Gene 537 (2014) 149–153

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short communication

Variable expression levels detected in the *Drosophila* effectors of piRNA biogenesis

Marie Fablet [a,b,*], Abdou Akkouche [a,b,1], Virginie Braman [a,b], Cristina Vieira [a,b,c,*]

[a] Université de Lyon, Université Lyon 1, F-69000 Lyon, France
[b] CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622 Villeurbanne, France
[c] Institut Universitaire de France

ARTICLE INFO

Article history:
Accepted 30 November 2013
Available online 18 December 2013

Keywords:
Transposable element
Retrotransposon
Piwi interacting small RNA
Drosophila simulans
Wild-type strains

ABSTRACT

piRNAs (piwi-interacting RNAs) are a class of small interfering RNAs that play a major role in the regulation of transposable elements (TEs) in *Drosophila* and are considered of fundamental importance in gonadal development. Genes encoding the effectors of the piRNA machinery are thus often thought to be highly constrained. On the contrary, as actors of genetic immunity, these genes have also been shown to evolve rapidly and display a high level of sequence variability. In order to assess the support for these competing models, we analyzed seven genes of the piRNA pathway using a collection of wild-type strains of *Drosophila simulans*, which are known to display significant variability in their TE content between strains. We showed that these genes exhibited wide variation in transcript levels, and we discuss some evolutionary considerations regarding the observed variability in TE copy numbers.

1. Introduction

In the last decade, major advances were made in our understanding of the epigenetic control of transposable elements (TEs), particularly regarding small RNAs (Saito and Siomi, 2010; Senti and Brennecke, 2010; Siomi et al., 2011). RNA interference is a widespread phenomenon, and the origin of its effectors dates back to the common ancestor of eukaryotes (Cerutti and Casas-Mollano, 2006). Several classes of small interfering RNAs were described, including piRNAs (piwi-interacting RNAs), which are the major regulators of TEs in *Drosophila* (Chambeyron et al., 2008; Pélisson et al., 2007; Vagin et al., 2006). In this study, we will refer to genes involved in the piRNA pathway as GIPPs. Analyses in *Drosophila melanogaster* revealed that mutations in GIPPs led to TE up-regulation (Kalmykova et al., 2005; Le Thomas et al., 2013; Vagin et al., 2004), causing abnormalities in germline development (Cook et al., 2004; Cox et al., 1998; González-Reyes et al., 1997; Li et al., 2009; Pane et al., 2007; Schüpbach and Wieschaus, 1991). The piRNA machinery is therefore considered a guardian of genome stability (Senti and Brennecke, 2010). In addition, its effectors seem to be involved in many other biological processes, such as splicing and DNA repair (Meister, 2013). Because of their biological relevance, GIPPs appear highly constrained and are described as conserved (Meister, 2013). We will refer to this as the development-like model.

In contrast, defense against TEs can be viewed as an immunological process, taking place at the genomic scale. Within this framework, evolutionary analyses of GIPPs revealed that they have been recurrently subject to positive selection, as is frequently observed for genes involved in immunity in a broad sense (Kolaczkowski et al., 2011; Obbard et al., 2009a, 2009b). As a consequence, GIPPs belong to the most rapidly evolving known coding sequences (Obbard et al., 2009a, 2009b). We will refer to this as the immunity-like model.

In order to address these apparently contradictory predictions—conservation (based on the development-like model) or rapid evolution (based on the immunity-like model), we used a collection of wild-type strains of *Drosophila simulans*, which are known to display variable TE contents between strains (Biémont et al., 2003; Vieira et al., 1999). We analyzed seven GIPPs: ago3, aub, and piwi, which are the direct effectors of the slicing step; the helicases spindle E (spnE) and armitage (armi); and the nucleases zucchini (zuc) and squash (squ). Although these genes were extensively studied at the DNA sequence level, a comparative analysis of their expression levels was never performed so far. Based on the development-like model, which suggests that GIPPs are highly constrained, their expression levels are expected to be conserved between strains. On the contrary, genes with the highest rates of sequence polymorphism are known to display the highest variability in expression levels (Lawniczak et al., 2008; Lemos et al., 2005; Nuzhdin

Abbreviations: piRNA, piwi-interacting RNAs; TE, Transposable elements; GIPP, genes involved in the piRNA pathway; ago3, argonaute 3; aub, aubergine; spnE, spindle E; armi, armitage; zuc, zucchini; squ, squash.

* Corresponding authors at: Université de Lyon, F-69000, Lyon, Université Lyon 1, CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, 43 bv 11 novembre 1918, F-69622, Villeurbanne, France. Tel.: +33 4 72 44 81 98.
E-mail addresses: marie.fablet@univ-lyon1.fr (M. Fablet), cristina.vieira@univ-lyon1.fr (C. Vieira).
[1] Present address: Institut de Génétique Humaine, CNRS, UPR1142; 141 rue de la Cardonille, F-34396, Montpellier, France.

0378-1119/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.11.095

et al., 2004). Thus, unlike the above expectation, based on the immunity-like model, GIPPs are predicted to present variable levels of expression between strains.

To test these opposing predictions, we quantified nucleotide polymorphism and present data on allozyme profiles for seven GIPPs in *D. simulans*. Further, we examined the variation in transcript level for the same GIPPs among 13 strains of *D. simulans*. Our data suggest that there is a high level of variation among strains, which supports the immunity-like model of evolution for GIPPs in *Drosophila*. We propose some evolutionary considerations regarding the associated variable TE contents of the strains.

2. Materials and methods

2.1. Drosophila stocks

We used wild-type strains of *D. simulans* which originated in Kenya (Makindu), Zimbabwe, Indian Ocean islands (Madagascar, Mayotte, Reunion), Atlantic Ocean islands (Madeira), Portugal (Chicharo), Russia (Moscow), Australia (Canberra, Eden, CannRiver), New Caledonia (Amieu, Noumea) and French Polynesia (Papeete). For the McDonald–Kreitman tests (MK tests), we also used two strains of *D. melanogaster* collected from Senegal and Portugal (Chicharo).

2.2. Sequence library for genes involved in the piRNA pathway

To focus exclusively on coding regions, we amplified sequences from cDNAs, which are devoid of introns. We extracted total RNAs from five adult females from each strain. PCR products were subsequently obtained from cDNAs, isolated using bacterial cloning and sequenced (Sanger sequencing). See Supplementary Material 1 for GenBank accession numbers.

The obtained sequences were translated and amino-acid sequences were analyzed using BLOSUM62 matrix scores (Henikoff and Henikoff, 1992). We considered substitutions with negative scores to belong to distinct allozymes.

2.3. Transcription level measurements

Twenty-five pairs of ovaries from two to four-day-old females were dissected in PBS. Total RNAs were extracted using the RNeasy Kit (Qiagen). cDNAs were produced using the ThermoScript RT-PCR system (Invitrogen) and oligo(dT) primers. The cDNAs were diluted 50-fold and quantified using SYBR Green 1 mix in a LightCycler 480 (Roche Diagnostics) using primers specific to each gene (Supplementary Material 2). Primers were designed in portions of sequences that were conserved between the variants we isolated. The transcript amounts were estimated relative to the amounts of the *rp49* gene, which showed the lowest variation among reference genes. The measurements were performed in three independent experiments.

3. Results and discussion

3.1. Sequences of GIPPs are variable between wild-type strains

Evolutionary studies of genes involved in immunity, particularly in the defense against viruses and TEs, have revealed that these are the most rapidly evolving genes in the genome and are repeatedly subject to positive selection (Kolaczkowski et al., 2011; Obbard et al., 2009b). Kolaczkowski et al. (2011) reported strong evidence of the adaptive evolution of *spnE* in *D. melanogaster*. They also identified evidence of adaptations in *aub* and *zuc*, but not in *piwi*. Obbard et al. (2009a) found significant deviations from neutrality for *aub* and *armi*. Our intention in this study was not to redo these analyses, however, we know that conclusions of neutrality tests depend on the sample of strains used. Therefore we tested whether data from our sample also led to neutrality rejection. For this purpose, we sequenced exonic portions of the genes in five wild-type strains of *D. simulans* and two wild-type strains of *D. melanogaster* and performed McDonald–Kreitman tests (MK tests). These analyses provided results congruent with the immunity-like model of sequence evolution (MK tests were significant for *armi*, *aub* and *spnE*) (Supplementary Material 3). This is also illustrated by the non-synonymous nucleotide diversity which was significantly larger for GIPPs than for the *alpha-tubulin at 84B* (*α-tub*) reference gene, whereas the synonymous nucleotide diversities were in the same range (Fig. 1). Again, this is in agreement with the immunity-like model for GIPPs. However, we cannot exclude the hypothesis of greater tolerance to segregating mildly deleterious mutations for GIPPs. In any case, sequence variability is higher for GIPPs.

The above sequences were translated in silico. They corresponded to distinct amino-acid sequences that could be clustered based on chemical profiles. The different clusters are referred to as allozymes (Supplementary Material 4). We could not find allozymes associated with particularly high (or low) expression level of GIPPs nor could we find allozymes associated with particularly high (or low) TE content. There were no obvious strain-specific associations between the allozymes of the different GIPPs, which led us to conclude that each strain had a unique combination of allelic variants of all seven genes.

3.2. Transcription levels of GIPPs are variable between wild-type strains

We quantified transcript levels in ovaries using RT-qPCR. We enlarged our sample to a panel of 13 wild-type strains of *D. simulans* and observed significant variation in the transcript levels of all GIPPs (Fig. 2). We performed the same experiment on four housekeeping genes that are often used as reference genes in expression experiments: *18S*, *adh*, *α-tub* and *CG13919*, and found lower variation. The coefficients of variation (square root of the variance divided by the mean) were significantly different between both categories of genes (Wilcoxon test, p-value = 0.012) (Fig. 2). We also performed ANOVA1 analyses to compute η^2 coefficients, which account for the amount of variability explained by the strains (η^2 equals sum of squares between groups divided by total sum of squares). ANOVAs were significant for all GIPPs, with high η^2 values (Fig. 2). These results indicate that GIPPs are significantly more variable in their transcript levels compared to reference genes. The transcription data from GIPPs were used to compute distance matrices among the strains. No significant correlations were established

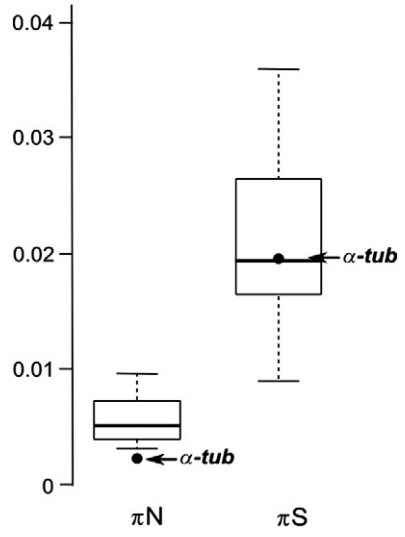

Fig. 1. Box plot of the nucleotide diversity (π) of GIPPs in *D. simulans*. The non-synonymous (N) and synonymous (S) nucleotide diversity (π) values calculated for seven GIPPs in wild-type strains of *D. simulans* are illustrated. The values obtained for *α-tub* are plotted in black circles.