



The mtDNA control region structure and preliminary phylogenetic relationships of the genus *Gampsocleis* (Orthoptera: Tettigoniidae)

YAN-XIA ZHANG, ZHI-JUN ZHOU¹, YAN-LIN CHANG, MING-RU YANG & FU-MING SHI

College of Life Sciences, Hebei University, Baoding, Hebei 071002, China

¹Corresponding author. E-mail: zhijunzhou@163.com

Abstract

The mitochondrial DNA control region (mtDNA CR) has been cloned and sequenced from 56 individuals representing five species of the genus *Gampsocleis*. The Neighbor-Joining (NJ) tree under the Kimura 2-parameter (K2P) model, *G. sedakovii* haplotypes cluster into two distinct clades A and B. It suggests there may be more than one species within this group. The macropterous species, *G. sinensis* grouped among *G. ussuriensis*. Two brachypterous species, *G. carinata* and *G. gratiosa* grouped as sibling species. Individuals of *G. carinata* and *G. gratiosa* grouped together in distinct clades, even when collected from different localities. MtDNA CR structural analysis found two conserved motifs in the primary sequences “(TA)_n” and poly-A stretch in the twelve ensiferan species examined, which may be involved in mtDNA transcription or replication control.

Key word: *Gampsocleis*, Mitochondrial DNA control region, Haplotype, genetic diversity

Introduction

The mitochondrial DNA control region (mtDNA CR) is well known for the mtDNA initiation of replication and transcription control in vertebrates (Shadel & Clayton 1997). In insects, the mtDNA CR is also known as an A+T-rich region because of the nearly 90% adenine (A) and thymine (T) in the earliest investigated species (Clary & Wolstenholme 1985; Beard *et al.* 1993; Crozier & Crozier 1993), and it is believed to have a similar role. The mtDNA CR is also the most variable segment and the major non-coding region in the maternally inherited mtDNA (Wolstenholme 1992; Boore 1999). It is frequently located between the *srRNA* and tRNA clusters (*trnI-trnQ-trnM*) in insects. The size of mtDNA CR varies in length from tens to several thousands of bases among species and also varies within species (Lewis *et al.* 1995; Kim *et al.* 2007; Zhou *et al.* 2007).

Previous studies of the mtDNA CR structural organization have classified the insect mtDNA CR into two different groups. In Group 1, such as *Drosophila* species, mtDNA CR was divided into two different domains: (1) a highly conserved domain adjacent to *trnI* that may be implicated in origin of replication and the control elements of transcription of mtDNA; (2) a variable domain containing the remainder of the mtDNA CR. In Group 2, found in grasshoppers, mosquitoes and possibly butterflies, short conserved sequences motifs are scattered through the whole mtDNA CR. Five conserved structural motifs were identified both in group 1 and group 2: (1) a poly-T stretch at the 5' end of the A+T-rich region (near *trnI*), which may be involved in the control of transcription and/or the initiation of replication; (2) a [TA(A)_n]-like stretch following the poly-T stretch; (3) a highly conserved stem-loop structure that may be associated with the origin of second strand replication; (4) a highly conserved flanking sequence within the stem-loop structure harboring one 5' consensus of “TATA” and one 3' consensus of “G(A)_nT”; and (5) a G+A-rich sequence motif located downstream of the stem-loop structure (Zhang *et al.* 1995; Zhang & Hewitt 1997). However, only some of these motifs has been reported in the oriental mole cricket, *Gryllotalpa orientalis* (Kim *et al.* 2007), such as three “[TA(A)_n” sequences and two poly-T stretches. Moreover, alignments of these two poly-T stretches to the poly-T stretch described previously by Zhang *et al.* (1995) yielded no similar sequence motifs (Kim *et al.* 2007). A poly-A stretch (~15 nucleotides) described in several *Drosophila* species without any interruption is interrupted by two G in *G. orientalis* (Kim *et al.* 2007).