Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo

Graphical Abstract

Highlights
- Aub binds to maternal mRNAs in early Drosophila embryos
- Interaction between Aub and maternal mRNAs depends on piRNAs
- aub mutants are defective in mRNA decay during the MZT
- Aub-dependent unstable mRNAs encode germ cell determinants

Authors
Bridlin Barckmann, Stéphanie Pierson, Jérémy Dufourt, ..., Jean-Pierre Desvignes, Tomaz Curk, Martine Simonelig

Correspondence
martine.simonelig@igh.cnrs.fr

In Brief
Using iCLIP, Barckmann et al. identify several hundred maternal mRNAs that interact with Aub in early embryos. A number of these mRNAs undergo Aub-dependent destabilization in the soma during the maternal-to-zygotic transition and encode germ cell determinants.

Accession Numbers
GSE70778
Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo

Bridlin Barckmann,1,6 Stéphanie Pierson,1,6 Jérémy Dufourt,1,6 Catherine Papin,1 Claudia Armenise,2 Fillip Port,3 Thomas Grenztinger,2 Séverine Chambeyron,2 Grégory Baronian,1 Jean-Pierre Desvignes,4 Tomaz Curk,5 and Martine Simonelig1,*

1mRNA Regulation and Development, Institut de Génétique Humaine, CNRS, 141 rue de la Cardonille, 34396 Montpellier Cedex 5, France
2RNA Silencing and Control of Transposition, Institut de Génétique Humaine, CNRS, 141 rue de la Cardonille, 34396 Montpellier Cedex 5, France
3Division of Cell Biology, Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, UK
4MGX-Montpellier GenomiX, c/o Institut de Génomique Fonctionnelle, 141 rue de la cardonille, 34094 Montpellier Cedex 5, France
5Faculty of Computer and Information Science, University of Ljubljana, 1000 Ljubljana, Slovenia
6Co-first author
*Correspondence: martine.simonelig@igh.cnrs.fr
http://dx.doi.org/10.1016/j.celrep.2015.07.030
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

The Piwi-interacting RNA (piRNA) pathway plays an essential role in the repression of transposons in the germline. Other functions of piRNAs such as post-transcriptional regulation of mRNAs are now emerging. Here, we perform iCLIP with the PIWI protein Aubergine (Aub) and identify hundreds of maternal mRNAs interacting with Aub in the early Drosophila embryo. Gene expression profiling reveals that a proportion of these mRNAs undergo Aub-dependent destabilization during the maternal-to-zygotic transition. Strikingly, Aub-dependent unstable mRNAs encode germ cell determinants. iCLIP with an Aub mutant that is unable to bind piRNAs confirms piRNA-dependent binding of Aub to mRNAs. Base pairing between piRNAs and mRNAs can induce mRNA cleavage and decay that are essential for embryonic development. These results suggest general regulation of maternal mRNAs by Aub and piRNAs, which plays a key developmental role in the embryo through decay and localization of mRNAs encoding germ cell determinants.

INTRODUCTION

In most species, the first steps of embryonic development depend on maternally loaded mRNAs and proteins. The control of development then switches from the maternal to the zygotic genome during the maternal-to-zygotic transition (MZT), during which maternal mRNAs are massively degraded (De Renzis et al., 2007). In Drosophila, the RNA-binding protein Smaug (Smg) plays a major role in this general maternal mRNA decay (Chen et al., 2014; Tadros et al., 2007). Smg binds to mRNAs containing Smg recognition elements (SREs) and induces their deadenylation and decay by recruiting the CCR4–NOT deadenylation complex (S eminentok et al., 2005; Zaessinger et al., 2006). Previously, we reported that the Piwi-interacting RNA (piRNA) pathway cooperates with Smg for maternal mRNA deadenylation and decay, prior to zygotic transcription. The piRNA pathway plays an essential role in the regulation of the posterior determinant Nanos (Nos) (Rouget et al., 2010). piRNAs are a specific class of small 23–30 nt non-coding RNAs loaded into Argonaute proteins of the PIWI clade (Guzzardo et al., 2013; Ishizu et al., 2012). A prominent function of the piRNA pathway is the repression of the expression and transposition of transposable elements (TEs) in the germline. A large proportion of piRNAs (~70%) derives from TE sequences in Drosophila ovaries. These piRNAs target TE mRNAs through complementarity and guide their cleavage by the cytoplasmic PIWI proteins Aubergine (Aub) and Argonaute 3 (Ago3) bound to the piRNAs.

Recently, further evidence has emerged for additional functions of the piRNA pathway (Peng and Lin, 2013; Watanabe and Lin, 2014; Weick and Miska, 2014). Variable fractions of piRNAs are produced from protein-coding mRNAs, with a bias toward 3′ UTR, both in mouse testes and in Drosophila ovaries and early embryos. The function of these genic piRNAs remains to be investigated, although their production might decrease the levels of mRNAs from which they originate (Robine et al., 2009; Saito et al., 2009).

Pachytene piRNAs, which account for more than 95% of piRNAs in the adult mouse testis, do not derive from TE sequences but mostly from intergenic regions (Beyret et al., 2012; Li et al., 2013). Recently, pachytene piRNAs have been shown to mediate global mRNA decay in spermatocytes and spermatids (Goh et al., 2015; Gou et al., 2014; Watanabe et al., 2015; Zhang et al., 2015). Pachytene piRNAs loaded into MIWI, the mouse homolog of Aub, target spermiogenic mRNAs with imperfect base pairing and induce their decay either through...
MIWI-dependent cleavage or by assembling a complex containing CAF1, a deadenylase in the CCR4-NOT complex.

Another compelling example of mRNA regulation by a piRNA has been reported for sex determination in Bombyx mori, where a female-specific piRNA plays a vital role by inducing cleavage of a cellular mRNA involved in masculinization of the embryos (Kuch et al., 2014).

In the Drosophila embryo, piRNAs target nos maternal mRNA and contribute to its decay in the somatic region. Nos is expressed as a gradient emanating from the posterior pole of the embryo and is essential for abdominal segmentation and germ cell development. nos mRNA is present throughout the embryo, although it is translationally repressed and degraded in the somatic part whereas stabilized and translated in the germ plasm (i.e., the cytoplasm localized at the posterior pole that is required for germ cell specification). Two piRNAs produced from TEs target the nos 3’ UTR with imperfect base pairing and guide interactions with Aub and Ago3, which in turn recruit the CCR4-NOT deadenylase complex, together with Smg. These interactions lead to nos mRNA translational repression and decay in the soma and are required for embryonic patterning (Rouget et al., 2010).

Here, we use Aub iCLIP (individual-nucleotide resolution UV crosslinking and immunoprecipitation) (König et al., 2010) to address a potential general role of Aub and piRNAs in maternal mRNA regulation in the early embryo. This approach identifies several hundred mRNAs that directly interact with Aub in 0- to 2-hr embryos. Gene-expression profiling reveals that one-third of these mRNAs undergo decay at the MZT in wild-type embryos and are stabilized in embryos from aub and spn-E mutant females. The RNA helicase Spn-E has a prominent role in germ line piRNA production (Malone et al., 2009). Strikingly, Aub- and Spn-E-dependent unstable mRNAs are enriched in mRNAs that are stabilized in the germ plasm and that encode germ plasm components involved in germ cell specification and development. This reveals a role of Aub in the localization of these mRNAs to the germ plasm through their selective decay in the somatic part of the embryo. Bioinformatic analyses, as well as iCLIP with a PAZ-domain mutant form of Aub that is unable to load piRNAs, are consistent with the requirement of piRNAs in Aub-mRNA interactions. In addition, base pairing of maternal mRNAs with piRNAs can induce the production of genic piRNAs by cleavage, in a process similar to the ping-pong occurring with TEs, which leads to functional mRNA downregulation.

These results reveal a general role for Aub and piRNAs in the regulation of maternal mRNAs in the embryo. This piRNA-dependent regulation plays a key role in the decay and localization of mRNAs involved in germ cell development.

RESULTS

Aub Directly Binds Maternal mRNAs in Embryos

To identify maternal mRNAs directly interacting with Aub, we performed iCLIP of GFP-Aub in 0- to 2-hr embryos using anti-GFP antibodies. This technique allows mapping RNA-protein interactions at individual nucleotide resolution (König et al., 2010). The GFP-Aub transgene (Harris and Macdonald, 2001) was able to rescue the aub mutant maternal phenotypes of embryonic lethality and fused dorsal appendages to a large extent, and it was not detrimental when expressed in wild-type embryos (Figure S1A). We performed three independent iCLIP experiments (Figure 1A). The protein-RNA complexes were absent when anti-GFP was replaced with rabbit serum, as well as when UV crosslinking was omitted, indicating that the protein-RNA complexes purified in the experiment were covalent complexes depending on UV. Crosslinked RNA was reverse transcribed and PCR amplified, and the resulting DNA was submitted to high-throughput sequencing. Sequence duplicates due to PCR amplification were eliminated by removing identical sequences linked to the same random barcode primer. High-throughput sequencing of the three biological replicates using Illumina HiSeq 2000 generated 32.9, 36.3, and 26.1 million reads, which identified, after removal of PCR duplicates 651,699, 265,646 and 2,620,195 reads mapping to the Drosophila genome, respectively (Table S1). Each of these reads represents a uniquely crosslinked RNA molecule.

Although crosslink efficiency was variable between replicates, the reproducibility of crosslinks was validated by a peak of crosslink positions reproduced in an independent experiment at the same exact position and with an offset of a few nt (Figure S1B). A similar offset by a few nt has previously been observed in iCLIP experiments. It was proposed to result from contacts of the protein to more than 1 nt of the RNA and/or from the steric hindrance of the remaining peptide on the RNA, which might cause an imprecise termination of reverse transcription (König et al., 2010). We also determined that the reproducibility between experiments increased with the incidence of crosslinks at the same position within one experiment (cDNA count), showing that the strongest crosslink sites were the most reproducible (Figure 1B).

Because Aub is thought to directly interact with piRNAs and mRNAs, sequence reads from both types of molecules were expected to be present in GFP-Aub iCLIPs. This technique is based on the truncation of cDNAs at the crosslink site; therefore, the recovery of complete piRNAs in Aub iCLIPs was not expected. However, length analysis of reads mapping to piRNA clusters revealed peaks of 23- to 27-nt reads, suggesting the presence of full-length piRNAs in GFP-Aub iCLIPs (Figures 1C and S1C). Analysis of nt distribution of reads 23–29 nt in length identified a 1U bias (Figures 1D and S1D), in agreement with a proportion of full-length piRNAs being recovered in GFP-Aub iCLIPs. These piRNAs would be crosslinked through their 5’ phosphate, consistent with the interactions between this phosphate and the MID domain of Argonaute proteins (Schirle and MacRae, 2012). The presence of reads corresponding to piRNAs validated the ability of GFP-Aub iCLIP to identify RNA molecules specifically interacting with Aub.

The mapping of unique-mapped and multi-mapped reads to the genome identified both sequences related to TE piRNAs (i.e., reads mapping to piRNA clusters and TEs referred to as piRNA reads), as well as sequences corresponding to cellular mRNAs (Figure S1E; Table S1). The proportion of reads in each of these categories was variable between iCLIP replicates, due in part to size variations in the RNA-protein complexes selected from the gels (Figure 1A). piRNA reads were enriched in short size reads (86%–96% of piRNA reads were 12–29 nt in length,
in iCLIP1–3), and the higher proportion of long reads in iCLIP3 (84% >29 nt versus 33% in iCLIP2 and 17% in iCLIP1) correlated with a smaller proportion of piRNA reads.

We focused on cellular mRNAs and identified a positive correlation (Pearson’s correlation coefficient r = 0.88) between the numbers of unique-mapped reads per gene in independent iCLIP experiments, indicating that the specificity of GFP-Aub iCLIP also applied to cellular mRNAs (Figure 1E). To differentiate robust reproducible from transient Aub-mRNA interactions, we considered crosslink sites with a cDNA count of at least five within an iCLIP experiment. Crosslink sites were reproduced with an offset of a few nt (Figure S1B); therefore, this cDNA count was scored within clusters of ±5 nt of each crosslink position (Ånkö et al., 2012). These criteria identified 1,778, 1,473, and 3,744 genes showing at least one cluster with a cDNA count greater than or equal to five in iCLIP1, 2, and 3, respectively (Table S2). Reproducibility between iCLIP biological replicates was then analyzed, and clusters with greater than or equal to five cDNA counts reproduced in at least two independent iCLIP experiments were selected (Table S3). This identified 1,594 reproduced clusters and a total of 634 genes (Table S4; Figure S1F). Among these genes, only five (ari-1, CG3812, CG8765, CG30497, and CG45186) had a TE insertion in the vicinity (±60 nt) of crosslink sites (see below), indicating that interactions between Aub and maternal mRNAs were largely independent of TE insertions. Analysis of reproduced crosslink distribution along genes revealed an enrichment in coding sequences and 3’ UTR and a depletion in 5’ UTR (Figure 1F).

We conclude that GFP-Aub iCLIP in early embryos allows specific identification of RNA directly bound by Aub. This includes complete piRNAs as well as several hundred maternal mRNAs.

Figure 1. Aub iCLIP Reveals Direct Interactions of Aub with piRNAs and Cellular mRNAs
(A) 32P-labeled Aub-RNA complexes showing decreasing size with increasing amounts of RNase I (+ to +++). No complexes were formed in the absence of UV or when rabbit serum was used for immunoprecipitation. The gel on the right represents an independent experiment in which embryos expressing GFP-Aub and GFP-AubAA were used in parallel. The asterisks indicate the size of Aub, whose presence was validated by western blot. The red squares indicate the regions of the membrane cut out for RNA extractions.
(B) Reproducibility of crosslink positions. Graph of the percentages of crosslinked nt with a given cDNA count that were reproduced in at least two biological replicates.
(C) Size distribution of reads from iCLIP1 that map to piRNA clusters.
(D) nt distribution in 23- to 29-nt reads mapping to piRNA clusters. The 1U bias is consistent with a proportion of these reads corresponding to full-length piRNAs.
(E) Plots of the number of uniquely mapped reads per gene in two iCLIP biological replicates. Each dot represents one gene. Pearson’s correlation coefficient (r) is indicated.
(F) Fold enrichment of reproduced crosslinks in gene regions, relative to the size of the corresponding regions in the whole genome.
See also Figure S1 and Tables S1–S3 and S4.
mRNAs Interacting with and Destabilized by Aub Are Involved in Germ Cell Development

Because Aub binding to nos mRNA leads to its decay in the somatic part of the embryo (Rouget et al., 2010), we sought to identify maternal mRNAs whose decay during the MZT depended on Aub or the piRNA pathway. We used microarrays to identify unstable mRNAs in wild-type embryos by comparing the transcriptomes derived from 0- to 2-hr and 2- to 4-hr embryos. mRNAs from 6,510 genes were present in 0- to 2-hr embryos, consistent with about half of Drosophila genes being maternally expressed (Thomsen et al., 2010). Microarray analyses identified 3,138 genes whose mRNAs were significantly destabilized in 2- to 4-hr embryos (fold change 1.2–151), representing 48.2% of maternally expressed genes. We then analyzed which of these unstable mRNAs are stabilized in embryos from aub or spn-E mutant mothers (aub or spn-E mutant embryos) by comparing mutant and wild-type embryonic transcriptomes. In 0- to 2-hr embryos, 547 and 571 genes produced unstable mRNAs that were stabilized in aub or spn-E mutant embryos, respectively. In 2- to 4-hr embryos, the number of stabilized mRNAs in aub and spn-E mutants increased (Figures 2A and 2B). This is consistent with maternal mRNA decay being more prominent after 2 hr of development (Thomsen et al., 2010) and with the role of Aub and the piRNA pathway in this decay, both directly and indirectly through developmental defects in the mutants, affecting mRNA decay.

The overlap between mRNAs bound by Aub and those stabilized in aub and spn-E mutant embryos identified mRNAs directly regulated by Aub. Among the 595 genes that were maternally expressed from microarray analysis and reproducibly crosslinked in iCLIP experiments, 182 (31%) produced mRNAs stabilized in Aub or the piRNA pathway. We used microarrays to identify unstable mRNAs whose decay during the MZT depended on Aub binding to nos mRNA leads to its decay in the somatic part of the embryo (Rouget et al., 2010), we sought to identify maternal mRNAs whose decay during the MZT depended on Aub or the piRNA pathway. We used microarrays to identify unstable mRNAs in wild-type embryos by comparing the transcriptomes derived from 0- to 2-hr and 2- to 4-hr embryos. mRNAs from 6,510 genes were present in 0- to 2-hr embryos, consistent with about half of Drosophila genes being maternally expressed (Thomsen et al., 2010). Microarray analyses identified 3,138 genes whose mRNAs were significantly destabilized in 2- to 4-hr embryos (fold change 1.2–151), representing 48.2% of maternally expressed genes. We then analyzed which of these unstable mRNAs are stabilized in embryos from aub or spn-E mutant mothers (aub or spn-E mutant embryos) by comparing mutant and wild-type embryonic transcriptomes. In 0- to 2-hr embryos, 547 and 571 genes produced unstable mRNAs that were stabilized in aub or spn-E mutant embryos, respectively. In 2- to 4-hr embryos, the number of stabilized mRNAs in aub and spn-E mutants increased (Figures 2A and 2B). This is consistent with maternal mRNA decay being more prominent after 2 hr of development (Thomsen et al., 2010) and with the role of Aub and the piRNA pathway in this decay, both directly and indirectly through developmental defects in the mutants, affecting mRNA decay.

The overlap between mRNAs bound by Aub and those stabilized in aub and spn-E mutant embryos identified mRNAs directly regulated by Aub. Among the 595 genes that were maternally expressed from microarray analysis and reproducibly crosslinked in iCLIP experiments, 182 (31%) produced mRNAs stabilized in Aub or the piRNA pathway. We used microarrays to identify unstable mRNAs whose decay during the MZT depended on Aub binding to nos mRNA leads to its decay in the somatic part of the embryo (Rouget et al., 2010), we sought to identify maternal mRNAs whose decay during the MZT depended on Aub or the piRNA pathway. We used microarrays to identify unstable mRNAs in wild-type embryos by comparing the transcriptomes derived from 0- to 2-hr and 2- to 4-hr embryos. mRNAs from 6,510 genes were present in 0- to 2-hr embryos, consistent with about half of Drosophila genes being maternally expressed (Thomsen et al., 2010). Microarray analyses identified 3,138 genes whose mRNAs were significantly destabilized in 2- to 4-hr embryos (fold change 1.2–151), representing 48.2% of maternally expressed genes. We then analyzed which of these unstable mRNAs are stabilized in embryos from aub or spn-E mutant mothers (aub or spn-E mutant embryos) by comparing mutant and wild-type embryonic transcriptomes. In 0- to 2-hr embryos, 547 and 571 genes produced unstable mRNAs that were stabilized in aub or spn-E mutant embryos, respectively. In 2- to 4-hr embryos, the number of stabilized mRNAs in aub and spn-E mutants increased (Figures 2A and 2B). This is consistent with maternal mRNA decay being more prominent after 2 hr of development (Thomsen et al., 2010) and with the role of Aub and the piRNA pathway in this decay, both directly and indirectly through developmental defects in the mutants, affecting mRNA decay.
embryos (mnk encodes the Chk2 kinase) or unfertilized wild-type and mutant eggs. mRNA stabilization in mnk aub and mnk spn-E double mutants and in aub and spn-E unfertilized eggs was consistent with the role of Aub and the piRNA pathway in the degradation of these mRNAs, independently of embryonic developmental defects (Figures 3A–3D and S2). mRNA stabilization throughout aub mutant embryos was also validated using in situ hybridization experiments for five mRNAs (Figure 3E).

Among the 413 mRNAs that are bound by Aub, but not stabilized in aub and spn-E mutant embryos, 18% (n = 74) were destabilized in wild-type embryos from the microarray analysis and 61% (n = 254) were stable according to the study of maternal mRNAs at the MZT (Thomsen et al., 2010) (Table S4). mRNAs from these 413 genes were enriched in GO terms involved in cell cycle, metabolic processes, and developmental processes occurring during later embryogenesis, such as neurogenesis or organ development (Figure 2D).

These results show that Aub directly binds a number of maternal mRNAs involved in germ cell development, which accumulate in the germ plasm and primordial germ cells. Aub participates in their localization by promoting their decay in the somatic part of the embryo. Aub also binds to a set of stable mRNAs and might mediate their regulation through a different mechanism.

Implication of piRNAs in the Binding of Aub to Maternal mRNAs
We have shown previously that Aub and Smg form a complex involved in nos mRNA destabilization in the somatic part of the embryo (Rouget et al., 2010). Smg binds to stem-loop structures, SREs which contain the consensus motif CNGGNN0–4 in the loop (Chen et al., 2014). Consistent with Aub and Smg co-regulating mRNAs that undergo Aub-dependent destabilization, calculation of SRE scores in Aub-bound mRNAs revealed an enrichment of SREs in mRNAs stabilized in aub and spn-E mutant embryos, with a median SRE score of 17 (n = 182 genes), compared to mRNAs that were not stabilized in mutant embryos whose median SRE score was 6.1 (n = 413 genes).

As a first approach to determine whether Aub is loaded with piRNAs for its role in maternal mRNA regulation, we examined whether piRNAs were present in Aub-Smg complexes by analyzing small RNAs that coprecipitated with Smg in 0- to 2-hr embryos. microRNAs were enriched in Smg immunoprecipitates (Figures 4A and 4B; Table S5), consistent with the interaction of Smg with Argonaute 1 in early embryos (Pinder and Smibert, 2013). piRNAs were also found in Smg immunoprecipitates and displayed the same features as piRNAs present in total embryos: a 1U bias and a bias toward antisense sequences for piRNAs matching TEs (Figures 4C, 4D, and S3). Because these are the same features as those of piRNAs loaded into Aub (Brennecke et al., 2007), these results indicate that at least a fraction of Aub in complex with Smg was loaded with piRNAs.

We then investigated the potential of embryonic piRNAs to target Aub-bound mRNAs in the vicinity of crosslinks. We analyzed potential piRNA targeting within ±60 nt from the crosslink sites, using varying complementarities. This identified a number of mRNAs potentially targeted by piRNAs for each

Figure 3. Validation of Maternal mRNA Regulation by Aub
(A–D) mRNA quantification using qRT-PCR in wild-type and mutant 2- to 3-hr embryos (A and B) or 3- to 4-hr unfertilized eggs (C and D). RpL32 was used as a control mRNA for normalization. The levels of mRNA at 2 to 3 hr or 3 to 4 hr were normalized to the levels at 0 to 1 hr set to 100% for each genotype. The genotypes of females are indicated. Females were either crossed or not with wild-type males. Means are from two to four biological replicates. The error bars represent SE. *p < 0.05; **p < 0.01; ***p < 0.001 with a one-tailed Student’s t test.
(E) In situ hybridization of mRNAs analyzed in (A)–(D) in 2- to 4-hr wild-type and aub mutant embryos. The scale bars represent 30 μm.
See also Figures S2 and S7.
complementarity (Figure 4E). To address whether this targeting could result from random complementarity, we performed negative controls consisting of the same targeting with shuffled piRNA sequences. For each complementarity, the number of mRNAs targeted by shuffled piRNA sequences was significantly lower than that of mRNAs targeted by piRNAs, suggesting a potential for embryonic piRNAs to target Aub-crosslinked mRNAs. Measurements of the distance between Aub crosslink sites and the 5' end of guide piRNAs revealed a peak of piRNAs overlapping Aub crosslink sites (0–29 nt) and within a few nt (−1 to −6 nt; Figure 4F). The peak at 0 nt suggested Aub potential interaction with the first nt of the sequence targeted by the guide piRNA. Strikingly, this result is consistent with recent data showing Aub capacity to select a specific nt (A) at this first position (Wang et al., 2014). We therefore analyzed the nt distribution at the first targeted position and identified a 1A bias, in agreement with the reported Aub preference for 1A on the target mRNA (Figure 4G).

Together, these data support the notion that piRNAs are involved as guides for the interaction of Aub with maternal mRNAs.

Unloaded Aub Is Defective for Maternal mRNA Binding
To confirm the role of piRNAs in the binding of Aub to maternal mRNAs, we produced a mutant form of Aub unable to load piRNAs. The 5' and 3' ends of small RNAs interact with the MID and PAZ domains of Argonaute proteins, respectively (Elkayam et al., 2012; Schirle and MacRae, 2012). We replaced by alanines two conserved tyrosines involved in the interaction with piRNAs, in the PAZ domain of the GFP-Aub transgene (Y345A and Y346A; Figure 5A). The resulting mutant transgene GFP-AubAA could not rescue the maternal effect embryonic lethality and dorso-ventral patterning defects of aub mutants (Figure S1A). We verified that GFP-AubAA was unable to bind piRNAs by performing GFP immunoprecipitations in 0- to 2-hr embryos (Figure 5B). Strikingly, in contrast to GFP-Aub, GFP-AubAA neither accumulated in the nuage, a structure involved in piRNA biogenesis that surrounds nurse cell nuclei in ovaries, nor in the germ plasm of oocytes and embryos (Figures 5C, S4A, and S4B). This shows that Aub accumulation in the nuage and germ plasm requires its loading with piRNAs. We performed an iCLIP experiment with GFP-AubAA, concomitantly with GFP-Aub iCLIP3. A low proportion of reads corresponded to piRNAs and mRNAs, whereas a high proportion of reads corresponded to rRNAs, suggesting that the sequences recovered in GFP-AubAA iCLIP were mostly unspecific (Table S1; Figure S1E). Consistent with this, the reproducibility of crosslinks indicates that nt 2–16 and 20–20 of piRNAs were considered, respectively. Statistical analysis was performed using the Fisher’s exact test.

Figure 4. Smg Is in Complex with piRNAs, and Embryonic piRNAs Have the Potential to Target Aub-Crosslinked mRNAs with Imperfect Base Pairing
(A and B) Length distribution of small RNAs that either coimmunoprecipitated with Smg in 0- to 2-hr embryos (A) or were present in 0- to 2-hr embryos (B). mRNAs were identified by their sequences. (C and D) Length distribution of TE-derived small RNAs in sense (blue) and antisense (red) orientations. (E) Capacity of embryonic piRNAs to target Aub-bound mRNAs with different complementarities within ±60 nt of crosslink sites. The first nt of piRNAs was not considered in the base pairing. Base pairing with 16-nm or 20-nm seed in
The formation of complexes between GFP-Aub AA and wild-type produced in Aub iCLIP experiments (Table S6). This could reflect type embryos (Figure S4 D; Table S2). Moreover, in Aub iCLIP in osk for the vast majority of Aub interactions with mRNAs. The production of a ping-pong signature and served as a proof of principle of this genic ping-pong (Figure S5). Among the five genes identified as containing a TE insertion within described in Figure 4 E (Table S7). Among the five genes identified as containing a TE insertion within (Figure 6 A). Although the occurrence of genic piRNAs was low, these ping-pong signatures indicated the potential of maternal mRNAs to be cleaved by Aub. In total, out of the 634 genes producing mRNAs bound by Aub, 55 showed a ping-pong signature between sense and antisense piRNAs, the ping-pong signature. Action catalyzed by PIWI proteins produces a 10-nt overlap between sense and antisense piRNAs. The cleavage reaction catalyzed by PIWI proteins produces a 10-nt overlap between sense and antisense piRNAs. Because GFP-Aub AA does not accumulate in the germ plasm, our results could reveal a weak potential for Aub to bind mRNAs independently of piRNAs. Because GFP-Aub AA does not accumulate in the germ plasm, we performed a control GFP-Aub iCLIP in osk mutant embryos that fail to accumulate GFP-Aub in the germ plasm (Figure S4C). In contrast to in Aub AA iCLIP, the reproducibility of crosslinks in Aub iCLIP in osk mutant embryos was similar to that in wild-type embryos (Figure S4D; Table S2). Moreover, in osk embryos, Aub interacted with mRNAs encoding germ cell determinants as it did in wild-type embryos (Figures 4E–4G). This showed that Aub interaction with these mRNAs could occur in the somatic part of the embryo and did not result from concentration in the germ plasm. We conclude that the loading of Aub with piRNAs is required for the vast majority of Aub interactions with mRNAs. Production of Genic piRNAs upon Targeting of mRNAs by piRNAs

In the germline, piRNA biogenesis involves a mechanism known as ping-pong, in which the targeting of an mRNA by a piRNA induces the production of a secondary piRNA. The cleavage reaction catalyzed by PIWI proteins produces a 10-nt overlap between sense and antisense piRNAs, the ping-pong signature. We noticed such ping-pong signatures between piRNAs targeting Aub-bound mRNAs and piRNAs produced by these mRNAs (Figure 6A). Although the occurrence of genic piRNAs was low, these ping-pong signatures indicated the potential of maternal mRNAs to be cleaved by Aub. In total, out of the 634 genes producing mRNAs bound by Aub, 55 showed a ping-pong signature within ±60 nt from crosslink sites, using piRNA base pairing described in Figure 4E (Table S7). Among the five genes identified as containing a TE insertion within ±60 nt of crosslink sites, four produced a ping-pong signature and served as a proof of principle of this genic ping-pong (Figure S5).

This suggested two possible levels of mRNA regulation by Aub that involved either the recruitment of the CCR4–NOT deadenylase complex as previously described for nos mRNA (Rouget et al., 2010) or endonucleolytic cleavage. We reasoned that, if Aub could destabilize mRNAs by involving either of these two mechanisms, one might be favored through impairment of the other. We therefore used the CCR4 deadenylase mutant twin and examined whether the levels of piRNAs produced from transcripts would increase when deadenylation is reduced. We sequenced small RNA libraries generated from wild-type
embryos and embryos from twin mutant females (Table S5; Figure S6). piRNAs from transcripts and transcript 3’ UTRs were more abundant in twin mutant than in wild-type embryos (Figures 6B and 6C). This is consistent with both possible modes of action for Aub—activator of deadenylation or endonuclease—in mRNA regulation.

We sought to functionally validate the targeting of dunce (dnc) mRNA by piRNAs from the R1 TE, which induces a ping-pong signature. R1 piRNA target site overlapped Aub crosslink in dnc 3’ UTR (Figures 6D and 6E). dnc mRNA was stabilized in aub and spn-E mutant embryos and unfertilized eggs and in mkn aub and mkn spn-E double-mutant embryos (Figures 6F and 6G), suggesting a direct role of Aub in dnc mRNA destabilization in the embryo. dnc encodes a cyclic AMP (cAMP)-specific phosphodiesterase acting in cAMP catabolism. A tight regulation of cAMP levels is required for female fertility; a decrease in cAMP phosphodiesterase acting in cAMP catabolism. A tight regulation of cAMP levels is required for female fertility; a decrease in cAMP

**DISCUSSION**

The function of the piRNA pathway in the regulation of TE expression is well established (Guzzardo et al., 2013; Ishizu et al., 2012). In contrast, other functions of this pathway in gene regulation are only now emerging and remain poorly understood. Here, we provide evidence for a widespread role of Aub and piRNAs in maternal mRNA decay in the early *Drosophila* embryo during the MZT. An essential biological function of this regulation is to participate in the localization of specific mRNAs involved in germ cell specification and development through their destabilization in the somatic part of the embryo. Furthermore, Aub directly binds another set of maternal mRNAs, which are stable and appear to be required at later steps of embryogenesis.

PIWI proteins are loaded with piRNAs, which serve as guides to target mRNAs by complementarity; this RNA-RNA association underlies PIWI protein mode of action in TE regulation. One important question concerning the role of Aub in cellular mRNA regulation was whether piRNAs were also involved. In mouse, MIWI has been shown to interact massively with cellular mRNAs in spermatids based on CLIP experiments (Gou et al., 2014; Vourekas et al., 2012; Zhang et al., 2015), and it was proposed to bind spermiogenic mRNAs independently of piRNAs in round spermatids (Vourekas et al., 2012). However, more-recent reports have validated the implication of pachytene piRNAs in the regulation of mRNAs involved in spermatogenesis during spermatocyte and spermatid stages (Goh et al., 2015; Gou et al., 2014; Watanabe et al., 2015; Zhang et al., 2015). piRNAs guide MIWI-dependent cleavage or deadenylation of these mRNAs, leading to their destabilization.

We find that unloaded Aub has a very weak capacity to bind mRNAs in iCLIP experiment, indicating that the majority of Aub interactions with cellular mRNAs depend on piRNAs. This result is strengthened by a number of mRNAs showing a ping-pong signature close to Aub crosslinks, thus implicating their targeting by piRNAs. The functional validation of one of these mRNAs establishes a role for piRNA targeting in mRNA destabilization and embryonic development.

How could Aub achieve a specific mRNA regulation? Maternal mRNAs are produced in nurse cell nuclei in the ovary and must transit through the nuage during their export to the cytoplasm.
We propose that, in the nuage, maternal mRNAs undergo a general scanning by piRNAs loaded into Aub, as in the case of TE mRNAs, because mRNAs from TEs and cellular genes might not be discriminated (Figure 7). TE mRNAs are highly targeted by piRNAs, whereas the targeting of maternal mRNAs should be more modest due to the lack of perfect base pairing with piRNAs. Aub binding to maternal mRNAs based on imperfect base pairing with piRNAs could contribute to regulation, as it has been validated for different PIWI proteins (Gou et al., 2014; Lee et al., 2012; Rouget et al., 2010). However, multiple interactions might be required to achieve effective regulation. Additional proteins might also contribute to potentiate piRNA-mediated regulation. Although Aub is expected to bind maternal mRNAs in the nurse cells, these mRNAs are not massively degraded in the ovary, whereas they are in the embryo. This strongly suggests that decay requires another component expressed in the embryo. We determined that SREs are enriched in mRNAs bound and destabilized by Aub in the embryo. This is consistent with the presence of piRNAs in Smg mRNPs and suggests that Smg, which is specifically expressed in early embryos, cooperates with Aub for the destabilization of these maternal mRNAs.

Our data reveal two possible mechanisms of mRNA regulation by Aub (Figure 7): the first involves Aub-dependent mRNA deadenylation through the recruitment of the CCR4-NOT complex, which we previously reported for nos mRNA (Rouget et al., 2010). Interestingly, this mechanism involving deadenylation and mRNA decay factors also underlie some aspect of piRNA-dependent repression of TE mRNAs (Lim et al., 2009). The second mechanism based on the endonucleolytic activity of Aub leads to the production of generic piRNAs. Both mechanisms are also implicated in cellular mRNA regulation by MIWI (Goh et al., 2015; Gou et al., 2014; Watanabe et al., 2015; Zhang et al., 2015). Our results reveal increased levels of generic piRNAs in mutant embryos for the CCR4 deadenylase, suggesting that both mechanisms might compete for the repression of the same mRNAs.

Aub is expressed throughout the whole syncytial embryo (Mani et al., 2014; Rouget et al., 2010) and accumulates in the germ plasm, where it is involved in primordial germ cell (PGC) formation (Harris and Macdonald, 2001). This function might be partly indirect because aub mutant embryos that do not develop PGCs also show DNA fragmentation resulting from earlier defects during oogenesis (Khurana et al., 2010). However, our results propose a more-direct function of Aub in PGC formation: we uncover a new role for Aub in the direct binding of a set of maternal mRNAs involved in PGC specification and development, which could underlie this Aub developmental function. Stabilization of these mRNAs in aub mutant embryos and unfertilized eggs indicates that Aub is involved in the decay of these mRNAs in the somatic part of the embryo. It is likely that the interaction between Aub and mRNAs encoding PGC determinants is maintained in the germ plasm where Aub accumulates, even if it does not lead to mRNA decay. In particular, Aub has been shown to colocalize with nos, pgc, and gcl mRNAs in RNA granules in PGCs (Rangan et al., 2009). Consistent with this, ectopic localization of Aub at the anterior pole of embryos through anterior localization of Osk (Ephrussi and Lehmann, 1992) resulted in the recruitment of mRNAs encoding PGC determinants at the anterior pole (Figure S7). The mechanism behind this switch in Aub function between soma and germ plasm in the early embryo remains to be addressed.

Strikingly, we found that unloaded Aub does not accumulate in the germ plasm, implying that Aub is loaded with piRNAs when it binds PGC determinant mRNAs in the germ plasm. Therefore, Aub would have two inter-related functions in PGC biology: the localization in the germ plasm of mRNAs encoding PGC determinants and the maternal transfer of a pool of piRNAs in the PGCs to start piRNA production and function in TE repression in the future germ cells.

It is likely that the general decay of maternal mRNAs by Aub and piRNAs in the somatic part of the embryo plays a substantial role in embryonic patterning. Indeed, ectopic anterior expression of PGC determinants such as osk or pgc is known to prevent anterior-posterior patterning of the embryo (Ephrussi and
Lehmann, 1992; Hanyu-Nakamura et al., 2008). We show here that piRNA-dependent regulation of dnc mRNA is important for embryonic development. This is consistent with the role of piRNA-dependent regulation of nos mRNA for embryonic patterning (Rouget et al., 2010). In both these cases, however, preventing this regulation did not lead to complete embryonic lethality, indicating that piRNA-dependent regulation would be partly redundant with other levels of post-transcriptional regulation and involved in fine-tuning gene expression. Because piRNA populations can evolve rapidly within and between species, mRNA regulation by piRNAs would also change rapidly, thereby providing a potential basis for adaptive responses.

EXPERIMENTAL PROCEDURES

CRISPR-Cas9 Genome Editing

The sgRNA was designed to target the R1 piRNA target site in the long 3’ UTR of dnc mRNA. The sgRNA was produced as previously described (Bassett et al., 2013) with the following modifications. PCR was performed with Super-prime Accu Mix I (Invitrogen) using the annealed primers dncCRISPR-F, which contains the dnc target sequence, and sgCRISPR-R. In vitro transcription from 1 μg of the obtained PCR fragment was performed with T7 RNA polymerase (Roche). The sgRNA was extracted with acid phenol chloroform (Ambion), precipitated, dissolved in injection buffer, and stored at −80°C. Two hundred picoliters of a 1 μg/μl sgRNA solution was injected in 0–20 min old act5C-cas9 embryos (Port et al., 2014). Surviving flies were crossed with a stock containing the FM7c balancer. Offspring were screened by T7 endonuclease I assay as described previously (Hwang et al., 2013), and mutations were determined by sequencing. The act5C-cas9 transgene was removed from dnc mutant stocks by recombination with a w* stock before further analysis.

ACCESSION NUMBERS

Deep-sequencing and microarray data generated in this study have been deposited to the NCBI GEO and are available under accession number GEO: GSE70778.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.07.000.

AUTHOR CONTRIBUTIONS

B.B., J.D., C.P., F.P., and T.G. performed biological experiments; S.P., C.A., and J.D. analyzed the data and wrote the manuscript; and all authors discussed the manuscript.

ACKNOWLEDGMENTS

We are very grateful to A. Ephrussi, P. Lasko, P. MacDonald, M. and H. Siomi, and C. Simbert for their gifts of fly stocks and antibodies; J. König and J. Ule for their advice on iCLIP; S. Bullock for his support and advice on CRISPR-Cas9 editing; and A.C. Meunier for technical assistance. This work was supported by the CNRS UPR1142 to M.S. and S.C.; the ANR Blanche (ANR-06-BLAN-0505) to T.C.; and the ANR JCJC (ESSOR-JCJC-1604 01) to S.C. B.B. was supported by ANR and DGF; J.D. by Fondation ARC; and S.P., T.G., and C.A. by FRM.

Received: February 24, 2015
Revised: June 17, 2015
Accepted: July 14, 2015
Published: August 6, 2015

REFERENCES

Ånkö, M.L., Müller-McNicoll, M., Brandl, H., Curk, T., Gorup, C., Henry, I., Ule, J., and Neugebauer, K.M. (2012). The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. Genome Biol. 13, R17.

Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.L. (2013). Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep. 4, 220–228.

Bellen, H.J., and Kiger, J.A. (1988). Maternal effects of general and regional specificity on embryos of Drosophila melanogaster caused by dunce and rutabaga mutant combinations. Dev. Genes Evol. 197, 258–268.

Beyret, E., Liu, N., and Lin, H. (2012). piRNA biogenesis during adult somatogonial genesis in mice is independent of the ping-pong mechanism. Cell Res. 22, 1429–1439.

Brennecce, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089–1103.

Chen, L., Dumeille, J.G., Li, X., Cheng, M.H., Yang, Z., Laver, J.D., Siddiqui, N.U., Westwood, J.T., Morris, Q., Lipshitz, H.D., and Smibert, C.A. (2014). Global regulation of mRNA translation and stability in the early Drosophila embryo by the Smaug RNA-binding protein. Genome Biol. 15, R4.

De Renzis, S., Elemento, O., Tavazozo, S., and Wieschaus, E.F. (2007). Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. PLoS Biol. 5, e117.

Elkayam, E., Kuhn, C.D., Tocilli, A., Haase, A.D., Greene, E.M., Hannon, G.J., and Joshua-Tor, L. (2012). The structure of human argonaute-2 in complex with miR-20a. Cell 150, 100–110.

Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. Nature 358, 387–392.

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

Goh, W.S., Falcioni, I., Tam, O.H., Burgess, R., Meikar, O., Kojima, K., and Hannon, G.J. (2015). piRNA-directed cleavage of meiotc transcripts regulates spermatogenesis. Genes Dev. 29, 1032–1044.

Gou, L.T., Dai, P., Yang, J.H., Xue, Y., Hu, Y.P., Zhou, Y., Kang, J.Y., Wang, X., Li, H., Hua, M.M., et al. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res. 24, 680–700.

Guzzardo, P.M., Muerdter, F., and Hannon, G.J. (2013). The piRNA pathway in flies: highlights and future directions. Curr. Opin. Genet. Dev. 23, 44–52.

Hanyu-Nakamura, K., Sonobe-Nojima, H., Tanigawa, A., Lasko, P., and Nakamura, A. (2008). Drosophila Pgc protein inhibits P-TEFb recruitment to chromatin in promordial germ cells. Nature 457, 730–733.

Harris, A.N., and Macdonald, P.M. (2001). Aubergine encodes a Drosophila polaron granule component required for pole cell formation and related to elf2C. Development 128, 2823–2832.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. 31, 227–229.

Ishizu, H., Siomi, H., and Siomi, M.C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev. 26, 2361–2373.

Jongens, T.A., Ackerman, L.D., Swedlow, J.R., Jan, L.Y., and Jan, Y.N. (1994). Germ cell-less encodes a cell type-specific nuclear pore-associated protein and functions early in the germ-cell specification pathway of Drosophila. Genes Dev. 8, 2123–2136.
Khurana, J.S., Xu, J., Weng, Z., and Theurkauf, W.E. (2010). Distinct functions for the Drosophila piRNA pathway in genome maintenance and telomere protection. PLoS Genet. 6, e1001246.

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.

Kuch, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Arai, Y., Ishihara, G., Kawaoka, S., Sugano, S., Shimada, T., et al. (2014). A single female-specific piRNA is the primary determiner of sex in the silkworm. Nature 509, 633–636.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat. Struct. Mol. Biol.

Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat. Struct. Mol. Biol.

Kiuchi, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature 461, 1296–1299.

Schirle, N.T., and MacRae, I.J. (2012). The crystal structure of human Argonaute2. Science 336, 1037–1040.

Schneider, L.E., and Spradling, A.C. (1997). The Drosophila G-protein-coupled receptor kinase homologue Gprk2 is required for egg morphogenesis. Development 124, 2591–2602.

Semotok, J.L., Cooperstock, R.L., Pinder, B.D., Vari, H.K., Lipshitz, H.D., and Smibert, C.A. (2005). Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early Drosophila embryo. Curr. Biol. 15, 284–294.

Song, Y., Fee, L., Lee, T.H., and Wharton, R.P. (2007). The molecular chaperone Hsp90 is required for mRNA localization in Drosophila melanogaster embryos. Genetics 176, 2213–2222.

Tadros, W., Goldman, A.L., Babak, T., Menzies, F., Vardy, L., Orr-Weaver, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell 131, 174–187.

Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte, D., Jr., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150, 78–87.

Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., and Zamore, P.D. (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol. Cell 50, 67–81.

Lim, A.K., Tao, L., and Kai, T. (2009). piRNAs mediate posttranscriptional retrome silencing and localization to pi-bodies in the Drosophila germline. J. Cell Biol. 186, 333–342.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137, 522–535.

Mani, S.R., Megosh, H., and Lin, H. (2014). PIWI proteins are essential for early Drosophila embryogenesis. Dev. Biol. 385, 340–349.

Peng, J.C., and Lin, H. (2013). Beyond transposons: the epigenetic and somatic functions of the Piwi-piRNA mechanism. Curr. Opin. Cell Biol. 25, 190–194.

Pinder, B.D., and Smibert, C.A. (2013). microRNA-independent recruitment of Argonaute 1 to nanos mRNA through the Smaug RNA-binding protein. EMBO Rep. 14, 80–86.

Port, F., Chen, H.M., Lee, T., and Bullock, S.L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc. Natl. Acad. Sci. USA 111, E2967–E2976.

Rangan, P., DeGennaro, M., Jaime-Bustamante, K., Coux, R.X., Martinho, R.G., and Lehmann, R. (2009). Temporal and spatial control of germ-plasm RNA. Curr. Biol. 19, 72–77.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3’UTR-directed primary piRNAs. Curr. Biol. 19, 2066–2076.

Rouget, C., Papin, C., Bourieux, A., Meunier, A.C., Franco, B., Robine, N., Lai, E.C., Pelisson, A., and Simonelig, M. (2010). Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. Nature 467, 1128–1132.

Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature 461, 1296–1299.

Song, Y., Fee, L., Lee, T.H., and Wharton, R.P. (2007). The molecular chaperone Hsp90 is required for mRNA localization in Drosophila melanogaster embryos. Genetics 176, 2213–2222.

Tadros, W., Goldman, A.L., Babak, T., Menzies, F., Vardy, L., Orr-Weaver, T., Hughes, T.R., Westwood, J.T., Smibert, C.A., and Lipshitz, H.D. (2007). SMAUG is a major regulator of maternal mRNA destabilization in Drosophila and its translation is activated by the PAN GU kinase. Dev. Cell 12, 143–155.

Thomsen, S., Anders, S., Janga, S.C., Huber, W., and Alonso, C.R. (2010). Genome-wide analysis of mRNA decay patterns during early Drosophila development. Genome Biol. 11, R93.

Varekakas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B.D., and Mourelatos, Z. (2012). Mill and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermogenesis. Nat. Struct. Mol. Biol. 19, 773–781.

Wang, W., Yishokaw, M., Han, B.W., Izumi, N., Tomari, Y., Weng, Z., and Zanmore, P.D. (2014). The initial uridine of primary piRNAs does not create the tenth adenine that is the hallmark of secondary piRNAs. Mol. Cell 56, 708–716.

Watanabe, T., and Lin, H. (2014). Posttranscriptional regulation of gene expression by Piwi proteins and piRNAs. Mol. Cell 56, 18–27.

Watanabe, T., Cheng, E.C., Zhong, M., and Lin, H. (2015). Retrotransposons and pseudogenes regulate mRNAs and IncRNAs via the piRNA pathway in the germline. Genome Res. 25, 368–380.

Weick, E.M., and Miska, E.A. (2014). piRNAs: from biogenesis to function. Development 141, 3458–3471.

Zaessinger, S., Busseau, I., and Simonelig, M. (2006). Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. Development 133, 4573–4583.
Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo

Bridlin Barckmann, Stéphanie Pierson, Jérémy Dufourt, Catherine Papin, Claudia Armenise, Fillip Port, Thomas Grentzinger, Séverine Chambeyron, Grégory Baronian, Jean-Pierre Desvignes, Tomaz Curk, and Martine Simonelig
A table showing the results of experiments with different treatments.

- **Embryonic lethality**
  - AubAA+; nos-Gal4: 3.2% (n=899)
  - AubAA+; nos-Gal4/+: 2.4% (n=1018)

- **Fused appendages**
  - AubAA+; nos-Gal4: 0.4% (n=899)
  - AubAA+; nos-Gal4/+: 0.5% (n=1018)

---

**Figure S1**

A histogram showing the distribution of nucleotides in iCLIP1, iCLIP2, and iCLIP3.

- **iCLIP2**
- **iCLIP3**

A bar chart showing the normalized number of reads for different RNA types.

- **Read length (nt)**
  - 0 500 1000 1500 2000 2500 3000 3500

- **Normalized number of reads**
  - 0 1000 2000 3000 4000 5000 6000

---

**Embryonic lethality**

- AubAA+; nos-Gal4: 3.2% (n=899)
- AubAA+; nos-Gal4/+: 2.4% (n=1018)

---

**Fused appendages**

- AubAA+; nos-Gal4: 0.4% (n=899)
- AubAA+; nos-Gal4/+: 0.5% (n=1018)
Figure S2
Figure S3

A  Smaug IP

B  Total embryos

nucleotide position

nucleotide distribution

%  

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

G
C
A
U

1 3 5 7 9 11 13 15 17 19 21 23 25 27 29

Smaug IP Total embryos
Overlap between mRNAs bound by Aub in osk^- and mRNAs stabilized in aub^- and spn-E^- embryos

- n=380 genes producing mRNAs reproducibly cross-linked in osk^- (1836 genes)
- n=1971 genes producing mRNA stabilized in aub^- and spn-E^- embryos

GO terms (135 genes)

- Biological process involved in reproduction: 6.89e-7
- Single organism reproductive process: 5.46e-5
- Mitotic cell cycle: 1.25e-4
- Germ cell development: 3.53e-4
- Neurogenesis: 7.93e-11

GO terms (218 genes)

- Biological process
  - Cellular process involved in reproduction: 2.46e-14
  - Multicellular organismal development: 5.79e-14
  - Nervous system development: 7.93e-11
- Neurogenesis: 3.76e-6
- Translation: 7.85e-5

Figure S4
Figure S6
Figure S7
Supplemental Figure Legends

Figure S1. GFP-Aub iCLIP in 0-2 h-embryos, Related to Figure 1.
(A) Genetic analysis of UASp-GFP-Aub and UASp-GFP-Aub\textsuperscript{AA} transgenes. UASp-GFP-Aub\textsuperscript{AA} has two point mutations in the PAZ domain that prevent loading with piRNAs (Figure 5A). Capacity of the transgenes to rescue \textit{aub\textsuperscript{NN2/aub\textsuperscript{QC42}}} mutant phenotypes of maternal effect embryonic lethality and fused dorsal appendages (left table), GFP-Aub\textsuperscript{AA} does not rescue \textit{aub} mutant phenotypes. None of the transgenes induce embryonic defects when expressed in a wild-type background (right table).
(B) Reproducibility of iCLIP between biological replicates. Black bars indicate the number of cross-linked nt in iCLIP1 that are reproduced with a given offset in iCLIP2.
(C) Size distribution of reads that map to piRNA clusters in iCLIP2 and iCLIP3.
(D) nt distribution in 23-29 nt-reads mapping to piRNA clusters in iCLIP2 and iCLIP3.
(E) Mapping of Aub and Aub\textsuperscript{AA} iCLIP reads to the Drosophila genome.
(F) Two examples of genes (\textit{nos} and \textit{germ cell-less} (\textit{gcl})) with reproduced cross-links. Thin boxes are 5'- and 3'UTRs, lines are introns, and thick boxes are exons. Cross-link clusters (≥5 nt) with cDNA counts ≥5 reproduced in at least two experiments are indicated in red. \textit{nos} mRNA, which was previously shown to be regulated by Aub (Rouget et al., 2010), was identified with clusters containing ≥5 cDNA counts in all three independent iCLIP experiments.

Figure S2. Validation of mRNA regulation by Aub, Related to Figure 3.
(A-D) Quantification of mRNA levels using RT-qPCR in wild-type and mutant 2-3 h-embryos (A, B) or 3-4 h-unfertilized eggs (C, D). \textit{adam} and \textit{mRPl43} are negative control mRNAs that are not cross-linked in Aub iCLIPs. \textit{RpL32} was used as a control mRNA for normalization. The levels of mRNA at 2-3 h or 3-4 h were normalized to the levels at 0-1 h set to 100% for each genotype. The genotypes of females are indicated. The females were crossed, or not, with wild-type males. For \textit{cycB} and \textit{smg} mRNAs, RT-qPCR in unfertilized eggs revealed a high contribution of the zygotic pathway of mRNA decay, which prevents recording the potential role of Aub or Spn-E. Means are from two to four biological replicates. Error bars represent standard error. * \textit{p}-value <0.05, ** \textit{p}-value <0.01, *** \textit{p}-value <0.001, ns: non-significant using the one-tailed Student's t-Test.
(E) \textit{aub} mutant embryos are fertilized. Because a proportion of \textit{aub} mutant embryos do not develop, we checked that they were fertilized using males expressing \textit{Don juan-GFP} (\textit{dj-GFP}), which marks the sperm flagellum. Immunostaining with anti-GFP showing that 53% (n=138) of embryos from \textit{aub\textsuperscript{NN2/aub\textsuperscript{QC42}}} females crossed with \textit{dj-GFP/CyO} males contain a sperm flagellum, indicating that they are fertilized (since \textit{dj-GFP} is heterozygous, 50% of the embryos are expected to contain a GFP-marked sperm flagellum if all embryos are fertilized). Scale bar: 30 μm.

Figure S3. Analysis of small RNA libraries from total 0-2 h-embryos and Smg immunoprecipitation (IP) in 0-2 h-embryos, Related to Figure 4.
(A, B) nt distribution in 23-29 nt-reads corresponding to piRNAs, showing a 1U-bias.

Figure S4. Characterization of GFP-Aub\textsuperscript{AA} mutated in the PAZ domain and GFP-Aub iCLIP in \textit{osk\textsuperscript{54}} mutant embryos, Related to Figure 5.
(A) Expression of GFP-Aub and GFP-Aub\textsuperscript{AA} in the ovaries using \textit{nos-Gal4}, showing that GFP-Aub\textsuperscript{AA} does not accumulate in the nuage in nurse cells (left panel), or at the posterior pole in the oocyte (right panel). Double staining with anti-GFP (green) and anti-Vasa (red) to mark the nuage (left panel), and with anti-GFP (green) and anti-Osk (red) to mark the germ plasm in stage 10 oocyte (right panel). DAPI (blue) was used to label DNA, the merge is shown on the right of each panel. Scale bars: 30 μm in (A) and (C).
(B) Quantification of GFP-Aub and GFP-Aub\textsuperscript{AA} localization at the posterior pole of stage 10-14 oocytes. The presence of GFP was determined using GFP fluorescence.
(C) Immunostaining with anti-GFP (green) of UASp-GFP-Aub/\textit{nos-Gal4} and \textit{osk\textsuperscript{54}} UASp-GFP-Aub /\textit{osk\textsuperscript{54}} \textit{nos-Gal4} embryos showing the lack of posterior accumulation of Aub in \textit{osk} mutant embryos.
(D) Reproducibility of cross-link sites within GFP-Aub iCLIP in \textit{osk\textsuperscript{54}} mutant embryos. iCLIP4 is a GFP-Aub iCLIP in wild-type (\textit{osk\textsuperscript{-}}) embryos performed in parallel to that in \textit{osk} mutant embryos. For these two iCLIPs the time of UV cross-link was reduced (three five-minute pulses), which might explain the low number of reads (Table S1). The reproducibility of cross-link sites was similar in iCLIP in \textit{osk\textsuperscript{54}} embryos (42% of cross-linked genes with cDNA counts ≥5) to that in iCLIP1-3 (26% to 44% of cross-linked genes with cDNA counts ≥5, Table S2), whereas it was very low in GFP-Aub\textsuperscript{AA} iCLIP (2.9% of cross-linked genes with cDNA counts ≥5, Table S2).
(E) Venn diagram of mRNAs interacting with Aub in \textit{osk} mutant embryos and mRNAs stabilized in \textit{aub} and \textit{spn-E} mutant embryos.
(F) GO term analysis of mRNAs bound by Aub in osk mutant embryos, stabilized (top) or not (bottom) in aub and spn-E mutant embryos, showing that mRNAs encoding germ plasm determinants interacted with Aub in the absence of germ plasm.

(G) mRNAs bound by Aub in osk mutant embryos and stabilized in aub and spn-E mutant embryos are enriched in mRNAs showing a posterior accumulation in wild-type embryos. Annotation from Fly-FISH and BDGP insitu expression patterns were used to determine mRNA localization.

**Figure S5. Ping-pong signature in Aub-bound mRNAs, Related to Figure 6.**

mRNAs containing a TE insertion and showing a ping-pong signature, both within ±60 nt from a cross-link site. Genes are in dark blue: Thick boxes are exons, thin boxes are UTRs, and lines are introns. The clusters of reproduced cross-links (≥5 cDNA counts in ±5 nt clusters, reproduced in ≥2 iCLIPs) are indicated in red above the genes. TE insertions are in black. The sequence of the regions showing a ping-pong signature is shown. The 10-nt overlap between sense and antisense piRNAs is in green. The cross-linked nt are in red. The occurrence of piRNAs is indicated.

**Figure S6. Analysis of small RNA libraries from wild-type and twin mutant embryos, Related to Figure 6.**

(A, B) Length distribution of small RNAs in 0-2 h-embryos from wild-type and twin^{12209} mutant females crossed with wild-type males. miRNAs were identified by their sequences.

(C, D) nt distribution in 23-29 nt reads corresponding to piRNAs. The 1U-bias is slightly less pronounced for piRNAs from twin^{12209} embryos; this is consistent with the increased amounts of transcript piRNAs in these embryos, since transcript piRNAs have a lower 1U-bias (66%) than TE piRNAs (Robine et al., 2009).

**Figure S7. Recruitment of germ plasm mRNAs with Aub protein to the anterior pole upon Osk localization at the anterior pole, Related to Figure 3.**

(A) Osk and Aub immunostaining in osk-bcd3'UTR embryos (embryos expressing osk coding sequence followed by bicoid 3'UTR (Ephrussi and Lehmann, 1992)), showing accumulation of Osk and Aub at the anterior pole. Scale bar: 30 µm.

(B) in situ hybrydization experiments in osk-bcd3'UTR embryos showing the accumulation of germ plasm mRNAs at the anterior pole (arrows), in the region of Aub accumulation. Anterior pole is to the left.
Table S1. Analysis of reads in iCLIP libraries, Related to Figures 1 and 5.

|                        | iCLIP1  | iCLIP2  | iCLIP3  | AubAA iCLIP | iCLIP4 * | iCLIP in osk * |
|------------------------|---------|---------|---------|-------------|----------|---------------|
| Total number of sequenced reads | 36 278 327 | 32 955 607 | 27 848 743 | 9 305 990 | 3 165 539 | 4 877 100 |
| Total number of mapped reads | 23 645 147 | 16 943 493 | 15 879 101 | 5 770 961 | 439 258 | 3 129 830 |
| Total number of mapped reads after removal of PCR duplicates | 651 699 | 265 646 | 2 620 195 | 107 437 | 2 301 | 19 661 |
| piRNA reads             | 277 319 | 72 327 | 184 869 | 674 | 69 | 1 136 |
| % piRNA reads           | 42.5%  | 27.2%  | 7.1%   | 0.6% | 3.0% | 5.8% |
| mRNA                   | 67 799 | 31501 | 1 033 732 | 12 036 | 457 | 4331 |
| % mRNAs from genes     | 10%    | 12%    | 39%    | 11%  | 20% | 22% |
| rRNA                   | 126 623 | 73670 | 1 125 182 | 89 250 | 1 427 | 10 693 |
| % rRNA                 | 19%    | 28%    | 43%    | 83%  | 62% | 54% |
| tRNA                   | 29 393 | 15001 | 74 358 | 444 | 14 | 65 |
| % tRNA                 | 5%     | 6%     | 3%     | 0%   | 1% | 0% |
| snoRNA                 | 6 380 | 4847 | 27 391 | 772 | 10 | 53 |
| % snoRNA               | 1%     | 2%     | 1%     | 1%   | 0% | 0% |
| snRNA                  | 2 372 | 1216 | 16 792 | 458 | 0 | 39 |
| % snRNA                | 0%     | 0%     | 1%     | 0%   | 0% | 0% |
| miRNA                  | 1 677 | 1607 | 2 493 | 34 | 0 | 6 |
| %miRNA                 | 0%     | 1%     | 0%     | 0%   | 0% | 0% |
| intergenic             | 17 961 | 8 413 | 19 738 | 765 | 128 | 1002 |
| %intergenic            | 3%     | 3%     | 1%     | 1%   | 6% | 5% |
| others                 | 122 175 | 57 064 | 135 640 | 3 004 | 197 | 2 336 |
| % others               | 19%    | 21%    | 5%     | 3%   | 9% | 12% |

*reduced UV cross-link
Table S2. Reproducibility of cross-link sites (cDNA counts) within iCLIP experiments, Related to Figures 1 and 5.

| cDNA counts in ±5 nt clusters | iCLIP1 | iCLIP2 | iCLIP3 | AubAA iCLIP |
|------------------------------|--------|--------|--------|-------------|
| ≥1                           | 6777   | 4315   | 8470   | 3379        |
| ≥3                           | 3111   | 1955   | 5408   | 467         |
| ≥5                           | 1778   | 1473   | 3744   | 97          |
| ≥10                          | 509    | 683    | 1586   | 4           |
| ≥15                          | 135    | 343    | 894    | 1           |
| ≥20                          | 50     | 170    | 609    | 1           |

Table S3. Reproducibility of cross-link sites between Aub iCLIP experiments, Related to Figure 1.

| cDNA counts in ±5 nt clusters | # of cross-links reproduced in ≥2 Aub iCLIP experiments |
|------------------------------|--------------------------------------------------------|
| ≥1                           | 2799                                                   |
| ≥3                           | 1214                                                   |
| ≥5                           | 634                                                    |
| ≥10                          | 215                                                    |

Table S4. Genes with cDNA counts ≥5 in ±5 nt clusters, reproduced in ≥2 Aub iCLIP experiments, Related to Figures 1 and 2. (see Table S4 Excel file).
Table S5. Analysis of reads in small RNA libraries, Related to Figures 4 and 6.

|                         | Smg IP   | wt embryos | twin^{12} embryos | wt embryos |
|-------------------------|----------|------------|-------------------|------------|
| Total number of sequenced reads | 55 478 974 | 11 328 536 | 46 226 523        | 42 337 509 |
| Processed reads (15-31nt) | 37 036 746 | 9 819 009  | 43 977 021        | 32 936 925 |
| Mapping to release 5     | 24 085 904 | 4 820 225  | 23 598 740        | 16 796 345 |
| % of mapped reads        | 65%      | 49%        | 54%               | 51%        |
| rRNA                    | 9 638 091 | 85 627     | 6 758 954         | 3 498 906 |
| tRNA                    | 470 342   | 85 684     | 812 408           | 440 638    |
| snRNA                   | 90 956    | 1 104      | 18 848            | 5 718      |
| snoRNA                  | 244 456   | 423        | 62 435            | 19 178     |
| miRNAs                  | 3 183 702 | 846 211    | 3 589 899         | 2 556 051  |
| piRNAs (23-29nt)        | 4 584 974 | 3 044 521  | 5 841 324         | 4 450 951  |
| others                  | 5 873 383 | 756 655    | 6 514 872         | 5 824 903  |
Table S6. Genes with cDNA counts ≥5 in ±5 nt clusters in Aub\textsuperscript{AA} iCLIP, reproduced in Aub iCLIPs, Related to Figure 5.

| Gene Symbol | FBgn          | Number of reproduced clusters in ≥2 Aub iCLIP replicates | Number of reproduced clusters in ≥2 Aub iCLIP replicates and in Aub\textsuperscript{AA} iCLIP |
|-------------|---------------|----------------------------------------------------------|-----------------------------------------------------------------------------------|
| aub         | FBgn0000146   | 36                                                       | 3                                                                                |
| CG13654     | FBgn0039290   | 1                                                       | 1                                                                                |
| CG1909      | FBgn0039911   | 1                                                       | 1                                                                                |
| CG32791     | FBgn0052791   | 3                                                       | 2                                                                                |
| CycB        | FBgn0000405   | 36                                                      | 4                                                                                |
| Ef1alpha48D | FBgn0000556   | 34                                                      | 4                                                                                |
| Rpl10       | FBgn0024733   | 4                                                       | 1                                                                                |
| Rpl41       | FBgn0066084   | 4                                                       | 1                                                                                |
| RpS2        | FBgn0004867   | 8                                                       | 1                                                                                |
| RpS3A       | FBgn0017545   | 8                                                       | 1                                                                                |
Table S7. mRNAs showing a ping-pong signature within ±60 nt of a cross-link site, Related to Figure 6.

| Gene Symbol | FBgn     | Gene Symbol | FBgn     |
|-------------|----------|-------------|----------|
| Act5C       | FBgn0000042 | MED26       | FBgn0039923 |
| alphaTub84B | FBgn0003884 | mod         | FBgn0002780 |
| ari-1       | FBgn0017418 | Nacalpha    | FBgn0086904 |
| awd         | FBgn0000150 | Not1        | FBgn0085436 |
| BicC        | FBgn0000182 | pnut        | FBgn0013726 |
| bif         | FBgn0014133 | porin       | FBgn0004363 |
| CG30497     | FBgn00050497 | RpL15     | FBgn0028697 |
| CG3812      | FBgn0030421 | RpL18       | FBgn0035753 |
| CG45084     | FBgn0266459 | RpL18A      | FBgn0010409 |
| CG45093     | FBgn0266526 | RpL29       | FBgn0016726 |
| CG8765      | FBgn0036900 | RpL3        | FBgn0020910 |
| CG9821      | FBgn0037636 | RpL31       | FBgn0025286 |
| Chro        | FBgn0044324 | RpL36A      | FBgn0031980 |
| ctp         | FBgn0011760 | RpL40       | FBgn0003941 |
| CycA        | FBgn0000404 | RpL5        | FBgn0064225 |
| CycB        | FBgn0000405 | RpL7        | FBgn0005593 |
| Cyp1        | FBgn0004432 | RpLP0       | FBgn0000100 |
| dom         | FBgn0020306 | RpS15       | FBgn0034138 |
| Ef1alpha48D | FBgn0000556 | RpS16       | FBgn0034743 |
| Ef1gamma    | FBgn0029176 | RpS27A     | FBgn0003942 |
| EF2         | FBgn0000559 | RpS28b      | FBgn0030136 |
| eIF-2beta   | FBgn0004926 | RpS30       | FBgn0038834 |
| eIF-5A      | FBgn0034967 | RpS3A       | FBgn0017545 |
| His2B:CG33868 | FBgn0053868 | sesB        | FBgn0003360 |
| His3.3B     | FBgn0004828 | spir        | FBgn0003475 |
| His4r       | FBgn0013981 | sta         | FBgn0003517 |
| Hsc70-4     | FBgn0266599 | Top1        | FBgn0004924 |
| lola        | FBgn0005630 | wisp        | FBgn0260780 |
**Supplemental Experimental Procedures**

**Drosophila** stocks and genetics

The w^118^ stock was used as a control. Mutant stocks were aub^{H2O} cn^1 bw'/CyO, aub^{QC2} cn^1 bw'/CyO (Schupbach and Wieschaus, 1991), mnk^{P6} aub^{H2O} CyO, mnk^{P6} aub^{QC2} CyO (Klitzenhoff et al., 2007), spn^{E'/TM3}, r^{506} spn^{E'-03987} e/TM3 (Gillespie and Berg, 1995), mnk^{P6}, spn^{E'/TM3}, r^{506} spn^{E'-03987} e/TM3, dj-GFP/CyO (Santel et al., 1997) and twin{12209} (Zaessinger et al., 2006). The stock osk-bcd3'UTR (Ephrussi and Lehmann, 1992) was used for localization of osk at the anterior pole. The act5C-cas9 stock was used for CRISPR-Cas9 genome editing (Port et al., 2014). GFP-Aub and GFP-Aub^{AA} were expressed following crosses between the germine driver nos-Gal4:VP16 (Rorth, 1998) and UASp-GFP-Aub (Harris and Macdonald, 2001) or UASp-GFP-Aub^{AA}. For GFP-Aub expression in osk^{54} mutant embryos, the stocks used were w; osk^{54} nos-Gal4:VP16/TM3 Sb and wv; osk^{54} e UASp-GFP-Aub/TM3 Sb. osk^{54} is a null allele bearing a nonsense mutation (Kim-Ha et al., 1991). The pUASp-GFP-Aub^{AA} construct was generated as follows. PCR mutagenesis was performed using the GFP-Aub transgene in which mGFP6 is fused in frame to the Aub start codon (Harris and Macdonald, 2001) and the two primer pairs 5'TTCCGGATCAGTTAGGGCC/5'GTATCCTCCTTAFCGCGCATCCAGTACG, and 5'CGTACGTGGATCGCGCTAAGAACGGATAC/5'TTTGGAGATCTTCATGACCC to amplify two PCR fragments. These fragments were annealed and re-amplified to produce a long Aub^{AA} mutant PCR fragment, which served to replace the corresponding fragment in the GFP-Aub plasmid using Sphl and BglII. The complete GFP-Aub^{AA} mutant sequence was then inserted into the pUASp vector (Rorth, 1998) using XbaI and KpnI. The Aub mutant sequence was verified by sequencing. Transgenic stocks were produced by BestGene.

**Small RNA libraries and iCLIP**

Small RNA libraries were prepared with 0-2 h wild-type embryos, embryos from twin^{12209} mutant females crossed with wild-type males, or following Smg immunoprecipitation. RNA was prepared using Trizol (Invitrogen). Small RNA libraries were prepared and sequenced by Fastersis (Switzerland), following 18-30 nt size selection on gel and using Illumina HiSeq 2000. The iCLIP libraries were prepared as published earlier (Konig et al., 2010, 2011) with the following modifications. Freshly collected 0-2 h UASp-GFP-Aub/nos-Gal4 or UASp-GFP-Aub^{AA}; nos-Gal4/+ embryos were dechorionated for two minutes in 50% bleach and resuspended in ice cold 0.1% Tween-20 in a 10 cm dish on ice. The UV cross-link was performed with six minute-pulses at 254 nm in a BIO-LINK (BLX with 5 8-watt bulbs). The embryos were stirred between each pulse and subsequently washed in PBS, pelleted and frozen at -80°C. Each immunoprecipitation consisted of 100 µl of embryos, 120 µl of protein A Dynabeads (100-02D, Invitrogen) and 6 µl of rabbit polyclonal anti-GFP (A-6455, Invitrogen). After embryo lysis, partial RNA digestion with RNase I (Ambion) and immunoprecipitation, the RNAs were ligated at their 3' ends to a preadénylated RNA adapter (miRNA cloning linker2, IDT) and radioactively labelled to allow visualization. The samples were analysed by NuPage gel electrophoresis (WG1401BOX, Invitrogen) and RNA-protein complexes were transferred to nitrocellulose. After cutting out the region of nitrocellulose containing RNA-Aub complexes, RNAs were removed from the membrane by Proteinase K digestion. The RNAs were reverse transcribed using oligonucleotides containing two adapter regions separated by a restriction site, and a barcode region containing a 3 nt experiment-specific barcode and a 4 nt random barcode to mark individual cDNA molecules. cDNA size selection, circularization and linearization were performed as previously described (Konig et al., 2010, 2011). Linearized cDNAs were then PCR-amplified using primers complementary to the adapter regions. The size-specific libraries from one experiment were mixed to obtain a 10 nM library. iCLIP libraries were sequenced by the Montpellier Genomix facility. Western blotting was performed according to standard protocols. Antibody dilutions for western blots were mouse anti-Aub (4D10, a gift from M. Siomi, 1/2500) and rabbit anti-GFP (A-6455 Invitrogen, 1/1000). Small RNA libraries and iCLIP sequencing data have been deposited to the NCBI GEO under accession number: GSE70778.

**Microarrays**

RNA was prepared with Trizol (Invitrogen) from 0-2 h and 2-4 h-embryos from wild-type, aub^{H2O}/aub^{QC2} or spn^{E'-03987}/spn^{E'} females crossed with wild-type males. Three biological replicates were prepared per condition. Each sample was compared to a common reference made of an equimolar mix of all eighteen samples. A technical replicate of each sample led to thirty-six hybridizations. For each sample, 1 µg of total RNA was amplified and labelled using the Amino Allyl Message Amp II aRNA Amplification Kit (Ambion), according to the manufacturer's instructions. The Cy3- and Cy5-labeled aRNA targets were mixed with Hybridization Buffer, hybridization component A and alignment oligonucleotides (Roche Nimblegen), denatured at 95° C for 3 min and applied to a 12x135K Nimblegen HG18_100718 microarray slide. Hybridization was performed at 42°C for 16 h in a Hybridization System 4 (Roche Nimblegen). Hybridized slides were washed according to the Nimblegen protocols. Microarrays were immediately scanned in Cy3.
and Cy5 channels at 1 μm resolution using Innoscan900 scanner (Innopsys) and variable photomultiplier tube (PMT) settings to obtain maximal signal intensities. NimbleScan v2.5 software (Roche Nimblegen) was used for feature extraction. Data were stored and visualized using BASE data management software (Saal et al., 2002). R packages from the BioConductor project (Gentleman et al., 2004) were used (R v.2.15.0, BioConductor v2.10) to perform data analysis. The data were normalized using the VSN (Variance Stabilizing Normalization) method (Huber et al., 2002) as implemented in the LIMMA (Linear Modeling of Microarray data) package. Differentially expressed probes were identified using LIMMA (Smyth, 2005) and FDR (False Discovery Rate) set to 0.1%. Microarray data set have been deposited to the NCBI GEO under accession number: GSE70778.

Immunoprecipitations and RNA analyses
Smg and GFP immunoprecipitations were performed as described previously (Rouget et al., 2010; Zaessinger et al., 2006) using 0-2 h-embryos and Guinea pig anti-Smg (a gift from C. Smibert) or rabbit anti-GFP (A-6455 Invitrogen), followed by RNA preparation with Trizol (Invitrogen) or phenol-chloroform for small RNAs radiolabeled with [γ-32P] ATP. Labelling was performed using the KinaseMax™ 5′ End-Labeling Kit followed by purification with the illustra MicroSpin G-25 Columns (GE Healthcare). Small RNAs were then separated by electrophoresis on a 15% acrylamide-urea gel. RT-qPCR with the LightCycler System (Roche Molecular Biochemical) were performed as previously described (Rouget et al., 2010) using the primers listed below, and were made from two to six independent RNA preparations.

Immunostaining, RNA in situ hybridization and cuticle preparations
Immunostaining was performed as previously reported (Rouget et al., 2010). Antibody dilutions for immunostaining were mouse anti-GFP (Roche IgG1k clones 7.1 and 13.1, 1/200), rabbit anti-Osk (a gift from P. Lasko, 1/1000) and rat anti-Vasa (Developmental Studies Hybridoma Bank, 1/200). Whole-mount in situ hybridization and cuticle preparation were performed by standard methods. Probes for in situ hybridization experiments were DIG-labelled antisense RNA transcribed in vitro from coding regions of the corresponding genes cloned into the Topo TA pcR™ II vector (Invitrogen).

mRNA localization using Fly-FISH and BDGP insitu
Fly-FISH (http://fly-fish.ccb.r.utoronto.ca) and BDGP insitu (http://insitu.fruitfly.org/cgi-bin/ex/insitu.pl) were used to analyse the localization of mRNAs bound and regulated by Aub in embryos. "Posterior localization" includes the terms: Pole plasm stage1-3; Posterior localization stage1-3; Pole buds stage1-3; RNA islands stage1-3; Perinuclear around pole cell nuclei stage1-3; All pole cell nuclei stage1-3; Pole cell nuclei stage1-3; Pole cell localization stage4-5; Pole cell enrichment stage4-5; Posterior localization stage4-5; Perinuclear around pole cell nuclei stage4-5; Pole cell nuclei stage4-5; Subset pole cell nuclei stage4-5; All pole cell nuclei stage4-5 in Fly-FISH, and Pole cells; Pole plasm; Posterior in BDGP insitu.

Calculation of SRE scores
SRE scores were determined as reported previously (Chen et al., 2014). Briefly, mRNAs were folded using RNAPifold, and instances where CNGG formed a loop at the apex of base-paired stem were identified. The probability of formation of an actual SRE at each candidate site was calculated. For each mRNA, all SRE probabilities were added up into an SRE score. SRE scores were determined for all potential transcripts of each gene.

Bioinformatic analyses of iCLIP and small RNA seq libraries
rRNA sequences (M21017.1) were downloaded from NCBI, tRNA, snoRNA and snRNA sequences were from modEncode (http://www.modencode.org/) (Celniker et al., 2009), miRNA sequences were from miRBase (http://www.mirbase.org/) (Kozomara and Griffiths-Jones, 2011), intergenic, gene coordinates and protein coding gene sequences were from Flybase (http://flybase.org/), and TE sequences were from Repbase (http://www.girinst.org/repbase/index.html) (Jurka et al., 2005).

iCLIP libraries. Sequence reads included a 3 nt barcode as experiment identifier and a 4 nt random barcode introduced during cDNA synthesis. Before mapping to the Drosophila melanogaster genome, adapter sequences were removed from both ends of the reads. Mapping of reads was performed using Bowtie, version 0.12.9 (Langmead et al., 2009). Reads were first mapped allowing unique hits and up to 2 mismatches (Bowtie parameters -v 2 -m 1 -a --best -strata), then, the remaining reads were mapped allowing multiple hits and no mismatches (Bowtie parameters –v 0 -a). PCR amplification was eliminated by merging into one, sequences that shared the same start and stop positions on the same strand, and the same random barcode. Cross-link sites were identified as the first nt upstream of the iCLIP reads. The number of sequences at each cross-link nt was added up into a "cDNA count". Reads were annotated in the following order to tRNA, rRNA, snoRNA, snRNA, miRNAs, and TEs (with no mismatch for TE reads.
from positive to negative around this point. When the number of cross-locations, we first identified cross-link positions that occurred within ±5 nt of each other. The cDNA count of each cluster was obtained by adding cDNA counts of each cross-link site in the cluster. Coordinates of reproduced clusters in ≥2 biological replicates were the coordinates of reproduced regions in 2 or 3 biological replicates. The reproducibility of cross-link positions (Figure 1B) and the offset to reproduced positions in biological replicates (Figure S1B) were calculated as previously described (Konig et al., 2010).

To analyse the distribution of reproduced cross-links in genes, six hierarchical categories were constructed, as follows: ncRNAs > 3'UTRs > 5'UTRs > CDS > introns > intergenic regions. Genomic regions were identified using the “subtractBED” function from BEDTools. Cross-links were assigned to a category using the “intersectBED” function from BEDTools with -wa -wb -s options. Cross-link enrichment in a given category was calculated by counting the number of cross-links in this category and dividing this number by the total number of nt in the Drosophila melanogaster genome for this category. Transcripts with reproduced cross-links were identified by mapping to transcripts, reads falling into reproduced clusters from the corresponding genes. This identified transcripts from 616 genes (out of 634 genes showing reproduced cross-links) because reads overlapping introns were not considered. Transcripts with reproduced cross-links were screened for TE insertions using Repeat Masker version 3.3.0 (http://www.repeatmasker.org) with option -div 30.

Small RNA seq libraries. Reads were stripped of the 3' linker and the resulting small RNA sequences were mapped to the Drosophila melanogaster genome (release 5.54, excluding the Uextra portion) using Bowtie, with up to one mismatch. Reads were annotated in the following order to tRNA, rRNA, snoRNA, snRNA and miRNAs. piRNAs were the remaining reads that were 23-29 nt in length. piRNAs were mapped to TE using Bowtie with up to 3 mismatches. Uniquely mapped piRNAs were mapped to piRNA clusters using cluster coordinates from (Brennecke et al., 2007). mRNA-derived piRNAs were uniquely mapped piRNAs that mapped in sense orientation to transcripts or 3'UTR sequences using Bowtie and allowing up to 1 mismatch. Small RNA counts were normalized to 1 Million mapped reads.

Prediction of piRNA target sites on cellular mRNAs. We used a pool of piRNAs from 0-2 h-embryos sequenced in previously published libraries (Brennecke et al., 2008) (GSM327625, GSM327626, GSM327627, GSM327628 and GSM327629), and in this study (two libraries from wild-type embryos, Table S5). This led to a total of 3,305,903 non-redundant piRNA sequences. For groups of piRNAs that had overlapping sequences with different 3' ends, the longest piRNA was selected. This resulted in 2,130,534 piRNA sequences that were used in the analysis. Removing related piRNA sequences was required to avoid producing more complex piRNA sets when generating control shuffled piRNAs (see below). Bowtie was used with different complementarities to identify piRNA target sites on transcripts with reproduced cross-links, as follows. Bowtie with option `-v 0`, `-v1`, `-v 2` or `-v3` was used to identify piRNAs that potentially target mRNAs with up to 0, 1, 2 or 3 mismatch(es), respectively. For complementarities with a seed, we did not use quality values, therefore the sum of the quality values at all mismatched read positions (-e/-maqerr) was set to an arbitrary value of 2000, which disabled the quality values. Furthermore, -l (length of the seed) and -n (number of mismatches within the seed) were set to different values. The option `-nowf` was used to search only for reverse-complementarity between piRNAs and mRNAs. Negative control piRNA sequences were generated by shuffling the original 2,130,534 piRNA sequences following two criteria, i) the same nt at position 1, and ii) similar mono- and di-nucleotide counts. The targeting with shuffled piRNAs was performed 100 times with 100 different shuffled libraries. Shuffled sequences that had the exact same sequence as true piRNAs were removed from the analysis. An average value was calculated using 100 randomizations and used as a negative control of the experiment. The Fisher's exact test (in the R package, http://cran.r-project.org/) was used to evaluate statistical significance.

Calculation of the distance between cross-links and targeting piRNAs. We used transcripts with reproduced cross-link sites. cDNA counts at each nt position from the 3 iCLIP replicates were added. For each cluster, we first identified cross-link peaks by cubic spline interpolation using Scipy (http://www.scipy.org). Peak location was determined using the derivative of the function at each point of the interpolation, locating the point where the derivative =0, and confirming that the derivative changes from positive to negative around this point. When the number of cross-link sites in a cluster was too low for
interpolation, the cross-link site with the highest cDNA count was selected. We then calculated the distances between the peaks or selected cross-link sites and the 5’ end of targeting piRNAs within ±60 nt of these sites. mRNAs with a TE insertion within ±60 nt of the cross-link sites were not included in the analysis.

**Calculation of genic ping.** Genic ping was calculated by a custom perl script using the approach described previously for TE (Brennecke et al., 2008). The genic ping was separately calculated for each *Drosophila melanogaster* transcript of interest.

**Primers used in this study**

**Primers used for CRISPR-Cas9 genome editing**

dncCRISPR-F: GAAATTAATACGACTCACTATAGgaagtagtaattcatcagTTTTTAGAGCTAGAAATAGC
  (nt in lower case letters indicate the target sequence in dnc)

sgCRISPR-R: AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAAAC

**Primers used to amplify fragments to generate probes for in situ hybridization**

Fw-insitu-osk: AATCAATTTGGCTCCACCTG
Rev-insitu-osk: GCATTCGCTTCGGATAAACT

Fw-insitu-gcl: TGGCACCTGCATAAGGTGTA
Rev-insitu-gcl: AAAATGCGAGTTTTCAC

Fw-insitu-pgc: ACCGAAATGTTGCAGCTAC
Rev-insitu-pgc: AACAATGCCAGTTTCAAGC

Fw-insitu-Hsp83: TCGGTGTGGGTCTTACTCC
Rev-insitu-Hsp83: TATGTTGTTGCTGCTTTC

**Primers used in RT-qPCR**

osk-fw-qPCR1: TGACCATCATCGAGAGCAAC
osk-rev-qPCR1: ATTGCCGCTCAGTTTTGC

gcl-fw-qPCR: CATTCACAACCACTGGGATG

gcl-rev-qPCR: CATTCACAACCACTGGGATG

pgc-fw-qPCR: CACTCCGGCATTACCAAATC
pgc-rev-qPCR: CGCACATTTTCGGGTCTTC

Hsp83-fw-qPCR: CAACAAGCAGCGTCTG
Hsp83-rev-qPCR: AGCCTGGAATGCAAAGGTC

aret-fw-qPCR: GCTGCCGCTAAATTGAAAGT
aret-rev-qPCR: GACTATTCGGCAGGCTGTTC

cycB-fw-qPCR: AGCCTCAAACAGCACGGTAC

cycB-rev-qPCR: TTCCGACCAAGACCACTGTAG

cycA-fw-qPCR: GACATGGGGGACTGAACATC

cycA-rev-qPCR: AACAACGATAACCACTGGTAC

smg-fw-qPCR: GAATCACGCCAGTCAACTCA
smg-rev-qPCR: GCAGGCAATTTAGCGAAAAG

adam-fw-qPCR: CGAGACCAAGGAGGAGTTCA
adam-rev-qPCR: TTTCGAGTGCAAGATTTCC

mRpL43-fw-qPCR: CCGGATTGTGCTCAGAGTTGA
mRpL43-rev-qPCR: GAGCTGCGTTTTCAGCTTGT
Primers used to generate iCLIP libraries

clipPCR(5P): AATGATACGGCGACCACGAGATCTCACTACTCTTTCTTACTACACGACGGCATTCTCTTCCGATCTTCTC
clipPCR(3P): CAAGCAGAAGACGCGATACGAGATCGGTCTCGGCATCTTGCTCTTGTGCCCGAGTGG
clipAnnealOligo: CAAGGAAGATCCACGCGCTCTTCCaaa
clipRT: NNNNxxxxAGATCGGAAGAGCGCTGAGGatCTCCTTGTGC
(N= random barcode; x= experiment specific barcode)

Supplemental references

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089-1103.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322, 1387-1392.

Celinker, S.E., Dillon, L.A., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M., et al. (2009). Unlocking the secrets of the genome. Nature 459, 927-930.

Chen, L., Dumelie, J.G., Li, X., Cheng, M.H., Yang, Z., Laver, J.D., Siddiqui, N.U., Westwood, J.T., Morris, Q., Lipshitz, H.D., et al. (2014). Global regulation of mRNA translation and stability in the early Drosophila embryo by the Smaug RNA-binding protein. Genome Biol 15, R4.

Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. Nature 358, 387-392.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5, R80.

Gillespie, D.E., and Berg, C.A. (1995). Homeless is required for RNA localization in Drosophila oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases. Genes Dev 9, 2495-2508.

Harris, A.N., and MacDonald, P.M. (2001). Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to elf52C. Development 128, 2823-2832.

Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18 Suppl 1, S96-104.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenetic and genome research 110, 462-467.

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.

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.

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nature structural & molecular biology 17, 909-915.

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2011). iCLIP--transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. J Vis Exp.
Rorth, P. (1998). Gal4 in the Drosophila female germline. Mech Dev 78, 113-118.
Rouget, C., Papin, C., Boureux, A., Meunier, A.C., Franco, B., Robine, N., Lai, E.C., Pelisson, A., and Simonelig, M. (2010). Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. Nature 467, 1128-1132.
Saal, L.H., Troein, C., Vallon-Christersson, J., Gruvberger, S., Borg, A., and Peterson, C. (2002). BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. Genome Biol 3, SOFTWARE0003.
Santel, A., Winhauer, T., Blumer, N., and Renkawitz-Pohl, R. (1997). The Drosophila don juan (dj) gene encodes a novel sperm specific protein component characterized by an unusual domain of a repetitive amino acid motif. Mech Dev 64, 19-30.
Schupbach, T., and Wieschaus, E. (1991). Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology. Genetics 129, 1119-1136.
Smyth, G. (2005). Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions using R and Bioconductor, R.C. Gentleman, V.J. Carey, S. Dudoit, R. Irizarry, and W. Huber, eds. (Springer, New York), pp. 397-420.
Zaessinger, S., Busseau, I., and Simonelig, M. (2006). Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. Development 133, 4573-4583.