Int J Legal Med (2007) 121:423–427
DOI 10.1007/s00414-006-0113-y

SHORT COMMUNICATION

Two universal primer sets for species identification among vertebrates

Takashi Kitano · Kazuo Umetsu · Wei Tian · Motoki Osawa

Received: 12 January 2006 / Accepted: 7 June 2006 / Published online: 15 July 2006

Abstract The aim of this study was to develop a simple method using universal primers for species identification based on direct PCR sequencing. Two primer sets were designed based on the conserved regions of the 12S and 16S rRNA loci detected by the comprehensive sequence comparison among 30 mammalian whole mitochondrial genomes. In humans, the expected sizes of PCR products of the 12S and 16S rRNAs were 215 and 244 bp, respectively. Both primer sets successfully amplified the expected PCR products from various kinds of vertebrates including mammals, birds, reptiles, amphibians, and fish, and the sequenced segments contained sufficient nucleotide differences to identify each animal species. A case example of the identification of a piece of buried bone of unknown species is presented, and the species was identified as a pig by this method.

Keywords Mitochondrial DNA · 12S rRNA · 16S rRNA · Species identification · Vertebrates

Introduction

Species identification is one of the important aspects of forensic science. Several methods using restriction fragment length polymorphism [1, 2], direct sequencing [3–8], or multiplex PCR using species-specific primers [9, 10] of mitochondrial DNA (mtDNA) have been widely used. Multiplex PCR methods using species-specific primers are very useful for samples in which there is a mixture of DNA from various animals. In contrast, sequence-based methods are useful for various kinds of species when suitable universal primers are used.

To design more suitable universal primers recognizing conserved regions, we made a comprehensive sequence comparison among 30 mammalian whole mitochondrial genomes in this study. Application of mtDNA is very useful for species identification, because sequences from various kinds of species have been deposited in databases. We found several conserved regions in the 12S and 16S ribosomal RNA (rRNA) loci and designed a universal primer set for each of two rRNA loci. For successful amplification of the DNA from severely degraded material, the size of these target regions for each primer set was designed to be relatively short.

Furthermore, we present a case example of a fragment of buried bone of unknown species origin, which was analyzed using the direct sequencing method with these universal primers.

Materials and methods

Multiple alignments

The whole mtDNA sequences of the 30 mammals (Supplementary Table) were retrieved from the Organelle Genome Resources of the National Center for Biotechnology Information web site (http://www.ncbi.nlm.nih.gov). ClustalW version1.8 was employed for the multiple alignment analysis [11]. The average number of nucleotide differences

Electronic supplementary material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00414-006-0113-y and is accessible for authorized users.

T. Kitano · K. Umetsu · W. Tian
Department of Experimental and Forensic Pathology,
Faculty of Medicine, Yamagata University,
Yamagata, Japan

M. Osawa (✉)
Department of Forensic Medicine,
Tokai University School of Medicine,
Isehara, Kanagawa 259-1193, Japan
e-mail: osawa@is.icc.u-tokai.ac.jp

per site (π) was calculated by sliding a 20-bp window [12]. A consecutive (>20 bp) region with less than 0.03 of the π value was defined as a conserved region among the mammals. It means that 0.6 nucleotide changes could be expected in 20 nucleotide sequences among mammals.

DNA preparation

DNA was extracted from peripheral blood or meat from samples (with asterisks listed in Supplementary Table) by using a QIAquick spin column (Qiagen, Hilden, Germany). For DNA preparation from a piece of buried bone of unknown species, the piece weighing 0.5 g was decalcified with 0.5 M EDTA overnight. It was then subjected to proteinase K treatment for 5 h at 56°C. After extraction of the supernatant with phenol, the DNA was purified by using the QIAquick spin column. For DNA preparation from hairs of human, rabbit, and guinea pig, DNA Extractor FM Kit (Wako, Osaka, Japan) was used.

PCR amplification and direct sequencing

PCR amplification was performed in a volume of 20 μl containing 10 ng of the extracted total DNA, 1× supplemented buffer; 0.2 mM concentration of each dNTP; 0.5 μM concentration of each oligonucleotide primer; and one unit of Ex *Taq* polymerase (Takara, Ohtsu, Japan). The primer sequences are given in Table 1. PCR was conducted for 35 cycles in a temperature block at 95°C for 30 s, 57°C for 15 s, and 72°C for 30 s. The PCR products were confirmed by 1.5% agarose gel electrophoresis and purified by using the QIAquick PCR purification kit (Qiagen). The purified PCR products were sequenced using BigDye terminator cycle sequencing kit version 1.1 (Applied Biosystems, Foster City, CA) and an ABI Prism 3100 genetic analyzer. Numbers of nucleotide substitutions per site were estimated using the two-parameter method [13]. A phylogenetic tree was constructed using the neighbor-joining method [14].

Results

Design of universal primer sets

Conserved regions spanning more than 20 consecutive base pairs were mapped on the whole genome through comprehensive multiple alignment of the 30 mammalian mtDNAs. A total of 15 regions, ranging from 20 to 40 bp, were identified as having an interspecies π value of less than 0.03. Among these conserved regions, five were located at the 12S rRNA locus and seven at the 16S rRNA locus (Fig. 1). The remaining three regions were located at the tRNA-Ile, tRNA-Met, and NADH dehydrogenase subunit 5 loci. In contrast, regions of the cytochrome b locus showed interspecies nucleotide differences with a π value greater than 0.03 (Fig. 1). To successfully amplify the DNA from severely degraded materials, we designed primers that could amplify the PCR products from 200 to 250 bp in size. As summarized in Table 1, two sets of oligonucleotide primers were designed, one for each locus, to yield products of 215 and 244 bp in humans.

PCR efficiency using the primer sets

Using the two universal primer sets, we performed PCR on the DNA from seven mammals (human, cow, horse, sheep, pig, dog, and mouse), nine nonmammalian vertebrates (chicken, ostrich, turkey, lizard, bullfrog, medaka, salmon, puffer fish, and carp), and two invertebrates (the sea squirt as a protochordate and the sea cucumber as an echinoderm). Products of a single band were obtained in all cases at the expected sizes except for the sea squirt and sea cucumber. Each DNA sequence of the products was identical with the one deposited in the database, except for two substitutions in lizard, two in medaka, and one in puffer fish in the sequenced fragment at the 12S rRNA. In addition, sequence comparisons of eight cartilaginous fishes, i.e., the rabbit fish (NC_003136), horn shark (NC_003137), sixgill stingray (NC_007230), gummy shark (NC_000890), ocellate spot skate (NC_007173), thorny skate (NC_000893), smaller spotted catshark (NC_001950), and spiny dogfish (NC_002012), retrieved from the database suggested that the two primer sets were appropriate even for them. These

Table 1 List of universal primers for vertebrate mtDNA, which was designed on the basis of a human sequence

Locus	Primer	Sequence (5′ to 3′)	Position	Temperature (°C)	Size (bp)
12S rRNA	L1085	CCCAAACTGGGATTAGATACCC	1064–1085	60.0	215
	H1259	GTTTGCTGAAGATGGCGGTA	1259–1278	61.1	
16S rRNA	L2513	GCCTGTTTACCAAAAACATCAC	2492–2513	58.5	244
	H2714	CTCCATAGGGTCTTCTCGTCTT	2714–2735	58.8	

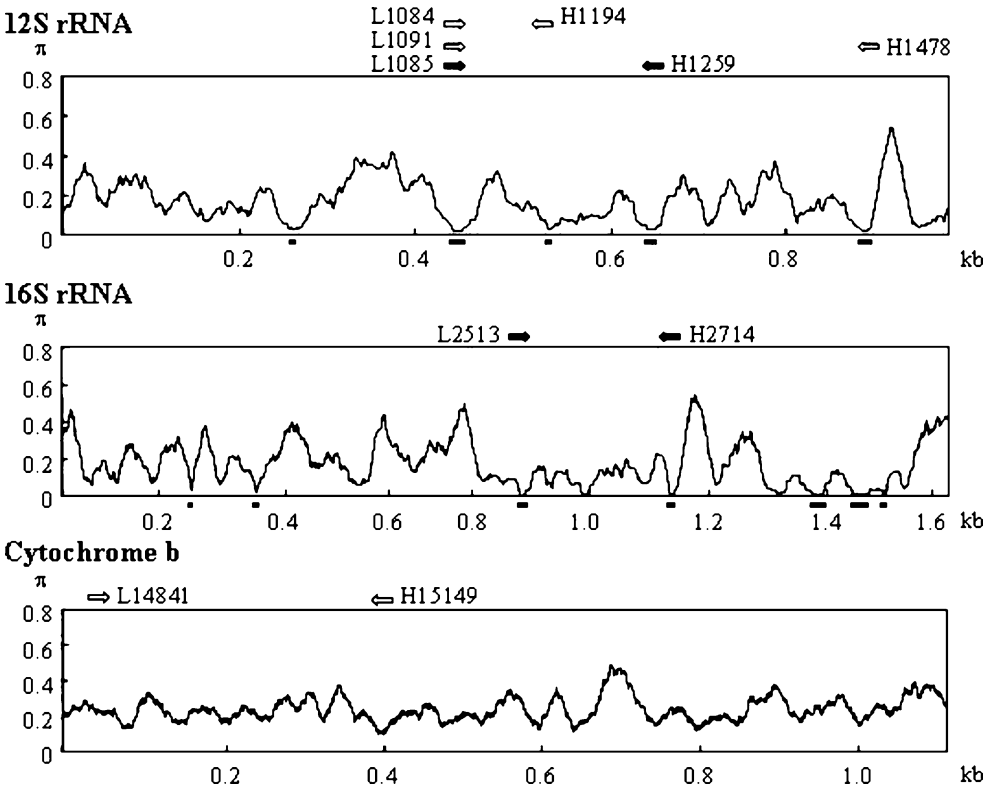

Fig. 1 Average numbers of pairwise nucleotide differences per site (π) among the 30 mammals for 12S rRNA, 16S rRNA, and cytochrome b loci. Window size of 20 bp and sliding range of 1 bp were used. *Bold black bars* indicate conserved regions. *Black arrows* indicate the positions of the present universal primers. *White arrows* indicate the position of primers used by Kocher et al. [3] and Balitzki-Korte et al. [8]

two sets of universal primers, therefore, appeared to be effective for almost all vertebrates.

Sequence analysis

The mean of the difference in the sequenced fragment at the 12S rRNA locus was calculated among 17 species, in which a chimpanzee sequence (NC_001643) from the database was also included. The mean of pairwise nucleotide differences in 215 bp was 29; the smallest one was 2 (between human and chimpanzee), and the largest was 46 (between lizard and medaka or puffer fish). Table 2 indicates the distance matrix for the sequenced fragment at the 16S rRNA locus. The mean of pairwise nucleotide differences in 244 bp was 39, with the smallest being 7 (between human and chimpanzee) and the largest 58

Table 2 A pairwise comparison of nucleotide differences (above the diagonal) and nucleotide substitutions per site by the two-parameter method (below the diagonal) for the sequenced segment in the 16S rRNA locus

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Human	–	7	25	21	29	28	27	28	58	47	58	49	39	48	40	44	42
2. Chimpanzee	0.04	–	21	18	26	26	24	28	57	48	56	46	37	48	39	44	42
3. Dog	0.15	0.12	–	14	12	14	16	29	50	40	50	48	43	45	40	44	39
4. Horse	0.12	0.10	0.08	–	13	16	14	28	49	42	50	40	35	48	39	44	43
5. Cow	0.17	0.15	0.07	0.07	–	13	14	32	50	43	51	49	41	48	41	47	44
6. Sheep	0.17	0.15	0.08	0.09	0.07	–	13	33	55	48	55	51	40	43	42	45	36
7. Pig	0.16	0.14	0.09	0.08	0.08	0.07	–	27	52	45	53	50	38	41	35	43	41
8. Mouse	0.16	0.16	0.17	0.16	0.19	0.20	0.16	–	54	43	53	43	44	49	42	52	46
9. Chicken	0.40	0.39	0.33	0.32	0.33	0.37	0.35	0.37	–	21	10	42	51	51	47	50	46
10. Ostrich	0.30	0.31	0.25	0.26	0.27	0.31	0.29	0.27	0.12	–	25	36	50	46	42	48	46
11. Turkey	0.41	0.39	0.34	0.33	0.34	0.38	0.36	0.36	0.06	0.15	–	41	53	52	50	51	44
12. Lizard	0.32	0.29	0.32	0.25	0.32	0.34	0.33	0.27	0.26	0.22	0.26	–	50	51	47	46	47
13. Bullfrog	0.24	0.23	0.27	0.21	0.26	0.25	0.23	0.28	0.33	0.32	0.35	0.32	–	38	31	41	36
14. Medaka	0.32	0.32	0.29	0.32	0.32	0.28	0.26	0.33	0.34	0.30	0.35	0.33	0.24	–	18	20	28
15. Puffer fish	0.25	0.24	0.25	0.24	0.26	0.27	0.21	0.26	0.30	0.26	0.33	0.30	0.19	0.10	–	19	24
16. Salmon	0.28	0.28	0.28	0.28	0.31	0.29	0.27	0.35	0.32	0.31	0.33	0.29	0.26	0.11	0.11	–	25
17. Carp	0.27	0.27	0.24	0.27	0.28	0.22	0.26	0.30	0.29	0.29	0.28	0.30	0.22	0.16	0.14	0.15	–

(between human and chicken or turkey). Although the latter region was more divergent than the former, the sequence obtained by either should contain sufficient information to determine the alleged species or taxon. As shown in Fig. 2, a phylogenetic tree was constructed by the neighbor-joining method for the sequenced fragment at the 16S rRNA locus. Even if we obtain a sequence that is not identical to any known sequence by this method, we expect to identify closely related species by constructing a phylogenetic tree.

Casework study

The present system for animal species identification was applicable for a practical casework. A part of fractured limb bone, approximately 14 cm long and 3 cm in diameter, was found in a yard during a construction project. More than a decade seemed to have passed after the remains had been buried. The bone looked to be that of a large mammal based on morphological observation, but the species could not be recognized by this means. PCR amplification of DNA extracted from a fragment of the bone yielded single products when the two sets of the universal primers were used. Both obtained nucleotide sequences were completely identical to the corresponding regions of pig mtDNA (Fig. 2). When one can obtain a sequence from an unknown species sample to be identified, the BLAST search can be used to find the sequence of a known species to which the sequence is homologous, although we constructed a phylogenetic tree in this study. In addition, we extracted DNA from hairs of human, rabbit, and guinea pig, and applied PCR direct sequencing them. Each DNA sequence of the products was identical with the one deposited in the database, except for one substitution in rabbit in the sequenced fragment at the 12S rRNA.

Discussion

Primers of broad reactivity are very beneficial to amplify DNA from an unknown origin. A number of studies have adopted the universal primers introduced by Kocher et al. [3] to address the cytochrome b locus [4–7]. However, it is evident that regions residing in the 12S and 16S rRNA loci in the mitochondrial genome among mammals are more strictly conserved than the region in the cytochrome b locus. These regions are located in both loop and stem portions in the secondary structure. The double-strand stem portion occupies 57 and 72% of the conserved regions of the 12S and 16S rRNA loci, respectively [15, 16]. It is likely that these conserved regions are functionally important. Using the highly conserved regions in this study, we were able to amplify DNA fragments as short as 215 and 244 bp in length. Although one primer set may not work well, another primer set might have a chance to obtain good results. It is for this reason that we designed two primer sets in this study.

In conclusion, we recommend this PCR direct sequencing system using either or both universal primer sets for the 12S and 16S rRNA loci for the identification of vertebrate species. Since using an automatic sequencer has become, PCR direct sequencing is not time-consuming and is even time-saving compared with the species-specific multiplex PCR [9, 10]. Direct sequencing of the PCR products allows comparative analysis of the unknown DNA sequences and those already deposited in the databases by using the BLAST program. These primer sets are universal from the standpoint of effective amplification of a wide variety of vertebrate mtDNAs and identification of highly degraded material.

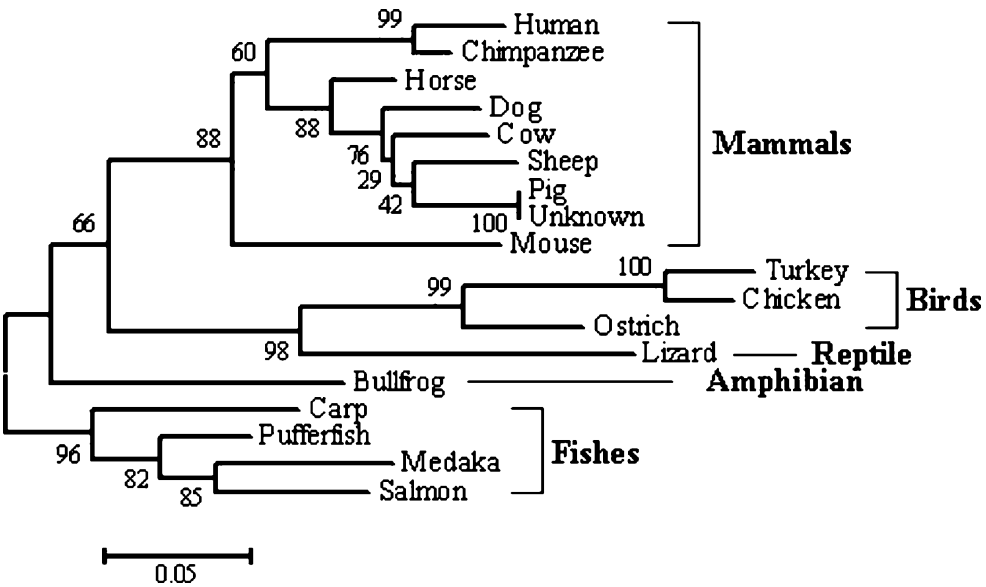

Fig. 2 A phylogenetic tree, constructed by the neighbor-joining method, for the sequenced fragment from the 16S rRNA locus. The *scale bar* corresponds to 0.05 nucleotide replacements per site. Percentage bootstrap values (based on 1,000 replications) appear at the *nodes* of the tree

References

1. Zehner R, Zimmermann S, Mebs D (1998) RFLP and sequence analysis of the cytochrome b gene of selected animals and man: methodology and forensic application. Int J Legal Med 111:323–327
2. Bravi CM, Liron JP, Mirol PM, Ripoli MV, Peral-Garcia P, Giovambattista G (2004) A simple method for domestic animal identification in Argentina using PCR-RFLP analysis of cytochrome b gene. Leg Med (Tokyo) 6:246–251
3. Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci USA 86:6196–6200
4. Parson W, Pegoraro K, Niederstatter H, Föger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. Int J Legal Med 114:23–28
5. Hsieh HM, Chiang HL, Tsai LC, Lai SY, Huang NE, Linacre A, Lee JC (2001) Cytochrome b gene for species identification of the conservation animals. Forensic Sci Int 122:7–18
6. Branicki W, Kupiec T, Pawlowski R (2003) Validation of cytochrome b sequence analysis as a method of species identification. J Forensic Sci 48:83–87
7. Bellis C, Ashton KJ, Freney L, Blair B, Griffiths LR (2003) A molecular genetic approach for forensic animal species identification. Forensic Sci Int 134:99–108
8. Balitzki-Korte B, Anslinger K, Bartsch C, Rolf B (2005) Species identification by means of pyrosequencing the mitochondrial 12S rRNA gene. Int J Legal Med 119:291–294
9. Bataille M, Crainic K, Leterreux M, Durigon M, de Mazancourt P (1999) Multiplex amplification of mitochondrial DNA for human and species identification in forensic evaluation. Forensic Sci Int 99:165–170
10. Dalmasso A, Fontanella E, Piatti P, Civera T, Rosati S, Bottero MT (2004) A multiplex PCR assay for the identification of animal species in feedstuffs. Mol Cell Probes 18:81–87
11. Thompson JD, Gibson TJ, Higgins DG (1994) ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680
12. Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, pp 254–286
13. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120
14. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425
15. Springer MS, Douzery E (1996) Secondary structure and patterns of evolution among mammalian mitochondrial 12S rRNA molecules. J Mol Evol 43:357–373
16. Burk A, Douzery EJP, Springer MS (2002) The secondary structure of mammalian mitochondrial 16S rRNA molecules: refinements based on a comparative phylogenetic approach. J Mamm Evol 9:225–252