



ELSEVIER

Gene 216 (1998) 149–153

GENE

AN INTERNATIONAL JOURNAL ON
GENES AND GENOMES

Cloning and characterization of a microsatellite in the mitochondrial control region of the African side-necked turtle, *Pelomedusa subrufa*

Rafael Zardoya *, Axel Meyer^{1,2}

Department of Ecology and Evolution and Program in Genetics, State University of New York, Stony Brook, NY 11794-5245, USA

Received 17 March 1998; accepted 25 May 1998

Abstract

The nucleotide sequence of the African side-necked turtle mitochondrial control region and its flanking tRNA genes was determined. This 73% A+T-rich region is 1194 bp long. Several conserved motifs involved in the regulation of the mitochondrial genome replication process, including one conserved sequence block (CSB1), and three termination-associated sequences were identified. The most remarkable feature found in this control region was the presence of six microsatellite-containing tandem repeats between the CSB1 motif and the tRNA^{Phe} gene. The potential usefulness of this microsatellite sequence for population-level studies is enhanced by its unique localization in the maternally inherited mitochondrial molecule. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Simple sequence repeats; Reptiles; D-loop

1. Introduction

Simple-sequence repeats (SSR) or microsatellites are DNA sequences made up of 2–5 bp tandemly repeated units. Microsatellites may be classified according to the nature of the repeat as perfect, imperfect (i.e. interrupted) or compound (adjacent tandem repeats of a different sequence) (Weber, 1990). Interrupted microsatellites seem to be less variable than perfect ones, and within the latter, longer repeats are expected to be more polymorphic (Jarne and Lagoda, 1996). Although the exact molecular mechanisms that create variation are not completely understood, the extremely large number of alleles exhibited per microsatellite locus is generally believed to be generated in a stepwise fashion (Valdes et al., 1993) by DNA polymerase replication slippage

events (Schlötterer and Tautz, 1992). Microsatellite loci appear to be widely distributed every 20–30 kb, on average, throughout each eukaryotic genome (Stallings et al., 1991). Di-, tri- and tetra-nucleotides have been found in a wide variety of eukaryotes (Hamada et al., 1982), as well as in the chloroplastic genome of plants (Vendramin et al., 1996). However, only one case of a microsatellite in a mitochondrial genome has been reported (Hoelzel et al., 1993).

Microsatellites are rapidly becoming the molecular marker of choice among evolutionary biologists who are interested in the analysis of population genetic structure (Jarne and Lagoda, 1996). The relative technical ease with which microsatellites can be obtained from a particular species (by constructing a genomic library of the species of choice, sequencing microsatellite-containing clones, and designing PCR primers in the flanking regions of the microsatellite), and their extremely high levels of polymorphism (with an average expected heterozygosity far above 50%) make these genetic markers very convenient and useful for problems in conservation biology, population genetics, and behavioral ecology studies (reviewed in e.g. Queller et al., 1993; Jarne and Lagoda, 1996). Moreover, there is recent evidence suggesting that, despite the extremely fast rates of microsatellite evolution, homologous microsatellite loci can persist for long evolutionary time spans because

* Corresponding author. Present address: Museo Nacional de Ciencias Naturales, José Gutiérrez Abascal, 2, 28006 Madrid, Spain.

Tel: +34 (91) 561-8607; Fax: +34 (91) 564-5078;

e-mail: mcnr154@pinar2.csic.es

¹ Present address: Department of Biology, University of Konstanz, D-78457 Konstanz, Germany.

² Hopkins Marine Station, Stanford University, Pacific Grove, CA 93950, USA.

Abbreviations: CSB, conserved sequence block; TAS, termination-associated sequence; SSR, simple-sequence repeat(s); mtDNA, mitochondrial DNA; tRNA, transfer ribonucleic acid.

of their relatively conserved flanking regions (FitzSimmons et al., 1995; Rico et al., 1996; Zardoya et al., 1996). This circumstance allows, through use of versatile PCR primers, screening of homologous microsatellite loci in a large number of species without having to invest time and expense in constructing and screening genomic libraries anew for each species (Zardoya et al., 1996).

Here, we report the unusual presence of a microsatellite sequence within the mitochondrial control region of the African side-necked turtle (*Pelomedusa subrufa*).

2. Materials and methods

Mitochondrial DNA (mtDNA) was extracted from the liver of an African side-necked turtle, *P. subrufa*, as described previously (Zardoya et al., 1995a). After homogenization, intact nuclei and cellular debris were removed by low-speed centrifugation, and the purified isolated mitochondria were subjected to a standard alkaline lysis procedure to extract their DNA. The isolated mtDNA was cleaved with the *ApaI* restriction enzyme and cloned into the pGEM-7f vector. A 5 kb *ApaI* fragment spanning the end of the *cyt b* gene to the end of the ND1 gene was found to contain the control region, as was expected based on the conserved mitochondrial genome order in vertebrates. Cloned DNA was used as template for Taq Dye Deoxy Terminator cycle-sequencing reactions (Applied Biosystems Inc.) following the manufacturer's instructions. Sequencing was performed with an automated DNA sequencer (Applied Biosystems 373A *Stretch*). Sequences were obtained using both M13 universal sequencing primers and three control-region-specific oligonucleotide primers (Tur d-loop F: 5'-GGCTATGTACGTCGTGCATTCAT-3'; Tur d-loop F1: 5'-TCTTCAGGATACCTCTGGCTGTT-3'; Tur d-loop R: 5'-GGAAGTGTATATGAAACCTGGGT-3'). The sequences obtained from both strands were about 450–550 bp in length, and each sequence overlapped the next contig by about 150 bp. In no case were differences in sequence observed between the overlapping regions. Sequence data were analyzed by use of the GCG program package (Devereux et al., 1984). The nucleotide sequence reported in this paper has been deposited in the EMBL/GenBank data libraries under the accession number AF039066.

3. Results and discussion

3.1. General features of the turtle mitochondrial control region

The control region of the African side-necked turtle mitochondrial genome is 1194 bp long and, as in most

vertebrates, it is localized between the tRNA^{Pro} and tRNA^{Phe} genes (Fig. 1). The overall base composition of the L-strand of the control region is A: 38%; T: 35%; C: 19%; and G: 8%. In other vertebrates, this region includes the origin of H-strand replication, the sites of initiation of both H- and L-strand transcription, and several motifs involved in the regulation of both processes. Interestingly, only one (CSB1) of the three conserved blocks involved in the initiation of the DNA synthesis (Walberg and Clayton, 1981; reviewed in Saccone et al., 1991) could be identified unambiguously at the 3' end of the control region (right domain) (Fig. 1). Additionally, an interrupted poly-C stretch, remarkably similar to the CSB2 motif (Walberg and Clayton, 1981; reviewed in Saccone et al., 1991), was found close to the 5' end of the control region (left domain) (Fig. 1). However, the homology and functionality of this stretch with the CSB2 motif remain tentative due to its unusual position. Additionally, up to three termination associated sequences (TASs) involved in the premature termination of the H-strand replication (Doda et al., 1981), as well as several copies of the conserved palindromic motif 5'-TACAT-3' (Saccone et al., 1991) were found at the 5' end of the control region (Fig. 1). No significant secondary structures that are found in other species (Saccone et al., 1991) could be identified in the left and central domains.

3.2. A microsatellite associated with longer tandem repeats

The most striking feature of the African side-necked turtle mitochondrial control region is the presence of six direct repeats localized in tandem at its 3' end, downstream of the CSB1, and close to the tRNA^{Phe} gene (Fig. 1). Each repeat is composed of a 45 bp sequence that is followed by a (TA)_n microsatellite with a variable number of repeat units ($n=10-11$). In total the repeat region is 453 bp long (i.e. it covers 38% of the whole control region). Interestingly, the microsatellite in the repeat that is closer to the tRNA^{Phe} gene is longer than the others, and it is composed of the repetition of seven 5'-TAA(TA)₃₋₄-3' units (Fig. 1). This type of pattern, in which the rightmost repeated unit of the array shows the higher level of divergence, has also been found in several mammals (e.g. Fumagalli et al., 1996 and references therein), and this could be related to an asymmetry in the replication of the mtDNA molecule (Fumagalli et al., 1996). The presence of direct repeats is normally associated with the 5' end of the control region (e.g. Wilkinson and Chapman, 1991; Zardoya et al., 1995b; Fumagalli et al., 1996), and has also been reported at the 3' end, either between the CSB1 and CSB2 motifs, or between the CSB2 and the tRNA^{Phe} gene (e.g. Fumagalli et al., 1996). Extensive size heteroplasmy in mitochondrial control regions due

a

```

tRNA-Thr→
1  GCATTTTCAGTAGTTTATCCAAAACCTTGGCCCTTGTAACCAAAAAATGGCAAAATCAGTTGCCCTGAAATATCAAAAAAAGGAGTAGACCCCTGTCCACC 100
      ←tRNA-ProControl region→
101  AACTCCCAAAGCTGGTATTTTCACTAAACTATTTTGGCCAAAAAAGAACAGTAAATCGCCTACATCGCCACCACGTCTCCAGCTATCCCCCCTAC 200
      TAS-1
201  CCCCCGTAGGGGATTTTGGCTATGTACGTCGTGCATTCATTTTTTGGCCCATGCATATCATTTGTATACATTAAATTTTCATGGTCACAGGACAATGT 300
      TAS-2      TAS-3
301  ACATTAAATCCATGATTTAGGACATACTATTAAATAATTCTTTCTCATGAATATTGCCACAGTACTTGGCTATTTCTTGATTGAGCTACTCACGTG 400
401  AACATGACATTCCTGTTATATAGGCACACGGTCCCAGCTTCTACGTCCATCATTTTAAATTCGTACCTTATAATGATCTTCAGGATACCTCTGGCTGT 500
501  TTTTTCAGGAACCTTACATCTAAGACGGGCATCGATCGTCTTCAGGATACCTCTCGTTGCATTTTTAGTCAGTGAGTTCTATACACACGTTCAAATAA 600
601  CAAGACATCAATTCCTTCCAGGCATATAGTACTCTTTTTATTTTTATCTCGCCTTCACAATGACACCTATTTATGGTCTCTCTCAATTCTCTGAAACTC 700
      CSB1
701  TGGCCCTTATGTAAGTTATTTACCGAACAACAAACCTTTCGCGTTGGCTTATTAATTAATGCTCACAGGACATAACACTTTTGTGTAAAACCTCAATA 800
801  GGAAAAATTCATTCGTTTTTACACTACTTATTAACATTATTACACAACATATATACAAACAAAGTACTAATTATACTAATCATATAAATATATATA 900
901  TATATATATATACAAACATATATACAAACAAAGTACTAATTATACTAATCATATAAATATATATATATATATATATATATATATATACAAACATA 1000
1001  AGTACTAATTATACTAATCCATATAAATATATATATATATATATATATATACAAACATATATACAAACAAAGTACTAATTATACTAATCATATAAATATA 1100
1101  TATATATATATATATACAAACATATATACAAACAAAGTACTAATTATACTAATCATATAAATATATATATATATATATATATATATATATACAAACATATATACAA 1200
1201  ACAAGTACTAATTATACTAATCATATAAATATATATACATATAAATATATATAAATATATATAAATATATATAAATATATATAAATATATATATAA 1300
      tRNA-Phe→
1301  TATATATATTACAAATATCTCTACTCTTAATCCAAATAAATTTTTCACATATACATGTTATTTGAGCTTATATTATTAAAGCATAGCACTGAAGATGC 1400
1401  TAAGATAGATTAAATAAATCTCAAAAACA 1430
    
```

b

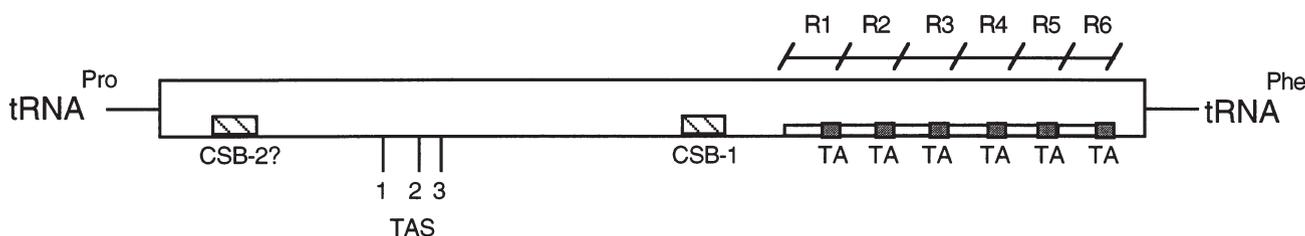


Fig. 1. Control region of the African side-necked turtle (*Pelomedusa subrufa*). (a) Nucleotide sequence of the 1430 bp mitochondrial genome fragment that includes the control region and flanking tRNA genes. Several conserved motifs were identified along this fragment: two conserved sequence blocks (CSB-1 and CSB-2), three termination-associated sequences (TASs), and six direct repeats. (b) Scheme showing the relative position of the above-mentioned control region features. The striking presence of a TA-microsatellite at the end of the direct repeats (R1–R6) is particularly emphasized.

to variation in the number of these tandem repeats has been reported in many species (e.g. Edwards and Wilson, 1990; Wilkinson and Chapman, 1991; Fumagalli et al., 1996). However, it is highly unusual to find a microsatellite associated with these mitochondrial control region tandem repeats. So far, mitochondrial microsatellites have been reported only in two elephant seal species

(Hoelzel et al., 1993). These species had a (AC)_n GT microsatellite not associated with longer tandem repeats in the 3' end of the mitochondrial control region. These mitochondrial microsatellites showed extensive heteroplasmy with up to three length variants present in single individuals (Hoelzel et al., 1993).

Tandemly repeated sequences in the mitochondrial

control region of non-microsatellite-type are believed to be generated by strand slippage and mispairing during replication (e.g. Fumagalli et al., 1996), which is the same mechanism proposed for the generation and maintenance of microsatellites (Schlötterer and Tautz, 1992). Moreover, it has been suggested that the presence of microsatellites may be, in some cases, directly related to the origin of longer tandem repeats (Wright, 1994). The fact that each of the larger repeats contains a conserved sequence and microsatellites of different sizes suggests that the origin of the microsatellite pre-dated the generation of the tandem array. The discovery of microsatellites in the mitochondrial control region may provide insight into the birth, evolution and properties of microsatellites. For instance, the strict Mendelian inheritance of microsatellites (e.g. Queller et al., 1993) is no longer tenable in light of this finding.

3.3. Potential usefulness of a mitochondrial microsatellite in population genetic studies

Traditionally, the mitochondrial control region has been widely used as a molecular marker in population-level studies because of its high mutation rates (Brown et al., 1979; Avise, 1994). The presence of a microsatellite in the control region of the African side-necked turtle provides a potentially very useful genetic marker which combines the properties of microsatellites (i.e. extremely large number of alleles) with those of mitochondrial genomes (i.e. maternal inheritance, lack of recombination, and haploidy). During the last years, the mitochondrial control region has been used to specifically study population structure, gene flow, migration, phylogeography, female nesting behavior assessment, and kinship of turtles (e.g. Norman et al., 1994; Encalada et al., 1996; Schroth et al., 1996 and references therein). These studies, as in most vertebrate species, have focused primarily on the 5' end of the mitochondrial control region because of the availability of versatile PCR primers (Kocher et al., 1989). However, microsatellites may be present in the understudied 3' end of the mitochondrial control region of other vertebrate species (e.g. Hoelzel et al., 1993), and may be more widespread than currently recognized. Our results reveal that the 3' end, due to the presence of microsatellites, may be a potentially informative molecular marker of population structure, and particularly female-mediated processes (e.g. fidelity of return to particular nesting sites in turtles) in vertebrates.

3.4. Conclusions

(1) We have cloned and characterized the mitochondrial control region of the African side-necked turtle, *P. subrufa*. Sequence comparison analyses reveal the

presence of typical control region conserved motifs (CSBs and TASs).

- (2) A (TA)_n microsatellite embedded within a larger repeat was found in the 3' domain of the control region. Our results support that the origin of the microsatellite pre-dated the generation of the larger tandem array.
- (3) A microsatellite in a mitochondrial genome may be a potentially very useful molecular marker for population genetic studies. Further research focusing on the 3' end of the control region of vertebrate mitochondrial genomes is encouraged by this finding.

Acknowledgement

The side-necked turtle specimen was a kind gift from Nicole Valenzuela of Stony Brook. R.Z. was sponsored by a postdoctoral grant of the Ministerio de Educacion y Ciencia of Spain. This work received partial financial support from grants from the National Science Foundation (BSR-9107838, BSR-9119867, DEB-9615178), a collaboration grant with the Max-Planck-Institut für Biology in Tübingen, and the University of Konstanz to A.M.

References

- Avise, J.C., 1994. *Molecular Markers, Natural History, and Evolution*. Chapman & Hall, New York.
- Brown, W.M., George, M.J., Wilson, A.C., 1979. Rapid evolution of mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 76, 1967–1971.
- Devereux, J., Haeberli, P., Smithies, O., 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.
- Doda, J.N., Wright, C.T., Clayton, D.A., 1981. Elongation of displacement loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proc. Natl. Acad. Sci. USA* 78, 6116–6120.
- Edwards, S.V., Wilson, A.C., 1990. Phylogenetically informative length polymorphism and sequence variability in mitochondrial DNA of Australian songbirds (*Pomatostomus*). *Genetics* 126, 695–711.
- Encalada, S.E., Lahanas, P.N., Bjørndal, K.A., Bolten, A.B., Miyamoto, M.M., Bowen, B.W., 1996. Phylogeography and population structure of the Atlantic and Mediterranean green turtle *Chelonia mydas*: a mitochondrial DNA control region sequence assessment. *Mol. Ecol.* 5, 473–484.
- FitzSimmons, N.N., Moritz, C., Moore, S.S., 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Mol. Biol. Evol.* 12, 432–440.
- Fumagalli, L., Taberlet, P., Favre, L., Hausser, J., 1996. Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. *Mol. Biol. Evol.* 13, 31–46.
- Hamada, H., Petrino, M.G., Kakunaga, T., 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. USA* 79, 6465–6469.
- Hoelzel, A.R., Hancock, J.M., Dover, G.A., 1993. Generation

- of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. *J. Mol. Evol.* 37, 190–197.
- Jarne, P., Lagoda, P.J., 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* 11, 424–429.
- Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals. *Proc. Natl. Acad. Sci. USA* 86, 6196–6200.
- Norman, J.A., Moritz, C., Limpus, C.J., 1994. Mitochondrial DNA control region polymorphisms: genetic markers for ecological studies for marine turtles. *Mol. Ecol.* 3, 363–373.
- Queller, D.C., Strassmann, J.E., Hughes, C.R., 1993. Microsatellites and kinship. *Trends Ecol. Evolution* 8, 285–288.
- Rico, C., Rico, I., Hewitt, G., 1996. 470 million years of conservation of microsatellite loci among fish species. *Proc. R. Soc. Lond. B* 263, 549–557.
- Saccone, C., Pesole, G., Sbisà, E., 1991. The main regulatory region of mammalian mitochondrial DNA: structure–function model and evolutionary pattern. *J. Mol. Evol.* 33, 83–91.
- Schlötterer, C., Tautz, D., 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Res.* 20, 211–215.
- Schroth, W., Streit, B., Schierwater, B., 1996. Evolutionary handicap for turtles. *Nature* 384, 521–522.
- Stallings, R.L., Ford, A.F., Nelson, D., Torney, D.C., Hildebrand, C.E., Moyzis, R.K., 1991. Evolution and distribution of (GT)_n repetitive sequences in mammalian genomes. *Genomics* 10, 807–815.
- Valdes, A.M., Slatkin, M., Freimer, N.B., 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* 133, 737–749.
- Vendramin, G.G., Lelli, L., Rossi, P., Morgante, M., 1996. A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Mol. Ecol.* 5, 595–598.
- Walberg, M.W., Clayton, D.A., 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Res.* 9, 5411–5421.
- Weber, J.L., 1990. Informativeness of human (dC-dA)_n dot (dG-dT)_n polymorphisms. *Genomics* 7, 524–530.
- Wilkinson, G.S., Chapman, A.M., 1991. Length and sequence variation in evening bat D-loop mtDNA. *Genetics* 128, 607–617.
- Wright, J.M., 1994. Mutation at VNTRs: are minisatellites the evolutionary progeny of microsatellites? *Genome* 37, 345–347.
- Zardoya, R., Garrido-Pertierra, A., Bautista, J.M., 1995a. The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. *J. Mol. Evol.* 41, 942–951.
- Zardoya, R., Villalta, M., Lopez-Perez, M.J., Garrido-Pertierra, A., Montoya, J., Bautista, J.M., 1995b. Nucleotide sequence of the sheep mitochondrial DNA D-loop and its flanking tRNA genes. *Curr. Genet.* 28, 94–96.
- Zardoya, R., Vollmer, D.M., Craddock, C., Streelman, J.T., Karl, S., Meyer, A., 1996. Evolutionary conservation of microsatellite flanking regions and their use in resolving the phylogeny of cichlid fishes (Pisces: Perciformes). *Proc. R. Soc. Lond. B* 263, 1589–1598.