Timing of a mtDNA gene rearrangement and intercontinental dispersal of varanid lizards

Sayed A. M. Amer^{1,2*} and Yoshinori Kumazawa^{2,3}

¹Zoology Department, Faculty of Science, Cairo University, Giza, Egypt

²Graduate School of Science, Nagoya University, Nagoya
464-8602, Japan

³Graduate School of Natural Sciences, Nagoya City
University, Nagoya, Japan

(Received 9 January 2008, accepted 27 May 2008)

The mitochondrial genomes of the Komodo monitor (*Varanus komodoensis*) and the Nile monitor (*V. niloticus*) were previously shown to have an extensive gene rearrangement. Here, we show that this gene arrangement widely occurs in varanid taxa originated from Africa, Asia and Australasia. Based on phylogenetic relationships of the varanids constructed using mitochondrial DNA sequences encoding the NADH dehydrogenase subunit 2 gene and seven flanking tRNA genes, we estimated their divergence times by the Bayesian method without assuming the molecular clock. The results suggested that the mitochondrial DNA gene rearrangement took place once in an ancestral varanid lineage in the Paleocene or earlier. Our results are more consistent with Cenozoic over-water dispersal of Southeast Asian varanids across the Indonesian Archipelago rather than the Cretaceous Gondwanan vicariance for the origin of Australasian varanids.

Key words: divergence time, gene organization, mitochondrial genome, over-water dispersal, Varanidae

INTRODUCTION

Varanidae is an anguimorph lizard family inhabiting Afro-Arabia, west to southeastern Asia, Indonesian Archipelago, Papua New Guinea and Australia (Fuller et al., 1998; Ast, 2001). Approximately 50 extant species are recognized in Varanidae as one genus *Varanus* or two subgenera *Varanus* and *Odatia*. Some species are threatened for extinction and are listed in the *Convention on International Trade in Endangered Species* (CITES).

Morphological (Böhme, 1988; Estes et al., 1988), biochemical (Holmes et al., 1975) and karyotypic (King and King, 1975) studies attempted to resolve phylogeny and biogeography of these remarkable lizards. Molecular studies based on the nucleotide sequences of the mitochondrial DNAs (mtDNAs) were also carried out (Baverstock et al., 1993; Fuller et al., 1998; Ast, 2001; Schulte et al., 2003; Fitch et al., 2006). The monophyly of the Varanidae and some of its sublineages was concluded by these studies using different tree-building methods.

The complete mitochondrial DNA (mtDNA) sequences for the Indonesian *V. komodoensis* (Kumazawa and Endo,

2004) and the African *V. niloticus* (Kumazawa, 2007) were found to have some unique features in their gene organizations. Genes between the NADH dehydrogenase subunit 6 (*ND6*) gene and proline tRNA gene were extensively shuffled and the control region was duplicated (Fig. 1). This finding prompted us to examine phylogenetic distribution of the unique gene organization within Varanidae in order to better understand evolutionary processes of the mitochondrial gene arrangements.

On the other hands, the historical biogeography of Varanidae is still under debates. While most of the above-mentioned studies supported Asian origin of varanids and their dispersal to Australasia, a recent molecular study has postulated the Cretaceous Gondwana vicariance for the origin of Australasian varanids (Schulte et al., 2003). Schulte et al. (2003) used an approximately 2 kbp segment of mtDNAs to construct the phylogeny of varanids and estimated their divergence times. However, their dating results have been questioned by some investigators due to potential methodological problems in their analyses (Hugall and Lee, 2004; Amer and Kumazawa, 2005). In the present study, we also attempted to reevaluate previous hypotheses of varanid divergences using the relaxed-clock Bayesian estimation of divergence times.

Edited by Etsuko Matsuura

* Corresponding author. E-mail: yasser92us@yahoo.com

MATERIALS AND METHODS

Samples, DNA extraction and amplification Extracted DNA samples of varanid taxa from the Afro-Arabia (*V. exanthematicus*), Asia (*V. salvator*, *V. olivaceus*, *V. rudicollis* and *V. dumerilli*), New Guinea (*V. jobiensis*) and Australia (*V. storri*, *V. gilleni* and *V. acanthurus*) were provided from the herpetological collection of Yoshinori Kumazawa. DNA samples were prepared

from dead wild individuals of *V. griseus* and *V. niloticus* of Egypt using a standard procedure (Kocher et al., 1989). The mtDNA segment between the *ND5* and the *ND6* genes was amplified with the primers rND5-2L and rND6-4H for various varanid taxa (see Fig. 1 and Table 1 for their sequences and locations). The primers rThr-2L and rND6-4H were also used to confirm the rearranged gene organization in this region. The polymerase chain reaction (PCR) was carried out as described in Amer and

Fig. 1. Relative arrangements of several genes and control regions (CRs) in the typical gene organization of vertebrate mtDNAs (A) and in the rearranged organization of the Komodo and Nile monitors (B). Correspondence of genes between the two organizations is shown with solid or dotted lines. The arrows below Fig. 1B refer to the fragments that were sequenced for *V. exanthematicus* (1), *V. rudicollis* (2) and *V. acanthurus* (3). The location and orientation of primers used for the amplification and sequencing are shown with their numbers in both A and B (see their sequences in Table 1). Genes which are encoded by the heavy strand are shown above the diagram and those which are encoded by the light strand are shown below. Gene abbreviations used are: *ND5* and *ND6* (NADH dehydrogenase subunits 5 and 6); *cytb* (cytochrome *b*); *CR* (control region); and *E*, *T*, *P* and *F* (tRNA genes specifying for glutamic acid, threonine, proline and phenylalanine, respectively).

Table 1. Sequences of the forward (L) and the reverse (H) primers used in PCR amplification and sequencing. Mixed bases are referred to as: R (A, G), Y (C, T), K (G, T), W (A, T), M (A, C) and D (G, A, T)

No	Name	Sequence (5' to 3')	Source
1	rND5-2L	GAACARGACATYCGAAAAATRGG	Kumazawa and Endo (2004)
2	rND6-4H	ATGTTAGTGGTDTTGTCKTATTC	Kumazawa and Endo (2004)
3	rGlu-1L	GAAAAACCRCCGTGTWATTCAACTA	Kumazawa and Endo (2004)
4	rThr-2L	YAAAGCMTTGRTCTTGTA	Kumazawa and Endo (2004)
5	rThr-1H	TCCCATCTTTGGYTTACAAG	Dong and Kumazawa (2005)
6	rPro-1L	CATCTCTAGYCCCCAAAAC	Amer and Kumazawa (2005)
7	rPro-1H	TTAAAATKCYRGYTTTGGGG	Kumazawa and Endo (2004)
8	rPhe-2H	GCATYTTCAGTGCCGTGCTT	Amer and Kumazawa (2005)
9	Vnil-ND5-2L	GAACAAGACCTGCGTAACATAGG	this study
10	V-ND6-1H	GGGATGGTYGTWGTWTTTGC	this study

Kumazawa (2005). The amplified products were electrophoresed on a 1% agarose gel, stained with ethidium bromide and photographed under the UV light. The amplified products were thereafter used as templates for sequencing with the same amplification primers and other appropriate internal primers shown in Fig. 1 and Table 1.

Tree reconstruction and divergence time estimation We took varanid nucleotide sequences of the mitochondrial *ND2* region (the complete NADH dehydrogenase subunit 2 gene plus several flanking tRNA genes) from the DDBJ database and aligned them with those of additional outgroup taxa (1 fish, 2 mammals, 2 amphibians, 2 birds, 2 crocodilians, 2 geckos and 2 non-varanid anguimorphans). The alignment was carried out by using the DNASIS 3.5 (Hitachi) and MacClade 4.03 (Sinauer Associates, Inc.) with manual adjustments. All unalignable and gap-containing sites were excluded from analyses so that 1433 sites were left for phylogenetic analyses. We primarily used Bayesian analyses for constructing phylogenetic trees and confirmed the robustness of phylogenetic relationships by using the maximum-likelihood method with 300 bootstrap replications. Bayesian trees were constructed using MrBayes version 3.0 (Huelsenbeck and Ronquist, 2001) using the GTR + I + G model for base substitution model. For more details in conditions for the Bayesian and maximum-likelihood bootstrap analyses, as well as the discussion in the utility of mtDNA sequences in deep branch phylogenetics, see Amer and Kumazawa (2007) and Kumazawa (2007).

The varanid relationships thus obtained was used for the divergence time estimation. The multidistribute package (Thorne et al., 1998) implements the Bayesian method for divergence time estimation without assuming the molecular clock (i.e., by allowing the rate variation along branches). Upper and/or lower time constraints at selected nodes were set for the Bayesian Markov chain Monte Carlo (MCMC) processes to estimate posterior distributions of divergence times and relative rates at ingroup nodes. We first used PAML (Yang, 1997) to optimize parameters for the F84 base substitution model and the gamma distribution for 8 categories to account for the site-heterogeneity. The *estbranches/multidivtime* programs were then used to estimate divergence times using five fossil-based reliable time constraints that had been used for previous studies successfully (see Kumazawa, 2007 for details of the paleontological evidence).

RESULTS AND DISCUSSION

Rearranged mtDNA gene order for Varanidae The control region duplication and the extensive gene shuffling at the vicinity of the *ND6* and proline tRNA genes were examined for the available varanid taxa. To do so,

PCR amplifications were conducted using two pairs of primers (rThr-2L + rND6-4H and rND5-2L + rND6-4H). The first primer pair gives amplification products only when threonine tRNA gene precedes the *ND6* gene in the rearranged gene organization (Fig. 1). Almost all varanid taxa examined gave rise to amplification products (Fig. 2A), suggesting a wide occurrence of the gene rearrangement in Varanidae. The size of the amplified fragments varied from 1.5 to 1.8 kbp for the examined varanid taxa possibly due to the length variation in the inserted control region (see Fig. 1B). For *V. niloticus*, the amplification product was not clearly detected (Fig. 2a). However, this turned out to be due to mismatching of the primers to the corresponding region of *V. niloticus* mtDNA (data not shown). The mtDNA of *V. niloticus* had been completely sequenced (Kumazawa, 2007) and does possess the gene rearrangement. We therefore designed a new primer V-ND6-1H (see Table 1) based on this sequence and obtained the amplified product of the expected size (Fig. 2C).

With the second primer pair, the size of amplification products was approximately 3 kbp for almost all varanids (Fig. 2B). This is consistent with the rearranged organization in which multiple genes are inserted between the *ND5* and the *ND6* genes (Fig. 1). In the typical organization of vertebrate mitochondrial genomes, the expected product size is approximately 1.2 kbp. Again, two African taxa (*V. griseus* and *V. niloticus*) did not provide amplification products of the reasonable size and this was interpreted to be due to the mismatching of the primers to the targets based on the complete mtDNA sequence of *V. niloticus* (data not shown). We thus designed another species-specific primer Vnil-ND5-2L (see Table 1) and obtained the amplified product of the expected size (Fig. 2C).

We partially sequenced the amplified products for the African *V. exanthematicus*, the Asian *V. rudicollis* and the Australian *V. acanthurus* (accession numbers AB301968–AB301971). *V. exanthematicus* sequence showed the gene arrangement of *cytb*, *tRNA^{Thr}*, *CR1*, *tRNA^{Glu}* and the *ND6* (Fig. 1B). The *V. rudicollis* sequence retained the gene arrangement of *tRNA^{Thr}*, *CR1*, *tRNA^{Glu}* and the *ND6* while *V. acanthurus* exhibited the arrangement of *tRNA^{Thr}* and *CR1*. The gene organizations found for these taxa are consistent with those of *V. komodoensis* and *V. niloticus* (Fig. 1B). Retention of the control region in the canonical position between the proline and phenylalanine tRNA genes was also confirmed by amplifying and partially sequencing this region using the primers rPro-1L and rPhe-2H (data not shown).

Timing of the gene rearrangement The phylogenetic relationships for Varanidae were constructed using the Bayesian method (Huelsenbeck and Ronquist, 2001). The resultant tree (Fig. 3) was basically consistent with those by previous authors (e.g., Ast, 2001; Schulte et al.,

Fig. 2. PCR assays for the gene rearrangements in varanid mtDNAs. Shown here are agarose gel electrophoresis profiles of the amplified products using the primers rThr-2L and rND6-4H (A) and rND5-2L and rND6-4H (B). The amplification profiles in this Fig. are consistent with the wide occurrence of the rearranged gene organization within Varanidae. The expected length of the DNA fragment in case where the primers 1 and 2 are used for amplification is 1.2 kbp for the typical gene organization. Unsuccessful amplification for *V. niloticus* in both assays (A and B) was presumably due to primer mismatching because the use of species-specific primers led to successful amplification for *V. niloticus* (see text and C below). The amplified fragments for *V. niloticus* using the primers rThr-2L and V-ND6-1H (1) and Vnil-ND5-2L and V-ND6-1H (2) are shown in C.

2003). Afro-Arabian taxa diverged most basally. One of two major lineages for the remaining taxa consists of exclusively Australasian taxa, whereas the other major lineage consists of taxa which are distributed in South/Southeast Asia and Pacific islands located in the eastern side of Wallace's Line. Based on the topological relationship of Fig. 3, the divergence times were estimated without assuming the molecular clock.

Figure 3 showed the Paleocene [approximately 60 million years ago (MYA)] or earlier origin of Varanidae as indicated by a basal divergence between the African clade

and the Asia-Australasian clade around 60 MYA. Because the rearranged gene organization was found in diverse varanid lineages (see Figs. 1 and 3 for taxa in which sequence evidence exists, as well as taxa used for the PCR assay in Fig. 2) and because mtDNA gene organizations of the anguimorph outgroups (the Gila monster and the alligator lizard) (Kumazawa, 2004, 2007) are identical to the typical vertebrate organization (Fig. 1A), it is likely that the gene rearrangement took place once in an ancestral varanid lineage in the Paleocene or earlier. This is another example of non-homoplasious

Fig. 3. The phylogeny and divergence times for varanids. A Bayesian tree with posterior probabilities at nodes was constructed as described in Materials and Methods using 1433 bp of the mtDNA segment of the *ND2* gene and seven flanking tRNA genes. The underlined values at nodes denote bootstrap proportions obtained by 300 replications of the maximum-likelihood analysis. Values in the open squares are five fossil-based reliable time constraints used for divergence time estimation (see Materials and methods for details). Estimated times at some relevant nodes are shown in means of millions of years ago $\pm$ standard deviations. Taxa with an asterisk were shown to have the rearranged gene organization by sequencing their complete or partial mtDNA sequences (Kumazawa and Endo, 2004, Kumazawa, 2007; this study). GenBank accession numbers for the nucleotide sequences used are: *Varanus acanthurus* (AF407488), *Varanus dumerilli* (AF407494), *Varanus gilleni* (AF407499), *Varanus griseus* (AF407503), *Varanus indicus* (AF407504), *Varanus jobiensis* (AF407507), *Varanus olivaceus* (AF407515), *Varanus komodoensis* (AF407510), *Varanus niloticus* (AF407514), *Varanus rudicollis* (AF407521), *Varanus varius* (AF407534), alligator lizard (AB080273), Gila monster (AB167711), banded gecko (AB114446), leopard gecko (AB308468), chicken (X52392), indigo bird (AF090341), alligator (Y13113), caiman (AJ404872), cow (J01394), platypus (X83427), toad (M10217), caecilian (AF154051), and coelacanth (U82228).

nature of mitochondrial gene rearrangements as has been seen for other vertebrates (see, e.g., Amer and Kumazawa, 2005, 2007).

Historical biogeography for Varanidae The historical biogeography of the varanid lizards has been traditionally explained by postulating Asian rather than Gondwanan origin of Varanidae. This is mainly because the earliest known varanoid fossils came from Laurasia-origin continents (Eurasia and North America) and because the majority of living varanid subgenera are found in Eurasia (Hecht, 1975; Fuller et al., 1998). Ancestors of extant varanids appeared in central Asia during Cretaceous times and they migrated throughout Europe and South Asia approximately 50 MYA during the Eocene (Ciofi, 1999). Based on morphological evidence, Böhme (1988) contributed much with a similar understanding of the rise and evolution of *Varanus*. King and King (1975) and Baverstock et al. (1993) used chromosomal structures and DNA sequences, respectively, to conclude that the genus originated between 25 and 40 MYA in Asia.

The divergence of the most basal varanid lineages around the Paleocene (Fig. 3) supports intercontinental dispersal of varanids, rather than their vicariant origin linked with continental drifts because major continental fragmentation involving Africa, Eurasia and Australia had already taken place in the early-mid Cretaceous (The Plates Project, 1998). Africa had been long isolated from other continents since approximately 100 MYA and it connected with Eurasia in the early Miocene at approximately 18 MYA (Rögl, 1998).

If varanids had originated from Gondwanaland-origin continents (Africa or Australia), it would have been quite difficult that they migrated to Eurasia before mid-Tertiary without assuming long-distance transmarine dispersal. However, varanid fossils have been found from Late Cretaceous deposits of Laurasia-origin continents (Benton, 1993). In addition, there is strong fossil evidence to support that Africa was colonized by *Varanus* within the last 20 million years (Clos, 1995).

Thus, *Varanus* probably originated in Eurasia in the Paleocene or earlier (> 60 MYA) and one of the basal lineages migrated westward and colonized Afro- Arabia after

Africa collided with Eurasia in the early Miocene. Another varanid lineage (or lineages) dispersed southward and colonized Australasia in the Eocene or more recently (Fig. 3). In order to reveal when and how many varanid lineages actually dispersed to Australasia, more thorough investigations with respect to their phylogenetic relationships and divergence times may be necessary.

A recent molecular study using mtDNA sequences (Schulte et al., 2003) proposed a deep (110–120 MYA) divergence between Southeast Asian and Australasian varanid lizards, which was consistent with ancient Gondwanan vicariance. However, some technical problems in their dating methods have been pointed out mostly due to the heavy reliance on the nonparametric rate smoothing for correcting rate variations between lineages and the use of discrete rates of sequence changes calibrated with outgroup taxa without explicitly treating rate variations between lineages (Hugall and Lee, 2004; Amer and Kumazawa, 2005). Based on the results presented in this study, we consider that divergence times between Australasian and Asian varanids were considerably overestimated by Schulte et al. (2003) and that the Gondwanan vicariance thus applied to these divergences was not probably the case. We favor the traditional view of the Cenozoic over-water dispersal for the origin of Australasian varanids. An independent fossil evidence (> 15 MYA: the earliest Australian varanid fossils; Hecht, 1975) is also consistent with our latest molecular dating.

We thank Mr. K. Yagi and Remix Peponi Co. for providing some animal samples. We also thank Mr. Y. Okajima for his kind cooperations in some of our experiments. Our gratitude is extended to Daiko Foundation (the Visiting Fellowship for Foreign Scholars number 10093) and the Japanese government (grant numbers 14902202 and 17570076) for financial supports.

REFERENCES

- Amer, S. A. M., and Kumazawa, Y. (2005) Complete mitochondrial genome of *Pogona vitticeps* (Reptilia; Agamidae): control region duplication and the origin of Australasian agamids. *Gene* **346**, 249–257.
- Amer, S. A. M., and Kumazawa, Y. (2007) The mitochondrial genome of the lizard *Calotes versicolor* and a novel gene inversion in South Asian draconine agamids. *Mol. Biol. Evol.* **24**, 1330–1339.
- Ast, J. C. (2001) Mitochondrial DNA evidence and evolution in Varanoidea (Squamata). *Cladistics* **17**, 211–226.
- Baverstock, P. R., King, D., King, M., Birrell, J., and Krieg, M. (1993) The evolution of species of the Varanidae: Microcomplement fixation analysis of serum albumins. *Aust. J. Zool.* **41**, 621–638.
- Benton, M. J. (1993) The fossil record 2. Chapman and Hall, London.
- Böhme, W. (1988) Zur Genitalmorphologie der Sauria: Funktionelle und stammesgeschichtliche Aspekte. *Bonn Zool. Monogr.* **27**, 1–176.
- Ciofi, C. (1999) “The Komodo Dragon.” *Sci. Amer.* **280**, 84–91.
- Clos, L. M. (1995) A new species of *Varanus* (Reptilia: Sauria) from the Miocene of Kenya. *J. Vertebr. Paleontol.* **15**, 254–267.
- Dong, S., and Kumazawa, Y. (2005) Complete mitochondrial DNA sequences of six snakes: phylogenetic relationships and molecular evolution of genomic features. *J. Mol. Evol.* **61**, 12–22.
- Estes, R., de Queiroz, K., and Gauthier, J. (1988) Phylogenetic relationships within Squamata. In: *Phylogenetic Relationships of the Lizard Families* (eds.: R. Estes, and G.K. Pregill), pp. 119–282. Stanford University Press, CA.
- Fitch, A. J., Goodman, A. E., and Donnellan, S. C. (2006) A molecular phylogeny of the Australian monitor lizards (Squamata: Varanidae) inferred from mitochondrial DNA sequences. *Aust. J. Zool.* **54**, 253–269.
- Fuller, S., Baverstock, P., and King, D. (1998) Biogeographic origins of goannas (Varanidae): A molecular perspective. *Mol. Phylogenet. Evol.* **9**, 294–307.
- Hecht, M. (1975) The morphology and relationships of the largest known terrestrial lizard, *Megalania prisca* Owen, found from the Pleistocene of Australia. *Proc. R. Soc. Victoria* **87**, 239–250.
- Holmes, R. S., King, M., and King, D. (1975) Phenetic relationships among varanid lizards based upon electrophoretic data and karyotypic analysis. *Biochem. Syst. Ecol.* **3**, 275–262.
- Huelsenbeck, J. P., and Ronquist, F. R. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
- Hugall, A. F., and Lee, M. S. Y. (2004) Molecular claims of Gondwanan age for Australian agamid lizards are untenable. *Mol. Biol. Evol.* **21**, 2102–2110.
- King, M., and King, D. (1975) Chromosomal evolution in the lizard genus *Varanus* (Reptilia). *Aust. J. Biol. Sci.* **28**, 89–108.
- Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., and Wilson, A. C. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Nat. Acad. Sci. USA* **86**, 6196–6200.
- Kumazawa, Y. (2007) Mitochondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations. *Gene* **388**, 19–26.
- Kumazawa, Y., and Endo, H. (2004) Mitochondrial genome of the Komodo dragon: Efficient sequencing method with reptile-oriented primers and novel gene rearrangements. *DNA Res.* **11**, 115–125.
- Rögl, F. (1998) Palaeogeographic reconstructions for Mediterranean and Paratethys seaways (Oligocene to Miocene). *Am. Naturhist. Mus. Wien* **99A**, 279–311.
- Schulte, J. A., II, Melville, J., and Larson, A. (2003) Molecular phylogenetic evidence for ancient divergence of lizard taxa on either side of Wallace’s Line. *Proc. R. Soc. Lond. B* **270**, 597–603.
- The Plates Project (1998) Atlas of paleogeographic reconstructions. Plates Progress Report No. 215, University of Texas Institute for Geophysics Technical Report No. 181. The University of Texas Institute for Geophysics, Austin, Texas.
- Thorne, J. L., Kishino, H., and Painter, I. S. (1998) Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657.
- Yang, Z. (1997) PAML: A program package for phylogenetic analyses by maximum likelihood. *CABIOS* **13**, 555–556 (<http://abacus.gene.ucl.ac.uk/software/paml.html>).