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[9,10,12], Guillem Ylla[1,11], Makoto Mizunami[4] and Cassandra G. Extavour[1,8,13*]

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. Current address: Howard Hughes Medical Institute, Chevy Chase, Maryland, USA
13. Howard Hughes Medical Institute, Chevy Chase, Maryland, USA

* Cassandra G. Extavour

Email: extavour@oeb.harvard.edu

Author Contributions: AK performed all qPCRs, bioinformatic analysis of CRE sites and LOTUS domains, 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; BEC performed whole mount gene expression and EdU assays; KT, YM, YL, and TW performed olfactory learning assays; GY performed small RNA library analysis, LOTUS domain-containing gene identification, and 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; MM and CGE obtained funding for the project.

Keywords: co-option, neuroblast, mushroom body, Kenyon cells, orphan gene

This PDF file includes: Main Text; Figures 1 to 4

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 hemimetabolous 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 holometabolous insects.

Main Text

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 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

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 in the brain 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-Lomassesse(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 (Fig. 2D; Suppl. Table S1) and small RNA profiling of osk^{RNAi} brains (Suppl. Tables S2, S3 and S4; Suppl. Fig. S1). 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. S2A). 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. S2B). 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 as previously described (17). Briefly, crickets were injected with double stranded RNA against the gene of interest (*oskar*) or a control gene (*DsRed* (35), and then subjected to odor preference tests (allowed to freely visit peppermint and vanilla odor sources, quantifying the time spent at each odor source) and conditioning trials (peppermint odor was paired with a water reward, and relative preference for this rewarded odor was compared before and after the conditioning) (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; Fig. 2A). Trained control crickets retained this learned preference ($p<0.01$) even at one day after training (Fig. 2A, "*DsRed*"), 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 (Fig. 2D; Suppl. Table S1) and small RNA sequencing (Suppl. Fig. S1, 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. S3A, $p>0.05$; Suppl. Fig. S3B) 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 S5). 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 S6). We performed *CrebA* and *CrebB* RNAi experiments and

discovered that *CrebA* (but not *CrebB*; Suppl. Fig. S3C) 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. S3C). 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. 3A), qPCR revealed no reduction of *CrebA* transcripts in $vasa^{RNAi}$ conditions (Suppl. Fig. 3B). 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* or *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 regulate each other's transcript levels. To evaluate the possibility that *osk* or *piwi* might be direct transcriptional targets of CrebA, we examined the genomic sequences within 10kb upstream of both loci and found two bioinformatically predicted CRE binding sites within the 6kb upstream of the transcription start sites for *osk* and *piwi* (Fig. 4C; Suppl. File 1). These predicted CRE binding sites 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 *osk* (Fig. 4D; Suppl. Table S7). 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*. 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. 4B), and that *osk* and *piwi* genomic loci contain predicted CRE binding sites (Fig. 4C), the results of our gel shift assay are consistent with the hypothesis that cricket *osk* is a direct transcriptional target of CrebA.

Discussion

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* disrupts 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 was recovered in an insertional mutagenesis screen for long term memory in in *D. melanogaster* (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 Pumilio, Staufen, and Orb (oo18 RNA-binding protein) 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*, including for example *staufen* (47) and *tudor* (48), also function in *G. bimaculatus* adult neuroblasts, which would suggest that *osk* acts with conserved molecular partners in different cellular contexts. Because mushroom bodies derived from neuroblasts are a conserved arthropod brain structure across and beyond insects (49–52), it

seems unlikely that *oskar* played a role in the evolution of insect mushroom body neurogenesis *per se*.

Both germ cells and neuroblasts are stem cells that give rise to highly specialized 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, many different highly conserved “germ line genes” including *vasa*, *nanos*, and *piwi* are found in a variety of multipotent cells in diverse animals(14, 53), raising the possibility that such genes were involved in establishing multipotency rather than specifying germ cell fate *per se*. In our previous examination of the distribution of *osk* orthologs across insects, we observed that crickets are not the only insects reported to express *osk* in the brain, as *osk* transcripts are detected in transcriptomes from the brains of cockroach, wasp and beetle species as well (16). This is consistent with the hypothesis that *osk* played an ancient neural role of some kind in insects. We previously showed evidence supporting the hypothesis that evolved changes in the biophysical characteristics of Oskar protein may have driven the evolution of a novel mechanism of germ line specification in the holometabolous insects (16). A broader understanding of the putative ancestral and derived 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 or acquiring 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. An alternative hypothesis to *de novo* evolution of CRE sites to explain *osk*’s expression in the cricket brain, is that *osk* inherited CRE binding sites from an ancestral sequence that contributed to *osk*’s genesis. Our previous work suggested that *osk* arose through a fusion of a eukaryotic LOTUS domain in the 5’ position, coding for *osk*’s N-terminal LOTUS domain, and a prokaryotic

SGNH hydrolase-like domain in the 3' position, coding for *osk*'s C-terminal OSK domain (54). In this scenario, a pre-existing LOTUS-domaining gene could have donated not just its LOTUS domain, but also some upstream 5' regulatory sequences, including CRE sites and/or neural expression elements, to *osk*. This hypothesis might predict that extant LOTUS domain-containing genes might display one or both of CRE binding sites, or expression in the brain. To test this hypothesis, we searched the *G. bimaculatus* genome for LOTUS domain-containing genes, and identified five such genes (Supplementary File 2). These were *osk*, *Tdrd5*, *Tdrd7*, *limkain b1*, and an uncharacterized gene with annotation ID GBI_15344 (Supplementary File 2). In transcriptomes previously generated from the adult brain (55), we detected levels of all four non-*oskar* genes at levels at least as high as that detected for *osk* (Suppl. Figure S4, Suppl. Table S8). Moreover, in the 10kb upstream of the first predicted codon of these genes, we detected a putative CRE binding site (Suppl. Table S9). With the caveat that our transcriptomes do not provide spatial or cell-type resolution for the expression data, both of these findings are consistent with the hypothesis that whatever eukaryotic LOTUS domain-containing sequence was the ancestor of *osk's* LOTUS domain, also contributed one or both of CRE-responsive or brain-expressed upstream regulatory sequences to *osk*. 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 further speculate that the biophysical properties of Osk protein that make it effective at sequestering RNAs and participating in translational control in the germ line(56–58), 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, 59), its regions of high predicted disorder(16, 46, 59), 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

Gryllus bimaculatus husbandry, in situ hybridization, immunostaining, olfactory learning assays, RNAi and qPCR were performed as previously described. See Extended Materials & Methods in Supporting Information for references and detailed protocols.

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 epifluorescence 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).

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 S2) 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 (Suppl. Table S3) were uploaded onto NCBI SRA database and are publicly available under the BioProject ID PRJNA837371.

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 (Suppl. Table S5) 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)(60). A maximum likelihood tree was created in RAxML using the PROTGAMMAWAG model (61) , and plotted with the FigTree package v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree) (Fig.3B).

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(62) (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 (63)), 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 randomly chosen *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 were 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 (64) 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 ~1.8 CRE sites for every 10Kb tested (Suppl. Files 1,2).

Bioinformatic analysis of LOTUS domain-containing genes in the G. bimaculatus *genome*

We searched the annotated *G. bimaculatus* genome (36) for genes whose protein product was predicted to contain PFAM motif PF12872, corresponding to the LOTUS domain. This search retrieved five genes (GBI_01840 "*oskar*", GBI_13502 "*TDRD5/tejas*", GBI_15344 "*uncharacterized*", GBI_15604 "*limkain b1*", GBI_03370 "*TDRD7/tapas*"). We assessed expression of these genes in brains and gonads using previously published RNA-seq libraries (55) available at NCBI (PRJNA564136). We analyzed the RNA-seq data as in (16), including removing adapters and reads shorter than 20 nucleotides with Cutadapt v3.4 (65) and quantifying the gene expression in transcripts per million (TPM) with RSEM v1.2.29 (66), using STAR

v2.7.0e1 (67) as read mapper against the *G. bimaculatus* genome (36) (Suppl. Table S8). For each of these genes we retrieved and searched the 10Kb upstream of the first codon annotated for CRE sites as described above in "*Bioinformatic prediction of CRE sites in* osk *and* piwi *upstream regulatory region*" (Suppl. Table S9).

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). Once both CRE sites were sequence confirmed by Sanger sequencing, the ~30bp fragments containing each CRE site were synthetically generated as duplexes (with 3'A overhangs) for use as EMSA pre-probes (Suppl. Table S7; CRE site in bold). 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-Fisher, 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(68). 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).

Acknowledgments and Funding Sources

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, and for financial support of SSP; 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. 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.

References

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

2. A. Ephrussi, L. K. Dickinson, R. Lehmann, oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66, 37–50 (1991).

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

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

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

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

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

8. J. Dubnau, *et al.*, The staufen/pumilio Pathway Is Involved in Drosophila Long-Term Memory. *Curr Biol* 13, 286–296 (2003).

9. T.-P. Pai, *et al.*, Drosophila ORB protein in two mushroom body output neurons is necessary for long-term memory formation. *Proc National Acad Sci* 110, 7898–7903 (2013).

10. P. N. Perrat, *et al.*, Transposition-Driven Genomic Heterogeneity in the Drosophila Brain. *Science* 340, 91–95 (2013).

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

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

13. B. Ye, *et al.*, nanos and pumilio Are Essential for Dendrite Morphogenesis in Drosophila Peripheral Neurons. *Curr Biol* 14, 314–321 (2004).

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

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

16. L. Blondel, S. Besse, E. L. Rivard, G. Ylla, C. G. Extavour, 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 (2021).

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

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

19. S. M. Farris, I. Sinakevitch, 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 (2003).

20. J. Malaterre, *et al.*, Development of cricket mushroom bodies. *J Comp Neurol* 452, 215–227 (2002).

21. R. Urbach, G. M. Technau, 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 (2003).

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

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

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

25. A. R. Simões, *et al.*, Damage-responsive neuro-glial clusters coordinate the recruitment of dormant neural stem cells in Drosophila. *Dev Cell* 57, 1661-1675.e7 (2022).

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

27. N. J. Strausfeld, *Arthropod Brains: Evolution, Functional Elegance, and Historical Significance* (Harvard University Press, 2012) (September 3, 2022).

28. S. Scotto-Lomassese, *et al.*, Suppression of Adult Neurogenesis Impairs Olfactory Learning and Memory in an Adult Insect. *J Neurosci* 23, 9289–9296 (2003).

29. T. Takahashi, *et al.*, Systemic RNA interference for the study of learning and memory in an insect. *J Neurosci Meth* 179, 9–15 (2009).

30. C. Klattenhoff, *et al.*, Drosophila rasiRNA Pathway Mutations Disrupt Embryonic Axis Specification through Activation of an ATR/Chk2 DNA Damage Response. *Dev Cell* 12, 45–55 (2007).

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

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

33. M. Mizunami, J. M. Weibrecht, N. J. Strausfeld, Mushroom bodies of the cockroach: Their participation in place memory. *J Comp Neurol* 402, 520–537 (1998).

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

35. R. E. Campbell, *et al.*, A monomeric red fluorescent protein. *Proc National Acad Sci* 99, 7877–7882 (2002).

36. G. Ylla, *et al.*, Insights into the genomic evolution of insects from cricket genomes. *Commun Biology* 4, 733 (2021).

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

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

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

40. C.-C. Chen, *et al.*, Visualizing Long-Term Memory Formation in Two Neurons of the Drosophila Brain. *Science* 335, 678–685 (2012).

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

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

43. K. Ito, Y. Hotta, Proliferation pattern of postembryonic neuroblasts in the brain of Drosophila melanogaster. *Dev Biol* 149, 134–148 (1992).

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

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

46. K. E. Kistler, *et al.*, Phase transitioned nuclear Oskar promotes cell division of Drosophila primordial germ cells. *Elife* 7, e37949 (2018).

47. D. S. Johnston, D. Beuchle, C. Nüsslein-Volhard, staufen, a gene required to localize maternal RNAs in the Drosophila egg. *Cell* 66, 51–63 (1991).

48. R. E. Boswell, A. P. Mahowald, tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. *Cell* 43, 97–104 (1985).

49. F. J. Maza, J. Sztarker, M. E. Cozzarin, M. G. Lepore, A. Delorenzi, A crabs' high-order brain center resolved as a mushroom body-like structure. *J Comp Neurol* 529, 501–523 (2021).

50. N. J. Strausfeld, G. H. Wolff, M. E. Sayre, Mushroom body evolution demonstrates homology and divergence across Pancrustacea. *Elife* 9, e52411 (2020).

51. N. Strausfeld, M. E. Sayre, Shore crabs reveal novel evolutionary attributes of the mushroom body. *Elife* 10, e65167 (2021).

52. C. Doeffinger, V. Hartenstein, A. Stollewerk, Compartmentalization of the precheliceral neuroectoderm in the spider Cupiennius salei: Development of the arcuate body, optic ganglia, and mushroom body. *J Comp Neurol* 518, 2612–2632 (2010).

53. C. E. Juliano, S. Z. Swartz, G. M. Wessel, A conserved germline multipotency program. *Development* 137, 4113–4126 (2010).

54. L. Blondel, T. E. M. Jones, C. G. Extavour, Bacterial contribution to genesis of the novel germ line determinant oskar. *Elife* 9, e45539 (2020).

55. C. A. Whittle, A. Kulkarni, N. Chung, C. G. Extavour, Adaptation of codon and amino acid use for translational functions in highly expressed cricket genes. *Bmc Genomics* 22, 234 (2021).

56. P. M. Macdonald, M. Kanke, A. Kenny, Community effects in regulation of translation. *Elife* 5, e10965 (2016).

57. C. Rongo, E. R. Gavis, R. Lehmann, Localization of oskar RNA regulates oskar translation and requires Oskar protein. *Dev Camb Engl* 121, 2737–46 (1995).

58. A. Kulkarni, C. G. Extavour, Convergent evolution of germ granule nucleators: A hypothesis. *Stem Cell Res* 24, 188–194 (2017).

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

60. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059–3066 (2002).

61. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313 (2014).

62. J. A. Castro-Mondragon, *et al.*, JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 50, gkab1113- (2021).

63. T. L. Bailey, J. Johnson, C. E. Grant, W. S. Noble, The MEME Suite. *Nucleic Acids Res* 43, W39–W49 (2015).

64. P. Stothard, The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *Biotechniques* 28, 1102–1104 (2000).

65. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet J* 17, 10–12 (2011).

66. B. Li, C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* 12, 323 (2011).

67. A. Dobin, *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).

68. Y.-W. Hsieh, A. Alqadah, C.-F. Chuang, An Optimized Protocol for Electrophoretic Mobility Shift Assay Using Infrared Fluorescent Dye-labeled Oligonucleotides. *J Vis Exp* (2016) https:/doi.org/10.3791/54863.

1

Figures and Tables

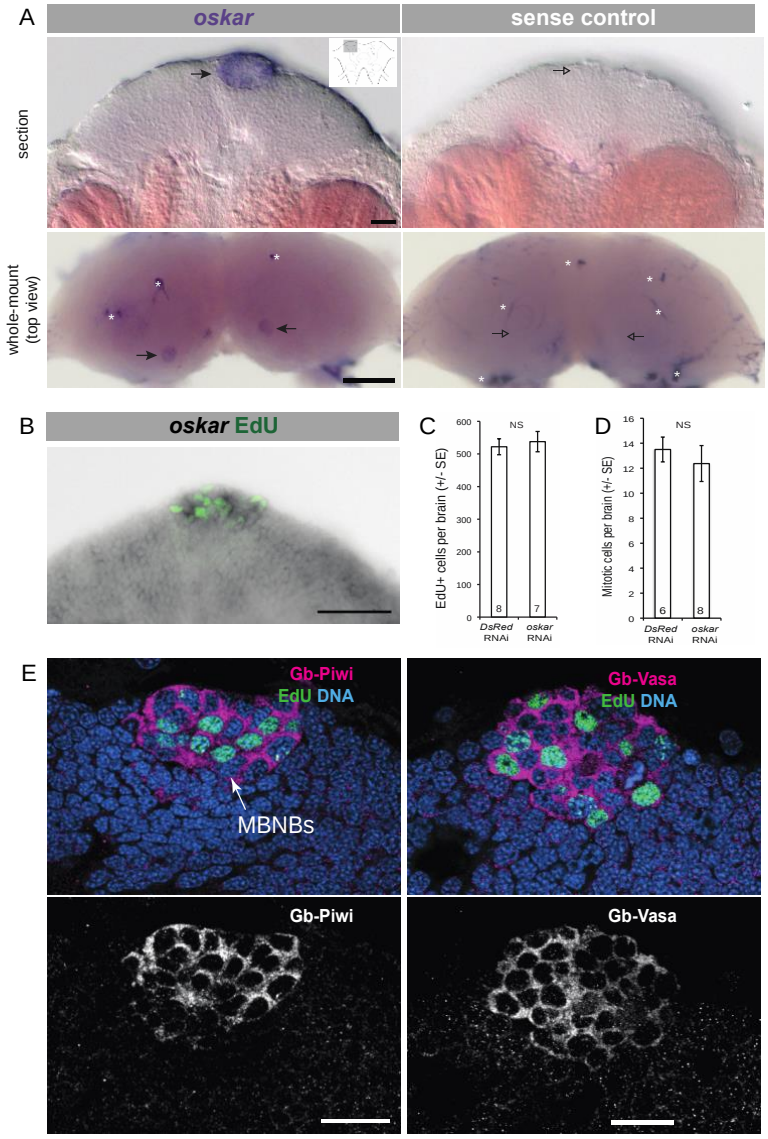

Figure 1. ***oskar*, Piwi and Vasa are expressed in *G. bimaculatus* adult mushroom body neuroblasts.** (**A**) In situ hybridization on adult *G. bimaculatus* brains detects *osk* transcripts in the cells of the mushroom body (arrows). Inset in top left panel shows the overall structure of the adult 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 (magenta) 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)**.

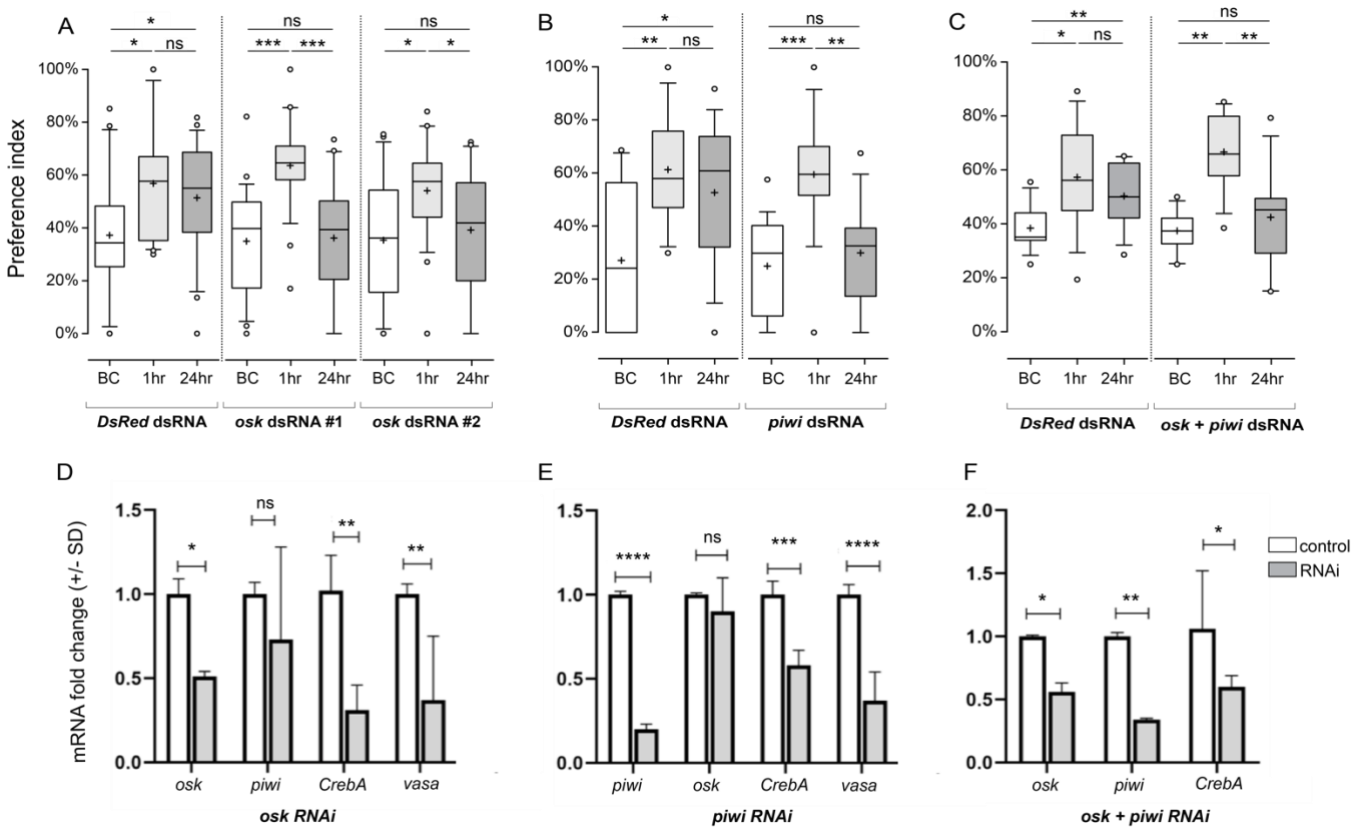

Figure 2. ***oskar^{RNAi}* and *piwi*^{RNAi} impairs cricket long-term memory.** (**A**) Results of olfactory memory assay in *osk^{RNAi}* animals. (**B**) Results of olfactory memory assay in *piwi^{RNAi}* animals. (**C**) Results of olfactory memory assay in *osk^{RNAi}/piwi^{RNAi}* double knockdown animals. For each assay, relative preference between the rewarded odor (peppermint) and control odor (vanilla) was tested before conditioning (BC), one hour post-training (1hr), and one day post training (24hr) for *DsRed^{RNAi}* controls and for *osk^{RNAi}* (using two different non-overlapping *osk* fragments #1 and #2), *piwi^{RNAi}* and *osk^{RNAi}/piwi^{RNAi}*. Boxes represent the 1st and 3rd quartiles surrounding the median (middle line). Whiskers extend to 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 statistically significant). (**D-F)**. qPCR results showing the extent of downregulation of different *G. bimaculatus* genes in *osk^{RNAi}*, *piwi^{RNAi}* and *osk^{RNAi}/piwi^{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).

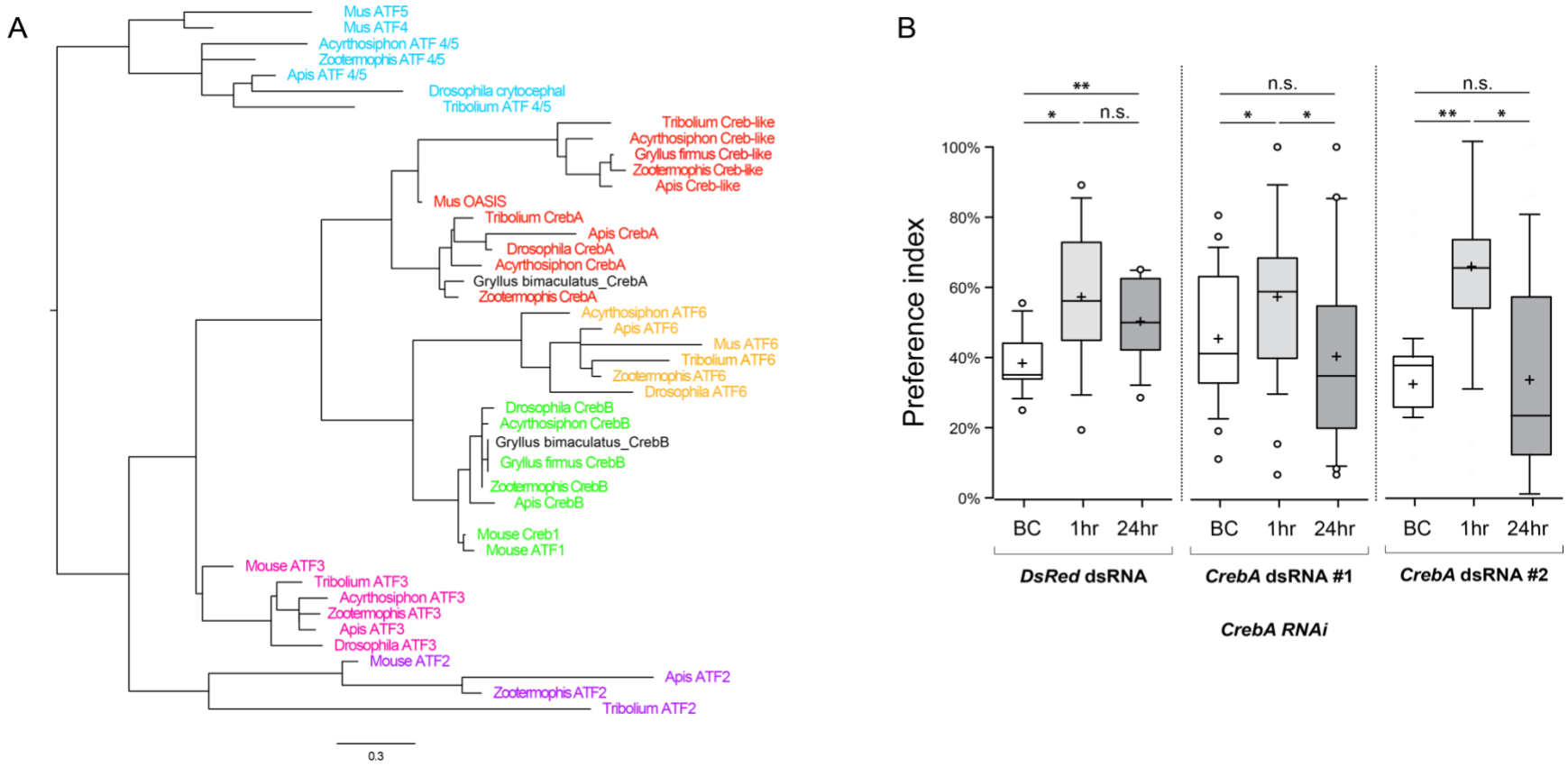

Figure 3. **Cricket *CrebA* is required for cricket long-term memory.** (**A**) GenBank IDs of *Creb/ATF* family member orthologs in mouse and insects (Suppl. Table S5) were used to construct a *Creb* phylogenetic tree to infer the evolutionary relationships between mammalian *Creb* proteins and their 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 conditioning (BC), one hour post-training (1hr), and one day post training (24hr) for *DsRed*^{RNAi} 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 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 statistically significant). n=9 for *CrebA* and n=10 for *DsRed*.

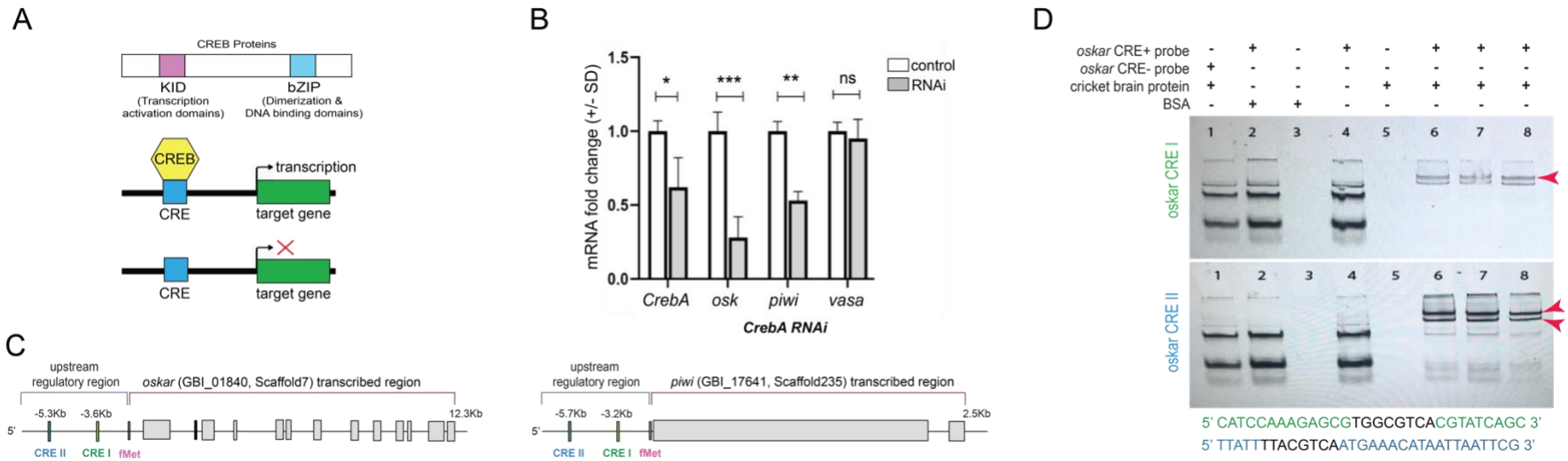

Figure 4. Cricket CrebA regulates *oskar*. (**A**) Schematic diagram of the transcription factor cAMP response element binding protein (Creb) protein (top) displaying only the two domains relevant to this study, the Kinase-Inducible Domain (KID) that can facilitate kinase-inducible transcription activation 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 genes. (**B**) qPCR results showing the relative expression levels of *G. bimaculatus osk*, *piwi*, and *vasa* in *CrebA*^{RNAi} knockdown conditions. The extent of *CrebA* transcript decrease 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* and *piwi* genes showing exons (depicted as rectangular grey boxes) along with their presumptive upstream regulatory regions, each containing two predicted CRE sites, which we call CRE I (~3.6Kb and ~3.2Kb upstream of the predicted transcription start site for *osk* and *piwi* respectively, marked by fMet) and CRE II (~5.3Kb and ~5.7Kb upstream of predicted fMet for *osk* and *piwi* respectively). (**D**) 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. Shift marked with red arrowheads.