oskar acts with the transcription factor Creb to regulate long-term memory in crickets

Arpita Kulkarni[1,2], Ben Ewen-Campen[1,3], Kanta Terao[4,5,], Yukihisa Matsumoto[4], Yaolong Li[6], Takayuki Watanabe[4,7], Jonchee A. Kao[8], Swapnil S. Parhad[9,10], Guillem Ylla[1,11], Makoto Mizunami[4] and Cassandra G. Extavour[1,8,12*]

1. Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, USA
2. Current address: Single Cell Core, Harvard Medical School, Boston, USA
3. Current address: Department of Genetics, Harvard Medical School, Boston, USA
4. Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan
5. Current address: Faculty of Liberal Arts, Tokyo Medical and Dental University, Tokyo, Japan
6. Graduate School of Life Science, Hokkaido University, Sapporo, Japan
7. Research Center for Integrative Evolutionary Science, School of Advanced Sciences, Sokendai-Hayama, Shonan Village, Hayama 240-0193, Japan
8. Department of Molecular and Cellular Biology, Harvard University, Cambridge, USA
9. University of Massachusetts Chan Medical School, Worcester, USA
10. Current address: Department of Cell Biology, Harvard Medical School, USA
11. Current address: Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Krakow, Poland
12. Howard Hughes Medical Institute, Chevy Chase, Maryland, USA

* correspondence to extavour@oeb.harvard.edu

Abstract

Novel genes have the potential to drive the evolution of new biological mechanisms, or to integrate into pre-existing regulatory circuits and contribute to the regulation of older, conserved biological functions. One such gene, the novel insect-specific gene *oskar*, was first identified based on its role in establishing the *Drosophila melanogaster* germ line. We previously showed that this gene likely arose through an unusual domain transfer event involving bacterial endosymbionts, and played a somatic role before evolving its well-known germ line function. Here, we provide empirical support for this hypothesis in the form of evidence for a novel neural role for *oskar*. We show that *oskar* is expressed in the adult neural stem cells of a basally branching insect, the cricket *Gryllus bimaculatus*. In these stem cells, called neuroblasts, *oskar* is required together with the ancient animal transcription factor *Creb* to regulate long-term (but not short-term) olfactory memory. We provide evidence that *oskar* positively regulates *Creb*, which plays a conserved role in long-term memory across animals, and that *oskar* in turn may be a direct target of Creb. Together with previous reports of a role for *oskar* in nervous system development and function in crickets and flies, our results are consistent with the hypothesis that *oskar*'s original somatic role may have been in the insect nervous system. Moreover, its co-localization and functional cooperation with the conserved pluripotency gene *piwi* in the nervous system may have facilitated *oskar*'s later co-option to the germ line in higher insects.

Keywords: *oskar*, memory, crickets, Creb, co-option, neuroblast, mushroom body, Kenyon cells

Introduction

Oskar (*osk*) is an insect-specific gene first discovered in *Drosophila melanogaster*, where it plays a critical role in germline specification[1]. Oskar mRNA is localized to the posterior of the developing *D. melanogaster* oocyte[2,3]. Local translation and anchoring of Oskar (Osk) protein leads to the posterior accumulation of the mRNA and protein products of several genes with conserved expression and function in animal germ lines, including *vasa* and *piwi*[2,4,5]. Collectively called germ plasm, these cytoplasmic contents act as necessary and sufficient determinants to specify embryonic germ cells[2,3]. The current model of Osk function in *D. melanogaster* germ plasm assembly is that it serves as a scaffolding protein, facilitating the assembly of the ribonucleoprotein complexes that contain germ plasm components[2,6,7].

Interestingly, *osk* and several other genes originally identified as *D. melanogaster* germ line genes, including *vasa*, *pumilio*, *staufen*, *orb*, and *piwi*-related genes including *aubergine* and *argonaute 3*, have since been shown to have a variety of roles in animal nervous systems[8–14]. For example, in *D. melanogaster*, *osk* RNAi in larval dendritic arborization neurons disrupts *nanos* mRNA localization, ultimately leading to a defect in dendrite morphogenesis and an associated defect in motor response to mechanical stimulation[12]. Furthermore, *osk* plays a role in the embryonic nervous system, but not in the germ line, in a basally branching hemimetabolous insect, the cricket *Gryllus bimaculatus*, where it is important for proper neuroblast divisions and subsequent axonal patterning[15]. Our recent analysis of hundreds of previously unidentified *osk* orthologs across insects showed that *osk* is expressed in at least a dozen somatic tissues in species across the insect tree[16]. This suggests that a somatic function of *osk* may be ancestral. However, the precise roles of *osk* in any somatic tissue, including the nervous system, remain largely unknown.

Here, we demonstrate a role for *osk* in the adult brain of the cricket *Gryllus bimaculatus*, in a population of neural stem cells in the mushroom body that persist throughout adult life. We show that *osk*, as well as Piwi and Vasa, are enriched in a population of adult neuroblasts in the mushroom body, and that RNA

interference (RNAi) targeting *osk* or *piwi* in adult crickets impairs long-term, but not short-term memory formation in an olfactory associative learning assay[17]. We also provide evidence that *osk* and *piwi* function in a regulatory feedback loop with the cyclic AMP response element binding protein (Creb), a transcription factor with well described conserved roles in long-term memory across metazoans[18]. Our data demonstrate a novel somatic role for *osk*, and shed light on how a novel gene may acquire critical roles by integrating with pre-existing gene regulatory systems comprising older, conserved genes.

Results and Discussion

osk is expressed in adult neuroblasts of the mushroom body

We previously showed that neuroblasts in the cricket embryo express *osk*, *vasa* and *piwi*, and that *osk* is required for correct neuroblast division and embryonic nervous system morphology[15]. Interestingly, in many insects, including crickets, a subset of embryonic neuroblasts persist throughout adulthood and continuously give rise to new neurons called Kenyon cells that comprise the mushroom body[19–21]. This contrasts with flies like *D. melanogaster*, in which neuroblasts die prior to adulthood[22], and in which adult brains are thus essentially devoid of neurogenesis[23] (although there are reports of potential stem cells in adult *D. melanogaster* brains[23,24] which may be damage-dependent rather than homeostatic in function[25], and which remain controversial[26].)

Given the role of *osk* in embryonic neuroblasts of crickets, we asked whether *osk* also plays a role in the adult mushroom body neuroblasts. We used *in situ* hybridization to examine *osk* expression in the adult brain, and found expression in a cluster of cells with the large, round nuclei and diffuse chromatin characteristic of stem cells, at the apex of each of the two lobes of the mushroom body, consistent with descriptions of adult neuroblasts in Orthopterans (Fig. 1A). EdU co-localization (Fig.1B) confirmed the identity of these cells as neuroblasts, the only proliferative cells in the adult brain[27]. We also found that mushroom body neuroblasts express high levels of Vasa and Piwi proteins (Fig. 1E).

Previous research has shown that mushroom body neuroblasts play an important role in long-term olfactory memory formation in Orthoptera[28]. Scotto-Lomassese[28] found that ablation of mushroom body neuroblasts using irradiation led to a dramatic reduction in olfactory, but not visual, learning after 24- and 48- hours, suggesting that newborn mushroom body neurons produced by those neuroblasts play a role in forming new olfactory memories. We therefore sought to test whether *osk*, expressed specifically in mushroom body neuroblasts, functions in these cells in the context of long-term memory formation.

We first tested whether Osk regulates the proliferation or survival of adult mushroom body neuroblasts. Using an established technique for systemic RNAi in the adult cricket brain[29], we injected double stranded *osk* RNA (dsRNA) into the head capsule, and confirmed efficiency of *osk* knockdown via qPCR and small RNA profiling of osk^{RNAi} brains (Fig. 2A, 2D, Suppl. Tables S1, S2, Suppl. Fig. S1, S2). osk^{RNAi} adult brains showed no gross anatomical defects relative to controls (data not shown). Moreover, neither the total number of neuroblasts ($p<0.05$), nor the number of neuroblasts undergoing mitosis as revealed by Edu labeling ($p<0.05$), were statistically significantly different between osk^{RNAi} adult brains and controls (Fig. 1C, 1D). We stained osk^{RNAi} and control brains with Cleaved Caspase-3, a marker for apoptosis, and did not observe any evidence of cell death (Suppl. Fig. S3A). We noted that one described role for PIWI in the *Drosophila* germ line is to prevent DNA damage caused by transposon mobilization[30]. However, we observed no detectable increase in γH2A staining, a marker for DNA damage, in osk^{RNAi} brains (Suppl. Fig. S3B). These data suggest that *osk* is not required for the proliferation, survival or genomic integrity of adult neuroblasts. However, the specific expression of *osk*, Piwi, and Vasa in the mushroom body neuroblasts suggested that some or all of these genes could play a role related to memory or learning.

***oskar* RNAi impairs long-term but not short-term memory**

The mushroom body is the anatomical substrate for olfactory memory and learning in insects[27,31,32], and ablation of the mushroom body or of the adult mushroom body neuroblasts impairs these processes[28,33,34]. Based on previous observations that mushroom body neuroblasts play a role in long-term olfactory memory formation in crickets[28], we hypothesized that *osk* might play a role in this process. To test this hypothesis, we assessed memory of osk^{RNAi} adult male crickets at one hour ("short-term memory") and one day ("long-term memory") post-training using well-established cricket olfactory behavior assays[17]. In control crickets (injected with dsRNA targeting DsRed[35]), four training sessions led to a significant ($p<0.05$) short-term preference for the rewarded odor (peppermint) at one hour after training (short-term). Trained control crickets retained this learned preference ($p<0.01$) even at one day after training (Fig. 2A, "DsRed" panel), demonstrating that long-term memory formation is intact in these controls. However, although osk^{RNAi} crickets formed and retained memory for the rewarded odor by one hour after training (short-term; Fig.2A *osk* dsRNA #1, $p<0.001$) this memory was lost by one day after training (long-term; Fig. 2A *osk* dsRNA #1, $p>0.05$), indicating a specific impairment of long-term memory formation. These results were reproducible in a second experiment using a non-overlapping fragment of *osk* dsRNA (Fig. 2A, *osk* dsRNA #2; $p<0.05$ for short-term, $p>0.05$ for long-term), suggesting the impact was specific to *osk* knockdown. Efficacy of the knockdown was confirmed via qPCR and small RNA sequencing (Suppl. Fig. S1, S2, Suppl. Tables S1, S2), indicating that *osk* is required for cricket long-term memory.

Since both *piwi* and *vasa* were co-expressed with *osk* in cricket mushroom body neuroblasts[15] (Fig. 1E), we also assessed the role of these two genes in olfactory memory. We found that *piwi* (Fig. 2B, $p<0.001$) but not *vasa* (Suppl. Fig. S4A, $p>0.05$; Suppl. Fig. S4B) was also required for cricket long-term memory. qPCR analyses showed that *osk* RNAi led to a significant decrease in *piwi* transcript levels (Fig. 2D), suggesting that *osk* positively regulates *piwi* in the cricket brain. However, *osk* transcript levels remained unaffected in $piwi^{RNAi}$ animals (Fig. 2E). Consistent with the phenotype of the single gene knockdowns, $osk^{RNAi}/piwi^{RNAi}$ double knockdown adults also showed a long-term memory impairment phenotype (Fig.

2C, p<0.01). Thus, *osk* and *piwi* do not globally disrupt olfaction, learning or short-term memory formation, but are required for consolidation of long-term memory in this species.

osk and *piwi* positively regulate the nuclear transcription factor *CrebA*

To understand how a novel gene like *osk* might have gained a role in an ancient animal function like long-term memory consolidation, we investigated the hypothesis that it might interact with conserved regulators of animal memory. To test this hypothesis, we asked whether we could detect a functional or regulatory interaction between *osk* and a highly conserved transcription factor with well-documented roles in long-term memory formation across animals, cyclic AMP response element binding protein (Creb)[18]. We first identified putative *Creb* orthologs in the *G. bimaculatus* genome[36] using a combination of BLAST searches and phylogenetic analyses (Fig. 3A, Suppl. Table S3). These analyses yielded two high-confidence *Creb* orthologs, which we called *CrebA* and *CrebB* based on their closest *D. melanogaster Creb* gene relative (Fig. 3A). Analysis of previously generated transcriptomes[37] showed that both genes are expressed in adult cricket brains (Suppl. Table S4). We performed *CrebA* and *CrebB* RNAi experiments and discovered that *CrebA* was required for long-term memory in crickets (Fig. 3B; p<0.003 and p<0.001 for dsRNA#1 and dsRNA#2 respectively; Suppl. Fig. S4C). Using quantitative PCR (qPCR), we then asked whether transcript levels of this memory regulator were altered in *osk* or *piwi* knockdown conditions, and found consistent downregulation of *CrebA* transcript levels in both single and double RNAi backgrounds (Fig. 2D-F). In contrast, and consistent with the observation that $vasa^{RNAi}$ had no long-term memory impact (Suppl. Fig.1B), qPCR revealed no reduction of *CrebA* transcripts in $vasa^{RNAi}$ conditions (Suppl. Fig. 1 C). This suggests that the long-term memory defects observed in osk^{RNAi} and $piwi^{RNAi}$ conditions (Fig. 2D-F) are due to a downregulation of *CrebA* in these animals.

osk and *piwi* are regulated by *CrebA*

Creb proteins are transcription factors that bind cyclic AMP response element (CRE) binding sites within the regulatory regions of target genes to initiate transcription[18] (Fig. 4A). Since target gene transcription

and new protein synthesis is crucial for long-term memory formation, and given the similarity in long-term memory phenotypes of osk^{RNAi}, $piwi^{RNAi}$ and $CrebA^{RNAi}$ animals, we asked whether *osk* and *piwi* might also be *Creb* target genes in this cricket. qPCR revealed that transcript levels of both *osk* and *piwi* are significantly decreased in $CrebA^{RNAi}$ conditions (Fig. 4B), suggesting that *osk*/*piwi* and CrebA may interact in a positive feedback loop to promote each other's transcription. To evaluate the possibility that *osk* or *piwi* might be direct transcriptional targets of CrebA, we examined the genomic sequences flanking both loci and found two bioinformatically predicted CRE binding sites within the 6kb upstream of the transcription start sites for *osk* (Fig. 4C) and *piwi* (Fig. 4D). These predicted CRE binding were found twice as frequently as we would expect to find such sequences in a randomly generated sequence of this length (see Methods). Electrophoretic mobility shift assays showed that protein(s) within the adult cricket brain bind specifically to the predicted CRE sites of both of *osk* and *piwi* (Fig. 4C, D). Given the current lack of species-specific CrebA reagents for this cricket species, we cannot rule out the interpretation that a protein(s) other than CrebA present in the adult cricket brain is causing the observed mobility shift by binding the predicted CRE sites of *osk* and *piwi*. However, given our functional data indicating that RNAi against *osk*, *piwi* and *CrebA* all yield long-term memory defects (Figs. 2, 3B), that *osk* and *piwi* transcript levels are reduced in $CrebA^{RNAi}$ brains (Fig. 2B), and that *osk* and *piwi* genomic loci contain predicted CRE binding sites (Fig. 4C, D), the results of our gel shift assays are consistent with the hypothesis that cricket *osk* and *piwi* are direct transcriptional targets of CrebA.

We have discovered a new role for *oskar* in the adult cricket brain (Fig.1). We have shown that *osk*, Piwi, and Vasa are co-expressed in mushroom body neuroblasts (Fig. 1A, B), a population of stem neural cells required for long-term olfactory memory formation[28], and that knockdown of *osk* and *piwi* disrupt olfactory long-term memory formation. The precise role that the mushroom body neuroblasts play in memory formation remains unknown, as does the molecular role of Osk in these cells. In *D. melanogaster*, where there are no adult neural stem cells in the mushroom body, olfactory long-term memory requires the ~2,500 differentiated Kenyon cell neurons of the mushroom body[38], which respond

with high selectivity to a small number of stimuli, allowing the mushroom body to house an explicit representation of a large number of olfactory cues[38,39]. Specific olfactory stimuli are associated with learned behavioral responses via specific sets of neurons connecting the mushroom body to other brain regions in a protein synthesis-dependent fashion, to form long-term memories[9,40]. Thus, one possibility is that adult-born Kenyon cells in *G. bimaculatus* (and other insects that display adult neurogenesis in the mushroom body) are recruited into an existing circuit and allow for a constantly increasing repertoire of olfactory associations. Our results suggest that *osk* could play a role in this process, as *osk* RNAi disrupts long-term memory. We note that of the two mammalian brain regions known to undergo adult neurogenesis, one (the subventricular zone) contributes to the olfactory bulb, and neurogenesis in this region is involved in olfactory memory[41].

Given that adult *D. melanogaster* lacks the mushroom body neuroblasts seen in *G. bimaculatus*[42], a straightforward test for a directly comparable *osk* function in this fruit fly is not possible. However, although *D. melanogaster* mushroom body stem cells are absent in adults, analogous mushroom body neuroblasts remain mitotically active late into pupal development[43]. Thus, it will be interesting to test whether *osk* functions in these neuroblasts during larval and/or pupal stages. We note that an insertion of an enhancer trap transposable element over 3 kb upstream of the *osk* transcription start site in *D. melanogaster* was recovered in an insertional mutagenesis screen for long term memory in this fruit fly[8]. However, this insertion has not been confirmed as compromising the sequence or function of the *osk* locus, nor has *osk* been tested directly to confirm a potential role in *D. melanogaster* learning or memory.

Although *D. melanogaster* lacks adult mushroom body neuroblasts, it is possible that *osk* could function in fruit fly olfactory long-term memory in a neuroblast-independent manner. A recent study of the mushroom body output neurons has suggested that long-term memory involves the activity-dependent de-repression of mRNAs localized to granules containing Pum, Staufen, and Orb proteins[9]. Given that Osk nucleates similar granules containing these proteins in the *Drosophila* oocyte[44,45], and Osk's ability to

nucleate phase-transitioned granules in *D. melanogaster* cells[46], it would be interesting to test whether Osk is involved in the formation and/or activity of these granules in the brain. Given our recent observations of highly similar molecular interactions of conserved molecules and animal germ lines and neural cells[14], future studies could test whether additional, traditionally known "germ line" genes other than *vasa* and *piwi* also function in *G. bimaculatus* adult neuroblasts, which would suggest that *osk* acts with conserved molecular partners in different cellular contexts.

Both germ cells and neuroblasts are stem cells that give rise to highly specialized, postmitotic daughter cells, while remaining proliferative for long periods of time. Thus, the original role of *osk* in both cell types could conceivably be related to stem cell maintenance and/or asymmetric division. Indeed, a variety of highly conserved "germ line genes" including *vasa*, *nanos*, and *piwi* are found in a variety of multipotent cells in diverse animals[14,47], raising the possibility that such genes were involved in establishing multipotency rather than specifying germ cell fate *per se*. A broader understanding of the function(s) of *osk* thus requires additional studies of phylogenetically diverse insects, as well as further detailed biochemical analysis in the context of *Drosophila* germ cells and neurons.

Our results provide an example of how newly arisen genes may find stable homes in pre-existing genetic regulatory circuits. In the case of *osk*, we hypothesize that by evolving binding sites responsive to the conserved transcription factor Creb, *osk* may have gained expression in the brain, opening the door for potential participation in neural roles. Future studies will be needed to elucidate the molecular mechanisms of *osk* gene products in the cricket brain, and specifically in learning and memory. We speculate that the biophysical properties of Osk protein that make it effective at sequestering RNAs and participating in translational control in the germ line[48–50], may have been advantageous in promoting the rapid translation needed the synaptic plasticity that underlies learning and memory. These include Osk's ability to form phase transitioned condensates[46,51], its regions of high predicted disorder[16,46,51], and its ability to achieve and maintain asymmetric subcellular localization, all of which are well known in the

germ line, and may have provided a selective advantage to Osk in the context of promoting neuronal function.

Materials and Methods

G. bimaculatus *husbandry*

For behavior experiments, *G. bimaculatus* crickets were maintained in the Mizunami laboratory at 27ºC on a 12:12 light cycle, with a diet of insect food pellets, as previously described[52]. For gene expression analysis, quantitative PCR, and cell proliferation experiments, crickets were maintained in the Extavour laboratory at 28ºC and 35% relative humidity on a 12:12 light cycle, with a diet of grain and cat food, as previously described[53].

In situ hybridization

For in situ hybridization, brains were dissected and de-sheathed in ice-cold 1x Phosphate Buffered Saline (1X PBS) as previously described[29,54]. Brains were fixed one hour in 4% paraformaldehyde in 1X PBS, followed by an additional overnight fixation in the same solution at 4ºC, or for an additional 3-4 hours at room temperature. *osk* transcripts were detected using a 788 bp probe, following standard protocols[53] with the following modifications to reduce background: 20-minute Proteinase K (Thermo Fisher Scientific, Cat# EO0491) treatment followed by a 30-minute fixation in 0.8% glutaraldehyde in 1X PBS and 4% paraformaldehyde in 1X PBS. The *osk* probe was used at 1.0 ng/µl concentration and hybridized at 69-70ºC. Brains were sectioned after in situ development was completed, by embedding in 4% low-melt agarose in distilled water, and sectioning at 50-90µM using a Leica VT1000S vibratome.

Immunostaining

For immunostaining, primary antibodies used were as follows: rabbit anti-Gb-Vasa and anti-Gb-Piwi[55] were used at 1:300, mouse anti-RNA polymerase II pSer 6 Mab H5 (Covance MMS-129R) 1:100, FITC-

conjugated anti-alpha Tubulin (Sigma F2168) 1:100 and rabbit anti-*Drosophila* Vasa (kind gift of Paul Lasko, McGill University) 1:500 following standard procedures as previously described[53]. Goat anti-rabbit secondary antibodies conjugated to Alexa 488, Alexa 555 or Alexa 568 (Invitrogen) were used at 1:500 or 1:1000. Counterstains used were Hoechst 33342 (Sigma B2261) at 0.1 to 0.05 mg/ml and FITC-conjugated phalloidin (Sigma P5282) at 1 U/ml. For antibody staining, brains were embedded in 4% low-melt agarose in distilled water, and sectioned at 50-90μM using a Leica VT1000S vibratome, prior to incubation with the primary antibody. For in situ hybridization, brains were sectioned after staining had been completed. For analysis of apoptosis and DNA damage, brains were fixed 4h after EdU injection and sectioned via vibratome, with EdU detection performed first (using Invitrogen's Click-iT protocol), followed by antibody staining and confocal analysis.

RNA extractions, cDNA preparation and cloning of gene fragments for production of dsRNA

Brains from unmated male adults were dissected in ice-cold 1X PBS, then immediately homogenized in TRIzol (ThermoFisher Scientific, catalog number 15596026). Total RNA was extracted following the manufacturer's instructions, including a 30-minute DNAse treatment. 1μg of total RNA was used as template for cDNA synthesis using SuperScript III (ThermoFisher Scientific, catalog number 18-080-044) with oligo-dT primers. cDNA was diluted 1:10 prior to PCR with gene specific primers, and 2 μL of template was used per 25 μL PCR reaction. PCR products were run on a 1% agarose gel and desired bands were gel eluted following IBI Scientific's PCR purification and gel elution kit (catalog number IB47030). Then, products were cloned into Zero blunt TOPO PCR cloning kit (ThermoFisher Scientific, catalog number 450245) using electro-competent *DH5α-E* cells (ThermoFisher Scientific, catalog number 11319019).

***Identification and annotation of Piwi Proteins in* G. bimaculatus**

We previously used a *G. bimaculatus* transcriptome[56] to identify two RNA fragments corresponding to two Piwi family proteins (*piwi*: JQ434103 and *piwi-2*: KC242806.1[15,55]). All previous published analyses

of *piwi* in *G. bimaculatus* were performed with "*piwi*" (JQ434103), as only this gene showed enriched expression in embryonic germ cells[55]. Since the time of our initial studies on *piwi*, we assembled and annotated an updated *G. bimaculatus* genome[36]. For the present study, we therefore performed new BLAST searches to clarify the status of *piwi* orthologs in this cricket (Suppl. Fig. S5). We found both previously identified fragmented *piwi* RNA sequences within the new gene annotations[36] with gene IDs GBI_17641 (containing "*piwi*" fragment JQ434103) and GBI_07509 (containing "*piwi-2*" fragment KC242806.1) respectively. We also identified two additional putative novel *piwi*-like genes annotated in the *G. bimaculatus* genome, with gene IDs GBI_09750 and GBI_09796[36].

Using InterProscan, we confirmed that the amino acid sequences of the proteins encoded by these four putative *piwi* genes contained the typical characteristics of the Argonaut/Piwi proteins[57] namely a Paz domain followed by a C-terminal Piwi domain. Additionally, we inferred the gene tree of the Argonaute/Piwi protein family using the putative *G. bimaculatus* Piwi and Argonaute protein amino acid sequences obtained from the genome, together with sequences of publicly available Piwi and Argonaute proteins from other insects (*Drosophila melanogaster*, *Apis mellifera*, *Bombyx mori*, *Tribolium castaneum*, *Blattella germanica*, *Zootermopsis nevadensis*, *Acyrthosiphon pisum*, and *Locusta migratoria)*. Protein sequence alignments were performed with MUSCLE[58,59] in Geneious (v3.8.425; www.geneious.com), and the gene tree was inferred with RAxML v8.2.11[60] with 100 bootstrap iterations to obtain the support values of each node. The tree was then visualized with ggtree[61,62]. The resulting tree differentiated four groups of sequences with bootstrap values above 90%, each of which contained different Piwi/Argonaute subfamilies as follows: Argonaute 1 proteins (AGO1; included GBI_02015), Argonaute 2 (AGO2; included GBI_13717), Argonaute 3 (AGO3; included GBI_01357), and Piwi proteins and their paralogs (Aub and Siwi, the Piwi paralogs in *D. melanogaster* and *B. mori* respectively; included GBI_17641) (Suppl. Fig. S5). This indicated that our previous analyses[55] had indeed targeted the true *piwi* ortholog in *G. bimaculatus*. Accordingly, all gene expression and function analyses in the present study were also performed on this true *piwi* ortholog (GBI_17641).

Analysis of the small RNA species present in *piwi*^{RNAi} animals indicated that our *piwi* RNAi experiments specifically targeted the true *piwi* (GBI_17641), and did not affect or impact expression of the other 3 *piwi* subfamily genes (Suppl. Table S2). The quality control of the eight small RNA-seq samples (3 *piwi*^{RNAi}, 2 *osk*^{RNAi}, 2 *DsRed*^{RNAi}, and 1 untreated or wild-type, WT) was performed with FastQC v0.11.8[63], and the adapters were trimmed with Cutadapt v1.8.1[64]. Clean reads were mapped to the *G. bimaculatus* genome assembly[36] with Bowtie 2 v2.3.4.1[65] using parameters "-L 18, -N 0". The numbers of sequenced reads, clean reads, and mapped reads are shown in Suppl. Table S1. The mapped reads were retrieved using samtools v1.9[66] for obtaining the read length distributions (Suppl. Fig. S2). The proportion of miRNAs and piRNAs in each sample was extrapolated as the percentage of reads of 22-23 nts and 28-29 nts respectively (Suppl. Fig. S2). The FeatureCounts function from the R package Rsubread v2.0.0[67] was used to count the number of reads mapped to all annotated genes and build a table of counts. The counts of o*sk* (GBI_0140) and the four *piwi* genes (GBI_09750, GBI_09796, GBI_07509, and GBI_17641) were retrieved (Suppl. Table S2). The small RNA-seq reads mapping to the target genes were assumed to be reads of the siRNAs produced from the dsRNA[68]. The thousands of such reads mapping to our targeted *piwi* (GBI_17641) and the absence of such siRNA reads mapping to other p*iwi* genes, suggest that no off-target effects impaired the expression of other *piwi* genes.

RNA interference

Unmated adult male crickets within one week of their final molt were injected with 2 µL of double-stranded RNA (dsRNA) through a hole pierced in the median ocellus[69] using a 10 µL syringe fitted with 26S gauge tip (WPI, Tokyo, Japan; Hamilton Inc., Nevada, USA). Behavioral tests were repeated using two non-overlapping fragments of *oskar* (742bp and 503bp), a 646bp fragment of *piw*i, a 541bp fragment of *vasa*, two non-overlapping fragments of *CrebA* (both 387bp fragments), a 384bp fragment of *CrebB*, and a 678bp fragment of *DsRed* as a negative control[15]. Double-stranded RNA concentrations used were

10 µM for *oskar*; 3.38 µg/µL for *piwi*; 2.71 µg/µL for *vasa*; 2.97µg/µL for *DsRed*; 6 µg/µL for *oskar/piwi* double knock down*;* 7µg/µL for *CrebA;* 7µg/µL for *CrebB; and* 7µg/µL for *DsRed* (Suppl. Table S5).

Construction, sequencing and analysis of small RNA libraries from* G. bimaculatus *adult brains

Unmated adult male crickets within one week of their final molt to adulthood were injected with dsRNA as described above (see *RNA interference)*. At 48h post RNAi injections, brains (Suppl. Table S5) were dissected in ice-cold 1x PBS, and transferred into Trizol, following which total RNA was extracted from them following manufacturers protocols. Next, RNA was size selected for 18-30nt size range after denaturing PAGE. A 2S rRNA specific oligo was used for 2S rRNA depletion. The small RNAs were ligated at the 3' and 5' ends by the respective adapters and purified by denaturing PAGE after each ligation. PCR was performed after reverse transcription. The PCR product was gel purified from an agarose gel to obtain the final library. The libraries were sequenced using Illumina NextSeq500 1x75bp. The resulting data were uploaded onto NCBI SRA database and are publicly available under the BioProject ID PRJNA837371.

Gene expression from mRNA-seq data

To check the expression of *G. bimaculatus* genes in nervous systems in wild-type animals, previously generated mRNA-seq libraries were used[37] and the complete CDS for genes of interest were obtained from the recently published genome[33]. Reads were trimmed with Cutadapt v3.4[64], and mapped to the full *G. bimaculatus* CDS using Geneious Read Mapper[70]. DESeq2 normalized counts in fragments per kilobase per exon per million mapped fragments (FPKMs) were then obtained for all genes of interest[73]. The expression of *osk* and *piwi* genes (Suppl. Table S4) shows that the depleted *piwi* (GBI_17641) is one of the two *piwi* genes expressed in wild-type brains of males and females. In the same way, we obtained the expression in FPKMs of *vasa* (GBI_17344), *CrebA* (GBI_04244), and *CrebB* (GBI_02305) (Suppl. Table S4).

379

380 ***Olfactory learning behavioral memory assays***

381 Adult male crickets at eight days after the final molt were used in all experiments. Three days before
382 conditioning, individual crickets were separated into 100-mL beakers and deprived of drinking water to
383 enhance their motivation to search for water. Two days before conditioning (ten days after the imaginal
384 molt), each cricket was injected with dsRNA as described above. Two days after dsRNA injection, each
385 cricket was subjected to an odor preference test, in which the animal was allowed to freely visit
386 peppermint and vanilla odors[71]. The time spent at each of the peppermint and vanilla odor sources was
387 measured cumulatively to evaluate relative odor preference[71]. Crickets were subjected to 4-trial
388 conditioning, in which an odor was paired with water reward, with an inter-trial interval of five
389 minutes[29,72]. For conditioning, a small filter paper was attached to the needle of a hypodermic syringe.
390 The syringe was filled with water reward (unconditioned stimulus), and the filter paper was soaked with
391 peppermint essence (conditioned stimulus). At one hour and one day after the end of the conditioning,
392 each cricket was subjected to an odor preference test. The relative odor preference of each conditioned
393 and control animal was measured using the preference index (PI) for rewarded odor (peppermint), defined
394 as tP/(tP+tV) *100 (%), where tP is the time spent exploring the peppermint source and tV is the time
395 spent exploring the vanilla (unrewarded) odor. Wilcoxon's test was used to compare odor preferences
396 before and after training. For multiple comparisons, Holm’s method was used to adjust the significance
397 level.

398

399 ***Quantitative PCR (qPCR)***

400 Two days after dsRNA injection, brains were dissected from unmated male adults (within a week of final
401 molt into adulthood) in ice-cold 1x PBS, then immediately homogenized in TRIzol (ThermoFisher
402 Scientific, catalog number 15596026). Total RNA was extracted from a total of six brains per treatment,
403 following the manufacturer’s instructions, including a 30-minute DNAse treatment. 1 µg of total RNA
404 was used as template for cDNA synthesis using SuperScript III (ThermoFisher Scientific, catalog number

18-080-044) with oligo-dT primers. cDNA was diluted 1:10 prior to qPCR, and 6 µL of template was used per 25 µL qPCR reaction. (PerfeCta SYBR Green SuperMix, Low ROX, Quanta Biosciences, catalog number 101414-158). qPCR reactions were conducted in triplicate, and fold change was calculated using the delta-delta Ct method[73], with standard deviation propagated following standard methods. *Beta-tubulin* was used as a reference gene[53]. Primers amplifying a 130bp fragment of *Gb-CrebA*, a 140bp fragment of *Gb-CrebB*, a 234bp fragment of *Gb-oskar*, a 166 bp fragment of *Beta-tubulin*[53], a 129bp fragment of *Gb-piwi*, a 150bp fragment of *Gb-vasa* and a 120bp fragment of *Gb-FGFR* (Fibroblast Growth Factor Receptor) were used (Suppl. Table S6).

EdU Assay

Cell proliferation was assayed using the Click-iT EdU Alexa 488 kit (Life Technologies, Cat# C10637). Crickets were injected with 10-15µl of EdU either into the abdomen or into the head capsule through the median ocellus (both methods successfully labeled dividing neuroblasts), and brains were dissected to visualize EdU incorporation four hours post-injection. Brains were dissected and de-sheathed in ice-cold 1x PBS. Calyces were removed with a microscalpel and incubated in 0.1M citric acid for 15-30 minutes on a poly-lysinated slide (Sigma Aldrich, Cat. No. P8920-100ML). Calyces were then spread into a monolayer by adding a Sigma cote-covered coverslip, and the entire slide was flash-frozen in liquid nitrogen. The coverslip was removed, leaving the mushroom body monolayer on the slide. Slides were air-dried and were then fixed for 15 minutes in 4% PFA. EdU detection was then carried out following manufacturer's instructions. EdU-positive cells were photographed under epifluoresence on a Zeiss AxioImager Z.1 compound microscope using Zen, and manually quantified in ImageJ. For any mushroom body where the EdU-positive cluster of cells was damaged or destroyed during preparation, that sample was discarded and not included in the analysis. For tissue double stained to visualize transcripts and EdU incorporation simultaneously, the in situ hybridization was conducted before the visualization of incorporated EdU. AxioImager Z.1, LSM 780 or LSM 880 confocal microscopes (Zeiss) were used for microscopy, driven by AxioVision or Zen (Zeiss).

Identification of G. bimaculatus *Creb genes and construction of Creb phylogenetic tree*

Putative orthologs of *Creb/ATF* family members from several animal species were initially identified by BLAST searches and then downloaded from NCBI. These sequences were then used to search for putative *G. bimaculatus CrebA* orthologs in the G. *bimaculatus* genome[36]. All identified sequences were then aligned with MAFFT (v 7.510)[74]. A maximum likelihood tree was created in RAxML using the PROTGAMMAWAG model[60] , and plotted with the FigTree package v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree) (Fig.3B, Suppl. Table S3).

Bioinformatic prediction of CRE sites in osk *and* piwi *upstream regulatory regions*

A position frequency matrix (PFM) for the full cyclic AMP response element (CRE) octameric palindrome was obtained from the JASPAR database (an open source database for transcription factor binding sites[75] (Suppl. File 1). In addition to CRE, PFMs for the TATA box were also obtained from the same database. We included TATA box proximity among our search criteria for putative CRE sites, since TATA boxes are often a feature of functional promoters, and functional promoter-proximal CRE sites are reported as often occupied by Creb. These raw PFM data (Suppl. File 1) were then used as an input in FIMO (Find Individual Motif Occurrence) in the MEME suite (a motif-based sequence analysis tool[76], and up to 10Kb of the gene sequences upstream of the predicted transcription start site for each of *G. bimaculatus oskar*, *piwi*, and *vasa* was scanned for the presence of the CRE and TATA motifs using the annotated *G. bimaculatus* genome[36], and using $p<0.0001$ as the stringency criteria. For comparison, *G. bimaculatus beta actin*, *alpha tubulin* and *FGFR* loci were subjected to the same analyses (also with $p<0.0001$ as stringency criteria) to assess the possibility that any *G. bimaculatus* gene would be predicted to have CRE sites in the 10Kb region upstream of their transcription start site using this method (Suppl. File 2). We found that for the latter three genes, there we no CRE predictions in their upstream regions (up to 10Kb from the transcription start site). Further, we bioinformatically generated one thousand 10Kb long DNA fragments of random sequence using the "random DNA sequence" tool in the Sequence

Manipulation Suite[77] and then tested them for CRE prediction. Our results indicate that a CRE site is expected to occur in a randomly generated sequence at a frequency of ~1.6 CRE sites for every 10Kb tested (Suppl. Files 1,2).

PCR amplification, sequence confirmation and cloning of CRE sites

Based on bioinformatic predictions of putative CRE sites, primers were designed in the upstream regulatory regions of *osk* (Suppl. Table S7 #1 and #2) and-*piwi* (Suppl. Table S7 #3 and #4). Once all four CRE sites were sequence confirmed by Sanger sequencing (2 CRE sites/gene), the ~30bp fragments containing each CRE site were synthetically generated as duplexes (with 3'A overhangs) for use as EMSA pre-probes (CRE site in bold) (Suppl. Table S7). The 3'A overhangs were then used to clone all EMSA pre-probes into a pGEM-T easy vector following manufacturer's instructions (Promega, catalog number A1360) using One-Shot chemically competent TOP10 *E. coli* cells (Thermo-Fischer, catalog number C4040-06).

Generation of 5'Cy5 labelled Electrophoretic Mobility Shift Assay (EMSA) Probes and EMSA

Once cloned, pGEM-T easy specific duplex forward primer (5'Cy5-ACGTCGCATGCTCCCGGCCATG, reverse complement 5'Cy5-CATGGCCGGGAGCATGCGACGT), and reverse primer (5'Cy5-GTCGACCTGCAGGCGGCCGCGAATT, reverse complement 5-Cy5-AATTCGCGGCCGCCTGCAGGTCGAC) were designed with 5'Cy5 modifications to amplify inserts and generate fluorescently labeled double stranded EMSA probes, using a two-step PCR program with the following conditions: (98°C for 60 seconds (x1cycle); 98°C for 15 seconds followed by 72°C for 30 seconds (x30 cycles); 72°C for 5 minutes (x1cycle) (Suppl. Table S7). The PCR product was loaded onto a 1% Agarose Gel and the desired bands were gel eluted following IBI Scientific's PCR purification and gel elution kit (catalog number IB47030) in 30µl water. A second round of PCR amplification following the conditions described above was performed using the eluted DNA from previous steps to increase probe yield. All steps starting with the first round of PCR were done in the dark to protect fluorescently

labeled probes. Probe concentrations were measured using Nanodrop, and diluted to a final concentration of 40 fmol/probe for use in EMSA[78]. 20% native PAGE gels were used to study gel shifts. Gels were imaged using an Azure Sapphire Biomolecular Imager (VWR).

Nuclear Protein extracts from unmated adult male G. bimaculatus *brains*

Brains were dissected from unmated *G. bimaculatus* males within one week of their final molt that were anesthetized briefly on ice prior to dissection in 1x PBS. Nuclear protein extracts were prepared from dissected brains following manufacturer's instructions (Abcam Nuclear Extraction Kit, catalog number ab113474).

Acknowledgements

Thanks to Taro Mito (University of Tokushima, Japan) for advice on cricket brain dissection; Venkatesh Murthy (Harvard University) for allowing us to use his laboratory's vibratome; Elena Kramer and Min Ya (Harvard University) for help with polyacrylamide gel electrophoresis setup for EMSAs; William E. Theurkauf (University of Massachusetts Chan Medical School) for guidance on cricket small RNA library preparation and sequencing; members of the Extavour lab for discussion; and the Faculty of Arts and Sciences Bauer Molecular Biology Core Facility at Harvard University for Illumina Sequencing.

Funding

This study was supported by National Science Foundation award #IOS-0817678, funds from Harvard University and from the Howard Hughes Medical Institute to CGE, a National Science Foundation Graduate Research Training Fellowship to BEC, and a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Culture, Sports, and Technology of Japan (no. 21K19245) to MM.

Author Contributions

AK performed all qPCRs, bioinformatic analysis of CRE sites, gel shift assays, cricket brain, nerve cord and other tissue transcriptomes referenced in this study, and prepared small RNA libraries with the assistance of SP in consultation with WET; BEC performed whole mount gene expression and EdU assays; KT, YM, YL, TW performed olfactory learning assays, GY performed small RNA library analysis and the cricket Piwi/AGO identification and phylogenetic analysis, JK provided technical support for gel shift assays; MM and CGE conceived of the project and experimental design; AK and CGE wrote the manuscript with input from all co-authors; WET, MM and CGE obtained funding for the project.

Figure Legends

Figure 1. ***oskar*, Piwi and Vasa are expressed in *G. bimaculatus* adult mushroom body neuroblasts.** (**A**) In situ hybridization on adult *G. bimaculatus* brains detects *osk* transcripts in the cells of the mushroom body (arrows). Inset in top left panel shows the overall structure of the adult brain; shaded box indicates a single mushroom body lobe, corresponding to the region shown in micrographs in top row. Bottom row: dorsal views of both mushroom body lobes, indicating *oskar* expression in neuroblast clusters (arrows). White asterisks indicate non-specific binding of probe to tracheal remnants in the brain. (**B**) Edu labelling (green) of the adult brain shows that *osk*-expressing cells (grey) are mitotically active, consistent with their identity as neuroblasts. **(C)** Quantification of EdU-positive cells shows no significant difference between osk^{RNAi} and control brains ($p<0.05$). **(D)** Quantification of total number of mitotically active cells shows no significant difference between osk^{RNAi} and control brains ($p<0.05$). Numbers within bars indicate sample sizes in (**C**) and (**D**). **(E)** Detection of Vasa & Piwi proteins in adult mushroom body neuroblasts (MBNBs). Scale bars = 50µm in top panels of (**A)** and (**B)** and in **(E)**, and 200µM in bottom panels of (**A)**.

Figure 2. ***oskar*^{RNAi} and *piwi*^{RNAi} impairs cricket long-term memory.** (**A**) Results of olfactory memory assay in $oskar^{RNAi}$ animals. (**B**) Results of olfactory memory assay in $piwi^{RNAi}$ animals. (**C**) Results of olfactory memory assay in $osk^{RNAi}/piwi^{RNAi}$ double knockdown animals. For each assay, relative preference between the rewarded odor (peppermint) and control odor (vanilla) was tested before training, one hour post-training, and one day post training for DsRed controls and for *osk* (using two different non-overlapping *osk* fragments #1 and #2), *piwi* and *osk/piwi* double RNAi. Boxes represent the 1st and 3rd quartiles surrounding the median (middle line). Whiskers extend to extreme values within 1.5x of interquartile range. Wilcoxon's test was used for comparison of preference before and after conditioning. For multiple comparisons, the Holm method was used to adjust the significance level (* $p < 0.05$, ** $p<0.01$, *** $p<0.001$, n.s. = not significantly different). (**D-F)**. Quantitative qPCR results showing the extent of downregulation of different *G. bimaculatus* genes in *osk*, *piwi* and *osk/piwi* double RNAi backgrounds. Effectiveness of RNAi per background is also shown in each case. Data is plotted as mRNA fold change (+/- standard deviation) based on the ΔΔCt method (* $p < 0.05$, ** $p < 0.01$, *** $p<0.001$, **** $p < 0.0001$, n.s. = not statistically significant).

Figure 3. **Cricket *CrebA* is required for cricket long-term memory.** (**A**) GenBank IDs of *Creb/ATF* family member orthologs in mouse and insects (Suppl. Table S3) were used to construct a *Creb* phylogenetic tree to infer the evolutionary relationships between mammalian *Creb* proteins and their insect counterparts. *G. bimaculatus CrebA* and *CrebB* are indicated in black in the tree (**B**) $CrebA^{RNAi}$ impairs long-term memory formation in crickets. Relative preference between the rewarded odor (peppermint) and control odor (vanilla) was tested before training, one hour post-training, and one day post training for *DsRed* controls and $CrebA^{RNAi}$ (using two different non-overlapping *CrebA* fragments #1 and #2 for independent confirmation). Boxes represent the 1st and 3rd quartiles surrounding the median (middle line). Whiskers extend to extreme values within 1.5x of interquartile range. Wilcoxon's test was used for comparison of preference before and after conditioning. For multiple comparisons, the Holm method was used to adjust the significance level. (* $p < 0.05$, ** $p<0.01$, *** $p<0.001$, n.s. = not significantly different). n=9 for *CrebA* and n=10 for *DsRed*.

Figure 4. Cricket CrebA regulates *oskar*. (**A**) Schematic diagram of the transcription factor cAMP response element binding protein (Creb) protein (top) displaying only the two domains relevant to this study, the Kinase-Inducible Domain (KID) that can facilitate kinase-inducible transcription activation and the basic leucine zipper (bZIP) domain that is important for dimerization and DNA binding. Creb proteins bind the cyclic AMP response element (CRE), a sequence present in the promoter regions of many cellular genes to increase (middle schematic) or decrease (bottom schematic) transcription of target gene. (**B**) Quantitative RT-PCR results showing the expression levels of *G. bimaculatus osk*, *piwi*, and *vasa* in

CrebA RNAi knockdown conditions. The extent of *CrebA* downregulation is also shown to assess efficiency of RNAi knockdown. Data is plotted as mRNA fold change (+/- standard deviation) based on the ΔΔCt method (* $p < 0.05$, ** $p < 0.01$, *** $p<0.001$, * $p < 0.0001$ n.s. = not statistically significant). (**C**) Schematic of *G. bimaculatus oskar* gene (top) showing exons (depicted as rectangular grey boxes) along with its presumptive regulatory region containing two predicted CRE sites, which we call CRE I (~3.6Kb upstream of predicted transcription start site marked by fMet) and CRE II (~5.3Kb upstream of predicted fMet). (Bottom) Electrophoretic Mobility Shift Assay (EMSA) to detect possible Creb binding to *osk's* CRE I (top, in green) and CRE II (bottom, in blue). "*oskar* CRE+ probe" indicates predicted CRE site-containing probe "*oskar* CRE- probe" indicates probe without predicted CRE sites; "cricket brain protein" indicates *G. bimaculatus* brain protein extract; "BSA" indicates 1% BSA as a non-specific protein control. The complete sequences of the EMSA probes used for CRE I and II experiments are indicated in green and blue text underneath the gel image, with black representing the predicted CRE sequence (**D**) Schematic of *G. bimaculatus piwi* gene (top) showing its exons and presumptive regulatory region containing two predicted CRE sites, CRE I (~3.2Kb upstream of predicted fMet) and II (~5.7Kb upstream of predicted fMet). (Bottom) EMSA results showing a mobility shift (marked with pink asterisk) for *piwi* CREs. Lane labels are as indicated in (C). The complete sequence of the EMSA probes are indicated in green (*piwi* CRE I) and blue (*piwi* CRE II) text underneath the gel image, with the CRE sequence in black.

Supplementary Figure Legends

Figure S1. Read Length Distribution: Percentage of mapped reads of each length from 18 to 29 nucleotides in each sequenced small RNA library colored by RNAi treatment. The two peaks at ~22 and ~28 mainly correspond to miRNAs and piRNAs.

Figure S2. miRNAs vs piRNA: Estimated proportion of reads corresponding to microRNAs and piRNAs in each sample (colored by condition) based on the percentage of reads of 22-23 nucleotides in length and 28-29 nucleotides respectively.

Figure S3. Apoptosis marker Cleaved caspase 3 (**A**) and DNA damage marker gamma H2AX immunostaining (**B**) in adult mushroom bodies, including neuroblasts of control and osk^{RNAi} brains.

Figure S4. ((**A**) *vasa* RNAi fails to recapitulate the long-term olfactory memory phenotype seen in *osk* and *piwi* RNAi. (**B**) qPCR on $vasa^{RNAi}$ brains shows significant up-regulation of *CrebA*. (**C**) *CrebB* RNAi does not recapitulate the long-term memory phenotype shown by $CrebA^{RNAi}$. N=9 for *CrebB*, and N=10 for *DsRed.*

Figure S5. *G. bimaculatus piwi* ortholog identification and phylogenetic analysis. Argonaute family gene tree generated with the PIWI, AUB, AGO1, AGO2, and AGO3 protein sequences from *Drosophila melanogaster*, *Apis mellifera*, *Bombyx mori*, *Tribolium castaneum*, *Blattella germanica*, *Zootermopsis nevadensis*, *Acyrthosiphon pisum*, and *Locusta migratoria*. Values at nodes represent bootstrap support, in boxes color-coded from dark (lowest) to light (highest) blue. *G. bimaculatus piwi* and *argonaut* genes indicated by red and green asterisks respectively.

Supplementary Table & File Legends

Table S1. Number of raw sequenced small RNA reads, number and percentage of clean reads, and number and percentage of reads mapped to the *G. bimaculatus* genome.

Table S2. Number of small RNA-seq reads mapped to *osk* and *piwi* genes to show specificity of RNAi knockdowns. These small RNA-seq reads come from the siRNA detected following injections of dsRNA for these respective genes. The targeted sequence in each library is highly enriched by small RNA-seq reads. In the *piwi*^{RNAi} experiments, only the targeted *piwi* (GBI_17641) shows a high number of mapped reads in the libraries generated from animals injected with the dsRNA against *piwi*, indicating that the other three *piwi* orthologs present in the *G. bimaculatus* genome were unlikely to be targeted by our approach.

Table S3. GenBank IDs of *Creb/ATF* family member orthologs in mouse and other insects including the cricket *G. bimaculatus*. This information was used to construct a Creb phylogenetic tree (Fig. 3B) to infer the evolutionary relationships between mammalian Creb proteins and their insect counterparts.

Table S4. Gene expression levels for *osk*, *piwi*, *vasa*, and *CrebA/B* (in FPKM per tissue) from brain and ventral cord transcriptomes of male and female adult *Gryllus bimaculatus*[37]. Gene IDs as per the annotated cricket genome[36].

Table S5. A total of 128 unmated, adult male cricket brains (16 brains from Control 1 "uninjected controls", 37 brains from Control 2 "*DsRed* injected", 45 brains from "*piwi*" dsRNA injected, and 30 brains from "*osk*" dsRNA injected) were dissected 48h post dsRNA injection and processed for making small RNA libraries.

Table S6. Primers used for quantitative PCR of all listed *G. bimaculatus* genes.

Table S7. Primers used for cloning and generation of EMSA probes for *G. bimaculatus oskar* and *piwi.*

Supplementary File 1: The Frequency Matrices (PFM) from JASPAR database used to predict CRE sites and TATA boxes in the presumptive regulatory regions of *G. bimaculatus* genes.

Supplementary File 2: FASTA file containing the simulated one thousand 10-Kb long DNA fragments generated to test the frequency of occurrence of CRE sites in randomly generated sequences.

References

1. Lehmann, R., and Nüsslein-Volhard, C. (1986). Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in drosophila. Cell *47*, 141–152. 10.1016/0092-8674(86)90375-2.

2. Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). oskar organizes the germ plasm and directs localization of the posterior determinant nanos. Cell *66*, 37–50. 10.1016/0092-8674(91)90137-n.

3. Kim-Ha, J., Smith, J.L., and Macdonald, P.M. (1991). oskar mRNA is localized to the posterior pole of the Drosophila oocyte. Cell *66*, 23–35. 10.1016/0092-8674(91)90136-m.

4. Hay, B., Jan, L.Y., and Jan, Y.N. (1990). Localization of vasa, a component of Drosophila polar granules, in maternal-effect mutants that alter embryonic anteroposterior polarity. Development *109*, 425–433.

5. Lasko, P.F., and Ashburner, M. (1990). Posterior localization of vasa protein correlates with, but is not sufficient for, pole cell development. Gene Dev *4*, 905–921. 10.1101/gad.4.6.905.

6. Jones, J.R., and Macdonald, P.M. (2007). Oskar controls morphology of polar granules and nuclear bodies in Drosophila. Development *134*, 233–236. 10.1242/dev.02729.

7. Lehmann, R. (2016). Germ Plasm Biogenesis - An Oskar-Centric Perspective. Curr Top Dev Biol *116*, 679–707. 10.1016/bs.ctdb.2015.11.024.

8. Dubnau, J., Chiang, A.-S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., et al. (2003). The staufen/pumilio Pathway Is Involved in Drosophila Long-Term Memory. Curr Biol *13*, 286–296. 10.1016/s0960-9822(03)00064-2.

9. Pai, T.-P., Chen, C.-C., Lin, H.-H., Chin, A.-L., Lai, J.S.-Y., Lee, P.-T., Tully, T., and Chiang, A.-S. (2013). Drosophila ORB protein in two mushroom body output neurons is necessary for long-term memory formation. Proc National Acad Sci *110*, 7898–7903. 10.1073/pnas.1216336110.

10. Perrat, P.N., DasGupta, S., Wang, J., Theurkauf, W., Weng, Z., Rosbash, M., and Waddell, S. (2013). Transposition-Driven Genomic Heterogeneity in the Drosophila Brain. Science *340*, 91–95. 10.1126/science.1231965.

11. Wharton, R.P., Sonoda, J., Lee, T., Patterson, M., and Murata, Y. (1998). The Pumilio RNA-Binding Domain Is Also a Translational Regulator. Mol Cell *1*, 863–872. 10.1016/s1097-2765(00)80085-4.

12. Xu, X., Brechbiel, J.L., and Gavis, E.R. (2013). Dynein-Dependent Transport of nanos RNA in Drosophila Sensory Neurons Requires Rumpelstiltskin and the Germ Plasm Organizer Oskar. J Neurosci *33*, 14791–14800. 10.1523/jneurosci.5864-12.2013.

13. Ye, B., Petritsch, C., Clark, I.E., Gavis, E.R., Jan, L.Y., and Jan, Y.N. (2004). nanos and pumilio Are Essential for Dendrite Morphogenesis in Drosophila Peripheral Neurons. Curr Biol *14*, 314–321. 10.1016/j.cub.2004.01.052.

14. Kulkarni, A., Lopez, D.H., and Extavour, C.G. (2020). Shared Cell Biological Functions May Underlie Pleiotropy of Molecular Interactions in the Germ Lines and Nervous Systems of Animals. Frontiers Ecol Evol *8*, 215. 10.3389/fevo.2020.00215.

15. Ewen-Campen, B., Srouji, J.R., Schwager, E.E., and Extavour, C.G. (2012). oskar Predates the Evolution of Germ Plasm in Insects. Curr Biol *22*, 2278–2283. 10.1016/j.cub.2012.10.019.

16. Blondel, L., Besse, S., Rivard, E.L., Ylla, G., and Extavour, C.G. (2021). Evolution of a Cytoplasmic Determinant: Evidence for the Biochemical Basis of Functional Evolution of the Novel Germ Line Regulator Oskar. Mol Biol Evol *38*, 5491–5513. 10.1093/molbev/msab284.

17. Matsumoto, Y., and Mizunami, M. (2004). Context-Dependent Olfactory Learning in an Insect. Learn Memory *11*, 288–293. 10.1101/lm.72504.

18. Silva, A.J., Kogan, J.H., Frankland, P.W., and Kida, S. (1998). CREB and Memory. Neuroscience *21*, 127–148. 10.1146/annurev.neuro.21.1.127.

19. Farris, S.M., and Sinakevitch, I. (2003). Development and evolution of the insect mushroom bodies: towards the understanding of conserved developmental mechanisms in a higher brain center. Arthropod Struct Dev *32*, 79–101. 10.1016/s1467-8039(03)00009-4.

20. Malaterre, J., Strambi, C., Chiang, A., Aouane, A., Strambi, A., and Cayre, M. (2002). Development of cricket mushroom bodies. J Comp Neurol *452*, 215–227. 10.1002/cne.10319.

21. Urbach, R., and Technau, G.M. (2003). Early steps in building the insect brain: neuroblast formation and segmental patterning in the developing brain of different insect species. Arthropod Struct Dev *32*, 103–123. 10.1016/s1467-8039(03)00042-2.

22. Pinto-Teixeira, F., Konstantinides, N., and Desplan, C. (2016). Programmed cell death acts at different stages of Drosophila neurodevelopment to shape the central nervous system. Febs Lett *590*, 2435–2453. 10.1002/1873-3468.12298.

23. Simões, A.R., and Rhiner, C. (2017). A Cold-Blooded View on Adult Neurogenesis. Front Neurosci-switz *11*, 327. 10.3389/fnins.2017.00327.

24. Fernández-Hernández, I., Rhiner, C., and Moreno, E. (2013). Adult Neurogenesis in Drosophila. Cell Reports *3*, 1857–1865. 10.1016/j.celrep.2013.05.034.

25. Simões, A.R., Neto, M., Alves, C.S., Santos, M.B., Fernández-Hernández, I., Veiga-Fernandes, H., Brea, D., Durá, I., Encinas, J.M., and Rhiner, C. (2022). Damage-responsive neuro-glial clusters coordinate the recruitment of dormant neural stem cells in Drosophila. Dev Cell *57*, 1661-1675.e7. 10.1016/j.devcel.2022.05.015.

26. Li, G., and Hidalgo, A. (2020). Adult Neurogenesis in the Drosophila Brain: The Evidence and the Void. Int J Mol Sci *21*, 6653. 10.3390/ijms21186653.

27. Strausfeld, N.J. (2012). Arthropod Brains: Evolution, Functional Elegance, and Historical Significance (Harvard University Press).

28. Scotto-Lomassese, S., Strambi, C., Strambi, A., Aouane, A., Augier, R., Rougon, G., and Cayre, M. (2003). Suppression of Adult Neurogenesis Impairs Olfactory Learning and Memory in an Adult Insect. J Neurosci *23*, 9289–9296. 10.1523/jneurosci.23-28-09289.2003.

29. Takahashi, T., Hamada, A., Miyawaki, K., Matsumoto, Y., Mito, T., Noji, S., and Mizunami, M. (2009). Systemic RNA interference for the study of learning and memory in an insect. J Neurosci Meth *179*, 9–15. 10.1016/j.jneumeth.2009.01.002.

30. Klattenhoff, C., Bratu, D.P., McGinnis-Schultz, N., Koppetsch, B.S., Cook, H.A., and Theurkauf, W.E. (2007). Drosophila rasiRNA Pathway Mutations Disrupt Embryonic Axis Specification through Activation of an ATR/Chk2 DNA Damage Response. Dev Cell *12*, 45–55. 10.1016/j.devcel.2006.12.001.

31. Belle, J. de, and Heisenberg, M. (1994). Associative odor learning in Drosophila abolished by chemical ablation of mushroom bodies. Science *263*, 692–695. 10.1126/science.8303280.

32. Heisenberg, M., Borst, A., Wagner, S., and Byers, D. (1985). Drosophila Mushroom Body Mutants are Deficient in Olfactory Learning: Research Papers. J Neurogenet *2*, 1–30. 10.3109/01677068509100140.

33. Mizunami, M., Weibrecht, J.M., and Strausfeld, N.J. (1998). Mushroom bodies of the cockroach: Their participation in place memory. J Comp Neurol *402*, 520–537. 10.1002/(sici)1096-9861(19981228)402:4<520::aid-cne6>3.0.co;2-k.

34. Malun, D., Plath, N., Giurfa, M., Moseleit, A.D., and Müller, U. (2002). Hydroxyurea-induced partial mushroom body ablation in the honeybee Apis mellifera: Volumetric analysis and quantitative protein determination. J Neurobiol *50*, 31–44. 10.1002/neu.10015.

35. Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (2002). A monomeric red fluorescent protein. Proc National Acad Sci *99*, 7877–7882. 10.1073/pnas.082243699.

36. Ylla, G., Nakamura, T., Itoh, T., Kajitani, R., Toyoda, A., Tomonari, S., Bando, T., Ishimaru, Y., Watanabe, T., Fuketa, M., et al. (2021). Insights into the genomic evolution of insects from cricket genomes. Commun Biology *4*, 733. 10.1038/s42003-021-02197-9.

37. Whittle, C.A., Kulkarni, A., and Extavour, C.G. (2021). Evolutionary dynamics of sex-biased genes expressed in cricket brains and gonads. J Evolution Biol *34*, 1188–1211. 10.1111/jeb.13889.

38. Heisenberg, M. (2003). Mushroom body memoir: from maps to models. Nat Rev Neurosci *4*, 266–275. 10.1038/nrn1074.

39. Caron, S.J.C., Ruta, V., Abbott, L.F., and Axel, R. (2013). Random Convergence of Olfactory Inputs in the Drosophila Mushroom Body. Nature *497*, 113–117. 10.1038/nature12063.

40. Chen, C.-C., Wu, J.-K., Lin, H.-W., Pai, T.-P., Fu, T.-F., Wu, C.-L., Tully, T., and Chiang, A.-S. (2012). Visualizing Long-Term Memory Formation in Two Neurons of the Drosophila Brain. Science *335*, 678–685. 10.1126/science.1212735.

41. Lazarini, F., and Lledo, P.-M. (2011). Is adult neurogenesis essential for olfaction? Trends Neurosci *34*, 20–30. 10.1016/j.tins.2010.09.006.

42. Cayre, M., Scotto-Lomassese, S., Malaterre, J., Strambi, C., and Strambi, A. (2007). Understanding the Regulation and Function of Adult Neurogenesis: Contribution from an Insect Model, the House Cricket. Chem Senses *32*, 385–395. 10.1093/chemse/bjm010.

43. Ito, K., and Hotta, Y. (1992). Proliferation pattern of postembryonic neuroblasts in the brain of Drosophila melanogaster. Dev Biol *149*, 134–148. 10.1016/0012-1606(92)90270-q.

44. Breitwieser, W., Markussen, F.H., Horstmann, H., and Ephrussi, A. (1996). Oskar protein interaction with Vasa represents an essential step in polar granule assembly. Gene Dev *10*, 2179–2188. 10.1101/gad.10.17.2179.

45. Chang, J.S., Tan, L., and Schedl, P. (1999). The Drosophila CPEB Homolog, Orb, Is Required for Oskar Protein Expression in Oocytes. Dev Biol *215*, 91–106. 10.1006/dbio.1999.9444.

46. Kistler, K.E., Trcek, T., Hurd, T.R., Chen, R., Liang, F.-X., Sall, J., Kato, M., and Lehmann, R. (2018). Phase transitioned nuclear Oskar promotes cell division of Drosophila primordial germ cells. Elife *7*, e37949. 10.7554/elife.37949.

47. Juliano, C.E., Swartz, S.Z., and Wessel, G.M. (2010). A conserved germline multipotency program. Development *137*, 4113–4126. 10.1242/dev.047969.

48. Macdonald, P.M., Kanke, M., and Kenny, A. (2016). Community effects in regulation of translation. Elife *5*, e10965. 10.7554/elife.10965.

49. Rongo, C., Gavis, E.R., and Lehmann, R. (1995). Localization of oskar RNA regulates oskar translation and requires Oskar protein. Dev Camb Engl *121*, 2737–2746.

50. Kulkarni, A., and Extavour, C.G. (2017). Convergent evolution of germ granule nucleators: A hypothesis. Stem Cell Res *24*, 188–194. 10.1016/j.scr.2017.07.018.

51. Bose, M., Lampe, M., Mahamid, J., and Ephrussi, A. (2022). Liquid-to-solid phase transition of oskar ribonucleoprotein granules is essential for their function in Drosophila embryonic development. Cell *185*, 1308-1324.e23. 10.1016/j.cell.2022.02.022.

52. Matsumoto, Y., and Mizunami, M. (2000). Olfactory learning in the cricket Gryllus bimaculatus. J Exp Biology *203*, 2581–2588. 10.1242/jeb.203.17.2581.

53. Kainz, F., Ewen-Campen, B., Akam, M., and Extavour, C.G. (2011). Notch/Delta signalling is not required for segment generation in the basally branching insect Gryllus bimaculatus. Development *138*, 5015–5026. 10.1242/dev.073395.

54. Hamada, A., Miyawaki, K., Honda-Sumi, E., Tomioka, K., Mito, T., Ohuchi, H., and Noji, S. (2009). Loss-of-function analyses of the fragile X-related and dopamine receptor genes by RNA interference in the cricket Gryllus bimaculatus. Dev Dynam *238*, 2025–2033. 10.1002/dvdy.22029.

785 55. Ewen-Campen, B., Donoughe, S., Clarke, D.N., and Extavour, C.G. (2013). Germ Cell Specification
786 Requires Zygotic Mechanisms Rather Than Germ Plasm in a Basally Branching Insect. Curr Biol *23*,
787 835–842. 10.1016/j.cub.2013.03.063.

788 56. Zeng, V., Ewen-Campen, B., Horch, H.W., Roth, S., Mito, T., and Extavour, C.G. (2013).
789 Developmental Gene Discovery in a Hemimetabolous Insect: De Novo Assembly and Annotation of a
790 Transcriptome for the Cricket Gryllus bimaculatus. Plos One *8*, e61479. 10.1371/journal.pone.0061479.

791 57. Zhou, X., Liao, Z., Jia, Q., Cheng, L., and Li, F. (2007). Identification and characterization of Piwi
792 subfamily in insects. Biochem Bioph Res Co *362*, 126–131. 10.1016/j.bbrc.2007.07.179.

793 58. Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space
794 complexity. Bmc Bioinformatics *5*, 113. 10.1186/1471-2105-5-113.

795 59. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput.
796 Nucleic Acids Res *32*, 1792–1797. 10.1093/nar/gkh340.

797 60. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
798 phylogenies. Bioinformatics *30*, 1312–1313. 10.1093/bioinformatics/btu033.

799 61. Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. Curr Protoc Bioinform *69*,
800 e96. 10.1002/cpbi.96.

801 62. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T. (2017). ggtree: an r package for visualization
802 and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol *8*,
803 28–36. 10.1111/2041-210x.12628.

804 63. Andrews, S. (2010). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput
805 Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

806 64. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads.
807 Embnet J *17*, 10–12. 10.14806/ej.17.1.200.

808 65. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*,
809 357–359. 10.1038/nmeth.1923.

810 66. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,
811 Durbin, R., and Subgroup, 1000 Genome Project Data Processing (2009). The Sequence Alignment/Map
812 format and SAMtools. Bioinformatics *25*, 2078–2079. 10.1093/bioinformatics/btp352.

813 67. Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and
814 better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res *47*, gkz114-.
815 10.1093/nar/gkz114.

816 68. Montañés, J.C., Rojano, C., Ylla, G., Piulachs, M.D., and Maestro, J.L. (2021). siRNA enrichment in
817 Argonaute 2-depleted Blattella germanica. Biochimica Et Biophysica Acta Bba - Gene Regul Mech *1864*,
818 194704. 10.1016/j.bbagrm.2021.194704.

819 69. Matsumoto, Y., Noji, S., and Mizunami, M. (2003). Time Course of Protein Synthesis-Dependent
820 Phase of Olfactory Memory in the Cricket Gryllus bimaculatus. Zool Sci *20*, 409–416.
821 10.2108/zsj.20.409.

822 70. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A.,
823 Markowitz, S., Duran, C., et al. (2012). Geneious Basic: An integrated and extendable desktop software
824 platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649.
825 10.1093/bioinformatics/bts199.

826 71. Matsumoto, Y., and Mizunami, M. (2002). Temporal determinants of long-term retention of olfactory
827 memory in the cricket Gryllus bimaculatus. The Journal of Experimental Biology *205*, 1429–1437.

828 72. Matsumoto, Y., Hatano, A., Unoki, S., and Mizunami, M. (2009). Stimulation of the cAMP system by
829 the nitric oxide-cGMP system underlying the formation of long-term memory in an insect. Neurosci Lett
830 *467*, 81–85. 10.1016/j.neulet.2009.10.008.

831 73. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-
832 Time Quantitative PCR and the 2−ΔΔC T Method. Methods *25*, 402–408. 10.1006/meth.2001.1262.

833 74. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple
834 sequence alignment based on fast Fourier transform. Nucleic Acids Res *30*, 3059–3066.
835 10.1093/nar/gkf436.

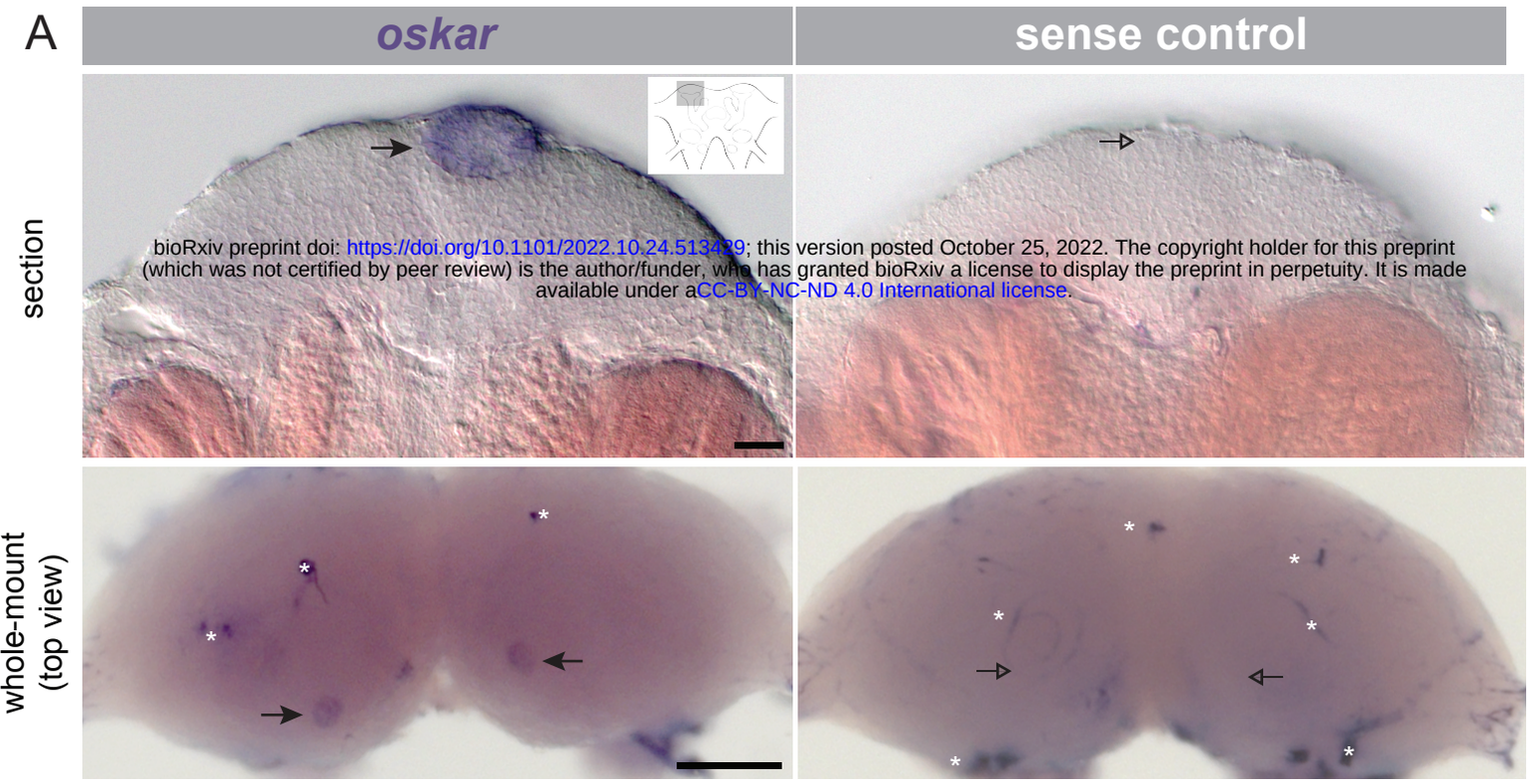
836 75. Castro-Mondragon, J.A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L.,
837 Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., et al. (2021). JASPAR 2022:
838 the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res *50*,
839 gkab1113-. 10.1093/nar/gkab1113.

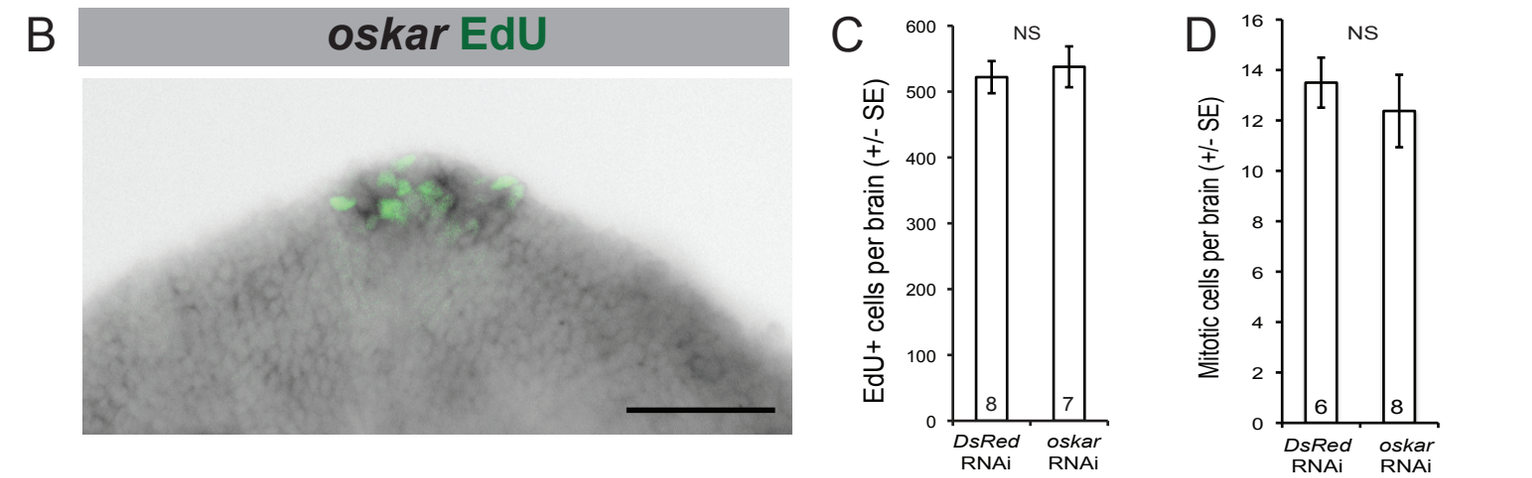
840 76. Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res
841 *43*, W39–W49. 10.1093/nar/gkv416.

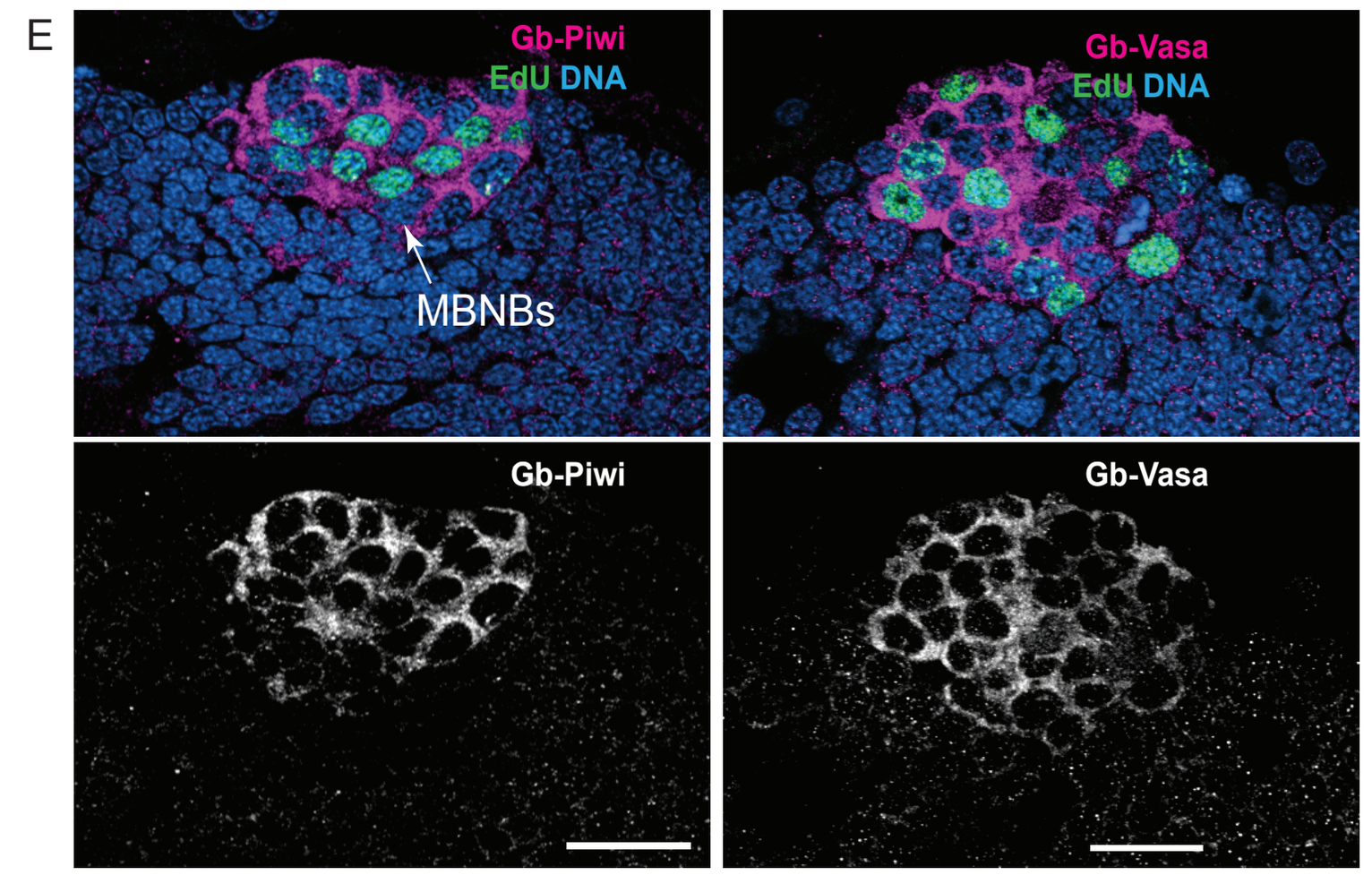
842 77. Stothard, P. (2000). The Sequence Manipulation Suite: JavaScript Programs for Analyzing and
843 Formatting Protein and DNA Sequences. Biotechniques *28*, 1102–1104. 10.2144/00286ir01.

844 78. Hsieh, Y.-W., Alqadah, A., and Chuang, C.-F. (2016). An Optimized Protocol for Electrophoretic
845 Mobility Shift Assay Using Infrared Fluorescent Dye-labeled Oligonucleotides. J Vis Exp.
846 10.3791/54863.

847







A

B

C

Preference index

100%
80%
60%
40%
20%
0%

before
1 hr after
1 day after

DsRed dsRNA
osk dsRNA #1
osk dsRNA #2
piwi dsRNA
osk dsRNA + *piwi* dsRNA

D

E

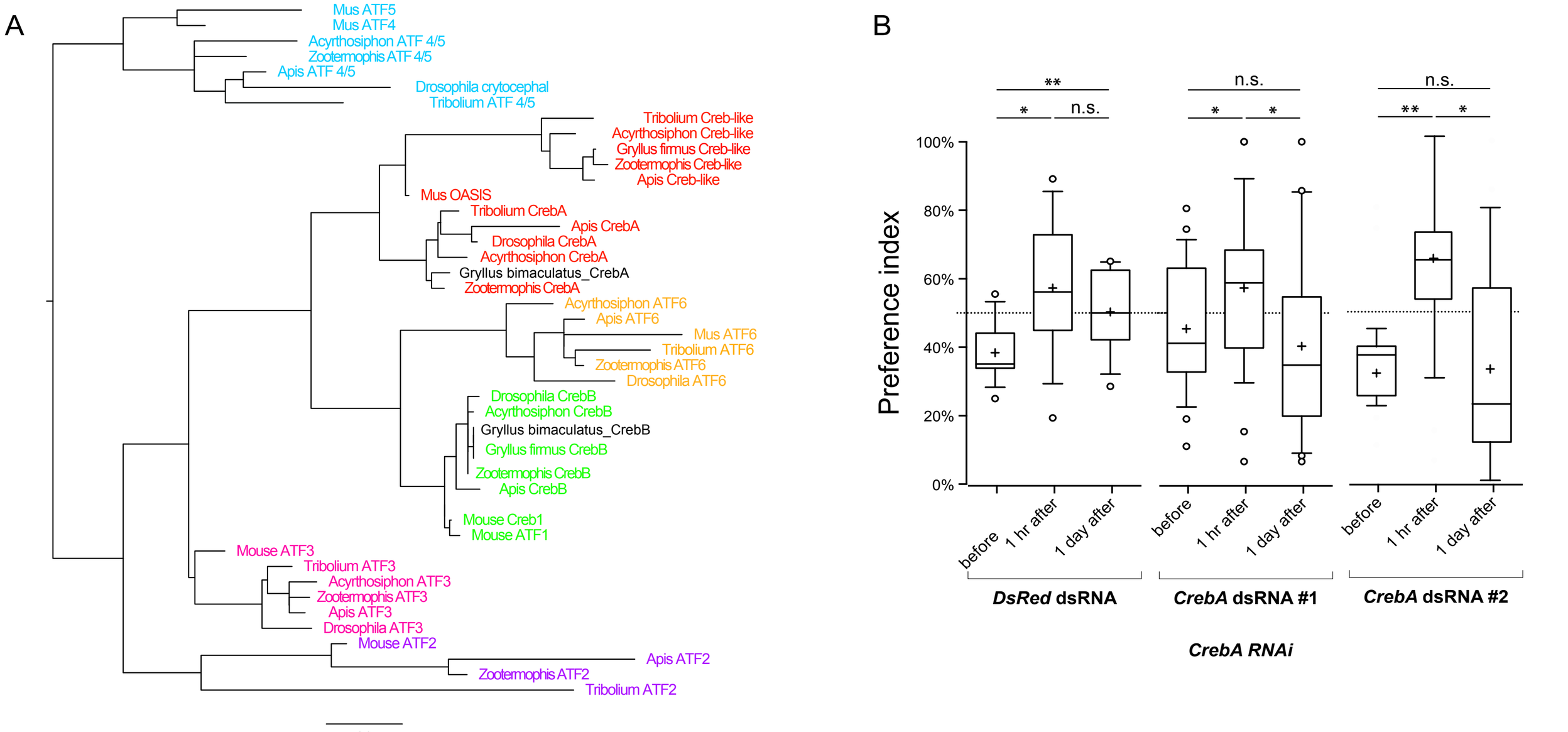
F

mRNA fold change (+/- SD)

control
RNAi

osk
piwi
CrebA
vasa

***osk* RNAi**
***piwi* RNAi**
***osk*/*piwi* RNAi**


A
Mus ATF5
Mus ATF4
Acyrthosiphon ATF 4/5
Zootermophis ATF 4/5
Apis ATF 4/5
Drosophila crytocephal
Tribolium ATF 4/5
Tribolium Creb-like
Acyrthosiphon Creb-like
Gryllus firmus Creb-like
Zootermophis Creb-like
Apis Creb-like
Mus OASIS
Tribolium CrebA
Apis CrebA
Drosophila CrebA
Acyrthosiphon CrebA
Gryllus bimaculatus_CrebA
Zootermophis CrebA
Acyrthosiphon ATF6
Apis ATF6
Mus ATF6
Tribolium ATF6
Zootermophis ATF6
Drosophila ATF6
Drosophila CrebB
Acyrthosiphon CrebB
Gryllus bimaculatus_CrebB
Gryllus firmus CrebB
Zootermophis CrebB
Apis CrebB
Mouse Creb1
Mouse ATF1
Mouse ATF3
Tribolium ATF3
Acyrthosiphon ATF3
Zootermophis ATF3
Apis ATF3
Drosophila ATF3
Mouse ATF2
Apis ATF2
Zootermophis ATF2
Tribolium ATF2
0.3
B
Preference index
100%
80%
60%
40%
20%
0%
**
*
n.s.
n.s.
*
*
n.s.
**
*
before
1 hr after
1 day after
DsRed dsRNA
CrebA dsRNA #1
CrebA dsRNA #2
CrebA RNAi