CSIRO PUBLISHING

Australian Journal of Zoology, 2015, **63**, 111–114
http://dx.doi.org/10.1071/ZO14092

Short communication

The first complete mitochondrial genome of Pygopodidae (*Aprasia parapulchella* Kluge)

Anna J. MacDonald[A,C], *Theresa Knopp*[A], *Mitzy Pepper*[B], *J. Scott Keogh*[B] *and Stephen D. Sarre*[A]

[A]Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601, Australia.
[B]Division of Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia.
[C]Corresponding author. Email: Anna.MacDonald@canberra.edu.au

Abstract. The Pygopodidae comprise an enigmatic group of legless lizards endemic to the Australo-Papuan region. Here we present the first complete mitochondrial genome for a member of this family, *Aprasia parapulchella*, from Australia. The mitochondrial genome of *A. parapulchella* is 16 528 base pairs long and contains 13 protein-coding genes, 22 tRNA genes, two rRNA genes and the control region, conforming to the typical vertebrate gene order. The overall mitochondrial nucleotide composition is 31.7% A, 24.5% T, 30.5% C and 13.2% G. This corresponds to a total A+T content of 56.3%, which is similar to that of other squamate lizard genomes.

Received 27 October 2014, accepted 7 April 2015, published online 28 April 2015

Introduction

The Pygopodidae are endemic to the Australo-Papuan region (Kluge 1976; Jennings *et al*. 2003). It is now well understood from both morphological (Kluge 1976; Kluge 1987) and molecular data (Donnellan *et al*. 1999; Jennings *et al*. 2003) that the Pygopodidae are closely allied to the Gekkonidae. The two groups share many unique features and the pygopodids are generally thought of as legless geckos. Phylogenetic studies based on nuclear and mitochondrial loci have placed the Pygopodidae as a monophyletic lineage within Gekkota; however, the exact relationship of the Pygopodidae to other gecko lineages is unclear. Based on analysis of two genes (the mitochondrial 12S rRNA and the nuclear c-mos), Donnellan *et al*. (1999) described the Pygopodidae as a monophyletic sister lineage to all Diplodactylinae. More recently, Mulcahy *et al*. (2012) and Wiens *et al*. (2012), using 25 and 44 nuclear loci respectively, identified Pygopodidae as a sister lineage to Diplodactylidae, with Carphodactylidae as a sister lineage to this group. In contrast, Pyron *et al*. (2013) analysed 12 nuclear and five mitochondrial genes and a much greater number of species, and described the Pygopodidae as a sister group to the Carphodactylidae, with this clade as a sister group to Diplodactylidae.

Complete mitochondrial genomes have been used to estimate phylogenetic relationships among taxa, as well as for studies of population genetics and molecular evolution (Alverson *et al*. 2010; Kan *et al*. 2010; Zhang *et al*. 2013). However, while 117 whole mitochondrial genomes are available for squamate reptiles as a whole, and 11 are available for geckos, there has been no published mitochondrial genome for any pygopodid. Here we present the complete mitochondrial genome for a member of the Pygopodidae, the pink-tailed worm lizard, *Aprasia parapulchella*.

Materials and methods

DNA extraction and sequencing

A genomic library was constructed from the tail snip of one individual of *A. parapulchella* collected within the Australian Capital Territory. Permission for sampling was granted by the ACT Government (Permit LT2012587) and the Committee for Ethics in Animal Experimentation at the University of Canberra (CEAE 11/12). DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Venlo, Limburg, Netherlands), according to the manufacturer's instructions. Approximately 5 µg of DNA was sent to the Australian Genome Research Facility Ltd (AGRF, Brisbane, Australia) for genomic library preparation and sequencing. We used standard methods for single-end 454 shotgun sequencing from non-model species (Gardner *et al*. 2011; Meglécz *et al*. 2012). DNA quality was determined using Quant-iT Fluorometry Assay, PicoGreen dsDNA quantification (Invitrogen, Grand Island, NY, USA). The DNA was fragmented by nebulisation, then size selected for fragments >350 bp. Adaptor ligation and library purification followed standard Roche (Basel, Switzerland) protocols, except that MID-tagged oligo adapters, generated as per standard Roche procedures and

ordered from IDT (Iowa, USA), were used to enable samples to be multiplexed without physical masking of the picotiter plate. The *A. parapulchella* library was sequenced using 1/8 of the capacity of a GS-FLX titanium run. Following sequencing, reads generated from the *A. parapulchella* library were extracted on the basis of the appropriate MID-tag. The Roche sffinfo program was used to generate raw FASTA files, trimmed of all sequencing adaptors and trimmed according to the default quality parameters.

Sequence assembly and annotation

De novo sequence assembly was conducted by AGRF using Newbler software (Roche) with the recommended default settings. Sequence and assembly quality of contigs were subsequently investigated using Geneious ver. 7 (Biomatters, Auckland, New Zealand, http://www.geneious.com: Kearse *et al*. 2012). The largest contig, at 16 528 bp in length, was putatively identified as the *A. parapulchella* mitochondrial genome. This sequence was queried against the NCBI nucleotide database using BLASTN 2.2.3 (Zhang *et al*. 2000) optimised for somewhat similar sequences, which confirmed the mitochondrial origin of our largest *Aprasia parapulchella* contig. To localise the approximate position of the *A. parapulchella* mitochondrial genes, we aligned the putative *A. parapulchella* mitochondrial contig with the *Coleonyx variegatus* (western banded gecko) mitochondrial genome (accession AB114446) using the 'map to reference' function with default parameters in Geneious ver. 7, with *C. variegatus* set as the reference sequence. We then used MITOS (Bernt *et al*. 2013) to conduct a *de novo* annotation of the *A. parapulchella* mitochondrial genome, using the general vertebrate mitochondrial code. We also used tRNAscanSE ver.1.21 (Lowe and Eddy 1997) to confirm the boundaries and orientation of the tRNAs. Genes and tRNAs were annotated onto the mitochondrial contig in Geneious and manually adjusted as needed to incorporate start and stop codons for the coding genes.

Results

Sequencing

The *Aprasia* shotgun library produced 156 852 sequence reads that passed quality filters, from which 1965 contigs of variable length were generated. The consensus sequence of the largest of these contigs was 16 528 bp in length. This contig was constructed from 639 reads with sizes ranging from 51 bp to 569 bp (mean = 379 bp). The mean coverage of each base in the consensus sequence was 14.5, with a standard deviation of 5.1 and a range of 1–34. Mean pairwise identity along the length of the contig was 98.5%. Sequence quality was high, with 99.5% of bases scored at Q40 or higher. BLAST analysis confirmed the mitochondrial origin of this contig. At 1217 bp, the BLAST hit with the highest identity to the query (96%) covered only 7% of the contig length, but originated from the *Aprasia pseudopulchella* mitochondrial ND2 gene (accession AY134577). The BLAST hit with the highest score, which had 93% coverage of the query sequence and 73% identity, was the *Coleonyx variegatus* complete mitochondrial genome (accession AB114446). The mitochondrial genome of *A. parapulchella* can be found in GenBank as accession KJ004564.

Genome organisation

The mitochondrial genome of *A. parapulchella* is 16 528 bp long (Fig. 1; Table S1 in the Supplementary Material), which is well within the size range of mitochondrial genomes of the phylogenetically close gekkotans (mean 16 814 bp), but shorter than the mitochondrial genomes of squamate lizards (Sauria) in general (mean 17 240 bp: Table S2 in the Supplementary Material). In contrast to some other lizards (Macey *et al*. 2004), in *A. parapulchella* the 13 protein-coding genes, 22 tRNA genes, two rRNA genes and the control region conform to the typical vertebrate gene order (Boore 1999). The majority of the mitochondrial genes are encoded on the H-strand, except for ND6 and eight of the tRNAs (Table S1). The genome has 12 intergenic regions varying between 1 and 27 bp in length (totalling 65 bp, excluding the control region) and 14 regions with overlapping genes (1–17 bp in length), including 4 pairs of overlapping protein-coding genes. The largest intergenic region, excluding the 1113-bp control-region, is between $tRNA^{Asn}$ and $tRNA^{Cys}$ and cannot be assigned to either of the tRNAs. This gap is common among squamates (Table S2) and is referred to as the putative L-strand origin (Seutin *et al*. 1994). All tRNAs can be folded into stem and loop structures. The overall mitochondrial nucleotide composition is 31.7% A, 24.5% T, 30.5% C and 13.2% G. This corresponds to a total A+T content of 56.3%, which is similar to that of other squamate lizard genomes (average 58.8%: Table S2).

The protein-coding genes

Ten of the 13 protein-coding genes in *A. parapulchella* start with the standard initiation codon ATG. The initiation codons for ND2 and ND5 are ATA and ATT respectively, which are both common in vertebrate mitochondrial genomes. The start codon for COXI is GTG (nucleotides 5332–5334), which has been identified as the start codon for this gene in a range of other reptiles, birds and fish (Seutin *et al*. 1994). However, there is also an ATA codon (nucleotides 5335–5337) immediately downstream of this GTG, meaning that there is some uncertainty regarding the COXI start codon. There is some ambiguity in the length of the coding region for the ND5 gene. Based on the position of $tRNA^{Leu}$ and through comparison with the *Coleonyx* mitochondrial genome, we suggest that the *ND5* start codon begins at nucleotide 11 757, meaning that the coding region is 1800 bp in length. A possible alternative start codon identified by MITOS at nucleotide 11 826 would produce a coding region 1731 bp in length, with a longer intergenic gap of 85 bp between $tRNA^{Leu}$ and *ND5*.

The termination codons are all complete, being TAA (8), TAG (3), AGA and AGG. With the exception of the ND5–ND6 pair, all adjoining protein-coding genes overlap with a minimum of 1 bp and a maximum of 10 bp. The total number of codons in the 13 protein-coding genes is 3780, including the initiation and stop codons. Detailed information on the protein-coding genes in *A. parapulchella* can be seen in Table S1.

Discussion

Vertebrate mitochondrial genomes were viewed as structurally conserved relative to many other taxonomic groups, which display much greater variation in gene content and order.

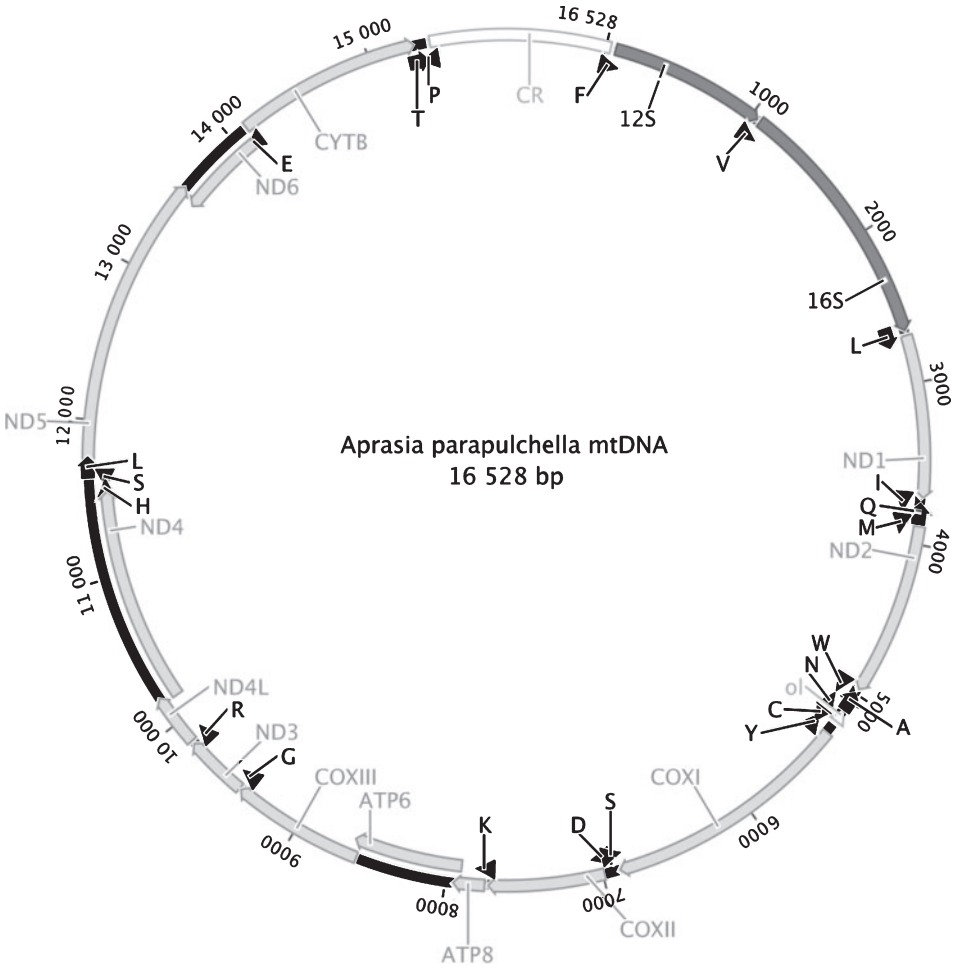

Fig. 1. Gene map of the *A. parapulchella* mitochondrial genome. Transfer RNAs are shown as black arrows, denoted by a single letter standing for the corresponding amino acid code. The rRNAs 12s and 16s are in dark grey. Abbreviations for the protein-coding genes (light grey) are as follows: ND1–6, NADH dehydrogenase subunits 1–6; COI–III, cytochrome oxidase subunits I–III; ATP6 and ATP8, ATPase subunits 6 and 8; CYTB, cytochrome *b*. The two longest non-coding regions, the control region (CR) and the putative origin of the light-strand replication (ol), are shown in white.

However, it is now thought that this perceived stability in part reflects uneven sampling coverage of many vertebrate taxa (Kumazawa *et al*. 1998; Gissi *et al.* 2008). As more mitochondrial genome sequences become available, there is increasing evidence that non-standard mitogenome structure is common in squamate lineages. For example, mitochondrial genome rearrangements and gene duplications have been observed in amphisbaenians, including convergent rearrangements that appear to have evolved independently within the same gene region in two distinct lineages (Macey *et al*. 2004). Lineage-specific rearrangements have also been observed in acrodontan mitogenomes (Okajima and Kumazawa 2010) and the Komodo dragon (Kumazawa and Endo 2004) and some species of snake have duplicated control regions (Kumazawa *et al*. 1996; Kumazawa *et al*. 1998). Other squamate rearrangements of the mitogenome include duplications of the tRNAs T and P, which occur twice in Amphisbaenia and once in Scincomorpha (Table S2). Within the Gekkonidae, species such as *Gekko gecko* (Zhou *et al*. 2006) and *Teratoscincus keyserlingii* (Macey *et al*. 2005) display the typical vertebrate mitochondrial gene order, but large tandem duplications of genes have been observed in some asexual (but not sexual) lineages of the gecko *Heteronotia binoei* (Moritz 1991; Zevering *et al*. 1991; Fujita *et al*. 2007). The Pygopodidae are closely related to the Gekkonidae, but their mitogenome structure has not previously been studied, despite the potential utility of such data for improving phylogenetic placement of this group. Here we present the mitochondrial genome of *Aprasia parapulchella*, the first pygopodid mitochondrial genome to be published. We demonstrate that this conforms to the typical vertebrate gene order (Boore 1999), and includes 13 protein-coding genes, 22 tRNA genes, two rRNA genes and a single control region.

Acknowledgements

We thank Will Osborne for help in collecting *A. parapulchella* samples and Tony Gamble and two anonymous reviewers for their comments on this

manuscript. We also thank those who have contributed whole mitochondrial genome data to GenBank for making this information available for us to use. This study was funded by the Academy of Finland and the ACT Government.

References

Alverson, A. J., Wei, X., Rice, D. W., Stern, D. B., Barry, K., and Palmer, J. D. (2010). Insights into the evolution of plant mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution* **27**, 1436–1448. doi:10.1093/molbev/msq029

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., and Stadler, P. F. (2013). MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69**, 313–319. doi:10.1016/j.ympev.2012.08.023

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research* **27**, 1767–1780. doi:10.1093/nar/27.8.1767

Donnellan, S. C., Hutchinson, M. N., and Saint, K. M. (1999). Molecular evidence for the phylogeny of Australian gekkonoid lizards. *Biological Journal of the Linnean Society* **67**, 97–118. doi:10.1111/j.1095-8312.1999.tb01932.x

Fujita, M. K., Boore, J. L., and Moritz, C. (2007). Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (*Heteronotia binoei*; Squamata, Gekkonidae). *Molecular Biology and Evolution* **24**, 2775–2786. doi:10.1093/molbev/msm212

Gardner, M. G., Fitch, A. J., Bertozzi, T., and Lowe, A. J. (2011). Rise of the machines – recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources* **11**, 1093–1101. doi:10.1111/j.1755-0998.2011.03037.x

Gissi, C, Iannelli, F, and Pesole, G (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity* **101**, 301–320. doi:10.1038/hdy.2008.62

Jennings, W. B., Pianka, E. R., and Donnellan, S. (2003). Systematics of the lizard family Pygopodidae with implications for the diversification of Australian temperate biotas. *Systematic Biology* **52**, 757–780. doi:10.1093/sysbio/52.6.757

Kan, X.-Z., Yang, J.-K., Li, X.-F., Chen, L., Lei, Z.-P., Wang, M., Qian, C.-J., Gao, H., and Yang, Z.-Y. (2010). Phylogeny of major lineages of galliform birds (Aves: Galliformes) based on complete mitochondrial genomes. *Genetics and Molecular Research* **9**, 1625–1633. doi:10.4238/vol9-3gmr898

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., and Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649. doi:10.1093/bioinformatics/bts199

Kluge, A. G. (1976). Phylogenetic relationships in the lizard family Pygopodidae: an evaluation of theory, methods and data. *Miscellaneous Publications, Museum of Zoology, University of Michigan* **152**, 1–7.

Kluge, A. G. (1987). Cladistic relationships in the Gekkonidae (Squamata, Sauria). *Miscellaneous Publications, Museum of Zoology, University of Michigan* **173**, 1–54.

Kumazawa, Y., and Endo, H. (2004). Mitochondrial genome of the Komodo dragon: efficient sequencing method with reptile-oriented primers and novel gene rearrangements. *DNA Research* **11**, 115–125. doi:10.1093/dnares/11.2.115

Kumazawa, Y., Ota, H., Nishida, M., and Ozawa, T. (1996). Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. *Molecular Biology and Evolution* **13**, 1242–1254. doi:10.1093/oxfordjournals.molbev.a025690

Kumazawa, Y., Ota, H., Nishida, M., and Ozawa, T. (1998). The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. *Genetics* **150**, 313–329.

Lowe, T. M., and Eddy, S. R. (1997). tRNA-scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**, 955–964. doi:10.1093/nar/25.5.0955

Macey, J. R., Papenfuss, T. J., Kuehl, J. V., Fourcade, H. M., and Boore, J. L. (2004). Phylogenetic relationships among Amphisbaenian reptiles based on complete mitochondrial genomic sequences. *Molecular Phylogenetics and Evolution* **33**, 22–31. doi:10.1016/j.ympev.2004.05.003

Macey, J. R., Fong, J. J., Kuehl, J. V., Shafiei, S., Ananjeva, N. B., Papenfuss, T. J., and Boore, J. L. (2005). The complete mitochondrial genome of a gecko and the phylogenetic position of the Middle Eastern *Teratoscincus keyserlingii*. *Molecular Phylogenetics and Evolution* **36**, 188–193. doi:10.1016/j.ympev.2005.03.025

Meglécz, E., Nève, G., Biffin, E., and Gardner, M. G. (2012). Breakdown of phylogenetic signal: a survey of microsatellite densities in 454 shotgun sequences from 154 non model eukaryote species. *PLoS One* **7**, e40861. doi:10.1371/journal.pone.0040861

Moritz, C. (1991). The origin and evolution of parthenogenesis in *Heteronotia binoei* (Gekkonidae): evidence for recent and localized origins of widespread clones. *Genetics* **129**, 211–219.

Mulcahy, D. G., Noonan, B. P., Moss, T., Townsend, T. M., Reeder, T. W., Sites, J. W. Jr, and Wiens, J. J. (2012). Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. *Molecular Phylogenetics and Evolution* **65**, 974–991. doi:10.1016/j.ympev.2012.08.018

Okajima, Y., and Kumazawa, Y. (2010). Mitochondrial genomes of acrodont lizards: timing of gene rearrangements and phylogenetic and biogeographic implications. *BMC Evolutionary Biology* **10**, 141. doi:10.1186/1471-2148-10-141

Pyron, R. A., Burbrink, F. T., and Wiens, J. J. (2013). A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology* **13**, 93. doi:10.1186/1471-2148-13-93

Seutin, G., Lang, B. F., Mindell, D. P., and Morais, R. (1994). Evolution of the WANCY region in amniote mitochondrial DNA. *Molecular Biology and Evolution* **11**, 329–334.

Wiens, J. J., Hutter, C. R., Mulcahy, D. G., Noonan, B. P., Townsend, T. M., Sites, J. W. Jr, and Reeder, T. W. (2012). Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. *Biology Letters* **8**, 1043–1046. doi:10.1098/rsbl.2012.0703

Zevering, C. E., Moritz, C., Heideman, A., and Sturm, R. A. (1991). Parallel origins of duplications and the formation of pseudogenes in mitochondrial DNA from parthenogenetic lizards (*Heteronotia binoei*; Gekkonidae). *Journal of Molecular Evolution* **33**, 431–441. doi:10.1007/BF02103135

Zhang, P., Liang, D., Mao, R.-L., Hillis, D. M., Wake, D. B., and Cannatella, D. C. (2013). Efficient sequencing of Anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. *Molecular Biology and Evolution* **30**, 1899–1915. doi:10.1093/molbev/mst091

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology* **7**, 203–214. doi:10.1089/10665270050081478

Zhou, K., Li, H., Han, D., Bauer, A. M., and Feng, J. (2006). The complete mitochondrial genome of *Gekko gecko* (Reptilia: Gekkonidae) and support for the monophyly of Sauria including Amphisbaenia. *Molecular Phylogenetics and Evolution* **40**, 887–892. doi:10.1016/j.ympev.2006.04.003

Handling Editor: Paul Cooper