Gene 420 (2008) 107–112

Contents lists available at ScienceDirect

Gene

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

Conservation of structural elements in the mitochondrial control region of *Daphnia*

Kerstin Kuhn, Bruno Streit, Klaus Schwenk *

Department of Ecology and Evolution, Goethe-Universität Frankfurt am Main, Siesmayerstraße 70, D-60054 Frankfurt am Main, Germany

ARTICLE INFO

Article history:
Received 13 December 2007
Received in revised form 11 March 2008
Accepted 29 May 2008
Available online 6 June 2008

Keywords:
Crustacea
Daphnia longispina complex
Mitochondrial DNA
Conserved elements
Heteroplasmy
Population genetics

ABSTRACT

Mitochondrial DNA is well characterized in vertebrates and insects and represents a prime target for phylogenetic and population genetic analyses. In particular the highly variable control region, responsible for regulation of replication and transcription, has been studied by molecular biologists and geneticists alike. However, we lack basic information on the structure and molecular evolution of the control region in major animal groups, such as the crustaceans. Here we present a genetic analysis of the mitochondrial DNA control regions of three microcrustacean species of the genus *Daphnia*. Their control regions are located between the tRNA-Ile and the 12S rRNA and exhibit conserved structural elements previously described only for insects. We identified the (TA(A))n-block, a putative hairpin structure with a 5′ conserved flanking region and the GA-block. Moreover, not only the presence of these elements, but also their relative location on the mtDNA is conserved among *D. longispina* and insects. A comparative analysis of Malacostraca and Branchiopoda, which form a sister clade of insects revealed that the presence of structural elements per se is conserved between insects and crustaceans, but the arrangement of the elements along the control region as well as the number of each element is variable.

Finally we discuss the usefulness of the mitochondrial control region for population genetic studies in *Daphnia*, with respect to sequence divergence and heteroplasmy.

1. Introduction

Animal mitochondrial DNA (mtDNA) is a circular double stranded molecule with a length ranging in most species from 14–16 kb. It usually contains 37 genes (13 protein-coding, two ribosomal and 22 transfer RNA genes) and one major non-coding region. This non-coding region is also called the control region, because it contains a number of regulatory elements responsible for replication and transcription of the mtDNA (Wolstenholme, 1992).

The primary DNA sequence of the control region is not conserved within arthropods, but some general features like a high A+T content (Lewis et al., 1994), repetitive elements (Lewis et al., 1994; reviewed in Zhang and Hewitt, 1997; Lunt et al., 1998), poly thymine stretches (Lewis et al., 1994; Zhang et al., 1995; Saito et al., 2005) and hairpin structures (Clary and Wolstenholme, 1987; Taylor et al., 1993) occur frequently. In a detailed analysis, based on the comparison of Dipteran and Orthopteran control regions, Zhang et al. (1995) and Zhang and Hewitt (1997) pointed out five, so-called structural elements, which are apparently conserved and their relative location on the mtDNA is the same:

A poly thymine stretch at the 5′ end of the control region, a (TA(A))n-like sequence between the poly T stretch and a secondary structure, a stem and loop structure, conserved 5′ and 3′ flanking regions of the stem and a G+A-rich sequence block downstream the secondary structure.

Although these five structural elements have been well described in several insect species, no comparative study has been conducted to identify them in other arthropods. Few studies reported on structural elements in crustacean taxa. Poly T stretches were found in *Artemia franciscana* (Valverde et al., 1994) and *Cherax destructor* (Miller et al., 2004). In *Daphnia pulex* (Van Raay and Crease, 1994), *Ligia oceanica*, *Panulirus japonicus* and *Squilla mantis* (Kilpert and Podsiadlowski, 2006) stem and loop structures were described. Moreover, Kilpert and Podsiadlowski (2006) also found homologies to the conserved motifs of the 5′ and 3′ flanking regions of the stems in insects. In contrast in *Tigriopus japonicus* (Machida et al., 2002) none of the conserved elements were identified.

Despite limited knowledge on their structure, control regions of crustaceans have been successfully used in phylogenetic (Lehman et al., 1995), phylogeographic and population genetic studies (McMillen-Jackson and Bert, 2003; Diniz et al., 2005). Based on the high level of genetic variation in control region sequences compared with protein protein-coding or ribosomal genes, they represent a suitable and

Abbreviations: A, adenine; bp, base pair(s); G, guanine; mtDNA, mitochondrial DNA; PCR, polymerase chain reaction; rRNA, ribosomal RNA; T, thymine; tRNA, transfer RNA; Ile, isoleucine.

* Corresponding author. Tel.: +49 69 7982 4775; fax: +49 69 7982 4910.
E-mail address: K.Schwenk@bio.uni-frankfurt.de (K. Schwenk).

0378-1119/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2008.05.020

powerful genetic marker for population genetic analyses (e.g. Dudycha, 2004; Place et al., 2005). However, several attempts to establish the control region as a molecular marker for species belonging to the *D. longispina* complex failed (personal communication D. Taylor and P. Spaak). Species of this group (*D. galeata*, *D. hyalina* and *D. cucullata*) form interspecific hybrids and recombinant genotypes in many large lake populations (Schwenk and Spaak 1995). These crustaceans belong to the most important key species in aquatic ecosystems and are widely distributed across the Holarctic. In order to effectively determine (1) the level of differentiation among populations, (2) the level of genetic diversity and (3) gene flow, highly variable mitochondrial DNA markers, such as the control region, are required.

The aim of our study was to characterize the control region of *D. galeata* and two other *Daphnia* species of the *D. longispina* complex, *D. cucullata* and *D. hyalina*. Secondly, we searched for conserved structural elements which have only been described in control regions of insects. In order to assess whether these elements are phylogenetically conserved among crustacean species, we have compared control regions of several branchiopodan and malacostracan species.

Thirdly, we evaluated the quality and discriminatory power of control region sequences for population genetic analyses of species belonging to the *D. longispina* complex.

2. Materials and methods

DNA was isolated from approximately 300 individuals of a *D. galeata* clone (G1; see Schwenk et al., 2001), a *D. cucullata* clone (C2) and four *D. hyalina* clones (H3, H2, DZ-01 and DZ-04) using the CTAB-method as described in Winnepenninckx et al. (1993). Different primer combinations (see Table 1) corresponding to the 5′ region of the 12S rRNA gene (reverse) and to the 5′ region of the tRNA-Ile or the tRNA-Met genes (forward) respectively, were used to amplify the putative control region.

PCR reactions were performed in 25 µl containing 1–10 ng template DNA, 5 pmol of each primer, 0.2 mM of each dNTP, 2 mM $MgCl_2$, 1× amplification buffer and 0.5 U Taq DNA polymerase (Invitrogen). The following temperature profile was applied: initial denaturation at 94 °C for 2:30 min; 35 cycles of 92 °C for 1 min, 52 °C for 1 min, 72 °C for 1 min; and a final extension at 72 °C for 3 min using a Tetrad (PTC-225) thermocycler (BIO-RAD). PCR products were visualized by electrophoresis using 1.4% agarose gels. Since all primer combinations (as well as nested PCR reactions) resulted in multiple PCR products, bands were excised from the gel and DNA was extracted and purified using a QIAgen Gel Extraction kit. 50 ng of DNA was cycle sequenced using the ABI Prism Big Dye terminator kit (Perkin Elmer, Norwalk, CT, USA) and analyzed on a CEQ 2000 (Beckman Coulter) automatic sequencer. Sequence data were deposited at GenBank: EU784141 (*D. galeata*), EU784142 (*D. cucullata*) and EU784143 (*D. hyalina*).

For southern blot analysis 1–5 µg of total DNA were digested using Hpa I and Nsp I in 60 µl at 37 °C over night. The restriction site of Hpa I is located at the 5′ end of the 12S rRNA and Nsp I cuts at the 5′ end of the tRNA-Ile. Digested DNA was electrophoresed on a 0.9% agarose gel, blotted onto a positively charged nylon membrane (Sambrock et al., 1989) and hybridized with a digoxigenin labeled probe. A sequenced fragment containing the putative control region (with only one repeat) was used as a template for the amplification of the probe. The PCR was performed as described above using an internal primer of the control region (CRint, see Table 1) and the 12S-3 primer which hybridizes at the 5′ end of the 12S rRNA. The amplified fragment was excised from the gel, purified and randomly labeled using the DIG High Prime DNA Labeling Kit (Roche). Hybridization conditions and detection reactions were performed according to the manufacturer's instructions (Roche). The program MFOLD (Zuker et al., 1999) was used to search for secondary structures.

Table 1
Primers were designed based on an alignment of mtDNA sequences of different crustaceans

Name	Forward primer	Name	Reverse primer
Ile-1	CTACCCTATCAAGGTAGTCCTT	12S-1	ACAGGGTATCTAATCCTGG
Ile-2	TTACCCTATCAAGGTAATCC	12S-2	TAACAGGGTATCTAATCCTGG
Met-1	TTGGGCATGAACCCACTA	12S-3	CGTCTAACCGCGACGGCTGGCACG
Met-2	CCCACTAGCTTAACTTAGCTTACC	CRint	CCTTCACCTTTTCTAAC
Met-3	GGGCATGAACCCACTAGCTT		

CRint represents an internal primer based on DNA sequence analysis of the *D. galeata* control region. Met-3 is described in Lehman et al. (1995).

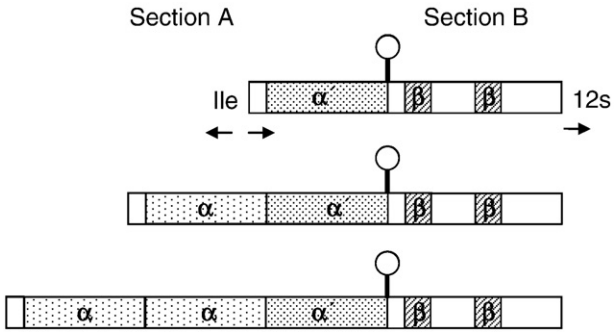

Fig. 1. Schematic diagram of the structure of the control region of *D. galeata*. Three length variants (394 bp, 573 bp and 752 bp), with one to three repeats in section A are shown. Tandem repeats α and α′ are indicated by spotted and tandem repeats β by striped boxes. ♀ = potential hairpin structure. Ile indicates the location of the isoleucin gene and 12S the position of the 12S ribosomal gene. Arrows indicate the direction of transcription according to Van Raay and Crease (1994).

The following control region sequences from further crustaceans available in GenBank were included in this study: Branchiopoda: *A. franciscana* (GenBank accession no. X69067), *Daphnia melanica* (GenBank accession no. EF471203), *Daphnia middendorffiana* (GenBank accession no. EF741204), *D. pulex* (GenBank accession no. AF117817), *Triops cancriformis* (GenBank accession no. AB084514) and *Triops longicaudatus* (GenBank accession no. NC_006079), Malacostraca: *Farfantepenaeus duorarum* (GenBank accession no. AF100736), *Fenneropenaeus chinensis* (GenBank accession no. DQ656600), *Gonodactylus chiragra* (GenBank accession no. DQ191682), *Harpiosquilla harpax* (GenBank accession no. NC_006916), *Litopenaeus vannamei* (GenBank accession no. EF584003), *Marsupenaeus japonicus* (GenBank accession no. AP006346), *P. japonicus* (GenBank accession no. AB071201), *Penaeus monodon* (GenBank accession no. NC_002184), *Squilla empusa* (GenBank accession no. DQ191684) and *S. mantis* (GenBank accession no. NC_006081).

3. Results and discussion

3.1. General characteristics of Daphnia control regions

The putative control region of *D. galeata*, *D. cucullata* and *D. hyalina* is located between the tRNA-Ile and the 12S rRNA gene, which is thought to represent the ancestral position in the insect–crustacean arrangement (Crease, 1999; Lavrov et al., 2000, 2002; Segawa and Aotsuka, 2005). It is also called the A+T-rich region, since in the genus *Drosophila* the adenine and thymine content ranges between 90% and 96% (Fauron and Wolstenholme, 1976). In general, insect control regions exhibit A+T contents above 85% (Zhang and Hewitt, 1997). The overall A+T content among the *Daphnia* species of the *D. longispina* complex is relatively low (64%–67%) but in the range of what is found in other crustacean species, e.g. 55.8% in *L. oceanica* (Kilpert and Podsiadlowski, 2006), 67% in *D. pulex* (Van Raay and Crease, 1994), 70.6% in *P. japonicus* (Yamauchi et al., 2002) and 81.5% in *P. monodon* (Wilson et al., 2000).

For a better description we divided the control region of *D. galeata* into sections A and B, of which both contain repetitive sequences (Figs. 1 and 2). Section A is located at the 5′ end of the control region adjacent to the tRNA-Ile. It is composed of a variable number of 179 bp