

Mitochondrial Genome of the Komodo Dragon: Efficient Sequencing Method with Reptile-Oriented Primers and Novel Gene Rearrangements

Yoshinori KUMAZAWA^{1,*} and Hideki ENDO²

Division of Material Science, Graduate School of Science, Nagoya University; Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan¹ and Department of Zoology, National Science Museum; 3-23-1 Hyakunin-cho, Shinjuku-ku, Tokyo 169-0073, Japan²

(Received 19 January 2004; revised 9 February 2004)

Abstract

The mitochondrial genome of the Komodo dragon (*Varanus komodoensis*) was nearly completely sequenced, except for two highly repetitive noncoding regions. An efficient sequencing method for squamate mitochondrial genomes was established by combining the long polymerase chain reaction (PCR) technology and a set of reptile-oriented primers designed for nested PCR amplifications. It was found that the mitochondrial genome had novel gene arrangements in which genes from NADH dehydrogenase subunit 6 to proline tRNA were extensively shuffled with duplicate control regions. These control regions had 99% sequence similarity over 700 bp. Although snake mitochondrial genomes are also known to possess duplicate control regions with nearly identical sequences, the location of the second control region suggested independent occurrence of the duplication on lineages leading to snakes and the Komodo dragon. Another feature of the mitochondrial genome of the Komodo dragon was the considerable number of tandem repeats, including sequences with a strong secondary structure, as a possible site for the slipped-strand mispairing in replication. These observations are consistent with hypotheses that tandem duplications via the slipped-strand mispairing may induce mitochondrial gene rearrangements and may serve to maintain similar copies of the control region.

Key words: lizard; reptile; mitochondrial DNA; polymerase chain reaction; tandem repeat

1. Introduction

Mitochondrial DNAs (mtDNAs) of vertebrates are 16–18 kbp double-stranded circular DNAs, which encode genes for 2 rRNAs, 22 tRNAs, and 13 respiratory proteins together with the major noncoding region or the control region which is responsible for replication and transcription of the mitochondrial genome (Fig. 1A).^{1,2} Unlike nuclear DNAs, the mtDNAs are conservative in the gene content, abundant (multicopied) in a cell, intronless, and free from frequent DNA recombination and gene duplication/deletion.³ Because of these advantages, molecular evolutionists have frequently chosen orthologous sets of partial mtDNA sequences to reconstruct the phylogenetic relationships of animals.⁴ In recent years, however, molecular techniques in sequencing and data analysis have been revolutionized so that the use of even complete mtDNA sequences for phylogenetic inference is

becoming a popular approach.^{5–9} In general, more sampling of nucleotide or amino acid sites would lead to more reliable phylogenetic conclusions.⁴

Squamates (lizards and snakes) are a group of reptiles that include a variety of species with intriguing evolutionary questions.¹⁰ However, complete mtDNA sequences have been reported from relatively few squamate species.^{5,11,12} It thus seems important to establish a method to sequence squamate mitochondrial genomes efficiently and accurately. This will also provide insights into how mitochondrial genes and genomes have evolved by comparing the genome structure between squamates and other vertebrates.

In this study, we describe an efficient sequencing method for squamate mitochondrial genomes by combining the long polymerase chain reaction (PCR) technology and a set of reptile-oriented primers for nested PCR amplifications. The mitochondrial genome of the Komodo dragon was sequenced by this method and found to have some unique features in its gene organization with implications on the evolutionary mechanisms of mitochondrial

Communicated by Masahiro Sugiura

* To whom correspondence should be addressed. Tel. +81-52-789-5799, Fax. +81-52-789-2811, E-mail: h44858a@nucc.cc.nagoya-u.ac.jp

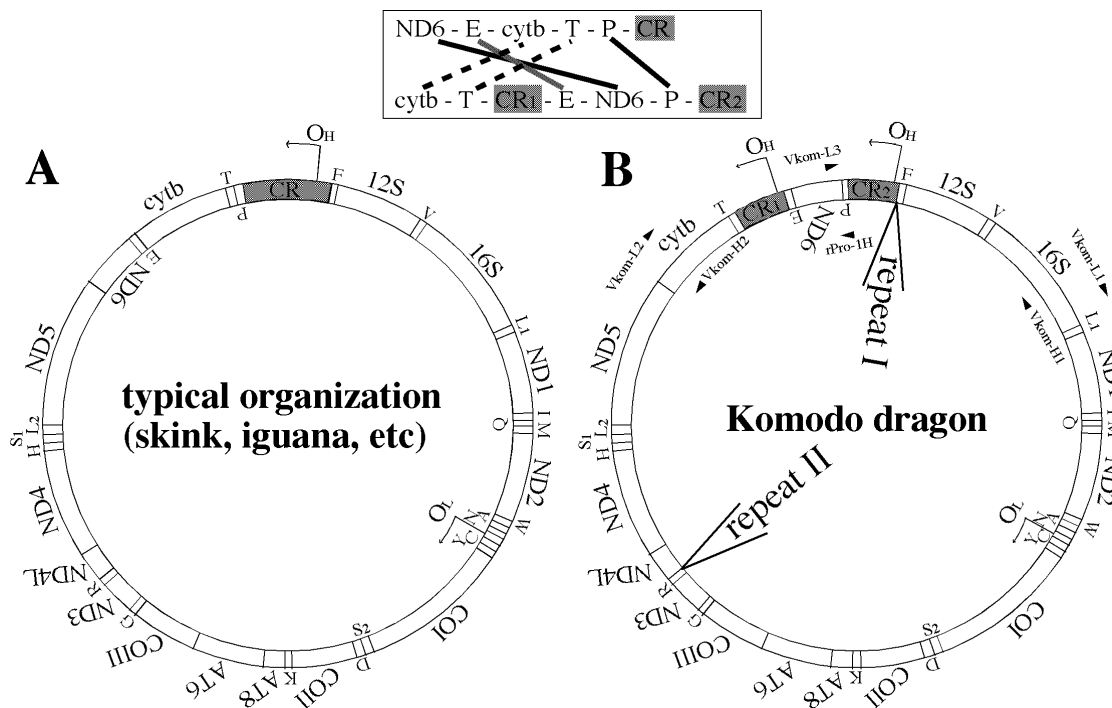


Figure 1. Gene rearrangements in the Komodo dragon mtDNA. *A*, the typical gene organization of vertebrates² found also in the mole skink⁵ and the green iguana.¹² *B*, the gene organization found for Komodo dragon mtDNA. Genes which are encoded by the heavy strand are shown outside the circle, whereas those encoded by the light strand are shown inside the circle. In *B*, the location and orientation of primers used for the LA-PCR amplifications are shown. Gene abbreviations used are: 12S and 16S, 12S and 16S rRNA genes; ND1–6, NADH dehydrogenase subunits 1–6; COI–III, cytochrome oxidase subunits I–III; AT6 and AT8, ATPase subunits 6 and 8; cytb, cytochrome *b*; and one-letter codes of amino acids, tRNA genes specifying them (L1 and L2 for leucine tRNA genes specifying, respectively, UUR and CUN codons and S1 and S2 for serine tRNA genes specifying, respectively, AGY and UCN codons). CR, O_H, and O_L stand for the control region, the heavy-strand replication origin, and the light-strand replication origin, respectively. Repeats I and II correspond to tandem repeat regions I and II shown in Table 2. *Inset*, relative arrangements of several genes and CRs in the typical organization (top) and in the rearranged organization of the Komodo dragon (bottom). Correspondence of genes between the two organizations is shown with various types of lines.

genomes.

2. Materials and Methods

2.1. Samples and basic experimental procedures

Tissue samples of the Komodo dragon (*Varanus komodoensis*) were obtained from an individual which died in the Ueno Zoo, Tokyo in 1999. Blood samples from a second individual were also donated by Drs. K. Fushitani and K. Imai. This sample was originally imported to Japan under the CITES permission by Mr. K. Igarashi. DNA extraction from these samples was carried out with a DNeasy tissue kit (Qiagen) and a QIAamp DNA blood mini kit (Qiagen), respectively.

PCRs were done in one of two different conditions. In order to amplify relatively long (2–15 kbp) regions, conditions for the long-and-accurate PCR (LA-PCR)¹³ were employed. LA-PCRs were set up in a total volume of 20 μ l that contained 2.0 μ l of 10X LA-PCR buffer, 2.0 μ l of 25 mM MgCl₂, 3.2 μ l of 2.5 mM dNTPs, 0.5 μ l each of 10 μ M primers, 0.2 μ l of LA Taq DNA polymerase (Takara), and 1.0 μ l of the total genomic DNA. Min-

eral oil was placed on the reaction mixture to prevent evaporation. Thirty-two PCR cycles were repeated with the Takara Thermal Cycler TP2000 with denaturation at 94°C for 30 sec, annealing at 55 or 60°C for 1 min, and extension at 72°C for 14 min. Amplified products were electrophoresed and recovered in a low-melting-temperature agarose gel (NuSieve GTG, FMC) with special care to avoid contamination by exogenous DNAs. Freshly prepared buffer was used for this electrophoresis and no LA-PCR products for different species were recovered from the same agarose gel.

Regular PCRs for targeting shorter regions were set up in a total volume of 25 μ l that contained 2.5 μ l of 10X Ztaq reaction buffer, 2.0 μ l of 2.5 mM dNTPs, 1.2 μ l each of 10 μ M primers, 0.25 μ l of Ztaq DNA polymerase (Takara), and 1.0 μ l of the template DNA. Thirty-five PCR cycles were usually repeated in an oil-free condition with the Takara Thermal Cycler SP with denaturation at 98°C for 1 sec, annealing at 55°C for 10 sec, and extension at 72°C for 30 sec. Amplified products were checked for their size and yield by agarose gel electrophoresis and purified with a High Pure PCR Product

Purification Kit (Roche). They were directly sequenced for both strands with the amplification primers and/or appropriate internal primers using a DYEnamic ET Terminator Cycle Sequencing Kit (Amersham) on the Applied Biosystems 373A DNA sequencer.

2.2. Sequencing the Komodo dragon mtDNA

Starting from the crude DNA extracted from the tissue sample, partial gene sequences for several mtDNA regions were first obtained by amplification and sequencing with well-conserved primers: rPhe-1L and H1858 for the 12S rRNA gene, r16S-2L and 16sbr-H for the 16S rRNA gene, uCO2-1L and rAT6-3H for the cytochrome oxidase subunit II (COII) and ATPase subunit 8 (AT8) genes, rND5-1L and rND5-2H for the NADH dehydrogenase subunit 5 (ND5) gene, and L14841¹⁴ and rcytb-1H for the cytochrome *b* (cytb) gene (see Table 1 and Fig. 2). Species-specific primers for the LA-PCR were then designed based on the partial gene sequences. Primers Vkom-L1 (5'-CAGCCGCTACTAAAGGTTTCGT-3') and Vkom-H1 (TTTGCACGGTTAGGATACCG) were designed for the 16S rRNA region, whereas primers Vkom-L2 (ATAGCTACAGCCTTCATAGG) and Vkom-H2 (TAGCTAGAAAGAGTCCTGT) were made for the cytb region (see Fig. 1B for their locations). Partial sequences obtained for the other genes (12S rRNA, COII, AT8, and ND5) were later used to reconfirm their identity to the assembled mtDNA sequence.

With the crude DNA as a template, LA-PCR carried out with Vkom-L1 and Vkom-H2 gave rise to a product of approximately 13 kbp, whereas amplification with Vkom-L2 and Vkom-H1 gave no detectable product for unknown reasons, such as effects of secondary structures in template DNA (data not shown). No product was also seen when using the GC buffers supplied with the LA Taq polymease (Takara) for structured templates. We then decided to divide the remaining half of the mtDNA into two parts. LA-PCR with primers Vkom-L2 and rPro-1H gave rise to a product of approximately 2.5 kbp. After sequencing the 3' end region of this product, another species-specific primer Vkom-L3 (GACCTCAAGCTCAACAAGCTTTCG) was synthesized. LA-PCR with Vkom-L3 and Vkom-H1 gave a product of approximately 6 kbp.

These LA-PCR products were carefully purified by the agarose gel electrophoresis as described above and used as a template for nested PCR amplifications using a variety of reptile-oriented primers. These primers are listed in Table 1 and their locations are shown in Fig. 2. As shown below, the gene organization of Komodo dragon mtDNA is different from the typical organization of vertebrates. Thus, various combinations of the primers for nested PCRs were carefully tested until all the results could be interpreted reasonably. The nested PCR products were sequenced and gaps were filled by the primer walking

strategy. These sequences were assembled with the Sequencher 3.1 (Gene Codes), giving rise to two continuous sequences (accession numbers AB080275 and AB080276; see Table 2). Because the Komodo dragon mtDNA contained two highly repetitive regions (Fig. 1B), we could not unambiguously sequence these repetitive regions.

2.3. Data analyses

Nucleotide sequences were analyzed with DNASIS-Mac ver. 3.5 (Hitachi Software Engineering) to identify coding regions based on the sequence similarity to mitochondrial genes from other lizards such as the blue-tailed mole skink⁵ and the green iguana.¹² Transfer RNA (tRNA) genes were identified with reference to the standard secondary structure model for mitochondrial tRNAs.¹⁵ The base composition, tandem repeats, secondary structures, and conserved sequence elements were assessed with DNASIS-Mac ver. 3.5.

3. Results and Discussion

3.1. Characteristics of the Komodo dragon mtDNA

The gene organization of the Komodo dragon mtDNA thus sequenced is depicted in Fig. 1B, and features on encoded genes and structural elements are summarized in Table 2. The Komodo dragon mtDNA contained all 37 genes typically found for other vertebrates,² but with a novel organization. Genes from ND6 to proline tRNA were extensively shuffled with duplicate control regions (Fig. 1). These two major noncoding regions were judged to be the control region because they contained conserved sequence blocks (CSBs) I and II, as well as extended termination associated sequence 1 (ETAS 1) that have been identified as conserved sequence elements for the control region of mammals^{16,17} (see Fig. 3). CSB II was suggested to be associated with initiation of heavy-strand replication,¹⁶ whereas ETAS 1 was shown to be a pausing site of nascent heavy-strand synthesis to make a displacement loop.¹⁸

Base frequency of the light-strand sequence of the Komodo dragon mtDNA was 29.6% for A, 29.5% for C, 26.9% for T, and 14.0% for G. The scarcity of G for the light strand is a common feature found in metazoan mtDNAs.¹⁹ Sizes and sequences of the encoded 37 genes were similar to those of the mole skink⁵ and the green iguana¹² (data not shown). Twelve protein genes other than ND2 appear to start from either of the two methionine codons of the mitochondrial genetic code¹ (i.e., ATA or ATG), whereas the ND2 gene starts from a noncanonical ATT codon (Table 2). In 6 of the 13 protein genes, a stop codon appears to be created by polyadenylation.²⁰ All transfer RNA genes could be folded into standard secondary structures for vertebrate mitochondrial tRNAs,¹⁵ except for the cysteine tRNA gene which appears to lack the dihydrouridine arm (data not shown). There was

Table 1. Primers designed to amplify and sequence squamate mtDNAs.

light-strand primers		heavy-strand primers		PCR amplification								
No.	name	sequence (length)	No.	name	sequence (length)	Socc	Eegr	Cwar	Agra	Scro	Vkom	Ldul
1	rPhe-1L	AAAGCACGGCACTGAARATGC (21)	32	H1858	TCGATTATAGRACAGGCTCCTCTAG (25)	*	*	*	*	*	*	*
2	rI2S-1L	AGGATTAGATACCTACTA (19)	33	rI6S-3H	CAKKTTCCTCTGCGGTACT (20)	*	I	*	*	*	*	*
3	uI2S-1L	GCGYACAYAYCGCCGTC (18)	34	rI6S-5H	TTTATYRRGYAACAGCTATC (21)	*	*	*	*	*	*	*
4	rI6S-3L	AACCCYYGTACTCTYTTGCATCATG (24)	35	rI6S-1H	TYCACAGGGTCTTYTCGTC (19)	*	*	*	*	*	*	*
5	rI6S-2L	CRACCTGTTACCAAAAAACAT (20)	36	l6sar-H	CCGGTCTGAACCTCAGATCACGT (22)	*	*	*	*	*	*	*
6	rI6S-4L	TACTCCAGGGATAACAGCGC (20)	37	rND1-1H	GCRTATTTGAGTTKAGAKCTCA (23)	*	*	*	*	*	*	*
7	rND1-1L	TACATRCAACTWCGAAAAAGG (20)	38	rND1-2H	TCAAATGGKGCCTGRITDGTTC (23)	*	*	*	*	*	*	*
8	rND1-2L	CAAACMATCTCMTAYGAAGT (20)	39	rMet-2H	GGTATGGGCCRAWAGCTT (19)	*	*	*	*	*	I	*
9	rND1-3L	CGATTCCGATAYGACCAACT (20)	40	rND2-2H	ATTGATGAGWAKGCTATRATTTTCG (26)	*	*	*	*	*	I	*
10	rND2-5L	TTACCCWCGAGCAACWGAAGC (21)	41	rAsn-1H	TGGGYGKTTAGCTGTTAAAYTA (21)	*	*	*	*	*	*	*
11	rND2-1L	GCCCCMYTMCCTCTCTGA (18)	42	rCOI-1H	GTAYAGGGTGCCRATRTCTTT (21)	I	*	*	*	*	*	*
12	rTrp-1L	TAAACCARGRGCCTTCAAAG (20)	43	rCOI-2H	GGGTGKCCAAAARAATCAGAA (20)	I	*	*	*	*	*	*
13	rCOI-1L	ATCGCGGRITTYGGAAACTG (20)	44	rCOI-1H	TAGTGAARTGKGTACTAC (20)	*	*	*	*	*	*	*
14	rCOI-2L	TCWGGCCACAATAATYATYGC (20)	45	rCO2-1H	TGGAAGTGWARTAGYCTTCTAT (23)	*	*	*	*	*	I	I
15	rCOI-4L	TACTCAGACTACCCAGAYGC (20)	46	uCO2-1H	CCGCAGATTTCTGAGCATTG (20)	*	*	*	I	*	I	*
16	uCO2-1L	GGMCAYCAATGATACTGA (18)	47	rAT6-3H	AAGYTTAKGGTCATGGTCA (19)	*	*	*	*	*	*	I
17	uLys-1L	AGCACTAGCCTTTAAGC (18)	48	uCO3-1H	AAYGTCTCGTCATCATTG (18)	*	I	*	*	*	I	*
18	uCO3-1L	ATAGTWGACCCMAGCCCATGACC (23)	49	uND3-2H	GGGTCRAAKCCRCATTCTA (20)	*	*	*	I	*	*	I
19	uCO3-3L	GAAGCMGWCCTGATACTGACA (23)	50	rND4L-2H	GCTAGGCCAGTRCYTGCTTCRCA (23)	*	*	*	*	I	R	*
20	rND4L-1L	TGCATTGAARGYATAATACT (20)	51	uND4-2H	CTACRTGKGTCTTTGGKARTCA (22)	*	I	*	*	*	*	*
21	rND4L-2L	TAACTTCTCMGCMGTGYGAAGC (22)	52	rND4-2H	GATGTAAKCCGTGGCRATTAT (23)	*	*	*	*	*	*	*
22	rND4-3L	CCAAAAGCCCAAYGTAGARGC (20)	53	rCUN-1H	CTTTACTTGGADTTGCACC (20)	*	*	*	*	*	*	*
23	rHis-2L	AACAAAAACAYTAGRGTGTG (20)	54	rND5-1H	ACWACTATTGTGCTKAGGTG (20)	*	*	*	*	*	I	*
24	rND5-1L	TCCAAGCMATYATCTAYAACCG (22)	55	rND5-2H	ATWGYGTCTTTGAGTARAACKC (23)	*	*	*	*	*	*	*
25	rND5-2L	GAACARGACATYCGAAAAATRGG (23)	56	rND6-4H	ATGTTAGTGGTDTTTGCKTATTC (23)	*	*	I	*	*	R	*
26	rND5-3L	YACMYMAACGCCTGAGCCCT (20)	57	rGlu-1H	ATTACAACGGYGGTTTTTC (19)	*	*	*	*	*	R	*
27	rND6-3L	GCAACWGAATAHGCAAAATAC (20)	58	ucytb-1H	GCCCCCAGAATGATATTTGCTCTCA (26)	*	*	*	*	*	R	*
28	rGlu-1L	GAAAAACRCCTGTGTWATCAACTA (26)	59	rcytb-1H	GCGTAGGCRAATAGGAAGTATCA (23)	*	*	*	*	*	R	*
29	rcytb-2L	TGAGGACAAATATCMTTCTGAGG (23)	60	rPro-1H	TTAAAATKCTAGTTTTGG (18)	I	*	I	*	*	R	*
30	rThr-2L	YAAAGCMTTGTCTTGTA (19)	61	rCONT-4H	CTCGKTTWGGGGTTGRCGA (21)	*	*	I	*	I	*	*
31	rCONT-4L	TCGYCAAACCCWAAAMCGAG (21)	62	rI2S-1H	TRTAACCGCGTKGCTGGCAC (21)	*	*	I	*	*	*	*

Thirty one typical combinations of primers for nested PCRs are shown in rows along with test results of their applicability to seven squamates. Location of each primer can be found in Fig. 2. Primer names starting with r and u mean that they were designed to be oriented for reptiles (mostly squamates) and universal for wider vertebrate groups, respectively. Names of previously published primers are shown in their original names: primer No. 32 from Sorenson et al.³⁰ and No. 36 from Palumbi et al.³⁴ The following are modified primers from original sequences: primer Nos. 2 and 58 modified from L1091 and H15149 of Kocher et al.,¹⁴ Nos. 5 and 59 modified from l6sar-L and CB3-H of Palumbi et al.,³⁴ Nos. 9 and 42 modified from L4160m and H5937m of Kumazawa and Nishida¹⁵, Nos. 12, 41, and 53 modified from L5549, H5689, and H12315 of Kumazawa and Nishida³⁵ and Nos. 8, 19, 38, 47, and 49 modified from L4500, L10647, H4644, H9233, and H10884 of Sorenson et al.³⁰ Interpretations of results on PCR amplifications are: excellent amplification (*), inferior amplification with either faint or multiple products (I), no expected product with reasonable size due to the gene rearrangement (R), and no amplification (blank). Taxon names tested are: Western fence lizard (Socc), Blue-tailed mole skink (Eegr), Warren's spinytail lizard (Cwar), Terrestrial arboreal alligator lizard (Agra), Chinese crocodile lizard (Scro), Komodo dragon (Vkom), and Texas blind snake (Ldul).

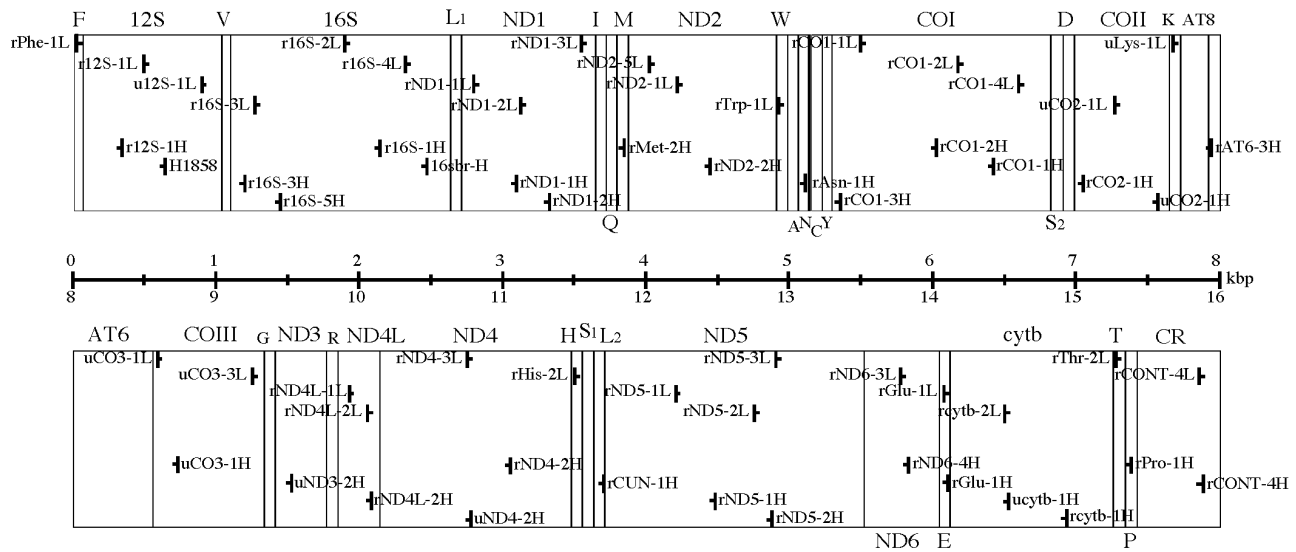


Figure 2. Positions of primers designed for amplifying and sequencing squamate mtDNAs. See Table 1 for nucleotide sequences of these 62 primers. The typical gene organization of vertebrate mtDNAs² is shown with columns which approximate the sizes of individual genes, although the size of the control region (CR) varies considerably among species. Heavy-strand-encoded and light-strand-encoded genes are shown above and below the column, respectively. See the legend of Fig. 1 for gene abbreviations.

a characteristic stem-and-loop structure between the asparagine and cysteine tRNA genes of the WANCY tRNA gene cluster (Fig. 4C), which was shown to be an origin for light-strand replication in mammals.²¹ Thus, replication of the Komodo dragon mtDNA may occur asymmetrically between two strands, as suggested for mammals.¹⁶ However, because there are two control regions with conserved sequence elements, heavy-strand replication may initiate efficiently from two replication origins (Fig. 1B).

Another feature of the Komodo dragon mtDNA was the considerable number of tandem repeats. There was an insertion of approximately 1.6 kbp between the genes for arginine tRNA and ND4L (Fig. 1B) as judged from the size of the PCR product amplified with taxon-specific primers designed for these genes (data not shown). Since the same long product was found from the second individual, this is not a length polymorphism occurring in a particular individual but a trait shared by multiple individuals of the species. Sequencing this PCR product with these primers together with the information on the size of the insert revealed that the insert consists of 15 tandem repeats of a 106-bp repeat unit (Fig. 4A). Tandem repeats in or adjoining the control region have been described for other vertebrates.^{22,23} However, to the best of our knowledge, the occurrence of such large tandem repeats outside the control region is quite unusual.

Although this insert was repeatedly sequenced from several taxon-specific primers (primer sequences not shown), the complete nucleotide sequence of this insert could not be determined unambiguously for the following reasons. First, electropherogram profiles by the sequencer always deteriorated for longer reading presum-

ably due to structural characteristics of the insert (i.e., simple repetition of almost identical sequences having a strong stem-and-loop structure; see Fig. 4A). Second, the primer walking strategy could not be applied for such repeated regions, nor could the inverse PCR method²⁴ be applicable because there was no appropriate restriction site in the repeated sequences.

Another tandem repeat was found at the 3' end of the second control region (Fig. 1B). This tandem repeat consists of arrays of an AT-rich 115-bp repeat unit (Figs. 3 and 4B). Although unambiguous sequences could not be obtained beyond the second repeat unit for the same reasons as described above, the electropherogram profiles suggested continuation of at least 6 repeat units in the 3' vicinity of the second control region (data not shown). It was also found that the same repeat unit continues at least 6 times in the 5' vicinity of the phenylalanine tRNA gene. If the space between the second control region and the phenylalanine tRNA gene consists entirely of the 115-bp repeats, there would be more than 20 repeated units in this region. We could not clarify this point due to the technical difficulties described above. The third tandem repeats were found at the 5' end portion of the two control regions (Fig. 3). A 36-bp repeat unit starts inside the threonine tRNA gene and continues 4 times in control region 1 (Fig. 3). In the case of control region 2, the 36-bp repeat starts immediately after the proline tRNA gene and continues 5 times (Fig. 3). This type of short tandem repeat inside the control region is common among vertebrates.²³

Table 2. Gene arrangements of the Komodo dragon mtDNA.

feature	positions	proteins		tRNAs
		start	end	anticodon
the first segment of mtDNA (accession No. AB080275) ¹				
tandem repeat region I	1 - 171	—	—	—
tRNA ^{Phe}	172-242	—	—	GAA
12S rRNA	243-1184	—	—	—
tRNA ^{Val}	1185-1251	—	—	TAC
16S rRNA	1252-2766	—	—	—
tRNA ^{Leu}	2767-2839	—	—	TAA
NADH dehydrogenase subunit 1 (ND1)	2841-3806	ATA	TAA	—
tRNA ^{Ile}	3808-3876	—	—	GAT
tRNA ^{Gln}	3875-3944 (H) ³	—	—	TTG
tRNA ^{Met}	3944-4011	—	—	CAT
NADH dehydrogenase subunit 2 (ND2)	4012-5048	ATT	TA*	—
tRNA ^{Trp}	5049-5117	—	—	TCA
tRNA ^{Ala}	5118-5186 (H)	—	—	TGC
tRNA ^{Asn}	5188-5260 (H)	—	—	GTT
light-strand replication origin (O _L)	5260-5297	—	—	—
tRNA ^{Cys}	5294-5349 (H)	—	—	GCA
tRNA ^{Tyr}	5350-5415 (H)	—	—	GTA
cytochrome oxidase subunit I (COI)	5417-7018	ATG	AGA	—
tRNA ^{Ser}	7012-7082 (H)	—	—	TGA
tRNA ^{Asp}	7084-7150	—	—	GTC
cytochrome oxidase subunit II (COII)	7151-7840	ATG	TAG	—
tRNA ^{Lys}	7841-7904	—	—	TTT
ATPase subunit 8 (AT8)	7905-8069	ATG	TAA	—
ATPase subunit 6 (AT6)	8060-8742	ATG	TA*	—
cytochrome oxidase subunit III (COIII)	8743-9526	ATG	T*	—
tRNA ^{Gly}	9527-9592	—	—	TCC
NADH dehydrogenase subunit 3 (ND3)	9593-9938	ATA	T*	—
tRNA ^{Arg}	9939-10004	—	—	TCG
tandem repeat region II	9979-10624	—	—	—

¹Numbering starts inside the tandem repeat region I and ends within the tandem repeat region II.²Numbering starts inside the tandem repeat region II and ends within the tandem repeat region I.³H in a parenthesis means that the corresponding gene is encoded by the light strand. Asterisks mean that polyadenylation can create a stop codon with the mitochondrial genetic code.²⁰

Table 2. Continued.

the second segment of mtDNA (accession No. AB080276) ²				
tandem repeat region II	1-671	—	—	—
NADH dehydrogenase subunit 4L (ND4L)	727-1023	ATA	TAA	—
NADH dehydrogenase subunit 4 (ND4)	1017-2390	ATG	TAA	—
tRNA ^{His}	2395-2464	—	—	GTG
tRNA ^{Ser}	2465-2525	—	—	GCT
tRNA ^{Leu}	2525-2595	—	—	TAG
NADH dehydrogenase subunit 5 (ND5)	2597-4398	ATA	TA*	—
cytochrome <i>b</i> (cytb)	4399-5531	ATG	TA*	—
tRNA ^{Thr}	5532-5599	—	—	TGT
control region 1	5600-6326	—	—	—
tandem repeat region III	5586-5729	—	—	—
extended termination associated sequence 1	5790-5860	—	—	—
conserved sequence block I (CSB I)	6246-6268	—	—	—
conserved sequence block II (CSB II)	6299-6314	—	—	—
tRNA ^{Glu}	6327-6394 (H)	—	—	TTC
NADH dehydrogenase subunit 6 (ND6)	6400-6930 (H)	ATG	AGG	—
tRNA ^{Pro}	6992-7058 (H)	—	—	TGG
control region 2	7059-7907	—	—	—
tandem repeat region III	7060-7257	—	—	—
extended termination associated sequence 1	7317-7387	—	—	—
conserved sequence block I (CSB I)	7773-7795	—	—	—
conserved sequence block II (CSB II)	7827-7842	—	—	—
tandem repeat region I	7908-8089	—	—	—

3.2. Evolution of the squamate mtDNAs

We found that the Komodo dragon mtDNA has two control regions separated by three genes (genes for glutamate tRNA, ND6, and proline tRNA). The duplicate control regions had a high sequence similarity to each other (99%; 7 changes out of 730 alignable sites; see Fig. 3). Kumazawa et al.²⁵ found that two control regions with nearly identical nucleotide sequences are present in separate locations of mtDNA (one in the typical position and the other after the isoleucine tRNA gene within the IQM tRNA gene cluster) of several snake species. However, mechanisms for the maintenance of the apparently redundant control regions, as well as those for the concerted sequence evolution of two control regions within species, remain to be elucidated. Duplicate control regions in Komodo dragon mtDNA are present in an entirely different arrangement from that in snakes, suggesting that independent duplication events occurred on lineages leading to snakes and the Komodo dragon.

Since there is no direct evidence for DNA recombination in animal mitochondria,²⁶ gene rearrangements of mtDNAs have been explained by the duplication-and-deletion model²⁷ in which tandem duplication of an mtDNA segment and sporadic deletion of redundant gene copies leads to rearranged organizations. Based on this model, at least two cycles of duplications and deletions are needed to explain the rearrangement that changed the typical gene organization to that of the Komodo dragon (Fig. 5).

We consider that the duplicate state of the control region may have played an important role in the gene rearrangement. If there are two control regions with similar sequences, this could be a hot target for the slipped-strand mispairing²⁸ that causes tandem duplication of the spacing region between the two control regions (refer to the process from state C to state D in Fig. 5). A similar process may also lead to the concerted sequence evolution of the two control regions (state E to state G in Fig. 5), as proposed by Kumazawa et al.¹¹ for snake mtDNAs.

```

          cytb ]
      K L L N W * ] tRNA(Thr)                                tRNA(Thr) ] ] control region 1
CR1   1  AACTACTTAA CTGATAGTCC TAGTAGCTTA ACAACAAAGC ACTGACCTTG TAAGCCAGAG AATGGAAGTT TATTCCTAG GACTCCCAT TTCACATTT
CR2   1  CTTAAACTAT CTTCTGCTT CACATTTTA GCTCTATCC CTAGGACTCT CCATTTTACA TTTTACGTC .....T.....
          tRNA(Pro) ] ] control region 2

CR1  101 TAGCTCTATT CCCTAGGACT CCCCATTTC CATTTTTACG TCTATTCCT AGGACTCCCC ATTTACACATT TTTAGCTCTA TTCCTAGGA CTCTCCATTT
CR2  101 .....T.....

CR1  201 CATATTTTAA TTTTTTCGGC CCCTCCTCCT TTAAGGTCAG CTTAGCTCAA CGTCTGAAT TTTCCCTTTT TTTTAAATTT TARGTCTTTC GAGCAACCAA
CR2  201 .....

          ETAS 1
CR1  301 ACGCGCCACC TCCTGGATAT CGTCCCGCCA ACAGCATTTC ATATTTTGT ACCTCTATTA AACTCGATT AAATGTCATT TTCAAGACAC TCAACTAAGC
CR2  301 .....

CR1  401 ACTGGCTACC CCTATCGGT CGCTACTGT TACCAGTCTC GTGGATCATA CCTATAGGTT GCACCTATTT AATGACCTTT CCAATACCTC TGGTGTGAG
CR2  401 .....

CR1  501 GCCCAGGAC TTTCTTTCAA GGTGACCACT CTTTTCTCTC TAAAGCACTT CGGGTTGGGT GAATCTCAGG AGCTTATCAC CTCATACTA CGGTACCCCT
CR2  501 .....

CR1  601 GGTATAGGCG CCTTCAGCCT TTTTCTTTT TTTTGGTGG ATCTCAGATC GCATGTTCGT CCGGGTCCAG CTTTACATTT ACCAAACGAT CATTGTGACA
CR2  601 .....T.....C.....

          CSB I                                CSB II
CR1  701 ATCGAACCTT TTATAATTAC ATTGGAGCTG GAATTTAATG GTCGCCGGAC ATACAATAAA TCAAAAAAAC ATAATTTTTT -AAAAAACCC CCAAACCCCC
CR2  701 ..... A.....

control region 1 ] ] tRNA(Glu)                                tRNA(Glu) ] ] ND6
CR1  801 TACACTCCCC ATTATTCTCA TTTGGATTTA CCAAACAACA AGGACACGAA AAATCCACGT TGTACTTCAA CTACAAGAAC CCTCCCTCTG ATACAGAACC
CR2  801 AGCAAAAAAA TAAAAATTTT TTTTAATTTT TTTTAAAAAA ATTAATAAAA AATTTTTTTT AAAAAAATTT TTTATTTTTT TTAATAATTT TTTTACGCGG
          control region 2 ] ] tandem repeat region 2

CR2  901 CAAGTTTAGG ACTTTATAAA TACTTTTATA TAACCTAAAA ATCACAAGTT TTTTTTAAAA AAACGGGATC CAAATATTGC AATTTTTTAT TTTTTAAAA

```

Figure 3. Nucleotide sequences of duplicate control regions and their flanking genes of the Komodo dragon mtDNA. All nucleotide sequences shown correspond to the light strand. The sequences are presented with the alignment between comparable regions of the two control regions (CR1 and CR2). Anticodon sequences of tRNA genes, conserved sequence elements, and repeat units are underlined.

The slipped-strand mispairing was considered to occur at regions with strong secondary structures, where the DNA polymerase stalls and a nascent DNA strand is temporarily displaced to reanneal to the secondary site of a template strand and replicate the same region twice.²⁸ We found that the Komodo dragon mtDNA is especially prone to tandem duplication. Consistent with the idea by Levinson and Gutman,²⁸ the repeat units for the two major tandem repeat regions had a strong secondary structure (Fig. 4).

3.3. Efficient and accurate sequencing of squamate mtDNAs

In the present study, we established an efficient and accurate method to sequence squamate mtDNAs by developing reptile-oriented primer sets. An outline of this method was briefly described in our previous paper⁵ when we sequenced complete mtDNAs of the blue-tailed mole skink and the green turtle. However, the akamata snake¹¹ was the only squamate taxon whose complete mtDNA sequence was known at that time, and we had to design conserved primers for nested PCRs based on this and other diverse vertebrate taxa including mammals, birds, and crocodylians. These primer sequences

were not published in Kumazawa and Nishida⁵ because they were judged not to be definitive sets of primers for sequencing reptilian mtDNAs from an observation that we had to design many taxon-specific primers for primer walking in order to sequence both strands of the skink mtDNA unambiguously.

Miya and Nishida²⁹ and Sorenson et al.³⁰ independently described similar methods for sequencing vertebrate (primarily fishes and birds, respectively) mtDNAs. Mitochondrial DNA sequence evolution may be relatively conservative for fish and birds;^{5,31,32} this is a trait which facilitates the identification of conserved regions for designing primers. In contrast, snake mtDNAs have considerably faster rates of sequence evolution,¹¹ which made it difficult to find conserved regions for squamates. In addition, mtDNA sequences of fishes and squamates are so diverged that fish-oriented primers²⁹ did not show excellent matching to squamate sequences in general (data not shown). Birds are more closely related to squamates than fishes. We found that many of the primers designed primarily based on avian sequences³⁰ worked also for nonavian amniotes including squamates, as suggested by Sorenson et al.³⁰ and Rest et al.,⁹ but other such primers did not show excellent matching to the squamate sequences (data not shown).

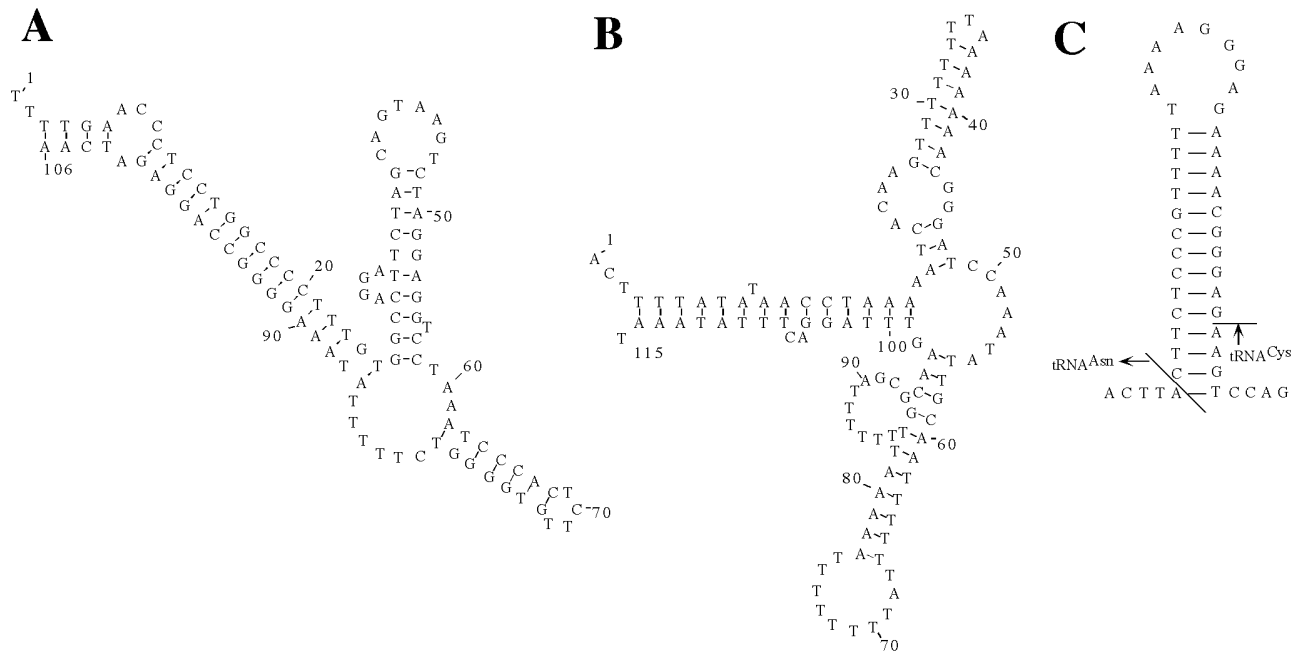


Figure 4. Characteristic secondary structures found for the 106-bp repeat unit between genes of arginine tRNA and ND4L (A), the 115-bp repeat unit at the 3' end of control region 2 (B), and the light-strand replication origin (C). All nucleotide sequences shown are light-strand sequences. Note that the repeat boundary in B is defined differently from that in Fig. 3 in order to stabilize the secondary structure.

Primers reported in the present study are much improved from the previous ones used to sequence skink mtDNA, as evidenced by the excellent PCR amplification for seven squamate taxa (Table 1). Some of the primers in Table 1 were actually used for sequencing mtDNAs of the seven squamates³³ but the others were modified on the basis of these mtDNA sequences. The 31 typical combinations of primers were designed so that sequences of adjacent amplified products can be assembled with overlap (Fig. 2). These primers cover both strands of the mitochondrial genomes with an interval of less than 700 bp (Fig. 2), which is suitable for a single sequencing reaction using modern DNA sequencers. However, there were some extremely variable regions (i.e., AT6 gene, 3' end portion of ND5 gene to ND6 gene, and the control region) where we had to design primers with somewhat longer intervals. These regions may still need to be sequenced with the aid of primer walking.

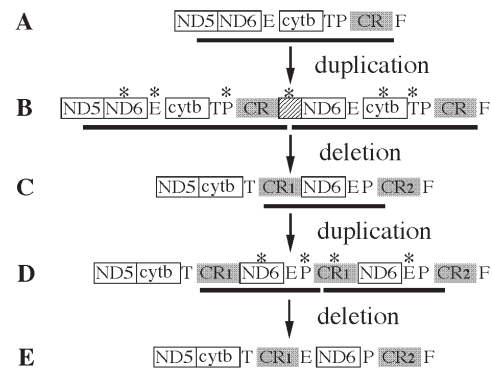
In the 1980s and 1990s, sequencing of an entire mitochondrial genome was mostly achieved by separation of closed circular forms of mtDNA, restriction enzyme digestion, cloning and subcloning, and assemblage of determined sequences into a continuous mtDNA sequence.¹ This is a solid strategy, but may lack general applicability to phylogenetic studies in which complete mtDNA sequences must be obtained for a number of taxa.^{5,29,30} The conventional method requires much time and effort for the subcloning step. In addition, when ecologically important animals are studied, such as the Komodo

dragon that is listed in CITES (Appendix I), it is difficult to obtain biological samples in an amount and quality sufficient for mtDNA purification.

The development of the LA-PCR technology¹³ made it possible to circumvent this difficulty. Because the method does not include any cloning procedure, errors cannot be introduced by Taq polymerase. Other merits of the LA-PCR procedure are the absence of contamination by exogenous DNAs by utilizing taxon-specific primers for at least one of paired primers for the LA-PCR, and the minimal risk of collecting nuclear copies of mtDNA-like sequences. If the method described in this study is utilized to sequence mtDNAs of more reptilian taxa, the obtained data will contribute to resolve intriguing evolutionary questions on phylogenetic relationships and genomic structures of mtDNAs.

Acknowledgements: We thank Mr. Kyohei Igarashi, Drs. Kenzo Fushitani and Kiyohiro Imai, and Ueno Zoo for providing animal samples. We also thank the Center for Gene Research of Nagoya University for providing experimental facilities. Gratitude is extended to Ms. Emiko Kato for her excellent technical assistance in sequencing experiments. This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (grants 12640680 and 14540641).

Gene rearrangement



Concerted evolution of CRs

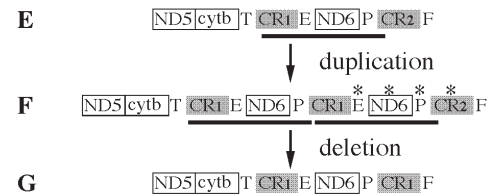


Figure 5. Possible mechanisms of the gene rearrangement from the typical organization (A) to the organization of the Komodo dragon (E) through intermediate states (B–D), and the concerted sequence evolution of the control regions (E–G). Note that the unequal crossing-over and gene conversion postulated to explain the concerted evolution of nuclear and bacterial genes⁴ may not be applicable, because DNA recombinational activities have not yet been proved in vertebrate mitochondria.²⁶ Genes are shown in the abbreviated form and the hatched box is a region of uncertain length added by duplication. Asterisks denote genes that should be deleted to give rise to the rearranged organization. Thick bars below genes denote a duplication unit whose endpoints are uncertain.

References

- Anderson, S., Bankier, A. T., and Barrell, B. G. et al. 1981, Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.
- Boore, J. L. 1999, Animal mitochondrial genomes. *Nucleic Acids Res.*, **27**, 1767–1780.
- Wolstenholme, D. R. 1992, Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.*, **141**, 173–216.
- Page, R. D. M. and Holmes, E. C. 1998, Molecular evolution. A phylogenetic approach. Blackwell Science, London.
- Kumazawa, Y. and Nishida, M. 1999, Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Mol. Biol. Evol.*, **16**, 784–792.
- Zardoya, R. and Meyer, A. 2000, Mitochondrial evidence on the phylogenetic position of caecilians (Amphibia: Gymnophiona). *Genetics*, **155**, 765–775.
- Nikaido, M., Kawai, K., and Cao, Y. et al. 2001, Maximum likelihood analysis of the complete mitochondrial genomes of eutherians and a reevaluation of the phylogeny of bats and insectivores. *J. Mol. Evol.*, **53**, 508–516.
- Miya, M., Takeshima, H., and Endo, H. et al. 2003, Major patterns of higher teleostean phylogenies: A new perspective based on 100 complete mitochondrial DNA sequences. *Mol. Phylogenet. Evol.*, **26**, 121–138.
- Rest, J. S., Ast, J. C., and Austin, C. C. et al. 2003, Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Mol. Phylogenet. Evol.*, **29**, 289–297.
- Zug, G. R. 1993, Herpetology. Academic Press, San Diego.
- 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.
- Janke, A., Erpenbeck, D., Nilsson, M., and Arnason, U. 2001, The mitochondrial genomes of the iguana (*Iguana iguana*) and the caiman (*Caiman crocodylus*): implications for amniote phylogeny. *Proc. R. Soc. Lond. B*, **268**, 623–631.
- Cheng, S., Chang, S.-Y., Gravitt, P., and Respass, R. 1994, Long PCR. *Nature*, **369**, 684–685.
- Kocher, T. D., Thomas, W. K., and Meyer, A. et al. 1989, Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 6196–6200.
- Kumazawa, Y. and Nishida, M. 1993, Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.*, **37**, 380–398.
- Clayton, D. A. 1992, Transcription and replication of animal mitochondrial DNAs. *Int. Rev. Cytol.*, **141**, 217–232.
- Sbisá, E., Tanzariello, F., Reyes, A., Pesole, G., and Saccone, C. 1997, Mammalian mitochondrial D-loop region

- structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene*, **205**, 125–140.
18. Doda, J. N., Wright, C. T., and 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. U.S.A.*, **78**, 6116–6120.
 19. Asakawa, S., Kumazawa, Y., Araki, T., Himeno, H., Miura, K., and Watanabe, K. 1991, Strand-specific nucleotide composition bias in echinoderm and vertebrate mitochondrial genomes. *J. Mol. Evol.*, **32**, 511–520.
 20. Ojala, D., Montoya, J., and Attardi, G. 1981, tRNA punctuation model of RNA processing in human mitochondria. *Nature*, **290**, 470–474.
 21. Wong, T. W. and Clayton, D. A. 1985, In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. *Cell*, **42**, 951–958.
 22. Moritz, C. and Brown, W. M. 1987, Tandem duplication in animal mitochondrial DNAs: variation in incidence and gene content among lizards. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 7183–7187.
 23. Lunt, D. H., Whipple, L. E., and Hyman, B. C. 1998, Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. *Mol. Ecol.*, **7**, 1441–1455.
 24. Sambrook, J. and Russell, D. W. 2001, Molecular cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
 25. 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. *Mol. Biol. Evol.*, **13**, 1242–1254.
 26. Erye-Walker, A. and Awadalla, P. 2001, Does human mtDNA recombine? *J. Mol. Evol.*, **53**, 430–435.
 27. Moritz, C., Dowling, T. E., and Brown, W. M. 1987, Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.*, **18**, 269–292.
 28. Levinson, G. and Gutman, G. A. 1987, Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.*, **4**, 203–221.
 29. Miya, M. and Nishida, M. 1999, Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. *Mar. Biotechnol.*, **1**, 416–426.
 30. Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T., and Mindell, D. P. 1999, Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol. Phylogenet. Evol.*, **12**, 105–114.
 31. Martin, A. P. and Palumbi, S. R. 1993, Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 4087–4091.
 32. Kumazawa, Y. and Nishida, M. 2000, Molecular phylogeny of osteoglossoids: a new model for Gondwanian origin and plate tectonic transportation of the Asian arowana. *Mol. Biol. Evol.*, **17**, 1869–1878.
 33. Kumazawa, Y. 2004, Mitochondrial DNA sequences of five squamates: phylogenetic affiliation of snakes. *DNA Res.*, **11**, 137–144.
 34. Palumbi, S., Martin, A., Romano, S., McMillan, W. O., Stice, L., and Grabowski, G. 1991, The simple fool's guide to PCR. version 2.0. Department of Zoology and Kewalo Marine Laboratory. University of Hawaii, Honolulu.
 35. Kumazawa, Y. and Nishida, M. 1995, Variations in mitochondrial tRNA gene organization of reptiles as phylogenetic markers. *Mol. Biol. Evol.*, **12**, 759–772.