



# Piwi/piRNAs control food intake by promoting neuropeptide F expression in locusts

Huimin Wang<sup>1,2,†</sup>, Feng Jiang<sup>1,2,†</sup> , Xiang Liu<sup>3,†</sup>, Qing Liu<sup>3,4</sup>, Yunyun Fu<sup>5</sup>, Ran Li<sup>1</sup>, Li Hou<sup>3</sup>, Jie Zhang<sup>1</sup>, Jing He<sup>3</sup> & Le Kang<sup>1,2,3,5,\*</sup> 

## Abstract

Animal feeding, which directly affects growth and metabolism, is an important physiological process. However, the contribution of PIWI proteins and PIWI-interacting RNAs (piRNAs) to the regulatory mechanism of animal feeding is unknown. Here, we report a novel function of Piwi and piRNAs in regulating food intake in locusts. Our study shows that the locust can serve as a representative species for determining PIWI function in insects. Knockdown of *Piwi1* expression suppresses anabolic processes and reduces food consumption and body weight. The reduction in food intake by knockdown of *Piwi1* expression results from decreased expression of neuropeptide *NPF1* in a piRNA-dependent manner. Mechanistically, intronic piRNAs might enhance RNA splicing of *NPF1* by preventing hairpin formation at the branch point sites. These results suggest a novel nuclear PIWI/piRNA-mediated mechanism that controls food intake in the locust nervous system.

**Keywords** feeding; neuropeptide; piRNA; PIWI protein

**Subject Categories** Metabolism; RNA Biology

**DOI** 10.15252/embr.202050851 | Received 10 May 2020 | Revised 8 December 2021 | Accepted 16 December 2021 | Published online 5 January 2022

**EMBO Reports (2022) 23: e50851**

## Introduction

Feeding is a ubiquitous behavior and a complicated physiological process among animals (Petrovich, 2018). Food intake, which is under the control of the central feeding system, is regulated by neuropeptides and neurotransmitters (Wei *et al.*, 2000; Falibene *et al.*, 2012; Van Wielendaele *et al.*, 2013; Tan *et al.*, 2019). Several studies suggest that genes associated with the small RNA pathway participate in regulating feeding. The loss of *Dicer* in mice leads to an imbalance in neuropeptide expression and severe hyperphagic obesity (Vinnikov *et al.*, 2014; Fiorenza *et al.*, 2016). In the diet-induced mice obesity model, inhibition of miR-128-1 expression led

to an increase in food intake of high-fat diet (HFD)-fed mice (Wang *et al.*, 2020). A cluster of six miRNAs, that is, miR-959-964, regulates feeding and metabolism in *Drosophila* (Vodala *et al.*, 2012). A total of 37 PIWI-interacting RNAs (piRNAs) are differentially expressed in sperm between lean and obese men, implying that altered piRNA expression coordinately modulates the expression of genes involved in food intake (Donkin *et al.*, 2016). However, the molecular function of piRNAs, another important class of small RNAs, in regulating feeding behavior has not yet been proven in animals.

In general, feeding behavior of animals is controlled by the central nervous system. The functions of PIWI proteins and piRNAs in the somatic nervous system have not been well investigated, even though their functions in the germline and its associated somatic cells are widely recognized (Ross *et al.*, 2014). PIWI proteins are known to enlist germline-specific piRNAs to suppress transposons. In the germline, the loss of or a decline in the expression of Piwi genes results in genomic instability and cellular dysfunction (Carmell *et al.*, 2007; Chen *et al.*, 2007). PIWI proteins are important for neurogenesis and nervous system function; for example, they control synaptic plasticity and play potential roles in neurodevelopmental and neurodegenerative diseases in nematodes, flies, mice, and humans (Rajasethupathy *et al.*, 2012; Kim *et al.*, 2018; Phay *et al.*, 2018; Sun *et al.*, 2018). Knockout of *Mili* in mouse hippocampal neurons leads to hyperactivity and reduced anxiety (Nandi *et al.*, 2016). Elevated expression of piRNA in the *Aplysia* central nervous system can enhance serotonin-dependent memory storage (Rajasethupathy *et al.*, 2012). However, few studies have focused on exploring the molecular mechanism underlying the regulatory role of the PIWI/piRNA pathway in the nervous system.

The migratory locust, *Locusta migratoria*, is the most widespread and dangerous locust species showing phenotypic plasticity (Jiang *et al.*, 2019a; Yang *et al.*, 2019a, 2019b, Liu *et al.*, 2021). It is estimated that a single swarm can consume the plant biomass equivalent of food consumed daily by 35,000 persons. The migratory locust is a model system for studying behavior, feeding, learning and memory, courting, mating, aggregation, and jumping (Van Wielendaele *et al.*, 2013; Vaudry *et al.*, 2013; Hou *et al.*, 2017, 2019;

1 Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing, China

2 CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing, China

3 State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

4 Sino-Danish College, University of Chinese Academy of Sciences, Beijing, China

5 College of Life Science, Hebei University, Baoding, China

\*Corresponding author. Tel: +86 10 6480 7219; Fax: +86 10 6480 7099; E-mail: lkang@ioz.ac.cn

†These authors contributed equally to this work

Zhou *et al*, 2019). Our previous studies revealed that small RNAs can regulate the expression of protein-coding genes to affect behavioral changes in locusts. For example, microRNA-133 can modulate behavioral phase changes by controlling dopamine synthesis (Yang *et al*, 2014), and microRNA-9a mediates locust attraction behavior by activating and inhibiting dopamine receptor 1 (Guo *et al*, 2018). Potential piRNAs, a class of small RNAs with length of 25–29 nt in locusts, were predicted *in silico* (Wei *et al*, 2009). Abundant protein-coding genes are differentially expressed after reducing *Piwi1* expression (Jiang *et al*, 2019b). Therefore, the PIWI/piRNA pathway may be involved in regulating the expression of protein-coding genes involved in locust behaviors.

In this study, we report the presence of somatic PIWIs/piRNAs in the locust brain and functional investigations on the mechanism of PIWIs/piRNAs in food intake. PIWIs/piRNAs in the brain exhibit the canonical characteristics of authentic PIWIs/piRNAs. We found that loss of *Piwi1* results in the suppression of food consumption and anabolic processes in locusts by inhibiting the expression of neuropeptide F (*NPF1*), which is regulated by piRNAs. Mechanistically, the piRNA piRs-3-13 may enhance the RNA splicing of the precursor mRNA transcripts of *NPF1* (*pre-NPF1*) by preventing hairpin formation at the branch point site. Overall, our results propose a novel nuclear PIWI/piRNA-mediated mechanism that maintains food intake in the somatic nervous system in animals.

## Results

### Locusts possess three PIWI proteins

To investigate the orthologous relationships of PIWI genes of the migratory locust, we screened the genome assemblies of 174 arthropod species. Arthropod Argonaut 3 (*Ago3*) genes were clearly identified because all *Ago3* genes clustered with *Drosophila Ago3* to form a monophyletic branch (Fig 1A). The split between the *Piwi* and *Aub* branches in *Drosophila* shows that *Aub* originates from lineage-specific gene duplication events only in a few insect genera within Diptera. Reciprocal best BLAST searches showed that one locust PIWI protein, LOCMI12271, displays the best hit to the *Drosophila Ago3* protein (~44% identity). According to their sequence similarities, two other locust *Piwi*-like proteins, *Piwi1* (~47%, LOCMI04032) and *Piwi2* (~43%, LOCMI1364), are approximately equivalent to the *Drosophila Piwi* and *Aub* proteins, respectively.

Instead of BLAST searches, the protein distance matrix method was applied to determine the overall similarities of *Piwi*-like sequences between *Drosophila* and the locust. In contrast to the BLAST search results, *Piwi1* (0.92) and *Piwi2* (0.91) in the locust showed the shortest distances to *Aub* and *Piwi* in *Drosophila*, respectively (Fig 1B). However, the distance between the *Piwi/Aub* proteins in *Drosophila* (0.94) was substantially larger than that between the *Piwi1/Piwi2* proteins in the locust (0.71). This distance was even equivalent to the distances between *Piwi1/Piwi2* in the locust and *Piwi/Aub* in *Drosophila* (0.95 for *Piwi1-Piwi*; 0.91 for *Piwi2-Piwi*; 0.92 for *Piwi1-Aub*; and 0.97 for *Piwi2-Aub*; Fig 1B). Two *Piwi*-like genes were detected in a large proportion of the 174 arthropods, and similar gene numbers were found in *Drosophila* (Fig 1C).

The divergence rates obtained by estimating pairwise  $K_a/K_s$  ratios between the *Piwi/Aub* duplicates are relatively free from the

influence of functional overlap with other duplicate copies (Ganko *et al*, 2007). The pairwise  $K_a/K_s$  ratio between *Piwi1* and *Piwi2* in the locust ( $K_a$ , 0.378;  $K_s$ , 3.66;  $K_a/K_s$ , 0.103) showed less divergence than that between *Piwi* and *Aub* in *Drosophila* ( $K_a$ , 0.508;  $K_s$ , 2.77;  $K_a/K_s$ , 0.184). Globally, the *Piwi/Aub* duplicates in *Drosophila* species display elevated  $K_a$  and decreased  $K_s$  values (Fig 1D); thus, the  $K_a/K_s$  ratios in *Drosophila* were significantly greater than those in other non-*Drosophila* insects. Consistent with the majority of the arthropods in the study, the locust has two *Piwi*-like genes and exhibits a lower  $K_a/K_s$  value than *Drosophila*. Taken together, these results suggest that the locust can serve as a representative species for determining PIWI function in non-*Drosophila* insects.

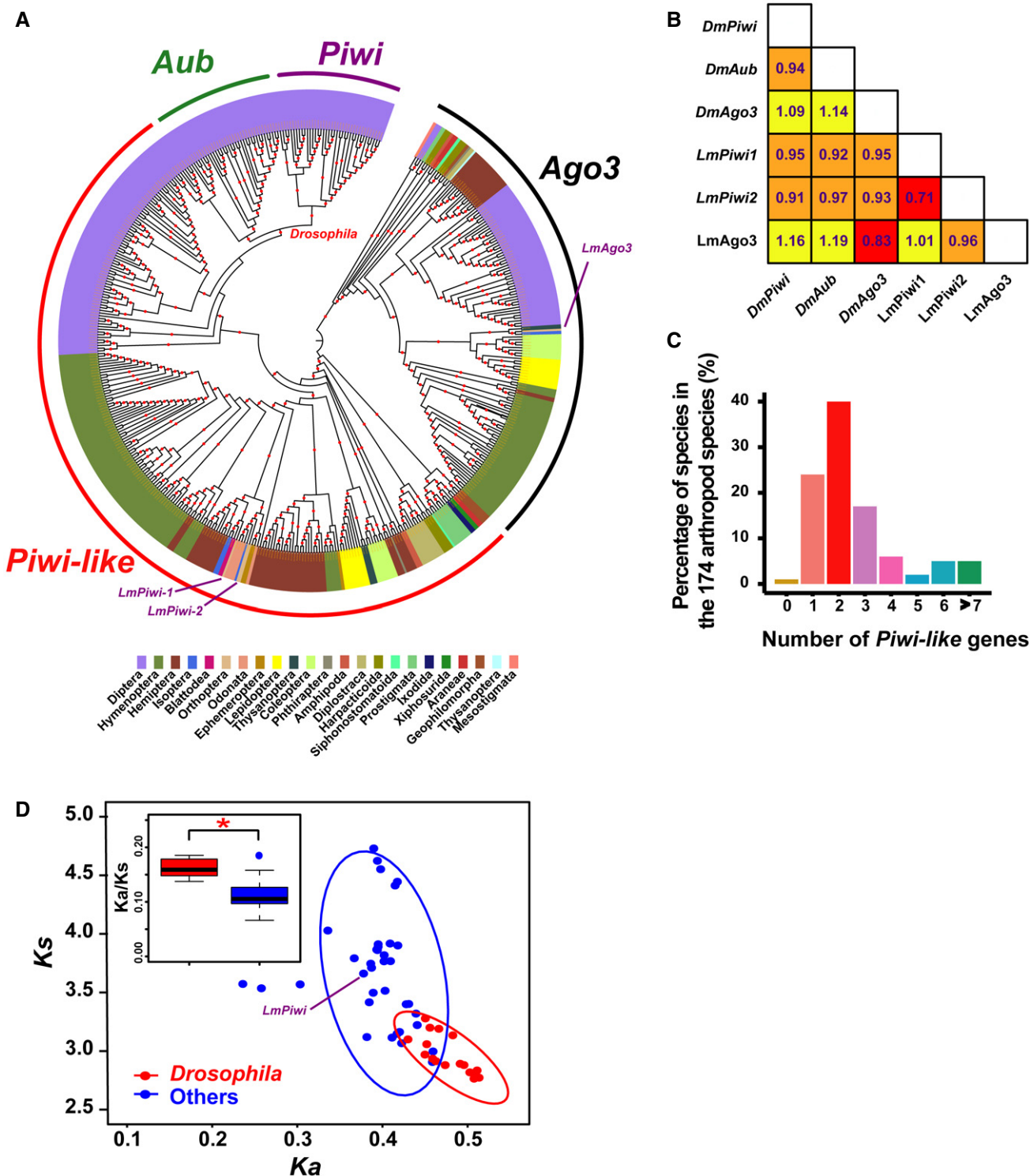
### The locust brain exhibits neuronal expression of the PIWI/piRNA pathway

We performed quantitative PCR (qPCR) analysis in relevant tissues and organs of locusts to determine the somatic expression of the three PIWI proteins. The locust brain exhibited the highest expression levels of *Piwi1* and *Ago3* among all tissues except for the ovary and testis (Fig 2A). Monoclonal antibodies against *Piwi1*, *Piwi2*, and *Ago3* were generated and verified using RNA interference (RNAi) assays to quantify the protein expression of three PIWI genes (Appendix Figs S1A–E and S2A). Consistent with the qPCR data, expression of the *Piwi1* and *Ago3* proteins was detected in the locust brain via Western blot analysis (Fig 2B). However, expression of neither *Piwi2* mRNA nor protein was detected in the locust brain by qPCR analysis or Western blot analysis (Fig 2A and B), respectively.

Immunofluorescence microscopy confirmed that *Piwi1* and *Ago3* showed predominantly nuclear and cytoplasmic localization, respectively (Fig 2C). Given that the presence of 2'-O methylation at the 3' terminus is a hallmark characteristic of piRNA biogenesis, we performed transcriptome sequencing of brain small RNAs subjected to periodate oxidation and  $\beta$ -elimination reactions. Unlike unoxidized small RNAs, oxidized small RNAs showed a shift toward increased expression in the size range larger than 22 nt, which is a typical size class of small interfering RNAs and microRNAs in locusts (Wei *et al*, 2009) (Fig 2D). We further verified the presence of three additional hallmark features that are critical for piRNA biogenesis and amplification: preferential matches with transposon elements (TEs), a 5' uracil bias and a ping-pong signature. The peak length of the oxidized small RNAs was 25–27 nt peak in the TE-derived portion (Fig 2D and E), and their 5' ends showed a strong preference for uracil (Fig 2F). The pairwise sequence comparison of the oxidized small RNAs confirmed the presence of a 10 nt overlap between their 5' ends, indicating that the piRNAs exhibited ping-pong signatures (Fig 2G). Taken together, these findings suggest that the locust brain exhibits neuronal expression of *Piwi1/Ago3* and piRNAs, implying a neuronal functional role for the PIWI/piRNA pathway in the locust brain.

### Knockdown of *Piwi1* expression results in food intake reduction and anabolic process suppression

We knocked down the *Piwi1* and *Ago3* genes by microinjecting double-stranded RNAs (*dsRNAs*) into the brains of fourth-instar locust nymphs to investigate their neuronal functions. The qPCR results showed that the mRNA expression levels of *Piwi1* and *Ago3*



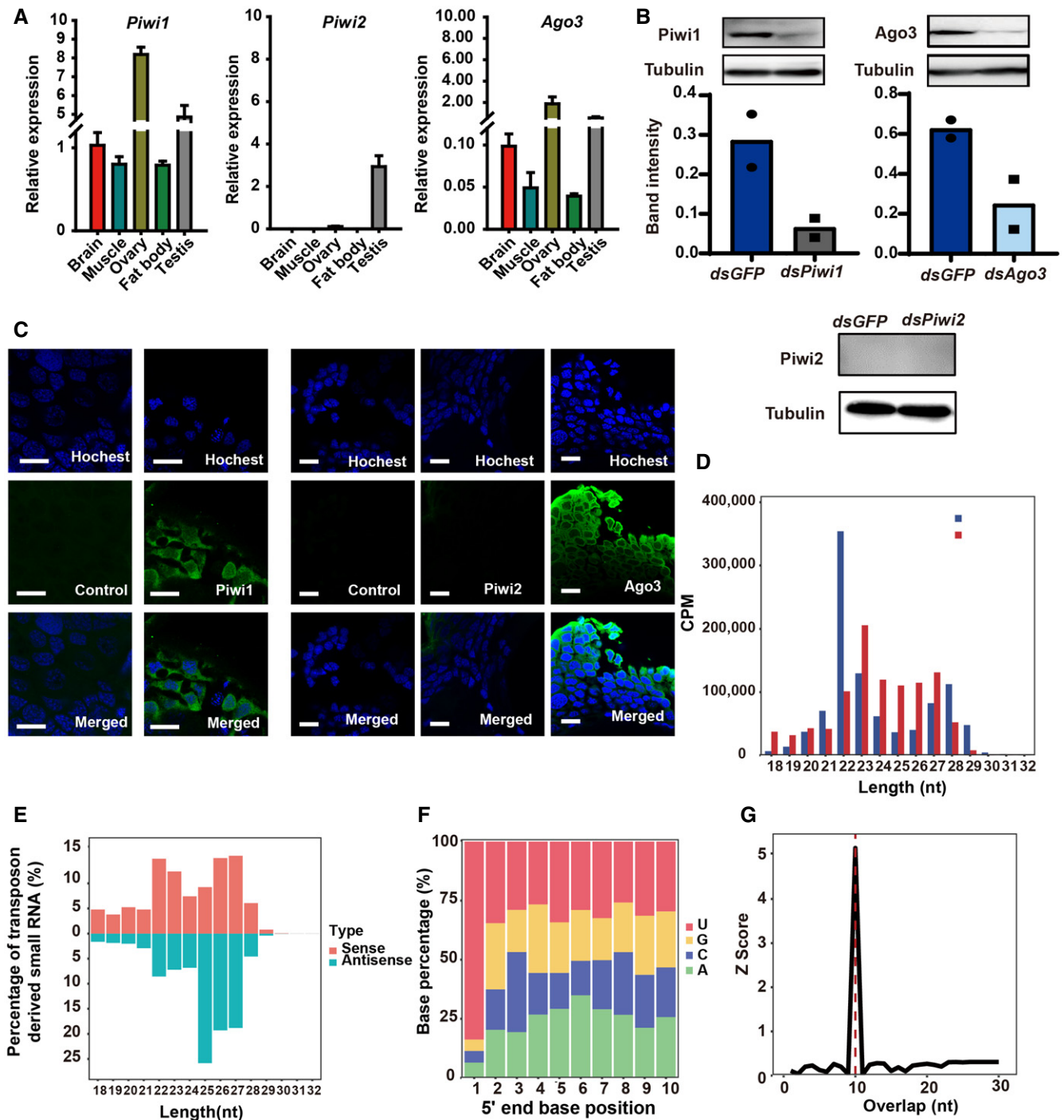
**Figure 1. Characterization of PIWI genes in arthropod species.**

**A** A phylogenetic analysis of PIWI genes in 174 arthropod genomes. A maximum-likelihood phylogenetic tree was generated on the basis of the protein sequences of PIWI genes using FastTree software version 2.1.10. The reliability of each tree branch was evaluated using the SH branch support test with 1,000 resamples. The red dots on a branch indicate genes with an SH test result > 0.5. The different colors outside of the phylogenetic tree represent different arthropod orders.

**B** Protein distance matrix for *Piwi*, *Aub*, and *Ago3* in *Drosophila* (Dm) and *Locusta migratoria* (Lm).

**C** The numbers of *Piwi*-like genes in the 174 arthropod species.

**D** Distributions of pairwise  $K_a/K_s$  (nonsynonymous-to-synonymous nucleotide substitution rates) ratios between PIWI duplicates in arthropods ( $n = 17$  in *Drosophila* and  $n = 37$  in others, Mann–Whitney *U*-test,  $*P < 0.05$ ). The center line of the boxplots represents median value, the bounds of the box represent 75<sup>th</sup> and 25<sup>th</sup> percentile, and the whiskers represent maximum and minimum value.



**Figure 2. Expression of PIWI proteins and piRNAs in the locust brain.**

A PIWI genes are present in the locust brain. Rp49 served as a loading control ( $n = 4$  biological replicates).

B Western blot analysis of PIWI proteins in the locust brain. Tubulin was used as an endogenous control ( $n = 2$  biological replicates). The bars represent the mean of band intensity.

C The fluorescence signals of PIWI proteins were determined in locust brains. Hochest33342 was used to label cell nuclei in locust brains. The scale bars represent 20  $\mu\text{m}$ .

D Size distribution of genome-matching small RNAs among total RNAs and oxidized smRNA-seq libraries. CPM, counts per million.

E Size distribution of transposon-matching sense (blue) and antisense (red) reads in oxidized smRNA-seq libraries.

F The nucleotide identity at each read position of transposon-matching piRNAs shows a first-position nucleotide bias for uracil.

G The pairwise read comparison of the oxidized small RNAs demonstrated an overlap of 10 nt between their 5' ends.

Data information: In (A) data are presented as mean  $\pm$  SEM.

were significantly decreased (Appendix Fig S2B). Consequently, compared with *dsGFP* control treatment, *dsPw1* treatment resulted in a striking decrease in locust body weight (Fig 3A). However, body weight losses were observed after *dsPw1* treatment only but not after *Ago3* treatment.

We assessed olfaction, locomotion, and food intake to determine whether *dsPw1* treatment affects feeding behavior. No significant change in olfactory choice related to food or locomotion performance was observed (Fig 3B and C). Because the exoskeleton is one of the major contributors to the body weight of insects (Smock, 1980), we measured head width, tergum length, and femur length. We found no significant differences in these traits between the *dsPw1* and *dsGFP* groups (Fig 3D–F). When locusts were deprived of food overnight for 16 h, the amount of food intake by the locusts was significantly decreased after *dsPw1* treatment for the following four successive days (Fig 3G).

Given that the fat body is the principal organ for the anabolic processing of digestive foods following absorption in insects, mRNA transcriptome sequencing was performed to analyze global expression changes in the fat bodies of locusts subjected to *dsPw1* treatment in the brain. The expression of a total of 852 and 817 genes was upregulated and downregulated, respectively, in the *dsPw1*-treated locusts compared with *dsGFP*-treated control locusts (DESeq2 method, *P*-value < 0.05, and adjusted *P*-value < 0.1). The results of gene ontology (GO) enrichment analysis for biological processes showed that the downregulated genes were enriched in metabolism-associated pathways, such as glycogen metabolism (GO: 0005978, glycogen biosynthetic process; GO: 0070873, regulation of glycogen metabolic process; and GO: 0006641, triglyceride metabolic process), fatty acid metabolism (GO: 0006633, fatty acid biosynthetic process; and GO: 0006635, fatty acid beta-oxidation), and acyl-CoA metabolism (GO: 0006084, acetyl-CoA metabolic process; Fig 4A and Appendix Figs S3A and B, and S4A). The fat content and lipid droplet size in fat bodies were dramatically decreased in the *dsPw1*-treated locusts compared with the *dsGFP*-treated control locusts (Fig 4B–D and Appendix Fig S5), suggesting a reduction in lipid accumulation and energy storage. Because glycogen and triacylglycerides (TAGs) are the major intracellular forms of stored energy in insects, we quantified their contents in fat bodies. The levels of glycogen and TAGs were significantly reduced in the *dsPw1*-treated locusts (Fig 4E and F). Therefore, *Pw1* expression knockdown results in decreased body weight and is strongly correlated with reduced food consumption and the suppression of anabolic processes.

### ***Pw1* controls food intake by regulating NPF1 signaling**

To identify the gene that acts downstream of *Pw1*, we performed mRNA transcriptome sequencing of the brains of the *dsPw1*-treated locusts and *dsGFP*-treated control locusts. We identified 349 upregulated genes and 380 downregulated genes in the *dsPw1* locusts compared with the *dsGFP*-treated control locusts (DESeq2 method, *P*-value < 0.05, and adjusted *P*-value < 0.1). GO enrichment analysis indicated that the GO terms enriched from the downregulated genes were associated with digestive tract development (GO: 0048546), intestinal epithelial structure maintenance (GO: 0060729), nymph development (GO: 0002168), and embryonic morphogenesis (GO: 0048598; Fig 5A and Appendix Fig S4B).

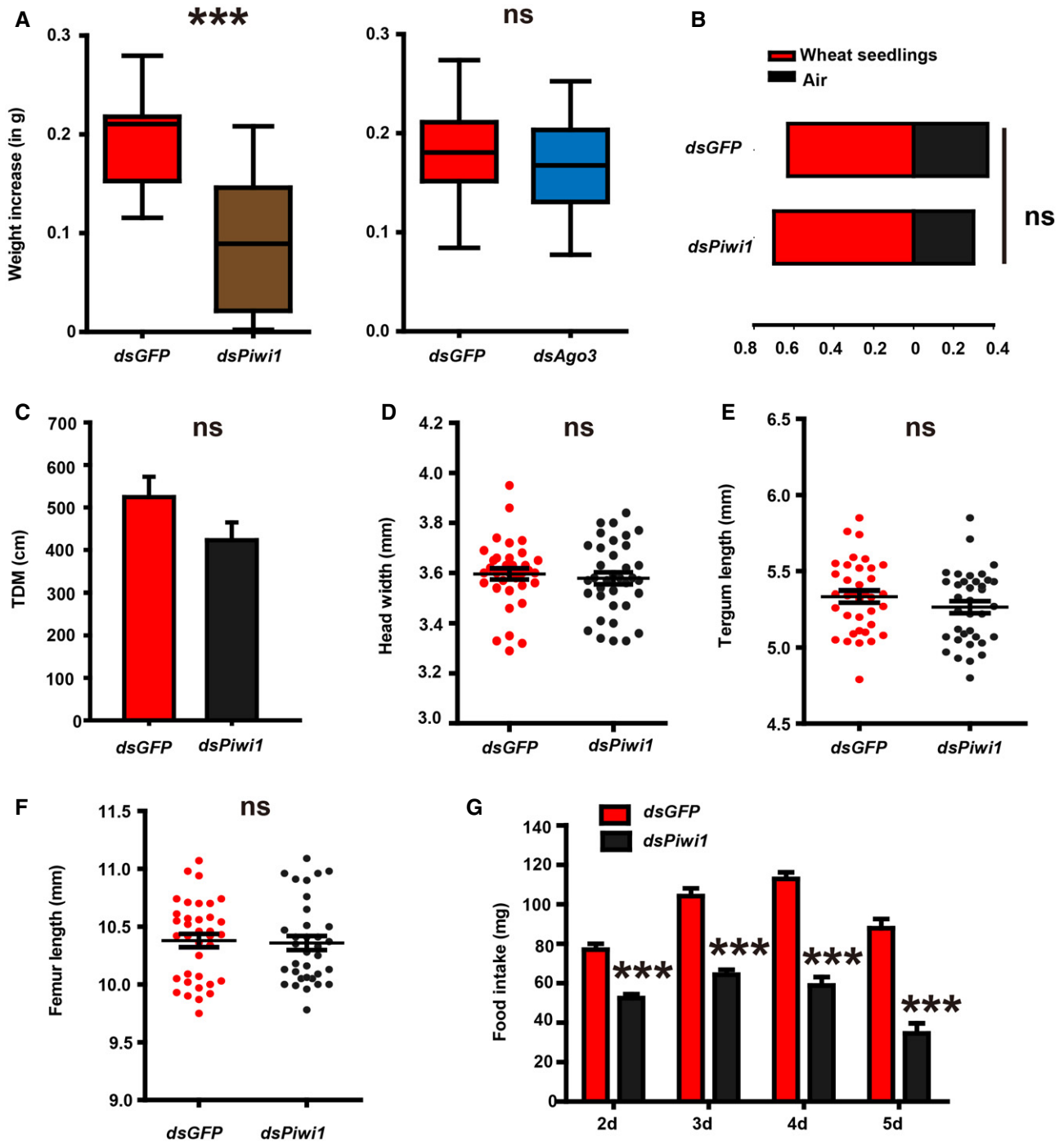
Among the downregulated genes, the top 10 most differentially expressed genes were sulfotransferase 2 (*St2*), transglutaminase (*Tg*), *Pw1*, homocysteine S-methyltransferase (*CG10621*), collagen type IV alpha 1 (*Col4a1*), N-acetyl-D-glucosamine kinase (CG6218), CG10407, major facilitator superfamily transporter 3 (*MFS3*), CG1368, and *NPF1* (Appendix Fig S6A and B and S7A–C). Given their potential association with body growth, lipid metabolism, and feeding in insects (Berger et al, 2012; Teng et al, 2012; Van Wieldaele et al, 2013), we selected *CG10621*, *MFS3*, and *NPF1* for further functional validation using RNAi assays. Compared with locusts injected with *dsGFP*, locusts injected with *dsNPF1* showed significantly decreased food intake (Fig 5B). However, the amount of food consumed by the *dsCG10621*- and *dsMFS3*-treated locusts did not change significantly (Fig 5B and Appendix Fig S8A–C). We further found that the mRNA level of *NPF1* in the brain was significantly increased after starvation for 12 h (Appendix Fig S9).

To confirm the specificity of *NPF1* acting downstream of *Pw1* in an independent experiment, we quantified the *NPF1* expression after *dsPw1* treatment via qPCR. Compared with locusts injected with *dsGFP*, locusts injected with *dsPw1* showed significant decrease in *NPF1* expression (Appendix Fig S7A). To further verify the *NPF1*-related downstream function of *Pw1*, we measured the body weight, lipid droplet size, and TAG content in *dsNPF1*-treated locusts. *dsNPF1* treatment led to a significant decrease in body weight (approximately 26.8% decrease compared with that of the *dsGFP*-treated control locusts; Fig 5C). Nile red staining of fat body lipid droplets revealed a significant decrease in lipid droplet size in the *dsNPF1*-treated locusts compared with the *dsGFP*-treated control locusts (Fig 5D and E). The significant decrease in TAG content in the *dsNPF1*-treated locusts suggested that TAG catabolism predominates over TAG biosynthesis (Fig 5F). To determine whether the *dsNPF1*-induced reduction in food intake is regulated downstream of *Pw1*, we performed food intake assays by increasing the *NPF1* content via injection of the synthetic *NPF1* into locusts pretreated with *dsPw1*. Food intake within 1 h was significantly increased after injection of synthetic *NPF1* into the *dsPw1*-treated locusts, and as the duration of food intake was prolonged (2 h), the feeding deficiency (approximately 88%) was significantly restored in *dsPw1*-pretreated locusts compared with *dsGFP*-treated control locusts (Fig 5G and H). Therefore, *NPF1* acts downstream of *Pw1* and reduces the food consumption of locusts.

### **Intronic piRNAs mediate the gene expression of NPF1**

Because the PIWI protein initiates piRNA biogenesis, we speculate whether piRNAs contribute to the regulation of *NPF1* expression in the locust brain. We performed small RNA transcriptome sequencing of the brains of *dsPw1*-treated locusts and *dsGFP*-treated control locusts. We mapped the sequencing reads to the locust genome and identified 21 differentially expressed nonredundant piRNAs located in the promoter and gene body regions of the *NPF1* gene (Fig 6A and Appendix Table S1). qPCR expression analysis showed that nine of the piRNAs were differentially expressed in *dsPw1*-treated locusts (Fig 6B and Appendix Fig S10). The nine piRNAs were classified into three groups on the basis of their genomic contexts: promoter region (piRs-1), the first (piRs-2) intron, and the second (piRs-3-I1, piRs-3-I2, and piRs-3-I3) intron (Fig 6A). With the exceptions of piRs-1-P1, piRs-1-P2, piRs-1-P3, and piRs-3-





**Figure 3. Loss of the piRNA pathway reduces food intake and body weight.**

**A** Body weight measurements of *dsPiwi1*-treated locusts and *dsGFP*-treated control locusts (*dsGFP* treated,  $n = 17$  locusts, *dsPiwi1* treated,  $n = 20$  locusts; *dsGFP* treated,  $n = 45$  locusts, and *dsAgo3* treated,  $n = 26$  locusts).  $***P < 0.001$  (Student's *t*-test). ns, not significant (Student's *t*-test). The center line of the boxplots represents median value, the bounds of the box represent 75<sup>th</sup> and 25<sup>th</sup> percentile, and the whiskers represent maximum and minimum value.

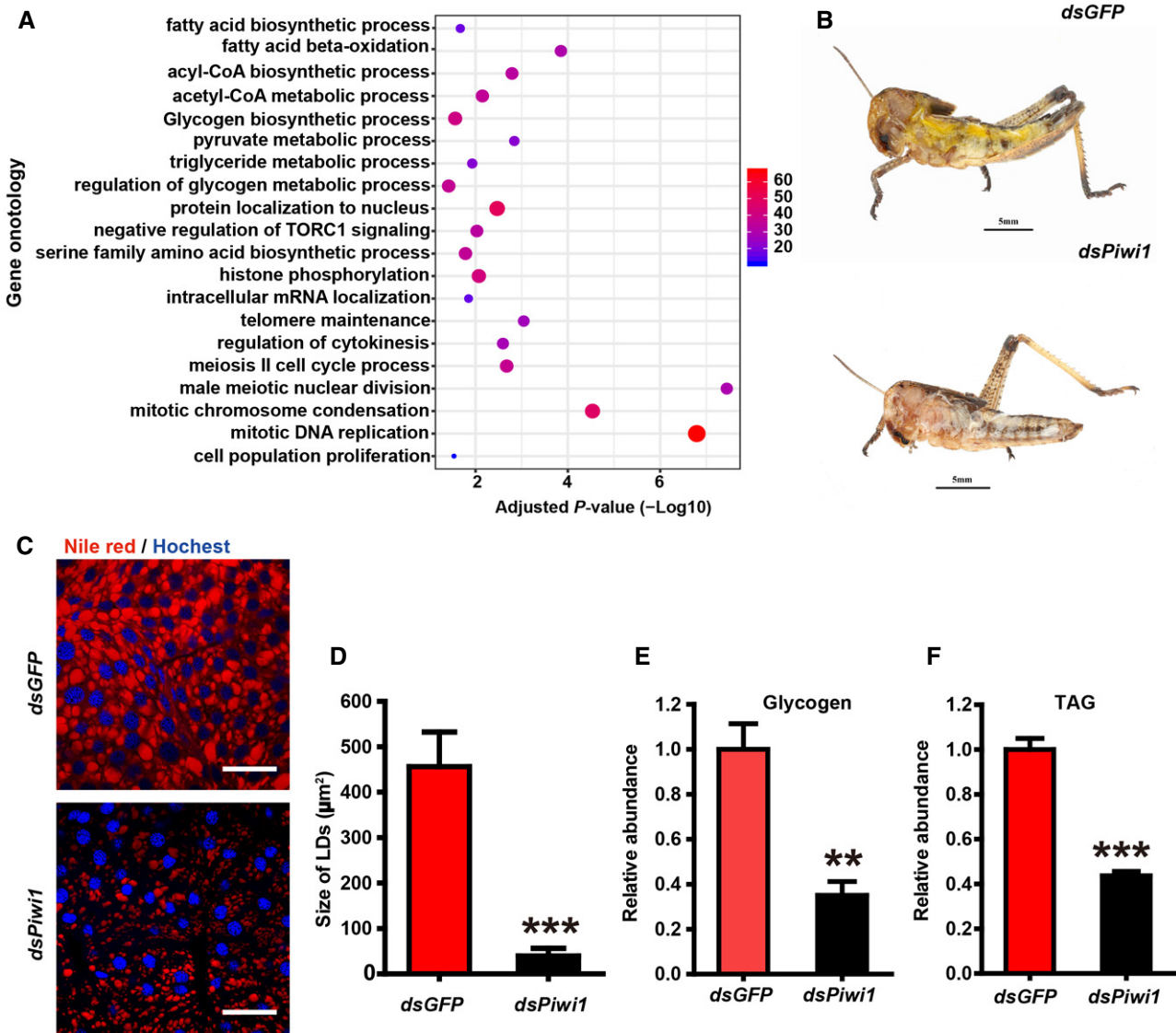
**B** Responses of locust nymphs in a dual-choice test in a Y-tube olfactometer after *Piwi1* knockdown ( $n = 46$  locusts), ns: not significant (Chi-square test).

**C** Analysis of the behavioral phenotypes of locusts in an arena after *Piwi1* knockdown ( $n = 18$  locusts), ns: not significant (Student's *t*-test). TDM, total distance moved.

**D–F** Analysis of the exoskeleton traits of locusts ( $n = 36$  locusts), ns: not significant (Student's *t*-test).

**G** Food intake by *dsPiwi1*-treated locusts and *dsGFP*-treated control locusts within 1 h ( $n = 5$  biological replicates),  $***P < 0.001$  (Student's *t*-test).

Data information: All assays were performed on fourth-instar locusts.



**Figure 4. Suppression of *Piwi1* expression reduces normal metabolic functions.**

- A** Gene ontology enrichments for biological processes in the downregulated genes in the fat body following *dsPiwi1* treatment. The heat map shows the gene ratio percentage and the number of differentially expressed genes for a given GO term/the number of all the locust genes for a given GO term.
- B** Representative vertical images of *dsPiwi1*-treated locusts and *dsGFP*-treated locusts on the fourth day of the fourth instar.
- C** Representative images of Nile red staining of lipid droplets in the fat bodies of *dsPiwi1*-treated locusts and *dsGFP*-treated control locusts ( $n = 5$  biological replicates). The scale bars represent 50  $\mu$ m.
- D** Quantification of lipid droplet staining in *dsPiwi1*-treated locusts and *dsGFP*-treated control locusts ( $n = 5$  biological replicates).
- E, F** Measurements of glycogen ( $n = 5$  biological replicates) and TAGs ( $n = 10$  biological replicates) in fat bodies.

Data information: All data are presented as the mean  $\pm$  SEM (Student's *t*-test, \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , ns: not significant). All assays were performed on 4-day-old fourth-instar locusts.

11, the locations of the five other piRNAs in the locust genome corresponded to the sense strand of *NPF1* (Appendix Table S1). After *dsPiwi1* treatment, the piRNAs in the piRs-1 and piRs-2 categories but not piRs-3-I1 and piRs-3-I2/piRs-3-I3 piRNAs displayed similar expression patterns.

To determine the effects of overexpression/silencing of these piRNAs on *NPF1* expression, we injected a mixture of piRNA mimics/inhibitors corresponding to the same genomic context into

the locust brain. *NPF1* expression was significantly reduced when locusts were injected with piRs-3-I2 and piRs-3-I3 inhibitors, whereas *NPF1* expression was unaffected in locusts injected with piRs-1 (piRs-1-P1, piRs-1-P2, and piRs-1-P3) mimics, piRs-2 (piRs-2-I1, piRs-2-I2, and piRs-2-I3) inhibitors, and piRs-3-I1 mimics (Fig 6C and Appendix Fig S11A–C). Moreover, compared with negative control injection, injection of piRs-3-I2 and piRs-3-I3 inhibitors led to a significant decrease in TAG content and the size of fat body

lipid droplets (Appendix Fig S12A and B). When piRs-3-I2 and piRs-3-I3 mimics were administered, *NPF1* expression was increased by approximately 50% (Fig 6C). piRs-3-I3 contains the typical piR-3-I3 and its piRNA variants that were likely generated by RNA editing and imprecise processing (Appendix Fig S13). The CPM (counts per million reads mapped) of piRs-3-I3 is 449.92 in the unoxidized library. To comprehensively explore the effects of piR-3-I3 and its

piRNA variants, we chose the longest variant as representative for function assays. To confirm the role of piRs-3-I2 and piRs-3-I3 in the regulation of food intake, we quantified the food intake and body weight of locusts following the administration of piRs-3-I2 and piRs-3-I3. Treatments with mimics of piRs-3-I2 and piRs-3-I3 resulted in a 24.5% increase in food intake and a 13.0% increase in body weight (Fig 6D and E), respectively. Conversely, injection of piRs-3-I2 and

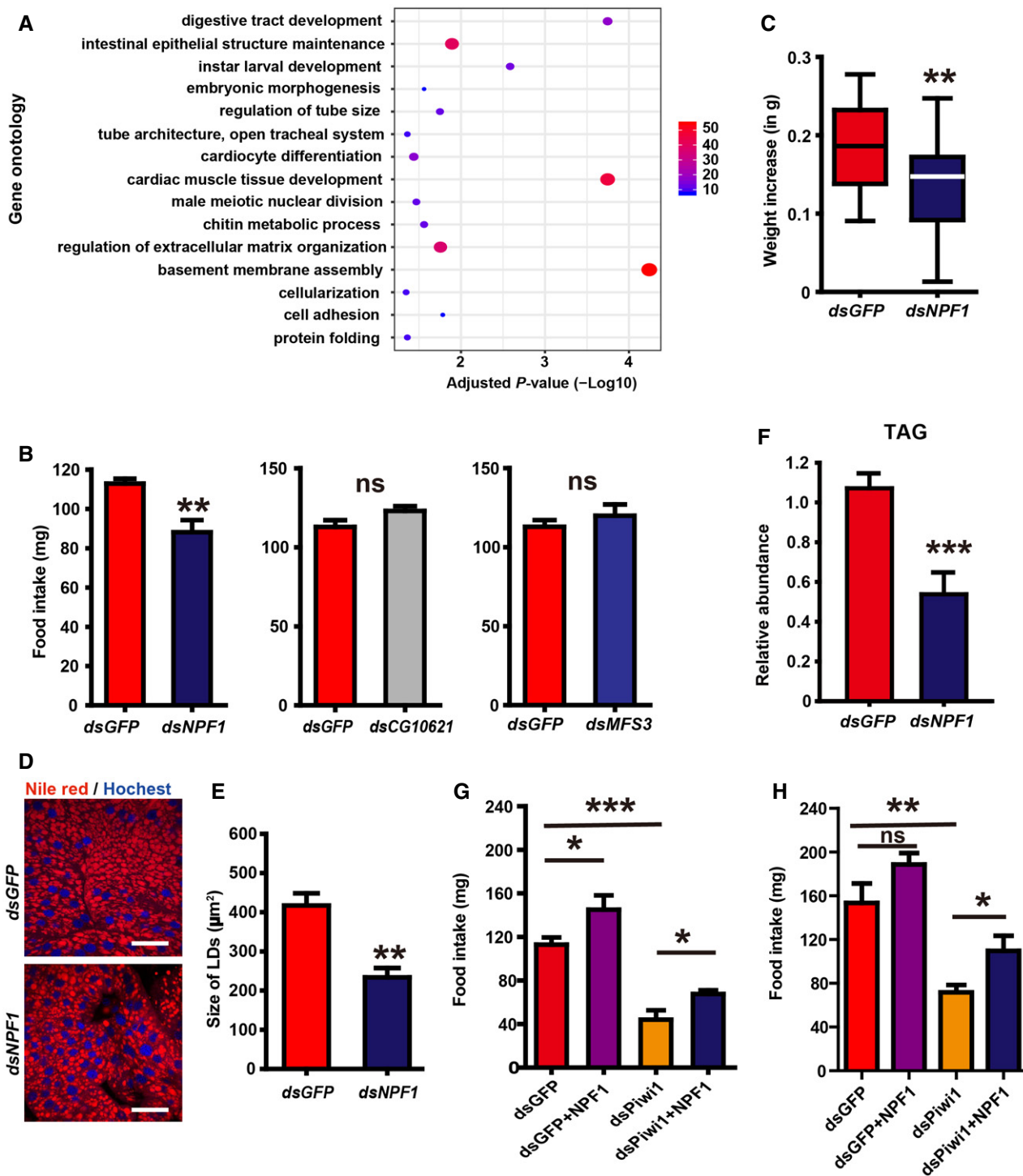


Figure 5.



**Figure 5. Identification of *NPF1* as a target associated with feeding.**

- A Gene ontology enrichments for biological processes in the downregulated genes in the locust brain following *dsPiwi1* treatment. The heat map shows the gene ratio percentage and the number of differentially expressed genes for a given GO term/the number of all locust genes for a given GO term.
- B Food intake within 1 h following treatment with candidate genes (*dsNPF1* treated,  $n = 4$  biological replicates; *dsCG10621* treated,  $n = 5$  biological replicates; and *dsMFS3* treated,  $n = 5$  biological replicates), the data are shown as the mean  $\pm$  SEM (Student's *t*-test,  $**P < 0.01$ ).
- C Measurement of the body weight of *dsNPF1*-treated locusts and the *dsGFP*-treated control locusts (*dsGFP*-treated,  $n = 32$  locusts; and *dsNPF1*-treated,  $n = 29$  locusts). The center line of the boxplots represents median value, the bounds of the box represent 75<sup>th</sup> and 25<sup>th</sup> percentile, and the whiskers represent maximum and minimum value.
- D, E Representative images (D) and quantification (E) of Nile red staining of fat body lipid droplets in *dsNPF1*-treated locusts and *dsGFP*-treated controls ( $n = 5$  biological replicates). The scale bars represent 50  $\mu$ m, the data are shown as the mean  $\pm$  SEM (Student's *t*-test,  $**P < 0.01$ ).
- F Measurement of TAG levels in the fat bodies of *dsNPF1*-treated locusts and *dsGFP*-treated control locusts ( $n = 5$  biological replicates), the data are shown as the mean  $\pm$  SEM (Student's *t*-test,  $***P < 0.001$ ).
- G Food intake within 1 h in the food intake assays ( $n = 5$  biological replicates), the data are shown as the mean  $\pm$  SEM (Student's *t*-test,  $*P < 0.05$ ,  $***P < 0.001$ ).
- H Food intake within 2 h in the food intake assays ( $n = 5$  biological replicates). Five locusts were included in each replicate, the data are shown as the mean  $\pm$  SEM (Student's *t*-test,  $*P < 0.05$ ,  $**P < 0.01$ ).

piRs-3-13 inhibitors decreased food intake and body weight by 42.3% and 28.9% (Fig 6D and E), respectively. To verify that piRs-3-12 and piRs-3-13 are indeed authentic piRNAs in locust brains, we performed the RNA-binding protein immunoprecipitation assay using a monoclonal antibody against the Piwi1 protein in locust brains. Compared with the negative controls, the levels of piRs-3-12 and piRs-3-13 were significantly increased in Piwi1-immunoprecipitated RNAs (Fig 6F). The expression of piRs-3-12 and piRs-3-13 was significantly decreased after *dsPiwi1* treatment (Fig 6B). Both piRs-3-12 and piRs-3-13 showed an uracil at their 5' end position and were detected in the oxidized small RNA library. These results suggested that the piRs-3-12 and piRs-3-13 are authentic piRNAs in locust brains and their biogenesis is Piwi1 dependent. Collectively, intronic piRs-3-12 and piRs-3-13 mediate *NPF1* expression and food intake.

**Intronic piRNAs may enhance the splicing of *NPF1* pre-mRNA**

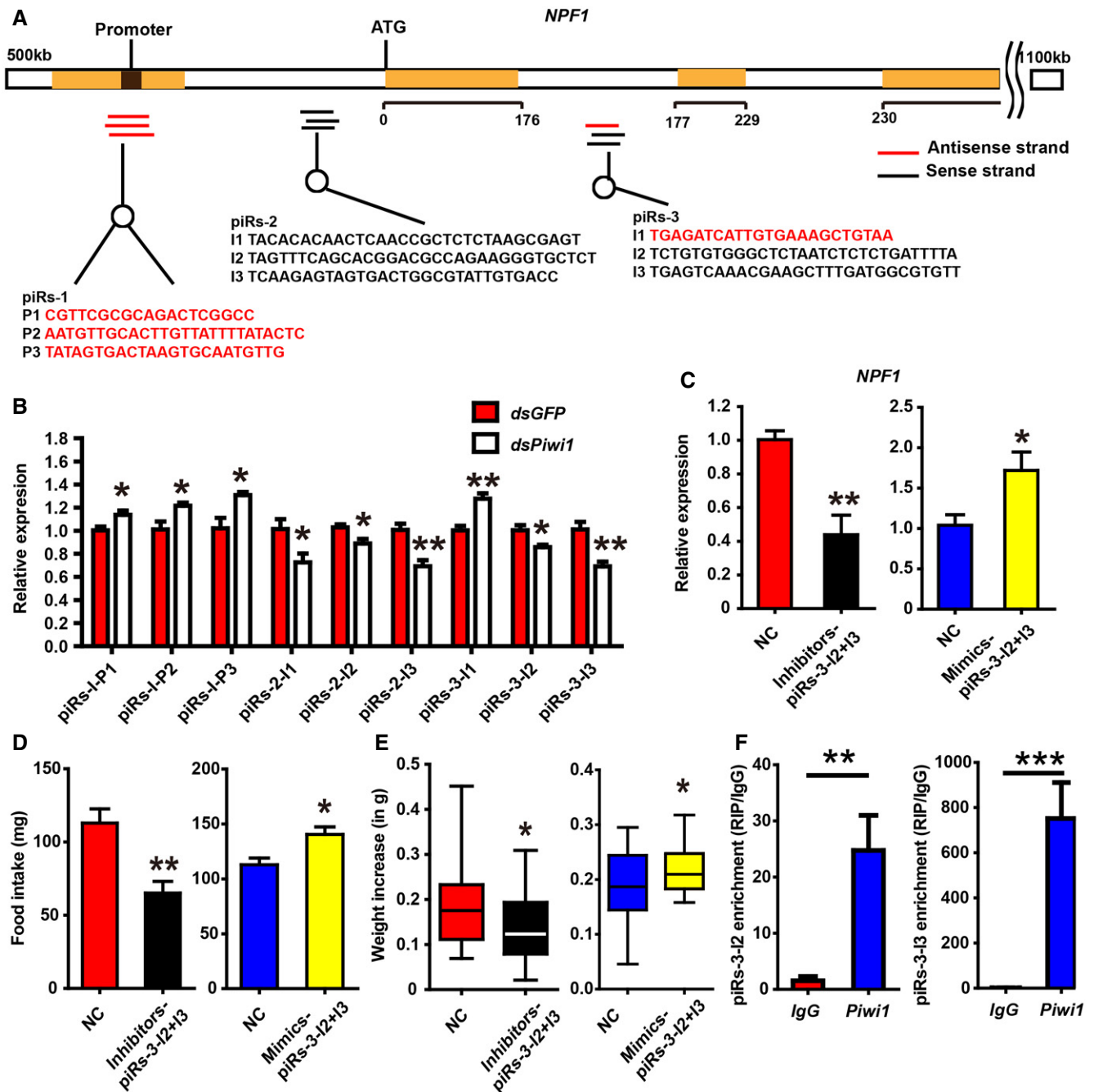
To explore the molecular mechanism underlying the ability of intronic piRNAs to mediate *NPF1* expression, we quantified the nuclear expression of piRs-3-12 and piRs-3-13 using subcellular fractionation experiments due to the absence of piRNA-binding sites on the mature RNAs of *NPF1* (Fig 7A). The qPCR data showed comparable nuclear expression of piRs-3-12 and piRs-3-13 (Fig 7B). Due to the potential nuclear functions of piRs-3-12 and piRs-3-13, we quantified the expression level of the precursor mRNA transcripts of *NPF1* (*pre-NPF1*). We verified that inhibitors of piRs-3-12 and piRs-3-13 could enter the nuclei of brain cells (Appendix Fig S14A and B). Inhibitor injection resulted in the accumulation of *pre-NPF1* in the nucleus (Fig 7C), which is consistent with *dsPiwi1* treatment (Fig 7D). A decrease in the level of mature *NPF1* RNA (Fig 6C) and the accumulation of *pre-NPF1* in the nucleus (Fig 7C) confirmed that piRs-3-12 and piRs-3-13 are probably required for the processing of precursor into mature mRNAs.

Because the cellular expression of piRs-3-13 was higher than that of piRs-3-12 in the brain (Appendix Fig S15), we chose piRs-3-13 to further explore the role of intronic piRNA in splicing *in vivo*. We transfected pcDNA 3.1(+) constructs fused to the exon-intron-exon structure flanking the target site of piRs-3-13 into 293T cells. Compared with treatment with mutated piRs-3-13 mimics, treatment with wild-type piRs-3-13 mimics resulted in a significant increase in the expression of spliced *NPF1* but had no significant effect on the expression of *pre-NPF1* (Fig 7E). These results suggest that piRs-3-13

plays direct regulatory roles at the posttranscriptional level but not at the pretranscription level.

The *NPF1* gene contains two isoforms in locusts and is alternatively spliced into two transcripts encoding distinct mature neuropeptides. The removal of a cassette exon in *pre-NPF1* in locust brains produced the mature neuropeptides that were studied in this study. To test whether piRs-3-13 participates in the processing of *pre-NPF1*, we measured its binding with heterogeneous nuclear RNA protein F/H (hnRNP F/H; a protein that comprises hnRNP complexes and is involved in alternative RNA splicing process; Appendix Figs S1F, S16, and S17A) using RNA-binding protein immunoprecipitation experiments. The levels of *pre-NPF1* and piRs-3-13 were significantly increased in hnRNP F/H-immunoprecipitated RNAs in the brains of locusts treated with piRs-3-13 mimics compared with those from the brains treated with the negative control (Fig 7F). This interaction of hnRNP complexes with piRs-3-13 suggests a potential regulatory role for *NPF1* in the RNA splicing process.

Using BPP software (Zhang *et al*, 2017), a branch point site for proper splicing by the spliceosome was identified in the intronic target site of piRs-3-13. The intron 2 of *NPF1* is a long intron (230 Kb), and its RNA processing is dependent on recursive splicing (RS) which is a common mechanism in long intron genes (Kelly *et al*, 2015). Branch point sequences were usually present in the 5–50 bp regions upstream of RS sites. A RS site, which was located downstream of the branch point site, was predicted using HOMER v4.9 based on the position frequency matrices retrieved from the previous study (Appendix Fig S18) (Burnette *et al*, 2005). Because U2AF is required for efficient recognition of ratchet points in recursive splicing (Duff *et al*, 2015), we experimentally validated the predicted RS sites by measuring its binding with U2AF65 using RNA-binding protein immunoprecipitation experiments (Appendix Fig S17B). The levels of the predicted RS sites were significantly increased in U2AF65-immunoprecipitated RNAs in the brains of locusts treated with piRs-3-13 mimics compared with those from the brains treated with the negative control (Fig 7G), suggesting that the predicted RS site was involved in recursive splicing. Furthermore, we found a hairpin-like structure of *pre-NPF1* in the intronic target site of piRs-3-13 (Fig 7H). To experimentally verify the formation of hairpin-like structures, we synthesized artificial *pre-NPF1* transcripts and incubated them with RNase III, an endoribonuclease that can cleave hairpin structures of RNAs. The level of *pre-NPF1* was significantly decreased after RNase III treatment (Fig 7I), indicating that



**Figure 6. piRs-3-I2 and piRs-3-I3 promote feeding in locusts.**

- A Representative genomic loci of differentially expressed piRNAs located within the *NPF1* gene (beige colored regions are exons).
- B mRNA levels of differentially expressed piRNAs located within the flanking region of the *NPF1* gene in *dsPiwi1*-treated locusts and *dsGFP*-treated control locusts ( $n = 5$  biological replicates).
- C The mRNA expression of *NPF1* was measured by qPCR after a mixture of inhibitors or mimics of piRs-3-I2 and piRs-3-I3 was injected ( $n = 5$  biological replicates).
- D Effects of inhibitors ( $n = 6$  biological replicates) and mimics ( $n = 4$  biological replicates) of piRs-3-I2 and piRs-3-I3 on food intake.
- E Body weight measurement of locusts administered inhibitors (NC,  $n = 25$  locusts; Inhibitors—piRs-3-I2 + I3,  $n = 23$  locusts) and mimics (NC,  $n = 37$  locusts; Inhibitors—piRs-3-I2 + I3,  $n = 39$  locusts) of piRs-3-I2 and piRs-3-I3. The center line of the boxplots represents median value, the bounds of the box represent 75<sup>th</sup> and 25<sup>th</sup> percentile, and the whiskers represent maximum and minimum value.
- F RNA-binding protein immunoprecipitation assay of piRs-3-I2 and piRs-3-I3 in Piwi1 immunoprecipitates from brain tissue extracts ( $n = 6$  biological replicates). Normal mouse IgG was used as a negative control.

Data information: The qPCR data are shown as the mean  $\pm$  SEM (Student's *t*-test, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ).

the intronic target site of piRs-3-I3 participates in hairpin structure formation. To further determine the effects of hairpin structure formation on splicing efficiency at branch point sites, we transfected both wild-type and mutated constructs fused with the exon-intron-exon structure flanking the target site of piRs-3-I3 into 293T cells. In the mutated construct, the complementary nucleotides located in the stem region were randomly mutated to abolish the hairpin-

forming potential. The expression of spliced *NPF1* was significantly increased in 293T cells transfected with the mutant construct and the piRs-3-I3 mimics compared with those transfected with the wild-type construct and the piRs-3-I3 mimics (Fig 7J), indicating that hairpin structure formation reduces the RNA splicing efficiency at the branch point site. As expected, the expression of *pre-NPF1* did not significantly vary between cells transfected with the wild-type

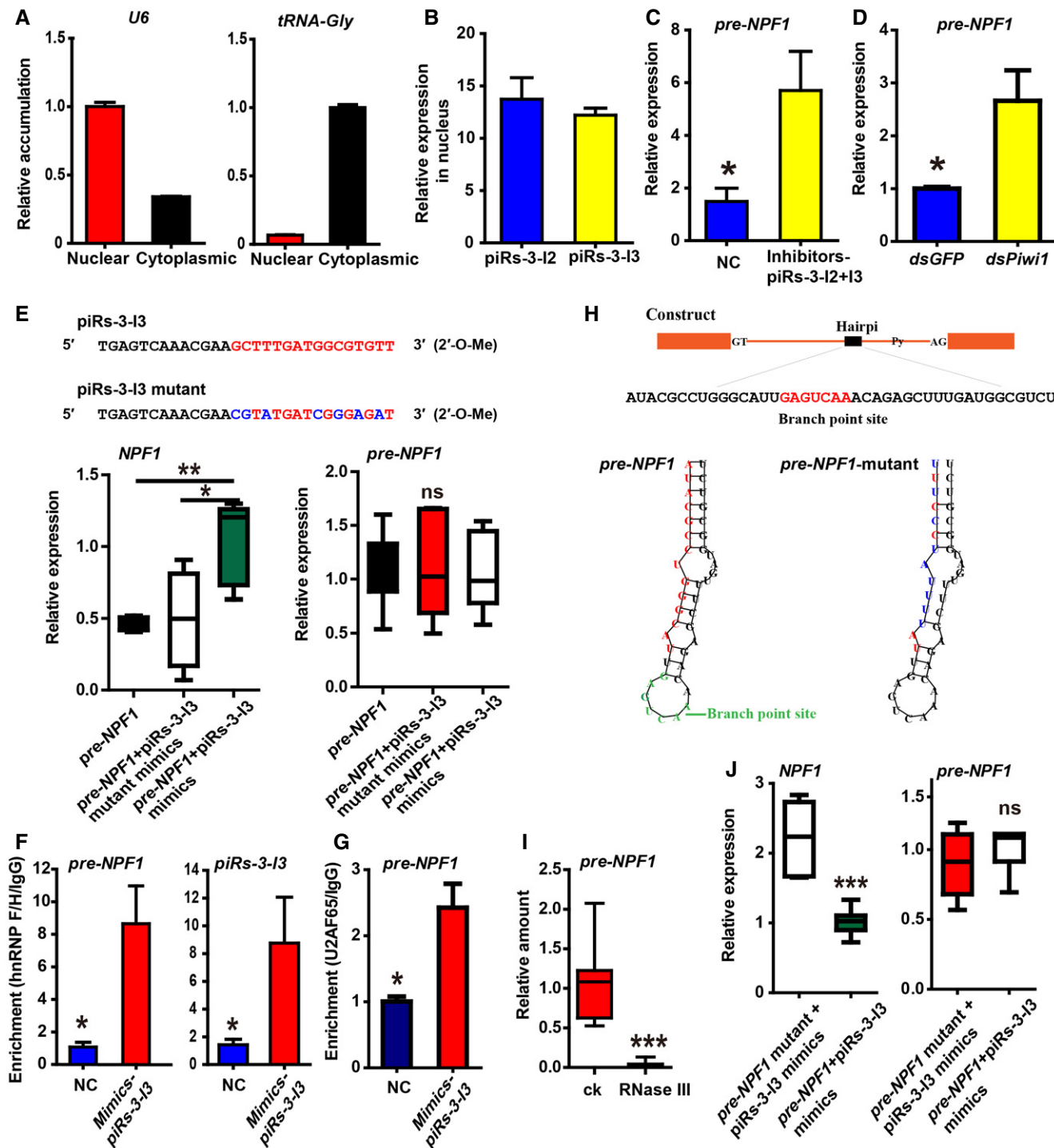


Figure 7.

### Figure 7. The expression of piRs-3-I2, piRs-3-I3, and pre-NPF1.

- A U6 RNA was used as a nuclear RNA control (relative value set to 1 in the nuclear fraction), glycine tRNA was used as the cytoplasmic RNA control (relative value set to 1 in the cytoplasmic fraction;  $n = 4$  biological replicates), and U6 and tRNA-Gly was used as an endogenous control.
- B Relative expressions of piRs-3-I2 and piRs-3-I3 determined by qPCR and high-throughput sequencing in the nucleus ( $n = 3$  biological replicates). The brain tissues were homogenized and then divided into two parts of equal volume; one was prepared for high-throughput sequencing and the other was used for piRNA quantification in nuclear and cytoplasmic fractions after subcellular RNA fractionation. The relative piRNA expression in nucleus was determined by multiplying the relative nuclear piRNA fraction estimate by qPCR quantification by the normalized count of piRNA in high-throughput sequencing libraries.
- C Expression level of *pre-NPF1* in the nucleus after a mixture of inhibitors (piRs-3-I2 and piRs-3-I3) was injected ( $n = 4$  biological replicates).
- D qPCR expression analyses of *pre-NPF1*. The precursor mRNA expression level was measured by qPCR after the *dsPiwi1* treatments ( $n = 6$  biological replicates).
- E Expression levels of the mature *NPF1* gene and *pre-NPF1* in 293T cells transfected with piRs-3-I3 and mutated piRs-3-I3 mimics ( $n = 5$  biological replicates).
- F Analysis was performed to amplify *pre-NPF1* and piRs-3-I3 in hnRNP F/H immunoprecipitates from extracts of brain tissue treated with piRs-3-I3 mimics compared to the negative controls (*pre-NPF1*  $n = 3$  biological replicates; piRs-3-I3  $n = 4$  biological replicates).
- G Analysis was performed to amplify *pre-NPF1* in U2AF65 immunoprecipitates from extracts of brain tissue treated with piRs-3-I3 mimics compared to the negative controls (*pre-NPF1*  $n = 3$  biological replicates).
- H The predicted secondary structures of WT and mutated *pre-NPF1*.
- I Expression level of *pre-NPF1* following ck and RNase III treatment ( $n = 7$  biological replicates). The center line of the boxplots represents median value, the bounds of the box represent 75<sup>th</sup> and 25<sup>th</sup> percentile, and the whiskers represent maximum and minimum value.
- J Expression level of the mature *NPF1* gene and *pre-NPF1* in 293T cells transfected with *pre-NPF1* and mutated *pre-NPF1* ( $n = 5$  biological replicates). NC means negative control. The center line of the boxplots represents median value, the bounds of the box represent 75<sup>th</sup> and 25<sup>th</sup> percentile, and the whiskers represent maximum and minimum value.

Data information: The qPCR data are shown as the mean  $\pm$  SEM (Student's *t*-test, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ).

and mutated constructs (Fig 7J). The lack of PIWI expression in 293T cells suggested the RNA splicing enhancement mediated by piRs-3-I3 was PIWI independent. Taken together, these results confirmed that piRs-3-I3 might enhance the RNA splicing processing of *NPF1* by preventing intronic hairpin formation at the branch point site.

## Discussion

Although PIWIs/piRNAs are thought to be germline specific, increasing evidence for their functions in somatic cells is emerging. Here, we discovered that PIWIs/piRNAs exist in several major organs in locusts and are expressed at significant levels in the nervous system, suggesting much broader roles for PIWIs/piRNAs outside the gonads than were previously identified. Furthermore, we proposed that Piwi1 is essential for maintaining food intake in locusts by establishing a novel nuclear function for piRNAs and *NPF1* interactions in the nervous system (Fig 8). Mechanistically, this nuclear function of piRNAs may prevent the formation of hairpin structures at the branch point site of *pre-NPF1*.

Locusts are agricultural pests that pose a substantial economic threat worldwide. The identification of a new function for piRNAs in insect feeding in this study expands the current understanding of methodological approaches for pest control and may stimulate the development of novel approaches for locust control. Throughout the world, locust control mainly involves pesticides, which endanger the environment and human health. In addition to the use of pesticides, RNAi approaches in which dsRNAs are designed to suppress target genes and then the target gene is processed into an siRNA have been considered as potential strategies for pest management (Kunte et al, 2020). However, sufficient nucleotide sequence similarity of the target gene raises the possibility of detrimental cross-species effect (Jackson et al, 2003; Allen, 2017). Approximately 79% similarity (15/19) between an siRNA and a target gene with 11 contiguous matched bases is sufficient to trigger the expression silencing of off-target transcripts. Thus, searching for highly

selective and species-specific target genes is a top priority for developing safe RNAi-based pest control approaches. Fortunately, because cross-species sequence conservation of piRNAs is remarkably low (Zhang et al, 2011), piRNA-based RNAi can be novel tools for suppressing the expression of vital genes for the selective control of targeted pests, thus preventing possible off-target effects on other insects in the ecosystem.

Here, we present several lines of evidence suggesting that the locust can serve as a representative species for determining PIWI function in non-*Drosophila* insects. First, ping-pong piRNA biogenesis is restricted to the germline in *Drosophila*, while it is active in the soma and germline in most other insects, such as locusts (Lewis et al, 2018). Second, although non-*Drosophila* insects have variable member numbers of two *Piwi-like* genes, the two *Piwi-like* genes in locusts represent a typical number in the majority of non-*Drosophila* insects. Finally, the split between the *Piwi* and *Aub* branches in *Drosophila* suggests that *Aub* originated from lineage-specific gene duplication events in only a few insect genera within Diptera. Another possibility related to the evolutionary origin of *Piwi-like* genes is that the emergence of *Piwi-like* duplicates appears to predate the last common ancestor of extant insects. The selective constraint acting on the *Piwi/Aub* duplicates in *Drosophila* is more relaxed than that acting on *Piwi1/Piwi2* duplicates in the migratory locust and beyond. This is consistent with the functional deprivation of ping-pong piRNA biogenesis in somatic tissues in *Drosophila* (Lewis et al, 2018). Therefore, the orthology of PIWI proteins in *Drosophila* cannot be directly extended to other non-*Drosophila* insects, and further functional examination of PIWI genes in non-*Drosophila* insects is required.

The PIWI/piRNA pathway is well known for its role in repressing the activity of harmful genetic elements such as transposons. In addition to repressing transposons, PIWIs/piRNAs appear to have the capacity to regulate the stability of several protein-coding genes (Saito et al, 2009; Rajasethupathy et al, 2012; Ross et al, 2014; Barckmann et al, 2015; Vourekas et al, 2016). However, the molecular mechanism of PIWIs/piRNAs in the regulation of protein-coding genes has remained an open question. We showed that knockdown





of *Piwi1* expression affects the expression of hundreds of protein-coding genes and that one of the significantly differentially expressed genes, *NPF1*, is associated with intronic piRNA-mediated regulation. The decreased *NPF1* expression caused by the *dsPiwi1* treatments confirmed the specificity of *NPF1* acting downstream of *Piwi1*. The administration of piRs-3-I2 and piRs-3-I3, of which biogenesis is dependent on *Piwi1* expression, resulted in the significant increases in *NPF1* expression and food intake. However, the *pre-NPF1* levels were not affected by the administration of piRs-3-I2 and piRs-3-I3, suggesting the regulatory mechanism operated at posttranscriptional level. After the administration of piRs-3-I2 and piRs-3-I3, the measurement of the ratios of *NPF1* and *pre-NPF1* *in vivo* should show the direct comparison between these conditions. However, it is technically challenging to measure the ratios of *NPF1* and *pre-NPF1* *in vivo*. Therefore, we turned to our experiments in cell lines to further prove the piRNA-mediated splicing mechanism we proposed. Our results also showed that the administration of *NPF1* increased food intake in locusts, suggesting a positive correlation of *NPF1* expression and food intake in a variety of conditions. The results provided evidence that the *Piwi1*/piRNA expression plays a role in maintaining *NPF1* expression, ultimately affecting food intake in locusts. Aside from gene regulation directly mediated by PIWIs/piRNAs, the activation of TE expression caused by PIWI knockdown may also affect gene regulation (Lin *et al.*, 2020). TEs can be mobilized and create *de novo* TE insertions in promoter regions and exonic regions, and insertions in promoter regions cause them to gain novel regulatory elements, while insertions in exonic regions results in disrupted protein function (Perrat *et al.*, 2013). For example, inactivation of TE-derived regulatory elements abolishes the expression of the *POMC* gene, which functions in the brain to control food intake (Lam *et al.*, 2015). It has even been proposed that PIWI proteins can regulate the expression of mRNAs in a piRNA-independent manner in human late-stage germline cells (Gou *et al.*, 2017). Therefore, the results of this study provide a novel clue for the multifaceted regulatory effects of PIWI proteins on protein-coding genes.

PIWIs/piRNAs can silence genes pretranscriptionally and posttranscriptionally in flies, silkworms, *Aplysia*, and mice (Rajasethupathy *et al.*, 2012; Kiuchi *et al.*, 2014; Pezic *et al.*, 2014; Watanabe & Lin, 2014). PIWI/piRNA-mediated pretranscriptional regulation represses the production of targets through chromatin modifications such as histone or DNA methylation (Aravin *et al.*, 2008; Sienski *et al.*, 2012). In contrast, PIWIs/piRNAs mediate posttranscriptional regulation by binding to the coding region or the 3'UTR of target mRNAs through endonucleolytic degradation and cleavage (Czech & Hannon, 2016). Recent studies have shown that PIWIs/piRNAs can enhance precursor mRNA splicing, thereby affecting host gene expression. Small RNA-dependent splicing regulation can be achieved in concert with piRNAs based on repressing chromatin states and siRNA based on *cis*-element masking (Wang & Chekanova, 2016; Teixeira *et al.*, 2017). However, in contrast to previous observations, our findings revealed a novel case of nuclear piRNA-mediated splicing enhancing at the posttranscriptional level. Consistent with previous results, piRNA-L-163\_igs aligned to intron 10 of the *LAMC2* gene can influence *LAMC2* gene expression (Mei *et al.*, 2015). Hairpin structures of precursor mRNA transcripts play an important role in the splicing of the RNA in a location-dependent manner, and their formation

upstream of branch point sites may improve the recognition of configuration by splicing machinery (Muh *et al.*, 2002; Buratti & Baralle, 2004; Liu *et al.*, 2010). In locusts, both the hairpin structure and branch point site were identified in the intronic target site of piRs-3-I3. Furthermore, disruption of hairpin formation at the branch point site combined with treatment with piRs-3-I3 mimics decreased the RNA splicing efficiency of *NPF1*. The migratory locust has a huge genome size, which is approximately two and 37 times larger than those of the human and fruit-fly genomes, respectively. The intron sizes in the locust genome were subjected to extensive expansion (Wang *et al.*, 2014). Due to dramatically intron size expansion, RNA processing in locusts includes recursive splicing (RS) which is a common mechanism in long intron genes (Kelly *et al.*, 2015). The 3' end of upstream exon is joined to segments within the following intron in a stepwise manner before splicing to the 5' end of downstream exon by the regeneration of 5' donor (GT) sites in RS sites. Although a vast majority of introns in fruit flies are not large intron, a specific RS motif was defined in large introns of fruit flies. In this study, the branch point sequence we studied was located upstream of a RS site, which was involved in recursive splicing. The presence of branch point site, polyT/C tract, and recursive splice site indicated this region was recognized and processed by splicing assembly complex. The competitive binding (Stem #1) of piRs-3-I3 forced the branch point site into an unpaired structure that would be bettered recognized by the splicing machinery. Therefore, mechanistically, the piRNA-mediated splicing processing proposed in this study is achieved through direct action on *pre-NPF1* rather than indirectly by piRNA-mediated changes in chromatin silencing. The accumulation of spliced transcripts for IVS3 is substantially increased in dysgenic germ cells. The reduction in H3K9me3 modification levels is also observed over the IVS3 transgenic reporter in dysgenic progeny when compared to nondysgenic progeny. These results suggest that H3K9me3 deposition by piRNA complexes is required for splicing regulation by the piRNA-mediated, chromatin-based mechanism (Teixeira *et al.*, 2017). We quantified the H3K9me3 deposition of the branch point site and its flanking region using CUT&Tag qPCR analysis. Compared to that in IgG, no significantly H3K9me3 enrichment in the brains of locusts showed the branch point site was devoid of H3K9me3 signals (Appendix Fig S19A and B). This result further excluded the possibility of piRNA-mediated splicing changes by chromatin-based mechanism in this study. Collectively, our results point to a novel mechanism by which piRs-3-I3 may compete for the binding site recognized by complementary nucleotides of the hairpin structure and thus may prevent hairpin formation at the branch point site.

## Materials and Methods

### Insect rearing

Locusts (the migratory locust, *L. migratoria*) were maintained at the Institute of Zoology, Chinese Academy of Sciences, China. Nymphs were reared in large cages (40 cm × 40 cm × 40 cm) at a density of 500–1,000 insects per container under a 14 h:10 h light: dark cycle at 30°C ± 2°C on a diet of fresh greenhouse-grown wheat seedlings and bran (Kang *et al.*, 2004).

## Identification and evolutionary analysis of PIWI genes

The official gene sets of 130 insect species, including 62 dipteran insects, 33 hymenopteran insects, 10 hemipteran insects, seven coleopteran species, nine lepidopteran insects, and representatives from Orthoptera, Phthiraptera, Phasmatoptera, Trichoptera, Thysanoptera, Isoptera, Blattodea, Ephemeroptera, and Odonata, were retrieved from their respective genome databases (Jiang *et al*, 2017). A total of 174 arthropod species, including 10 chelicerate species, five crustacean species, and two noninsect hexapod species, were also included as basal arthropod species. The reciprocal BLAST hit criterion was used to determine the orthologous relationships of PIWI genes (Altschul *et al*, 1997). Multiple alignments of PIWI were generated using MAFFT alignment software version 7.215 (Kato *et al*, 2009). The protein distance matrices were calculated using the distmat program in EMBOSS version 6.6.0. The maximum-likelihood phylogenetic tree was generated on the basis of the protein sequences of PIWIs by using FastTree software version 2.1.10 (Price *et al*, 2010). The reliability of each tree branch was evaluated using the Shimodaira–Hasegawa (SH) branch support test with 1,000 resamples. The protein distance matrix for *Piwi*, *Aub*, and *Ago3* was constructed using the ProtDist program of the PHYLIP package version 3.695 with the Dayhoff PAM matrix. The ratios of nonsynonymous-to-synonymous nucleotide substitution rates were determined using the program *KaKs\_Calculator* version 2.0 by the MA method (Wang *et al*, 2010).

## RNA extraction and quantitative RT-PCR (qRT-PCR)

Total RNA was extracted from locust brains by using TRIzol (Invitrogen). cDNA was synthesized using Moloney murine leukemia virus reverse transcriptase (Promega) with an oligo (dT) primer. Small RNAs were isolated using the mirVana miRNA isolation kit (Ambion). For piRNA quantification, the polyadenylated RNAs were generated using *Escherichia coli* poly(A) polymerases (Tiangen) and were subjected to first-strand cDNA synthesis using the modified Oligo-d(T) primers, which contain a specific sequence in the 5' end. The specific sequence was used for designing upstream PCR primers. qRT-PCR was performed using SYBR Green for mRNAs and piRNAs in a LightCycler 480 instrument (Roche). The ribosomal protein RP49 and U6 snRNA were used as endogenous controls for normalization of mRNA and piRNA expression levels, respectively. The primers used for qRT-PCR are shown in Appendix Table S2.

## High-throughput transcriptome sequencing of small RNAs and mRNAs

Total RNA was resuspended in 5× borate buffer (148 mM borax and 148 mM boric acid, pH 8.6) and treated with 200 mM sodium periodate for 10 min at room temperature. Glycerol was added to quench the unreacted NaIO<sub>4</sub>, and the treated RNA was collected by ethanol precipitation.

Small RNAs with a size of 18–33 nt were excised and used to generate small RNA libraries with a TruSeq small RNA sample preparation kit (Illumina) following the manufacturer's protocol. Raw sequencing data were filtered using Trimmomatic 0.36 software (Bolger *et al*, 2014) and mapped to the locust genome allowing

two mismatches using Bowtie 1.1.2 software (Langmead, 2010). snoRNAs, microRNAs, tRNAs, and rRNAs were filtered in accordance with a previous study (Wang *et al*, 2015). For piRNA expression quantification, piRNAs showing identical 5' start sites in the locust genome were collapsed together.

Total RNA from locust brains was extracted from the locusts in two independent replicates of *dsGFP*-treated and *dsPiwi1*-treated experiments. An Agilent 2100 Bioanalyzer (Agilent) was used to assess the quality of the RNA. For mRNA sequencing, cDNA libraries were prepared in accordance with Illumina protocols and sequenced on the Illumina NovaSeq platform. Raw sequencing data were filtered using Trimmomatic 0.36 software (Bolger *et al*, 2014) and mapped to the locust genome by using HISAT 22.1.0 software (Kim *et al*, 2019). Differential gene expression between samples was quantified at the gene level by using the read summarization program featureCounts v2.0.0 (Liao *et al*, 2014). Differential analysis of the count data was performed with the DESeq2 v1.18.1 package (Love *et al*, 2014). Biological process enrichment analysis of differentially expressed genes was performed using the “clusterProfiler” and “GO. db” packages in the R v3.5.0 program (Yu *et al*, 2012).

## RNAi

The primers used for *dsRNA* synthesis were fused with the T7 promoter sequence, and RNA was prepared using the RiboMax express RNAi system (Promega). The brains of third-instar locusts were microinjected with *dsRNA*, and a second injection was performed on day 2 of the fourth-instar locusts (3.45 µg/locust for each injection) to improve RNAi efficiency and specificity. Green fluorescence protein (GFP) was used as a negative control. The brains of the test locusts were harvested 48 h after the second injection, snap frozen, and stored at –80°C.

## Western blot analysis

Locust brains (8–10 individuals/sample) were collected and homogenized in lysis buffer (CWBIO) containing protease inhibitor (CWBIO). The total protein content was determined using a bicinchoninic acid protein assay kit (Thermo Scientific). The extracts were reduced, denatured, and separated by gel electrophoresis on a 10% SDS-PAGE gel and transferred to a polyvinylidene difluoride membrane (Millipore). The membrane was incubated separately with specific antibodies against Piwi1, Piwi2, and Ago3 (mouse anti-Piwi1 serum, 1:500; mouse anti-Piwi2 serum, 1:500; and mouse anti-Ago3 serum, 1:500, respectively), hnRNP F/H (Mouse monoclonal antibody, 1:2,000, Abcam), and U2AF65 (Mouse monoclonal antibody, 1:200, Santa Cruz Biotechnology). Tubulin was used as an endogenous control (rabbit polyclonal antibody, 1:5,000, CWBIO). Goat anti-rabbit IgG was used as the secondary antibody (1:10,000, CWBIO). The intensities of the Western blot bands were quantified using densitometry with Quantity One software.

## Immunoprecipitation

Immunoprecipitation was performed according to a previous study (Zhao *et al*, 2021). Approximately 15 brains were homogenized using 1× PBS buffer (pH 7.2) supplemented with a protease inhibitor cocktail (Thermo Fisher Scientific). Then, 50 µl of Dynabeads Protein G

(Novex, Thermo Fisher Scientific) was mixed with 5  $\mu\text{g}$  of anti-Piwi1 monoclonal antibody or anti-Mouse IgG (Merck Millipore) antibody before being incubated with 200  $\mu\text{l}$  of protein extracts for 15 min at room temperature. Approximately 10% of the total protein was reserved as input. After washing three times with washing buffer (Novex), the antibody-protein complex was disassociated from the beads with elution buffer (Novex) for Western blot analysis.

### Food intake and body weight measurement

The effects of RNAi-mediated knockdown of genes on food intake and body weight were analyzed. All locusts were starved for 16 h prior to the measurements. On the first day of RNAi treatment, the locusts were weighed, and a second measurement was performed on day 4 of the fourth instar. The increase in body weight was determined by calculating the difference between the weight of each locust at the beginning and the end of the experiment. The food intake of fourth-instar locust nymphs was measured after 1 h. Fresh wheat seedlings (4 g) were fed to 10 locusts kept in small plastic containers. The unconsumed wheat seedlings were weighed after 1 h and the weight was recorded. The experiment was repeated three times.

### Body size measurement and behavioral assay

A Vernier caliper was used to measure the body size of each locust. The head width was measured from one side to the other across the head. The tergum length was measured in accordance with a previously described method (Yang *et al*, 2019a, 2019b). The femur length was recorded as the length of the posterior femur. The number of locusts measured in each group was as follows: *dsPiwi1*-treated group,  $n = 36$  locusts; and *dsGFP*-treated group,  $n = 36$  locusts.

A Y-tube olfactometer was used to analyze the behavioral responses of individual locusts to volatiles from wheat seedlings in the absence of any visual cues. A response of “first choice” for wheat seedlings or air (when the locust moved more than 5 cm into either arm) or “no choice” (N.C.) within a period of 5 min was recorded for each locust.

A rectangular plexiglass arena (60 cm  $\times$  30 cm  $\times$  30 cm) containing two chambers was used for the behavioral assay. Locust behavior was recorded for 10 min by using a video-tracking system and analyzed with EthoVision XT software (version 11.5; Noldus Information Technology) in accordance with the method used in our previous work (Wei *et al*, 2019). The total distance traveled (traveled distance in cm) by locusts presented with 20  $\mu\text{g}$  of fresh wheat seedlings and control locusts was objectively measured without bias. In the behavioral assay, 16–35 locusts from each treatment group were tested, which is in accordance with the sample size reported in previous studies (Ma *et al*, 2011; Hou *et al*, 2017).

### Immunofluorescence assays and lipid staining

Whole-mount double immunohistochemistry in the locust brain was performed by using affinity-purified monoclonal mouse antibodies against Piwi1, Piwi2, and Ago3 (1:100, BGI), hnRNP F/H (1:100, Abcam), and U2AF65 (1:50, Santa Cruz Biotechnology). An Alexa Fluor 488-conjugated goat anti-mouse antibody (1:500, Life Technologies Cat. A-11008) was used as the secondary antibody for Piwi1, Piwi2, and Ago3 staining. For lipid staining, lipids were

visualized by staining the locust fat body with Nile red (0.5 mg/ml, Thermo Fisher N-1142) for 1 h at room temperature.

### TAG and glycogen level measurements

TAG levels were measured using a triglyceride assay kit (Applygen Technologies Inc.) following the manufacturer's protocols (Ding *et al*, 2018). Fat bodies were homogenized in 100  $\mu\text{l}$  of lysis buffer and heated at 70°C for 10 min to inactivate the endogenous enzymes. The samples were incubated with triglyceride reagent for 30 min at 37°C and read with a SpectraMax Plus 384 instrument at a wavelength of 540 nm. The protein concentration was measured using a bicinchoninic acid protein assay kit (Thermo Fisher), and the TAG content was calculated on the basis of a standard curve for TAG, with the standard samples being run in parallel with the experimental samples.

Glycogen levels were determined using the Glycogen Colorimetric Assay Kit II (BioVision Inc.). In brief, fat bodies were homogenized in 200  $\mu\text{l}$  of ddH<sub>2</sub>O on ice and boiled for 10 min to inactivate enzymes. The samples were centrifuged at 13,000  $g$  for 15 min. The supernatant was used for the glycogen assay. The measured values were normalized to lysate protein levels.

### Peptide injection

The effects of NPF1 peptide injection, RNAi-mediated knockdown of the *Piwi1* transcript, and restoration treatment (*Piwi1* knockdown combined with NPF1 peptide injection) on food intake were analyzed. The concentrations of peptides used were in accordance with previous studies (Hou *et al*, 2017). The commercially synthesized peptide (BGI, peptide, AEAQQADGNKLEGLADALKYLQELDRYYSQVARPRF-NH<sub>2</sub>) was dissolved in ddH<sub>2</sub>O to generate a stock solution (20 mg/ml). A working solution with a concentration of 2 mg/ml was injected into the hemolymph of fourth-instar locusts via the abdomen by using a microinjector (2  $\mu\text{l}$ /locust). Food intake was measured as described above.

### piRNA inhibition or overexpression *in vivo*

piRNA mimics were chemically modified; the effects of single-strand, stable piRNA mimics were similar to the effect of overexpression of endogenous piRNAs. piRNA inhibitors were chemically modified, and single-strand RNA analogs whose sequences were the reverse complements of piRNAs could efficiently and specifically silence endogenous piRNAs (see Appendix Supplementary Text for nonspecific effects of piRNA inhibitors). The locusts were coinjected with 1.65  $\mu\text{g}$  of piRNA mimics or inhibitors (GenePharma) using Entanster™ *in vivo* (Engreen) with a Nanoliter Injector 2000 (World Precision Instruments). Because only a portion of piRNA artifacts is able to be delivered to nuclei, the effective molar mass of piRNA artifacts in the cell nuclei is much lesser than 161 pmol (1.65  $\mu\text{g}$ ). Scrambled piRNA mimics/inhibitors were used as negative controls. Food intake and body weight were measured after 48 h following treatment as described above.

### Subcellular RNA fractionation

Nuclear and cytoplasmic fractions were separated as described previously (Csorba *et al*, 2014). The brains were homogenized in

PBS containing 0.2% Nonidet P-40. The homogenates were centrifuged at 1,500 g for 10 min at 4°C. The nuclei were in the pellet, whereas the cytoplasm remained in the supernatant. The nuclei were resuspended three times in PBS to remove the residual cytoplasm. The cytoplasmic fraction was centrifuged at 12,000 g for 20 min at 4°C to remove the residual nuclei. Nuclei and cytoplasm RNA were extracted as described above. The RNA was pretreated with DNase (Invitrogen). As quality controls for fractionation, U6 was used as a nuclear marker, and tRNA-Gly was used as a cytoplasmic marker.

### RNA-binding protein immunoprecipitation

RNA-binding protein immunoprecipitation experiments were performed using a Magna RIP Quad kit (Millipore) with slight modifications. Briefly, approximately 25 brains were homogenized in ice-cold RIP lysis buffer and stored at –80°C overnight for thorough tissue lysis. A total of 5 µg portion of an anti-hnRNP F/H (Mouse monoclonal antibody, Abcam)/anti-U2AF65 (Mouse monoclonal antibody, Santa Cruz Biotechnology)/anti-IgG antibody (negative control) was incubated with magnetic beads for 30 min. Then, the lysate was thawed and centrifuged, and the supernatant was coincubated with the bead–antibody complex at 4°C overnight. Each lysate was divided into two samples, which were incubated with the anti-hnRNP F/H- (or anti-U2AF65-) and anti-IgG-conjugated beads. RNA was extracted from the immunoprecipitates and input and reverse transcribed into cDNA using a High Capacity RNA-to-cDNA kit (ABI). The quality and specificity of the hnRNP F/H and U2AF65 antibodies were validated before the experiment (Appendix Fig S1G and S9). Using these cDNAs as templates, the level of *pre-NPF1* in the immunoprecipitates was measured as the relative enrichments in the anti-hnRNP F/H or anti-U2AF65 antibody-treated sample compared with the anti-IgG antibody-treated sample. The input and the IgG control-treated samples were used for normalization of the relative expression levels and to ensure the specificity of the RNA-hnRNP F/H or RNA-U2AF65 interaction.

The RIP assay for piRs-3-I2 and piRs-3-I3 was performed as above. Magnetic beads were preincubated with 5 µg of Piwi1/hnRNP F/H or with 5 µg of normal IgG. The RNA in the immunoprecipitates and inputs was extracted by TRIzol (Invitrogen) and reverse transcribed into cDNA as above. Then, qPCR was performed using these cDNAs as templates to analyze the expression levels of piRs-3-I2 and piRs-3-I3. Input samples were used for normalization of the relative expression of piRs-3-I2/piRs-3-I3 and IgG controls were assayed for subtraction of the nonspecific interactions of piRNA-Piwi1 or piRNA-hnRNP F/H.

### Plasmid construction, transfections, and expression assays

The wild or mutated partial sequence of *pre-NPF1* was cloned into the EcoRI and EcoRV sites of pcDNA3.1(+). The presence of both precursor RNAs and mature RNAs was verified by Sanger sequencing to ensure the success of the splicing process. 293T cells cultured in DMEM supplemented with 10% FBS were cotransfected with the plasmid expression vectors along with wild-type or mutated piRs-3-I3 mimics using Lipofectamine 3000 reagent (Invitrogen) according to the manufacturer's instructions. The transfected amounts of piRNA mimics were dosed according to the manufacturer's

instructions of Lipofectamine 3000 reagent ([https://tools.thermofisher.com/content/sfs/manuals/lipofectamine3000\\_scaling.pdf](https://tools.thermofisher.com/content/sfs/manuals/lipofectamine3000_scaling.pdf)). The mRNA levels of *NPF1* and *pre-NPF1* in the cotransfected 293T cells were measured after 36 h of transfection. Briefly, one third of every sample was used for RNA extraction, and the other two third of each sample was separated into nuclear and cytoplasmic fractions as described above. The plasmid was extracted from the nuclear fraction and used as a control. The RNA was pretreated with DNase (Invitrogen). cDNA was reverse transcribed from 1 µg of total RNA using M-MLV reverse transcriptase (Promega).

### In vitro RNase III cleavage assay

*Pre-NPF1* transcripts were generated using the RiboMax express RNAi system (Promega). The RNase III cleavage assay was performed using a ShortCut<sup>®</sup> RNase III Kit (NEB) in accordance with previously described methods (Anacker *et al*, 2018). Briefly, the artificial *pre-NPF1* substrate was heated at 70°C and cooled slowly to room temperature to allow it to adopt a hairpin structure. Substrate RNA (2 µg) and RNase III (0.5 µl) were combined in cleavage buffer and incubated at 37°C for 5 min. The reaction was stopped by adding EDTA. The reaction products were analyzed by qPCR as described above.

## Data availability

The RNA-seq data and small RNA transcriptome data were deposited in the Sequence Read Archive Database of NCBI under BioProject accession number PRJNA605316 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA605316>).

**Expanded View** for this article is available online.

### Acknowledgements

This research was supported by grants from the National Natural Science Foundation of China (32088102, 31672353, and 31702060) and the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDPB16). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Liping Wang and Zizheng Zhang for experimental assistance, and Prof. Zhangwu Zhao for insightful discussion. The computational resources were provided by the Research Network of Computational Biology and the Supercomputing Center at the Beijing Institutes of Life Science, Chinese Academy of Sciences.

### Author contributions

**Huimin Wang:** Data curation; Formal analysis; Validation; Investigation; Methodology; Writing – original draft. **Feng Jiang:** Conceptualization; Formal analysis; Supervision; Funding acquisition; Writing – original draft; Writing – review & editing. **Xiang Liu:** Formal analysis; Investigation. **Qing Liu:** Formal analysis. **Yunyun Fu:** Formal analysis. **Ran Li:** Formal analysis. **Li Hou:** Formal analysis. **Jie Zhang:** Formal analysis. **Jing He:** Formal analysis. **Le Kang:** Conceptualization; Supervision; Funding acquisition; Writing – review & editing.

In addition to the CRediT author contributions listed above, the contributions in detail are FJ and LK conceived the project and designed the study. HW designed and performed the experiments. XL and QL performed the bioinformatics analyses. YF, RL, LH, JZ, and JH contributed to the experiments. FJ, HW, and LK wrote the manuscript. All authors read and approved the final manuscript.

## Disclosure statement and competing interests

The authors declare that they have no conflict of interest.

## References

- Allen ML (2017) Comparison of RNAi sequences in insect-resistant plants to expressed sequences of a beneficial lady beetle: a closer look at off-target considerations. *Insects* 8: 27
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402
- Anacker ML, Drecktrah D, LeCoutre RD, Lybecker M, Samuels DS (2018) RNase III processing of rRNA in the Lyme disease spirochete *Borrelia burgdorferi*. *J Bacteriol* 200: e00035–e118
- Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* 31: 785–799
- Barckmann B, Pierson S, Dufourt J, Papin C, Armenise C, Port F, Grentzinger T, Chambeyron S, Baronian G, Desvignes J-P et al (2015) Aubergine iCLIP reveals piRNA-dependent decay of mRNAs involved in germ cell development in the early embryo. *Cell Rep* 12: 1205–1216
- Berger JH, Charron MJ, Silver DL (2012) Major facilitator superfamily domain-containing protein 2a (MFS2A) has roles in body growth, motor function, and lipid metabolism. *PLoS One* 7: e50629
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120
- Buratti E, Baralle FE (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* 24: 10505–10514
- Burnette JM, Miyamoto-Sato E, Schaub MA, Conklin J, Lopez AJ (2005) Subdivision of large introns in *Drosophila* by recursive splicing at nonexonic elements. *Genetics* 170: 661–674
- Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* 12: 503–514
- Chen Y, Pane A, Schubach T (2007) Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in *Drosophila*. *Curr Biol* 17: 637–642
- Csorba T, Questa JI, Sun Q, Dean C (2014) Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc Natl Acad Sci USA* 111: 16160–16165
- Czech B, Hannon GJ (2016) One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. *Trends Biochem Sci* 41: 324–337
- Ding D, Liu G, Hou L, Gui W, Chen B, Kang L (2018) Genetic variation in *PTPN1* contributes to metabolic adaptation to high-altitude hypoxia in Tibetan migratory locusts. *Nat Commun* 9: 4991
- Donkin I, Versteyhe S, Ingerslev L, Qian K, Mechta M, Nordkap L, Mortensen B, Appel E, Jørgensen N, Kristiansen V et al (2016) Obesity and bariatric surgery drive epigenetic variation of spermatozoa in humans. *Cell Metab* 23: 369–378
- Duff MO, Olson S, Wei X, Garrett SC, Osman A, Bolisetty M, Plocik A, Celniker SE, Graveley BR (2015) Genome-wide identification of zero nucleotide recursive splicing in *Drosophila*. *Nature* 521: 376–379
- Falibene A, Rossler W, Josens R (2012) Serotonin depresses feeding behaviour in ants. *J Insect Physiol* 58: 7–17
- Fiorenza A, Lopez-Atalaya JP, Rovira V, Scandaglia M, Geijo-Barrientos E, Barco A (2016) Blocking miRNA biogenesis in adult forebrain neurons enhances seizure susceptibility, fear memory, and food intake by increasing neuronal responsiveness. *Cereb Cortex* 26: 1619–1633
- Ganko EW, Meyers BC, Vision TJ (2007) Divergence in expression between duplicated genes in *Arabidopsis*. *Mol Biol Evol* 24: 2298–2309
- Gou LT, Kang JY, Dai P, Wang X, Li F, Zhao S, Zhang M, Hua MM, Lu Y, Zhu Y et al (2017) Ubiquitination-deficient mutations in human Piwi cause male infertility by impairing histone-to-protamine exchange during spermiogenesis. *Cell* 169: 1090–1104.e1013
- Guo X, Ma Z, Du B, Li T, Li W, Xu L, He J, Kang L (2018) Dop1 enhances conspecific olfactory attraction by inhibiting miR-9a maturation in locusts. *Nat Commun* 9: 1193
- Hou L, Yang P, Jiang F, Liu Q, Wang X, Kang L (2017) The neuropeptide F/nitric oxide pathway is essential for shaping locomotor plasticity underlying locust phase transition. *Elife* 6: e22526
- Hou L, Li B, Ding D, Kang L, Wang X (2019) CREB-B acts as a key mediator of NPF/NO pathway involved in phase-related locomotor plasticity in locusts. *PLoS Genet* 15: e1008176
- Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 21: 635–637
- Jiang F, Liu Q, Wang Y, Zhang J, Wang H, Song T, Yang M, Wang X, Kang L (2017) Comparative genomic analysis of SET domain family reveals the origin, expansion, and putative function of the arthropod-specific SmydA genes as histone modifiers in insects. *Gigascience* 6: 1–16
- Jiang F, Liu Q, Liu X, Wang X, Kang L (2019a) Genomic data reveal high conservation but divergent evolutionary pattern of Polycomb/Trithorax group genes in arthropods. *Insect Sci* 26: 20–34
- Jiang F, Zhang J, Liu Q, Liu X, Wang H, He J, Kang L (2019b) Long-read direct RNA sequencing by 5'-Cap capturing reveals the impact of Piwi on the widespread exonization of transposable elements in locusts. *RNA Biol* 16: 950–959
- Kang L, Chen X, Zhou Y, Liu B, Zheng W, Li R, Wang J, Yu J (2004) The analysis of large-scale gene expression correlated to the phase changes of the migratory locust. *Proc Natl Acad Sci USA* 101: 17611–17615
- Katoh K, Asimeno G, Toh H (2009) Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol* 537: 39–64
- Kelly S, Georgomanolis T, Zirkel A, Diermeier S, O'Reilly D, Murphy S, Langst G, Cook PR, Papantonis A (2015) Splicing of many human genes involves sites embedded within introns. *Nucleic Acids Res* 43: 4721–4732
- Kim KW, Tang NH, Andrusiak MG, Wu Z, Chisholm AD, Jin Y (2018) A Neuronal piRNA pathway inhibits axon regeneration in *C. elegans*. *Neuron* 97: 511–519
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37: 907–915
- Kiuchi 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
- Kunte N, McGraw E, Bell S, Held D, Avila LA (2020) Prospects, challenges and current status of RNAi through insect feeding. *Pest Manag Sci* 76: 26–41
- Lam DD, de Souza FSJ, Nasif S, Yamashita M, López-Leal R, Otero-Corchon V, Meece K, Sampath H, Mercer AJ, Wardlaw SL et al (2015) Partially redundant enhancers cooperatively maintain mammalian pomc expression above a critical functional threshold. *PLoS Genet* 11: e1004935
- Langmead B (2010) Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* Chapter 11: Unit 11.7
- Lewis SH, Quarles KA, Yang Y, Tanguy M, Frézal L, Smith SA, Sharma PP, Cordaux R, Gilbert C, Giraud I et al (2018) Pan-arthropod analysis reveals



- somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol* 2: 174–181
- Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923–930
- Lin K-Y, Wang W-D, Lin C-H, Rastegari E, Su Y-H, Chang Y-T, Liao Y-F, Chang Y-C, Pi H, Yu B-Y et al (2020) Piwi reduction in the aged niche eliminates germline stem cells via Toll-GSK3 signaling. *Nat Commun* 11: 3147
- Liu W, Zhou Y, Hu Z, Sun T, Denise A, Fu XD, Zhang Y (2010) Regulation of splicing enhancer activities by RNA secondary structures. *FEBS Lett* 584: 4401–4407
- Liu Q, Jiang F, Zhang J, Li X, Kang L (2021) Transcription initiation of distant core promoters in a large-sized genome of an insect. *BMC Biol* 19: 62
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550
- Ma Z, Guo W, Guo X, Wang X, Kang L (2011) Modulation of behavioral phase changes of the migratory locust by the catecholamine metabolic pathway. *Proc Natl Acad Sci USA* 108: 3882–3887
- Mei Y, Wang Y, Kumari P, Shetty AC, Clark D, Gable T, MacKerell AD, Ma MZ, Weber DJ, Yang AJ et al (2015) A piRNA-like small RNA interacts with and modulates p-ERM proteins in human somatic cells. *Nat Commun* 6: 7316
- Muh SJ, Hovhannisyants RH, Carstens RP (2002) A Non-sequence-specific double-stranded RNA structural element regulates splicing of two mutually exclusive exons of fibroblast growth factor receptor 2 (FGFR2). *J Biol Chem* 277: 50143–50154
- Nandi S, Chandramohan D, Fioriti L, Melnick AM, Hebert JM, Mason CE, Rajasethupathy P, Kandel ER (2016) Roles for small noncoding RNAs in silencing of retrotransposons in the mammalian brain. *Proc Natl Acad Sci USA* 113: 12697–12702
- Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S (2013) Transposition-driven genomic heterogeneity in the *Drosophila* brain. *Science* 340: 91–95
- Petrovich GD (2018) Feeding behavior survival circuit: anticipation & competition. *Curr Opin Behav Sci* 24: 137–142
- Pezic D, Manakov SA, Sachidanandam R, Aravin AA (2014) piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev* 28: 1410–1428
- Phay M, Kim HH, Yoo S (2018) Analysis of piRNA-like small non-coding RNAs present in axons of adult sensory neurons. *Mol Neurobiol* 55: 483–494
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5: e9490
- Rajasethupathy P, Antonov I, Sheridan R, Frey S, Sander C, Tuschl T, Kandel ER (2012) A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell* 149: 693–707
- Ross RJ, Weiner MM, Lin H (2014) PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* 505: 353–359
- Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC (2009) A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* 461: 1296–1299
- Sienski G, Donertas D, Brennecke J (2012) Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* 151: 964–980
- Smock LA (1980) Relationships between body size and biomass of aquatic insects. *Freshw Biol* 10: 375–383
- Sun W, Samimi H, Gamez M, Zare H, Frost B (2018) Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. *Nat Neurosci* 21: 1038–1048
- Tan S, Li A, Wang Y, Shi W (2019) Role of the neuropeptide F 1 in regulating the appetite for food in *Locusta migratoria*. *Pest Manag Sci* 75: 1304–1309
- Teixeira FK, Okuniewska M, Malone CD, Coux RX, Rio DC, Lehmann R (2017) piRNA-mediated regulation of transposon alternative splicing in the soma and germ line. *Nature* 552: 268–272
- Teng YW, Ellis JM, Coleman RA, Zeisel SH (2012) Mouse betaine-homocysteine S-methyltransferase deficiency reduces body fat via increasing energy expenditure and impairing lipid synthesis and enhancing glucose oxidation in white adipose tissue. *J Biol Chem* 287: 16187–16198
- Van Wielendaele P, Dillen S, Zels S, Badisco L, Vanden Broeck J (2013) Regulation of feeding by neuropeptide F in the desert locust, *Schistocerca gregaria*. *Insect Biochem Mol Biol* 43: 102–114
- Vaudry H, Dillen S, Zels S, Verlinden H, Spit J, Van Wielendaele P, Vanden Broeck J (2013) Functional characterization of the short neuropeptide F receptor in the desert locust, *Schistocerca gregaria*. *PLoS One* 8: e53604
- Vinnikov IA, Hajdukiewicz K, Reymann J, Beneke J, Czajkowski R, Roth LC, Novak M, Roller A, Dorner N, Starkuviene V et al (2014) Hypothalamic miR-103 protects from hyperphagic obesity in mice. *J Neurosci* 34: 10659–10674
- Vodala S, Pescatore S, Rodriguez J, Buescher M, Chen YW, Weng R, Cohen SM, Rosbash M (2012) The oscillating miRNA 959–964 cluster impacts *Drosophila* feeding time and other circadian outputs. *Cell Metab* 16: 601–612
- Vourekas A, Alexiou P, Vrettos N, Maragkakis M, Mourelatos Z (2016) Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature* 531: 390–394
- Wang D, Zhang Y, Zhang Z, Zhu J, Yu J (2010) KaKs\_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinformatics* 8: 77–80
- Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, Li B, Cui F, Wei J, Ma C et al (2014) The locust genome provides insight into swarm formation and long-distance flight. *Nat Commun* 5: 2957
- Wang Y, Jiang F, Wang H, Song T, Wei Y, Yang M, Zhang J, Kang L (2015) Evidence for the expression of abundant microRNAs in the locust genome. *Sci Rep* 5: 13608
- Wang HL, Chekanova JA (2016) Small RNAs: essential regulators of gene expression and defenses against environmental stresses in plants. *Wiley Interdiscip Rev RNA* 7: 356–381
- Wang L, Sinnott-Armstrong N, Wagschal A, Wark AR, Camporez J-P, Perry RJ, Ji F, Sohn Y, Oh J, Wu SU et al (2020) A microRNA linking human positive selection and metabolic disorders. *Cell* 183: 684–701.e614
- Watanabe T, Lin H (2014) Posttranscriptional regulation of gene expression by Piwi proteins and piRNAs. *Mol Cell* 56: 18–27
- Wei Z, Baggerman G, J. Nachman R, Goldsworthy G, Verhaert P, De Loof A, Schoofs L (2000) Sulfakinins reduce food intake in the desert locust, *Schistocerca gregaria*. *J Insect Physiol* 46: 1259–1265
- Wei Y, Chen S, Yang P, Ma Z, Kang L (2009) Characterization and comparative profiling of the small RNA transcriptomes in two phases of locust. *Genome Biol* 10: R6
- Wei J, Shao W, Cao M, Ge J, Yang P, Chen L, Wang X, Kang L (2019) Phenylacetone nitrile in locusts facilitates an antipredator defense by acting as an olfactory aposematic signal and cyanide precursor. *Sci Adv* 5: eaav5495
- Yang M, Wei Y, Jiang F, Wang Y, Guo X, He J, Kang L (2014) MicroRNA-133 inhibits behavioral aggregation by controlling dopamine synthesis in locusts. *PLoS Genet* 10: e1004206

- Yang M, Wang Y, Liu Q, Liu Z, Jiang F, Wang H, Guo X, Zhang J, Kang L (2019a) A  $\beta$ -carotene-binding protein carrying a red pigment regulates body-color transition between green and black in locusts. *Elife* 8: e41362
- Yang P, Hou L, Wang X, Kang L (2019b) Core transcriptional signatures of phase change in the migratory locust. *Protein Cell* 10: 883–901
- Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16: 284–287
- Zhang Y, Wang X, Kang L (2011) A k-mer scheme to predict piRNAs and characterize locust piRNAs. *Bioinformatics* 27: 771–776
- Zhang Q, Fan X, Wang Y, Sun MA, Shao J, Guo D (2017) BPP: a sequence-based algorithm for branch point prediction. *Bioinformatics* 33: 3166–3172
- Zhao W, Zhu J, Lu H, Zhu J, Jiang F, Wang W, Luo L, Kang L, Cui F (2021) The nucleocapsid protein of rice stripe virus in cell nuclei of vector insect regulates viral replication. *Protein Cell* 6: 1–19
- Zhou F, Kang L, Wang XH (2019) JumpDetector: an automated monitoring equipment for the locomotion of jumping insects. *Insect Sci* 27: 613–624



**License:** This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.