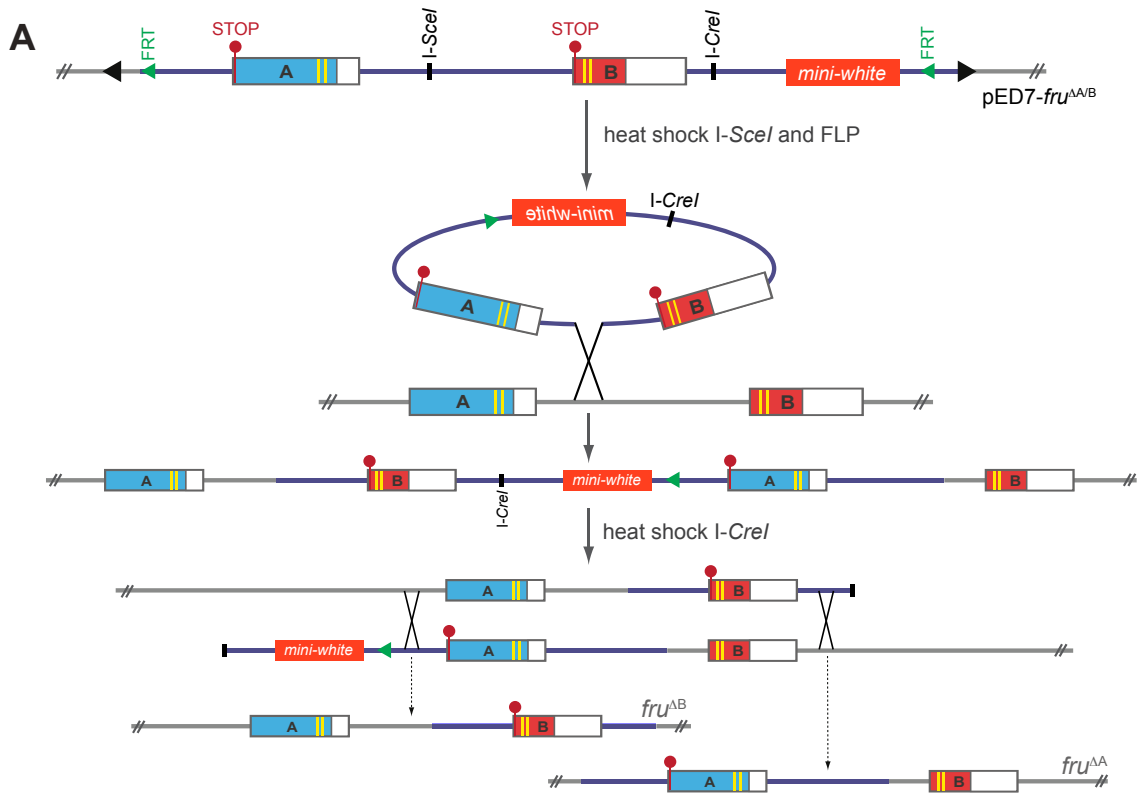
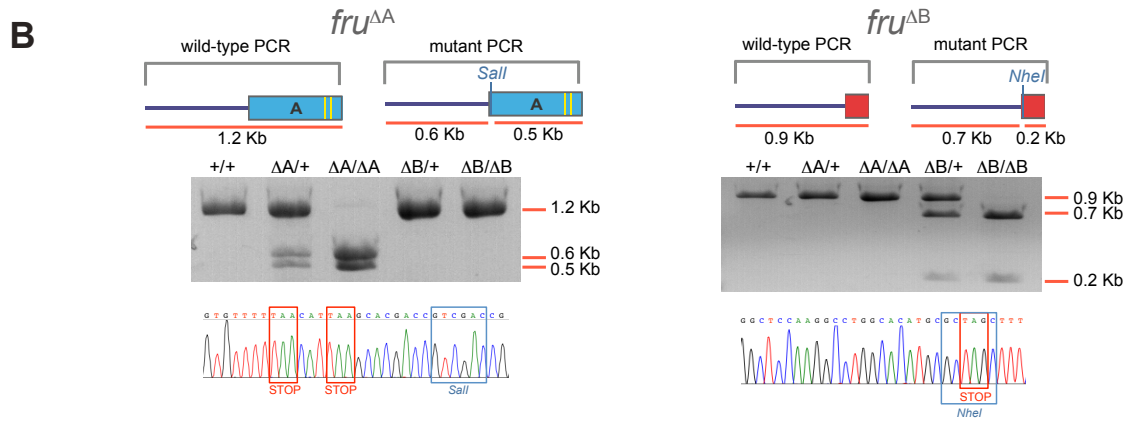
Current Biology, Volume *24*

Male-Specific Fruitless Isoforms Target Neurodevelopmental Genes to Specify a Sexually Dimorphic Nervous System

Megan C. Neville, Tetsuya Nojima, Elizabeth Ashley, Darren J. Parker, John Walker, Tony Southall, Bram Van de Sande, Ana C. Marques, Bettina Fischer, Andrea H. Brand, Steven Russell, Michael G. Ritchie, Stein Aerts, and Stephen F. Goodwin


A
FRT
STOP
I-SceI
STOP
I-CreI
mini-white
FRT
pED7-fruΔA/B
heat shock I-SceI and FLP
mini-white
I-CreI
heat shock I-CreI
mini-white
fruΔB
fruΔA


B
fruΔA
wild-type PCR
mutant PCR
SalI
1.2 Kb
0.6 Kb
0.5 Kb
+/+
ΔA/+
ΔA/ΔA
ΔB/+
ΔB/ΔB
1.2 Kb
0.6 Kb
0.5 Kb
STOP
STOP
SalI
fruΔB
wild-type PCR
mutant PCR
NheI
0.9 Kb
0.7 Kb
0.2 Kb
+/+
ΔA/+
ΔA/ΔA
ΔB/+
ΔB/ΔB
0.9 Kb
0.7 Kb
0.2 Kb
STOP
NheI

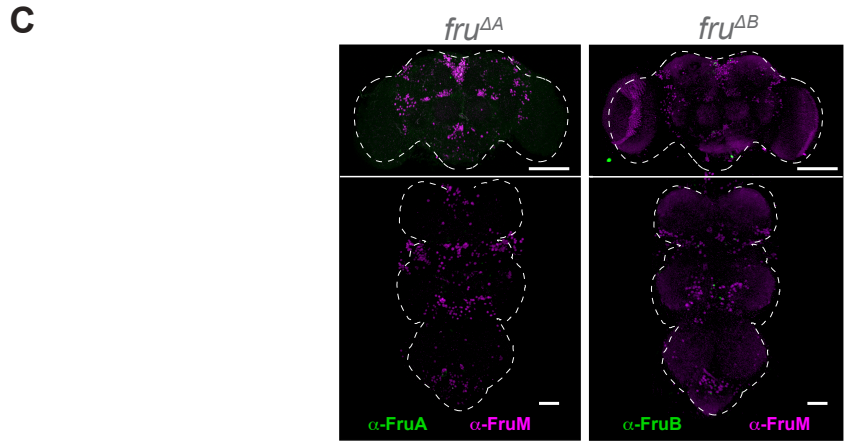

C
fruΔA
fruΔB
α-FruA
α-FruM
α-FruB
α-FruM

Figure S1. Ends-in homologous recombination at the *fru* locus. Related to Figure 1.

(A) To create novel $fru^{\Delta A}$ and $fru^{\Delta B}$ mutants, we generated a transgenic fly line containing a *fru P*-element targeting construct (donor) located on the second chromosome. The donor *P*-element contains the two FLP-recombinase recognition FRT sequences (green arrows), the I-*Cre*I recognition site, the I-*Sce*I recognition, and three regions of homology to the *fru* locus, including premature stop codon mutation (red line with circle) in exons A and B as well as the novel recognition sequences *Sal*I or *Nhe*I (see below). The donor sequence was next excised and linearized using *hs-FLP* and *hs-I-SceI*, respectively. A recombination event at the *fru* locus was detected by the visibility of a third chromosomal expressing $white^{+}$ marker in an *ey-FLP* background. This recombinant contained a duplication at the *fru* locus. The duplication was resolved by the expression of heat shock *I-CreI*, causing a double-stranded break at the locus, which was then repaired, leading to alternative resolution events selected for by the loss of the $white^{+}$ marker. The inclusion of either the A or B mutant exons were subsequently selected by the inclusion of the recognition sequences *Sal*I or *Nhe*I, respectively.

(B) Molecular Confirmation of $fru^{\Delta A}$ and $fru^{\Delta B}$ mutants. Genomic DNA from heterozygous and homozygous $fru^{\Delta A/B}$ mutants was PCR amplified and restriction enzyme digested (with *SalI* for $fru^{\Delta A}$ and *NheI* for $fru^{\Delta B}$). Chromatograms of sequencing reactions across mutations introduced into *fru* exons A and B using template DNA from homozygous mutant animals.

(C) Fru^{M} expression in $fru^{\Delta A}$ and $fru^{\Delta B}$ mutants. Five-day old $fru^{\Delta A}$ or $fru^{\Delta B}$ mutant adult male CNSs stained with anti-Fru^{M} and either anti-Fru^{A} or anti-Fru^{B}, respectively. Brain scale bars represent 100 μm, and VNC scale bars represent 50 μm.

A

Genotype	$fru^{\Delta A}$ ♂	$fru^{\Delta A}$ ♀	$fru^{\Delta B}$ ♂	$fru^{\Delta B}$ ♀	$fru^{\Delta C}$ ♂	$fru^{\Delta C}$ ♀
Canton S	97 (105)	100 (132)	100 (121)	100 (101)	99 (52)	100 (63)
fru^{4-40}	113 (105)	99 (121)	100 (101)	110 (136)	105 (96)	105 (118)
fru^{sat15}	98 (125)	96 (146)	0 (90) ***	9[1] (98) ***	18 (140) ***	12 (122) ***
$fru^{\Delta A}$	117 (78)	109 (94)	115 (163)	107 (131)	102 (163)	107 (131)
$fru^{\Delta B}$	107 (131)	115 (163)	0 (56) ***	0 (54) ***	96 (78)	104 (92)
$fru^{\Delta C}$	102 (64)	100 (121)	94 (103)	105 (75)	4 (67) ***	7 (94) ***

B

Genotype	# examined	# with MOL	# without MOL	% with MOL
Canton S ♂	26	26	0	100
Canton S ♀	25	0	25	0
fru^{sat15}/ fru^{4-40} ♂	22	0	22	0 ***
$fru^{\Delta A}$ / fru^{4-40} ♂	27	27	0	100
$fru^{\Delta B}$ / fru^{4-40} ♂	25	25	0	100
$fru^{\Delta C}$ / fru^{4-40} ♂	25	0	25	0 ***

C

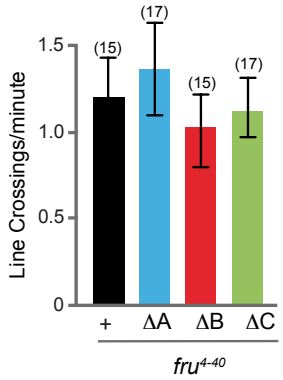


Figure S2. *fru* mutant analyses, related to Figure 1 and 2.

(A) Viability of $fru^{\Delta A/B/C}$ mutants; homozygotes, or in combination with other *fru* variants, related to Figure 2. Viability is shown as percentage of the expected/actually number of offspring observed in the progeny produced by the corresponding genetic cross. The lethal stage of the $fru^{\Delta B}$ mutants was late pupal (data not shown). [1]A few female escapers survived to adulthood from the $fru^{\Delta B}/fru^{sat15}$ genotype, however they failed to live beyond a few hours. ***$p < 0.001$ (Fisher exact test).

(B) Formation of the male-specific Muscle of Lawrence (MOL) in *fru* mutants, related to Figure 1. ***$p < 0.001$ (Fisher exact test).

(C) Locomotion of male fru^{M} mutants. No significant changes (Kruskal-Wallis ANOVA test).

Allele 1	Allele 2	IPI	pulse song frequency	sine song frequency	*n*
$fru^{\Delta A}$	$fru^{\Delta A}$	34.74 ± 0.574	283.9 ± 7.60	169.8 ± 5.04	24
$fru^{\Delta A}$	fru^{4-40}	36.26 ± 0.628	262.2 ± 8.32	169.2 ± 5.42	20
$fru^{\Delta A}$	fru^{sat15}	38.1 0± 0.645	282.5 ± 8.54	182.6 ± 5.58	19
$fru^{\Delta A}$	+	35.76 ± 0.664	283.6 ± 8.79	159.0 ± 5.92	18
$fru^{\Delta A2}$	$fru^{\Delta A2}$	34.91 ± 0.587	267.6 ± 7.77	166.6 ± 5.44	23
$fru^{\Delta A2}$	fru^{4-40}	38.62 ± 0.630	250.5 ± 8.35	163.5 ± 5.44	20
$fru^{\Delta A2}$	fru^{sat15}	38.32 ± 0.629	279.1 ± 8.33	175.8 ± 6.13	20
$fru^{\Delta A2}$	+	36.79 ± 0.646	253.8 ± 8.55	157.7 ± 5.43	19
$fru^{\Delta B}$	fru^{4-40}	**39.45 ± 0.703** **	219.1 ± 9.30	170.5 ± 6.82	16
$fru^{\Delta B}$	+	36.97 ± 0.630	263.7 ± 8.34	153.7 ± 5.44	20
$fru^{\Delta B2}$	fru^{4-40}	**39.08 ± 0.630** **	224.6 ± 8.34	155.7 ± 5.44	20
$fru^{\Delta B2}$	+	36.61 ± 0.630	265.6 ± 8.35	154.3 ± 5.58	20
$fru^{\Delta C}$	$fru^{\Delta C}$	**45.31 ± 0.631** ***	**244.4 ± 8.35** *		20
$fru^{\Delta C}$	fru^{4-40}	**45.92 ± 0.645** ***	299.9 ± 8.55		19
$fru^{\Delta C}$	fru^{sat15}	**45.54 ± 0.733** ***	281.6 ± 9.17		15
$fru^{\Delta C}$	+	37.18 ± 0.615	272.9 ± 8.15	155.1 ± 5.18	21
fru^{4-40}	+	35.83 ± 0.629	246.5 ± 8.32	164.6 ± 5.44	20
fru^{sat15}	+	35.50 ± 0.631	258.3 ± 8.35	177.1 ± 5.30	20
+	+	36.63 ± 0.429	261.7 ± 5.68	159.3 ± 3.94	43

Figure S3. Analysis of courtship song in *fru* mutants, related to Figure 3.

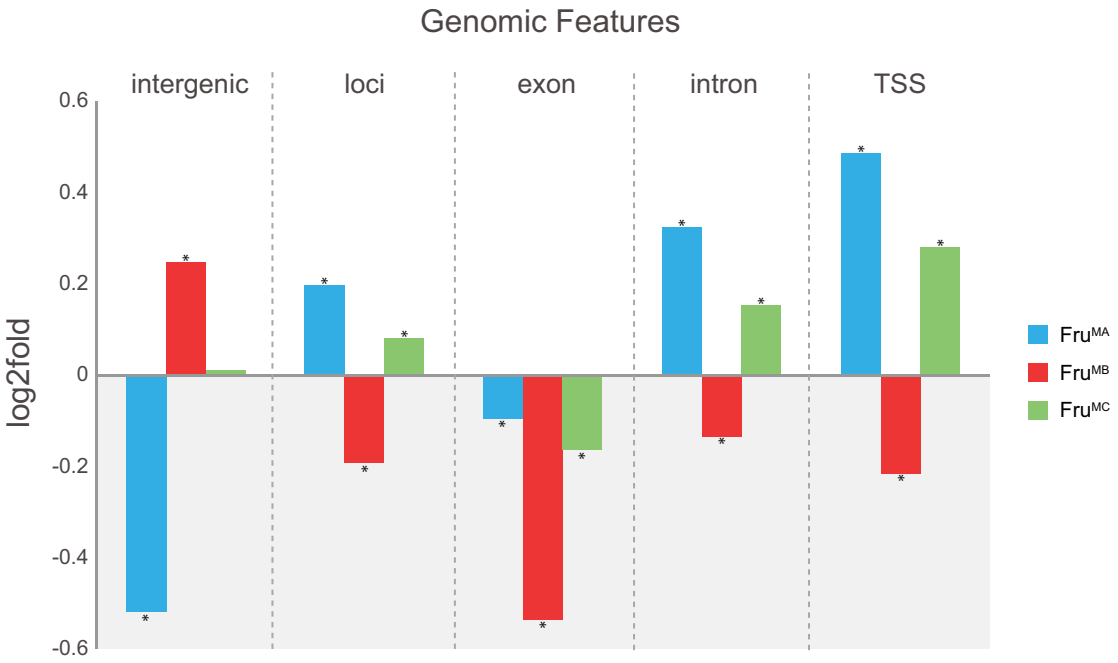
Song analysis of $fru^{\Delta A/B/C}$ mutant males as homozygotes, or in combination with other *fru* variants is shown (allele1/allele2). Independently isolated $fru^{\Delta A2}$ and $fru^{\Delta B2}$ mutant alleles generated by homologous recombination (as described in Figure S1) have been included in this analysis for comparison. The mean inter-pulse intervals (IPIs) are shown ± Standard Error of the Mean (SEM). Mean pulse song frequencies and sine song frequencies (in Hz) are shown ± SEM. *n* represents the number of males tested for each genotype. Statistically significant number are shown in bold, $*p < 0.01$; $**p < 0.001$; $***p < 0.0001$ (Kruskal-Wallis ANOVA test).

A

	FDR				
Dataset	**1**	**5**	**10**	**11 or 13**	**25**
MA larvae (1,2)	389	835	1169	nd	2308
MA pupae (1,2)	0	0	16	1750	14639
MA adult	32	62	160	2594	12289
MB larvae (2,3)	2109	5791	7165	nd	10299
MB pupae	2104	3219	3972	nd	5549
MB adult	809	1886	2495	nd	12398
MC larvae	1420	5302	7560	nd	12274
MC pupae	2813	4268	5194	nd	6957
MC adult (1,2)	3618	9516	11023	nd	13835

Half window size: 300 bp (400 bp for MC pupae)
Maximum gap separating peaks: 200 bp (350 bp for MC pupae)
Minimum number of probes: 5 (3 for MC pupae)

B



C

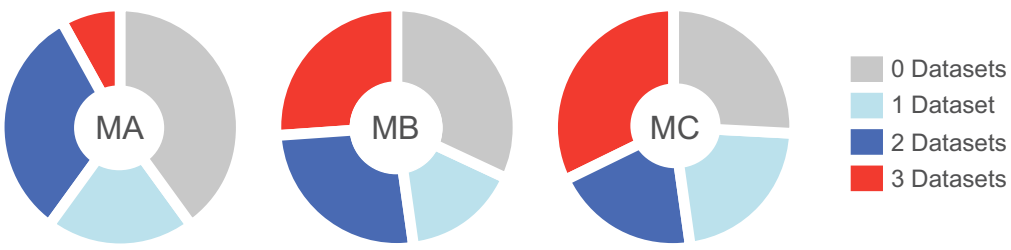


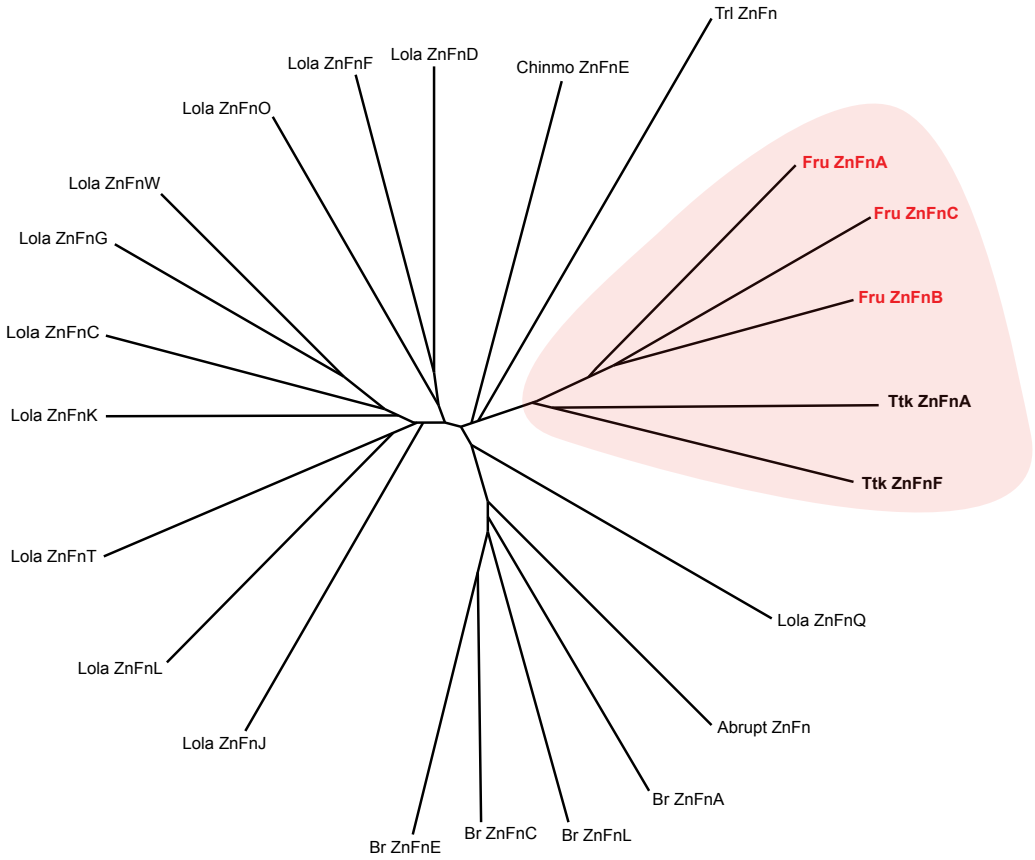
Figure S4. Enrichment of Fru^M peaks associated with genomic regions, related to Figure 4 and 5.

(A) Number of genes identified at varying FDR cut offs for each DamID dataset (as described in experimental procedures).

(B) Log2-fold enrichment of Fru^M peaks associated with the following genomic regions: intergenic regions, transcriptional start site (TSS) including 1 kb of sequence upstream, exonic regions, intronic regions [S1].

(C) Overrepresentation of direct Dsx targets in Fru^M datasets [S2]. Each portion of the circles shown represents the percent of direct Dsx target genes present in 0, 1, 2 or 3 DamID datasets for each Fru^M isoform throughout development.

A



B

```
Fru ZnFnA   LYRCV--SCN-----KIVSN----RWHHANIHRP----QSHECPVCGQKFTRRDNMKAHCKIKHAD-
Fru ZnFnB   LHRCK--LCG-----KVVTH----IRNHYHVHFP----GRFECPLCRATYTRSDNLRTHCKFKHPM-
Fru ZnFnC   MWRCR--SCG-----KEVTN----RWHHFHSHTA----QRSMCPYCPATYSRIDTLRSHLRVKHPD-
Ttk ZnFnA   VYRCT--ECAKENMQKTFKNKYSFQRHAFLYHEG-KHRKVFPCPVCSKEFSRPDKMKNHLKMTHEN-
Ttk ZnFnF   TYRCK--VCS-----RVYTHISNFCRHYVTSHKR--NVKVYPCPFCFKEFTRKDNMTAHVKIIHKI-
Trl ZnFn    --------AK-----RAKHPPGTEKPRSRSQSE-----QPATCPICYAVIRQSRNLRRHLELRHFAK
Chinmo ZnFnE--KCL--YCD-----RLYGYETNLRAHIRQRHQG----IRVPCPFCERTFTRNNTVRRHIAREH---
Br ZnfnA    PQECP--YCR-----RTFSCYYSLKRHFQDKHEQ--SDTLYVCEFCHRRYRTKNSLTTHKSLQHRG-
Br ZnfnE    DFRCN--PCN-----KNLSSLTRLKRHIQNVHMR--PTKEPVCNICKRVYSSLNSLRNHKSIYHRNL
Br ZnfnL    LFSCQ--LCG-----KLLCSKASLKRHIADKHAV--RQEEYRCAICERVYCSRNSLMTHIYTYHKS-
Lola ZnFnA  SFTCP--QCY-----RTYRRHGTLRRHLRQECGK---GKSMVCSVCGHRTKRADHLRQHVRKKHPE-
Lola ZnFnJ  QLLC---QCG-----RYYNTLSRLMLHQREECQ---DFKRFQCDFCLKWFKRRSHLNRHKKLHDAE-
Lola ZnFnD  RFHCA--VCN-----KSYLRKRHLQRHMRDECIG--IPPRFNCEFCSSRFRRKYHMVRHLVSKHGI-
Lola ZnFnO  -MQCK--HCN-----RYYKSHQKLQEHVRKYCLK---QKKYKCVSCEYRSRRKDHVLRHAKRKHCM-
Lola ZnFnC  GHPCP--VCG-----RVYKLKSSLRNHQKWECG---KEPQFQCPFCVYRAKQKMHIGRHMERMHKE-
Lola ZnFnT  SHICP--RCE-----KAYTYKKNLSRHLRYECG---QLPTEKCRHCSYVARYKHSLNMHVKTQHPE-
Lola ZnFnQ  PFVCQ--TCG-----RRYQVLGTLRRHMRKECNQ---PKKYVCRMCERRFHYNFKLQDHYYYVHKG-
Lola ZnFnK  -HACD--QCG-----KSYKTRKSLSRHRRFECRFTTERPIFQCPSCNYAAKRSDNLTKHIK------
Lola ZnFnG  -YECR--HCG-----KKYRWKSTLRRHENVECGG--KEPSHQCPYCPYKSKQRGNLGVHVRKHHTD-
Lola ZnFnU  SHICP--RCE-----KAYTYKKNLSRHLRYECG---QLPTEKCRHCSYVARYKHSLNMHVKTQHPE-
Lola ZnFnL  KLQCP--QCP-----NAYTRLSALKRHLEFECG---MLENFRCQVCDAGFKRKDSLNRHCKVKKHN-
Lola ZnFnW  KYACN--VCG-----KTYKIKGSLKRHKNYECG---VEPNLKCPHCPHKCKYRSDLRKHMNQKHAD-
Lola ZnFnY  QLLC---QCG-----RYYNTLSRLMLHQREECQ---DFKRFQCDFCLKWFKRRSHLNRHKKLHDAE-
Lola ZnFnF  RYQCPRFNCL-----KSYKDASSLQRHIRYECGG---QKKFRCLMCGKAFSQSSHLKRHLESGVCV-
```

Figure S5. Relationship between C_2H_2 Zn-fingers found in Fru and Ttk isoforms, related to Figure 6.

(A) Phylogenetic tree of Fru-related C_2H_2 Zn-fingers domains. The unrooted tree was visualized using open source Dendroscope software [S3]. The pink shaded area represents the branch containing the related Fru and Ttk domains.

(B) ClustalW Alignment of Fru-related C_2H_2 Zn-fingers domains.

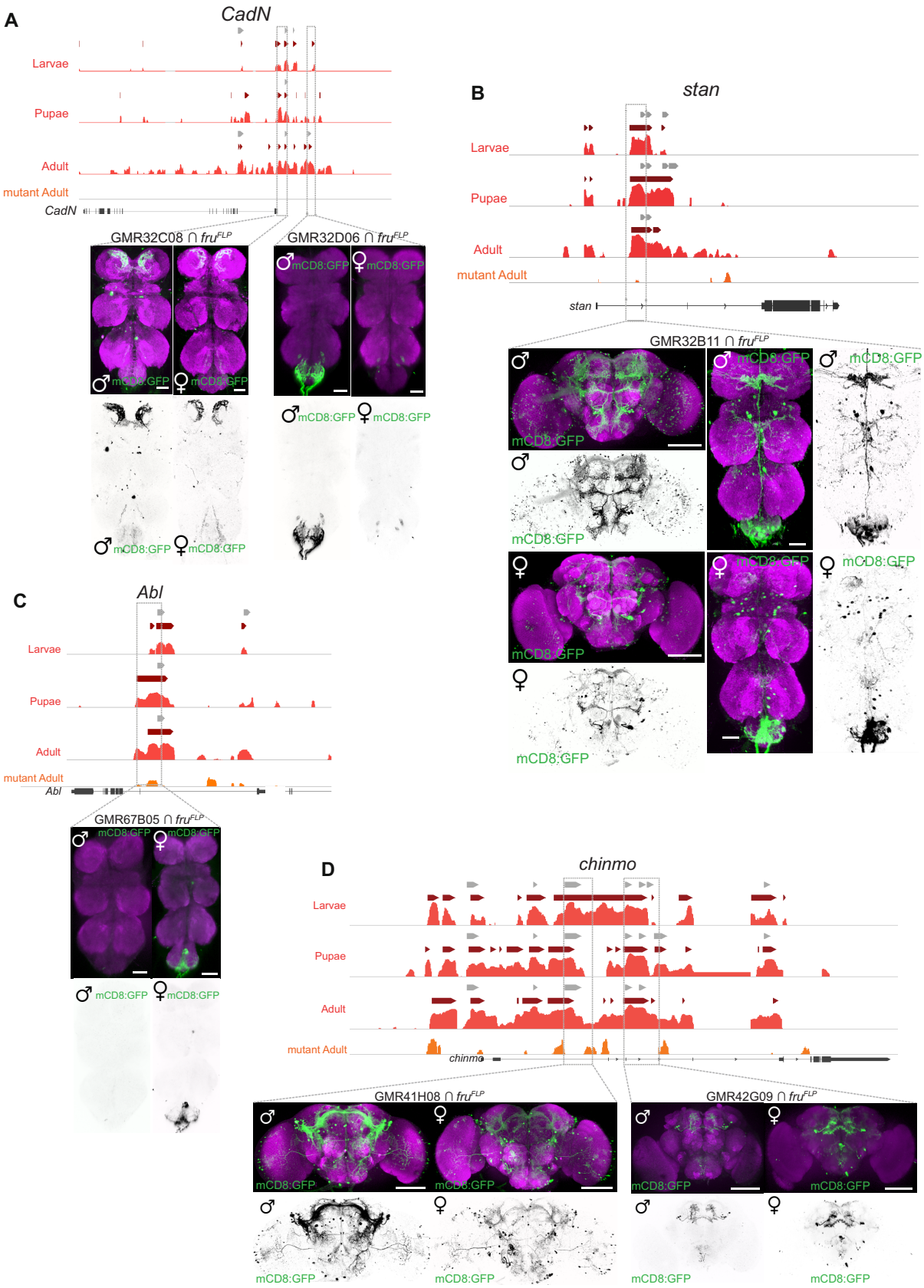

A
CadN
Larvae
Pupae
Adult
mutant Adult
CadN
GMR32C08 ∩ fruFLP
GMR32D06 ∩ fruFLP
mCD8:GFP
B
stan
Larvae
Pupae
Adult
mutant Adult
stan
GMR32B11 ∩ fruFLP
C
Abl
Larvae
Pupae
Adult
mutant Adult
Abl
GMR67B05 ∩ fruFLP
D
chinmo
Larvae
Pupae
Adult
mutant Adult
chinmo
GMR41H08 ∩ fruFLP
GMR42G09 ∩ fruFLP

Figure S6. Sexually dimorphic expression of Fru^{MB} motif-containing genomic enhancers in *fru* neurons, related to Figure 7.

(A) Dimorphic expression of *CadN* Fru^{MB}-CRM-GAL4s in *fru* neurons in the CNS. DamID binding profiles of the Fru^{MB} isoforms through development at the *CadN* locus are shown. FlyLight-GAL4 lines GMR32C08 and GMR32D06 (boxed in grey).

(B) Dimorphic expression of *stan* Fru^{MB}-CRM-GAL4 in *fru* neurons in the CNS. DamID binding profiles of the Fru^{MB} isoforms through development at the *stan* locus are shown. FlyLight-GAL4 line GMR32B11 (boxed in grey).

(C) Dimorphic expression of *Abl* Fru^{MB}-CRM-GAL4 in *fru* neurons in the CNS. DamID binding profiles of the Fru^{MB} isoforms through development at the *Abl* locus are shown. FlyLight-GAL4 line GMR32B11 (boxed in grey).

(D) Dimorphic expression of *chinmo* Fru^{MB}-CRM-GAL4s in *fru* neurons in the CNS. DamID binding profiles of the Fru^{MB} isoforms through development at the *chinmo* locus are shown. FlyLight-GAL4 lines GMR41H08 and GMR42G09 (boxed in grey).

(A)-(D) fru^{FLP} UAS>stop>mCD8-GFP male and female brains stained with anti-GFP (green) and the nc82 anatomical reference (magenta) are shown with inverted views of the GFP signal shown in black and white below. Brain scale bars represent 100 μm, and VNC scale bars represent 50 μm.

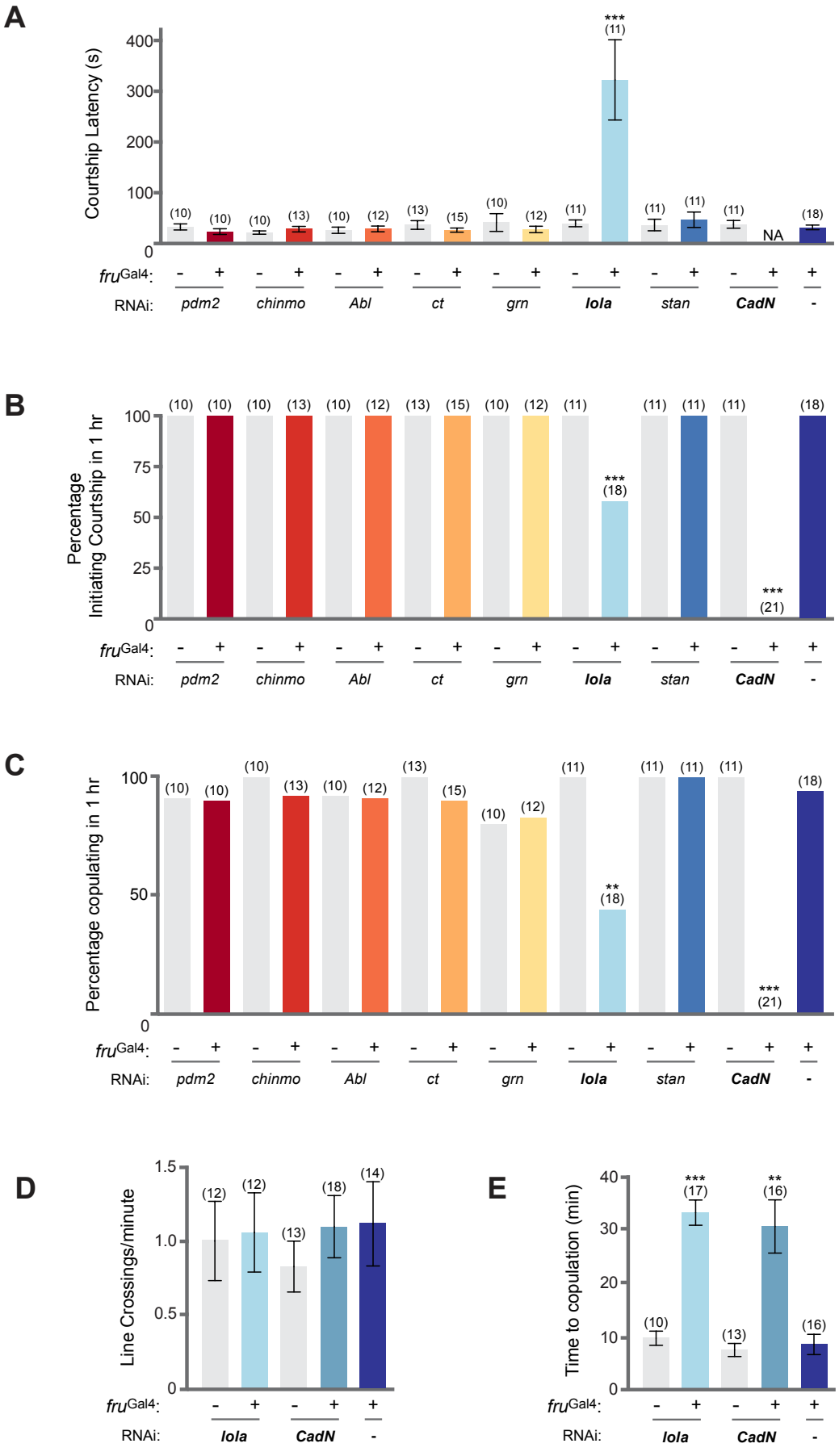

A
Courtship Latency (s)
400
300
200
100
0
(10) (10) (10) (13) (10) (12) (13) (15) (10) (12) (11) *** (11) (11) (11) (11) NA (18)
fruGal4: - + - + - + - + - + - + - + - + - + +
RNAi: pdm2 chinmo Abl ct grn Iola stan CadN -
B
Percentage Initiating Courtship in 1 hr
100
75
50
25
0
(10) (10) (10) (13) (10) (12) (13) (15) (10) (12) (11) *** (18) (11) (11) (11) *** (21) (18)
fruGal4: - + - + - + - + - + - + - + - + - + +
RNAi: pdm2 chinmo Abl ct grn Iola stan CadN -
C
Percentage copulating in 1 hr
100
50
0
(10) (10) (10) (13) (10) (12) (13) (15) (10) (12) (11) ** (18) (11) (11) (11) *** (21) (18)
fruGal4: - + - + - + - + - + - + - + - + - + +
RNAi: pdm2 chinmo Abl ct grn Iola stan CadN -
D
Line Crossings/minute
1.5
1.0
0.5
0
(12) (12) (13) (18) (14)
fruGal4: - + - + +
RNAi: Iola CadN -
E
Time to copulation (min)
40
30
20
10
0
(10) *** (17) (13) ** (16) (16)
fruGal4: - + - + +
RNAi: Iola CadN -

Figure S7. Fru^{MB} target genes *lola* and *CadN* are required in *fru* neurons for normal male courtship behaviour, related to Figure 7.

UAS-RNAi lines against selected Fru^{MB} targets genes were expressed specifically in *fru* neurons using fru^{GAL4} [S4]. All genotypes indicated are males; target females are wild-type. Error bars indicate ± SEM. *n* values shown in parentheses.

(A) Courtship Latencies, ***$p < 0.001$ (Kruskal-Wallis ANOVA test). The following RNAi lines were used:

(B) Percentage males initiating courtship in 1 hr, ***$p<0.0001$ (Fisher Exact Test).

(C) Percentage males mating in 1 hr, ****$p < 0.0001$ (Fisher Exact Test).

(D) Locomotion. No significant changes in locomotion were detected in the *lola* and *CadN* RNAi males shown in Figure 7A (Kruskal-Wallis ANOVA test).

(E) Delayed time to copulation using additional *lola* and *CadN* RNAi lines; **$p < 0.01$, ***$p < 0.001$ (Kruskal-Wallis ANOVA test).

Table S1. Pairwise comparisons between Fru^{M} datasets using Genome Association Tool (GAT). Log2 fold values were plotted in Figure 5A (see experimental procedures).

Table S2. Lists of targets genes associated with the significantly occupied-regions for each Fru^{M} DamID dataset. Gene lists generated at the described false discovery rate (FDR) for each isoform (Fru^{MA}, Fru^{MB} and Fru^{MC}) are shown at each developmental stage (larvae, pupae and adult). Each gene is shown as a secondary identifier, symbol and full name. Related to Figure 5.

Table S3. Gene ontology analysis of each Fru^{M} DamID dataset. Individual tables are included for each dataset. GO terms are listed along with their ID (GO_ID), the level associated with each term and the term description. *p*-values associated with

enrichment/depletion are shown (P_VALUE) along with a list of the genes from the dataset associated with each term (GENES_IN_SET). Related to Figure 5.

Table S4. MICRA analysis of Fru^{M} DamID Genomic Occupancy. The full results from the MICRA analysis of 8mer motifs, as described in experimental procedures, are shown. Related to Figure 6.

Table S5. Full i-*cis*Target Motif analysis of Fru^{M} DamID datasets. Motif IDs are shown along with the enrichment scores for each motif in each dataset. Fru^{MA} datasets are shown in blue, Fru^{MB} (and Fru^{MBmut}) in red, and Fru^{MC} in green. Related to Figure 6.

Table S6. CRMs associated with the putative Fru^{MB}-motif (flyfactorsurvey-ttk-PA_SANGER) in the Fru^{MB} and Fru^{MC} datasets. CRMs are shown as Region IDs, along with their rank and associated genes. Related to Figure 6.

Table S7. Expression of FlyLight Lines in *fru* neurons. The FlyLight lines examined along with their associated GMR numbers and associated genes. Observed expression of FlyLight-Gal4 in *fru* neurons in male and female CNSs is shown as Yes (Y) or No (N). Sexually dimorphic expression is shown as Yes (Y) or No (N). Related to Figure 7.

Supplemental Experimental Procedures

Anti-Fru^{B} Antibody Production

Fru Exon B was PCR amplified with forward primer ZnFB-for (GGG G**GA ATT C**TC CAA GGC CTG GCA CAT GCG G) containing a *EcoR*I site (bold) and reverse primer ZnFB-rev (CCG **CTC GAG** GAC GCC AGG CGC CGT AGT TCC) containing a *Xho*I site (bold) and cloned into complementary sites of the GST expression plasmid pGEX-6P1 (Amersham Biosciences). The GST-FruB fusion antigen was purified from *E. coli* following the manufacturers protocol and sent to Diagnostics Scotland for injection into rat.

Generation of *fru$^{\Delta A}$* and *fru$^{\Delta B}$* Mutants by Ends-in Homologous Recombination

fru$^{\Delta A}$ and *fru$^{\Delta B}$* mutants were generated at the *fru* locus by ends-in homologous recombination as previously described [S5] and as illustrated in Figure S1. Three fragments with homology to the *fru* genomic region were cloned sequentially into the targeting vector pED7 (provided by B. Dickson) using the following primers (5' to 3'): Fragment I GGA TCC CAC TAA ACG AGA ATC AAA AAC and TCT AGA GAG AAA GAC CCT GCT GGC TAT CAG, Fragment II CCT AGG CTT AAA GTG GCT CTT CGG AGG and GGA TCC GAT TCC AGC TTC TGA TAT CCT A, Fragment III GAG GCC GCA CAA AAC TTT CAC CCA ACT CAA TA and GAG GCC GCT TGT CTG CCA GGG GTT TTT CAG. Annealed complementary primers were then used to introduce an additional *I-CreI* site into the *BamH*I site of the targeting vector. The resulting construct was injected into a *w^{1118}* strain. One of the ensuing second chromosome transformant lines, was crossed (2500 virgin females) to males of the genotype *y,w/Y,hs-hid; Sco,hs-I-SceI,hs-FLP/CyO*; first instar larvae were heat shocked for 1.5 hrs at 38°C on the 3rd day following the cross and again on the following day for 1hr. The following fly stocks were used to identify and balance all third chromosomal *white^{+}* recombinants: *y,w; ey-FLP, y,w,ey-FLP; Pin/CyO, y,w,ey-FLP;;Ly/TM3,Sb* (provided by B. Dickson). Approximately 200,000 flies were

screened, and 25 independent targeted recombination events were recovered. PCR, sequencing, and restriction enzyme digestion analysis were used to confirm the predicted recombination event in multiple lines (data not shown). Using a *hs-CreI* expressing line (*w*^{1118}; *P{hs-ICreI.R}1A,Sb1/TM6*), we then reduced the duplication at the *fru* locus, and screened for the inclusion either of the two desired outcomes: the inclusion of the *Sal*I site into the *fru* A exon or the inclusion of the *Nhe*I site into the *fru* B exon. PCR, restriction enzyme digestion and sequencing analysis were performed to ensure the isolation of the desired *fru* mutant. Two independently isolated *fru*$^{\Delta A}$ and *fru*$^{\Delta B}$ mutants were back-crossed eight times into a wild-type *Canton S* genetic background. After initial expression and behavioral analysis showed not differences between the independent mutants (data not shown), one mutant was chosen for each isoform for further analysis. Genomic DNA was extracted from individual homozygous *fru*$^{\Delta A}$ adult or *fru*$^{\Delta B}$ pupae, PCR amplified using the following primers: FruAF: 5'-CGA AAT GGA AAT GTC CCG AAT C-3', FruAR: 5'-GCT ATT GAA ACC GAG ACT GAA CCC-3' (for *fru*$^{\Delta A}$) and FruBF: 5'-TTC CCA TCC CGT TCA AAA GTG-3', FruBR: 5'-GCC GAC GAG GAC ATT AGG-3' (for *fru*$^{\Delta B}$). Purified PCR products were directly sequenced or subjected to restriction enzyme digestion analysis.

Generation of Dam-*fru*^{M} Constructs.

To generate the pUAST-NDam-*fru*$^{MA/MB/MC}$ DNA constructs, a 1092 bp fragment containing the 5′ coding end of *fru* was PCR amplified using a *fru* male cDNA template [S6] with forward primer (5′-GGG **AGA TCT** GAT GAT GGC GAC GCT ACA GGA T-3′) containing a *Bgl*II site (bold), and reverse primer (5′-CGA GGA CAG CGA GTT ATC AC-3′), this primer sequence is downstream of an endogenous *Pvu*I site in *fru.* The resulting *Bgl*II-*Pvu*I fragment was then ligated to individual *Pvu*I-*Kpn*I isoform-specific 3′ end fragments excised from previously described pUAS-*fru*$^{MA/MB/MC}$ constructs [S7]. Final constructs were generated via a three-way ligation into the pUAST-NDam vector [S8]. To generate pUAS-

NDam-*fru*^{mutMB}, a Quikchange™ SDM kit (Stratagene) was used to generate site-directed mutants (Cys>Ala) in three cysteine codons of the C_2H_2 Zn-finger. Transgenic flies for pUAST-NDam--*fru*$^{MA/MB/MC}$ (*UAS-NDam-fru*^{MA}, *UAS-NDam-fru*^{MB} and *UAS-NDam-fru*^{MC}), pUAS-NDam-*fru*^{mutMB} (*UAS-NDam-fru*^{mutMB}) and pUAST-NDam (*UAS-NDam*), were generated as described previously [S9].

CNS Tissue Dissection and DNA Extraction.

Larvae: L3 'wandering' larvae were collected and sexed. Male CNSs were dissected in DamID buffer (5 mM EDTA, 100 mM Tris-HCl pH 8.5, 137 mM NaCl), and flash frozen on dry ice (~ 45 larval CNSs were used per experimental replicate). Pupae: L3 wandering larvae were sexed and males were collected and transferred to a Petri dish lined with Whatman paper soaked in 1xPBS. After 46 hrs, pupal CNSs were dissected in DamID buffer and flash frozen on dry ice (~ 45 pupal CNSs were used per experimental replicate). Adult: Virgin males were isolated from females and left to mature in small groups of 3-5 for 48 hrs. At this point, the flies were anesthetised with CO_2 gas, and dipped briefly in ethanol before CNSs were dissected in DamID buffer, and flash frozen on dry ice (~ 30 adult brains and VNC were used per experimental replicate). DNA was extracted from pooled dissected CNSs using the Qiagen DNEasy Blood and Tissue kit. Due to our inclusion of EDTA in the DamID dissection buffer (used to inhibit DNase activity), RNaseA was omitted from the initial DNA extraction procedure, and was instead added after genomic DNA was ethanol precipitated and resuspended in 100µl dH_2O (containing 0.5µg/ml RNaseA).

Digestion and Genomic DNA processing for DamID

For selective PCR amplification of methylated DNA fragments, 2.5 µg of genomic DNA for each replicate was digested for 16 hr at 37°C with ten units *Dpn*I (NEB) in a total volume of 100 µl buffer 4 (NEB). After inactivation of *Dpn*I at 80°C for 20 min, *Dpn*I and RNaseA

were removed using the QIAquick PCR purification kit (Qiagen) and eluted in 30 µl dH_2O. 1.25 µg of the *Dpn*I digested genomic DNA was ligated to 40 pmol of a double-stranded unphosphorylated adaptor (top strand, 5′-CTA ATA CGA CTC ACT ATA GGG CAG CGT GGT CGC GGC CGA GGA-3′; bottom strand, 5′-TCC TGC GCC-3′) for 16 hr at 16°C with five units T4-DNA Ligase (Roche) in a total volume of 20µl ligation buffer (Roche). To prevent amplification of DNA fragments containing unmethylated GATCs, the adaptor-ligated DNA was cut with five units *Dpn*II (NEB) for 1 hour at 37°C in a total volume of 80 µl *Dpn*II buffer (NEB). Next, amplification was performed with 20 µl *Dpn*II-cut DNA (313 ng), 1.6 µl PCR Advantage enzyme mix (Clontech), 16 nmoles of dNTPs and 100 pmoles primer (5′-GGT CGC GGC CGA GGA tc-3′) in 80 µl total volume of PCR Advantage reaction buffer under the following cycling conditions: activation of the polymerase and nick translation for 10 min at 68°C, followed by one cycle of 1 min at 94°C, 5 min at 65°C, and 15 min at 68°C; three cycles of 1 min at 94°C, 1 min at 65°C, and 10 min at 68°C; and 19 cycles of 1 min at 94°C, 1 min at 65°C, and 2 min at 68°C. Final PCR samples were pooled and purified (Qiaquick PCR purification). Samples were eluted in 25µl dH_2O and quantified using a NanoDrop (Thermo Scientific). Amplified DNA from experimental and Dam-only controls were labelled with either Cy3 or Cy5 and co-hybridized to NimbleGen 2.1M *Drosophila* whole genome tiling arrays (performed at Roche NimbleGen, Inc., USA). The exception being Fru^{MC} pupae, which was the first experiment to be completed, and was performed on a customised tiling array [S10]. For the 2.1 M arrays, probe length was 45bp, for the custom array they were 60bp. In each instance, one of the three replicates was dye swapped.

Normalization and Peak Finding

To determine meaningful conclusions from these datasets it was important to carefully normalise the data, removing false positives and identifying a reliable set of targets for

each isoform at each developmental time point, before proceeding to bioinformatic analysis. NimbleGen files were analyzed using the open-source R package Ringo [S11]. All analyses were based on Release 5 of *Drosophila melanogaster* genome. Quantile Normalization was performed across all samples and controls together (quant-all). Probe intensities were smoothed into window scores using a half window size of 300 bp (400 bp for Fru^{MC} pupae), replacing intensities with the median inside the sliding window. Peaks were identified as five (three for Fru^{MC} pupae) or more probes identified at the respective FDR, with a maximum gap of 200bp separating peaks (350bp for Fru^{MC} pupae). All datasets were analyzed at a 1% FDR, with the exception of Fru^{MA} pupae and adult which were analyzed at an 11% or 13% FDR, respectively. This accommodated the lower number of identified peaks in these data sets, as too few peaks would skew future analysis. Binding profiles were visualized with the Integrated Genome Browser browser (IGB). For visualization purposes IGB generated figures have been shown as log2 transformed on a scale between 0 and 2.5. After analysis with Ringo, FDR FBgn lists were matched to gene positions to provide a list of genes associated with peaks. For each significant bound-region, surrounding target genes (FlyBase genes from UCSC database) were assigned to the bound-region. A gene was assigned to a bound-region if it directly overlapped with the region; otherwise the closest gene was assigned to the region. To determine the closest gene, the genomic distance between the centre of the bound-region and the end of each annotated gene 3′ or 5′ to the peak was used.

GO Enrichment Analysis

Genes were functionally classified with Gene Ontology terms using GOToolbox. Over- or under-representation of the GO terms was statistically determined using the binomial test and *p*-values corrected for multiple testing using the Bonferroni method. A corrected *p*-value better than 0.05 was regarded as significant. Full analysis is presented in Table S3.

The GO clustering heatmap was generated using the heatmap.2 function in the gplot package of the stats package R (http://stat.ethz.ch/R-manual/R-patched/library/stats/html/hclust.html).

Network Analysis

Cytoscape 3.0.0 was used to generate gene networks using the GeneMania (*D. melanogaster* Version 2012-08-02) database application. After imputing the appropriate gene list, interaction networks were created using the input networks: genetic interactions, physical interaction and predicted; weighting was set at automatic. The largest connected network was visualized using the following topological parameters: edge colour/width represents edge 'betweenness' (a score which measures the number of shortest path flows through a certain edge), node size represents 'betweenness centrality' (reflects the amount of control that this node exerts over the interactions of other nodes in the network).

MICRA analysis

For each binding site identified by DamID, 1000 bp of sequence was extracted and filtered for DNA sequences that possessed phastCon scores [S12] equal to or greater than 0.6 across at least 8bp. Then every 8 mer in the conserved DNA sequences was logged and its relative frequency (occurrence every 1000 bp) calculated by comparing it to its background frequency in all non-exonic DNA (pre-computed) [S13]. To generate position weight matrices, sequences from the top 10 enriched motifs were outputted with a padding of 10 bp each side, for input into MEME.

i-*cis*Target Analysis

The identification of enriched DNA motifs along with corresponding cis-regulatory modules (CRMs) was carried out using i-*cis*Target (http://med.kuleuven.be/lcb/i-cisTarget).

Significant peak regions for each dataset were input (.bed files). These peak regions are automatically mapped to the candidate regulatory regions defined in i-*cis*Target (136K regions covering the entire non-coding genome). For all parameters the default settings were used. Motifs identified with an enrichment score >3 are shown in Table S5.

Fru^{M} DamID Dataset Comparisons and Genomic Region Enrichment Analysis

Genome Association Tool (GAT) [S14] was used to test the overlap between genomic regions enriched in binding of the different Fru isoforms, pairwise comparisons were made between datasets. To account for the differences in peaks called across experiments we randomly sampled. A random sampling approach of datasets using the smallest size, was taken to test the significance of overlap between two datasets (data shown in Table S1) and to compare their overlap relative to a genomic background. We used genomic region enrichment analysis in GAT to test for the enrichment of Fru^{M} peak segments. Associations were inferred between Fru^{M} peak segments across the following annotation sets: obtained from the USCS Genome Browser: Transcriptions Starts Site (TSS) + 1kb upstream segments, Exon segments, Intron segments and Intergenic segments. All genome annotations were downloaded from the USCS Genome Browser [S15].

Histology

Immunohistochemistry of *Drosophila* brains was carried out as previously described [S9] using the following antisera: anti-GFP (rabbit polyclonal; Molecular Probes) and nc82 (mouse monoclonal; Developmental Studies Hybridoma Bank) at the dilution of 1:1000 and 1:10, respectively, for 72-96 hr at 4 ºC. After washing, the tissues were treated with Alexa Fluor 488 goat anti-rabbit IgG (Invitrogen) and Alexa Fluor 546 goat anti-mouse IgG (Invitrogen) at a 1:500 dilution for 48 hr at 4 ºC. The observation of the MOL was carried

out as per [S16]. Images were acquired with a Leica SP5 II confocal microscope. Figures and illustrations were assembled using Adobe Photoshop and Adobe Illustrator.

Behavioral Assays

Flies were raised at 25°C in a 12hr:12hr light:dark cycle. Individual virgin males were collected and aged for 5-7 days post-eclosion at 25°C and assays carried out at 25°C. To score courtship, individual naïve males were introduced into a round chamber (19 mm diameter × 4 mm height) with an individual wild-type Canton-S female. The following parameters were measured during a 60 min observation period (or until mating occurred): courtship initiation (first bout lasting over 3 seconds that included two or more behavioral displays including following, tapping, wing extension, licking, and attempted copulation), time to copulation (in minutes), copulation success (% copulating in 1 hr), and courtship index. The wing extension index (WEI) was measured as the percent of the time during courtship the male unilaterally extended his wing. Locomotion was measured with individual males during a 20 min observation period. In order to accurately measure the WEI, high-resolution courting pairs were filmed in small round observation chambers (10 mm diameter x 3 mm height). Fertility was measured as per [S6].

Statistics

Behavioral means were compared using Kruskal-Wallis ANOVA test and Dunn's post hoc statistical test where indicated. For Fisher's exact test, two-tail *p* values were compared with controls. Statistical analyses were performed with GraphPad Prism (version 6.0).

Song analysis

For song analysis virgin males were reared in isolation for 2-8 days, at 23°C in a 12hr:12hr light:dark cycle. Males were recorded when courting a sexually immature Canton-S female

with her wings removed (to prevent female rejection sounds). Recordings were made for four minutes with two INSECTAVOX microphones [S17, S18] using Audacity software (Audacity-win-1.2.6, available from: http://audacity.sourceforge.net/download/windows) after bandpassing at around 150 Hz to 1 kHz. Temperature inside the recording chamber was monitored continuously with a digital thermometer (mean 24.5^{o}C, SD ±1.1). Song parameters (Interpulse Interval, IPI, Intrapulse frequency, IPF and Sine song frequency, SSF) were measured using customised routines within the DataView analysis package (version 8.6 available from: http://www.st-andrews.ac.uk/~wjh/dataview/), after Fourier transform for frequency variables. No parameters were found to vary with male age, but significance among mutations was assessed after adjusting for recording temperature (significant for IPI ($p<0.001$) and IPF ($p=0.007$) and microphone (significant for SSF, $p=0.014$).

D. melanogaster strains

The Janelia Farm FlyLight-Gal4 lines and the following *UAS*-TRiP were obtained from the Bloomington *Drosophila* Stock Center: *Abl* (TRiP.HMS02275); *pdm2* (TRiP.HMC03066); *stan* (TRiP.HMS01464); *chinmo* (TRiP.HMS00036); *ct* (TRiP.HMS00924); *grn* (TRiP.HMS00980); *lola* (TRiP.GLV21087) and (TRiP.JF02254); *CadN* (TRiP.HMS02380) and (TRiP.JF02653). The *fru*^{GAL4} and *UAS>stop>mCD8::GFP; fru*^{FLP} lines were generously provided by B. Dickson (IMP).

Accession numbers

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE52247.

Supplemental References

S1. Kutter, C., Watt, S., Stefflova, K., Wilson, M.D., Goncalves, A., Ponting, C.P., Odom, D.T., and Marques, A.C. (2012). Rapid turnover of long noncoding RNAs and the evolution of gene expression. PLoS genetics *8*, e1002841.

S2. Luo, S.D., Shi, G.W., and Baker, B.S. (2011). Direct targets of the D. melanogaster DSXF protein and the evolution of sexual development. Development *138*, 2761-2771.

S3. Huson, D.H., Richter, D.C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: An interactive viewer for large phylogenetic trees. BMC bioinformatics *8*, 460.

S4. Stockinger, P., Kvitsiani, D., Rotkopf, S., Tirian, L., and Dickson, B.J. (2005). Neural circuitry that governs Drosophila male courtship behavior. Cell *121*, 795-807.

S5. Rong, Y.S., and Golic, K.G. (2000). Gene targeting by homologous recombination in Drosophila. Science *288*, 2013-2018.

S6. Billeter, J.C., Villella, A., Allendorfer, J.B., Dornan, A.J., Richardson, M., Gailey, D.A., and Goodwin, S.F. (2006). Isoform-specific control of male neuronal differentiation and behavior in Drosophila by the fruitless gene. Current biology : CB *16*, 1063-1076.

S7. Song, H.J., Billeter, J.C., Reynaud, E., Carlo, T., Spana, E.P., Perrimon, N., Goodwin, S.F., Baker, B.S., and Taylor, B.J. (2002). The fruitless gene is required

for the proper formation of axonal tracts in the embryonic central nervous system of Drosophila. Genetics *162*, 1703-1724.

S8. Choksi, S.P., Southall, T.D., Bossing, T., Edoff, K., de Wit, E., Fischer, B.E., van Steensel, B., Micklem, G., and Brand, A.H. (2006). Prospero acts as a binary switch between self-renewal and differentiation in Drosophila neural stem cells. Dev Cell *11*, 775-789.

S9. Rideout, E.J., Dornan, A.J., Neville, M.C., Eadie, S., and Goodwin, S.F. (2010). Control of sexual differentiation and behavior by the doublesex gene in Drosophila melanogaster. Nat Neurosci *13*, 458-466.

S10. van Steensel, B., and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nature biotechnology *18*, 424-428.

S11. Toedling, J., Sklyar, O., and Huber, W. (2007). Ringo - an R/Bioconductor package for analyzing ChIP-chip readouts. BMC bioinformatics *8*.

S12. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M.M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research *15*, 1034-1050.

S13. Southall, T.D., and Brand, A.H. (2009). Neural stem cell transcriptional networks highlight genes essential for nervous system development. The EMBO journal *28*, 3799-3807.

S14. Heger, A., Webber, C., Goodson, M., Ponting, C.P., and Lunter, G. (2013). GAT: a simulation framework for testing the association of genomic intervals. Bioinformatics.

S15. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome research *12*, 996-1006.

S16. Nojima, T., Kimura, K., Koganezawa, M., and Yamamoto, D. (2010). Neuronal synaptic outputs determine the sexual fate of postsynaptic targets. Current biology : CB *20*, 836-840.

S17. Gorczyca, M.G., and Hall, J.C. (1987). The Insectavox, an integrated device for recording and amplifying courtship songs of Drosophila. Dros Inf Serv *66*, 157-160.

S18. Ritchie, M.G., Yate, V.H., and Kyriacou, C.P. (1994). Genetic-Variability of the Interpulse Interval of Courtship Song among Some European Populations of Drosophila-Melanogaster. Heredity *72*, 459-464.