene expression in fragments per kilobase (FPKM) across treatment libraries. Inset silhouette lacking fill color for increased visibility between treatments. **E)** Percent distribution of biotypes according to current annotated Flybase 'feature' types across treatment libraries.

	Total Sequences	Uniquely Mapped	Exonic / Intronic / Intergenic	Transcript Counts (FPKM)
Air	33844558	85.87%	97% / 1.5% / 0.7%	15784305
	37211293	87.31%	97% / 1.9% / 0.7%	15262082
	46395515	76.84%	97% / 1.8% / 0.8%	17714446
EtOH	65772590	85.91%	93% / 4.1% / 2.5%	12129006
	25243988	87.39%	97% / 1.6% / 0.9%	14165946
	29887470	83.14%	97% / 1.7% / 0.9%	15497176
Odors	30059906	89.08%	97% / 1.4% / 0.8%	15415177
	25082770	90.57%	97% / 1.4% / 0.8%	14377060
	41694314	89.92%	97% / 1.4% / 0.8%	21702932
	29669861	89.92%	97% / 1.6% / 0.9%	15614864
Trained	27938257	89.86%	97% / 1.4% / 0.8%	16104880
	24496938	89.84%	97% / 1.7% / 1%	14242031
	29179965	90.72%	97% / 1.4% / 0.8%	15455975
	33911395	89.6%	97% / 1.5% / 0.8%	17410191
AVERAGE	34313487	87.5%	96.7% / 1.7% / 0.94%	15776862

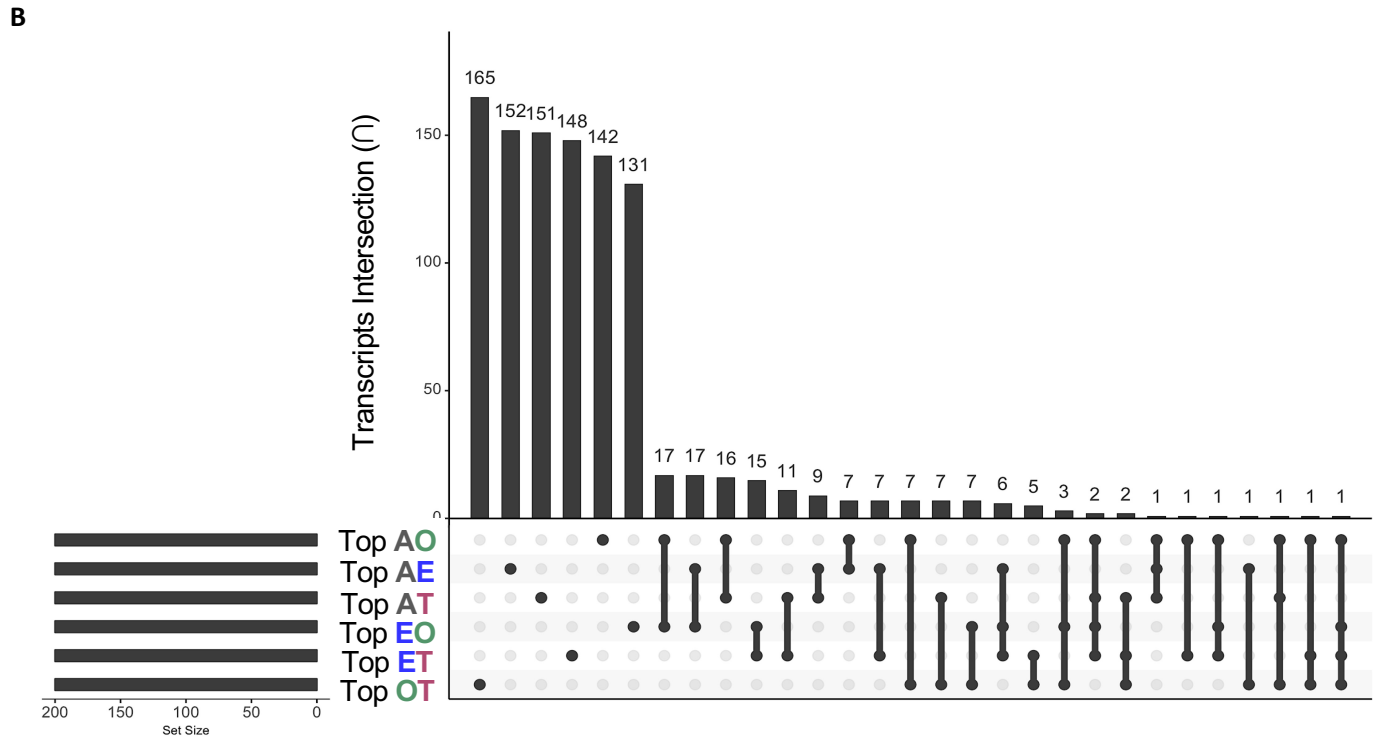
**Supplemental Figure 2. QualiMap (BAM QC) analysis of the RNA-seq libraries.** The total sequences, uniquely mapped reads %, genomic biotype, and normalized transcript counts (FPKM) are displayed for all 14 libraries.



**Supplemental Figure 3.** Expression of genes-of-interest, highest expressed, and most variable genes across behavioral conditions. A-D) Density histogram and normalized FPKM expression (blue, yellow, and red color representing relative levels) of (A) MB-associated, (B) ethanol-associated, (C) the highest 20 genes expressed genes without filtering, and (D) 100 most variable genes without filtering. All gene expression represented as condition means (log2(FPKM+1)) and ordered by mean expression.

**A**

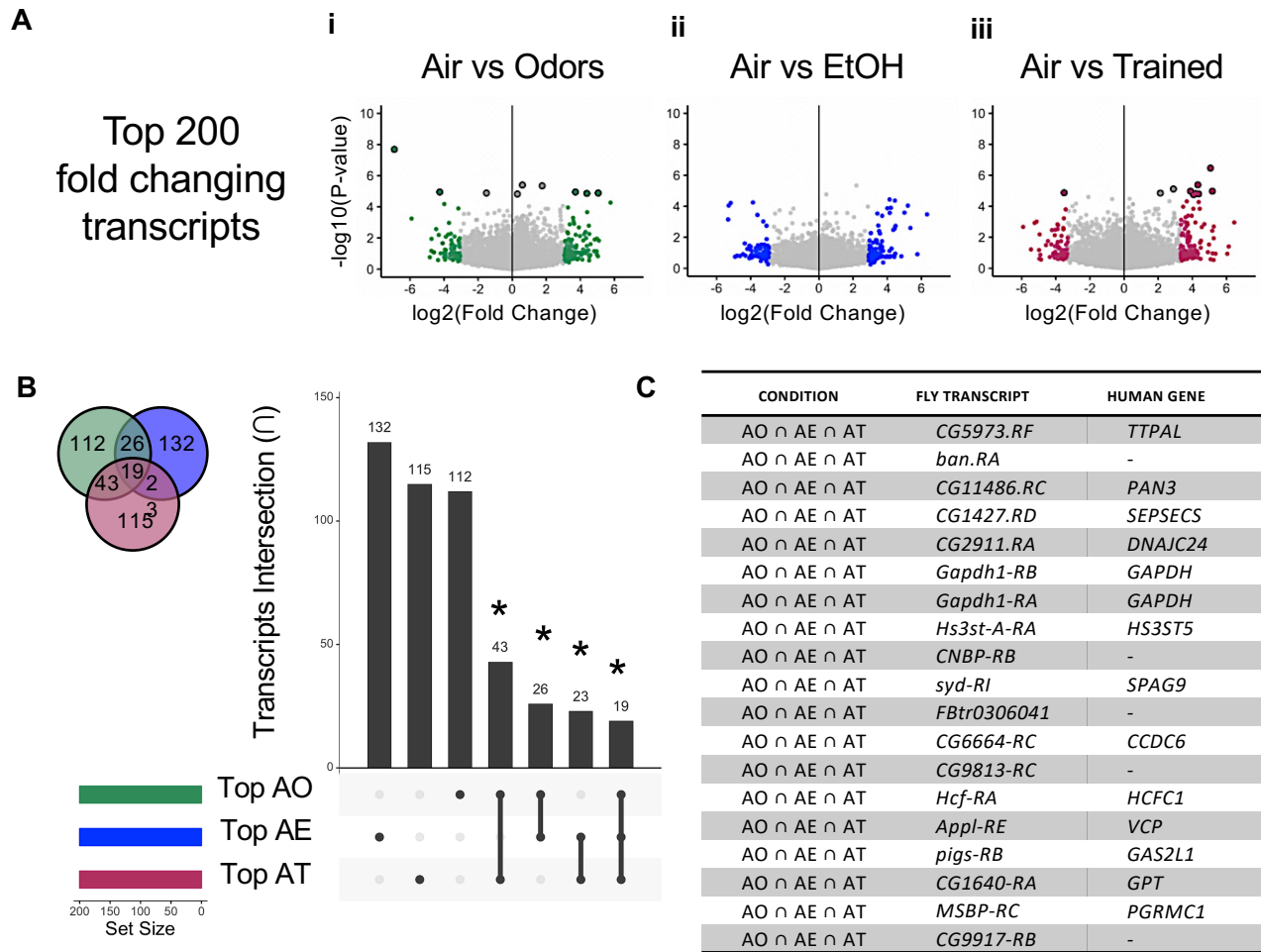
	Air vs Odors	Air vs Ethanol	Air vs Trained	Odors vs Ethanol	Ethanol vs Trained	Odors vs Trained
<b>Genes</b>	1	0	0	1	0	0
<b>Transcripts</b>	9	0	10	7	4	8



**C** **Hypergeometric distribution test p-values**

	Top 200 p-value transcripts	Top 200 Fold Changer transcripts
Interaction	p-value	p-value
AO ∩ AE	5.6e-4	4.5e-46
AO ∩ AT	2.4e-14	9.3e-73
AE ∩ AT	2.3e-5	9.6e-42

**Supplemental Figure 4. Comparison of top differentially expressed transcripts across treatments, as compared to 'Air'.** **A)** The number of genes or transcripts reaching statistical significance at FDR < 0.05 between all pairwise comparisons: 'Air vs Odors', 'Air vs EtOH', 'Air vs Trained', 'Odors vs EtOH', 'EtOH vs Trained', and 'Odors vs Trained'. Colors depict treatment condition. **B)** An upgraded Venn Diagram plot generated by an R package called 'UpSetR' demonstrating the intersection between top 200 p-value transcripts (abbreviated by first letter in treatment). **C)** The phyper function in R was used to determine hypergeometric p-values. The 'universe' test number was 34,741 transcripts from the current Dm6 transcriptome. Values are as follows: phyper(q, m, n, k, lower.tail=FALSE) ... eg. phyper(overlap#, 200, (34741-200), 400, lower.tail=FALSE).



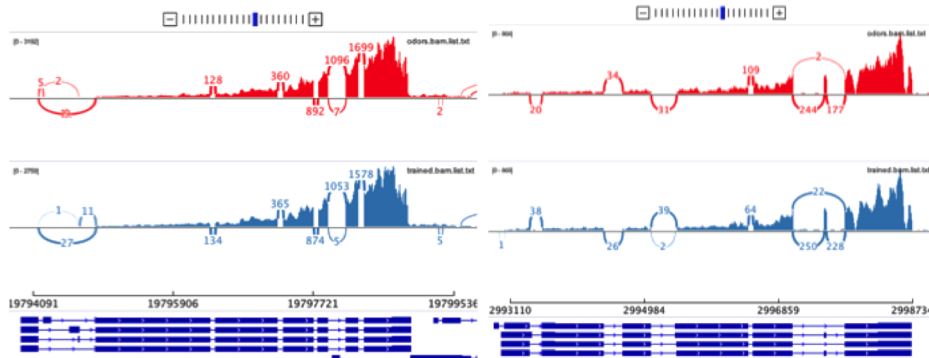
**Supplemental Figure 5. Comparison of top 200 folding-changing transcripts in response to odor, ethanol or trained (odor-ethanol) treatment.**

**A)** Volcano plots showing fold change of transcript expression ( $\log_2(\text{fc}+1)$ ) compared to the inverse of statistical significance ( $-\log_{10}(\text{p-value})$ ) (dark outline, FDR < 0.05). Plots for **(i)** Odors, **(ii)** EtOH and **(iii)** Trained (ethanol-odor pairing) compared to Air (colors depict the top 200 fold changing transcripts). **B)** An upgraded Venn Diagram plot generated by an R package called 'UpSetR' demonstrating the intersection in top 200 fold changing transcripts (abbreviated by first letter in treatment). **C)** The 19 transcripts found to be similar across all pairwise intersectional analyses are listed, along with the corresponding highest DIOPT-scored human genes.

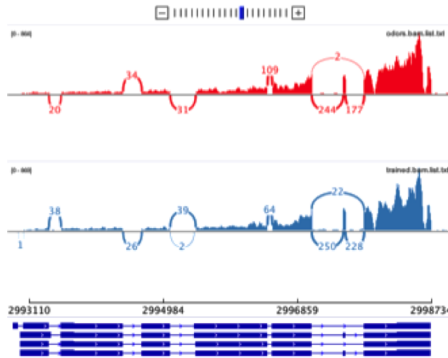
	DAVID GO terms	Adjusted p-value
<b>Air vs Odor</b>	phototransduction	0.86
	NADH dehydrogenase (ubiquinone) activity	0.91
	protein binding	0.92
<b>Air vs Ethanol</b>	<b>Alternative splicing</b>	<b>0.0002</b>
	<b>Phosphoprotein</b>	<b>0.041</b>
	<b>Coiled Coil</b>	<b>0.047</b>
<b>Air vs Trained</b>	Alternative splicing	0.062
	Developmental protein	0.2
	Phosphoprotein	0.25

**Supplemental Figure 6. DAVID analysis of gene ontology (GO) of top 200 p-value genes.** Enrichment of current GO terms within the top 200 p-value genes found between Air vs Odor, Air vs Ethanol, and Air vs Trained pairwise comparisons (Boded adjusted p-value < 0.05).

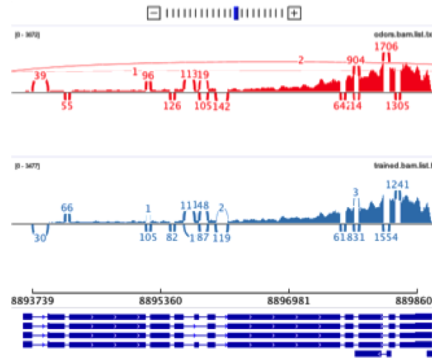
HLH106



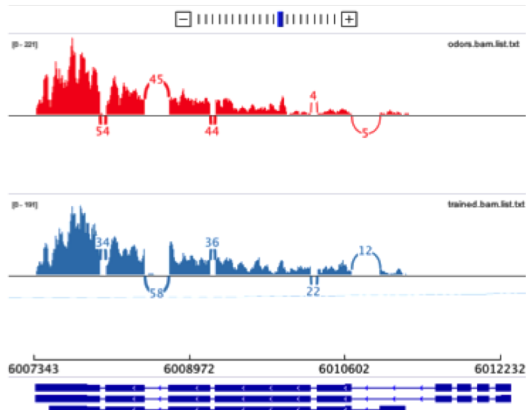
CG3558



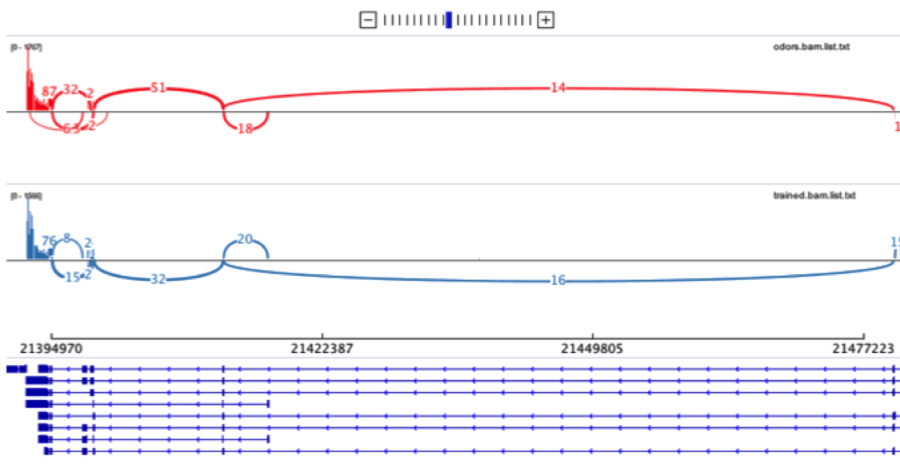
Sec31



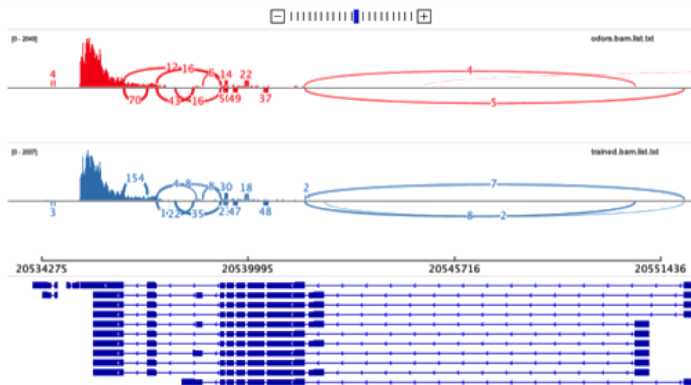
CG9098



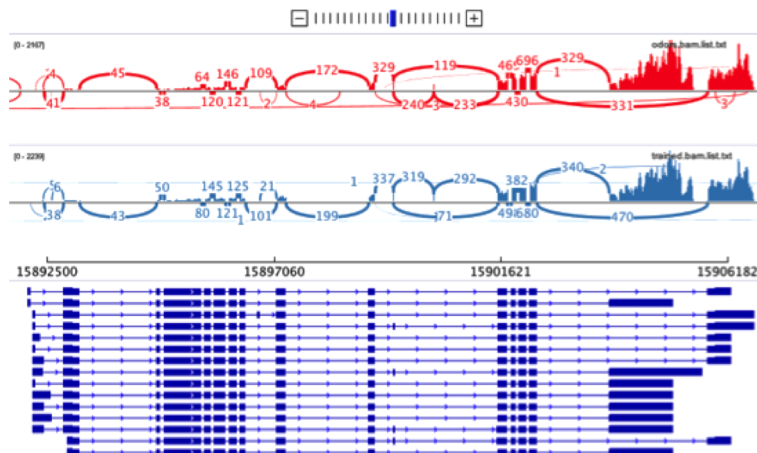
Sdc



Stat92E

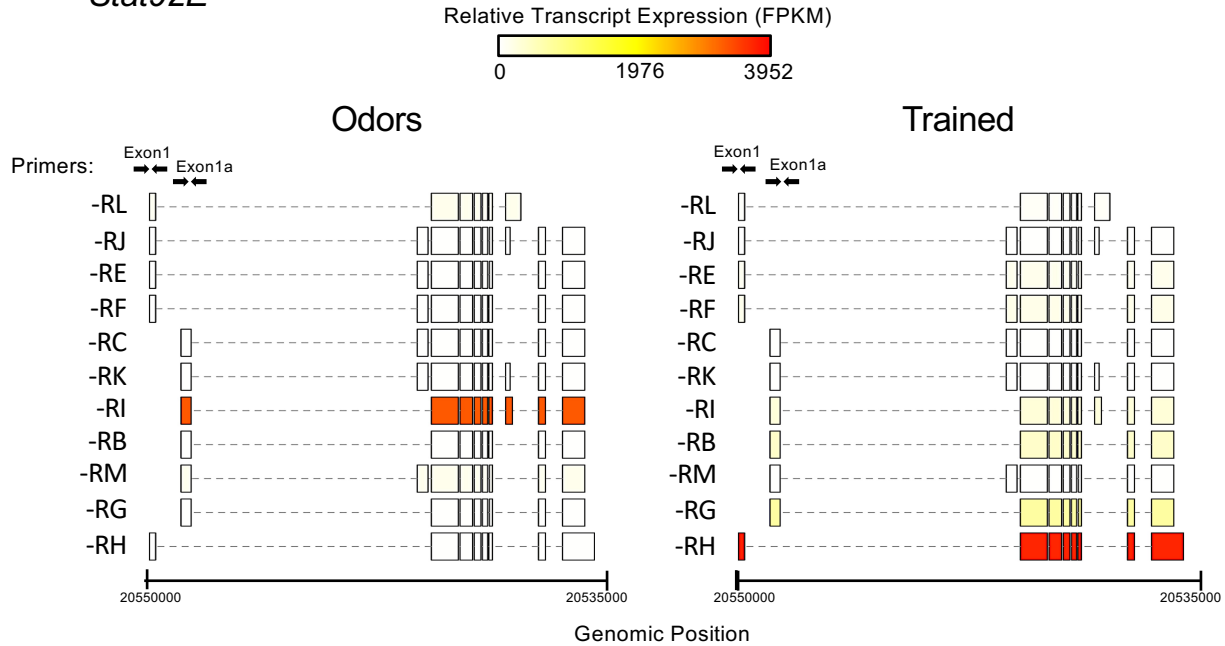


shi

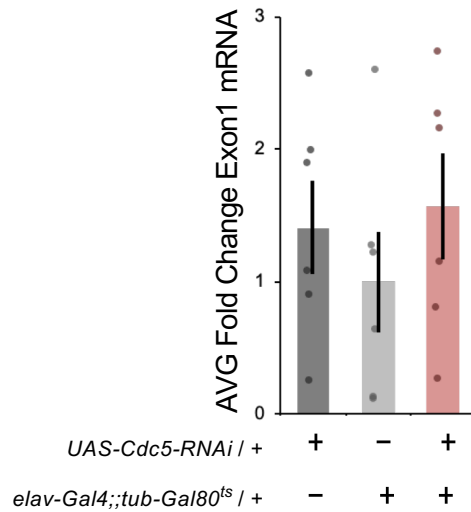


**Supplemental Figure 7. Sashimi plots (IGV, Broad Institute) for alternatively expressed transcripts.** Sashimi plots help to visualize splice junctions across merged 'Odors' and 'Trained' samples for differentially expressed transcripts.

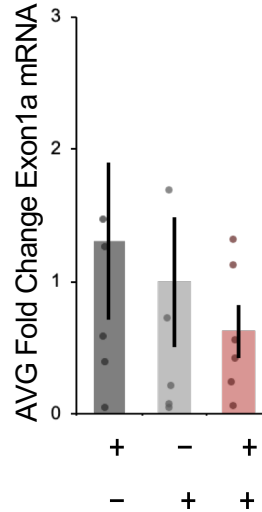
## A *Stat92E*



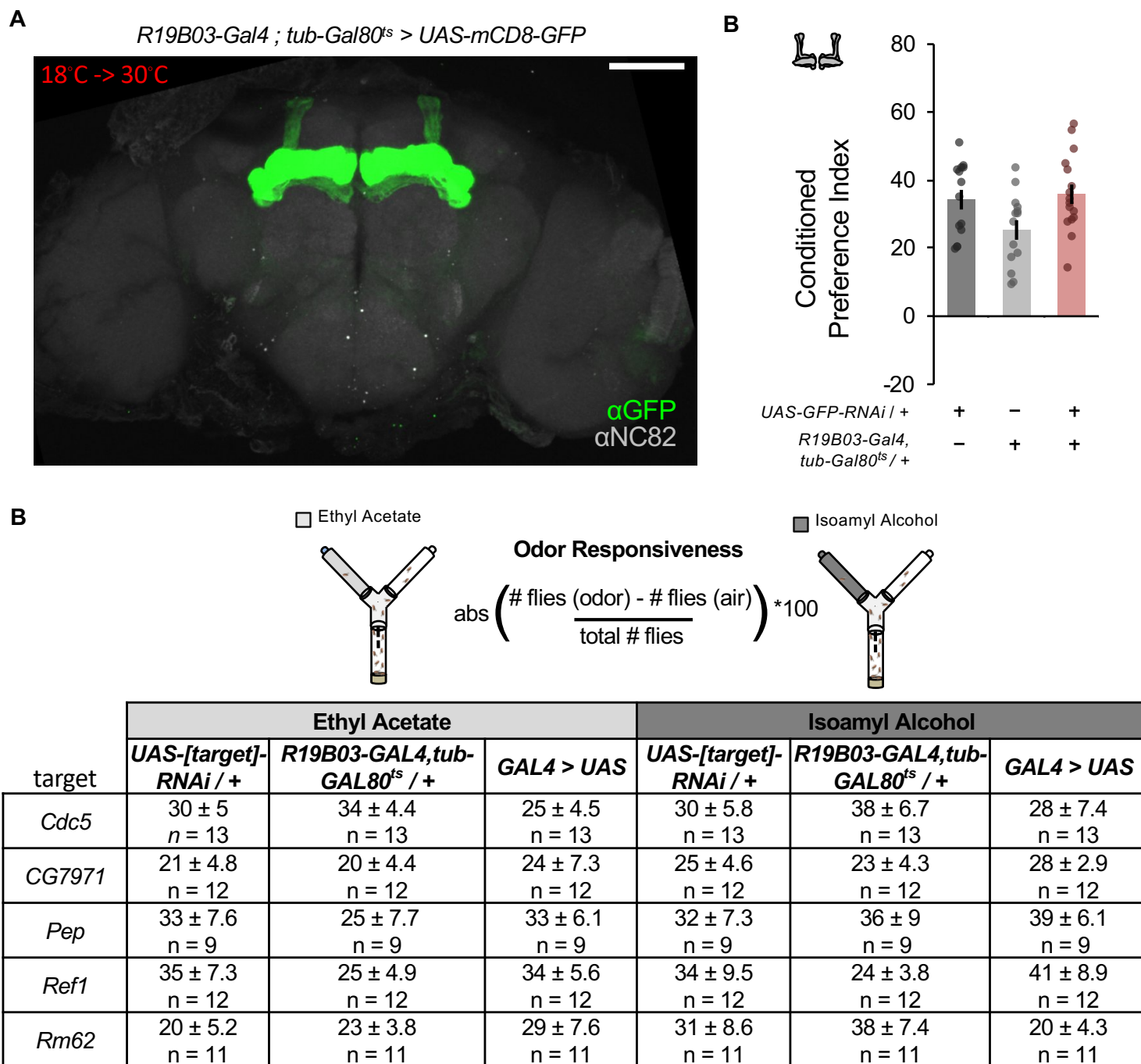
## B *i*



## *ii*



**Supplemental Figure 8. Decreasing expression of *Cdc5* in adult neurons affects *Stat92E* expression.** **A)** All transcripts of *Stat92E* expressed in MB neurons of flies treated with odors, or odor-ethanol pairings. *Stat92E-RI* was most highly expressed in MB nuclei of control flies whereas *Stat92E-RH* was more highly expressed in trained flies. **B)** Average fold change of *Stat92E* Exon1 transcripts (including *Stat92E-RH*) and Exon 1a transcripts (including *Stat92E-RI*). Data for both graphs are normalized to *elav-Gal4*; *tub-Gal80<sup>ts</sup>* / +. Decreasing *Cdc5* in adult neurons did not significantly affect *Stat-92E* Exon 1 ( $F(2,16)=2.56$ ,  $p=0.45$ ) or Exon 1a ( $F(2,16)=0.99$ ,  $p=0.40$ ) transcript expression alone.

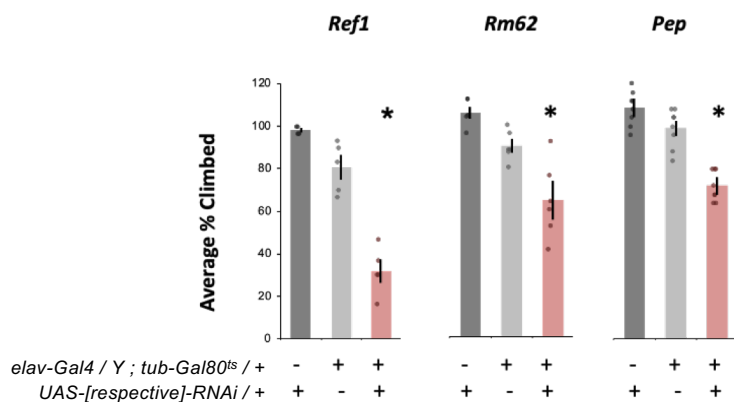


**Supplemental Figure 9. Knockdown of splice-associated targets in adult MB neurons does not affect odor responses. A)** *R19B03-Gal4* is expressed in all MB neurons when using a *UAS-mcd8-GFP* reporter. Temporal expression was achieved by rearing flies in 18°C, then permitting expression by shifting one-day-old flies to 30°C for 3 days. (Scale bar 50uM) **B)** Expressing GFP-RNAi in all MB cells using *R19B03-Gal4* did not reduce ethanol-odor memory ( $F(2,41)=4.45$ ,  $p=0.02$  with Student's t-test *UAS/Gal4* vs *Gal4,G80/+*  $p=0.007$ , and *UAS/Gal4* vs *UAS/+*  $p=0.54$ ). **C)** Odor responses as defined by preference between an odor or no odor in a Y-maze test to 1:36 ethyl acetate in mineral oil or 1:36 isoamyl alcohol in mineral oil were not significantly affected by expressing RNAi's for *Cdc5*, *CG7971*, *Pep*, *Ref1* or *Rm62* in adult MB neurons, as compared to their respective genetic controls.

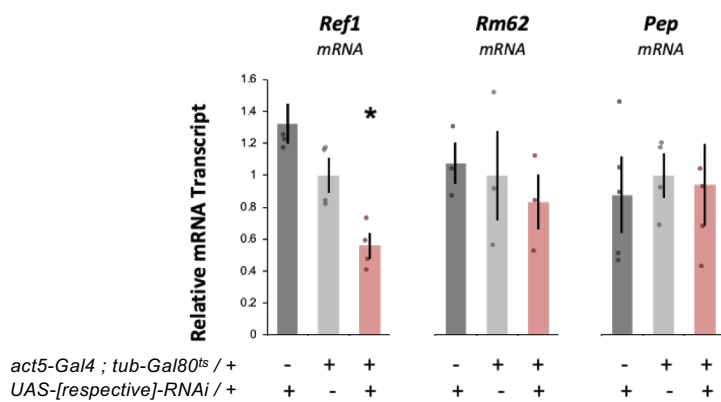
**A**

Gene	#Flies Eclosed			% Flies eclosed from pupae		
	<i>elav-Gal4/+</i>	<i>UAS-RNAi/+</i>	<i>elav-Gal4&gt;UAS-RNAi</i>	<i>elav-Gal4/+</i>	<i>UAS-RNAi/+</i>	<i>elav-Gal4&gt;UAS-RNAi</i>
<i>Ref1</i>	63 ± 4 (n=9)	48 ± 3 (n=9)	22 ± 6 (n=9)*	100	76	39*
<i>Rm62</i>	54 ± 4 (n=9)	62 ± 5 (n=9)	39 ± 3 (n=9)*	100	100	100
<i>Pep</i>	66 ± 4 (n=9)	42 ± 4 (n=9)	27 ± 4 (n=9)*	100	72	49*
<i>CG7971</i>	50 ± 9 (n=9)	53 ± 8 (n=9)	0 (n=9)*	100	100	0*
<i>Cdc5</i>	54 ± 7 (n=9)	29 ± 4 (n=9)	0 (n=9)*	100	85	0*

**B**



**C**



**Supplemental Figure 10. Efficacy of RNAis of spliceosome associated genes. A)** Expressing RNAi for *Ref1*, *Pep*, *CG7971* or *Cdc5* in all neurons throughout the life of the flies significantly affected the number of non-balanced flies eclosed, resulting in decreased survival to adulthood. Expressing RNAi for *Rm62* significantly decreased the number of non-balanced flies eclosed. Wilcoxon rank sum tests were performed for experimental lines vs respective heterozygous controls.  $p < 0.05$  are depicted in a red color with an \*. **B)** Expressing RNAi for *Ref1* ( $p < 0.001$ ), *Rm62* ( $p < 0.001$ ) or *Pep* ( $p < 0.001$ ) in all neurons using *elav-Gal4* significantly reduced the ability to climb to a 5 cm mark in a plastic vial within 10 seconds after tap down. **C)** Expressing RNAi in all adult cells using *actin5c-Gal4* significantly reduced expression of *Ref1* in adult fly heads ( $p < 0.001$ ).

Publication and study details	Gene lists
<p><b>Petrucelli et al, 2019</b> (this study)</p> <p>Time: 24hr after training  Tissue: Mushroom body nuclei  Condition: appetitive alcohol LTM, 3x 10min (sequential odors) interval spaced 60min  Technique: INTACT  Analysis: RNA-seq (HISAT2, StringTie, Ballgown)</p>	<p><b>Genes associated with top 200 transcripts 'Odors vs Trained':</b>  Aats-trp, aay, Act5C, Adgf-E, Amnionless, AnxB9, asparagine-synthetase, Atg8b, aurA, BBS4, beg, ben, Cad96Cb, CBP, Cdk1, Cdk8, CG10170, CG10428, CG10512, CG10839, CG11131, CG11148, CG11425, CG12116, CG12746, CG12909, CG13484, CG13563, CG13607, CG13708, CG13872, CG14075, CG14275, CG14285, CG14435, CG14621, CG14688, CG14717, CG14879, CG14926, CG15093, CG1513, CG15925, CG1673, CG16756, CG16926, CG16953, CG17570, CG17612, CG17816, CG17821, CG18130, CG2316, CG2911, CG3040, CG30430, CG31419, CG31431, CG3176, CG31787, CG31882, CG32533, CG33098, CG33111, CG33120, CG33795, CG3402, CG34126, CG34166, CG34376, CG34417, CG3505, CG3558, CG42394, CG42502, CG42540, CG42559, CG42577, CG42724, CG42752, CG43093, CG43172, CG43347, CG43630, CG4462, CG4462, CG4842, CG4962, CG5174, CG5446, CG5953, CG5958, CG6067, CG6770, CG8216, CG8508, CG8726, <b>CG9098</b>, CG9149, CG9263, CG9300, CG9663, CG9723, CG9777, CG9784, CG9975, CR32205, CR43653, CR43887, CR44265, CR44301, CR44886, CR45113, CR45170, CR45172, CR45884, Crk, crol, cu, Cyp12c1, Cyp4e3, cype, Cyt-b5-r, dbo, Desat2, DhpD, Dip-B, dor, endos, endos, fabp, fit, fus, Glut3, Gpdh, gus, <b>HLH106</b>, HmgZ, HPS4, Hr39, Hr4, hubl, Idh, Idh, intr, jnj, l(2)k09913, LKR, mav, mir-4963, mod, mod(mdg4), mRps18B, mtsh, NaCP60E, net, Not1, nsl1, Obp57b, odd, Or85f, Pask, Pitslre, pix, Rb97D, Reps, Reps, Rm62, Rpl1, Rpl3, <b>Sdc</b>, <b>Sec31</b>, Sh, shi, sif, SkpA, snoRNA:Me28S-A2634c, snoRNA:Me28S-C3420b, Sodh-2, <b>Stat92E</b>, su(r), Syt1, tau, TFAM, Tgi, Thor, TM45F, TotA, Trpm, UK114, viaf, Xport</p> <p style="text-align: right;"><b>*bold FDR &lt; 0.05</b></p>
<p><b>Crocker et al, 2016</b></p> <p>Time: 30min after training  Tissue: Kenyon Cells, male only?  Condition: aversive shock olfactory LTM, 8x 1min interval spaced 10min  Technique: Single Cell RNA-seq  Analysis: TopHat2, HTseq-count, DESeq2</p>	<p>no differentially expressed changes found in KCs, likely due to pooling</p>
<p><b>Widmer et al, 2018</b></p> <p>Time: T2 (12-24hr) and T3 (24-48hr) after training  Tissue: Heads  Condition: appetitive sucrose olfactory LTM, starve 19-21hr then 2min (sequential odors spaced 1min)  Technique: Dam-Pol II  Analysis: Bowtie2, damidseq_pipeline with FDR via 50,000 simulations</p>	<p><b>T2:</b>  Acox57D-p, Acp54A1, Adam, Adk2, Asciz, ATPsynC, Cdk7, CG10344, CG10562, CG10581, CG10584, CG11069, CG11137, CG11597, CG11668, CG11693, CG11777, CG11852, CG12118, CG13026, CG13031, CG13082, CG13239, CG13640, CG13731, CG13748, CG14096, CG14147, CG14572, CG14646, CG14661, CG14968, CG15098, CG15515, CG15545, CG15546, CG16884, CG17068, CG17224, CG17337, CG17549, CG17919, CG18067, CG2604, CG30101, CG30280, CG31626, CG31717, CG31812, CG32170, CG32579, CG32650, CG3301, CG33170, CG33946, CG34010, CG3552, CG3566, CG3700, CG42269, CG42568, CG42615, CG43367, CG4407, CG44286, CG45676, CG5390, CG5397, CG5867, CG5885, CG5934, CG6272, CG6429, CG6908, CG7457, CG7872, CG7896, CG8317, CG8319, CG8525, CG8788, CG9372, CG9521, CG9928, CHKov2, CIAPIN1, Cpr47Ea, Cpr49Aa, Cpr49Ah, Cpr64Aa, Cpr64Ad, Cpr78Cb, CR31144, CR43196, CR44107, CR44168, CR44264, CR44383, CR44468, CR44469, CR44621, CR44756, CR44888, CR44893, CR45150, CR45350, CR45501, CR45541, CR45682, CR45827, CR45828, CR45925, CR46108, CREG, Cyp6a15Psi, Cyp6d5, Cyp6g1, Cyp9f2, <b>Dip-B</b>, Dnal-60, E(spl)mbeta-HLH, Eflbeta, Fmo-2, FucT6, Gapdh1, GstE8, GstO1, GstO2, GstZ1, Hsf, Hsp70Bb, Lcp65Aa, Lsm10, mir-282, Mmp1, mRps25, Myb, ND-B16.6, Nxt1, Obp56d, obst-A, olf186-M, Osi19, PCB, Pcd, PGRP-LF, pinta, Prosbeta3, Prx2540-1, Ptr, remPA, Rpl15, Rps15Ab, scu, snmRNA:838, snoRNA:Psi18S-301, Syx8, tal-1A, tal-2A, tal-3A, tal-AA, Taldo, tRNA:CR32289, tRNA:H:48F, Ugt35b, vajk-1, ver, wal, yip2, yip2</p> <p><b>T3:</b>  Acox57D-p, Acp54A1, Adk2, B-H1, Cdk7, CG10257, CG10344, CG10602, CG10754, CG11137, CG11191, CG11668, CG11852, CG12118, CG13082, CG13299, CG13731, CG13751, CG13928, CG14096, CG14572, CG14968, CG15515, CG15545, CG17337, CG17549, CG30280, CG32039, CG32170, CG3226, CG32579, CG32650, CG32806, CG32815, CG33170, CG33946, CG3566, CG3939, CG42239, CG44286, CG5397, CG5867, CG6272, CG6429, CG7457, CG7872, CG8500, CG8788, CG9319, CG9521, CG9928, CG9960, CR43700, CR44371, CR44420, CR45008, <b>CR45172</b>, CR45350, CR45554, CR45911, CR45919, CR45925, CR46108, CREG, Cyp6a15Psi, Eflalpha100E, Faf2, Flo1, Fmo-2, Gapdh1, GstO2, Hsp70Bb, jdp, mei-5332, mir-282, mRpl27, mRpl35, mRps25, ND-B16.6, Nup54, Obp56d, olf186-M, pinta, pncr015:3L, Prx2540-1, Pu, put, remPA, scu, Snapin, snoRNA:lola-a, snoRNA:Psi18S-301, Syx8, Taldo, Tes, TRAM, tRNA:CR32287, tRNA:CR32289, tRNA:H:48F, Vps4, yip2</p> <p style="text-align: right;"><b>*bold similar in this study</b></p>
<p><b>Pacifico et al, 2018</b></p> <p>Time: 5-, 20-, 30-, and 40-day-old  Tissue: Sex- and age-matched brains  Condition: Age-related cognitive decline, aversive shock olfactory STM, 1min (sequential odors spaced 30sec)  Technique: RNA-seq  Analysis: STAR, HTSeq-count, DESeq2 with default filtering and WGCNA</p>	<p><b>Male, likelihood ratio test (LRT) q-value &lt; 0.05 (1076 genes identified):</b> too many to list here</p> <p>Similar in this study: <b>aay, aurA, ben, CG11425, CG13607, CG14879, CG1673, CG42394, CG5174, CG5446, CG5953, CG9098, cype, Gpdh, Rpl1, Sodh-2, Thor, TM45F, TotA</b></p> <p style="text-align: right;"><b>*bold similar in this study</b></p>
<p><b>Signor and Nuzhdin, 2018</b></p> <p>Time: 10, 20, 30min post exposure  Tissue: Male and female heads, heterozygous inbred lines  Condition: Ethanol vapor exposure from food  Technique: RNA-seq  Analysis: Exon delineation, a priori gene list filtering and abundance cutoffs</p>	<p><b>Genes associated with exonic P_Treatment &lt; 0.05 (1941 unique genes of 2402 identified):</b>  too many to list here</p> <p>Similar in this study: <b>CG12116, CG13607, CG17816, CG2911, CG34166, CG3505, CG42540, CG9098, Crk, dor, endos, fus, jnj, l(2)k09913, mod(mdg4), NaCP60E, odd, Pitslre, Rpl1, Sec31, su(r), Trpm</b></p> <p style="text-align: right;"><b>*bold similar in this study</b></p>
<p>Other important studies with different methodologies: <b>Kong et al, 2010 (microarray); Morozova et al, 2009 (microarray); Morozova et al, 2015 (QTL analysis).</b></p>	

**Supplemental Figure 11. RNA-seq comparison to previous studies.** The gene lists from previous studies (in no particular order), most relevant to this study. Study details are provided to bear in mind contextual relevance.