Drosha, Dicer-1 and Argonaute-1 in the desert locust: Phylogenetic analyses, transcript profiling and regulation during phase transition and feeding

Niels Wynant *, Dulce Santos, Sudheendra Hebbar Subramanyam, Heleen Verlinden, Jozef Vanden Broeck

Molecular Developmental Physiology and Signal Transduction, KU Leuven, Naamsestraat 59, P.O. Box 02465, B-3000 Leuven, Belgium

A B S T R A C T

In this article, we identify and characterise the miRNA machinery components Drosha, Dicer-1 and Argonaute-1 of the desert locust. By means of phylogenetic analyses, we reveal important insights in the evolutionary context of these components. Our data illustrate that insect Argonaute-1 proteins form a monophyletic group with ALG-1 and ALG-2 of *Caenorhabditis elegans* and with the four (non-Piwi) Argonaute proteins present in humans. On the other hand, humans apparently lack clear homologues of the insect Argonaute-2 proteins. In addition, we demonstrate that *drosha*, *dicer-1* and *argonaute-1* display wide transcript tissue-distribution in adult desert locusts, and that during locust phase transition and feeding of starved locusts the expression levels of the miRNA pathway are regulated at the transcript level.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Micro (mi)RNAs are a class of non-coding small RNAs that regulate gene expression at the post-transcriptional level. They derive from long endogenous RNA fragments that are trimmed to a stem–loop structure by an RNase III enzyme called Drosha. This pre-miRNA is transported to the cytoplasm, where another RNAse III domain-containing enzyme, Dicer, cleaves it into a mature miRNA. Finally, the mature miRNA is loaded into the RNA induced silencing complex (RISC) that contains an Argonaute (Ago) protein. Following Watson–Crick base pairing with a complementary transcript sequence, protein production is prevented by translational inhibition or degradation of mRNA targets (Hammond, 2005).

The miRNA-mediated pathway belongs to a larger dsRNA dependent regulatory system, known as RNA interference (RNAi), which also comprises small interfering (si)RNA- and Piwi-interacting (pi)RNA-mediated gene silencing. Although diverse proteins are involved in the biogenesis of these noncoding small RNAs, the processing and effector steps of the RNAi-response are mediated in a common mode in plants, worms, insects and vertebrates, with a central role for Dicer and Argonaute proteins (Ghildiyal and Zamore, 2009). Endogenous (endo-) and exogenous (exo-)siRNAs originate from longer dsRNA-fragments, and play a central role in the protection of the organism against foreign invading nucleic acids, such as viruses and transposons (Sijen and Plasterk, 2003; Umbach and Cullen, 2009; Wang et al., 2006). Similarly to the processing of pre-miRNAs, long dsRNAs are cleaved into siRNAs by cytoplasmic Dicer enzymes that are subsequently loaded into Argonaute (Ago) containing RISC complexes, which result in degradation of complementary cellular transcripts. Finally, several studies demonstrated that endogenously encoded piRNAs play a crucial role in the control of the integrity of the genome, mainly by transposon silencing (Aravin et al., 2007). Although their biogenesis is very different from this of siRNAs and miRNAs, members of the Argonaute superfamily (AGO) are key players in

---

* Corresponding author. Tel.: +32 16324260; fax: +32 16323902.
E-mail address: Niels.Wynant@bio.kuleuven.be (N. Wynant).

piRNA-directed RNAi (Saito et al., 2006; Vagin et al., 2006). The Argonaute proteins involved in piRNA-directed RNAi are in general known as Piwi proteins and form a different AGO subfamily from the Argonaute proteins involved in siRNA- and miRNA-directed gene silencing (Ghildiyal and Zamore, 2009).

Although most investigated animal species encode a single Drosha enzyme, the number of Dicer and Argonaute enzymes can differ between organisms. Where nematodes and vertebrates have only one Dicer protein, insects typically encode two Dicer proteins (Dcr1 and Dcr2), which mainly produce miRNAs and siRNAs, respectively (Lee et al., 2004). The situation for the AGO super-family is more complex, which is illustrated by the fact that *Caenorhabditis elegans* possesses 27 annotated AGO genes (Yigit et al., 2006). Different forward and reverse genetic studies eluci-dated many of the functions of individual AGO proteins. First, these studies illustrated that two highly conserved members of the *C. elegans* AGO family, ALG-1 and ALG-2, have overlapping func-tions in miRNA-mediated RNAi and are essential for normal devel-opment (Grishok et al., 2001). Secondly, exogenous (exo-)siRNAs (originating from long foreign dsRNA-fragments) are mainly bound by the Argonaute protein RDE-1 (Yigit et al., 2006). Next, the RDE-1 complex will recruit RNA-dependent RNA polymerases (RdRPs), which will use the target mRNA as a template for primer-independent synthesis of new dsRNA (see also Boisvert and Simard, 2008). Then, a second Dicer complex will process these RdRP-products into secondary siRNAs, which will guide secondary Argonautes, including SAGO-1, SAGO-2 and other related proteins, for sequence specific gene silencing (Yigit et al., 2006). Furthermore, forward genetic studies have shown that endogenous siRNAs (endo-siRNAs) will primarily be incorporated into ERGO-1 containing RISC complexes (Yigit et al., 2006). *C. elegans* also encodes two Piwi Argonautes, PRG-1 and PRG-2, of which PRG-1 was shown to interact with piRNAs (Batista et al., 2008).

Insects and vertebrates lack RdRP homologous sequences in their genome and, probably for this reason, they do not have homologs for secondary AGO members. Most insects possess two different (primary) Ago proteins, Ago1 and Ago2, which mainly act in miRNA- and siRNA-dependent RNAi, respectively (Carthew and Sontheimer, 2009), and two or three Piwi Argonautes (Ago3, Piwi and Aubergine) that mediate piRNA-dependent gene silencing (Ross et al., 2014). In flies, loading of the small RNA duplexes is uncoupled from their loading into Ago1 or Ago2, but is directed by the structure of the duplex. Duplexes that contain a greater dsRNA structure will be loaded into Ago2, while small RNAs with mismatches and bulks are sorted into Ago1 (Tomari et al., 2007).

Several studies have demonstrated that miRNAs regulate impor-tant biological processes, including aging, apoptosis, development, neurodegeneration and metabolism in *Drosophila* (reviewed by Lucas and Raikhel, 2013). In addition, stage- and tissue-specific expression of miRNAs during development was reported in non-drosophilidae, *e.g.*, *Aedes albopictus* and *Culex quinquefasciatus* (Diptera) (Skalsky et al., 2010), *Bombyx mori* (Lepidoptera) (Yu et al., 2008) and *Blattella germanica* (Dictyoptera) (Cristino et al., 2011). Moreover, a role for miRNAs in phenotypic plasticity has been suggested, with different expression levels for several miRNAs in the solitary and gregarious phases of the migratory locust, *Locusta migratoria* (Orthoptera) (Wei et al., 2009), in heads of different honeybees, *Apis mellifera* (Hymenoptera) (Greenberg et al., 2012; Li et al., 2012) and in different morphs of the pea aphid, *Acyrthosiphon pisum* (Hemiptera) (Legeai et al., 2010).

Since many centuries, huge swarms of the desert locust, *Schistocerca gregaria*, caused great losses to agriculture in Africa, the Middle East and the Indian subcontinent. The success of this locust species is partially due to its ability to occur in two different phenotypic 'phases', a solitary and a gregarious one, that can gradually pass into each other and result in obvious differences in behaviour, development, morphology, physiology and colouration (Pener and Simpson, 2009; Uvarov, 1966; Verlinden et al., 2009). Increasing population density triggers the transformation into the gregarious phase, which is induced by stimuli from other locusts in close proximity (Ott et al., 2012; Verlinden et al., 2010). Besides phase transition, the desert locust serves as an important insect model organism for many other different biological processes, including phase transition, digestion (Dillen et al., 2014; Van Wielendaele et al., 2013), neurobiology (Badisco et al., 2011; Dillen et al., 2013) and the (systemic) siRNA machinery (Wynant et al., 2012, 2014a,b,c,d) in insects. Nevertheless, so far, the role of miRNAs in these locusts has not been investigated. Moreover, the expression level of components of the miRNA machinery in response to physiological changes remains to be investigated in insects. Therefore, in this paper, we first identify and characterise the miRNA machinery components Drosha, Dcr1 and Ago1 of the desert locust and next assess their regulation during locust phase transition and feeding of starved locusts, two processes that are characterised by the regulation of many different components simultaneously.

2. Materials and methods

2.1. Retrieving the sequences and protein domain prediction

By using the *drosha*, *dicer* and *argonaute* sequences of other insects (including *L. migratoria*, *Tribolium castaneum* and *Drosophila melanogaster*) as a query, transcript sequence informa-tion for *S. gregaria drosha*, *dcr1*, *dcr2*, *ago1*, *ago2*, *piwi1*, *piwi2* and *piwi3* was retrieved from the transcriptome database of *S. gregaria* (unpublished Illumina sequencing data) with reciprocal tBLASTn. The deduced amino acid sequence was determined by *in silico* translation using Prosite (ExPASy), which was used as a query for the prediction of protein domains using NCBI Conserved Protein Domain Search and PROSITE (ExPASy).

2.2. Phylogenetic analysis

Nucleotide sequence information for (other) insect, nematode and mammalian *drosha*, *dicer* and *argonaute* family members were retrieved from Genbank (NCBI) and *in silico* translated into the cor-responding amino acid sequence (Prosite, ExPASy). The identity of these fragments was confirmed by reciprocal tBLASTn (NCBI) and by verifying the present protein domains (NCBI Conserved Protein Domain Search). Next, the entire ORF sequences of Drosha, Dicer and Argonaute proteins, respectively, were aligned using Muscle alignment software (MEGA6.06). The protein domain with highest sequence similarity was selected using Jalview software. This corre-sponded to the first RNase III domain for Drosha, the second RNase III domain for Dicer and the PIWI domain for Argonaute proteins. Alignment using Muscle software (MEGA5.1) of the latter, as well as the entire ORF amino acid sequence was used for the construction of a maximum likelihood (ML) phylogenetic tree with 100 boot-straps (MEGA5.1). In addition, in order to confirm these data, the sequences were also aligned with T-coffee alignment software using the BLOSUM matrix (EMBL-EBI). The T-coffee alignment was subse-quently used for the construction of ML phylogenetic tree using the MEGA6.06 software. Similar results were obtained using the Muscle and T-coffee alignment software, as well as using the selected pro-tein domains and the entire ORF. The tree with the highest bootstrap values was selected for further analyses.

2.3. Rearing of animals

Gregarious desert locusts were reared under crowded condi-tions with controlled temperature (32 ± 1 °C), light (14 h photope-riod) and ambient relative humidity (40–60%). The gregarious