Mitochondrial DNA analysis for the African elephant (Loxodonta africana) and the Asian elephant (Elephas maximus). : Genetic structure of the control region and population phylogeny

岡山, 俊直
九州大学比較社会文化研究科日本社会文化専攻

https://doi.org/10.11501/3159075

出版情報：九州大学, 1999, 博士（理学）, 課程博士
バージョン：
権利関係：

Mitochondrial DNA analysis for the African elephant (*Loxodonta africana*) and the Asian elephant (*Elephas maximus*).

---Genetic structure of the control region and population phylogeny---

（アフリカゾウ（*Loxodonta africana*）と アジアゾウ（*Elephas maximus*）の ミトコンドリアDNA解析

---コントロール領域の遺伝的構造と個体識別について---

Graduate School of Social and Cultural Studies,
Kyushu University

Toshinao Okayama

Mitochondrial DNA analysis for the African elephant (*Loxodonta africana*)

Mitochondrial DNA analysis for the African elephant (*Loxodonta africana*) and the Asian elephant (*Elephas maximus*).

--- Genetic structure of the control region and population phylogeny ---

(アフリカゾウ(*Loxodonta africana*)と
アジアゾウ (*Elephas maximus*)の
ミトコンドリアDNA解析
---コントロール領域の遺伝的構造と個体群系統について---)

**Graduate School of Social and Cultural Studies,
Kyushu University**

Toshinao Okayama

Abstract

The African elephant (*Loxodonta africana*) and Asian elephant (*Elephas maximus*) are endangered species, listed in Appendix I and II of CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). In this study, the genetic variabilities within both species, the population phylogeny of *L. africana* and the genetic conditions of the population of *L. africana* in South Africa were examined by sequencing the control region of their mitochondrial DNA.

Because *L. africana* and *E. maximus* are very large mammals and endangered species, it is not easy to obtain fresh samples for DNA analysis. So, feces, saliva, hair or ivory samples were used as samples for DNA analysis. In particular, feces samples were useful, and DNA was amplified even from feces preserved in very dry conditions.

Nucleotide sequences of the entire control region were determined for one individual of each species. The control region of the elephants were divided into three domains as for other mammals, the left domain which contains TAS (termination associated sequence) and CS (conserved sequence)-L, the central conserved domain which contains subsequence A-C, and the right domain which contains CSB (conserved sequence block) 1-3. A repeat sequence was detected in the right domain. The left domain which has the largest number of variable sites is expected to be the most suitable for population analysis.

A total of 213 samples of *L. africana* (50 individuals from zoo in Japan, 4 individual pieces of ivory imported to Japan, and 38 ivory and feces samples collected at Kruger National Park in 1996 and 1997, 24 feces samples collected at Addo Elephant National Park in 1997 and 97 feces samples collected at Tembe Elephant park in 1997 in South Africa) as well as 86 samples of *E. maximus* (52 individuals from zoo in Japan, and 34 individuals from Pinnawale Orphanage Park in Sri Lanka) were analyzed for population analysis.

L. africana showed 36 variable sites defining 18 haplotypes, from 351 bp of the left domain in the control region. These haplotypes were divided into two clusters (LI and LII), and cluster LII, which could be divided into two subclusters (LIIa and LIIb). Cluster LI consisted of haplotypes from forest elephants in central Africa and savanna elephants in eastern and South Africa. Cluster LII was composed of haplotypes from only savanna elephants. This seemed to be caused by the great mobility and high adaptability to different environments by the elephant. There are no geographical barriers within African continent which interfere with the movement of the elephant populations. *E. maximus* showed 23 variable sites defining 19 haplotypes. The haplotypes from Sri Lanka and Indian subcontinent overlapped with each other, suggesting that the Sri Lankan elephant is not clearly separated from the Indian elephant, according to mitochondrial control region analysis.

Four haplotypes from Kruger, 3 from Addo and 2 from Tembe were found. Haplotype diversities of each population were low, because the number of haplotypes detected was small and their frequencies were biased. In contrast, nucleotide diversities of each population were relatively high, because the sequence divergences of haplotypes within each park were relatively large. Elephants from all three parks had unique haplotypes which were not found for other parks. This may imply that a large population in the past was divided into small populations and experienced a population bottleneck effect. The population of *L. africana* in South Africa has increased satisfactorily over several decades due to successful conservation plans, the populations are presumed to have experienced a population bottleneck in the past, resulting low level of genetic diversity. For further conservation plans, programs should be followed that increase the genetic diversity in each of the elephant populations. Haplotype diversity in Sri Lanka was relatively high because seven haplotypes were found for Sri Lankan elephants. However, the nucleotide diversity in Sri Lanka was low, because the sequence divergences of haplotypes from Sri Lanka was small. This contrasts with that of *L. africana* in South Africa.

要旨

アフリカゾウ*Loxodonta africana*およびアジアゾウ*Elephas maximus*は，CITES（Convention on International Trade in Endangered Species of Wild Fauna and Flora）の付属書ⅠおよびⅡに記載されている絶滅危惧種である。本研究では，ミトコンドリアDNAコントロール領域を用いて，両種内の分子系統関係，アフリカゾウにおける個体群系統および南アフリカ共和国におけるアフリカゾウの遺伝的多様性の現状を把握することを試みた。

アフリカゾウおよびアジアゾウは大型哺乳類でありかつ絶滅危惧種であるため，DNA分析のための新鮮な材料を入手するのは容易ではない。そこで，糞，唾液および象牙をDNA分析の材料として用いる方法を検討した。これら全ての材料からDNAの抽出および増幅が可能で，特に糞はかなり乾燥した状態のものでも，DNAの増幅が可能であった。

分析領域を選定するため，両種各1個体に関してミトコンドリアDNAコントロール領域の全領域の塩基配列を決定した。両種のコントロール領域は他の哺乳類と同様にTAS（termination associated sequence)およびCS（conserved sequence）-Lを含むレフトドメイン，サブシーケンスA-Cを含むセントラルコンサーブドドメインおよびCSB（conserved sequence block）1-3を含むライトドメインの3つのドメインに分けられた。またライトドメインには，反復配列が認められた。コントロール領域のうちレフトドメインが最も多くの変異座位を含んでおり，集団内の遺伝的分析には最も分解能が高いと期待される。よって全サンプルについてのハプロタイプ検出には，レフトドメイン全体を含む351bpを用いた。

分析された試料はアフリカゾウ213試料（日本国内の動物園飼育個体50個体，日本に輸入された象牙片4試料，1996年および1997年に南アフリカ共和国のクルーガー国立公園において採取した象牙片および糞試料38試料，1997年にアドゥ国立公園において採取した糞試料24試料，および1997年にテンベ国立公園において採取した糞試料97試料）およびアジアゾウ86試料（日本国内の動物園飼育個体52個体およびスリランカのピンナワラ公園34個体）である。

アフリカゾウからは18のハプロタイプが得られた。それらは２つのクラスター（LIおよびLII）に分けられ，クラスターLIIはさらに２つのサブクラスター（LIIaおよびLIIb）に分けられた。クラスターLIIは，サバンナゾウの生息地からの個体のみから成っていたが，クラスターLIにはシンリンゾウの生息地からの個体とともに，サバンナゾウの生息地である東アフリカと南アフリカ共和国の個体が含まれた。これはアフリカゾウが，非常に高い移動能力と環境適応能力を持っており，かつアフリカ大陸には個体群の移動を妨げるような地理的障壁がないためであろうと推測された。アジアゾウからは19のハプロタイプが検出されたが，今回の結果では，スリランカ産個体から得られたハプロタイプとインド亜大陸産個体から得られたハプロタイプは互いに重複しており，コントロール領域の塩基配列からスリランカ産個体とインド亜大陸産個体を明確に区別することは困難であった。

クルーガー国立公園からは4つの，アドゥ国立公園からは3つの，テンベ国立公園からは2つのハプロタイプが検出された。ハプロタイプ数が少ないため，ハプロタイプ多様度は分析されたアフリカゾウ全体に比べて低いものの，いずれの公園からも別のサブクラスターに属する遺伝的に離れたハプロタイプが検出されたため，塩基多様度はそれほど低くないことが明らかになった。さらに３つの国立公園からは，それぞれ他では見られない独自のハプロタイプが検出された。これらのことは，かつての大きな連続していた個体群が分断化され，またボトルネックを経たことに対する遺伝的な反映であろうと推測された。南アフリカ共和国では過去数十年間にわたるアフリカゾウの保全策が成功し，個体数は順調に増加してきたが，遺伝的には個体数が増加した現在でもボトルネックの影響が残っていることが明らかになった。今後は遺伝的多様性を増加させる方向での保全策が必要と思われた。スリランカにおけるアジアゾウからは7つのハプロタイプが検出されたためハプロタイプ多様度は，かなり高い値となったが，全てのハプロタイプ間の塩基置換率が小さかったため，塩基多様度は比較的低い値となり，この傾向は南アフリカ共和国におけるアフリカゾウとは対照的であった。

Contents

List of tables

List of figures

General introduction

The African elephant *Loxodonta africana* (Blumenback, 1797) once occurred throughout African continent, from the Mediterranean Sea to the Cape of Good Hope, and was absent only in the Sahara and other desert regions (Nowak, 1991). *L. africana* occupies a variety of habitats including savanna, rain forests, swamps, seashores and mountains. Over the past decade, populations of *L. africana* have declined dramatically. Habitat destruction and hunting over the last century has severely fragmented many of the formerly large populations. Significant populations are now confined to well protected in national parks and reservation areas (Said and Chunge, 1994).

About 6,000 years ago, the range of the Asian elephant *Elephas maximus*, Linnaeus, 1758 extended from the Euphrates - Tigris river system in the west, eastward through Asia, south of the Himalayas to Indochina and most of China as far north as the Yellow river (Olivier, 1978). Since the turn on the century, *E. maximus* declined both in range and number and according to Santiapillai and Jackson (1990) the species survives as a number of discontinuous populations in Sri Lanka, India, Nepal, Bhutan, Bangladesh, Myanmar, Thailand, Laos, Cambodia, Vietnam, Malaysia (both Peninsular Malaysia and Sabah), Indonesia (both Sumatra and Kalimantan) and China.

In 1989, the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) determined the transfer of *L. africana* and *E. maximus* from Appendix II to Appendix I. Populations of *L. africana* in southern Africa currently appear to be stable or increasing (Republic of South Africa, 1994), and in 1998's CITES meeting Zimbabwe, Namibia and Botswana proposed for *L. africana* to be transferred from Appendix I to Appendix II to export tasks,

which were already stocked in their countries, to Japan after approvement with some qualifications.

The population of *L. africana* in South Africa is now nearly 10,000 individuals but formerly the numbers of population in South Africa decreased to less than 100 (Hall-Martin, 1992). Therefore, *L. africana* in South Africa went through population bottleneck. It is interesting to observe the genetic conditions for such populations, and genetic data is important to make a conservation plan for *L. africana* in South Africa.

Mitochondrial DNA (mtDNA) has been widely used for molecular genetic studies (Avise *et al.* 1994, 1996), since sequence analysis of nuclear DNA is complicated because of recombination that rearranges genes. Because mtDNA is maternally inherited and does not recombine, mtDNA sequence analysis can provide precise divergence of the species phylogeny. Furthermore, mtDNA evolves rapidly at the sequence level, due in part to lack of known repair mechanisms for mutations that arise during replication (Wilson *et al.* 1985), therefore nucleotide substitutions accumulate quickly even between species that diverged quite recently.

The control region is characterized by rapid change in sequence and length apart from particular sequence blocks which is conserved among taxa (Saccone *et al.* 1991). This region evolves three to five times faster than the rest of the mtDNA genome (Aquadro and Greenberg 1983). Therefore, this region is commonly used to intraspecies and population genetics. For example, peopling of human being is extensively analyzed based on the variation of the mitochondrial control region (Horai *et al.* 1991, 1996). But until now there are a few reports on the control region of the elephants and no population survey with this molecular marker has been made in the elephants.

In the part 1 in this study, the entire nucleotide sequences of the control region for *L. africana* and *E. maximus* were determined to find the most suitable region for population analysis of both species in the control region. Because *L. africana* and *E. maximus* are very large mammals and endanger species, it is not easy to obtain fresh samples for DNA analysis. So, feces, saliva, hair or ivory samples were used as samples for DNA analysis. In the part 2, to examine the genetic variabilities within *L. africana* and *E. maximus*, the population phylogeny of *L. africana*, and the genetic conditions of the population of *L. africana* in South Africa, samples from zoo in Japan, three national parks in South Africa were analyzed. In comparison with *L. africana*, *E. maximus* from zoo in Japan, and Sri Lanka were also analyzed.

Part 1. Structure of the mitochondrial control region for the elephants

1. Introduction

For genetic analysis on wild populations of large mammalian species, it is essential to use appropriate ways not to be harmful to animal themselves. In that case, it might be preferable to use non-harmful materials that once discarded from their bodies, such as hairs, feces and saliva. These materials have been already successfully used in populational genetic studies in various mammalian species (Höss *et al.* 1992; Kohn *et al.* 1995). Ivory of the African elephants (Koike *et al.* 1998) is available in reasonable quantities in their native countries, and the establishment of genetic examination with such materials would be highly desired.

An important point for the genetic survey is the use of appropriate DNA markers for population phylogeny. The mitochondrial control region is known to evolve rapidly compared to other regions of mitochondrial DNA (mtDNA) (Brown 1985), the control region is likely to be the best candidate for that purpose (Horai *et al.* 1991, 1996). To avoid contamination of DNA from other species in the materials, it needs to develop PCR primer sets specific to each species or species group. Another advantage is that mtDNA is abundant even in these samples only small amounts or fragments of DNA are preserved. This underscores the importance of mtDNA analysis for which no living tissue is required. These circumstances facilitate designing of original primers and drawing of phylogenetic tree.

Animal mtDNA contains the control region that is essential for the replication of the mtDNA. A new synthesis of the heavy (H) strand of the

mtDNA starts from an initiation site of the control region, H-strand replication origin (OH), and terminates at a site ca. 700 nucleotides downstream from OH. This leads to a partial relaxation of parental supercoiled molecules because the short newly synthesized H strand remains associated with the template, thus creating a triple-strand structure known as a displacement loop (D-loop) (Doda *et al.* 1981; Mackay *et al.* 1986). The mitochondrial control region is known to have several conserved sequence elements that are associated to the initiation and termination of the replication, and also have other ubiquitous sequence blocks, a long stretch of highly conserved domain at a central part of the control region. The central conserved domain (CCD) demarcates the control region into three parts, CCD and flanking left and right domains (Walberg and Clayton 1981; Saccone *et al.* 1991; Gemmell *et al.* 1996).

The initiation of H-strand DNA synthesis is thought to involve the OH and conserved sequence blocks (CSBs) (Brown *et al.* 1986). There are three CSBs (CSB 1-3), which locate in the right domain, are well defined through comparison of rat, mouse and human sequences (Walberg and Clayton 1981). The CSBs are the targeting sites of mitochondrial transcriptional factor 1 (mtTF1) and the cleavage sites by RNase for mitochondrial RNA processing (RNase MRP) (Clayton 1991). Then H-strand DNA replication is primed by the transcribed RNA segment. Till now, such primary functional significance of the CSBs is likely to be well assigned but the meaning of the evolutionary variability is not well understood. The number of the CSBs apparently differs from species to species. For example, CSB-3 is missing in platypus, cow and cetacea (Southern *et al.* 1988; Hoelzel *et al.* 1991; Dillon and Wright 1993) or fused with CSB-2 in sheep (Wood and Phua 1996), cervidae (Douzery and Randi 1997), cow (Steinborn *et al.* 1998), giraffe and buffalo (Ghivizzani *et al.* 1993). Thus, what

all the CSBs have some function is open to question.

The termination of the H-strand replication is governed by moderately conserved repetitious sequences, termed termination associated sequences (TASs), which are located in the left domain (Doda *et al.* 1981). The stop signals of D-loop strand, TAS, showed the consensus sequence as TACATtAAAaYYYAAT (Foran *et al.* 1988). Gemmell *et al.* (1996) suggested that first five nucleotide (TACAT) are comparatively well conserved but the latter half is not so similar between each TASs with indels and nucleotide substitutions. In addition to the sequence variation, there are considerable variation in number of the repeats among mammalian species. Till now, however, the evolutionary trends of the variation of the TAS elements and their functional significance are not well understood.

The CCD (ca. 250 bp long) is one of the best conserved parts of mtDNA. This domain contains three conserved sequence (subsequence A, B and C), each of which could be formed stable secondary structures (Gemmell *et al.* 1996) and exhibits conservative amino acid motif between species (Saccone *et al.* 1987). Therefore, certain function can be afforded to the CCD but no positive evidence of functionality has been assigned.

In contrast to the highly conservative nature of CCD, regions for repetitive sequences (RS) are highly evolutionarily unstable. Such repetitive sequences can be recognizable in the left and / or right domain of the control region in many animals including mammals (rabbit, Mignotte *et al.* 1990 and Saccone *et al.* 1991; Cetacea, Hoelzel *et al.* 1991; Japanese monkey, Hayasaka *et al.* 1991; bat, Wilkinson and Chapman 1991 and Wilkinson *et al.* 1997; harbor seal, Árnason and Johnsson 1992; pig, Ghivizzani *et al.* 1993; 18 carnivore spp., Hoelzel *et al.* 1994; shrew, Stewart and Baker 1994 and Fumagalli *et al.* 1996).

There is a conspicuous difference between individuals with regard to the size and organization of this region. In addition, it has been revealed that the RSs are involved in the generation of high degree of mitochondrial heteroplasmy in various animal species. Evolutionary mode of the simple tandem repeats and their functional significance are of large concern of molecular phylogenetists. It is also desired to find reasonable explanatory way for the biological meaning and differentiation mode of the heteloplasmic variation.

Here, in the study of the part 1, to know the evolutionary trends of the control region and to assess the functional significance of the four sequence members, CSBs, TASs, CCD, and RSs, I determined the entire sequences of the control region in the two extant species of elephants, the African (*Loxodonta africana*) and Asian (*Elephas maximus*) elephants. Further I examined the mode of heteroplasmic variation of the RSs in *L. africana* using the PCR mediated subcloning method. Comparison of these sequences obtained here and those of other mammalian species from literature would provide a new perspective on the evolutionary rends of the control region and functional significance of the four sequence members, CSBs, TASs, CCD, and RSs.

2. Materials and methods

2-1. DNA extraction (Appendix B - 1)

The materials used in this study were ivories, feces, saliva and hair roots of *L. africana* and *E. maximus*. Under a microscopic observation of thin sections of the ivory, it was found that main composition of the ivory was dentine without living cells but cementum layer which is the surface of the dentine, contained living cells. About 20 mg of powdered ivory scraped from the surface near tooth root was used for DNA extraction.

For the fresh fecal samples which had wet surfaces, the mucous layer on the surface of feces was scraped with a cotton bud, and the cotton bud was put into a microtube with extraction buffer (10 mM Tris-HCl, 10 mM NaCl, 25 mM EDTA). For the dried fecal samples, about 20 mg of the dried mucous layer scraped from the surface was used. For saliva samples, mucous layer in the mouth was scraped with a cotton bud. In the case of fair samples, only hair root with follicle was cut and placed in a microtube with buffer.

For the protein digestion, the sample was suspended with 310 ml of extraction buffer, 15 ml of 10% SDS and 25 ml of 20 mg / ml Proteinase K and incubated at 55°C for 2 hours on a rotator. For ivory samples EDTA was added to 50 mM in final concentration. Then the nucleic acids were extracted using the IsoQuick Nucleic Acid Extraction Kit (ORCA Research Inc.; U. S. A.). The extracted nucleic acids were precipitated with iso-propanol or ethanol.

An example of electrophoresis images of the extracted DNA was shown in the Fig. 1-1. In the most of the fecal samples, sizes of the nucleic acids ranged from 20 k bp to 100 bp. In the ivory samples, DNA less than 800 bp were obtained. High molecular weight DNA around 20 kb seen in fecal samples was possibly obtained from other materials such as intestinal bacteria.

2-2. Designing of the PCR primers for the elephants

To avoid amplifications of the contaminated DNA from other organisms by the polymerase chain reaction (PCR) method, elephant-specific primers were designed. For the control region, an universal primer H651 (Irwin *et al.* 1991) was effective for the elephants but L15926 (Irwin *et al.* 1991) was not effective. So, we designed L15926.ele referring L15926. After the entire control region was amplified with L15926.ele and H651 for the elephants, the primers Lox L1,

Lox L2, Lox L3, Lox L4, Lox H1, Lox H2 and Lox H3 were designed by walking method (Fig. 1-2). These specific primers for elephants were used for the amplifications and sequencing.

2-3. DNA amplification using PCR (Appendix B - 2)

The PCR were conducted with a PCR thermal cycler (Takara, model MP) in 30 cycles of denature at 95°C for 15 sec, annealing at 55°C for 30 sec and extension at 72°C for 60 sec. When the first PCR produced poor amplification, one tenth of the first PCR product was used as templates for the second PCR. PCR products were purified with Bio-Spin Disposable Chromatography Columns (Bio-Rad).

Fig. 1-3 shows an example of an electrophoresis image of the PCR products using feces and ivory samples of *L. africana*. Using two primer sets of L15926. ele and H 651, and L15926.ele and Lox-H2, about 1700 bp and 800 bp fragments were amplified, respectively. The 800 bp fragments were amplified in all the samples, feces, saliva, hair or pieces of ivory. The 1700 bp fragments were amplified in cases of fresh feces samples.

2-4. Cloning of the control region using T vector (Appendix B - 3)

Many of the mammalian species has a heteroplasmy in the repetitive region (e. g. Mignotte *et al.* 1990; Hayasaka *et al.* 1991; Hoelzel *et al.* 1994). For the repetitive region of the elephants, nucleotide sequences could not be determined by direct sequencing. This might be also because of the heteroplasmy like other mammals. Therefore, DNA cloning of the control region containing the repetitive region was conducted with TOPO TA Cloning Kit (Invitrogen). PCR products between L15926.ele and H651 which contain whole control region

from one individual of *L. africana* and *E. maximus* were subcloned into a pCR 2.1 TOPO-vector and transformed into TOP 10 strain of *E. coli* (Fig. 1-4).

To test the possibility of slippage or recombination during amplification, clones were re-amplified by the PCR.

2-5. Sequencing with an automatic sequencer (Appendix B - 4, 5)

Sequencing was conducted by R.O.B. DNA Processor (Pharmacia) with the Thermosequenase Cycle Sequencing Kit (Amersham) using Cy5-labeled fluorescence primers with the same sequences that used in the PCR. The cycle was repeated 25 times with denaturation at 95°C for 30 sec, annealing and extension at 65°C for 30 sec. Sequencing was conducted with ALFred DNA Autosequencer (Pharmacia) at 47°C, current of 26 mA and voltage of 1200 V.

Cy5 - labeled L15926.ele and Lox - H1 were used to read sequences of all the samples for L - and H - strand, respectively. The fragments of the control region ca. 500-600 bp were sequenced with these primers. To determine the sequence of the entire control region for one individual for each species, Cy5 - labeled L15926.ele, Lox L1 - L4 and LoxH1 - H3 were used. An example of output of the DNA sequencing for the repetitive region using the feces sample is shown in Fig. 1-5.

3. Results

To compare the control region sequence, it is expected that important informations such as distributions of variable sites and putative functional region are revealed. Generally, the control region is so diverged that it is not easy to align the sequences with distantly related species (Saccone *et al.* 1991). It is possible to align the sequences of control region only with closely related species.

But now there are no closely related species with the elephants. As a result of a homology search to data bases, only the control region sequence of manatee (AF04615, Gracia-Rodoriguez *et al.* 1998) could be aligned with those of the elephants .

3-1. General structure of the control region of the elephants

The length of complete control regions encompassing from $tRNA^{pro}$ to $tRNA^{phe}$, without purine-pyrimidine altering regions, in *L africana* (93EL50) and *E. maximus* (93EL25) were 1040 bp and 1041 bp, respectively (Fig. 1-6). In this region an indel (insertion or deletion) was found in only one site, position 7. Base substitutions between *L. africana* (93EL50) and *E. maximus* (93EL25) were 22 in the left domain, one in the CCD and six in the right domain, respectively.

In the control regions of *L. africana* and *E. maximus*, CCD and both left and right domains were distinguished. In this study the start and end points of the CCD are followed with Gemmell *et al.* (1996) and Saccone *et al.* (1991), respectively. In *L. africana* left domain was located at position 1-333, CCD was located at position 334-575 and right domain was located at position 576-1455.

3-2. Central conserved domain (CCD)

The CCD is the region which is about 250 bp in length and very conserved between species. The sequences of the CCD in *L. africana* and *E. maximus* showed high affinity with those of other mammals (Fig. 1-7). Subsequences A, B and C (Gemmell *et al.* 1996) were also detected in *L. africana* and *E. maximus*.

3-3. Termination associated sequence (TAS)

TAS element in which sequences are moderately conserved, are

ubiquitously observed in vertebrate mtDNA. In *L. africana* and *E. maximus*, there were four TAS-like sequences, TAS-1 (positions 50-61), TAS-2 (positions 86-97), TAS-3 (positions 241-252) and TAS-4 (positions 268-279) (Fig. 1-8). For the comparison of the sequences of 213 samples of *L. africana* and 86 individuals of *E. maximus*, 44 variable sites were detected along with the sequence of the left domain. TAS-1 and -2 have no such variable site but TAS-3 in both species and TAS-4 in *E. maximus* have variable sites. Comparison of TAS positions in the sequences of the elephants and manatee (AF04615, Gracia-Rodoriguez *et al.* 1998), indicated that TAS-1 was absent from manatee (Fig. 1-9). The alignment of the TAS sequences for *L. africana* and *E. maximus* with other mammals were shown in Fig. 1-10.

Saccone *et al.* (1991) indicated that short mirror symmetries (TACAT, ATGTA) repeated several times in the left domain. In this study 213 samples of *L. africana* and 86 individuals of *E. maximus* were analyzed. As a result, four or five of TACAT and three of ATGTA were found in *L. africana* and four or five of TACAT and two of ATGTA were found in *E. maximus*. One-pair of these palindromic sequence was found to form a stable secondary structure (position 86-105) (Fig. 1-11A).

3-4. Conserved sequence L (CS-L)

Repetitive sequence block were found in the left domain of both *L. africana* and *E. maximus* . These repetitive block were relatively conserved. The repetitive block, designated as CS-L, is about 33 bp in length and start with ATGTA which is symmetry of TACAT (Fig. 1-12). The CS-Ls were repeated two times in the left domain of *L. africana* and *E. maximus* (CS-L1; 16-49, CS-L2; 101-133). Although CS-L1 has one variable site in *L. africana* and two in *E. maximus*,

CS-L2 is the same sequence in both species. The alignment of the sequences of the CS-Ls for *L. africana* and *E. maximus* with other 16 species was presented in Fig. 1-13. Similar sequence blocks were found even in birds.

In the species which has repetitive sequences (RS2, Hoelzel *et al.* 1994) in left domain, the structure of the repetitive sequences and CS-Ls were closely related (Fig. 1-14). In dasypus which has repetitive sequence of three times and half, repetitive sequence starts and ends with CS-L. In cat, incomplete 4th repeat terminated just before CS-L. In crocidura, CS-L was located the end of repeat motif.

In many species, CS-L has short mirror symmetry (TACAT, ATGTA) (Saccone *et al.* 1991) or TGCAT and ATGTA. These short symmetry or sub-symmetry formed relatively stable secondary structure (Fig. 1-10B).

The sequence of the termination of D-loop strand is TCCCC in pig, GCCCC in cow (MacKay *et al.* 1986; Madsen *et al.* 1993) and GCCCC or ACCCC in Cervidae (Douzery and Randi 1997), the homologous sequence is included at 3' end of CS-L. In the case of the elephants, these sequences were TCCCC in CS-L1 and ACCCC in CS-L2.

3-5. Conserved sequence block (CSB)

In the right domain there are conserved sequence blocks, CSB-1 to 3, which were found by Walberg and Clayton, (1981). These three CSBs were also found in *L. africana* and *E. maximus*, CSB-1: 683-709, CSB-2: 1184-1202, CSB-3: 1248-1265. The alignment of the sequence of the CSBs for *L. africana* and *E. maximus* with other mammals were shown in Fig. 1-15.

3-6. Repetitive sequence

The repeat regions were found between CSB-1 and CSB-2 and composed of purine-pyrimidine alternation To determine the nucleotide sequences of the repetitive region, the entire control region containing the repetitive region was cloned and sequenced. In *L. africana* 10 clones from 93EL50 and three clones from 96LX79 were determined. All these clones had different sequences with length variation from 142 bp to 430 bp. In *E. maximus* 14 clones from 93EL25 were determined. Fourteen clones were classified into 10 types.

To make it simply, AT, AC, GT and GC were replaced with "1", "2", "3", and "4", respectively and after that "412" and "322" were replaced with "5" and "6". (Fig. 1-16 and 17). As a result repetitive sequence of *L. africana* was divided into six domains (A-F) (Fig. 1-18). Domain A was basically "4621255", domain B was basically three repeat of "666462", domain C was a large number of "6", domain D was irregular insertion found only in three clones, domain E was a large number of "5" and domain F was "663123". The repetitive sequence of *E. maximus* was divided into five domains (A-E) (Fig. 1-19). Domain A was "432254115", domain B was three to six times repeat of "3222" or "32222", domain C was basically a large number of repeat of "42_n" or "32_n", domain D was a large number of "5" and domain E was basically "3123123".

The repeat motif 5 which were found in both *L. africana* and *E. maximus* and motif 6 which was found in *L. africana* could be made relatively stable secondary structure (Fig. 1-20). Models of addition and loss of repeat units were infrared by secondly structure of double strand during replication (Fig. 1-21). Addition of repeat motif of the newly synthesized daughter strand would be formed when the length of the hairpin loop of the daughter strand is longer than that of template strand. Loss of repeat motif should be formed when the length of hairpin loop of the newly synthesized daughter strand is shorter than that of

template strand. Consequently, variation of tandem repeat was developed exactly with repeat motif through successive replication.

4. Discussion

4-1 Nonrepetitive region of the elephant control region

4-1-1 Central Conserved domain (CCD)

The CCD of the elephants were also able to align with those of other mammals. Since the CCD is the only long conserved sequence in control region, some functions were presumed. But the idea as to function of the CCD has not confirmed yet.

Saccone *et al.* (1987) suggested that there is a open reading frame (ORF) in the CCD and the RNA or peptide coded here participated in transcription or replication of mtDNA. It was possible to set the ORF in the CCD of the elephants. But in comparison with many mammalian ORFs, the conserved motif of amino acid was only "LSFLRAH" (Saccone *et al.* 1987), furthermore the start and stop points of the ORF were different depend on species. Therefore, it may be difficult to set up the functional ORF (Hoelzel *et al.* 1991).

4-1-2 Termination associated sequence (TAS)

Madsen *et al.* (1993) reported that seven TAS-like sequences exist in the bovine control region but only one is selectively bound to 48 kDa protein. They said it seems likely that TAS binding is species specific. So, it is expected that the protein binding TAS element has no mutations within species. In this study I tried to ascertain the functional one among TASs of *L. africana* and *E. maximus* to identify the variable sites.

In this study four TAS-like sequences were found in the left domain of each species. Among these TAS-like sequences TAS-1 and TAS-2 had no variable site but TAS-3 in *E. maximus* and TAS-4 in both species had variable sites. And the manatee which is one of the most closely related species with the elephants, lacks TAS-1 exactly. Consequently, TAS-2 which has no variable site in the elephants and also found in manatee has highest possibility among four TAS-like sequences to have some functions in the control region of the elephants.

4-1-3 Conserved sequence L (CS-L)

CS-Ls were start with ATGTA which is one of the short mirror symmetries and about 33 bp long. This was found not only in mammals from many taxa but also in birds. The position of CS-L in cat, dasypus and shrew seemed to have some relation with repeat (RS2 in Hoelzel *et al.* 1994) structure. Short mirror symmetries in CS-L were able to form relatively stable secondary structure. These secondary structures associated with the CS-L might have some relation with the formation of RS2.

Generally, the sequence of the left domain can be aligned among the animals belongs to the same family or order. But it is very difficult to align the sequence of the left domain among interfamilial or interordinal animals because large indels specific to the taxa (Saccone *et al.* 1991) rather than nucleotide substitutions. So, the left domain seems to have the nature that it is easy to happen large scale insertion and deletion when two species diverged some extent.

4-1-4 Distributions of the variable sites in nonrepetitive region

The CCD, highly conserved during evolution, behave as a good molecular clock and reliable estimates of the times of divergence between closely and

distantly related species were obtained form CCD (Saccone *et al.* 1991). However, the variable site between *L. africana* and *E. maximus* was only one, therefore the CCD is too conservative to evaluate the subspecies or population level for the elephants.

The right domain without repetitive region which had only six variable sites was also highly conserved between *L. africana* and *E. maximus*.

Although the left domain had some conserved sequences such as TASs and CS-Ls which are putative functional region, other parts in the left domain had many variable sites within and between *L. africana* and *E. maximus*. The left domain was most variable in the three domain of the control region without repetitive region. Therefore, the left domain is most suite to investigate genetic variety within species or inter populations in the elephants.

4-2 Repetitive sequences of the elephant control region

4-2-1 Heteroplasmy of the repetitive sequence

In case of PCR for repetitive sequences, it should be remind the artifacts by amplifications. In carnivore DNA amplified from whole cell preparations migrated as a smear of various-size fragments, while DNA amplified from clones migrated as a single band in agarose gel electrophoresis (Hoelzel *et al.* 1994). These experiments do not rule out the possibility of slippage or recombination during amplification, but suggest that such events are rare. In this study the PCR product including repetitive region were isolated by cloning and after that the PCR product using the clones as templates showed clear single bands (Fig. 1-22). Therefore, the sequences of each clone determined in this study were considered in nature, not the artifacts.

It was found that there are repetitive sequences in mitochondrial control

region in many animals including vertebrates and invertebrates (reviewed Lunt *et al.* 1998). These repetitive sequences were classified into five types, RS1 to RS5, according its location and the length of the motives (Hoelzel *et al.* 1994).

The repetitive sequence of the elephants were between CSB-1 and CSB-2, which is RS3 type. The species which has RS3 were rabbit (Mignotte *et al.* 1990; Dufresne *et al.* 1996; Casane *et al.* 1997), seals (Árnason and Johnsson 1992; Hoelzel *et al.* 1993), pig (Ghivizzani *et al.* 1993), 18 carnivore (Hoelzel *et al.* 1994) and so on.

A significant feature of the structure of the repetitive region in *L. africana* and *E. maximus* were 1) it was composed of complete alteration of purine - pyrimidine sequence, 2) it was not a complete tandem repeat of a same motif but a imperfect repeat of several motives, 3) the repeat motives were relatively short. These characteristics were similar to carnivore and seals.

4-2-2 Repetitive sequences as genetic markers

Árnason and Johnsson (1992) suggested that repeat region has a potential to provide a most powerful genetic marker in population studies because of high mutation rates expected than other parts in control region. On the other hand, the nature of mitochondrial repeats have not fully elucidated yet. For example, the number of repeat were different between somatic and germ cells in rabbit (Dufrense *et al.* 1996; Casane *et al.* 1997). Therefore, although there is a considerable potential for mitochondrial repeat pilot studies to assess genetic diversity, the utility of the technique to resolve broader questions in molecular ecology should be treated cautiously until such a time as the system is better understood (Lunt *et al.* 1998).

To use the mitochondrial repetitive regions for the analysis of population

or family group phylogeny of the elephants, it is necessary to make the nature of the repeat clear first, the extent of the heteroplasmy, the variation of the number of the repeat depend on the tissue, is it a mother inheritance with repeat number and so on. After that it should be established the method to analyze the repetitive region.

Part 2. Population phylogeny of the elephants

1. Introduction

1-1. Historical background of the elephants

Population history for *L. africana*

L. africana once occurred throughout African continent except for the Sahara and other desert regions. But now large populations are confined to protected areas like national parks or game reserve (Fig. 2-1, A). Now the total population is estimated as about 500,000 individuals (Said and Chunge, 1994).

In South Africa: The recorded distribution of *L. africana* (Hall-Martin, 1992) begins with a report of their occurrence on the Cape south coast by Vasco da Gama. In about 1650's, *L. africana* are reported to have occupied much of South Africa except the interior of the country (Fig. 2-2, A). The arid central lands are unlikely to have supported large or permanently resident populations. The savanna and woodland regions of Transvaal and the high rainfall areas on the east Coast as far south as the Cape of Good Hope, however, supported large populations of *L. africana.*

By about the end of last century, the large populations of *L. africana* had been wiped out. The main force eliminating *L. africana* was the shooting for ivory trade and crop protection. The distributions of the wild populations in early years of this century were only four, Kruger, Addo, Tembe and Knysna

(Fig. 2-2, B). Since then the distribution of *L. africana* has changed dramatically as populations increased.

In Kruger National Park: The population estimates are plotted in Fig. 2-3, A. First record of *L. africana* in Kruger was in 1905 at Oliphant bridge. The population increase during 1900-1960 indicated mostly a reflection of natural recruitment. From 1960-1970 the increase was very high which is accounted for by recruitment, and also massive immigration due to heavy hunting pressure and drought in Mozambique. Kruger National Park has been fenced off from Mozambique since 1974. After the 1970's the population has been kept fairly stable at around 7,500 with hundreds of individuals were culled per year until 1996.

In Addo Elephant National Park: The population size from 1931 to 1991 is shown in Fig. 2-3, B. As the earliest record about populations of *L. africana* in Addo was 11 in 1931, which is the known lowest number in Addo up to now. The population of *L. africana* in Addo (Hall-Martin 1980) has been intensively studied and most of the animals are individually known. After the completion of the fence in 1954 which confined *L. africana* to the Park, the rate of increase had been gradually raised and the population size had grown to 162 in 1991 (Said and Chunge, 1994).

In Tembe Elephant Park: The population data is summarized in Fig. 2-3, C. There were reports of breeding herds in Tembe before 1970's, and only bulls

and no breeding herds in the middle of the 1970's (Hall-Martin 1980). The known minimum number of *L. africana* in Tembe was 16 individuals in 1971. After the middle of the 1970's, heavy poaching of *L. africana* in Mozambique caused a movement of *L. africana* from Mozambique to Tembe Elephant Park. Now the border to Mozambique was sealed with fences which stopped movements of *L. africana*. The number of *L. africana* was 80 in 1990. The apparent decline between 1978 and 1986 is due to population counting method on minimum helicopter counts.

Population history for *E. maximus*

Formerly the range of *E. maximus* extended from the Euphrates - Tigris river system in the west, eastward through Asia, south of the Himalayas to Indochina and most of China as far north as the Yellow river (Olivier, 1978). At the turn of the century, there could have been as 200,000 elephants in Asia, but today it is estimated that only about 42,750 elephants survive in the wild (Santiapillai, 1998). Of the 439,000 km^2 of habitat inhabited by *E. maximus*, only 132,000 (30%) km^2 are protected. Since 1900, about a quarter of the elephants in the wild had been captured, domesticated and used in forestry, agriculture, wildlife, ecotourism, and in religious ceremonies. Sri Lanka is estimated to have been 3,000 - 4,000 elephants in the wild, and about 400 in captivity.

1-2. Research history of molecular genetics of Elephantidae

Interspecies phylogeny of Elephantidae

Dene *et al.* (1980) determined the complete amino acid sequence of myoglobin for *E. maximus*, compared it with those of other mammals and discussed on proboscidean phylogeny. Johnson *et al.* (1985) isolated DNA from ancient tissue samples of *Mammuthus primigenius* (wooly mammoth) dated between 10,000-53,000 years old and suggested DNA from mammoth tissue significantly enriched in elephant - related DNA sequence by DNA - DNA hybridization. Drysdale *et al.* (1989) compared the allozyme variations between *L. africana* and *E. maximus* and suggested the two elephant species have been separated for 2.4 mys (millions of yeas).

Höss *et al.* (1994) and Hagelberg *et al.* (1994) first determined the nucleotide sequences of *M. primigenius* for 93 bp of the mitochondrial 16S rRNA gene and 375 bp of the mitochondrial cytochrome b gene, respectively. To know the phylogenetic relationships within Elephantidae and between Tethytheria, Yang *et al.* (1996) analyzed 228 bp of cytochrome b gene for *L. africana*, *E. maximus*, *M. primigenius* and *Mammut americanum* (American mastodon) using museum specimens. They emphasized the importance of appropriate outgroup for resolving phylogenies of highly derived or early radiating linages and suggested a monophyly of *E. maximus* - *M. primigenius* using *M. americanum* as an outgroup. Ozawa *et al.* (1997) determined the 1,005 bp segments of cytochrome b gene for *L. africana*, *E. maximus*, *M. primigenius*, *Hydrodamalis gigas* (Steller's sea cow), *Trichechus manatus* (western Indian manatee) and *Procavia capensis* (hyrax). They used 1st and 2nd codon positions, and amino acid sequences deduced from the nucleotide to construct phylogenetic tree and suggested that *M. primigenius* was more closely related to *E. maximus*

than to *L. africana*. On the other hand, Noro *et al.* (1998) suggested that *M. primigenius* was more closely related to *L. africana* than to *E. maximus* based on the complete sequences of cytochrome b and 12S ribosomal RNA genes. Noro *et al.* (1998) used all codon positions and argued it is reasonable to use all codon positions for analysis since nucleotide differences among the three elephant species are small enough. Recently, Shoshani *et al.* (1998) reviewed both morphological and molecular data for Elephantidae species in detail. They suggested that phylogeny of morphological and molecular data both supported the hypothesis that *Mammuthus* and *Elephas* are more closely related to each other than either is to *Loxodonta*. They also suggested that it is extremely important to employ an outgroup which is closer to the ingroup.

Intraspecies variations for *L. africana* and *E. maximus*

Coetzee *et al.* (1993) examined allozyme variations for 61 individuals of *L. africana* from Kruger National Park in South Africa. They found 7 of 25 loci displayed polymorphic and average hetelozygosity was 0.047, suggesting that the population condition of *L. africana* in the Kruger National Park is relatively health. They recommended a continuous monitoring on population's genetic variation over the period.

Georgiadis *et al.* (1993) analyzed mitochondrial ND 5 - 6 region for 270 savanna elephants from 10 sampling locations using restriction enzymes. Ten haplotypes detected were divided into two representative clades. They indicated protractive gene flow because one relatively derived haplotype was found at all the sampling locations, and exceptionally divergent haplotypes were found to

coexist at distant sampling locations. Such a pronounced mitochondrial divergence must be maintained by gene flow and mobility of elephants with a sufficiently large effective population size persisted by chance within a population based on neutral coalescent theory

Tiedemann *et al.* (1998) determined the nucleotide sequences of cytochrome b gene for 23 individuals of *L. africana* and found eight haplotypes. They suggested that genetic variation is high and very divergent haplotypes coexist within *Loxodonta*, which showed the absence of bottleneck in population history and a relatively large long term effective population size of *Loxodonta*.

Hartl *et al.* (1996) determined the nucleotide sequences of cytochrome b gene for 53 individuals of *E. maximus* and found eight haplotypes. They showed these haplotypes separated out into two major groups. They suggested that no indication of a major separation between *E. maximus* from Sri Lanka and the Asian mainland, previously assumed to represent two different subspecies. They also suggested that *E. maximus* have formed a coherent population after invasion of southern and southeastern Asia and the present local differentiation can be explained by varying human impact on population structure and size. Tiedemann *et al.* (1998) showed genetic variation in *Elephas* is moderate and *Elephas* has a star like phylogeny of haplotypes using the data of their previous study (Hartl *et al.* 1996), which showed population bottleneck in the past or smaller long term effective population size of *Elephas* than that of *Loxodonta*.

For intraspecific analysis, the mitochondrial control region must most effective and there are no work for the elephants using control region until now. In this study nucleotide sequences of the control region were determined to find

the population phylogeny and genetic conditions for regional populations.

2. Materials

The samples analyzed in this study has two categories. Samples from zoo in Japan and imported ivory can be regard as the samples collected randomly from various habitats for *L. africana* and *E. maximus*. These samples were not enough to estimate the total genetic status of each species, but it is still possible to show genetical relationship of each species. The other category is feces and ivory samples from collected in small limited areas: Kruger National Park, Tembe Elephant Park and Addo Elephant National Park in South Africa (Fig. 2-1, B) and Pinnawala Orphanage Park in Sri Lanka. Numbers of the samples are large enough to discuss the genetic conditions in each region.

2-1 Samples for *L. africana* (Table 2)

Feces, saliva or hair samples from zoo in Japan

DNA samples for 50 individuals of *L. africana* were provided by The Japanese Association of Zoological Gardens and Aquarium (Appendix A - 1). Feces, saliva and/or hair roots were collected for each individual. Geographical origins of these elephants were inferred from exported countries, and divided into three categories: 5 individuals from the eastern Africa (exported from Kenya / Tanzania), 16 individuals from the southern Africa (exported from South Africa), and 29 individuals which could not be determined the exported countries.

Ivory samples imported to Japan before CITES agreement

Four ivory samples which were imported to Japan before the CITES agreement were provided by the Japan Ivory Association (Appendix A - 3). These ivory samples were imported form central Africa, Congo, Zaire and Gabon, and so-called "hard ivories" in the Japanese ivory industries.

Ivory and feces samples from Kruger National Park

A total of 30 ivory samples (Appendix A - 5) were collected from the ivory storage room at Kruger National Park in 1995. Detailed informations on curing date and area, sex, age and weight of task are described. Localities of these ivory samples were distributed throughout the park (Fig. 2-4).

A total of 8 feces samples (Appendix A - 6) were collected mainly along about 37 km of the road around Skukuza at the south region of Kruger National Park in 1997 (Fig. 2-4).

Feces samples from Tembe Elephant Park

A total of 24 samples were collected at Tembe Elephant Park in 1997 (Appendix A - 7). Feces were collected along about 6 km of the road at the southern region of the park and at around the Mahlasela Pan (Fig. 2-5). The feces samples were collected considering age and sex presumed by the size of feces and ecological situations (e. g. feeding traces) and fresh / dried conditions of the feces. The feces condition of freshness depends on a temperature, humidity, sunlight and activities of dung beetle. The “fresh feces” in this study were supposed to drop early in the morning or former evening. Both fresh and dried

feces were used for the analysis.

Feces samples from Addo Elephant National Park

A total of 97 samples were collected at Addo Elephant National Park in 1997 (Appendix A - 8). Feces were collected from the western half of the park (Fig. 2-6). Both fresh and dried feces found along about 7.5 km of the road and around the Hapoor water hole were used for the analysis. Feces conditions were also recorded as in Tembe.

Although the individuality for the dropped feces is not completely distinguished and has some extent of duplicate, in Tembe Elephant Park and Addo Elephant National Park the samples analyzed in this study were enough to discuss the frequencies of haplotypes in each population .

2-2 Samples for *E. maximus* (Table 3)

Feces, saliva or hair samples from zoo in Japan

A collections of DNA samples of feces, saliva and/or hair root for 52 individuals of *E. maximus* were provided by The Japanese Association of Zoological Gardens and Aquarium, which 35 individuals were known for the exported country as Thailand, Malaysia, Myanmar, India or Sri Lanka, 18 individuals were unknown for the exported countries (Appendix A - 2).

Feces samples from Pinnawala Orphanage Park in Sri Lanka

A total of 30 feces samples of were collected at the Pinnawala Orphanage Park in Sri Lanka in 1996 (Appendix A - 4) The feces samples were collected

from each individuals which were caught in the wild as orphanage elephant, and their captured places were almost all over the Sri Lanka. Consequently, the samples collected in Pinnawala Orphanage Park must be enough to estimate the genetic diversity in Sri Lanka.

3. Methods

DNA extraction, amplification and sequencing of the mitochondrial control region

DNA extraction was followed as described in part 1. A PCR product of a part of the control region was obtained with the primer L15926. ele and Lox H1. DNA sequences of this fragment were determined by direct sequencing with dye primer cycle sequencing reactions using R. O. B. DNA Processor (Pharmacia) and ALFred automatic sequencer (Pharmacia) with following the manufacture's protocols. To detect haplotypes 351 bp nucleotide portion from the beginning of the control region was sequenced in all samples.

Data analysis

The sequence data were aligned by clustal V (Higgins *et al.* 1992) using Alignment Editor BIORESEARCH / AE (FUJITSU LIMITED 1995). The sequence divergence was estimated by the 2 - parameter method (Kimura 1980) and phylogenetic tree was constructed by the neighbor-joining method (Saitou and Nei 1987) and UPGMA (unweighted pair group method with arithmetic mean) using Molecular Evolutionary Analysis System BIORESEARCH / SINCA

(FUJITSU LIMITED, 1995). Bootstrap analysis of DNA sequence data involved 1,000 replications. As an index for genetic diversity, haplotype diversity (h) and nucleotide diversity (π) were used. Haplotype diversities (h) were calculated using the formula $h=n\ (1-\sum x_i^2) / (n-1)$, where n is the sample size and x_i is the frequency of haplotype i (Nei and Tajima 1981). Nucleotide diversities (π) were calculated using the formula $\pi=n\ \sum x_i x_j \pi_{ij} / (n-1)$, where n is the sample size, x_i and x_j is the frequency of haplotypes i and j, respectively and π_{ij} is the nucleotide divergence between haplotype i and haplotype j (Nei and Jin 1989).

Phylogenetic network represent the evolutionary relationship between haplotypes, indicating where each haplotypes differed in evolutionary process. It was constructed according to the following procedures. (1) the clusters were set up according to the topology of the neighbor-joining tree. (2) connecting closest haplotypes with one or two substitutions in a cluster. (3) connecting haplotypes belonging to the same cluster forming a network for the parallel substitutions.

4. Results

4-1. Haplotype variations

The 351 bp sequences of the control region for 213 samples of *L. africana*, and 86 samples of *E. maximus* showed 54 variable sites, 36 and 23 variable sites within *L. africana* and *E. maximus*, respectively (Table 4). All the substitutions except for the position 178 were transitions, indicating the high transition bias in mtDNA (Brown, 1985). Variable site at the position 178 contained all the four nucleotides; A in all the haplotypes from *L. africana*, and C, T or G in the

haplotypes from *E. maximus*. One indel (insertion or deletion) was detected at position 8.

Among the total of 54 variable sites, the sites which completely differ between *L. africana* and *E. maximus* were only three, position 5 (C for *L. africana* and T for *E. maximus*), position 13 (C for *L. africana* and T for *E. maximus*) and position 178 (A for *L. africana* and C, T or G for *E. maximus*). While, a total of seven sites, positions 9, 100, 200, 204, 216, 264 and 300 were variable in both species. Fourteen sites (positions 66, 67, 78, 185, 190, 201, 215, 259, 275, 323 for *L. africana*, and 154, 168, 220, 246 for *E. maximus*) were identified only in a single haplotype.

4-1-1. Haplotype variations of *L. africana*

L. africana showed 36 variable sites which defined 18 haplotypes (Table 4). There were only one substitution differences between haplotypes L8 and L9, L9 and L10, L10 and L11, L13 and L14, L15 and L17, L16 and L17, indicating that haplotypes L8, L9 L10 and L11, and haplotypes L13, L14 L15, L16 and L17 were belonging to closer clusters. While, haplotype L3 was relatively isolated from other haplotypes, located at 6 substitution differences from haplotype L1. Haplotype L12 was also located at 7 substitution differences from haplotype L14 and L15.

4-1-2. Haplotype variations of *E. maximus*

E. maximus showed 23 variable sites which defined 19 haplotypes (Table 4). There were only one substitution differences between haplotypes E3 and E4,

E3 and E5, E3 and E7, E5 and E8, E7 and E8, E8 and E9, E8 and E10, E9 and E11, E10 and E11, E10 and E12, E11 and E13, E12 and E13, E17 and E18, E18 and E20 and E20 and E21, indicating that haplotypes E3, E4, E5, E7, E8, E9, E10, E11, E12 and E13, and haplotypes E17, E18, E19, E20 and E21 were belonging to closer clusters. While, haplotype E14 was relatively isolated from other haplotypes, located at five substitution differences from haplotype E13.

4-2. Molecular phylogeny

Neighbor-joining (Fig. 2-7) and UPGMA (Fig. 2-8) trees were constructed using the 18 haplotypes for *L. africana* and the 19 haplotypes for *E. maximus*. Haplotypes of *L. africana* distinguished clearly from those of *E. maximus* with high bootstrap values in the neighbor-joining (98%) and UPGMA (95%) trees. The sequence divergence of interspecies was 0.039 - 0.084 (average 0.063) and those within each species were 0.0029 - 0.064 (average 0.033) in *L. africana* and 0.0029 - 0.048 (average 0.024) in *E. maximus*.

4-2-1 Molecular phylogeny for *L. africana*

Both neighbor-joining and UPGMA tree showed a similar topology for *L. africana*. They were divided into two clusters; cluster LI consisting of haplotypes L1 to L7, and cluster LII consisting of haplotypes L8 to L18. The sequence divergence between the clusters LI and LII was 0.024 - 0.064 (average 0.042) and those within each clusters were 0.006 - 0.042 (average 0.022) in cluster LI, and 0.003 - 0.042 (average 0.021) in cluster LII. The bootstrap values of the cluster LI are 41% and 61% in the neighbor-joining and the UPGMA tree,

respectively. The bootstrap values of the cluster LII are 64% and 56% in the neighbor - joining and the UPGMA tree, respectively. Therefore, the clustering of LI and LII were not significant. But all the haplotypes in the cluster LII had a deletion at position 8.

In the cluster LI consisted of haplotypes L1 to L7. The bootstrap value of the clade composed of haplotypes L4 to L7 was 83% in the neighbor-joining tree. Haplotype L1 was the nearest to the hypothetical ancestral node of *L. africana* and L7 was the most distant to the hypothetical ancestral node of *L. africana.*

In the cluster LII consisted of haplotypes L8 to L18, haplotype L18 was first separated from other haplotypes in the neighbor-joining tree. The clade composed of haplotypes L8 to L11 had a relatively high bootstrap values (89%) in the neighbor-joining tree. So cluster LII could be divided into subcluster LIIa consisting of haplotypes L8 to L11 and subcluster LIIb consisting of haplotypes L12 to L14.

4-2-2 Molecular phylogeny for *E. maximus*

Both neighbor-joining and UPGMA tree showed a similar topology for *E. maximus*. They were divided into two clusters; cluster EI consisting of haplotypes E1 to E14, and cluster EII consisting of haplotypes E16 to E21. The sequence divergence between the clusters EI and EII was 0.023 - 0.048 (average 0.037) and those within each clusters were 0.003 - 0.039 (average 0.014) in cluster EI and 0.003 - 0.012 (average 0.006) in cluster EII. The bootstrap values of the cluster EI are 65% and 79% in the neighbor-joining and the UPGMA tree,

respectively. The bootstrap values of the cluster LII are 97% and 99% in the neighbor-joining and the UPGMA tree, respectively, suggesting that the clustering of EII was significant. All the haplotypes in the cluster EII had a transversion at position 178 which was the only transversion among *L. africana* and *E. maximus* analyzed.

In the cluster EI which consisted of haplotypes E1-E14, the clade composed of E1 and E2 was separated from other haplotypes with high bootstrap values of 96 replications in the neighbor-joining tree. So cluster EI could be divided into subcluster EIa consisting of haplotypes E1 and E2, and subcluster EIb consisting of haplotype E3 to E14

In the cluster EII consisting of haplotypes E16-E21, E16 was an outergroup of other haplotypes in the neighbor-joining tree.

4-3. Network analysis

Network for *L. africana* was composed of cluster LI, LIIa, LIIb and haplotype L18 (Fig. 2-9). In the cluster LI, haplotypes L1 and L2, L4 and L5, and L6 and L7 differed by two substitutions from each other. Variable sites 42, 276 and 157 had parallel substitutions forming a double cubic network. Haplotype L3 was isolated with a distance of six substitutions from haplotype L1.

The cluster LII was divided into subcluster LIIa consisting of haplotypes L8 to L11, subcluster LIIb consisting of L13 to L17, and haplotype L18. All the haplotypes within subcluster LIIa were connected by one substitution. In subcluster LIIb haplotypes L13 and 14, and haplotypes 15, 16 and 17 were also connected by one substitution.

Network for *E. maximus* was composed of cluster EIa, EIb and EII (Fig. 2-10). The cluster EI was divided into subcluster EIa consisting of haplotypes E1 and E2, subcluster EIb consisting of E3 to L14. All the haplotypes except for E3 and E14 were connected by one substitution distance. Cubic network was formed by parallel substitution of position 156 in the first dimension, positions 167 and 147 in the second dimension, positions of 204, 327 and 269 in the third dimension. In the cluster EII all the haplotypes were connected by one substitution distance forming a network by positions 100 and 269.

4-4. Geographical distribution of haplotypes

4-4-1 Geographical distribution of haplotypes of *L. africana* (Table. 5)

Feces, saliva or hair samples from zoo in Japan

A total of 50 zoo samples were divided into three geographical categories, 5 individuals from the eastern Africa, 16 individuals from South Africa and 29 individuals of countries unknown. Haplotype frequencies of the 5 individuals from the eastern Africa were composed of one individual of haplotype L6, one individual of L13, two individuals of L14 and one individual of L17. The 16 individuals from the South Africa made up of 7 individuals of L10, three individuals of L15 accompanied with L4, L9, L8 and L14. A total of 29 individuals of countries unknown samples showed a similar frequency with those from the South Africa, consisting of 7 individuals of L10, 6 individuals of L14, accompanied with L3, L4, L5, L8, L9, L12, L13, L15 and L16.

Consequently, the haplotype L10 was the most common in South Africa (44%), and also in unknown samples (24%). The L14 was the second most

common and found in all three geographical categories.

Ivory samples imported to Japan

Haplotypes for the four ivory samples which were imported from central Africa were haplotype L1 for one individual from Congo, L2 for one individual from Gabon and L7 for two individuals from Zaire. All these haplotypes were belonged to the cluster LI.

Ivory samples from the Kruger National Park

Haplotype frequencies of the 30 ivory samples from Kruger National Park showed that haplotype L14 was the most common reaching at 50% of the individuals. The L10 was the second most common at 40% of the individuals. These common haplotypes were distributed throughout the Park. Another two haplotypes, L18 in one individual from Marelane in the south region of the park and L9 in two individuals bred in Skukuza, were also found from the ivory samples. All these four haplotypes belonged to the cluster LII.

The number of samples in this study were not enough considering the area and population size in Kruger. If more samples cover a wide range of the Park will be analyzed, there is a possibility to detect regional differences of haplotype distribution.

Feces samples from the Kruger National Park

Haplotype frequencies of the 8 feces samples from the Kruger National Park consisted of 7 samples with haplotype L10 and one sample with L14. Both

L10 and L14 were also common in the ivory samples from the Kruger National Park. All these haplotypes were placed in cluster LII.

All the feces were collected for the only fresh ones. The seven feces samples identified as haplotype L10 were collected within 15 km on the road from Skukuza to Malelane gate.

Feces samples from the Tembe Elephant Park

Haplotype frequencies of the 24 feces samples from the Tembe Elephant Park were composed of 12 samples with haplotype L14 (50%), 7 samples with L16 (29%) and 5 samples with L10 (21%). All the haplotypes from Tembe were located at the cluster LII. Haplotypes L14 and L10 were also common in the Kruger National Park, however L16 was not detected from Kruger.

The haplotypes when plotted distinguishing fresh and dried feces were distributed evenly within about 6 km of the road. There are no difference for haplotype distributions between the feces on the road and around Mahlasela Pan (Fig. 2-5). These feces were all composed of solitarily males, perhaps resulting mixture of haplotypes L10, L14 and L16.

Feces samples from the Addo Elephant National Park

Haplotype frequencies of the feces samples in about 7.5 km of the road in western half of the Addo Elephant National Park showed that 97 samples were made up of 77 samples with haplotype L11 (79%) and 20 samples with L14 (21%). The L14 was also common in the Kruger National Park and the Tembe Elephant Park. However L11 was unique haplotype at the Addo Elephant National

Park, not detected from other two parks. The haplotypes L11 and L14 were belonged to the cluster LII.

The clear difference was indicated in haplotype frequencies between fresh and dried feces. Both haplotypes L11 and L14 from dried feces distributed evenly on the road and around Hapoor waterhole, however all the haplotypes from fresh feces were composed of L11 with only one exception (L14 in 96LX62). These feces were inferred to be produced by family groups of elephants, suggesting that these feces came from the last herds which moved across the sampling area.

Number of the samples analyzed in Tembe and Addo is sufficient considering the area and population size of the parks, it is expected that there is a little possibility to find new haplotypes and big change in the haplotype frequencies.

4-4-2 Geographical distribution of haplotypes of *E. maximus* (Table. 6)

Feces, saliva or hair samples from zoo in Japan

Haplotype frequencies of *E. maximus* from the zoo samples showed that the 21 individuals from Thailand consisted of seven individuals with haplotype E13 and four individuals with E20, with accompanied by haplotypes E1, E2, E4, E7, E8, E11, E14, E19 and E20. One individual from Malaysia was identified as haplotype E1. Three individuals from Myanmar were identified as haplotypes E7, E19 and E21. Seven individuals from India were identified as haplotypes E4, E13, E17 and E20. Consequently, haplotypes E4, E7, E13 and E20 were found both in Southeast Asia and South Asia.

Geographical distribution of the haplotypes suggested that subcluster EIa

(haplotypes E1 and E2) were composed of individuals only from Southeast Asia, while cluster EIb (haplotypes E3 - E14) and cluster EII (haplotypes E16 - E21) were mixed with those from Southeast Asia and South Asia.

Feces samples from Pinnawala Orphanage Park in Sri Lanka

Haplotype frequency of *E. maximus* from Pinnawala Orphanage Park in Sri Lanka showed that 34 individuals were made up of 13 individuals with haplotype E4 (38%), 10 individuals with haplotype E10 (29%), and five individuals with haplotype E13 (15%) with accompanied by haplotypes E3, E7, E9 and E12. All these haplotypes were placed in cluster EI. Haplotype E4 was the most common in Sri Lanka and also found in Thailand and India. While, the haplotype E10 which was the second common in Sri Lanka was only found in Sri Lanka.

4-5 Genetic diversities

To characterize a genetic diversity of populations, both haplotype diversity (h) and nucleotide diversity (π) are essential (Hartl *et al.* 1996). The haplotype diversities and nucleotide diversities within each population (Table 7) were calculated using the samples from zoo in Japan and three national parks in South Africa for *L. africana*, and the samples from zoo in Japan and Pinnawala Orphanage Park in Sri Lanka for *E. maximus* .

4-5-1 Haplotype diversity (h) of *L. africana* and *E. maximus*

The haplotype diversity (h) of samples from zoo in Japan was 0.87 in *L africana* and 0.91 in *E. maximus*, both of which were very high. At Pinnawala

Orphanage Park in Sri Lanka, h = 0.76 for *E. maximus*, which was relatively high because many haplotypes (n=7) were found from the Park. At Kruger National park and Tembe Elephant Park in South Africa, h = 0.58 and 0.65 for *L. africana*, respectively, which were reasonably high. The number of haplotypes in Tembe (n=3) was lesser than in Kruger (n=4) but the haplotype diversity was greater in Tembe than in Kruger. This is because the frequencies of three haplotypes appeared from Tembe were relatively uniformity. At Addo Elephant National park in South Africa, h =0.33 for *L. africana*, which was very low (h=0.33) because in Addo only two haplotypes were appeared and haplotype frequencies were heavily biased toward L11.

4-5-2 Nucleotide diversity (π) of *L. africana* and *E. maximus*

The nucleotide diversity (π) of samples from zoo in Japan was 2.14 in *L. africana* and 2.17 in *E. maximus*, both of which were very high. At Kruger National park, Tembe Elephant Park and Addo Elephant National Park in South Africa, π = 1.30, 1.06 and 0.88 for *L. africana*, respectively. Because each parks had diverged haplotypes, the nucleotide diversities (π) of each park were considerably high though the number of haplotypes were small. Even in Addo which has only two haplotypes, the nucleotide diversities (π) was larger than Pinnawala Orphanage Park in Sri Lanka which had seven haplotypes. Although many haplotypes were found from Pinnawala Orphanage Park in Sri Lanka, π = 0.65, which was very low. This was because the sequence divergences of haplotypes from Pinnawala were small and all the haplotypes from Pinnawala were belonged to cluster EIb.

5. Discussion

5-1 Establishment of DNA analysis for the elephants

The feces samples have a drawback that complete individuality was not known, but it is extremely useful to study the endangered species that is difficult to encounter and acquire the adequate samples for DNA analysis. In European brown bear the mitochondrial control region and nuclear SRY gene were amplified using PCR from feces and a single hair (Kohn *et al.* 1995). Also in *L. africana* and *E. maximus* it is possible to extract and amplify the DNA using feces, saliva, hairs and pieces of ivory. Especially in feces which is relatively fresh, more than 1500 bp containing whole control region could be amplified even if the feces collected in the field.

5-2. Nucleotide variations

Sequence divergence between control region and cytochrome *b* gene

In general, it is said that the evolutionary rate of the control region is three to five times faster than the rest of the mitochondrial genome (Aquadro and Greenberg 1983). For elephant species Hartl *et al.* (1996) determined the 335 bp nucleotide sequences of cytochrome b gene for 53 individuals of *E. maximus* from Thailand, Malaysia, Burma, India and Sri Lanka. Since their studies covers various places in Asia, it is possible to estimate the difference of evolutionary rate between the control region and cytochrome b gene. They found 7 variable sites in 335 bp of cytochrome b gene and determined eight

haplotypes. The sequence divergences of these eight haplotypes were 0.003 - 0.012 with average of 0.009. The 351 bp of control region in this study exhibited 23 variable sites defining 19 haplotypes in *E. maximus*. The sequence divergences of these 19 haplotypes were 0.029 - 0.048 with average of 0.024. Consequently, the control region had about 3.3 times of variable sites and 2.8 times of average sequence divergence, presuming that evolutionary rate of the control region is about three times faster than cytochrome b gene in *E. maximus*. There are no data of cytochrome b gene for *L. africana* from various places, the divergence of the control region in this study was 0.029 - 0.064 (average 0.033) .

Sequence divergence between *L. africana* and *E. maximus*

In this study, the sequence divergences of the control region were 0.039 - 0.084 (average 0.063) between species, while 0.029 - 0.064 (average 0.033) within *L. africana* and 0.029 - 0.048 (average 0.024) within *E. maximus*. The divergences of interspecies were very much smaller than those within species unexpectedly. To interpret this phenomenon it is necessary to consider the evolutionary rate of the control region. In this study total of 37 haplotypes and 54 variable sites were found from 351 bp of the control region for *L. africana* and *E. maximus*. But the sites which completely differ between *L. africana* and *E. maximus* were only three.

Tiedemann *et al*. (1998) examined 369 bp nucleotide sequences of cytochrome b gene for 23 individuals of *L. africana* from three regions, 18 individuals from northwest Namibia, 4 individuals from northeast Namibia and one individual from Kruger National Park, in southern part of Africa and compared

with Harlt *et al*. (1996). They found six and eight haplotypes from *L. africana* and *E. maximus*, respectively. The average sequence divergence of cytochrome b gene of interspecies was 0.082 and within each species were 0.030 in *L. africana* and 0.009 in *E. maximus*.

Since the sampling points and range of Tiedemann *et al.* (1998) were different from those of my present study, the following comparison are outlines. The divergences of intraspecies of cytochrome b gene for each species were smaller than those of the control region, which was reasonable because control region has a higher evolutionary rate than cytochrome b gene. On the other hand in interspecies the divergence of cytochrome b gene was larger than that of the control region. This indicates that there is a possibility that the nucleotide substitutions in 351 bp of the control region has been close to saturated between *L. africana* and *E. maximus*. Accordingly, the divergence between *L. africana* and *E. maximus* based on the 351 bp of the control region may be underestimated. The 351 bp of the control region is expected that it does not suit to analysis of phylogeny and divergence for interspecies. In the following sections, phylogenies and divergences for populations within each species were discussed.

It is a point at issue in molecular phylogenetics in Elephantidae that species of the extant elephants is more closely related with *Mammuthus*, *Loxodonta* or *Elephas*. Yang *et al.* (1998), Ozawa *et al.* (1997) and Shoshani *et al.* (1998) suggested that *Elephas* is more closely related to *Mammuthus* than *Loxodonta*. On the other hand, Noro *et al.* (1998) suggested that *Loxodonta* and *Mammuthus* are more closely related to each other than to *Elephas*. The differences of analyzed region and data analysis in each study maybe supposed to cause this

inconsistency. As shown in this study, the genetic variation within both *L. africana* and *E. maximus* are large enough to be considered. Georgiadis *et al.* (1994) found two exceptionally divergent clades within savanna elephant. Tiedemann *et al.* (1998) also confirmed high levels of genetic variation to be present in *L. africana.* To discuss the phylogeny of these three species hereafter, it might be necessary to consider the genetic variations within species.

Psuedogenes of mtDNA in nuclear DNA

One point to notice in determining the nucleotide sequences of mtDNA using PCR is that large fragments of mitochondrial genome sometimes transposed to nuclear genome during the course of evolution (Lopez *et al.* 1994). In fact, Greenwood and Pääbo (1998) revealed the presence of psuedogenes of the mitochondrial control region in nuclear genome in *E. maximus* by sequencing the true control region (AF071098) and four putative nuclear insertion sequences of the control region (AF071099 - AF071102). The true control region is most close to haplotype E10 with one substitution in position 216. The sequences of the four putative nuclear insertion were extremely different from those of the haplotypes presented in this study. Therefore, the haplotypes detected in this study might be originated from true mitochondrial control region, not from nuclear DNA.

5-3. Subspecies and population phylogeny

5-3-1. Population phylogeny for *L. africana*

General classification for subspecies

Generally, the subspecies of *L. africana* has been classified into two major divisions, the savanna elephant division and the forest elephant division (Laursen and Bekoff 1978). The savanna elephant differs from the forest elephant in larger body size, sparser hair covering, triangular-shaped ears rather than smaller round ears, gray skin rather than brown skin. The tusks of the savanna elephant are horizontal, thick, curved which are opposed to the straight, narrow, downward-pointing of the forest elephant (Morrison - Scott 1948). The savanna elephant include four subspecies, *L. a. africana*, *L. a. knochenhaueri*, *L. a. orleansi* and *L. a. oxyotis*. The forest elephant include two subspecies, *L. a. cyclotis* and *L. a. pharaohensis* (Laursen and Bekoff 1978). *Loxodonta. africana. africana* is distributed from central to southern part of Africa. *L. a. knochenhaueri* is distributed in eastern part of Africa, mainly Kenya and Tanzania. *L. a. cyclotis* is distributed in tropical forest from central to western part of Africa. *L. a. pharaohensis* and *L. a. orleansi* were extinct subspecies. Another subspecies were also recognized such as "pygmy" elephant (Morrison - Scott 1948) or "water" elephant (Trouessart 1991). But these were not proven.

The relation between subdivision based on control region and the general classification

In this study the haplotypes from *L. africana* were divided into two clusters, LI and LII. The cluster LII was further divided into subcluster LIIa, subcluster LIIb, and haplotype L18. The cluster LI include the samples from eastern Africa, central Africa and South Africa. All the haplotypes from 4 ivory samples came from central Africa belonged to cluster LI. These 4 samples were assumed to be

forest elephant since these were so called "hard ivory" and the countries of origin were the habitats of forest elephant. Haplotypes L1 and L2 detected in ivory samples from eastern Africa were very close to the ancestral node of *L. africana*.

The subcluster LIIa was composed of only haplotypes from South Africa at the moment. Though the representation of the samples analyzed was biased to South Africa, there is a possibility that haplotypes in subcluster LIIa are restricted in southern part of Africa. The haplotype L10 which was located at the center in this subcluster in the network was the most common haplotype in Kruger National Park, and also appeared in Tembe Elephant Park and zoo samples. Accordingly, it is suggested that the haplotype L10 is an ancestral haplotype in southern part of Africa. The subcluster LIIb was composed of haplotypes from South Africa and Kenya / Tanzania. The haplotype L14 which was located at the center in this subcluster in the network was found not only in three national parks in South Africa but also in Kenya / Tanzania, suggesting that L14 was common ancestral haplotype in evolutionary history in eastern and southern Africa.

There is a general agreement that *L. africana* is divided into two divisions, savanna elephant and forest elephant. There is a large sequence divergence between cluster LI and LII. Haplotypes from eastern Africa, or habitat of forest elephant were located in cluster LI. Cluster LI also include the zoo samples from Kenya / Tanzania and South Africa. Consequently, cluster LI is a mixture of both the forest and savanna elephant and cluster LII is a savanna elephant cluster if the exported countries of the zoo samples showed the true habitat of the animals. *L. africana* has great mobility and high adaptability to different

environments, and there are no geographical barriers within the African continent which interfere with the movement of the populations. Consequently, to confirm the relationship between the two clusters based on control region in this study (LI and LII) and two general divisions (savanna and forest elephant), further analysis will be needed.

Georgiadis *et al.* (1994) examined 270 individuals of savanna elephants using restriction enzyme method for 2,450 bp of ND5 - 6 regions in mtDNA. They divided the savanna elephants into two clusters. From my present data it was also evident that savanna elephants can be divided into two groups from mitochondrial DNA (LI and LII or LIIa and LIIb).

Subclusters in South Africa

In this study, the haplotypes belonging to both subclusters (LIIa and LIIb) were found in South Africa. Tiedemann *et al.* (1998) also found very divergent haplotypes of cytochrome b gene coexisted in Namibia. The coexistence of divergent mtDNA in small area seems distinctive features of *L. africana*. The sequence divergence between the haplotypes of two subclusters in this study is 0.021 - 0.042 (average 0.029). In mammals the coexistence of individuals bearing such divergent haplotypes was not so common (Cronin 1990; Wayne *et al.* 1990).

The coexistence of distant haplotypes in South Africa might be the reason of ecological characters of *L. africana*. Because *L. africana* has typical k-strategy in reproduction, it is more difficult to create bush-like tree radiation as r-strategy type of animals which have many haplotypes with a few substitution. While

k-strategy animals has a discontinuous lineage of haplotypes with several substitution differences. Besides, *L. africana* has great mobility, resulting coexistence of haplotypes with distance in South Africa.

5-3-2. Population phylogeny for *E. maximus*

Genetical originality of Sri Lankan population

Though there are several ways for classification of subspecies in *E. maximus*, Sri Lankan population is always regarded as different subspecies form those of Asian continent. In this study about 43 % of the samples of *E. maximus* were from Sri Lanka and about 37 % of the samples were from Asian continent. Considering the distribution of *E. maximus*, it is expected that no considerable change would be occurred in the current level of the mitochondrial DNA variation, because the samples were collected from wide range of habitats of *E. maximus*.

Five haplotypes from Sri Lanka, E3, E5, E9, E10 and E12, could be assumed to be the original haplotypes in Sri Lanka. However, haplotypes E4, E7 and E13, which were main haplotypes in Sri Lanka, were also found in India and Southeast Asia. Therefore, the haplotypes from Sri Lanka and Indian subcontinent overlapped with each other, suggesting that *E. m. maximus* is not clearly separated as a subspecies from *E. m. indicus* according to mitochondrial control region analysis. Hartl *et al.* (1996) analyzed 335 bp of cytochrome b gene for 53 individuals of *E. maximus* from India, Sri Lanka, Myanmar, Thailand and Vietnam. They found *E. maximus* was separated into two clusters and there was no indication of a major separation between elephant from Sri Lanka and the Asian mainland. The present study agree with their results. This phenomenon

reflect the historical event that the Sri Lankan Island separated from the Indian subcontinent during the Holocene sea- separation and the human impact on the elephant genetic structure; tamed elephants were transported between the Sri Lankan Island and the Indian subcontinent. Tiedemann *et al.* (1998) suggested that the haplotypes of cytochrome b gene for *E. maximus* formed star like phylogeny. Based on the control region in this study similar star like phylogeny was found in *E. maximus* (Fig. 2-9). The star like phylogeny seems to be a distinctive feature of *E. maximus*.

Haplotype distributions in Sri Lanka

The geographical distribution of the haplotypes in Sri Lanka shows that haplotypes E4 and E10 are the most common and are distributed throughout the country, suggesting that the local population of *E. maximus* in Sri Lanka is now genetically homogenous. Mobility of elephants must be large enough to be mixed together throughout the Sri Lankan Island after separation from the continent during the Holocene sea-transgression. Another explanation is human impact on the elephant genetic structure; tamed elephants in Sri Lanka were transported by human, and might be mixed with wild population though time.

5-4. Genetic diversity of local populations and conservation for the elephants

Haplotype diversity (h) and nucleotide diversity (π) are used as indices for genetic diversity of populations. It is expected that haplotype diversity of the population decrease after the population bottleneck because the number of the haplotypes get fewer and haplotype frequencies is biased. The large nucleotide

diversity means that the common ancestor of the population is more retrospective. So, high nucleotide diversity could be indicated high haplotype diversity. Both haplotype diversity and nucleotide diversity are high and distribution of haplotypes is uniformity means genetically healthy condition.

Population history in South Africa

In South Africa before 1650 *L. africana* were in all over the country except the interior of the land. After that the habitat and population size decreased sharply by the human impact (Hall-Martin, 1992). At the beginning of this century the wild population had remained in only four regions, Kruger National Park, Tembe Elephant Park, Addo Elephant National Park and Knysna Forest. A minimum population number was 11 at Addo Elephant National Park in 1931 (helicopter census). Though there are no accurate data in Kruger National Park and Tembe Elephant Park, it is said that the population is nearly zero in 1903 at Kruger National Park, and 16 individuals in 1971 at Tembe Elephant Park. It is known that there was influx of animals from Mozambique to Kruger National Park in 1960's, but in other three habitats there were no immigration of the animals from other regions. The population of Kruger, Tembe and Addo had increased their size and in 1990 the population size was estimated to be 7278 in Kruger, 80 in Tembe and 162 in Addo.

Bottleneck event and its effect to genetic diversity

The Asian elephant population decreased rapidly over the last 200 years in Sri Lanka. The nucleotide diversity (π) was considerably low (0.65), because

the sequence divergences of haplotypes from Sri Lanka was small. All the haplotypes from Sri Lanka were belonged to the cluster EIb and settled closely each other in phylogenetic tree and network. But the haplotype diversity (h) was relatively high (0.76), which indicates that the population decrease in Sri Lanka is not fatal damage from genetical perspective. Although Sri Lankan population still maintained a reasonable genetic diversity, it may merely reflect the large population size in ancient times. Fortunately, the genetic diversity of the elephant in Sri Lanka may not be severely hit by the population decrease from the point of view of the mitochondrial control region. It is important to prevent the population decrease hereafter.

On the other hand, haplotype diversity in each population in South Africa were low (0.33 - 0.65), because the number of haplotypes was small and their frequencies were biased. In contrast, nucleotide diversities of each population were relatively high (0.88 - 1.30), because the sequence divergences of haplotypes with each park were relatively large. In Kruger National Park, the number of haplotypes were only four but the haplotypes were very diverged and scattered in phylogenetic tree and network. Therefore haplotype diversity was low but nucleotide diversity was two times greater than Pinnawala. This may be due to haplotype extinction through the shrink of population size. This is the same in Tembe and Addo, in which haplotype diversity was low, while nucleotide diversity was high compared to the case of Pinnawala. In particular, the extent of π is low in Addo which possessed only two haplotypes (Table 2-6).

Elephants from all three parks in South Africa had unique haplotypes which were not found for other parks: L9 and L18 in Kruger National Park, L16

in Tembe Elephant Park and L11 in Addo Elephant National Park. It is noteworthy that there is substantial differences in composition and frequency of the haplotypes between the populations from such small area in South Africa. Especially, the distance between Kruger and Tembe is about only 100 km. Before 1650, wide area of South Africa except interior dry region was a habitat of *L. africana* (Fig. 2-2). So formerly the populations in Kruger, Tembe and Addo were likely to be connected genetically. Considering the mobility of the elephants and the current state of habitat in South Africa, it is unusual that the three parks exhibit such differences in the genetic constitutions of haplotypes. Therefore, it is unlikely that such haplotype variation is due to populational differentiation from the past. It would simply reflect the random linage soaring after the recent bottleneck events.

The elephants in three national parks in South Africa has two genetic properties: 1) the haplotype diversity (h) in each parks was low and nucleotide diversities (π) was high compared to those for all the *L. africana* combined; 2) they have their own haplotypes. These may imply that a large population in the past was divided into small populations and experienced a population bottleneck effect.

Implications for conservation in South Africa

The populations of *L. africana* in South Africa has increased satisfactorily over several decades due to successful conservation plans. In Kruger, Tembe and Addo the population sizes of *L. africana* are likely to reach the upper limit. But as I mentioned above, the populations in South Africa are presumed to have

experienced a population bottleneck in the past, resulting low level of genetic diversity. For further conservation plans, therefor, programs should be followed that increase the genetic diversity in each of the elephant populations. For example, it seems a valid way to transport the individuals which have unique haplotypes to another park. It will not cause any genetical disturbance because Kruger, Tembe and Addo populations exchanged genetic elements freely until recent time, 100 or 200 years ago.

However, the diversity of the mtDNA and nuclear DNA are different matter. The diversity of nuclear DNA is also important. To raise the amount of genetic diversity in the whole genome, it is the best way to make a conservation plan after examine the nuclear DNA.

Genetic sourcing using control region analysis

At the present time, the demand for resume to trade ivories has been increased in southern African countries. On occasion for resume to trade ivories, the identifications of the place of origin is important subject to prevent illegal trade. In this study, there were haplotypes which found in particular region with high frequency (e. g. L11 in Addo and L16 in Tembe), while widely distributed haplotypes (e. g. L14) were also exist. There seems to be two types of haplotypes in *L. africana*, local type and common type. This tendency agree with the result of Georgiadis *et al.* (1996). They examined savanna elephants from ten locations in Africa and divided into two categories, unique haplotypes restricted to Kenya, Tarangire in Tanzania and Sengwa in Zimbabwe, common the haplotype appeared from all the ten locations. They found two haplotypes in Kruger National Park,

both of them appeared outside the Kruger National Park. Among four haplotypes found in this study from Kruger National Park, major two haplotypes (L10 and L14) were also found in another regions with high frequencies.

The samples analyzed in this study were mainly from South Africa. If many samples from other regions are analyzed in the future, many more local haplotypes will probably appear. For such local haplotypes the genetic sourcing of the elephants will be possible using the mitochondrial control region analyzed after some improvements as to the analyzed region.

Conclusions

1. The method of DNA analysis for the elephants using feces and ivory samples were established.

2. The nucleotide sequences of the entire control region for *L. africana* and *E. maximus* were determined and the distribution of variable sites in the control region for the elephants was investigated.

3. The nucleotide sequences of 351 bp of the control region were determined for 213 samples of *L. africana* and 86 samples of *E. maximus*. *Loxodonta africana* was divided into two or three groups and *E. maximus* was divided into two groups based on the analysis for the control region.

4. The genetic diversities for wild elephant populations, three parks in South Africa and one region in Sri Lanka, were examined and a unfavorable genetic state caused by recent population bottleneck in South Africa was suggested.

5. It is expected that DNA analysis for the elephant will progress using the method established in this study, which will be of assistance to make effective conservation plan for the elephant.

References

Aquadro CF, Greenberg BD (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences for seven individuals. Genetics 103:287-312

Árnason U, Johnsson E (1992) The complete mitochondrial DNA sequence of the harbor seal, Phoca vitulina. Journal of Molecular Evolution 34:493-505

Avise JC (1994) Molecular markers, Natural history and evolution. Chapman and Hall, pp 511. New York

Avise JC, Hamrick JL (1996) Conservation genetics case histories from nature. Chapman and Hall, pp 512. New York

Brown WM (1985) The mitochondrial genome of animals. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum Press, New York, Pp. 95-130

Brown GG, Gadaleta G, Pete G, Saccone C, Sbisa E (1986) Structural conservation and variation in the D-loop-containing region of Vertebrate mitochondrial DNA. Journal of Molecular Biology 192:503-511

Casane D, Dennebouy N, Rochambeau H, Mounolou JC, Monnerot M (1997) Nonneutral evolution of tandem repeats in the mitochondrial DNA control region of lagomorphs. Molecular Biology and Evolution 14:779-789

Clayton DA (1991) Replication and transcription of vertebrate mitochondrial DNA. Annual Review of Cell Biology 7:453-478

Coetzee EN, Bank FH, Critser JK (1993) Allozyme variation in a wild African elephant (*Loxodonta africana*) population from Kruger National Park, South Africa. Comparative Biochemistry and Physiology 106B:109-114

Cronin MA (1990) Intraspecific variation in mitochondrial DNA of north American cervids. Journal of Mammalogy 73:70-82

Dene H, Goodman M, Romero-Herrera AE (1980) The amino acid sequence of elephant (*Elephas maximus*) myoglobin and the phylogeny of Proboscidea. Proceedings of Royal Society of London [Biology] 207:111-127

Dillon MC, Wright JM (1993) Nucleotide sequence of the D-loop region of the sperm whale (*Physeter macrocephalus*) mitochondrial genome. Molecular Biology and Evolution 10:296-305

Doda JN, Wright CT, Clayton DA (1981) Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequence. Proceedings of National Academy of Science of The United States of America 78:6116-6120

Douzery E, Randi E (1997) The mitochondrial control region of cervidae: Evolutionary patterns and phylogenetic content. Molecular Biology and Evolution 14:1154-1166

Drysdale TA, Florkiewicz RF (1989) Electrophoretic variation within and between the two extant elephant species (Mammalia: Proboscidea). Journal of Mammalogy 70:381-383

Dufresne C, Mignotte F, Guéride M (1996) The presence of tandem repeats and the initiation of replication in rabbit mitochondrial DNA. European Journal of Biochemistry 235:593-600

Foran DR, Hixson JE, Brown WM (1988) Comparisons of ape and human sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. Nucleic Acids Research 16:5841-5861

Fumagalli L, Taberlet P, Favre L, Hausser J (1996) Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region

of shrews. Molecular Biology and Evolution 13:31-46

Garcia-Rodriguez AI, Bowen BW, Domning D, Mignucci-Giannoni AA, Marmontel M, Montoya-Ospina RA, Morales-Vela B, Rudin M, Bonde RK, McGuire PM (1998) Phylogeography of the West Indian manatee (*Trichechus manatus*): How many populations and how many taxa ? Molecular Ecology 7:In press

Gemmell NJ, Western PS, Watson JM, Graves JAM (1996) Evolution of the mammalian mitochondrial control region - Comparisons of control region sequence between monotreme and therian mammals. Journal of Molecular Evolution 13:798-808

Georgiadis N, Bischof L, Templeton A, Patton J, Karesh W, Western D (1994) Structure and history of African elephant populations: I. Eastern and Southern Africa. Journal of Heredity 85:100-104

Ghivizzani SC, Mackay SLD, Madsen CS, Laipis PJ, Hauswirth WW (1993) Transcribed heteroplasmic repeated sequence in the porcine mitochondrial DNA D-loop region. Journal of Molecular Evolution 37:36-47

Greenwood AD, Pääbo S (1998) Nuclear insertion sequences of mitochondrial DNA predominate in hair but not in blood of elephants. Molecular Ecology in press:

Hagelberg E, Thomas MG, Cook EC Jr, Sher AV, Baryshnikov GF, Lister AM (1994) DNA from ancient mammoth bones. Nature 370:333-334

Hall-Martin AJ (1980) Elephant survivors. Oryx 15:355-362

Hall-Martin AJ (1992) Distribution and status of the African elephant *Loxodonta Africana* in South Africa, 1965-1992. KOEDOE 35:65-88

Hartl GB, Kurt F, Tiedemann R, Gmeiner C, Nadlinger K (1996) Population

genetics and systematics of Asian elephant (*Elephas maximus*): A study based on sequence variation at the Cyt b gene of PCR-amplified mitochondrial DNA from hair bulbs. Zeitschrift Für Säugetierkunde. 61:285-294

Hayasaka K, Ishida T, Horai S (1991) Heteroplasmy and polymorphism in the major noncoding region of mitochondrial DNA in Japanese monkeys: Association with tandemly repeated sequence. Molecular Biology and Evolution 8:399-415

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: improved software for multiple sequence alignment. Computer Applications in the Biosciences 8:189-191

Hoelzel AR, Hancock JM, Dover GA (1991) Evolution of the cetacean mitochondrial D-loop region. Molecular Biology and Evolution 8:475-493

Hoelzel AR, Hancock JM, Dover GA (1993) Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. Journal of Molecular Evolution 37:190-197

Hoelzel AR, Lopez JV, Dover GA, O'Brien SJ (1994) Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. Journal of Molecular Evolution 39:191-199

Horai S, Kondo R, Murayama K, Hayashi S, Koike H, Nakai N (1991) Phylogenetic affiliation of ancient and contemporary human inferred from mitochondrial DNA. Philosophical transactions of the Royal Society of London. B 333:409-417

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, Park KS, Omoto K, Pan I-H (1996) mtDNA polymorphism in east Asian populations, with special reference to the peopling of Japan. American Journal of Human Genetics 59:579-590

Höss M, Kohn M, Pääbo S, Knauer F, Schroder W (1992) Excrement analysis by PCR. Nature 359:199

Höss M, Pääbo S, Vereshchagin NK (1994) Mammoth DNA sequence. Nature 370:333

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the Cytochrome b gene of mammals. Journal of Molecular Evolution 32:128-144

Johnson PH, Olson CB, Goodman M (1985) Isolation and characterization of deoxyribonucleic acid from tissue of the wooly mammoth, *Mammuthus primigenius*. Comparative Biochemistry and Physiology 81B:1045-1051

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequence. Journal of Molecular Evolution 16:111-120

Kohn M, Knauer F, Stoffella A, Schroder W, Pääbo S (1995) Conservation genetics of the European brown bear - a study using excremental PCR of nuclear and mitochondrial sequences. Molecular Ecology 4:95-103

Koike H, Okayama T, Baba Y (1998) Conservation and genetics of CITES animal species. Mertensiella 9:37-44

Laursen L, Bekoff M (1978) *Loxodonta africana.* Mammalian Species vol. 92 The American Society of Mammalogist, pp 1-8.

Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ (1994) Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. Journal of Molecular Evolution 39:174-190

Lunt DH, Whipple LE, Hyman BC (1998) Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology.

Molecular Ecology 7:1441-1456

MacKay SLD, Olivo PD, Laipis PJ, Hauswirth WW (1986) Template-directed arrest of mammalian mitochondrial DNA synthesis. Molecular and Cellular Biology 6:1261-1267

Madsen CS, Ghivizzani SC, Hauswirth WM (1993) Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Molecular and Cellular Biology 13:2162-2171

Mignotte F, Gueride M, Champagne A-M, Mounolou J-C (1990) Direct repeats in the non-coding region of rabbit mitochondrial DNA. European Journal of Biochemistry 194:561-571

Morrison-Scott TCS (1948) A revision of our knowledge of African elephant's teeth, with notes on forest and "pygmy" elephants. Proceedings of The Zoological Society of London 117:505-527

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163

Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. Molecular Biology and Evolution 6:290-300

Noro M, Masuda R, Dubrovo IA, Yoshida MC, Kato M (1998) Molecular phylogenetic inference of the woolly mammoth Mammuthus primigenius, based on complete sequences of mitochondrial cytochrome b and 12S ribosomal RNA genes. Journal of Molecular Evolution 46:314-326

Nowak RM (1991) In: Walker's mammals of the world. The Johns Hopkins University Press, Pp. 1614

Olivier RCD (1978) Distribution and status of the Asian elephant. Oryx 14:379-424

Ozawa T, Hayashi S, Mikhelson VM (1997) Phylogenetic position of mammoth and Steller's sea cow within tethytheria demonstrated by mitochondrial DNA sequence. Journal of Molecular Evolution 44:406-413

Republic of South Africa (1994) A proposal for the transfer of South African's African elephant population from Appendix I to Appendix II. pp 34.

Saccone C, Attimonelli M, Sabisa E (1987) Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. Journal of Molecular Evolution 26:205-211

Saccone C, Pesol G, Sbisa E (1991) The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. Journal of Molecular Evolution 33:83-91

Said M, Chunge R (1994) African elephant database A preliminary update: November 1994. IUCN, pp 169

Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Journal of Molecular Evolution 4:406-425

Santiapillai C, Jackson P (1990) The Asian elephant: An action plan for its conservation. IUCN, Gland

Santiapillai C (1998) Distribution, number and status of the Asian elephant in the wild. A paper presented at the World Wilderness Congress, Bangalore, India

Shoshani J, Golenberg EM, Yang H (1998) Elephantidae phylogeny: morphological versus molecular result. Acta Theriologica, Suppl. 5:89-122

Southern SO, Southern PJ, Dizon AE (1988) Molecular characterization of a cloned dolphin mitochondrial genome. Journal of Molecular Evolution 28:32-42

Steinborn R, Müller M, Berm G (1998) Genetic variation in functionally important domains of the bovine mtDNA control region. Biochimica et Biophysica Acta 1397:295-304

Stewart DT, Baker AJ (1994) Patterns of sequence variation in the mitochondrial D-loop region of shrews. Molecular Biology and Evolution 11:9-21

Tiedemann R, Kurt F, Haberhauer A, Hartl GB (1998) Differences in spatial genetic population structure between African and Asian elephant (*Loxodonta africana* and *Elephas maximus*) as revealed by sequence analysis of the mitochondrial Cyt b gene. Acta Theriologica, Suppl. 5:123-134

Trouessart EL (1991) L'elephant d'eau. La Nature 39:98-101

Walberg MW, Clayton DA (1981) Sequence and properties of the human KB cell and mouse L cell D-loop region of mitochondrial DNA. Nucleic Acids Research 9:5411-5421

Wayne RK, Meyer A, Lehman N, Valkenburgh BV, Kat PW, Fuller TK, Girman D, O'Brien SJ (1990) Large sequence divergence among mitochondrial DNA genotypes within populations of eastern African black-backed jackals. Proceedings of National Academy of Science of The United States of America 87:1772-1776

Wilkinson GS, Chapman AM (1991) Length and sequence variation in evening bat D-loop mtDNA. Genetics 128:607-617

Wilkinson GS, Mayer F, Kerth G, Petri B (1997) Evolution of repeated sequence arrays in the D-loop region of bat mitochondrial DNA. Genetics 146:1035-1048

Wilson AC, Cann RL, Carr SM, George M Jr, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985)

Mitochondrial DNA and two perspectives on evolutionary genetics. Biological Journal of Linneous Society 26:375-400

Wood NJ, Phua SH (1996) Variation in the control region sequence of the sheep mitochondrial genome. Animal genetics 27:25-33

Yang H, Golenberg EM, Shoshani J (1996) Phylogenetic resolution within the Elephantidae using fossil DNA sequence from the American mastodon (*Mammut americanum*) as an outgroup. Proceedings of National Academy of Science of The United States of America 93:1190-1194