Received: 4 February 2018 | Revised: 15 May 2018 | Accepted: 20 June 2018

DOI: 10.1002/jez.b.22815

RESEARCH ARTICLE

Diversity of piRNA expression patterns during the ontogeny of the German cockroach

Natalia Llonga[1]* | Guillem Ylla[1,3]* | Josep Bau[2] | Xavier Belles[1] | Maria-Dolors Piulachs[1]

[1]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain

[2]Department of Biosciences, University of Vic – Central University of Catalonia, Vic, Barcelona, Spain

[3]Department of Microbiology and Cell Science, Institute for Food and Agricultural Sciences, Genetics Institute, University of Florida, Gainesville, Florida

Correspondence

Maria-Dolors Piulachs, Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Passeig Maritim de la Barceloneta, 37, Barcelona 08003, Spain.

Email: mdolors.piulachs@ibe.upf-csic.es

Funding information

Spanish Ministry of Economy and Competitiveness, Grant/Award Numbers: CGL2015-64727-P, CGL2012-36251, CGL2016-76011-R; Departament d'Innovació, Universitats i Empresa, Generalitat de Catalunya, Grant/Award Number: 2017 SGR 1030; Ministerio de Ciencia e Innovación, Grant/Award Number: BFU2011-22404; European Fund for Economic and Regional Development (FEDER); Agencia Estatal de Investigación (AEI)

*Both authors contributed equally to this work.

Abstract

The Piwi-interacting RNA (piRNA) system is an evolutionarily conserved mechanism involved in the control of transposable elements and maintenance of genomic stability, especially in germ line cells and in early embryo stages. However, relevant particularities, both in mechanism and function, exist across species among metazoans and even within the insect class. As a member of the scarcely studied hemimetabolan group, *Blattella germanica* can be a suitable reference model to study insect evolution. We present the results of a stringent process of identification and study of expressed piRNAs for *B. germanica* across 11 developmental stages, ranging from unfertilized egg to nymphs and adult female. Our results confirm the dual origin of piRNA in this species, with a majority of them being generated from the primary pathway, and a smaller but highly expressed set of sequences participating in the secondary ("ping-pong") reamplification pathway. An intriguing partial complementarity in expression is observed between the piRNA of the two biogenesis pathways, with those generated in the secondary pathway being quite restricted to early embryo stages. In addition, many piRNAs are exclusively expressed in late embryo and nymphal stages. These observations point at piRNA functions beyond the role of transposon control in early embryogenesis. Our work supports the view of a more complex scenario, with different sets of piRNAs acting in different times and having a range of functions wider than previously thought.

KEYWORDS

Blattella germanica, embryo development, hemimetabola, maternal–zygotic transition, small noncoding RNA

1 | INTRODUCTION

Piwi-interacting RNAs (piRNAs) are small noncoding RNAs (sncRNAs) of 26 to 31 nucleotide (nt) in length (Aravin et al., 2006) that bind to the Piwi clade of Argonaute (Ago) protein. In *Drosophila melanogaster*, these proteins comprise Piwi, Aubergine (Aub), and Ago3 (Ghildiyal & Zamore, 2009).

piRNAs are involved in the control of transposable elements (TEs), as part of an evolutionarily conserved mechanism that defends the genome from viral or parasitic threats and preserves genomic stability (Aravin, Hannon, & Brennecke, 2007; Brennecke et al., 2007; Ernst, Odom, & Kutter, 2017).

Most piRNAs originate from loci known as piRNA clusters. These loci are enriched in inactive transposon sequences and are necessary to prevent active TEs from spreading throughout the genome. Studies on flamenco, a major piRNA producing locus in *D. melanogaster*, has revealed that most of TE sequences in such loci are present as a unique copy. This has led to think that these occasionally called "transposon graveyards" may be in fact "transposon traps" that enable the cell to arrest the propagation of newly acquired mobile elements in a way similar to the bacterial CRISPR system, which provides adaptive immunity against invasive threats (Goriaux, Théron, Brasset, & Vaury, 2014).

Unlike other types of sncRNAs like miRNAs, the sequence of transcribed piRNAs is poorly conserved even in closely related species (Aravin et al., 2007). Thus millions of unique piRNA sequences have been reported, which only share a bias toward a uracile at 5′ end (1U; Brennecke et al., 2007). piRNA length comprises between 26 to 31 nt, which is longer than that of miRNAs and siRNAs (21 to 23 nt) being used, thus as one of the general piRNA distinctive features. The diversity of piRNA sequences is a consequence of their biogenesis pathways that, albeit sharing some common traits with other sncRNAs, show many peculiarities (Huang, Tóth, & Aravin, 2017).

piRNAs are transcribed as long RNA transcripts, generally produced from piRNA clusters with a strand bias that favors antisense sequences

 © 2018 Wiley Periodicals, Inc.

from transposons mapped within the cluster. The piRNA transcription does not seem to occur through canonical promoters. It has been proposed that the Rhino-Deadlock-Cutoff (RDC) complex would participate in regulating the transcription of piRNA clusters in germ cells, inhibiting the transcription of adjacent genes in the cluster or allowing the functionality of noncanonical promoters (Gleason, Anand, Kai, & Chen, 2018). Moreover, the localization of some piRNA clusters in heterochromatic regions led to the study of them in relation to histone modifications, which has revealed an important role of H3K9me3 on cluster transcription (Mohn, Sienski, Handler, & Brennecke, 2014; Molla-Herman, Vallés, Ganem-Elbaz, Antoniewski, & Huynh, 2015). In somatic cells, the mechanisms regulating piRNA transcription are less known, but it has been suggested that the control exerted by the RDC complex would not be necessary in these cells (see Gleason et al., 2018).

Maternally inherited piRNAs seem to play a part in defining the piRNA production loci and there is evidence in *D. melanogaster* that these maternal loadings are essential to target paternally transmitted transposons, which would protect the genome in early embryogenesis (Czech & Hannon, 2016). After being exported to the cytoplasmic perinuclear nuage region (Le Thomas, Tóth, & Aravin, 2014), piRNA precursors are processed into mature piRNAs by means of the protein Zucchini, thus diverging from miRNA and siRNA enzymatic processes that use Dicer and Drosha machinery. Mature piRNAs are eventually loaded into Piwi proteins that will act directly upon active cytoplasmic transposon transcripts or repress their transcription at a nuclear level (Huang et al., 2017).

Besides the primary pathway, the alternative secondary ("ping-pong") pathway is responsible for the generation of piRNAs that are loaded into Aub and Ago3 in *D. melanogaster*. The secondary pathway acts as an amplification loop that, upon RNA recognition from an expressed TE by the antisense sequence of a primary piRNA, triggers the swift generation of more piRNAs from the cluster transcript, which eventually target active transposons (Huang et al., 2017; Luteijn & Ketting, 2013). A specific signature of the piRNAs generated through the secondary pathway is the production of piRNA pairs showing sequence complementarity along their 10 first base pairs, plus a tendency to have an adenine in the 10th position (10A) in the sense strand. The pairing of an antisense piRNA-Aub complex with the Ago3-loaded sense target and the subsequent Ago3 slicer activity explain these particular features, being the 1U bias of the antisense piRNA what accounts for the corresponding 10A bias in its sense piRNA counterpart (Luteijn & Ketting, 2013).

In *D. melanogaster*, the secondary pathway is restricted to germ line cells (Goriaux et al., 2014), thus piRNAs and Piwi protein are predominantly expressed in gonads. For this reason, the Piwi system has been mainly studied in such tissues and during early stages of the embryo development (Czech & Hannon, 2016; Girard, Sachidanandam, Hannon, & Carmell, 2006; Huang et al., 2017). However, more ubiquitous expression patterns, including somatic tissues, have also been reported (Lewis et al., 2018; Ninova, Griffiths-Jones, & Ronshaugen, 2017; Ross, Weiner, & Lin, 2014; Yan et al., 2011). Somatic piRNA expression has been poorly studied, although recent studies point at an ancestral occurrence in insect evolution and it may also account for specific adaptations of individual arthropod species (Lewis et al., 2018).

Current knowledge on piRNA expression and function in arthropods suffers a large bias toward studies on holometabolan species, especially on *D. melanogaster*. However, the data from other species, although scarce, suggests that relevant differences exist across different insect groups, though the general piRNA mechanisms are conserved.

Studies on disease vectors, like *Aedes aegypti* and *Anopheles gambiae*, indicate that there are piRNA divergences from *D. melanogaster* (Arensburger, Hice, Wright, Craig, & Atkinson, 2011). Moreover, in the hymenopteran *Apis mellifera*, the piRNA system is essential to ensure genomic stability to the vulnerable haploid male genome and reproductive castes despite of having one of the lowest TEs contents in the animal kingdom (Wang, Ashby, Ying, Maleszka, & Forêt, 2017). Important contributions have been recently reported for the red flour beetle, *Tribolium castaneum*, where maternally deposited and zygotically expressed piRNAs have been characterized (Ninova et al., 2017). This indicates that the piRNA pathway is not restricted to the germ line, at least in this species, thus reinforcing the idea that there is a certain diversity of piRNA pathways and functions in different insect lineages. Seemingly, new and surprising relationships in piRNA function and evolution are being discovered, as in the case of the lepidopteran, *Bombyx mori*, where the primary sex-determination mechanism has been attributed to the gene-silencing activity of a piRNA (Kiuchi et al., 2014).

Although some piRNA data of hemimetabolan species have been reported (Chen et al., 2012; Wei, Chen, Yang, Ma, & Kang, 2009; Zhang, Wang, & Kang, 2011), in-depth studies concerning these insects are still lacking. The present work focuses on the piRNA expression in the German cockroach, *Blattella germanica*, along development, covering all key stages from the unfertilized egg (NFE) to the adult. It is, therefore, a first attempt to unveil the piRNA expression patterns along the ontogeny of a hemimetabolan insect.

Interestingly, the *B. germanica* genome has a high amount of repetitive DNA, which can reach more than 50% of the genome in Blattodean species (Harrison et al., 2018). This suggests that cockroaches can be suitable models for studying piRNAs, which could help to unveil new roles for them, and provide insights into genome evolution and the functional relationships between TE and genome size. Previous studies already highlight the importance of sncRNAs, especially in developmental transitions between embryonic stages of development in *B. germanica*, where miRNAs play a relevant role (Ylla, Piulachs, & Belles, 2017). Now, our current results reveal striking differences in piRNA expression that suggest specific roles for piRNA, not only in the early embryo but also in postembryonic development.

2 | MATERIALS AND METHODS

2.1 | Insects and small RNA libraries

Small RNA libraries were obtained from a *B. germanica* colony reared in the dark at 29 ± 1°C. Two replicates of 11 different developmental stages were prepared using the NEBNext® small RNA libraries

kit (New England Biolabs, Ipswich, MA) and later sequenced with NextSeq® platform (Illumina, San Diego, CA). Detailed procedures for the generation of theses libraries are described elsewhere (Ylla et al., 2017).

The studied stages included: NFE, 8, 24, 48, 144, and 312 hr after oviposition (ED0, ED1, ED2, ED6, and ED13); first, third, fifth, and sixth (last) nymphal instars (N1, N3, N5, and N6); and adult female (Adult). The timing of late embryonic and nymphal stages was selected to match the period of ecdysone production (Supporting Information Figure S1).

RNA extraction was performed on isolated eggs and embryos in the case of NFE and embryo samples, while the whole insect body was used in nymph and adult stages.

Small RNA-seq data are publicly available at GEO GSE87031 (Ylla et al., 2017).

2.2 | piRNA identification

Small RNA libraries were preprocessed to remove adapters and low quality reads and merge paired reads, as described in Ylla et al. (2017), and aligned to the *B. germanica* genome assembly (Harrison et al., 2018) with Bowtie2 (Langmead & Salzberg, 2012). The read length distribution was calculated for the mapping reads spanning between 19 and 31 nt, which covers the length of miRNAs (20–25 nt), siRNAs (~22 nt), and piRNAs (26–31 nt). The mapped reads ranging between 26 and 31 nt length, corresponding to the piRNA fraction was retrieved with cutadapt (Martin, 2011), and merged from the 22 small RNA libraries to obtain 36,600,717 reads, which accounted for the 20.6% of the total number of reads. After the removal of a small fraction of reads mapping to known miRNA sites, the number of reads was reduced to 36,591,344, which were collapsed in 6,389,510 unique sequences and mapped again to the *B. germanica* genome using Bowtie (Langmead, Trapnell, Pop, & Salzberg, 2009) reporting all best positions, and allowing for a single mismatch along the read. Finally, 4,396,518 sequences were retrieved, reporting 158,077,447 alignments. To exclude short fragments from RNA degradation, overlapping alignments were collapsed using GenomicRanges (Lawrence et al., 2013). Only those ranges between 26 and 31 nt length, corresponding to the piRNA fraction, were considered. Finally, the sequences aligning to potentially contaminated regions in the genome, and low complexity sequences (Harrison et al., 2018; Rosenkranz & Zischler, 2012), were excluded (Supporting Information Figure S2).

In order to identify piRNA pairs generated in the secondary ("ping-pong") pathway, all alignments showing a 10-nt overlapping sequence pair mapping to a given genomic locus with opposite orientation, were selected using GenomicRanges (Lawrence et al., 2013). Finally, those piRNAs participating in the secondary pathway that were also found in loci where no overlap with the partner piRNA was present (indicating possible primary production site), were considered as primary produced piRNAs entering the secondary pathway (piRNA-PS). On the contrary, those piRNAs that were only found in overlapping loci were considered to be exclusively generated from expressed transposon sequences through the secondary pathway (piRNA-exS). The remaining loci were considered piRNA production sites of exclusively primary piRNAs (piRNA-exP).

Clusters of piRNAs were identified using proTRAC 2.4.0 (Rosenkranz & Zischler, 2012), with a sliding window of 5-kb with 1-kb increments. The proportion of piRNA reads within clusters was calculated by counting the number of reads mapping piRNAs within a given cluster.

The localization of piRNA-exP and piRNA-exS in genic regions was carried out by discriminating those piRNAs that mapped to gene CDS (sense and antisense), intronic regions, as well as TEs.

Identified *B. germanica* piRNAs were compared with the annotated piRNA from *D. melanogaster* and *B. mori* in piRBase (Zhang, Wang & Kang, 2014), with *Chilo supressalis* piRNA from InsectBase (Yin et al., 2016) and *T. castaneum* piRNA from Ninova et al. (2017). The blastn against piRNA were performed with a 100% identity in the alignments.

2.3 | piRNA and cluster expression

The expression of piRNAs from the secondary pathway was quantified by counting the number of mapped reads to a piRNA-exS loci, by means of the R package FeatureCounts (Liao, Smyth, & Shi, 2014), and normalized to reads per million mapped reads. piRNAs with a low variance across the developmental stages (var < 0.05) were excluded.

The number of reads mapping to a given cluster were counted, and the cluster expression was normalized to reads per kb of cluster length per million mapped reads. The heatmaps obtained in both cases were plotted after normalizing by rows.

3 | RESULTS

3.1 | Identification of *B. germanica* piRNAs

The general length distribution of the *B. germanica* small RNAs is bimodal (Figure 1a). Reads between 20 and 25 nt in length, with a sharp peak at 22 nt, basically correspond to miRNA sequences (Ylla, Fromm, Piulachs, & Belles, 2016), whereas those between 26 and 31 nt that peak at 28 to 29 nt, correspond to the typical piRNA length in insects. Although miRNAs and piRNAs are the most abundant molecules in all our small RNA datasets, their relative proportions vary across the different developmental stages. Taken as a whole, the piRNA group is generally larger than that of miRNAs, especially in the early stages of embryogenesis (from ED0 to ED6), where reads corresponding to piRNAs are significantly more abundant than miRNAs. This trend, however, is somewhat reversed toward the end of the embryogenesis (ED13) and especially at the first nymphal instar (N1), where piRNAs are less abundant with respect to miRNAs. Subsequently, the proportion between piRNAs and miRNA remains similar in nymphal stages and in adult females. In correspondence with these observations, the abundance of the piRNA group (26–31 nt) is clearly most abundant in the early embryogenesis, decreasing almost by half in postembryonic stages (Figure 1b).

A total of 4,396,518 out of the initial 6,389,510 unique 26 to 31 nt sequences, could be aligned to the *B. germanica* genome. All alignments

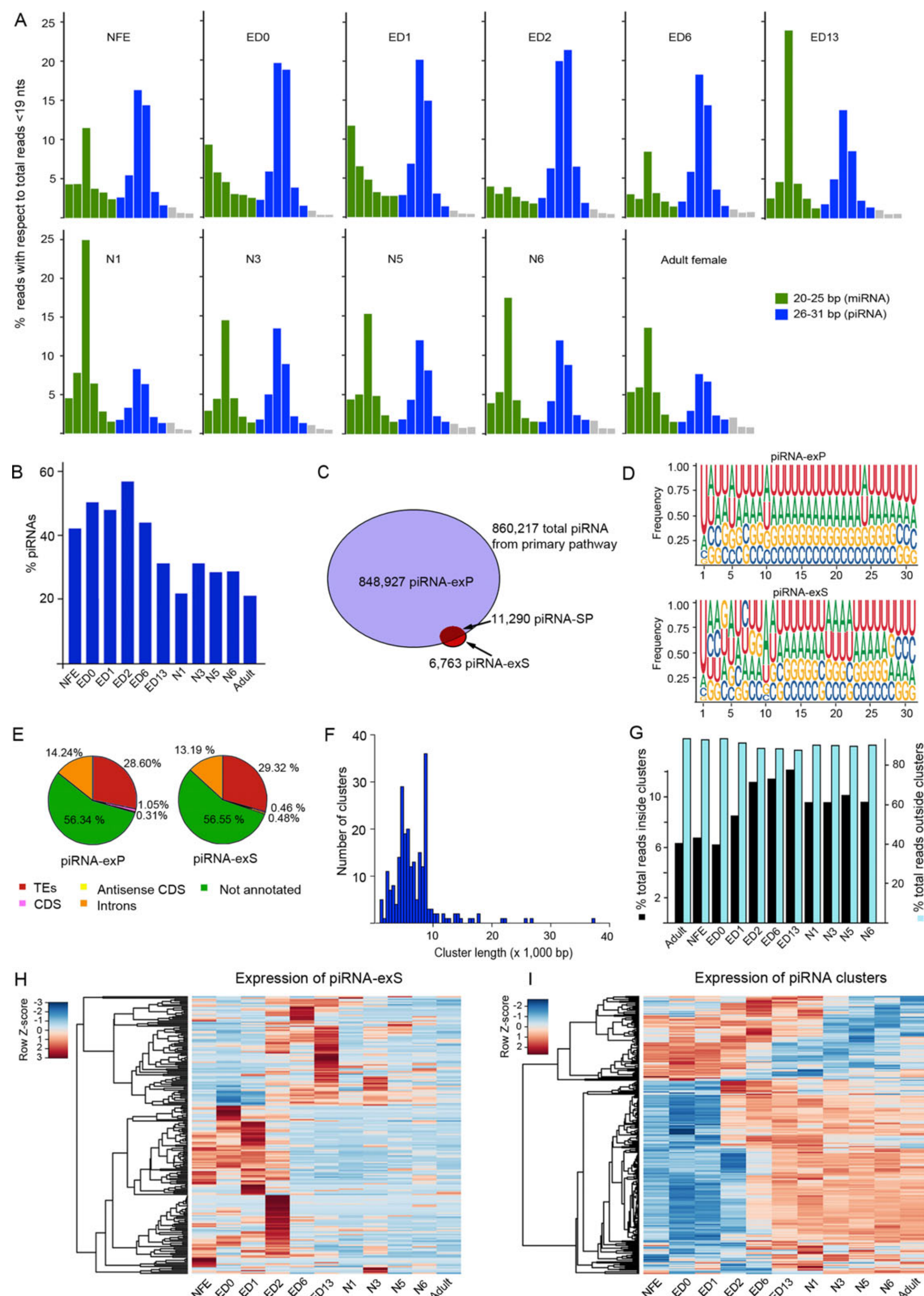

FIGURE 1 Piwi-interacting RNAs (piRNAs) of *Blattella germanica*. (a) Length distribution of reads in each stage-library. Only reads longer than 19 nucleotides were considered. (b) Relative abundance of the piRNA fraction (reads between 26 and 31 nucleotides) in the stage-libraries. (c) Venn diagram showing the number of piRNA generated in each pathway and the primary piRNAs that participate in the secondary pathway. (d) Sequence logos showing the nucleotide frequency in each position of the exclusively primary piRNA (piRNA-exP) and for the secondary pathway piRNA (piRNA-exS). (e) Number of piRNA clusters in relation to their length; minimum cluster length was set to 1 kb. (f) Distribution of piRNA reads within and outside the clusters (percentage of total piRNA reads) in each stage-library; in this panel, the adult female library is displayed beside the unfertilized egg in order to highlight the correlation between these two data sets. (g) Pie chart showing the location of primary (upper) and secondary (lower) piRNA loci in the *B. germanica* genome. (h) Heatmap representing the differential expression in each stage-library of the 2,661 piRNA-exS; piRNA with low variance (var < 0.05) across development stages were excluded for clarity. (i) Heatmap representing the differential expression in each stage-library of the 239 piRNA clusters identified

TABLE 1 Conserved Piwi-interacting RNA (piRNA) among insect species. The number of piRNA identified in each species is indicated below the species name. Only these piRNA with 100% identity were considered

	Blattella germanica (866,980)	*Tribolium castaneum* (7,179,768)	*Bombyx mori* (1,174,963)	*Drosophila melanogaster* (21,027,419)
B. germanica (866,980)	-	77	47	75
T. castaneum (7,179,768)		-	843	12,922
B. mori (1,174,963)			-	462
D. melanogaster (21,027,419)				-

mapping to potentially contaminated regions (Harrison et al., 2018) and those that can result from RNA degradation were excluded, thus leaving 866,980 identified *B. germanica* unique piRNAs (see Supporting Information Table S1). To identify the piRNA participating in the secondary pathway, the piRNAs were mapped again to the *B. germanica* genome, selecting those that mapped in the same loci displaying the characteristic 10 nt overlap between the 5′ end of the sense and antisense piRNA partners. A total of 24,887 loci accounting for 18,053 different piRNAs were found, most of them being 28–29 nt in length. No overlapping piRNA partners were found for the remaining 848,927 sequences, thus we considered them to be exclusively related to the primary pathway (piRNA-exP; Figure 1c). A total of 11,290 out of the secondary pathway piRNAs were also found isolated (with no overlapping partner) in other genomic loci, which probably corresponds to the production site of primary piRNAs that enter in the secondary pathway (Figure 1c). We refer to these sequences as piRNA-PS since they participate in both, primary and secondary pathways. The remaining 6,763 piRNA were always found in conjunction with an overlapping piRNA partner and are therefore considered to be generated exclusively in the secondary pathway from active TEs (piRNA-exS). In support of this interpretation, these piRNA-exS show a clear bias for a U in the first 5′ position (70%) and for an A in the 10A (65%), as expected (Figure 1d).

Comparing the 866,980 piRNA of *B. germanica*, with the publicly available piRNA annotated in other insect species, we found that only 75 piRNA are conserved in *D. melanogaster*, 47 in *B. mori*, and 77 in *T. castaneum* (Table 1). Only five of these conserved piRNA are present in *B. germanica*, *D. melanogaster*, and *B. mori*, while *T. castaneum* only shares four piRNA with these three species. However, when comparing the piRNA within holometabolan species, *D. melanogaster* has 12,922 piRNA identical to *T. castaneum* and 462 identical to *B. mori*, and, at the same time, *B. mori* has 843 identical piRNA to *T. castaneum* (Table 1), indicating a higher degree of conservation between them.

3.2 | Genome distribution of piRNAs and expression during ontogeny

Genome location of primary and secondary piRNAs (Figure 1e) revealed that around 30% of the piRNAs (29.3% of piRNA-exS and 28.6% of piRNA-exP) are found neighboring TEs. Less than 1% of the piRNAs (both piRNA-exP and piRNA–exS) locate in known gene CDs, whereas 13% and 14% locate in introns (Figure 1e).

We identified 239 piRNA clusters in the genome of *B. germanica*, with lengths ranging from 1 to 38 kb, most of them between 1 and 10 kb (Figure 1f). Most clusters show unidirectional expression, while only 22 clusters are bidirectional. Indeed, a total of 96,229 piRNAs (around 8.33% of all piRNA found in *B. germanica*) are grouped in clusters, representing a 39.7% of total piRNAs reads. Only 45 loci inside the clusters showed piRNAs with and overlapping piRNA partner, suggesting that piRNA-exS are mostly generated outside the clusters. Furthermore, and as it would be expected from secondary pathway biogenesis, the expression of out-of-cluster piRNAs is higher in adult females, in NFE, and in the zygote (ED0 and ED1), while it is comparatively lower in embryo, and in nymphal stages (Figure 1g).

Concerning expression during ontogeny, we can observe a high expression of the piRNA-exS during embryo development while only piRNA-exS groups are expressed in certain nymphal stages (Figure 1h). Conversely, considering the expression of piRNA clusters as a representation of the primary piRNAs, we can observe that most of these are highly expressed in postembryonic stages (Figure 1i). A small group of clusters is highly and specifically expressed in NFE and in early embryo stages. However, a surprisingly high number of clusters are highly expressed in late embryogenesis and during postembryonic stages, whereas they are virtually absent in NFE and in the zygote.

4 | DISCUSSION

The evolution of RNA interference (RNAi) pathways that drive the production of siRNAs, miRNAs, and piRNAs in insects is a gradual and complex process (Belles, Cristino, Tanaka, Rubio, & Piulachs, 2012). The RNAi system of the insect last common ancestor diversified and expanded through time across the different insect lineages (Dowling et al., 2017). Although the canonical functions of piRNAs are associated with genome protection from transposon activity during embryogenesis, evidences of additional roles (like regulation of gene expression) are growing (Gebert, Ketting, Zischler, & Rosenkranz, 2015; Gleason et al., 2018; Le Thomas et al., 2013; Lewis et al., 2018; Peng & Lin, 2013; Pritykin, Brito, Schupbach, Singh, & Pane, 2017; Sarkar, Volff, & Vaury, 2017). As the taxonomic spectrum of studied species broadens, the complexity of the piRNA system appears to expand. In insects, data available is still scarce and refers mostly to holometabolan species (see Lewis et al., 2018, and the bibliography included). Our work on *B. germanica* fills an important gap as it provides information on piRNAs of a hemimetabolan species, and expands the usual study range of embryogenesis to encompass postembryonic development as well.

A major concern during the study has been to avoid false positives and consider only an extremely reliable set of expressed piRNAs. With this aim in mind, we applied a higher-than-usual degree of

stringency that, consequently, returned a relatively low number of *B. germanica* piRNAs (866,980), which is a really conservative figure when compared to the 12 to 21 million candidates reported in other species like *Aedes albopictus*, *A. aegypti*, and *D. melanogaster* (Arensburger et al., 2011; Brennecke et al., 2007; Liu et al., 2016). The number of reported piRNAs is strikingly different in different species, and, furthermore, the number of piRNA shared between them is extremely low. This indicates different RNAs from which they originate and suggests a low degree of structural constraints for functioning.

The dual origin of the expressed piRNAs has also been confirmed in *B. germanica* through the identification of 6,763 piRNA exclusively generated in the secondary pathway. These hereby-called piRNA-exS are produced in 24,887 loci in the genome, frequently in regions neighboring TEs as well in gene introns, which would be consistent with the active transposon origin of secondary piRNAs. Similar results have been found for piRNAs produced through the primary pathway (piRNA-exP).

It is well known that piRNAs are grouped in genomic clusters mainly localized in heterochromatic regions (Brennecke et al., 2007), although the data provided for cluster size and the number of piRNAs contained in them is highly divergent in different species. This is mainly due to the different methodologies and parameters used to identify insect piRNAs (Brennecke et al., 2007; Liu et al., 2016; Malone et al., 2009; Ninova et al., 2017). In our work with *B. germanica* we followed severely stringent conditions in order to avoid false positives, which led to the identification of 239 clusters that contain 39.7% of total piRNA reads. In this context, these figures can be hardly compared with equivalent data obtained in other insect species under less astringent conditions.

As occurs in other species, the analysis of *B. germanica* clusters shows a unidirectional expression. The piRNA-exS were preferentially located outside clusters, as only 45 loci for these piRNAs were mapped in clusters. This may be a consequence of the high number of active transposon-derived piRNAs generated through the secondary pathway, which seems to be active mainly during early embryogenesis. The highly stringent approach used in the identification of piRNA candidates in *B. germanica* has probably enhanced the discrimination of this previously unreported effect.

Expression analyses showed that a group of piRNA-exS is expressed from NFE to early embryo development (ED2). Additionally, several groups of piRNA-exS are sharply expressed in temporally specific windows in different stages of development (Figure 1h). In NFE, the piRNA-exS come from maternal loading, whereas those expressed in ED0 and ED1 would operate in the maternal to zygotic transition. Both, ED0 and ED1 stages, show specific piRNAs and we cannot rule out the possible influence of transposons from a paternal origin in these cases. There is again a new batch of piRNA-exS expressed in ED2, when the transition, zygote–embryo, occurs. It is noteworthy that this transition coincides with the time at which maternal piRNA degradation and zygotic transcriptional activation occurs (Ninova et al., 2017). Later in embryogenesis and in the postembryonic development, the importance of the secondary pathway seems less relevant as only small groups of piRNA-exS are expressed, again looking specific of given developmental stages. A high expression of the secondary pathway piRNAs in early stages of embryo development has been also reported in several holometabolan species (Kawaoka et al., 2011; Liu et al., 2016; Ninova et al., 2017), which suggests that the protective role over germinal cells of the secondary pathway piRNAs is evolutionary conserved.

It seems that the expression of the piRNAs from both pathways in *B. germanica* is somewhat complementary. Some piRNAs groups seem crucial in particular developmental stages, particularly coinciding with key transitions (i.e., before and after egg fecundation, in the maternal to zygotic transition and then in the embryo). The piRNA clusters expressed in the first stages of embryo development are few in number, some of them are from a maternal origin, and they are also expressed during early embryogenesis in a stage-specific manner. Clusters expressed in these early stages would again be most probably related with the canonical piRNA function, which is to protect germinal cells from TEs. However, a remarkable finding is the high expression of certain piRNA clusters exclusively during late embryogenesis and postembryonic development. This unexpected result suggests specific piRNA functions in somatic cells, possibly unrelated with the gonads. They might be associated with the formation of new tissues, related to the molting processes, when circulating ecdysone levels are high, which coincides with the situation in most of the stage-libraries used that were staged in periods of high ecdysone production. Thus, it seems plausible that piRNAs would be needed to protect from TEs the stem cells that are giving rise to new tissues during molting. Our study opens a number of intriguing questions that should be addressed in future works, notably, whether there are piRNA functions beyond protection from transposon activity, in embryonic and postembryonic development.

ACKNOWLEDGMENTS

Thanks are due to Elena Navas for technical assistance during the preparation of the small RNA libraries. Genomic analyses were performed using the *B. germanica* genome available at https://www.hgsc.bcm.edu/arthropods/german-cockroach-genomeproject, as provided by the Baylor College of Medicine Human Genome Sequencing Center.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

ORCID

Xavier Belles http://orcid.org/0000-0002-1566-303X
Maria-Dolors Piulachs http://orcid.org/0000-0002-4500-2718

REFERENCES

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., ... Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, *442*, 203–207.

Aravin, A., Hannon, G. J., & Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science*, *318*, 761–764.

Arensburger, P., Hice, R. H., Wright, J. A., Craig, N. L., & Atkinson, P. W. (2011). The mosquito *Aedes aegypti* has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. *BMC Genomics*, *12*, 606.

Belles, X., Cristino, A. S., Tanaka, E. D., Rubio, M., & Piulachs, M.-D. (2012). Insect microRNAs: From molecular mechanisms to biological roles. In L. I. Gilbert (Ed.), *Insect molecular biology and biochemistry* (pp. 30–56). Amsterdam, the Netherlands: Elsevier.

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

Chen, Q., Lu, L., Hua, H., Zhou, F., Lu, L., & Lin, Y. (2012). Characterization and comparative analysis of small RNAs in three small RNA libraries of the brown planthopper (*Nilaparvata lugens*). *PloS One*, *7*, e32860.

Czech, B., & Hannon, G. J. (2016). One loop to rule them all: The ping-pong cycle and piRNA-guided silencing. *Trends in Biochemical Sciences*, *41*, 324–337.

Dowling, D., Pauli, T., Donath, A., Meusemann, K., Podsiadlowski, L., Petersen, M., ... Niehuis, O. (2017). Phylogenetic origin and diversification of RNAi pathway genes in insects. *Genome Biology and Evolution*, *8*, 3784–3793.

Ernst, C., Odom, D. T., & Kutter, C. (2017). The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity. *Nature Communications*, *8*, 1411.

Gebert, D., Ketting, R. F., Zischler, H., & Rosenkranz, D. (2015). piRNAs from pig testis provide evidence for a conserved role of the Piwi pathway in post-transcriptional gene regulation in mammals. *Plos One*, *10*, e0124860.

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews Genetics*, *10*, 94–108.

Girard, A., Sachidanandam, R., Hannon, G. J., & Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, *442*, 199–202.

Gleason, R. J., Anand, A., Kai, T., & Chen, X. (2018). Protecting and diversifying the germline. *Genetics*, *208*, 435–471.

Goriaux, C., Théron, E., Brasset, E., & Vaury, C. (2014). History of the discovery of a master locus producing piRNAs: The flamenco/COM locus in *Drosophila melanogaster*. *Frontiers in Genetics*, *5*, 257.

Harrison, M. C., Jongepier, E., Robertson, H. M., Arning, N., Bitard-Feildel, T., Chao, H., … Bornberg-Bauer, E. (2018). Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nature Ecology & Evolution*, *2*, 557–566.

Huang, X., Tóth, F. K., & Aravin, A. (2017). piRNA biogenesis in *Drosophila melanogaster*. *Trends in Genetics*, *33*, 882–894.

Kawaoka, S., Arai, Y., Kadota, K., Suzuki, Y., Hara, K., Sugano, S., ... Katsuma, S. (2011). Zygotic amplification of secondary piRNAs during silkworm embryogenesis. *RNA*, *17*, 1401–1407.

Kiuchi, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Arai, Y., ... Katsuma, S. (2014). A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature*, *509*, 633–636.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359.

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*, R25.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., ... Carey, V. J. (2013). Software for computing and annotating genomic ranges. *Plos Computational Biology*, *9*, e1003118.

Le Thomas, A., Rogers, A. K., Webster, A., Marinov, G. K., Liao, S. E., Perkins, E. M., ... Tóth, K. F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes & Development*, *27*, 390–399.

Le Thomas, A., Tóth, K. F., & Aravin, A. A. (2014). To be or not to be a piRNA: Genomic origin and processing of piRNAs. *Genome Biology*, *15*, 204.

Lewis, S. H., Quarles, K. A., Yang, Y., Tanguy, M., Frézal, L., Smith, S. A., ... Jiggins, F. M. (2018). Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nature Ecology & Evolution*, *2*, 174–181.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930.

Liu, P., Dong, Y., Gu, J., Puthiyakunnon, S., Wu, Y., & Chen, X.-G. (2016). Developmental piRNA profiles of the invasive vector mosquito *Aedes albopictus*. *Parasites & Vectors*, *9*, 524.

Luteijn, M. J., & Ketting, R. F. (2013). PIWI-interacting RNAs: From generation to transgenerational epigenetics. *Nature Reviews Genetics*, *14*, 523–534.

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

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*, 10.

Mohn, F., Sienski, G., Handler, D., & Brennecke, J. (2014). The Rhino-Deadlock-Cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell*, *157*, 1364–1379.

Molla-Herman, A., Vallés, A. M., Ganem-Elbaz, C., Antoniewski, C., & Huynh, J.-R. (2015). tRNA processing defects induce replication stress and Chk2-dependent disruption of piRNA transcription. *EMBO J*, *34*, 3009–3027.

Ninova, M., Griffiths-Jones, S., & Ronshaugen, M. (2017). Abundant expression of somatic transposon-derived piRNAs throughout *Tribolium castaneum* embryogenesis. *Genome Biology*, *18*, 184.

Peng, J. C., & Lin, H. (2013). Beyond transposons: The epigenetic and somatic functions of the Piwi-piRNA mechanism. *Current Opinion in Cell Biology*, *25*, 190–194.

Pritykin, Y., Brito, T., Schupbach, T., Singh, M., & Pane, A. (2017). Integrative analysis unveils new functions for the Drosophila Cutoff protein in noncoding RNA biogenesis and gene regulation. *RNA*, *23*, 1097–1109.

Rosenkranz, D., & Zischler, H. (2012). proTRAC: A software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics*, *13*, 5.

Ross, R. J., Weiner, M. M., & Lin, H. (2014). PIWI proteins and PIWI-interacting RNAs in the soma. *Nature*, *505*, 353–359.

Sarkar, A., Volff, J.-N., & Vaury, C. (2017). piRNAs and their diverse roles: A transposable element-driven tactic for gene regulation? *FASEB Journal*, *31*, 436–446.

Wang, W., Ashby, R., Ying, H., Maleszka, R., & Forêt, S. (2017). Contrasting sex- and caste-dependent piRNA profiles in the transposon depleted haplodiploid honeybee *Apis mellifera*. *Genome Biology & Evolution*, *9*, 1341–1356.

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 Biology*, *10*, R6.

Yan, Z., Hu, H. Y., Jiang, X., Maierhofer, V., Neb, E., He, L., ... Khaitovich, P. (2011). Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Res.*, *39*, 6596–6607.

Yin, C., Shen, G., Guo, D., Wang, S., Ma, X., Xiao, H., ... Li, F. (2016). InsectBase: A resource for insect genomes and transcriptomes. *Nucleic Acids Research*, *44*, D801–D807.

Ylla, G., Fromm, B., Piulachs, M. D., & Belles, X. (2016). The microRNA toolkit of insects. *Scientific Reports*, *6*, 37736.

Ylla, G., Piulachs, M. D., & Belles, X. (2017). Comparative analysis of miRNA expression during the development of insects of different metamorphosis modes and germ-band types. *BMC Genomics*, *18*, 774.

Zhang, P., Si, X., Skogerbø, G., Wang, J., Cui, D., Li, Y., ... Huang, D.-W. 2014. piRBase: A web resource assisting piRNA functional study. *Database: The Journal of Biological Databases and Curation*, *2014*, bau110.

Zhang, Y., Wang, X., & Kang, L. (2011). A k-mer scheme to predict piRNAs and characterize locust piRNAs. *Bioinformatics*, *27*, 771–776.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Llonga N, Ylla G, Bau J, Belles X, Piulachs M-D. Diversity of piRNA expression patterns during the ontogeny of the German cockroach. *J Exp Zool (Mol Dev Evol)*. 2018;1–8. https://doi.org/10.1002/jez.b.22815