



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

A DOCTORAL DISSERTATION

Molecular analysis of Korean *Apodemus*
mitochondrial genome and possible historical
migration routes of *A. chejuensis*

GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

Department of Life Science

Dae-Ju Oh

August, 2012

한국산 붉은쥐속의 미토콘드리아 유전체 분석
및 제주등줄쥐의 역사적 이주경로 추정

지도교수 김 세 재

오 대 주

이 논문을 이학 박사학위 논문으로 제출함

2012년 8월

오대주의 이학 박사학위 논문을 인준함

심사위원장	오 흥 식	
위 원	김 원 택	
위 원	김 동 환	
위 원	한 상 권	
위 원	김 세 재	

제주대학교 대학원

2012년 8월

**Molecular analysis of Korean *Apodemus*
mitochondrial genome and possible historical
migration routes of *A. chejuensis***

Dae-Ju Oh

(Supervised by Professor Se-Jae Kim)

**A dissertation submitted in partial fulfillment of the requirement for
the degree of Doctor of Science**

August, 2012

This dissertation has been examined and approved by

Hongshik Oh

Chairperson of the Committee

Dontack Kim

Yonghwan JUNG

Sanghyun Ham

Se Jae Kim

Aug. 14 2012

Date

**Department of Life Science
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY**

CONTENTS

LIST OF TABLES	iii
LIST OF FIGURES	iv
ABBREVIATION	vi
BACKGROUND	11
ABSTRACT	18
Chapter 1. Comparative analysis of the complete mitochondrial genome sequences of Korean <i>Apodemus</i>	
1.1. INTRODUCTION	23
1.2. MATERIALS AND METHODS	26
1.2.1. Sampling and DNA extraction	26
1.2.2. PCR, cloning, and sequencing	26
1.2.3. Sequence analysis of mt genomes of Korean <i>Apodemus</i>	30
1.2.4. Phylogenetic analysis	30
1.3. RESULTS	32
1.3.1. Mitochondrial genomes of Korean <i>Apodemus</i>	32
1.3.2. Protein-coding genes	32
1.3.3. rRNA and tRNA encoding genes	32
1.3.4. Control region	47
1.3.5. Phylogenetic relationships	47
1.4. DISCUSSION	51
1.4.1. Comparative analysis of mitochondrial genome sequences	51
1.4.2. Phylogenetic relationships of Korean <i>Apodemus</i>	54

Chapter 2. Possible historical migration routes of *Apodemus chejuensis*

2.1 INTRODUCTION	57
2.2. MATERIALS AND METHODS	63
2.2.1. Sampling and DNA extraction	63
2.2.2. PCR and sequencing	64
2.2.3. Sequence analysis	68
2.2.4. Divergence time estimation	68
2.3. RESULTS	70
2.3.1. Haplotype and demographic analysis	70
2.3.2. Divergence time	78
2.4. DISCUSSION	79
2.4.1. Haplotype analysis of <i>Apodemus agrarius</i> and <i>A. chejuensis</i>	79
2.4.2. Possible migration histories of <i>Apodemus chejuensis</i>	80
LITERATURE CITED	84
국문 초록	102
Appendix 1. The sequence of mitochondrial genome of <i>Apodemus agrarius</i>	106
Appendix 2. The sequence of mitochondrial genome of <i>Apodemus chejuensis</i>	117
Appendix 3. The sequence of mitochondrial genome of <i>Apodemus peninsulae</i>	128

LIST OF TABLES

Chapter 1. Comparative analysis of the complete mitochondrial genome sequences of Korean *Apodemus*

Table 1.1. PCR primers used in this study	1
Table 1.2. List of Rodentia mitochondrial genomes	31
Table 1.3. Location of feature in the mt genome of <i>A. agrarius</i>	1
Table 1.4. Location of feature in the mt genome of <i>A. chejuensis</i>	1
Table 1.5. Location of feature in the mt genome of <i>A. peninsulae</i>	1
Table 1.6. Nucleotide frequencies of 13 protein-coding genes	1
Table 1.7. Codon usage based on 13 protein-coding genes of <i>A. agrarius</i>	39
Table 1.8. Codon usage based on 13 protein-coding genes of <i>A. chejuensis</i>	40
Table 1.9. Codon usage based on 13 protein-coding genes of <i>A. peninsulae</i> ...	41

Chapter 2. Possible historical migration routes of *Apodemus chejuensis*

Table 2.1. Specimen information used in this study	1
Table 2.2. Genetic distance matrix among haplotype sequences	1
Table 2.3. Population analysis of <i>A. agrarius</i> and <i>A. chejuensis</i>	1

LIST OF FIGURES

Chapter 1. Comparative analysis of the complete mitochondrial genome sequences of Korean *Apodemus*

Figure 1.1. The mitochondrial genome structure of Korean <i>Apodemus</i>	33
Figure 1.2. The putative structures of mt tRNAs of <i>A. agrarius</i>	43
Figure 1.3. The putative structures of mt tRNAs of <i>A. chejuensis</i>	44
Figure 1.4. The putative structures of mt tRNAs of <i>A. peninsulae</i>	45
Figure 1.5. The putative structures of mt <i>tRNA-Ser^(AGY)</i> found in Korean <i>Apodemus</i>	46
Figure 1.6. Common features mt L-strand replication origin sequences of Korean <i>Apodemus</i>	48
Figure 1.7. Conserved sequences of the mitochondrial control region of Korean <i>Apodemus</i>	49
Figure 1.8. Phylogenetic tree of Korean <i>Apodemus</i>	50

Chapter 2. Possible historical migration routes of *Apodemus chejuensis*

Figure 2.1. Origination and expansion of ancestral lineage of <i>A. agrarius</i> into Korean Peninsula	60
Figure 2.2. Collection sites of <i>A. agrarius</i> and <i>A. chejuensis</i>	57
Figure 2.3. The variable sites of the <i>CYTB</i> haplotype sequences of <i>A. agrarius</i> and <i>A. chejuensis</i>	71
Figure 2.4. Mismatch distribution analysis of the <i>CYTB</i> sequences	76
Figure 2.5. Median network tree inferred from <i>CYTB</i> haplotypes	77
Figure 2.6. Possible migration routes of <i>A. chejuensis</i>	83

ABBREVIATION

A or Ala	alanine
AA	amino acid
<i>ATPase</i>	<i>ATPase subunit gene</i>
bp	base pairs
<i>BRCA1</i>	<i>breast cancer susceptibility gene 1</i>
C or Cys	cysteine
<i>COI-3</i>	<i>cytochrome c oxidase subunit 1-3 gene</i>
CR	control region
CSB	conserved sequence block
<i>CYTB</i>	<i>cytochrome b gene</i>
D or Asp	aspartic acid
DHU	dihydrouridine
D-loop	displacement loop
E or Glu	glutamic acid
EtBr	ethidium bromide
F or Phe	phenylalanine
G or Gly	glycine
<i>GHR</i>	<i>growth hormone receptor gene</i>
H or His	histidine
H-strand	heavy strand
I or ILe	isoleucine
<i>IRBP</i>	<i>interphotoreceptor retinoid-binding protein gene</i>
K or Lys	lysine
kbp	kilobase pairs
L or Leu	leucine
<i>LCAT</i>	<i>lecithin cholesterol acyl transferase gene</i>

L-strand	light strand
M or Met	methionine
ML	maximum likelihood
mt	mitochondria(l)
Mya	million years ago
N or Asn	asparagine
<i>ND1-6</i>	<i>NADH dehydrogenase subunit 1-6 gene</i>
O _L	light strand replication origin
P or Pro	proline
Q or Gln	glutamine
R or Arg	arginine
<i>RAG1</i>	<i>recombination activating gene 1</i>
rDNA	ribosomal DNA
rRNA	ribosomal RNA
S or Ser	serine
SNP	single nucleotide polymorphism
T or Thr	threonine
TAS	termination-association sequence
tRNA	transfer RNA
UV	ultraviolet
V or Val	valine
<i>vWF</i>	<i>von Willebrand factor gene</i>
W or Trp	tryptophan
Y or Tyr	tyrosine

BACKGROUND

Rodentia is the largest order of mammals, encompassing 2,277 species (Musser and Carleton, 2005). Forty percent of mammal species are rodents, and they are found on all continents but Antarctica. Most rodents eat various plants or seeds, though some have more varied diets. Because they are fast growing mammals and vectors of fatal diseases such as the plague; humans have studied them for ecology, evolution, and genetics.

Since the mid-20th century, views on the size of the Rodentia have appreciated substantially from 1,591 species (Corbet and Hill, 1980) to 2,277 species (Musser and Carleton, 2005). Rodents include mice, rats, porcupines, beavers, squirrels, hamsters, and guinea pigs. Before the study by Graur *et al.* (1991), Rodentia was believed to be a monophyletic order of Mammalia. They questioned whether the guinea pig, *Cavia porcellus*, is a rodent. They analyzed amino acid sequences and suggested that the guinea pig diverged before the separation of the primates and artiodactyls from the myomorph rodents. In 1996, analysis of the complete guinea pig mitochondrial (mt) genome indicated that the species was not a rodent (D'Erchia *et al.*, 1996). Furthermore, hystricognaths including guinea pigs were treated as a separate mammalian order, and the Rodentia was conventionally viewed as polyphyletic (Graur *et al.*, 1992; Li *et al.*, 1992). However, many molecular studies strongly supported the Rodentia as a monophyletic order (Frye and Hedges, 1995; Cao *et al.*, 1997; Adkins *et al.*, 2001; Corneli, 2002; Waddell and Shelley, 2003), and Luckett and Hartenberger (1993) proposed rodent monophyly based on morphological and molecular data. Rodentia is currently believed to be a monophyletic order.

Rodentia consists of 5 suborders; Sciuromorpha, Castorimorpha,

Myomorpha, Anomaluromorpha, and Hystricomorpha (Musser and Carleton, 2005). Sciuromorpha has 3 families (Aplodontiidae, Sciuridae, and Gliridae), Castorimorpha comprises 3 families (Castoridae, Geomyidae, and Heteromyidae), Myomorpha consists of 7 families (Dipodidae, Calomyscidae, Cricetidae, Muridae, Nesomyidae, Platacanthomyidae, and Spalacidae), Anomaluromorpha includes only 2 families (Anomaluridae and Pedetidae), and Hystricomorpha contains 18 families (Ctenodactylidae, Bathyergidae, Hystricidae, Petromuridae, Thryonomyidae, Heptaxodontidae, Abrocomidae, Capromyidae, Caviidae, Chinchillidae, Ctenomyidae, Dasyproctidae, Cuniculidae, Dinomyidae, Echimyidae, Erethizontidae, Myocastoridae, and Octodontidae) and an *incertae sedis* family (Diatomyidae). Among 5 suborders, Myomorpha is the largest, containing 1,137 species of mouse-like rodents, i.e., nearly a quarter of all mammal species. This suborder contains mice, rats, hamsters, lemmings, gerbils, and voles. Most myomorph species belong to the Muroidea superfamily, which is classified into 6 families, 310 genera, and 1,518 species (Musser and Carleton, 2005). Among the 6 families of the Muroidea, Muridae is the largest with 150 genera, with over 700 species (Musser and Carleton, 2005). In the Korean Peninsula and Jeju Island, 4 genera (*Apodemus*, *Micromys*, *Mus*, and *Rattus*) of Muridae are found (Yoon *et al.*, 2004). Muridae are relatively small mammals, generally around 10 cm long excluding the tail length. Some species are larger; the giant bushy-tailed cloud rat, *Crateromys schadenbergi*, is the largest at over 70 cm.

Although many researchers have tried to establish phylogeny and evolution of rodents (Nedbal *et al.*, 1996; Huchon *et al.*, 1999, 2000; Michaux *et al.*, 2001; Steppan *et al.*, 2004), the evolution of many small mammals including murids is unclear due to the incongruence between the molecular data and morphological information. Against traditional classification, many molecular taxonomists researched various nuclear and mt genes such as *growth hormone receptor (GHR)* gene, *breast cancer susceptibility (BRCA1)* gene,

*von Willebrand factor (vWF) gene, lecithin cholesterol acyl transferase (LCAT) gene, interphotoreceptor retinoid-binding protein (IRBP) gene, recombination activating gene 1 (RAG1), 12S ribosomal RNA (rRNA) gene, cytochrome c oxidase subunits 1-3 (COI-3) gene, and cytochrome b (CYTB) gene (Conroy and Cook, 2000; Robinson-Rechavi et al., 2000; Suzuki et al., 2000; Adkins et al., 2001; Michaux et al., 2001; Stepan et al., 2005). These molecular studies broadened our knowledge of rodent phylogeny and evolutionary history. For instance, CYTB sequence analysis suggested that North American *Microtus* are derived from multiple invasions from Asia or, alternatively, as a single invasion followed by autochthonous speciation (Conroy and Cook, 2000), and combined analysis of nuclear and mt genes suggested that the murines originated in Southeast Asia and then rapidly expanded across the Old World (Stepan et al., 2005).*

With the rapid development of sequencing technology, phylogenetic and evolutionary studies using mt genome sequences have explosively increased. Since the first mt genome was sequenced in humans (Anderson et al., 1981), mt genome sequences have been reported for various taxa, including fish (Miya et al., 2003; Oh et al., 2007b), mammals (Mouchaty et al., 2001; Horner et al., 2007), insects (Stewart et al., 2003; Covacin et al., 2006), reptiles (Dong and Kumazawa, 2005; Podnar et al., 2009), birds (Valverde et al., 1994; Harlid and Arnason, 1999), and amphibians (Zhang et al., 2006; Oh et al., 2007a), as well as other species (Milbury and Gaffney, 2005; Sinniger et al., 2007). Remarkable studies were conducted using mass mt genome sequences for vertebrate phylogeny. Miya et al. (2003) sequenced 54 mt fish genomes and established higher teleostean phylogeny with 100 mt genome sequences. In amphibians, Zhang et al. (2006) sequenced 15 mt genomes of Asian salamanders and proposed the Asian hynobiid originated from North China. Arnason et al. (2008) analyzed 109 mt genome sequences including 11 newly sequenced mt genomes for placental mammalian phylogeny. Their analyses

identified 4 basal monophyletic groups, Afroplacentalia, Xenarthra, Archontoglires, and Laurasiaplacentalia. They also estimated divergence time for basal placental divergence. As in various vertebrates, complete mt genome sequences are rapidly becoming available; nevertheless, only 33 mt genomes of rodent have been sequenced (37 mt genomes including below subspecies) until now.

The mt genome has several merits, such as its compact size (Cantatore and Saccone, 1987), lack of recombination (Clayton, 1982, 1992), maternal inheritance (Kondo *et al.*, 1990), fast evolutionary rate compared to nuclear sequences (Brown *et al.*, 1979), and multicopy status (Michaels *et al.*, 1982; Robin and Wong, 1988), for research in population genetics, phylogenetics, and evolution (e.g., Serizawa *et al.*, 2000; Liu *et al.*, 2004; Suzuki *et al.*, 2008). Comparisons of complete mt genome sequences are much more powerful for phylogenetic reconstruction than comparative analysis of single gene sequences (Ingman *et al.*, 2001; Parsons and Coble, 2001; Boore, Medina, and Rosenberg, 2004). In fact, phylogenetic analysis with enough sequences decreases stochastic uncertainty (Cao *et al.*, 1994). Accordingly, more mt genomes provide higher resolution for phylogenetic analysis.

The mt genomes are generally circular (Boore, 1999), although Cubozoa, Scyphozoa, and Hydrozoa have linear mt genomes (Bridge *et al.*, 1992). Animal mt genomes are usually around 15,000 - 20,000 base pairs (bp) in size and contain the same set of 37 genes, encoding 13 proteins, 2 rDNAs, 22 transfer RNAs (tRNA), and a control region (Boore, 1999). Commonly, mt gene contents are constant and gene arrangements are nearly identical in mammals. Various mammalian mt protein-coding genes end with an incomplete termination codon as TA or T, rather than a complete termination codon (TAA or TAG). Such incomplete termination codons can be modified by post-transcriptional polyadenylation to a complete termination codon (Ojala *et al.*, 1981). Among 13 protein-coding genes, *NADH dehydrogenase subunit*

(ND) 6 is encoded by the light strand (L-strand); the others are encoded by the heavy strand (H-strand). *tRNA-Phe*, *tRNA-Val*, *tRNA-Leu*, *tRNA-Ile*, *tRNA-Met*, *tRNA-Trp*, *tRNA-Asp*, *tRNA-Lys*, *tRNA-Gly*, *tRNA-Arg*, *tRNA-His*, *tRNA-Ser*^(AGY), and *tRNA-Thr* are encoded by the H-strand and the remainder are encoded by the L-strand (Wolstenholme, 1992). A major noncoding control region called the displacement loop (D-loop) is located between *tRNA-Pro* and *tRNA-Phe*. This sequence has several elements such as a termination-associated sequence (TAS) upstream of the 3' end of the control region in mammals (Doda *et al.*, 1981; Mackay *et al.*, 1986) and conserved sequence blocks (CSBs) associated with mt DNA replication and transcription (Clayton, 1991; Shadel and Clayton, 1997). The L-strand replication origin (O_L) is between *tRNA-Asn* and *tRNA-Cys* and comprises approximately 30 nucleotides (Wolstenholme, 1992). The O_L sequence forms a stable stem-loop structure, and a conserved motif is located in the base of the stem of *tRNA-Cys*. This motif is found in various vertebrate mt genomes (Oh *et al.*, 2008; Oh *et al.*, 2010).

Although the mt genome structure of higher animals such as mammals is almost constant, some organisms, closer to the lower animals, have unique mt genome structures (Okimoto *et al.*, 1991; Le *et al.*, 2000; Helfenbein *et al.*, 2004). The mt gene arrangements are almost the same in major groups, especially the mammals (Boore, 1999). Nevertheless, several rearrangements have been found in nematodes (Okimoto *et al.*, 1991, 1992), arthropods (Boore *et al.*, 1995), bivalves (Hoffman *et al.*, 1992), pulmonate mollusks (Yamazaki *et al.*, 1997), echinoderms (Cantatore *et al.*, 1987; Jacobs *et al.*, 1988), and fruit flies (Clary and Wolstenholme, 1985). These mt gene rearrangements are considered rare evolutionary events and are used to reconstruct the phylogeny of various organisms (Singh, 2008).

Since the development of sequence analysis tools, many evolutionary studies using mt genes or genomes estimated divergence time between

species. The molecular clock is used to estimate the time of speciation or radiation and significantly advanced the neutral theory of molecular evolution (Kimura, 1968). The divergence time estimation of rodents revealed their evolutionary histories (Jacobs and Downs, 1994; Adkins *et al.*, 2001; Douzery *et al.*, 2003; Liu *et al.*, 2004; Horner *et al.*, 2007; Arnason *et al.*, 2008; Robins *et al.*, 2008). The divergence time between Lagomorpha and Rodentia was estimated at 93 million years ago (Mya) and the deepest Rodentia divergence at around 88 Mya (Arnason *et al.*, 2008). Bayesian dating inferred from mt genomes indicate a rapid radiation within Glires around 60 Mya, and that within Rodentia occurred approximately 63 Mya (Horner *et al.*, 2007). Horner *et al.* (2007) also suggested that the divergence of the Muroidea + Dipodidae clade from the Sciuridae + Myoxidae clade occurred 53.4 Mya (protein) or 51.8 Mya (DNA). In addition, they estimated the divergence between *Mus* and *Rattus*, representative species of Muridae, occurred 15.1 (protein) or 15.9 (DNA) Mya, while the split between *Mus* and *Rattus* was approximately 13 - 19 Mya (Douzery *et al.*, 2003), 23 Mya (Adkins *et al.*, 2001), or 12 Mya (Jacobs and Downs, 1994). Molecular analyses of *Rattus* suggest that the deepest divergence within *Rattus* occurred 3.5 Mya (Robins *et al.*, 2008). For *Apodemus*, *A. peninsulae* diverged from the *A. speciosus* + *A. chevrieri* + *A. agrarius* clade about 8.39 Mya, and *A. agrarius* diverged from *A. chevrieri* about 2.68 Mya, during the first glacial period of the Quaternary (Liu *et al.*, 2004).

The whole species number is estimated from 10 to 100 million (Wilson, 1998). This wide range is due to difficulties in species definition. A species is one of the primary units of biological classification and a taxonomic rank. A biological species concept was used first by Jordan (1896) and later Mayr (1969). They defined a species as “group of interbreeding populations which are reproductively isolated from other groups”. Later, the biological species concept was redefined by Mayr (1982) as “a reproductive community of

populations (reproductively isolated from others) that occupies a specific niche in nature.” Baker and Bradley (2006) define genetic species as a group of genetically compatible interbreeding natural populations that is genetically isolated from other such groups. Since rapid development of DNA sequencing technology, many researchers have recognized and described species using DNA sequence data (Roca *et al.*, 2001; Baker *et al.*, 2002; Matocq 2002; Piaggio *et al.*, 2002), and the accumulation of DNA data solidify the adequacy of the genetic species concept.

ABSTRACT

Chapter 1. Comparative analysis of the complete mitochondrial genome sequences of Korean *Apodemus*

The complete mitochondrial (mt) genomes of *Apodemus agrarius*, *A. chejuensis*, and *A. peninsulae* were determined to be 16,260, 16,261, and 16,266 base pairs (bp) long, respectively, and have been deposited in the GenBank database (Accession Nos. HM034866, HM034867, and HQ660074, respectively). Like other vertebrate mt genomes, the 3 mt genomes contain 13 protein-coding genes, 22 tRNAs, 2 rDNAs and a noncoding control region. The mt genomes of *A. agrarius* and *A. chejuensis* had protein-coding genes that use the most common initiation codon, ATG, and 2 unusual initiation codons: GTG and ATA. However, *A. peninsulae* had an additional unusual ATC codon for initiation in the *NADH dehydrogenase subunit (ND) 2* gene. Overall, the nucleotide frequencies of all analyzed mt protein-coding genes were similar. Codon usage patterns based on mt protein-coding genes were similar in *A. agrarius* and *A. chejuensis*, while the mt protein-coding genes of *A. peninsulae* were slightly different. The *12S rRNA* gene sequences were 955, 955, and 956 bp long in *A. agrarius*, *A. chejuensis*, and *A. peninsulae*, respectively, and the *16S rRNA* gene sequences were 1,572, 1,573, and 1,571 bp long, respectively. *tRNA-Leu* and *tRNA-Ser* were identified two distinct forms, respectively. The Three tRNA clusters, IQM (isoleucine, glutamine, and methionine), WANCY (tryptophan, alanine, asparagine, cysteine, and tyrosine), and HSL (histidine, serine, and leucine), were conserved in *A. agrarius*, *A. chejuensis*, and *A. peninsulae*, as in the typical mt genomes of Rodentia.

Generally, the mt genomes of mammals have 2 noncoding regions: O_L and CR, which is also called the displacement loop (D-loop). The CR sequences spanned 854, 857, and 866 bp in *A. agrarius*, *A. chejuensis*, and *A. peninsulae*, respectively, and was positioned between the *tRNA-Pro* and *tRNA-Phe* genes. The phylogenetic analysis of Rodentia using mt 13 protein-coding genes suggested that *Sciurus* and *Myoxus* was diverged earlier than other rodents, and *Rattus* consisted of two distinct clades. Korean *Apodemus* was closely related to *Mus* as sister group. Although phylogenetic relationships of *Apodemus* spp. have been studied by many researchers, some species remain controversial owing to wide distribution, subspecies problems, and fragmentary data. Thus, the mt genome of Korean *Apodemus* spp. will be useful information to clarify rodent, especially the *Apodemus*, phylogeny.

Chapter 2. Possible historical migration routes of *Apodemus chejuensis*

Nineteen *CYTB* sequence haplotypes were defined in 24 *Apodemus agrarius* individuals, and 25 were defined in 49 *A. chejuensis* individuals. The most frequent haplotype group was Hap27, which was shared by 6 *A. chejuensis* individuals. The most frequent haplotype group in the *A. agrarius* was Hap15, which was shared by 3 individuals. Fu's *F_s* test showed that COR-II and CHE had significantly negative *F_s* values (-4.462 and -5.703, respectively), whereas the COR-I had a positive value (0.613). Mismatch distribution analyses suggested population expansion in COR-II and CHE and a stable COR-I population. Median-joining network data indicated that *A. chejuensis* is a clearly separate species containing 4 subgroups, but these subgroups did not correspond to local distribution patterns on Jeju Island. Interestingly, Hap42 of the CHE population was closely related to Hap9 of the COR-I. In addition, molecular clock calculations that estimate the time of divergence of *A. agrarius* and *A. chejuensis* suggest that those species diverged at approximately 0.3 million years ago (Mya). Based on these data, three historical migration routes were hypothesized. (1) The ancestral lineage of *A. agrarius* might migrate into the Korean Peninsula via North Korea; then, a group migrated to Jeju Island via the Yellow Sea basin when the sea level was lower. (2) The ancestral lineage of *A. chejuensis* might migrate to Jeju Island via the Yellow Sea basin, followed by a second migration into the western islet of the Korean Peninsula via the Yellow Sea basin when the sea level was lower. (3) The common ancestor of *A. agrarius* and *A. chejuensis* might have originated in eastern China, where it split into 2 lineages; one lineage moved to the western islet of the Korean Peninsula, and the other moved to Jeju Island. The estimation of divergence time suggested that *A.*

agrarius and *A. chejuensis* diverged at approximately 0.3 Mya. Because *A. chejuensis* has been geographically and reproductively isolated from *A. agrarius* since the last glacial age, many researchers have proposed that populations diverged as species. This study supports the notion that *A. chejuensis* should be treated as a species; however, further study of North Korea and Chinese samples is needed to understand their biogeographic history fully.

CHAPTER 1

Comparative analysis of the complete mitochondrial
genome sequences of Korean *Apodemus*

1.1. INTRODUCTION

Rodentia is the largest order of mammals, encompassing 2,277 species (Musser and Carleton, 2005). Among them, Muridae is the largest family with over 700 species (Musser and Carleton, 2005) and is often employed as a laboratory model for studies of human diseases. On the Korean Peninsula and Jeju Island, 4 genera (*Apodemus*, *Micromys*, *Mus*, and *Rattus*) of Muridae are found (Yoon *et al.*, 2004a).

The genus *Apodemus* contains at least 20 rodent species (Musser and Carleton, 2005), including several of the most common small rodent species and is widely distributed in the Palearctic region (the zoogeographical region consisting of Europe, Africa north of the Sahara, and most of Asia north of the Himalayas). *Apodemus* species are distributed in a variety of habitats, including woodlands, forests, and grasslands. Some species are often found in the same habitat, showing resource partitioning among syntopic species. This ecological feature could be associated with the biogeographic history of the genus and its speciation processes. The striped field mouse *Apodemus agrarius* and Korean Field mouse *A. peninsulae* are distributed throughout the Korean Peninsula and are the most common rodent species there, whereas the Jeju Striped Field mouse, *A. chejuensis*, is found only on Jeju Island (Won and Smith, 1999). Generally, *A. agrarius* inhabits rural flat lands, rivers, and agricultural farmlands, while the Jeju Striped Field mouse, *A. chejuensis*, is often found in bushy areas rich with small trees in mountainous areas. They are distributed from the coastal regions to the high-altitude regions of Mt. Halla.

The wide distribution, subspecies problems, and fragmentary data of *Apodemus* spp. were so complicated that even experts have been unable to

comprehend its phylogenetic relationships (Suzuki *et al.*, 2008). Molecular and morphological analyses improved the taxonomic situation of *Apodemus* spp.; however, the difficult cases continue to complicate the relationships between closely related species (Musser and Carleton 2005). In particular, the taxonomic status of some species inhabiting the western Palearctic region including subgenus *Sylvaemus* are complicated owing to several new species in southern Europe and the Middle East (Musser and Carleton 2005). Their phylogenetic relationships are not established, and the species *A. sylvaticus* and *A. flavicollis* remain uncertain (Suzuki *et al.*, 2008). Previous phylogenetic results inferred from morphological and molecular data (Serizawa *et al.*, 2000; Filippucci *et al.*, 2002; Michaux *et al.*, 2002; Suzuki *et al.*, 2003; Liu *et al.*, 2004; Suzuki *et al.*, 2008), suggested that the genus *Apodemus* consists of 8 distinct lineages; nevertheless, previous data are insufficient to solve their phylogenetic complexity.

The low resolution of molecular phylogenies inferred from fragmentary data has prompted many researchers to analyze mitochondrial (mt) genome sequences. The mt are small organelles found in most eukaryotic cells and play an important role in metabolism (Brand, 1997), cell death *via* apoptosis (Kroemer *et al.*, 1998), various diseases (Graeber and Muller, 1998), and aging (Wei, 1998). The mt is essential for ATP production and has its own genome. In general, the vertebrate mt genome consists of 16 - 18 kilobase pairs (kbp) and contains 22 transfer RNAs (tRNAs) necessary for the translation of mt proteins, 2 ribosomal DNAs (rDNAs), 13 protein-coding genes related to oxidative phosphorylation, and a major noncoding region for replication and transcription initiation (Anderson *et al.*, 1981; Boore, 1999). Of the 13 protein-coding genes, 12 are encoded on the heavy strand (H-strand), and 1 gene, *NADH dehydrogenase subunit 6 (ND)* gene, is encoded on the light strand (L-strand). Fourteen mt tRNAs are encoded on the H-strand and the rest are encoded on the L-strand (Wolstenholme, 1992). The mt

protein-coding genes in most animals have incomplete termination codons such as TA or T. These abnormal termination codons can be modified to a complete termination codon by post-transcriptional polyadenylation (Ojala *et al.*, 1981). The mt genome has 2 noncoding regions: the major noncoding control region (CR) called the displacement loop (D-loop) and the L-strand replication origin (O_L), which mediates replication of the mt genome (Clayton, 1991; Shadel and Clayton, 1997).

The mt genome has several merits for research in population genetics; phylogenetics; and evolution, such as its compact size; fast evolutionary rate; short coalescence time; and maternal inheritance (e.g., Serizawa *et al.*, 2000; Liu *et al.*, 2004; Suzuki *et al.*, 2008). Mammalian mt DNA has been used for phylogenetic and evolutionary studies since the development of DNA sequencing techniques. In the last decade, mt genome sequences have been reported for various taxa, including fish (Miya *et al.*, 2003; Oh *et al.*, 2007b), mammals (Mouchaty *et al.*, 2001; Horner *et al.*, 2007), insects (Stewart *et al.*, 2003; Covacin *et al.*, 2006), reptiles (Dong and Kumazawa, 2005; Podnar *et al.*, 2009), and amphibians (Zhang *et al.*, 2006; Oh *et al.*, 2007a), as well as other species (Milbury and Gaffney, 2005; Sinniger *et al.*, 2007). Although Rodentia contains the most mammal species and the mt genomes of the mouse and rat were sequenced first (Bibb *et al.*, 1981; Gadaleta *et al.*, 1989), and only a small proportion of mt genome sequences of Rodentia have been reported. The mt genome sequences of 33 species (excluding subspecies) in Rodentia have been determined (or 1.5% of the 2,277 species of Rodentia).

To provide the basal data set for *Apodemus* and all rodent phylogeny, the complete mt genome sequences of Korean *Apodemus* were sequenced and their general features analyzed, including nucleotide composition, nucleotide frequency, codon usage, secondary cloverleaf tRNA structures, the O_L , and CR.

1.2. MATERIALS AND METHODS

1.2.1. Sampling and DNA extraction

Apodemus agrarius and *A. peninsulae* were collected in traps in Cheongdo, Gyeongsangbuk-do; *A. chejuensis* was caught at Mt. Halla, Jeju, Korea. Genomic DNA was extracted from small pieces of tail or muscle using NucleoSpin[®] Tissue Kit (Macherey-Nagel, Germany). First, small pieces of tail or muscle was placed in a microcentrifuge tube with buffer T1 and proteinase K solution, then vortexed and incubated at 56°C until lysis was complete. The lysed samples were vortexed and buffer B3 was added. The tubes were vortexed and incubated at 70°C for 10 min. Ethanol 1 was added and vortexed. For each sample, a NucleoSpin[®] Tissue Column was placed into a collection tube and the samples applied to the column. The assembly was centrifuged for 1 min at 11,000g. The flow-through was discarded, and the column was returned to the collection tube. To wash the silica membrane, buffer BW and buffer B5 were added in turn. The column-collection tube assembly was centrifuged for 1 min at 11,000g after each wash step. The silica membrane was dried by centrifugation. Last, the column was placed into a new 1.5 ml tube and prewarmed buffer BE was added, then incubated at room temperature for 1 min. The tube was centrifuged for 1 min at 11,000g, and the extracted DNA was stored at -20°C.

1.2.2. PCR, cloning, and sequencing

To amplify the complete mt genome, long PCR was performed with primers complementary to the aligned and complete nucleotide sequences of the mt genomes of several mouse species (Table 1.1). The long PCR reactions were carried out using Maxime[™] PCR PreMix (iNtRON

Biotechnology, Korea) according to the manufacturer's protocol. 50–100 ng of genomic DNA and 1 μ l of each primer was added into Maxime™ PCR PreMix tubes. Reaction conditions included an initial 3 min denaturation at 95°C, followed by 35 cycles with denaturation at 95°C for 30 s, annealing for 30 s at 55°C and extension at 72°C for 2–5 min, and a final extension of 10 min at 72°C. Amplification products were confirmed by electrophoresis with 100 bp Plus DNA Ladder (Bioneer, Korea) as a size marker in 1% agarose gels, stained with ethidium bromide (EtBr), and visualized under ultraviolet (UV) light.

The amplified fragments were cloned using TOPO TA Cloning Kit (Invitrogen, USA). PCR products were ligated into the pCR2.1 TOPO vector according to the manufacturer's protocol. The ligation mixtures were mixed gently and incubated for 5 min at room temperature. The reaction mixtures were put on ice immediately and incubated for 2 min. Following ligation, 30 μ l of competent cells were defrosted in ice and added to the reaction, which was heat shocked at 42°C for 30 s then put on ice for 2 min. SOC medium was added and the reaction mixtures were shaken at 37°C for 90 min. The reaction mixtures were spread on LB agar plates containing X-Gal, IPTG, and ampicillin. The plates were incubated at 37°C for 18 h. White colonies were picked into LB medium containing ampicillin and incubated at 37°C with shaking overnight.

Positive clones were verified by PCR with the original amplification primers and 1 μ l of cultured cells as template. Reaction conditions included an initial 3 min denaturation at 95°C, followed by 25 cycles with denaturation at 95°C for 30 s, annealing for 30 s at 55°C, and extension at 72°C for 2–5 min, and a final extension of 10 min at 72°C. Amplification products were confirmed by electrophoresis with 100 bp Plus DNA Ladder (Bioneer, Korea) as size standard marker in 1% agarose gels, stained with EtBr, and visualized under UV light. Plasmid DNA was extracted from the positive clones.

Plasmid DNA was extracted with the *AccuPrep*[®] Nano-Plus Plasmid Mini Extraction Kit (Bioneer, Korea). The cultured cells were centrifuged at 13,000 rpm for 1 min and the media was removed by pipetting. Resuspension buffer containing nanoparticles was added and the pellet was completely resuspended by pipetting. Cell lysis buffer was added and mixed by gently inverting the tube 3-4 times. Neutralization buffer was added and immediately mixed by gently inverting the tube 3-4 times. Cell debris was removed by centrifugation at 4°C and 13,000 rpm for 1 min and the cleared lysate was transferred to the DNA binding column. The columns were centrifuged at 13,000 rpm for 1 min and the flow-through was discarded. To wash the silica membrane, washing buffer was added to the column tube and centrifuged at 13,000 rpm for 1 min. Additional centrifugation was carried out at 13,000 rpm for 1 min to dry the silica membrane in the column. The DNA binding filter column was transferred to new 1.5 ml microcentrifuge tube and elution buffer was added and incubated at room temperature for 1 min. To recover the eluted plasmid DNA, the tube was centrifuged at 13,000 rpm for 1 min, and the DNA was transferred to a new 1.5 ml microcentrifuge tube.

Each cloned fragment was sequenced with TOPO vector inner primers (Invitrogen, USA) using an automatic DNA sequencer (ABI 3730xl; Applied Biosystems, USA) according to the manufacturer's protocol, and nested flanking sequencing primers were used to walk in both directions.

Table 1.1. PCR primers used in this study

Name	Sequences (5' → 3')
Apomt-1F	TACCCTCACCTGAATTGGAGGC
Apomt-1R	GGGTTAATCGTATGACCGCG
Apomt-2F	CGCGGTCATACGATTAACCC
Apomt-2-1F*	GCAATGAAGTACGCACACACC
Apomt-2R	AGATAGAAACCGACCTGGATTGC
Apomt-3F	ACATCCCAATGGTGCAGAAG
Apomt-3-1F*	CCACTACCAATACCTCACCCCTC
Apomt-3-2F*	TCATAGCGAGTATCCCACCAT
Apomt-3-3F*	ATTGGAGCATGAGGTGGACT
Apomt-3-2R*	CGTGGAATGCTATGTCTGG
Apomt-3-1R*	GGCAGATGTGAAGTAAGCTCG
Apomt-3R	GGGGTTCGAWTCCTTCCTTTC
Apomt-4F	GAAAGGAAGGAATCGAACCCC
Apomt-4-1R*	GAATGAGTGAAGGCGGTTGT
Apomt-4R	GCATGRGTTTGGTGKGCAT
Apnmt-5F	ATGACACACCAAACCCATGC
Apomt-5-1F*	GGATTTGAAGCAGCAGCATG
Apomt-5-1R*	CTGTTTGGCGTAAGCAGATTG
Apomt-5R	GGYTCCTAAGACCAAYGGAT
Apomt-6F	ATCCATTGGTCTTAGGAACC
Apomt-6-1F*	TGGACGATCAGACGCAAATAC
Apomt-6-2F*	CAACGCCTGAGCCCTAATAAT
Apomt-6-2R*	ATGAGCGGGTTTGTAGGTTG
Apomt-6-1R*	AGAAGCCCCCTCAGATTTCAT
Apomt-6R	GCCTCCAATTCAGGTGAGGGTA

Asterisk indicates nested flanking primers.

1.2.3. Sequence analysis of mt genomes of Korean *Apodemus*

The complete mt genome sequences of *Apodemus agrarius*, *A. chejuensis*, and *A. peninsulae* were compared with the mt genome sequences of other rodents in the GenBank database. The locations of the 13 protein-coding genes were determined by comparing nucleotide or amino acid sequences and identified 2 rDNAs and the CR. The 22 tRNA genes were identified using tRNAscan-SE (Lowe and Eddy, 1997) and ARWEN (Laslett and Canbäck, 2008).

The size of complete mt genome, gene lengths, and initiation and termination codons of the 13 protein-coding genes were compared. In addition, nucleotide composition, nucleotide frequency, and codon usage of the 13 protein-coding genes, O_L , and the putative tRNA cloverleaf structures of 3 Korean *Apodemus* mt genomes were compared.

1.2.4. Phylogenetic analysis

Phylogenetic relationships were analyzed by maximum likelihood (ML) methods. The mt 13 protein-coding genes of other Rodentia were available in GenBank database (Table 1.2) and aligned with those of Korean *Apodemus* using MEGA5 software (Tamura *et al.*, 2011). The best model of nucleotide evolution was determined using jmodeltest (Posada, 2008). The GTR+I+G model was determined based on the Akaike Information Criterion (Akaike, 1973). Tree reconstruction was conducted using MEGA5 software and bootstrap analysis were conducted with 1,000 replications. *Oryctolagus cuniculus* (AJ001588; Lagomorpha, Leporidae) was used as outgroup species.

Table 1.2. List of Rodentia mitochondrial genomes

Order	Family	Species	Size	Accession Num.
Lagomorpha	Leporidae	<i>Oryctolagus cuniculus</i>	17,245	AJ001588
Rodentia	Anomaluridae	<i>Anomalurus sp.</i>	16,923	AM159537
	Caviidae	<i>Cavia porcellus</i>	16,801	AJ222767
	Cricetidae	<i>Cricetulus griseus</i>	16,284	DQ390542
	Cricetidae	<i>Eothenomys chinensis</i>	16,362	FJ483847
	Cricetidae	<i>Eothenomys regulus</i>	16,379	JN629046
	Cricetidae	<i>Mesocricetus auratus</i>	16,264	EU660218
	Cricetidae	<i>Microtus kikuchii</i>	16,312	AF348082
	Cricetidae	<i>Microtus levis</i>	16,283	DQ015676
	Cricetidae	<i>Proedromys sp.</i>	16,296	FJ463038
	Cricetidae	<i>Tscherskia triton</i>	16,488	EU031048
	Cricetidae	<i>Microtus fortis</i>	16,310	JF261174
	Dipodidae	<i>Jaculus jaculus</i>	16,546	AJ416890
	Gliridae	<i>Myoxus glis</i>	16,602	AJ001562
	Muridae	<i>Leggadina lakedownensis</i>	16,262	EU305668
	Muridae	<i>Mus musculus</i>	16,299	AY172335
	Muridae	<i>Mus terricolor</i>	16,310	EU352649
	Muridae	<i>Pseudomys chapmani</i>	16,249	EU305669
	Muridae	<i>Rattus exulans</i>	16,307	EU273710
	Muridae	<i>Rattus fuscipes</i>	16,301	GU570664
	Muridae	<i>Rattus leucopus</i>	16,297	GU570659
	Muridae	<i>Rattus norvegicus</i>	16,313	AY172581
	Muridae	<i>Rattus praetor</i>	16,293	EU273708
	Muridae	<i>Rattus rattus</i>	16,305	EU273707
	Muridae	<i>Rattus sordidus</i>	16,309	GU570665
	Muridae	<i>Rattus tanezumi</i>	16,306	EU273712
	Muridae	<i>Rattus tunneyi</i>	16,292	GU570662
	Muridae	<i>Rattus villosissimus</i>	16,303	GU570663
	Sciuridae	<i>Sciurus vulgaris</i>	16,507	AJ238588
	Spalacidae	<i>Spalax ehrenbergi</i>	16,408	AJ416891
	Thryonomyidae	<i>Thryonomys swinderianus</i>	16,626	AJ301644
	Muridae	<i>Apodemus agrarius</i>	16,260	This study
	Muridae	<i>Apodemus chejuensis</i>	16,261	This study
	Muridae	<i>Apodemus peninsulae</i>	16,266	This study

1.3. RESULTS

1.3.1. Mitochondrial genomes of Korean *Apodemus*

The complete mt genomes of *Apodemus agrarius*, *A. chejuensis*, and *A. peninsulae* were determined to be 16,260, 16,261, and 16,266 bp long, respectively (Fig. 1.1 and Table 1.3-5), and have been deposited in the GenBank database (Accession Nos. HM034866, HM034867, and HQ660074, respectively). Like other vertebrate mt genomes, the *Apodemus* mt genomes contain 13 protein-coding genes, 22 tRNAs, 2 rDNAs, and a noncoding control region (Fig. 1.1).

1.3.2. Protein-coding genes

The mt genomes of *Apodemus agrarius* and *A. chejuensis* had protein-coding genes that use the most common initiation codon, ATG, and 2 unusual initiation codons: GTG and ATA (Table 1.3-5). *A. peninsulae* had an additional unusual ATC initiation codon in the *ND2* gene (Table 1.5).

Four termination codons were found in the *Apodemus agrarius*, *A. chejuensis*, and *A. peninsulae* mt protein-encoding genes: TAA, TAG, incomplete TA, and T (Table 1.3-5).

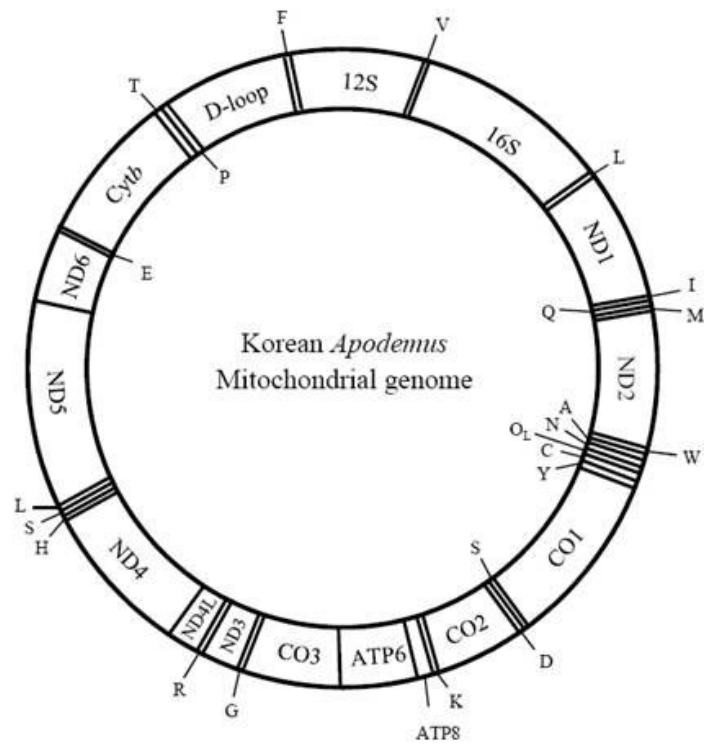


Figure 1.1. The mitochondrial genome structure of Korean *Apodemus*. The complete mt genome have 13 protein-coding genes, 22 tRNAs, 2 rDNAs, and control region. The mt gene order is identical to those of other mammalian mt genomes.

Table 1.3. Location of feature in the mt genome of *A. agrarius*

Feature ^a	Position		Size (bp)	Codon	
	5'	3'		Initiaton	Termination
<i>tRNA-Phe</i>	1	67	67		
<i>12S rRNA</i>	68	1022	955		
<i>tRNA-Val</i>	1023	1091	69		
<i>16S rRNA</i>	1092	2663	1572		
<i>tRNA-Leu</i> ^(<i>UUR</i>)	2664	2738	75		
<i>ND1</i>	2739	3695	957	GTG	T--
<i>tRNA-Ile</i>	3694	3762	69		
<i>tRNA-Gln</i>	3760	3834	75		
<i>tRNA-Met</i>	3835	3903	69		
<i>ND2</i>	3904	4941	1038	ATA	T--
<i>tRNA-Trp</i>	4940	5007	68		
<i>tRNA-Ala</i>	5009	5077	69		
<i>tRNA-Asn</i>	5079	5149	71		
<u><i>O_L</i></u>	5150	5180	31		
<i>tRNA-Cys</i>	5181	5248	68		
<i>tRNA-Tyr</i>	5249	5315	67		
<i>CO1</i>	5317	6861	1545	ATG	TAA
<i>tRNA-Ser</i> ^(<i>UCN</i>)	6859	6927	69		
<i>tRNA-Asp</i>	6931	6999	69		
<i>CO2</i>	7000	7683	684	ATG	TAA
<i>tRNA-Lys</i>	7687	7750	64		
<i>ATPase 8</i>	7752	7955	204	ATG	TAA
<i>ATPase 6</i>	7913	8593	681	ATG	TAA
<i>CO3</i>	8593	9377	785	ATG	T--
<i>tRNA-Gly</i>	9377	9445	69		
<i>ND3</i>	9445	9792	348	ATA	TAA
<i>tRNA-Arg</i>	9794	9861	68		
<i>ND4L</i>	9864	10160	297	ATG	TAA
<i>ND4</i>	10154	11531	1378	ATG	T--
<i>tRNA-His</i>	11532	11599	68		
<i>tRNA-Ser</i> ^(<i>AGY</i>)	11600	11658	59		
<i>tRNA-Leu</i> ^(<i>CUN</i>)	11658	11729	72		
<i>ND5</i>	11729	13558	1830	ATA	TAA
<u><i>ND6</i></u>	13536	14054	519	ATG	TAA
<i>tRNA-Glu</i>	14055	14123	69		
<i>CYTB</i>	14127	15270	1144	ATG	T--
<i>tRNA-Thr</i>	15271	15339	69		
<i>tRNA-Pro</i>	15340	15406	67		
Control region	15407	16260	854		

^a*ND1-6*, *NADH dehydrogenase subunit 1-6* gene; *CO1-3*, *cytochrome c oxidase subunit 1-3* gene; *ATPase 6* and *8*, *ATPase subunit 6* and *8* gene; *CYTB*, *cytochrome b* gene; *O_L*, L-strand replication origin; CR, control region. Underlines indicate a gene encoded on the L-strand.

Table 1.4. Location of feature in the mt genome of *A. chejuensis*

Feature ^a	Position		Size (bp)	Codon	
	5'	3'		Initiaton	Termination
<i>tRNA-Phe</i>	1	67	67		
<i>12S rRNA</i>	68	1022	955		
<i>tRNA-Val</i>	1023	1091	69		
<i>16S rRNA</i>	1091	2663	1573		
<i>tRNA-Leu^(UUR)</i>	2664	2738	75		
<i>ND1</i>	2739	3695	957	GTG	T--
<i>tRNA-Ile</i>	3694	3762	69		
<i>tRNA-Gln</i>	3760	3830	71		
<i>tRNA-Met</i>	3835	3903	69		
<i>ND2</i>	3904	4941	1038	ATA	T--
<i>tRNA-Trp</i>	4940	5007	68		
<i>tRNA-Ala</i>	5009	5077	69		
<i>tRNA-Asn</i>	5079	5149	71		
<u><i>O_L</i></u>	5150	5180	31		
<i>tRNA-Cys</i>	5181	5248	68		
<i>tRNA-Tyr</i>	5249	5315	67		
<i>CO1</i>	5317	6861	1545	ATG	TAA
<i>tRNA-Ser^(UCN)</i>	6859	6927	69		
<i>tRNA-Asp</i>	6931	6998	68		
<i>CO2</i>	7000	7683	684	ATG	T--
<i>tRNA-Lys</i>	7687	7750	64		
<i>ATPase 8</i>	7752	7955	204	ATG	TAA
<i>ATPase 6</i>	7913	8593	681	ATG	TA-
<i>CO3</i>	8593	9377	785	ATG	T--
<i>tRNA-Gly</i>	9377	9444	68		
<i>ND3</i>	9445	9792	348	ATA	TAG
<i>tRNA-Arg</i>	9794	9861	68		
<i>ND4L</i>	9864	10160	297	ATG	TAA
<i>ND4</i>	10154	11531	1378	ATG	T--
<i>tRNA-His</i>	11532	11599	68		
<i>tRNA-Ser^(AGY)</i>	11600	11658	59		
<i>tRNA-Leu^(CUN)</i>	11658	11728	71		
<i>ND5</i>	11729	13558	1830	ATA	TAA
<u><i>ND6</i></u>	13536	14054	519	ATG	TAA
<i>tRNA-Glu</i>	14055	14123	69		
<i>CYTB</i>	14127	15270	1144	ATG	T--
<i>tRNA-Thr</i>	15271	15338	68		
<i>tRNA-Pro</i>	15339	15404	66		
Control region	15405	16261	857		

^a*ND1-6*, NADH dehydrogenase subunit 1-6 gene; *CO1-3*, cytochrome c oxidase subunit 1-3 gene; *ATPase 6* and *8*, ATPase subunit 6 and 8 gene; *CYTB*, cytochrome b gene; *O_L*, L-strand replication origin; CR, control region. Underlines indicate a gene encoded on the L-strand.

Table 1.5. Location of feature in the mt genome of *A. peninsulae*

Feature ^a	Position		Size (bp)	Codon	
	5'	3'		Initiaton	Termination
<i>tRNA-Phe</i>	1	67	67		
<i>12S rRNA</i>	68	1023	956		
<i>tRNA-Val</i>	1024	1092	69		
<i>16S rRNA</i>	1093	2663	1571		
<i>tRNA-Leu^(UUR)</i>	2664	2738	75		
<i>ND1</i>	2739	3695	955	GTG	TAG
<i>tRNA-Ile</i>	3694	3762	69		
<i>tRNA-Gln</i>	3760	3830	71		
<i>tRNA-Met</i>	3834	3902	69		
<i>ND2</i>	3903	4940	1038	ATC	TAG
<i>tRNA-Trp</i>	4939	5006	68		
<i>tRNA-Ala</i>	5008	5076	69		
<i>tRNA-Asn</i>	5078	5148	71		
<u>O_L</u>	5149	5179	31		
<i>tRNA-Cys</i>	5180	5245	66		
<i>tRNA-Tyr</i>	5246	5312	67		
<i>CO1</i>	5314	6858	1545	ATG	TAA
<i>tRNA-Ser^(UCN)</i>	6856	6924	69		
<i>tRNA-Asp</i>	6928	6995	68		
<i>CO2</i>	6997	7680	684	ATG	TAG
<i>tRNA-Lys</i>	7684	7747	64		
<i>ATPase 8</i>	7749	7952	204	ATG	TAA
<i>ATPase 6</i>	7910	8590	681	ATG	TAA
<i>CO3</i>	8590	9374	785	ATG	TA-
<i>tRNA-Gly</i>	9374	9441	68		
<i>ND3</i>	9442	9789	348	ATA	TAA
<i>tRNA-Arg</i>	9791	9858	68		
<i>ND4L</i>	9861	10157	297	ATG	TAA
<i>ND4</i>	10151	11528	1378	ATG	T--
<i>tRNA-His</i>	11529	11595	67		
<i>tRNA-Ser^(AGY)</i>	11596	11654	59		
<i>tRNA-Leu^(CUN)</i>	11654	11724	71		
<i>ND5</i>	11725	13554	1830	ATA	TAA
<u><i>ND6</i></u>	13532	14050	519	ATG	TAA
<i>tRNA-Glu</i>	14051	14119	69		
<i>CYTB</i>	14123	15266	1144	ATG	T--
<i>tRNA-Thr</i>	15267	15333	67		
<i>tRNA-Pro</i>	15334	15400	67		
Control region	15401	16266	866		

^a*ND1-6*, NADH dehydrogenase subunit 1-6 gene; *CO1-3*, cytochrome *c* oxidase subunit 1-3 gene; *ATPase 6* and *8*, ATPase subunit 6 and 8 gene; *CYTB*, cytochrome *b* gene; O_L, L-strand replication origin; CR, control region. Underlines indicate a gene encoded on the L-strand.

Overall, the nucleotide frequencies of the analyzed mt protein-encoding genes were similar (Table 1.6). In Korean *Apodemus*, the nucleotide frequencies of *A. agrarius* and *A. chejuensis* were similar. However, those of *A. peninsulae* were different. The protein-coding genes of *A. peninsulae* had the greatest adenine and thymine content. In particular, *A. peninsulae* had the most thymine and adenine in the second and third position, respectively, in Korean *Apodemus*. Adenine frequency in the first position was relatively low but guanine in same position was much more frequent. In the second position, *A. peninsulae* had the most thymine. The third codon position in *A. peninsulae* was the most variable. Guanine frequency in the third position was lowest in *A. peninsulae*.

Codon usage based on mt protein-coding genes showed that the patterns of *A. agrarius* and *A. chejuensis* were similar, while those of *A. peninsulae* slightly differed (Table 1.7-9). With respect to the mt protein-coding genes, the most frequent codon was CUA for leucine (252 times and 254 times, respectively) in *A. agrarius* and *A. chejuensis* and AUU for isoleucine (277 times) in *A. peninsulae* whereas the most rare codons were ACG for threonine and CGG for arginine in *A. agrarius* (2 times, respectively) and UCG for serine and AAG for lysine in *A. chejuensis* (2 times, respectively) and CGG for arginine (0 time) in *A. peninsulae*. Of all the codon groups that encode a single amino acid, leucine was used most frequently (600 times/15.8%, 598 times/15.7%, and 590 times/15.5% in *A. agrarius*, *A. chejuensis*, and *A. peninsulae*, respectively). Conversely, the least used amino acid was the cysteine group [29 times/0.76% in *A. agrarius* and *A. chejuensis*, and 28 times/0.74% in *A. peninsulae*].

Table 1.6. Nucleotide frequencies of 13 protein-coding genes

Position Num.	<i>A. agrarius</i>				<i>A. chejuensis</i>				<i>A. peninsulae</i>			
	T	C	A	G	T	C	A	G	T	C	A	G
1st	28.4	24.0	42.9	4.7	28.5	23.8	42.9	4.8	24.8	22.6	32.5	20.1
2nd	24.4	23.9	31.7	20.0	24.2	24.0	31.9	19.9	42.3	25.7	20.0	12.0
3rd	41.9	25.4	20.7	12.0	41.9	25.3	20.8	12.0	30.0	21.3	45.5	3.2
Total	31.5	24.5	31.8	12.2	31.5	24.4	31.9	12.2	32.3	23.2	32.7	11.8

Table 1.7. Codon usage based on 13 protein-coding genes of *A. agrarius*

AA ^a	Codon	N ^b	AA	Codon	N	AA	Codon	N	AA	Codon	N
Phe	UUU	148	Ser	UCU	51	Tyr	UAU	74	Cys	UGU	9
	UUC	97		UCC	54		UAC	57		UGC	20
Leu	UUA	128	Pro	UCA	136	Stop	UAA	11	Trp	UGA	93
	UUG	9		UCG	3		UAG	2		UGG	13
	CUU	110		CCU	43	His	CAU	42	Arg	CGU	10
	CUC	80		CCC	30		CAC	54		CGC	7
	CUA	252		CCA	121	Gln	CAA	78		CGA	44
	CUG	21		CCG	5		CAG	5		CGG	2
	Ile	AUU		220	Thr	ACU	80	Asn	AAU	81	Ser
AUC		148	ACC	61		AAC	88		AGC	35	
Met	AUA	197	Ala	ACA	159	Lys	AAA	94	Stop	AGA	0
	AUG	35		ACG	2		AAG	3		AGG	0
Val	GUU	50	Ala	GCU	69	Asp	GAU	31	Gly	GGU	43
	GUC	27		GCC	65		GAC	38		GGC	41
	GUA	84		GCA	93	Glu	GAA	78		GGA	101
	GUG	8		GCG	4		GAG	15		GGG	26

^aAmino acid^bFrequency of codon

Table 1.8. Codon usage based on 13 protein-coding genes of *A. chejuensis*

AA ^a	Codon	N ^b	AA	Codon	N	AA	Codon	N	AA	Codon	N
Phe	UUU	149	Ser	UCU	49	Tyr	UAU	80	Cys	UGU	9
	UUC	97		UCC	52		UAC	51		UGC	20
Leu	UUA	120	Pro	UCA	139	Stop	UAA	12	Trp	UGA	96
	UUG	13		UCG	2		UAG	1		UGG	8
	CUU	106		CCU	44	His	CAU	44	Arg	CGU	11
	CUC	82		CCC	29		CAC	52		CGC	6
	CUA	254		CCA	123	Gln	CAA	78		CGA	41
	CUG	23		CCG	4		CAG	5		CGG	6
	Ile	AUU		225	Thr	ACU	83	Asn	AAU	78	Ser
AUC		148	ACC	58		AAC	89		AGC	37	
Met	AUA	199	Ala	ACA	156	Lys	AAA	97	Stop	AGA	0
	AUG	33		ACG	4		AAG	2		AGG	0
Val	GUU	46	Ala	GCU	66	Asp	GAU	32	Gly	GGU	42
	GUC	28		GCC	67		GAC	39		GGC	39
	GUA	86		GCA	91	Glu	GAA	73		GGA	104
	GUG	8		GCG	5		GAG	20		GGG	24

^aAmino acid^bFrequency of codon

Table 1.9. Codon usage based on 13 protein-coding genes of *A. peninsulae*

AA ^a	Codon	N ^b	AA	Codon	N	AA	Codon	N	AA	Codon	N		
Phe	UUU	144	Ser	UCU	62	Tyr	UAU	76	Cys	UGU	10		
	UUC	97		UCC	40		UAC	57		UGC	18		
Leu	UUA	167	Pro	UCA	143	Stop	UAA	12	Trp	UGA	100		
	UUG	11		UCG	1		UAG	1		UGG	4		
	CUU	111		CCU	43		His	CAU		50	Arg	CGU	9
	CUC	57		CCC	25		CAC	50		CGC	12		
Ile	CUA	228	Thr	CCA	131	Asn	CAA	82	Ser	CGA	44		
	CUG	16		CCG	2		CAG	1		CGG	0		
	AUU	277		ACU	78		AAU	68		AGU	17		
	AUC	96		ACC	66		AAC	94		AGC	33		
Met	AUA	206	Ala	ACA	161	Lys	AAA	95	Stop	AGA	0		
	AUG	33		ACG	3		AAG	8		AGG	0		
Val	GUU	45	Glu	GCU	66	Asp	GAU	39	Gly	GGU	42		
	GUC	20		GCC	63		GAC	36		GGC	43		
	GUA	86		GCA	90		GAA	81		GGA	111		
	GUG	15		GCG	3		GAG	11		GGG	15		

^aAmino acid^bFrequency of codon

1.3.3. rRNA and tRNA encoding genes

The *12S rRNA* gene was 955, 955, and 956 bp long in *Apodemus agrarius*, *A. chejuensis*, and *A. peninsulae*, respectively, and the *16S rRNA* gene was 1,572, 1,573, and 1,571 bp long, respectively. As in other vertebrates, including Rodentia, these genes were located between the *tRNA-Phe* and *tRNA-Leu^(UUR)* genes and are separated by the *tRNA-Val* gene (Table 1.3-5).

The mt genome of Korean *Apodemus* contained 22 tRNA genes dispersed between rDNAs and protein-coding genes, as in other rodents (Fig. 1.2-4 and Table 1.3-5). I identified 2 forms of *tRNA-Leu* (UUR and CUN) and *tRNA-Ser* (UCN and AGY; Fig. 1.2-4 and Table 1.3-5). The 3 tRNA clusters, IQM (isoleucine, glutamine, and methionine), WANCY (tryptophan, alanine, asparagine, cysteine, and tyrosine), and HSL (histidine, serine, and leucine), were conserved in Korean *Apodemus*, as in the typical Rodentia mt genome. The putative structure of *tRNA-Ser^(AGY)* was compared to the mt tRNAs of Rodentia. The abnormal structure of *tRNA-Ser^(AGY)* was identified in the dihydrouridine (DHU) arm in the mt genome sequences of 3 Korean *Apodemus* (Fig. 1.5).

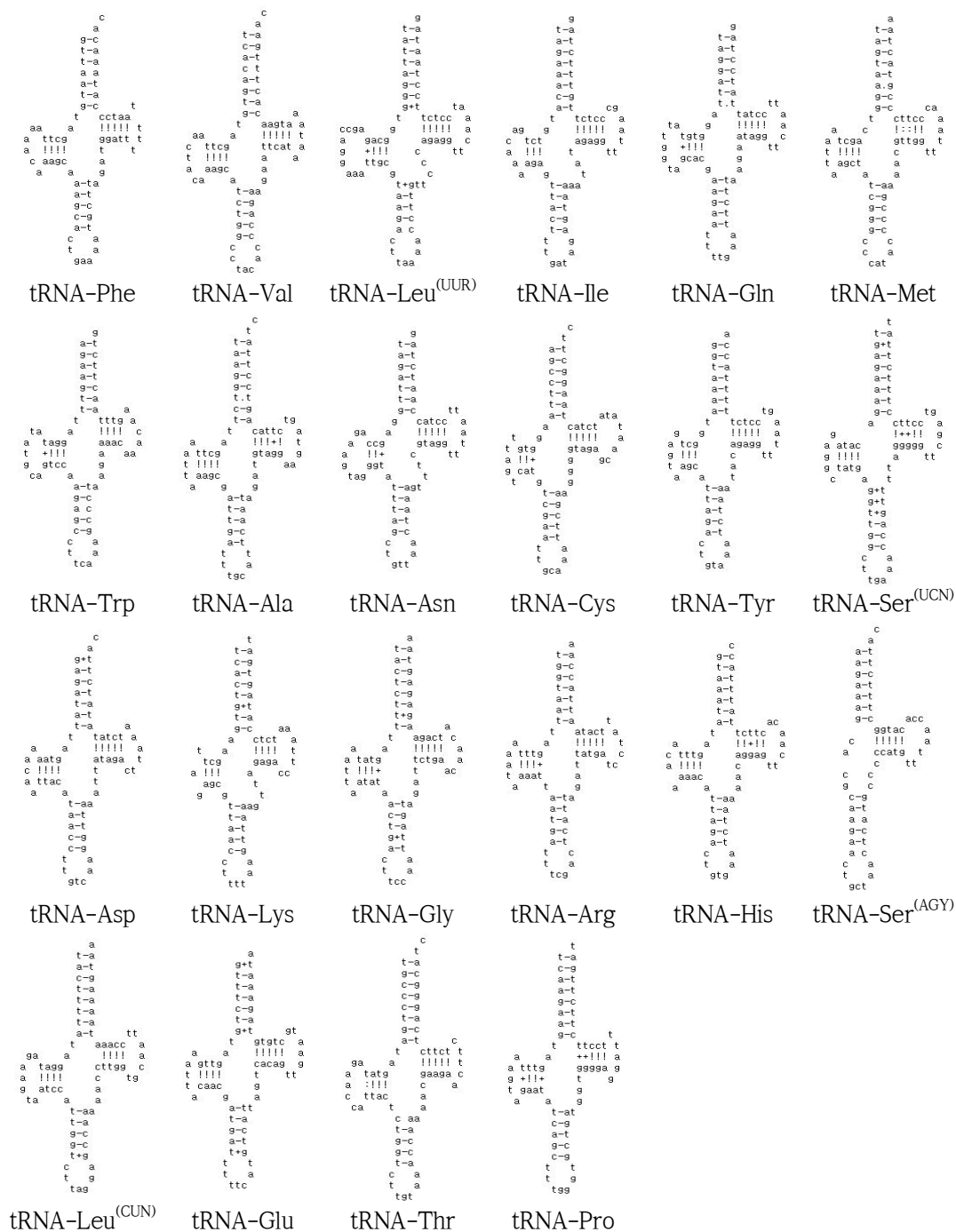


Figure 1.2. The putative structures of the 22 tRNAs of *Apodemus agrarius*, inferred from nucleotide sequences of the tRNA genes. The tRNAs are labeled with the abbreviations of their corresponding amino acids.

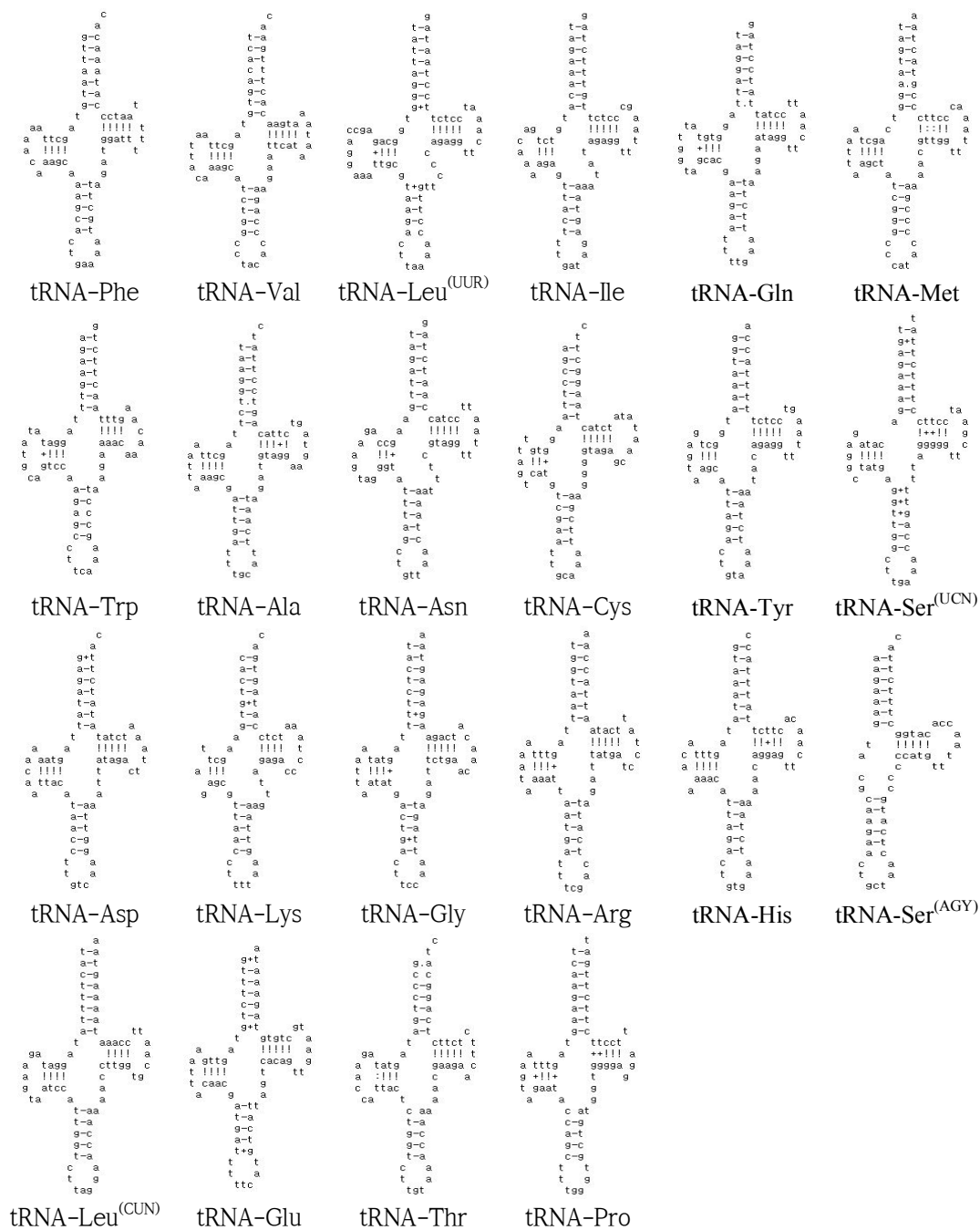


Figure 1.3. The putative structures of the 22 tRNAs of *Apodemus chejuensis*, inferred from nucleotide sequences of the tRNA genes. The tRNAs are labeled with the abbreviations of their corresponding amino acids.

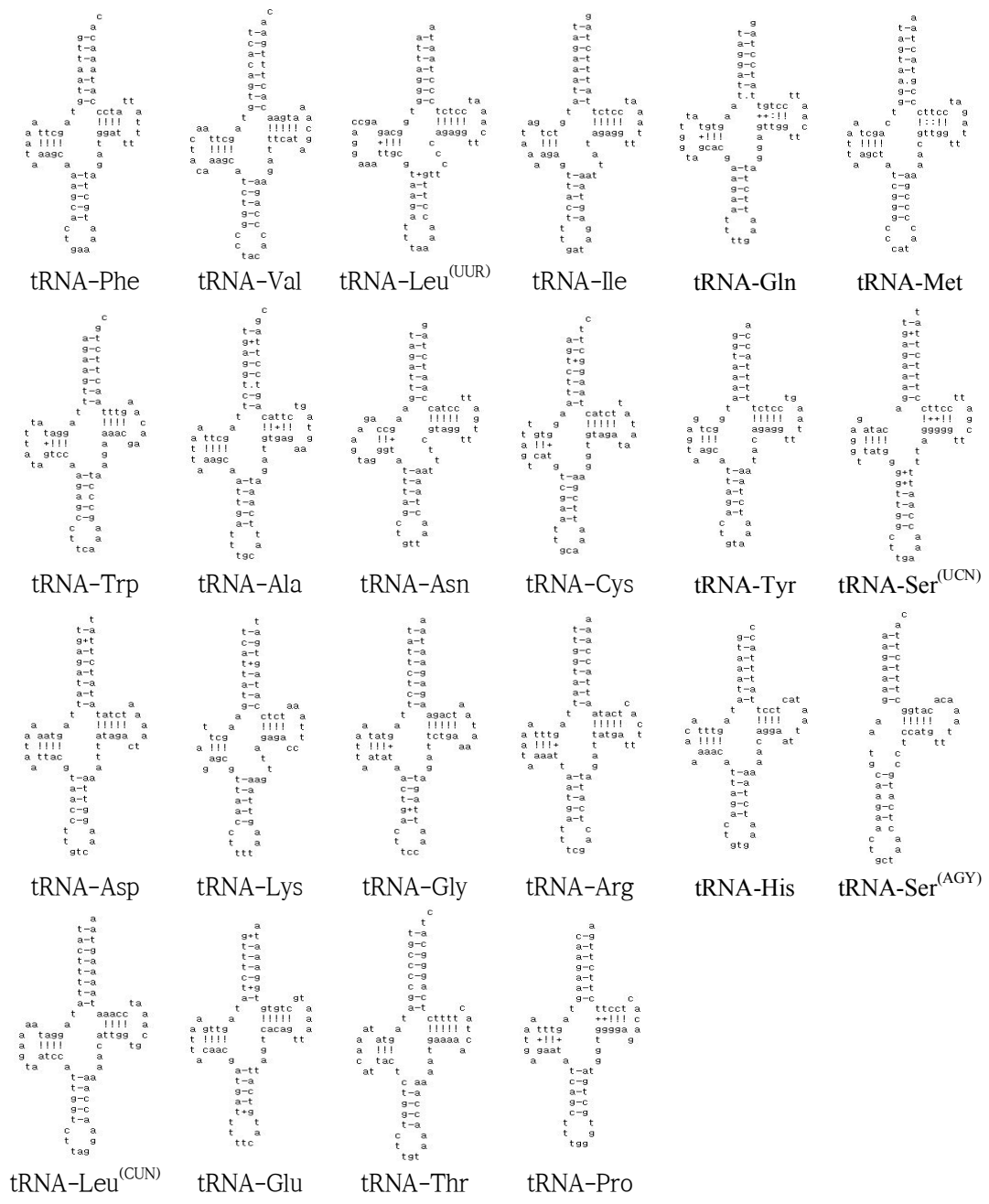
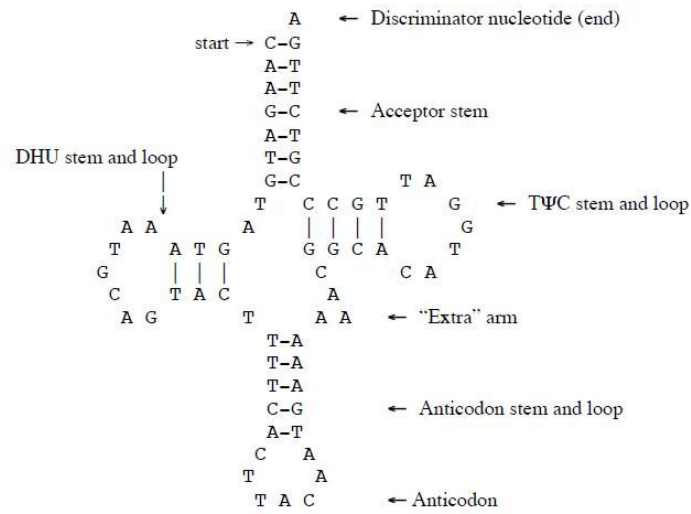
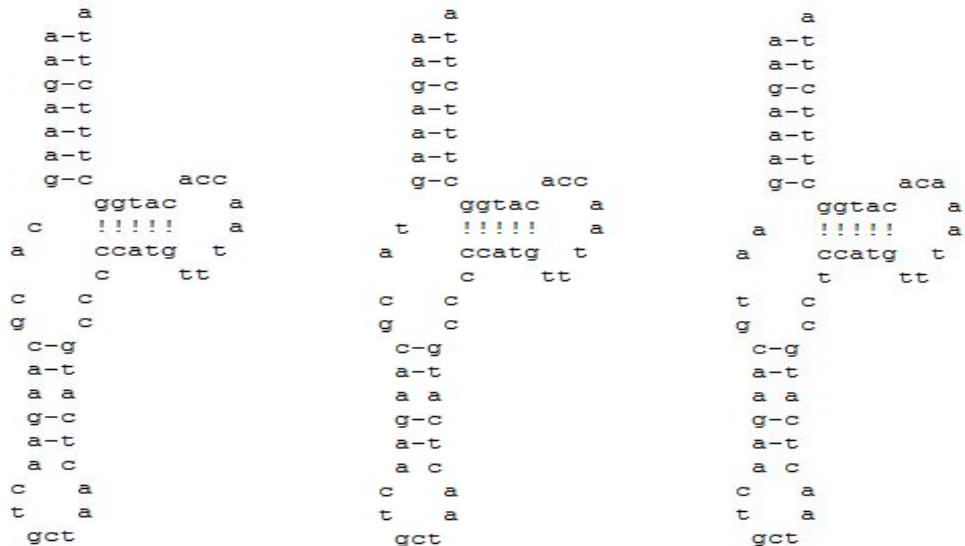


Figure 1.4. The putative structures of the 22 tRNAs of *Apodemus peninsulae*, inferred from nucleotide sequences of the tRNA genes. The tRNAs are labeled with the abbreviations of their corresponding amino acids.



(A)



(B)

(C)

(D)

Figure 1.5. The putative structure of the tRNA-Ser in Korean *Apodemus*. The abnormal structure of *tRNA-Ser*^(AGY) was identified in DHU arm. (A) shows secondary structure of general mt tRNAs (Wyman and Boore, 2003). General mt tRNAs have normal DHU arm with stem-loop structure. (B), (C), and (D) has abnormal structures with four nucleotides in DHU arm of mt *tRNA-Ser*^(AGY) of *A. agrarius*, *A. chejuensis*, and *A. peninsulae*, respectively.

1.3.4. Control region

Generally, the mt genomes of mammals have 2 noncoding regions: O_L and CR, which is also called the D-loop. The O_L of Korean *Apodemus* was 31 bp long and was located between *tRNA-Asn* and *tRNA-Cys* (Table 1.3-5), consistent with other mammals. The stem-loop region, conserved motif (5' -TAAGG-3'), and the L-strand replication initiation site (GGC) are presented in Fig. 1.6.

The CR of *A. agrarius*, *A. chejuensis*, and *A. peninsulae* spanned 854 bp, 857 bp, and 866 bp, respectively, and was positioned between the *tRNA-Pro* and *tRNA-Phe* genes (Table 1.3-5). The CR contains several conserved sequences (Fig. 1.7): the termination-associated sequence (TAS) and conserved sequence blocks (CSBs). These elements show some sequence variation but are generally conserved among Rodentia.

1.3.5. Phylogenetic relationships

Generally, the ML tree shows that phylogenetic relationships are highly supported by sequences of 13 protein-coding genes of Rodentia mt (Fig. 1.8). Muridae was grouped with Cricetidae as sister taxon. *Rattus* was divided into distinct two clades, and Korean *Apodemus* and *Mus* derived from common ancestors. The branch support value of *Anomalurus+Cavia+Thryonomys* was lower than 50%. *Microtus levis* and *Proedromys* sp. was grouped with very low support value (10%). The monophyly of Muroidea (Muridae, Cricetidae, and Spalacidae) was well supported with 100% bootstrap value.

A. agrarius
A. chejuensis
A. peninsulae

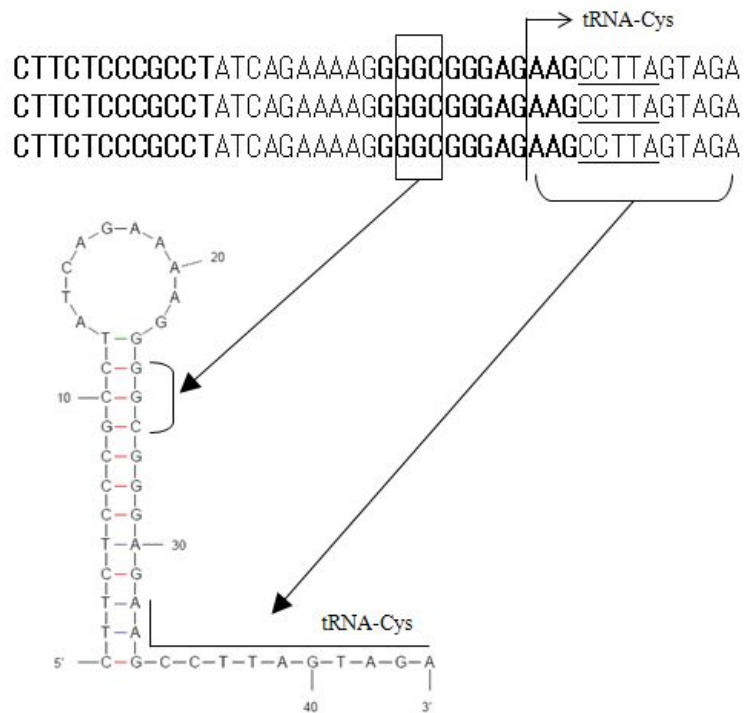


Figure 1.6. Common features of L-strand replication origin of Korean *Apodemus*. The O_L sequences of Korean *Apodemus* mt DNA was identical. The O_L has the potential to form a stable stem-loop structure. The initiation site "GGC" for L-strand replication is located on stem region. Bold and underlines indicate stem region and vertebrate conserved motif, respectively.

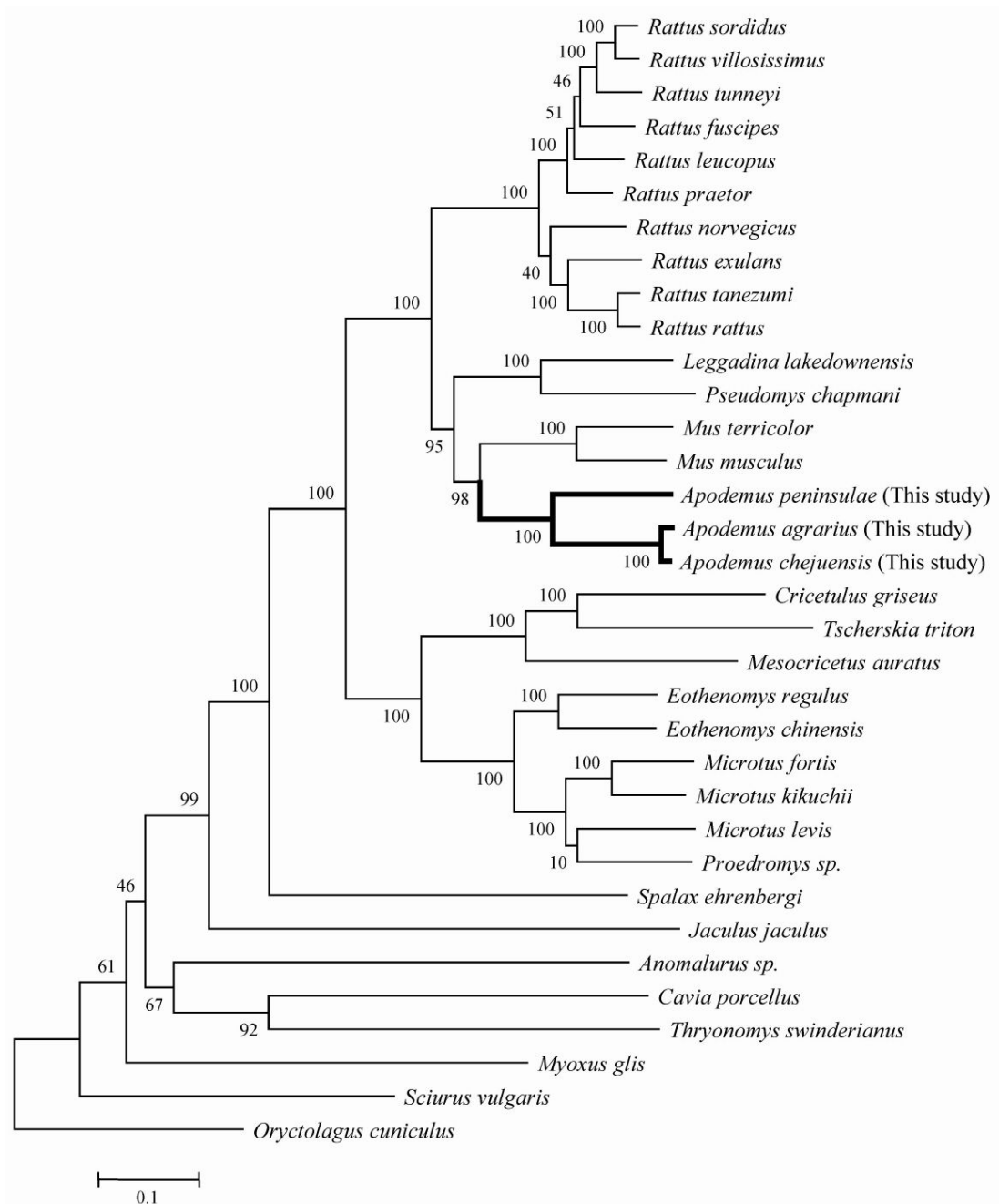


Figure 1.8. Phylogenetic tree of Korean *Apodemus*. The ML tree (lnL = -168970.343) was reconstructed using GTR+I+G model. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

1.4. DISCUSSION

1.4.1. Comparative analysis of mitochondrial genome sequences

Korean *Apodemus* contain relatively small mt genomes in the order Rodentia, according to the GenBank database (Table 1.2). The higher animals exhibit significantly less variation in mt genome size and tend to have smaller mt genomes than lower animals (Rand, 1993). Rand (1993) thought these patterns are due to differential metabolic rates and reduced mutation pressures. This theory would suggest that Korean *Apodemus* spp. might have evolved late in the evolutionary history of Rodentia.

General features of mt genome of the Korean *Apodemus* are similar to those of other Rodentia. The gene rearrangements of mt genomes were reported in various vertebrates such as marsupials (Pääbo *et al.*, 1991), birds (Bensch and Härlid. 2000), reptiles (Yan *et al.*, 2008), fishes (Mabuchi *et al.*, 2004), and amphibians (Mueller and Boore, 2005). However, gene rearrangements have not been identified in mt genomes of Rodentia, including *Apodemus* spp. This indicates that mt genomes of Rodentia are very stable and highly conserved.

The mt protein-coding genes of Korean *Apodemus* used various initiation and termination codons. The mt protein-coding genes of *A. agrarius* and *A. chejuensis* used 3 types of initiation codons (ATG, ATA, and GTG) and those of *A. peninsulae* used 4 types of initiation codons (ATG, ATA, ATC and GTG). ATG is the most common initiation codon, but various vertebrate mt genes use non-canonical starts (Mouchaty *et al.*, 2001; Mabuchi *et al.*, 2004; Mueller and Boore, 2005; Horner *et al.*, 2007). These alternative

initiation codons are translated as methionine or formylmethionine (Touriol *et al.*, 2003). Five types of initiation codons (ATG, GTG, ATC, ATA, and ATT) have been identified in the Rodentia. ATT is not used in Korean *Apodemus*.

The mt protein-coding genes of *A. agrarius*, *A. chejuensis*, and *A. peninsulae* used 2 (TAA and T), 4 (TAA, TAG, TA, and T), and 3 (TAA, TAG, and T) types of termination codons, respectively. Korean *Apodemus* mt genes also used incomplete termination codons such as TA and T, as occurs in other vertebrates (Oh *et al.*, 2007a; 2007b; Yan *et al.*, 2008; Oh *et al.*, 2010). These incomplete termination codons are completed via post-transcriptional polyadenylation (Ojala *et al.*, 1981).

Overall, the nucleotide frequencies of the analyzed mt protein-coding genes were slightly different in Korean *Apodemus*. Especially, guanine frequency at the first codon position in *A. peninsulae* mt was remarkably higher than in *A. agrarius* and *A. chejuensis*, while guanine frequency at the third codon position of *A. peninsulae* mt was much lower. The relatively low frequencies of guanine at the first codon position of *A. agrarius* and *A. chejuensis* have not been reported in Muridae. For instance, guanine at the first codon position of *Rattus norvegicus* mt genes were relatively more frequent at 19% (Gadaleta *et al.*, 1989). In the protein coding genes, the proportion of guanine at the third codon position is relatively low, as previously reported (Reyes *et al.*, 1998; Inoue *et al.*, 2000; Bayona-Bafaluy *et al.*, 2003). The variation in base composition at the 3 codon positions is related to gene expressivity (Gutiérrez *et al.*, 1996). Accordingly, the differential nucleotide preference for codon position may affect mt gene expression across Korean *Apodemus*.

The codon usage pattern was similar to those of other rodents (Lin *et al.*, 2002; Horner *et al.*, 2007; Partridge *et al.*, 2007). Codon usage also might be related to translation of mt genes. Jia and Higgs (2008) discovered very little evidence for translational accuracy selection in mt genes. They suggested that the selection for translational efficiency might lead to codon preferences.

Moreover, Knight *et al.* (2001) suggested low codon frequencies can be related to codon reassignment, and highly expressed genes have a strong preference for a subset of codons, while low-expression genes have a more uniform pattern of codon usage (Ikemura, 1981; Gouy and Gautier, 1982). Therefore, some frequently used codons such as CUA (*tRNA-Leu*), AUC (*tRNA-Ile*), AUA (*tRNA-Met*), etc., of Korean *Apodemus* mt may be associated with highly expressed genes while infrequently used codons like UCG (*tRNA-Ser*), ACG (*tRNA-Thr*), CGG (*tRNA-Arg*), etc., might be related to rarely expressed genes.

The mt genomes of Korean *Apodemus* contain 22 tRNA genes interspersed between rRNAs and protein-coding genes. The general features of the mt tRNA genes were identical to those found in other rodents (Lin *et al.*, 2002; Horner *et al.*, 2007; Partridge *et al.*, 2007). The tRNA genes ranged from 59 to 75 nucleotides.

The *tRNA-Ser*^(AGY) found in the mt genomes of Korean *Apodemus* had no discernible DHU stem, similar to other rodents (Gadaleta *et al.*, 1989; Horner *et al.*, 2007; Partridge *et al.*, 2007). The *tRNA-Ser*^(AGY) lacking the DHU stem works in the mt translation system but has low translational activity in an *in vitro* translation system of bovine mitochondria (Hanada *et al.*, 2001). Accordingly, *tRNA-Ser*^(AGY) of Korean *Apodemus* also might work with some molecular disadvantages on ribosomes in the mt translation system.

The O_L sequence has the potential to form a stable stem-loop structure, and a conserved motif (5'-TAAGG-3') is located in the base of the stem within *tRNA-Cys*. This motif has been reported to exhibit some sequence variation in vertebrates such as *Lialis* (5'-ACCGG-3'), *Varanus* (5'-CCCTG-3'), *Bipes* (5'-GCCAG-3') (Macey *et al.*, 1997), *Halichoeres* (5'-GCCGG-3' Oh *et al.*, 2008), and *Hynobius* (5'-TCCGG-3' Oh *et al.*, 2007a). The pentanucleotide motif is conserved among vertebrates (Noack *et al.*, 1996). This sequence may play an important role in *in vitro* replication of

the L-strand (Hixson *et al.*, 1986). The O_L of Korean *Apodemus* mt contained an A-rich region in the loop. Other rodents are not A-rich or have only a single A in the loop in *Jaculus jaculus* (AJ416890) and *Mus musculus* (AY172335). These dissimilarities may result in slight differences in the structural recognition efficiency of mt DNA primase (Wong and Clayton, 1985; Hixson *et al.*, 1986). In addition, the stem region of O_L has the initiation site (GGC) for L-strand replication (Brennicke and Clayton, 1981), which was identical in all Korean *Apodemus*.

Within the CR, several conserved features were found such as TAS and CSB-1-3. The TAS element was identified upstream from the 3' end of the control region in mammals (Doda *et al.*, 1981; Mackay *et al.*, 1986). This is associated with the termination of D-loop strand synthesis (Broughton *et al.*, 2001). The TAS elements slightly differed between *A. agrarius*, *A. chejuensis*, and *A. peninsulae*. The CSB sequences were identical and/or slightly different. The CSB sequences are associated with mt DNA replication and transcription (Clayton, 1991; Shadel and Clayton, 1997). Accordingly, replication and/or transcription efficiency might differ between Korean *Apodemus* spp.

1.4.2. Phylogenetic relationships of Korean *Apodemus*

The ML analysis placed *Sciurus* and *Myoxus* at the base of the phylogenetic tree of rodents showing similar to the phylogenetic tree of major rodent clades inferred from six nuclear genes (Blanga-Kanfi *et al.*, 2009). The result showed that major rodent clades comprised a squirrel-related clade, a mouse-related clade, and Ctenohystrica suggested that the basal position of a squirrel-related clade might be the most suitable for rodent evolutionary scenario (Blanga-Kanfi *et al.*, 2009).

Rattus consisted of two distinct clades with 100% bootstrap support value,

which phylogeny corresponded to previous study (Robins *et al.*, 2008; Robins *et al.*, 2010). Robins *et al.* (2008) suggested that the divergence between *Rattus* and *Mus* was occurred at about 12 Mya, and the deepest divergence within *Rattus* at 3.5 Mya, however subsequent research using 16 mt genomes of *Rattus* proposed the most recent common ancestor of *Rattus* diverged at about 2.7 Mya (Robins *et al.*, 2010). The phylogenetic relationships among *Rattus* spp. is very complex so that ML analysis using mt genome sequences of *Rattus* could not fully resolved their relationships.

The phylogeny of two native Australian rodents, *Leggadina lakedownensis* and *Pseudomys chapmani* showed that these two species related to *Mus* as a sister group (Nilsson *et al.*, 2010). However, the phylogenetic analysis including *Apodemus* showed that *Mus* was the closest related to *Apodemus* (Fig. 1.8). The closed relationships between *Mus* and *Apodemus* was reported in previous researches (Suzuki *et al.*, 2000; Stepan *et al.*, 2005), however, those studies contained just a few species of Muridae.

To date, the mt genome sequences of Rodentia were reported only 33 species (37 mt genomes including below subspecies). Accordingly, the phylogenetic analysis of Rodentia using mt genome sequences could not clarify the their relationships. However, the sequence data of single gene have accumulated via continued efforts to explain the phylogenetic relationships of rodents using nuclear and mt gene sequences. Especially, DNA barcoding project contributed to accumulate DNA sequences of mt gene such as *COI* and *CYTB*. Accordingly, mt *COI* and/or *CYTB* gene sequences might be very useful for rodent phylogeny and species identification (Robins *et al.*, 2010).

Chapter 2

Possible historical migration routes of *Apodemus chejuensis*

2.1 INTRODUCTION

The genus *Apodemus* contains at least 20 species (Musser and Carleton, 2005), including several of the most common small rodent species, and is widely distributed in the Palearctic region.

The striped field mouse *A. agrarius* is distributed throughout nearly the entire Korean Peninsula and is the most common rodent species in Korea; another striped field mouse, *A. chejuensis*, is found only on Jeju Island off the southern coast of the Korean Peninsula (Won and Smith, 1999; Yoon *et al.*, 2004a). Generally, *A. agrarius* inhabits rural flat lands, rivers, and agricultural farmlands. They breed 3 - 4 times a year and have 2 - 8 offspring per litter (Yoon *et al.*, 1997). They eat seeds of rice, fruits, and grasses (Kang, 1971). *A. agrarius* are known to be a vector of the fatal epidemic hemorrhagic fever and tsutsugamushi disease (Lee and Lee, 1976; Lee *et al.*, 1991). The Jeju Striped Field mouse, *A. chejuensis*, is often found near bush and mountainous area populated by shrubbery. They are distributed from the coastal region to the high-altitude region of Mt. Halla. They breed 3 - 4 times per year and usually have 4 - 5 offspring per litter. Similar to *A. agrarius*, they eat rice, fruits, and grass plants.

Although Jones and Johnson (1965) originally classified *A. chejuensis* as one (*A. agrarius chejuensis*) of the 4 subspecies of *A. agrarius*, subsequent molecular studies suggested that *A. chejuensis* is a distinct species (Han *et al.*, 1996; Koh *et al.*, 2000). The ability of offspring to hang to the mother's papilla, known as tenacious nipple attachment, is observed only in *A. chejuensis* (Oh and Mori, 1998b); the results of crossbreeding experiments led to the conclusion that *A. chejuensis* are separate biological species (Oh and Mori, 1998a). However, researchers are still performing molecular genetic

studies of *A. agrarius* and *A. chejuensis* (Koh *et al.*, 2000; Yoon *et al.*, 2004b), and the taxonomic status of *A. chejuensis* still remains a matter of debate (Musser and Carleton, 2005).

Molecular evolutionary and phylogenetic studies have been conducted by several researchers focused on Korean *Apodemus*. Koh and Yoo (1992) analyzed mitochondrial (mt) DNA polymorphisms of *A. agrarius* and *A. chejuensis* by using 8 restriction enzymes. Thirty-one fragments were recognized and 15 clones were identified in 15 specimens of *A. a. coreae* and 6 specimens of *A. a. chejuensis*. Thus, *A. agrarius* were grouped into 4 major subgroups, but the taxonomic status of the subgroup remained unclear. Koh and Yoo (1992) suggested that the *A. chejuensis* specimens are indicative of a distinct species. In 1996, Han *et al.* also conducted restriction fragment analysis by using 11 restriction enzymes with nuclear ribosomal RNA spacers and mt DNA and suggested that geographic conditions may preserve various mtDNA haplotypes. Since the 2000s, molecular studies of DNA sequences have been used to establish the biogeographic, phylogenetic, and evolutionary status of Korean *Apodemus*. The analysis of mt *cytochrome b* (*CYTB*) sequences of *A. agrarius* and *A. chejuensis*. from the Korean Peninsula, southern coastal islets, and Jeju Island showed significant genetic structuring by geographical area and suggested that the hypothesized oldest haplotypes were obtained by the Jeju population and *A. agrarius* might be derived from *A. chejuensis* (Yoon *et al.*, 2004b).

Suzuki *et al.* (2008) suggested that the European lineage is a likely descendant of the Asian lineage and some genetic elements from Asia contributed to the mt DNA found in Europe. Phylogenetic analyses of mt and nuclear gene sequences indicated that *A. agrarius* and *A. chevrieri* are the most closely related sister species (Suzuki *et al.*, 2003; Liu *et al.*, 2004). Moreover, *A. agrarius* diverged from *A. chevrieri* about 2.68 million years ago (Mya), during the first glacial age of the Quaternary (Liu *et al.*, 2004). The

most reasonable theory of the origin of Asian *Apodemus* was suggested by Xia (1984). He proposed that Asian *Apodemus* occurred in the Hengduan Mountain region. This region contains extreme topographic and climatic complexity (Zang, 2002) that led to species differentiation during the glacial period of the Quaternary (Liu *et al.*, 2004). If the ancestor of *A. agrarius* inhabitants in the Korean Peninsula originated from the Hengduan Mountains, they may have moved into the peninsula or the North Korea via the Yellow Sea basin during the glacial age (Fig. 2.1).

The mt *CYTB* gene is commonly located between 2 transfer RNA (tRNA) genes, *tRNA-Glu* and *tRNA-Thr*, in the mammalian mt genome. This gene is generally used in as a phylogenetic marker in vertebrate population genetics and phylogenetic inferences (Iwasa *et al.*, 2000; Koh *et al.*, 2001; Hodgkinson *et al.*, 2003; Yoon *et al.*, 2004b). The *CYTB* gene contains both conserved and variable codon positions and its molecular phylogenetic utility has been demonstrated in vertebrates (Irwin *et al.*, 1991; Graybeal, 1993; Zardoya and Doadrio, 1999). Therefore, this gene is suitable for investigations of animal phylogenetic status (Meyer, 1994; Lovejoy and de Araújo, 2000).

In 2009, Dubey *et al.* suggested that the *CYTB* pseudogene can lead to phylogenetic misinterpretations in Muridae. The existence of a nuclear copy of the *CYTB* pseudogene had been not widely recognized. The nuclear sequences can cause major problems in systematic analyses due to different selection pressures and rate of nucleotide substitutions. Therefore, such sequence analysis can lead to inaccurate interpretations (Zhang and Hewitt, 1996, 2003; Triant and DeWoody, 2007).

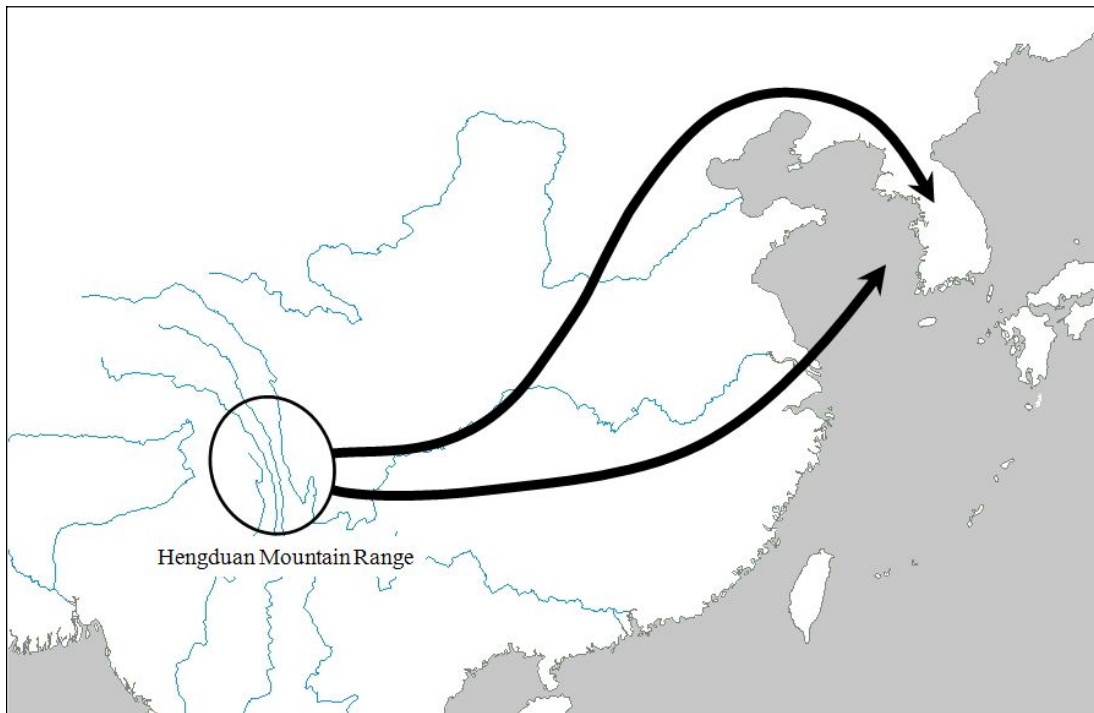


Figure 2.1. Origination and expansion of ancestral lineage of *A. agrarius* into the Korean Peninsula. An open circle is Hengduan Mountain Range considered as origination of Asian *Apodemus* (Xia, 1984). Arrows refer to possible migration routes of ancestor lineages of *A. agrarius*. The gray region represents the sea.

Despite the various data regarding the distribution, morphology, and molecular genetics of *A. agrarius* (Han *et al.*, 1996; Serizawa *et al.*, 2000; Liu *et al.*, 2004; Yoon *et al.*, 2004b; Suzuki *et al.*, 2008) and *A. chejuensis* (Han *et al.*, 1996; Koh *et al.*, 2000; Yoon *et al.*, 2004b), their historic migration routes into the Korean Peninsula and Jeju Island, respectively, are unclear. The biogeographic history of *A. chejuensis* has not been established. The most reasonable theory regarding *A. chejuensis*, based on recent genetic data, is that *A. agrarius* have descended from a population on Jeju Island (Yoon *et al.*, 2004b). Jeju Island is considered to have been a “refugium” during the last glacial age, so that the maternal lineages that occur on Jeju Island could have originated there or elsewhere.

Jeju Island is located southwest of the Korean Peninsula and consists mainly of basaltic and pyroclastic flows (Yoon *et al.*, 1995). The island formed via volcanic eruption about 2 Mya (Yoon, 1997). Yoon (1997) summarized the formation history of Jeju Island. After the formation of the first Jeju volcanic edifice (about 2 Mya), the volcanic edifice subsided, and marine transgression occurred about 1.8 Mya. The main volcanic edifice was formed about 0.89–0.47 Mya. At this time, marine regression occurred, and the older Jeju Island area became land. Subsequent volcanic activity led to the formation of various major topographic features of Jeju Island. About 0.07 Mya, the summit of Mt. Halla was formed, and the central part of the main volcanic edifice was uplifted by rising magma. During this time, marine transgression occurred again and the Jeju volcanic edifice became Jeju Island.

Sea level changes affect biogeographic distributions (Mercer and Roth, 2003; Won and Renner, 2006). During the glacial epoch, global sea levels fluctuated repeatedly (Voris, 2000; Zhong *et al.*, 2004), and researchers have suggested that some species migrated from the Korean Peninsula to Jeju Island or vice versa during the glacial epoch (Shim and Park, 1998). Therefore, integrated analysis of the distribution of terrestrial animals and

geological changes is important to understanding the biogeographic history.

This chapter has been attempted to explain the population history of *A. agrarius* and *A. chejuensis* and migration of *A. chejuensis* on the basis of the mt *CYTB* gene sequences of *A. agrarius* and *A. chejuensis* from the Korean Peninsula and Jeju Island.

2.2. MATERIALS AND METHODS

2.2.1. Sampling and DNA extraction

Forty-nine *Apodemus chejuensis* individuals and 24 *A. agrarius* individuals were collected in the wild at various sites on Jeju Island and the Korean Peninsula, respectively (Table 2.1 and Fig. 2.2). Among all tissue samples of *A. agrarius*, 13 samples (samples Agr11 - 23) from various locations were obtained from the Conservation Genome Resource Bank for Korean Wildlife.

Genomic DNA was extracted from small pieces of tail or muscle using a NucleoSpin Tissue Kit (Macherey-Nagel, Germany). First, small pieces of tail or muscle was placed in a microcentrifuge tube with buffer T1 and proteinase K solution and then vortexed. The tubes were incubated at 56°C until complete lysis was obtained. The lysed samples were vortexed, and buffer B3 was added. The tubes were vortexed and incubated at 70°C for 10 min. Ethanol was added and the samples were vortexed vigorously. For each sample, a NucleoSpin[®] Tissue Column was placed into a collection tube, and the samples applied to the column. The assembly was centrifuged for 1 min at 11,000g. The flow-through was discarded and the column returned to the collection tube. To wash the silica membrane, buffer BW and buffer B5 were added in turn. The column-collection tube assembly was centrifuged for 1 min at 11,000g after each buffer addition. The silica membrane was dried by centrifugation. Finally, the column was placed into a new 1.5 ml tube; prewarmed buffer BE was added, and then incubated at room temperature for 1 min. The tube was centrifuged for 1 min at 11,000g, and the extracted DNA was stored at -20°C.

2.2.2. PCR and sequencing

A 1.2 kbp fragment of the *CYTB* gene was amplified from the mt DNA by PCR using primers L14115 and H15288 (Martin *et al.*, 2000). PCR was carried out in 50 μ l reaction mixtures containing 5 μ l of 10 \times *Pfu-X* buffer (Solgent, Korea), 2 μ l of dNTP mix (2.5 mM of each dNTP), 1 μ l of each primer (10 μ M), 0.5 μ l of *Pfu-X* DNA polymerase (2.5 U/ μ l), 50-100 ng of genomic DNA, and sterile distilled water. Amplified products were examined by electrophoresis on 1% agarose gels, stained with EtBr, visualized under UV light, and purified using an AccuPrep PCR purification kit (Bioneer, Korea). The PCR products were directly sequenced using a BigDye Terminator Sequencing Kit (Applied Biosystems) on an ABI 3730xl automatic DNA sequencer with the newly designed inner primers ApoCYTB-F (5' -CAAATCTCCTCTCAGCCATCC-3') and ApoCYTB-R (5' -CTGTTGG GTTGTGTTGAGCCTG-3').

Table 2.1. Specimens and collection sources

Species	Specimen Num.	Collection site	Accession Num.	Haplotype Num.
<i>A. agrarius</i>	Agr1	Boreumdo, Incheon-si	HM034868	Hap1
	Agr2	Cheongdo, Gyeongsangbuk-do	HM034869	Hap2
	Agr3	Cheongdo, Gyeongsangbuk-do	HM034870	Hap3
	Agr4	Cheongdo, Gyeongsangbuk-do	HM034871	Hap2
	Agr5	Cheongdo, Gyeongsangbuk-do	HM034872	Hap4
	Agr6	Yangsan, Gyeongsangnam-do	HM034873	Hap5
	Agr7	Yangsan, Gyeongsangnam-do	HM034874	Hap6
	Agr8	Yangsan, Gyeongsangnam-do	HM034875	Hap5
	Agr9	Yangsan, Gyeongsangnam-do	HM034876	Hap7
	Agr10	Yangsan, Gyeongsangnam-do	HM034877	Hap8
	Agr11	Kanghwa, Incheon-si	HM034878	Hap9
	Agr12	Kanghwa, Incheon-si	HM034879	Hap10
	Agr13	Samcheok, Kangwondo	HM034880	Hap11
	Agr14	Macheon, Gyeongsangnam-do	HM034881	Hap12
	Agr15	Macheon, Gyeongsangnam-do	HM034882	Hap12
	Agr16	Seokmodo, Incheon-si	HM034883	Hap13
	Agr17	Seokmodo, Incheon-si	HM034884	Hap14
	Agr18	Seokmodo, Incheon-si	HM034885	Hap15
	Agr19	Seokmodo, Incheon-si	HM034886	Hap15
	Agr20	Seokmodo, Incheon-si	HM034887	Hap15
	Agr21	Hwacheon, Gangwon-do	HM034888	Hap16
	Agr22	Hwacheon, Gangwon-do	HM034889	Hap17
	Agr23	Hwacheon, Gangwon-do	HM034890	Hap18
	Agr24	Naju, Jeonranam-do	HM034891	Hap19
<i>A. chejuensis</i>	Che1	Hallim, Jeju-do	HM034892	Hap20
	Che2	Seongsan, Jeju-do	HM034893	Hap21
	Che3	Seongsan, Jeju-do	HM034894	Hap22
	Che4	Seongsan, Jeju-do	HM034895	Hap23
	Che5	Seongsan, Jeju-do	HM034896	Hap24
	Che6	Hallim, Jeju-do	HM034897	Hap25
	Che7	Jocheon, Jeju-do	HM034898	Hap26
	Che8	Andeok, Jeju-do	HM034899	Hap27
	Che9	Ara, Jeju-do	HM034900	Hap25
	Che10	Yonggang, Jeju-do	HM034901	Hap28
	Che11	Yonggang, Jeju-do	HM034902	Hap29
	Che12	Yonggang, Jeju-do	HM034903	Hap28
	Che13	Gujwa, Jeju-do	HM034904	Hap30
	Che14	Sanghyo, Jeju-do	HM034905	Hap31
	Che15	Sanghyo, Jeju-do	HM034906	Hap32
	Che16	Pyoseon, Jeju-do	HM034907	Hap33
	Che17	Pyoseon, Jeju-do	HM034908	Hap34
	Che18	Pyoseon, Jeju-do	HM034909	Hap27

Che19	Pyoseon, Jeju-do	HM034910	Hap27
Che20	Pyoseon, Jeju-do	HM034911	Hap33
Che21	Ara, Jeju-do	HM034912	Hap35
Che22	Ara, Jeju-do	HM034913	Hap36
Che23	Ara, Jeju-do	HM034914	Hap37
Che24	Ara, Jeju-do	HM034915	Hap22
Che25	Ara, Jeju-do	HM034916	Hap32
Che26	Ara, Jeju-do	HM034917	Hap35
Che27	Ara, Jeju-do	HM034918	Hap36
Che28	Yeongpyeong, Jeju-do	HM034919	Hap23
Che29	Yeongpyeong, Jeju-do	HM034920	Hap38
Che30	Ara, Jeju-do	HM034921	Hap39
Che31	Ara, Jeju-do	HM034922	Hap35
Che32	Ara, Jeju-do	HM034923	Hap37
Che33	Hallim, Jeju-do	HM034924	Hap20
Che34	Hallim, Jeju-do	HM034925	Hap28
Che35	Sanghyo, Jeju-do	HM034926	Hap31
Che36	Sanghyo, Jeju-do	HM034927	Hap27
Che37	Sanghyo, Jeju-do	HM034928	Hap40
Che38	Seongsan, Jeju-do	HM034929	Hap39
Che39	Pyoseon, Jeju-do	HM034930	Hap27
Che40	Pyoseon, Jeju-do	HM034931	Hap41
Che41	Pyoseon, Jeju-do	HM034932	Hap33
Che42	Pyoseon, Jeju-do	HM034933	Hap27
Che43	Pyoseon, Jeju-do	HM034934	Hap33
Che44	Pyoseon, Jeju-do	HM034935	Hap42
Che45	Pyoseon, Jeju-do	HM034936	Hap20
Che46	Pyoseon, Jeju-do	HM034937	Hap43
Che47	Pyoseon, Jeju-do	HM034938	Hap33
Che48	Seogwipo, Jeju-do	HM034939	Hap44
Che49	Seogwipo, Jeju-do	HM034940	Hap20

(Continued)

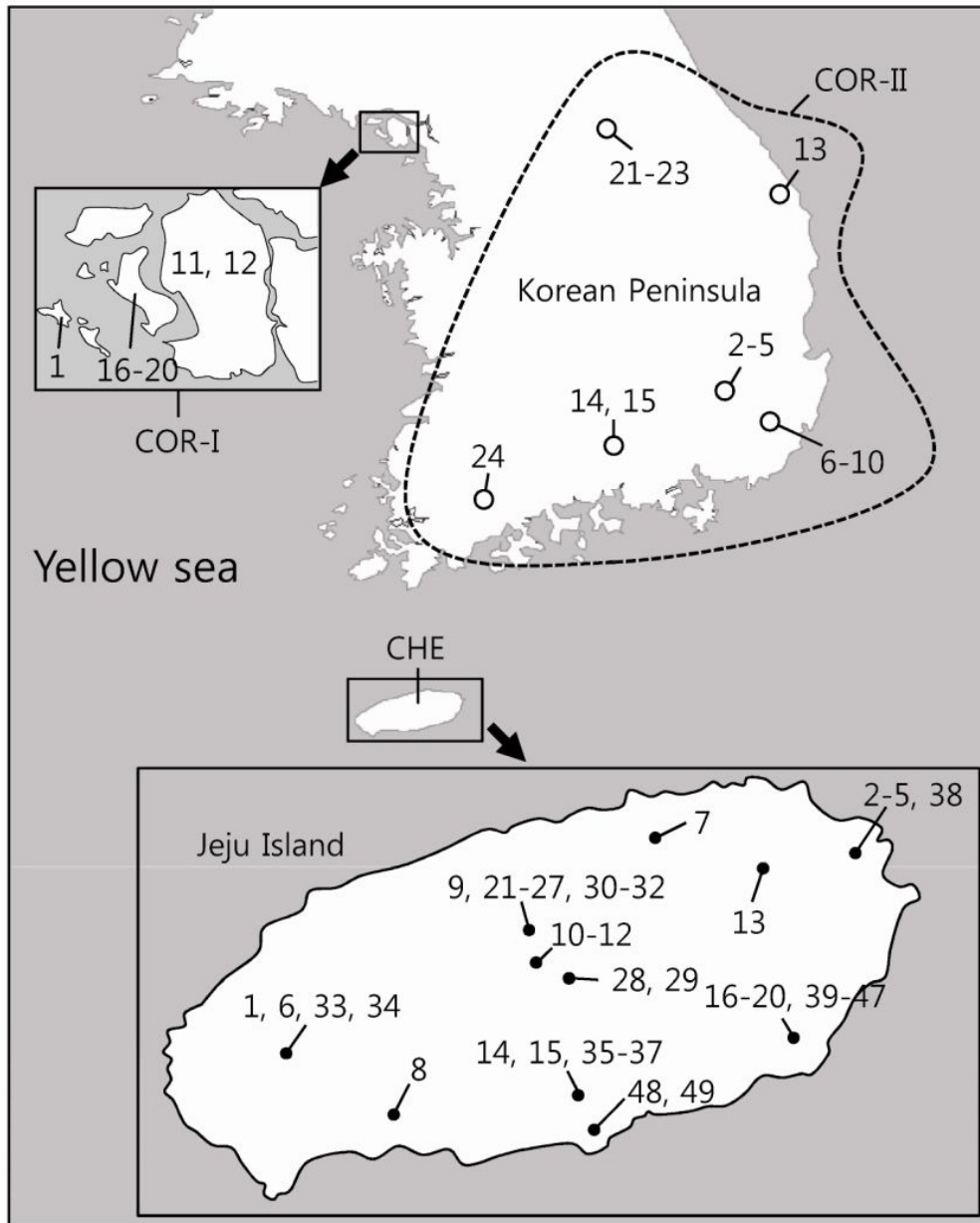


Figure 2.2. Collection sites of *A. agrarius* and *A. chejuensis*. Each specimen is indicated by a number on the map. Small open and closed circles indicate *A. agrarius* and *A. chejuensis* collection sites, respectively. Each local population of *A. agrarius* is indicated by a dotted line and a square. COR, Korean Peninsula; CHE, Jeju Island. Open squares indicate COR-I (western islet of the Korean Peninsula) collection sites. Details are shown in the magnified area. The gray region represents the sea.

2.2.3. Sequence analysis

The sequence analysis was conducted using 73 new *CYTB* sequences from *A. agrarius* and *A. chejuensis* (Table 2.1). Collection sites were classified for biogeographic analysis as follows: COR-I and COR-II, the western islets and inland of the Korean Peninsula, respectively; and CHE, Jeju Island (Fig. 2.2). All complete *CYTB* sequences were aligned using the ClustalX program (Thompson *et al.*, 1997) and detected single nucleotide polymorphism (SNP) sites. For *CYTB* haplotype analysis, MEGA5 software (Tamura *et al.*, 2011) and DNAsp 5.0 software (Librado and Rozas, 2009) were used to estimate the Kimura two-parameter genetic distances (Kimura, 1980), the number of haplotypes and polymorphic sites, the average number of nucleotide differences, and the level of haplotype and nucleotide diversity. Fu's F_s test was used to assess evidence of population expansion (Fu, 1997) by using Arlequin 3.5 software (Excoffier and Lischer, 2010). A mismatch distribution analysis using DNAsp 5.0 software was conducted to estimate population dynamics (Librado and Rozas, 2009).

To construct a phylogenetic median-joining network, Network 4.5.1.6 software (<http://www.fluxus-engineering.com>) was used with haplotype data.

2.2.4. Divergence time estimation

The divergence time between the *A. agrarius* and *A. chejuensis* was calculated according to the method of Liu *et al.* (2004). The maximum likelihood distances were calculated using the TN93+G model (Tamura and Nei, 1993), which was selected as the best model for nucleotide evolution, based on the Akaike Information Criterion (AIC; Akaike, 1973) using jModeltest software (Posada, 2008). To calibrate the divergence time, 2.68 Mya was employed as the divergence time between *A. agrarius* and *A. chevrieri* (Liu *et al.*, 2004). Because excessive rate heterogeneity affects the

molecular clock, the heterogeneity of evolutionary change along the different lineages was estimated. Using χ^2 tests (Felsenstein, 1988), I compared the likelihood trees created with and without an enforced molecular clock. Rate constancy among lineages was assumed for the tree with the enforced molecular clock.

2.3. RESULTS

2.3.1. Haplotype and demographic analysis

Nineteen *CYTB* sequence haplotypes were defined in 24 *A. agrarius* individuals, and 25 were defined in 49 *A. chejuensis* individuals (Table 2.1). The most frequent haplotype group was Hap27, which was shared by 6 *A. chejuensis* individuals. The most frequent haplotype group in the *A. agrarius* population was Hap15, which was shared by 3 individuals.

The variable sites of *CYTB* sequence haplotypes were estimated. In 1,140 sites of *CYTB* sequence, 105 sites were variable and 43 sites were singleton variable sites (Fig. 2.3). In *A. agrarius*, variable and singleton variable sequences were found in 56 sites and 33 sites, respectively, while 53 variable sites and 20 singleton variable sites were found in *A. chejuensis*. Among variable sites, 6 were completely different between the *A. agrarius* and *A. chejuensis*.

The genetic distances of *CYTB* haplotypes showed that Hap9 is closest to Hap42, while Hap7 is farthest from Hap36 (Table 2.2). Each haplotype of the COR-I population, excluding Hap10 (Agr12), were relatively closer to each other of COR-I than to haplotypes of the COR-II population. In *A. chejuensis*, all haplotypes were closer to each other than to haplotypes of *A. agrarius*.

0.072
0.040 0.072
0.094 0.106 0.094
0.072 0.083 0.072 0.019
0.105 0.116 0.105 0.093 0.071
0.072 0.083 0.072 0.061 0.040 0.050
0.083 0.094 0.083 0.029 0.010 0.082 0.050
0.072 0.019 0.072 0.106 0.083 0.116 0.083 0.094
0.072 0.083 0.072 0.106 0.083 0.116 0.083 0.094 0.083
0.061 0.094 0.083 0.142 0.118 0.153 0.118 0.130 0.118 0.118
0.072 0.106 0.072 0.130 0.106 0.141 0.106 0.118 0.106 0.106 0.118
0.050 0.083 0.010 0.106 0.083 0.116 0.083 0.094 0.083 0.083 0.094 0.083
0.116 0.128 0.116 0.105 0.082 0.010 0.060 0.093 0.128 0.128 0.166 0.153 0.128
0.072 0.083 0.072 0.106 0.083 0.116 0.083 0.094 0.083 0.019 0.118 0.106 0.083 0.128
0.061 0.072 0.061 0.094 0.072 0.105 0.072 0.083 0.072 0.029 0.106 0.094 0.072 0.116 0.029
0.072 0.083 0.072 0.106 0.083 0.116 0.083 0.094 0.083 0.019 0.118 0.106 0.083 0.128 0.019 0.029
0.092 0.104 0.092 0.081 0.060 0.070 0.039 0.070 0.104 0.104 0.139 0.127 0.104 0.081 0.104 0.092 0.104
0.072 0.083 0.072 0.106 0.083 0.116 0.083 0.094 0.083 0.019 0.118 0.106 0.083 0.128 0.019 0.029 0.019 0.104
0.050 0.061 0.050 0.083 0.061 0.093 0.061 0.072 0.061 0.040 0.094 0.083 0.061 0.105 0.040 0.029 0.040 0.081 0.040
0.083 0.094 0.083 0.072 0.050 0.060 0.029 0.061 0.094 0.094 0.130 0.118 0.094 0.071 0.094 0.083 0.094 0.049 0.094 0.072
0.010 0.061 0.050 0.106 0.083 0.116 0.083 0.094 0.083 0.083 0.050 0.083 0.061 0.128 0.083 0.072 0.083 0.104 0.083 0.061 0.094

(Continued)

The overall haplotype and nucleotide diversity values of *A. agrarius* were 0.978 and 0.007, respectively, while those of *A. chejuensis* were 0.962 and 0.007, respectively (Table 2.3). The haplotype diversity of the COR-I was the lowest of all biogeographic populations. The COR-II was the most diverse. Although this analyses included more CHE than COR individuals, the COR exhibited greater genetic diversity (Table 2.3).

Fu's F_s test showed that the COR-II and CHE had significantly negative F_s values (-4.462 and -5.703, respectively), whereas the COR-I had a positive value (0.613). Mismatch distribution analyses suggested population expansion in the COR-II and CHE but stable COR-I, respectively (Fig. 2.4).

Median-joining network data indicated that *A. chejuensis* is a clearly separate species (Fig. 2.5) containing 4 subgroups, but these subgroups did not correspond to local distribution patterns on Jeju Island. The Hap42 (Che44) haplotype was located in the closest linear descendant position of the common ancestor of *A. chejuensis*. In *A. agrarius*, all COR-II individuals shared a single maternal lineage, and most COR-I individuals were close to the ancestral position of the *A. agrarius*. In particular, Hap9 (Agr 11) was located in the ancestral lineage position of the *A. agrarius* (Fig. 2.5).

Table 2.3. Population analysis of *A. agrarius* and *A. chejuensis*

Subspecies	Population	S	N	<i>v</i>	<i>k</i>	<i>h</i> (SD)	π (SD)	<i>F_s</i>
	COR-I	8	6	23	8.321	0.893 (0.111)	0.007 (0.001)	0.613
<i>A. agrarius</i>	COR-II	16	13	37	6.067	0.975 (0.029)	0.005 (0.000)	-4.462*
	All	24	19	56	8.413	0.978 (0.019)	0.007 (0.001)	-6.207*
<i>A. chejuensis</i>	CHE	49	25	53	7.776	0.962 (0.012)	0.007 (0.000)	-5.703*

S, sample size; N, number of haplotypes; *v*, variable sites; *k*, average nucleotide difference; *h*, haplotype diversity; π , nucleotide diversity; SD, standard deviation; *F_s*, Fu's *F_s*. Asterisks mean a significant difference between populations ($P < 0.01$).

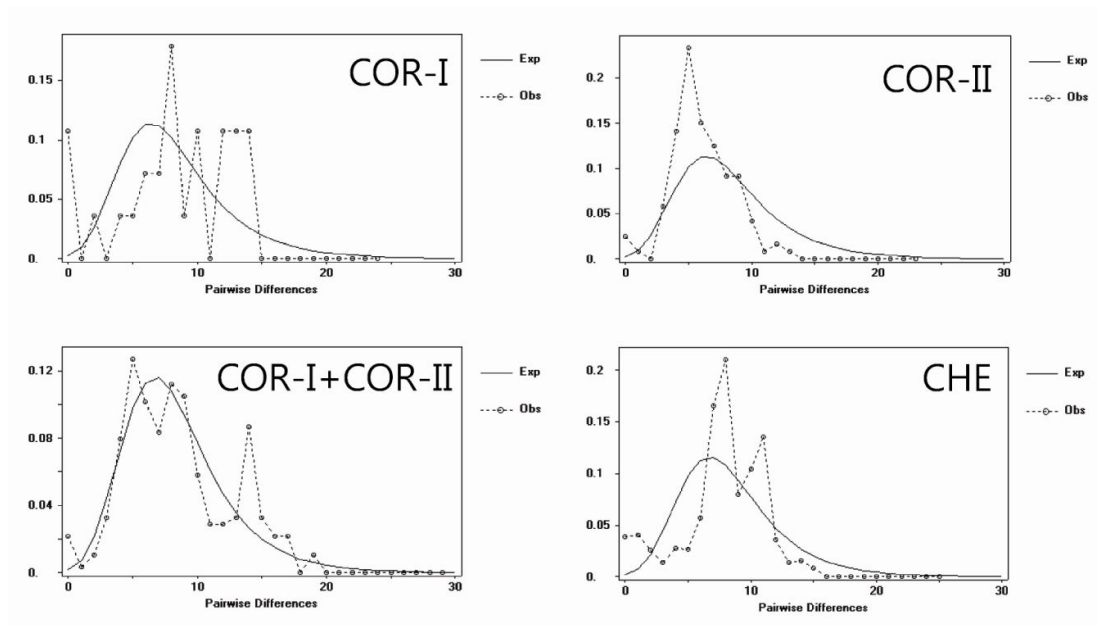


Figure 2.4. Mismatch distribution analysis in the *CYTB* gene sequences in *Apodemus* populations. The location of the COR-I, COR-II, and CHE is presented in Fig. 2.2. The dotted line with open circle is the observed distribution, the solid line is the expected in an expanding population.

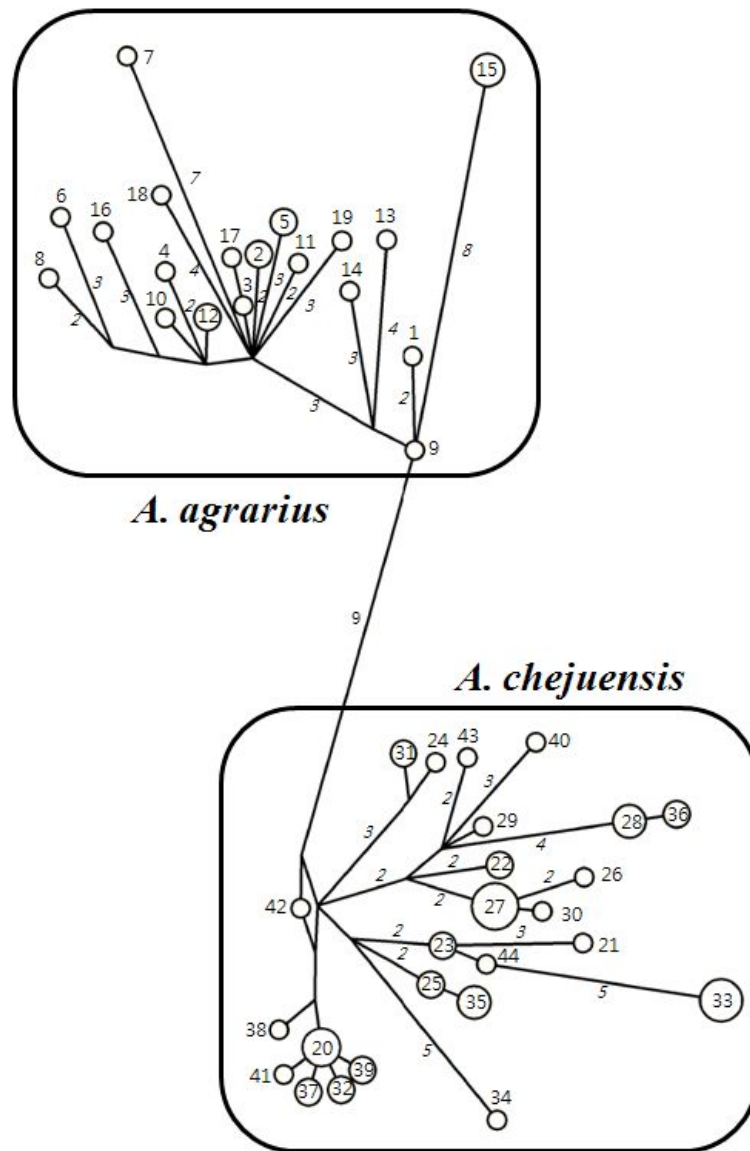


Figure 2.5. Median network tree inferred from *CYTB* haplotypes. Each circle indicates a haplotype (represented as a number), and the size of each circle is proportional to the number of individuals with that haplotype. Mutational steps between haplotypes are represented by a line. More than one mutational step is represented by italic numbers.

2.3.2. Divergence time

The maximum likelihood tree assuming a constant rate for the lineage ($\ln L = -2451.999$) presented the same topology as the tree without an enforced molecular clock ($\ln L = -2451.966$). The likelihood ratio test failed to reject the hypothesis of rate constancy between these 2 models ($\chi^2 = 0.033$, $df = 7$, $P > 0.05$). Thus, rate heterogeneity among lineages did not appear to affect the molecular clock calibration. The TN93+G distance between *Apodemus agrarius* and *A. chevrieri* was about 0.079, which corresponds to an evolution rate of 0.015 per million years, a result similar to that of Liu *et al.* (2004). Thus, I suggest that the split between *A. agrarius* and *A. chejuensis* occurred at approximately 0.3 Mya, in the late Pleistocene.

2.4. DISCUSSION

2.4.1. Haplotype analysis of *Apodemus agrarius* and *A. chejuensis*

The *CYTB* haplotype sequence alignment showed that the sequences of *Apodemus chejuensis* are distinct from those of *A. agrarius*. Among the sequence differences, 6 nucleotides could be used as a SNP marker to discriminate between *CYTB* haplotype sequences of *A. agrarius* and *A. chejuensis*. Distinguishing between *A. agrarius* and *A. chejuensis* in the wild is difficult because of their extremely similar morphology. Molecular markers inferred from sequence differences are often used in molecular systematics (van Embden *et al.*, 1993; Chu *et al.*, 2001; Gupta *et al.*, 2001). Consequently, those SNP sites can be used as a marker for species or subspecies identification from wild samples.

The genetic distances among haplotype sequences suggest that the CHE population is genetically close to the COR-I population. The geographical distance between the COR-I region and Jeju Island is farther than the distance from the southern region of the Korean peninsula to Jeju Island. Smaller genetic distances indicate a close genetic relationship (Machado *et al.*, 2000). Accordingly, this result suggests that the common ancestor of the CHE population is derived from the ancestor of the COR-I population.

Haplotype diversity was very high and nucleotide diversity was very low in *A. agrarius* and *A. chejuensis*, suggesting that the populations of two species expanded rapidly after a period of low effective population size (Grant and Bowen, 1998). Although the analyses included more *A. chejuensis* than *A. agrarius* individuals, the *A. agrarius* exhibited greater genetic diversity.

Assuming that *A. agrarius* and *A. chejuensis* have a common ancestor and diverged recently, this finding suggests that *A. chejuensis* diverged from an *A. agrarius* ancestor. In addition, network analysis showed that *A. agrarius* migrated inland and dispersed from the COR-I population, and that *A. chejuensis* diverged from an ancestral lineage of the COR-I population and migrated to Jeju Island; these populations dispersed rapidly to the Korean peninsula and Jeju Island, respectively. *F_s* estimation and mismatch distribution analysis support this hypothesis.

2.4.2. Possible migration histories of *Apodemus chejuensis*

It is estimated that Jeju Island was formed by volcanic activity about 2 Mya, and is younger than the Korean Peninsula (Yoon, 1997). The previous reports have been postulated that *A. chejuensis* migrated through the southern region of the Korean Peninsula because Jeju Island is located closer to the southern Korean peninsula than eastern China or Japan, and is the youngest in terms of geographic history, scientists suggested that the Southern Sea acted as a bridge between the Korean Peninsula and Jeju Island (Han *et al.*, 1996; Koh *et al.*, 2000; Yoon *et al.*, 2004b).

However, the results of this study showed that the closer relationship between the population (CHE) of *A. chejuensis* and COR-I of *A. agrarius* than that between both populations (COR-I and -II) of *A. agrarius*. Although Jeju Island is much closer to the southern area of the Korean peninsula than it is to the western islets, indicating that these were inconsistent with previous opinions on the migration routes of *A. chejuensis*.

Based on the results of the present study, new hypotheses can be postulated (Fig. 2.6). At first, the ancestors of *A. agrarius* probably immigrated from northeastern China through the northern part of the Korean Peninsula and reached around the COR-I region, then they arrived on Jeju

Island through the Yellow Sea shelf about 0.3 Mya (Fig. 2.6A). Although Jeju Island is remarkably closer to the southern peninsula than to COR-I, the genetic distance between the CHE and COR-I is higher; it is also supposed that the geographical conditions of southern parts of the Korean Peninsula was unsuitable for migration or habitation of their ancestors.

At second, the ancestor of *A. chejuensis* might have come from the eastern coastal regions of China via the Yellow Sea basin during glacial period when the sea level was enough low to pass by walking. Part of the CHE, an ancestral lineage of *A. chejuensis*, moved across the Yellow Sea basin from Jeju Island to the western islets of the Korean Peninsula about 0.3 Mya (Fig. 2.6B).

Finally, since *A. agrarius* might migrate from eastern China to the Yellow Sea basin when the sea level was lower (Fig. 2.6D), the species diverged into 2 lines about 0.3 Mya. As the sea level dropped further (Fig. 2.6D), one of these lines could reach the western islet of the Korean peninsula, and the other might reach Jeju Island (Fig. 2.6C). The ancestor of the COR-II scattered rapidly throughout the Korean Peninsula. Differing from the previous reports proposed the southern parts of the Korean Peninsula as a bridge for migration of *A. chejuensis* between the Korean Peninsula and Jeju Island, in all hypotheses of this study, the Yellow Sea basin may have acted as a migration route of *A. chejuensis*. Because the Yellow Sea has an average depth of 55 m (maximum, 100 m) (Yang *et al.*, 2003), small animals were able to cross the Yellow Sea when the sea level was low (Park *et al.*, 2004). Han *et al.* (1996) suggested that *A. agrarius* and *A. chejuensis* diverged 1.2 Mya, whereas Koh *et al.* (2000) estimated that these species diverged about 7000 - 500,000 years ago, although both referred to the last ice age as a critical period. Therefore, the ancestral lineage of *A. chejuensis* possibly migrated to Jeju Island through the Yellow Sea basin and dispersed quickly after colonizing the island; this lineage eventually became the distinct

subspecies *A. chejuensis*.

There has been considerable controversy about the taxonomic status of *A. agrarius* and *A. chejuensis*. The population found on Jeju Island was classified as a subspecies of *A. agrarius* by Jones and Johnson in 1965. However, subsequent molecular studies suggested that *A. chejuensis* is a species distinct from *A. agrarius* (Han *et al.*, 1996; Koh *et al.*, 2000). Crossbreeding experiments conducted by Oh and Mori (1998a) led them to conclude that the 2 populations are separate biologic species. The results of this study support the suggestion that *A. chejuensis* is a separate species (Han *et al.*, 1996; Oh and Mori, 1998a; Koh *et al.*, 2000).

The results of this study suggest that *A. agrarius* and *A. chejuensis* are clearly distinct populations, and that *A. chejuensis* migrated to Jeju Island through the Yellow Sea basin after divergence from the common ancestral lineage of *A. agrarius*. Assessment of additional mouse samples from North Korea and China may help to understand the evolutionary and biogeographic history of these 2 species

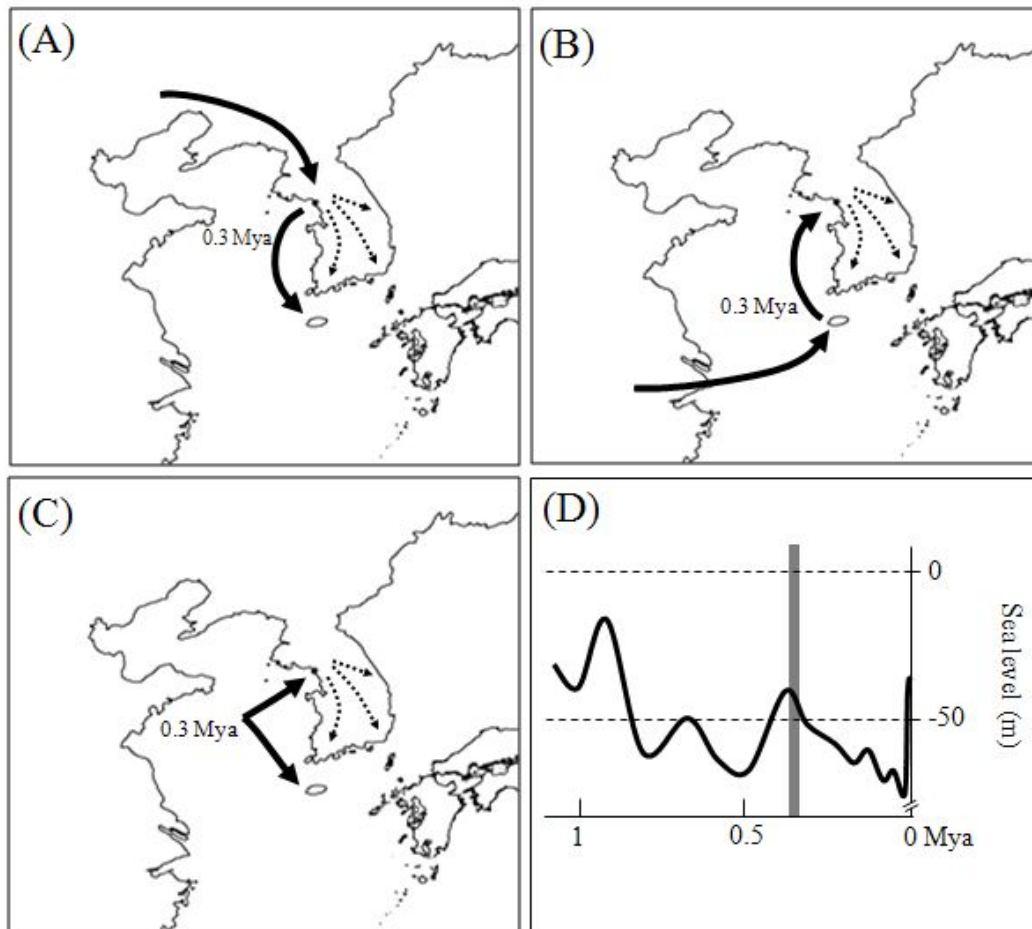


Figure 2.6. Hypothetic migration routes of *A. agrarius* and *A. chejuensis* in Korea. (A) The ancestral lineage of *A. agrarius* might migrate into the Korean Peninsula via North Korea and then could migrate to Jeju Island via the Yellow Sea basin when the sea level was low. (B) The ancestral lineage of *A. chejuensis* might migrate to Jeju Island via the Yellow Sea basin and then could migrate to the western islet of the Korean Peninsula via the Yellow Sea basin when the sea level was low. (C) The common ancestor of *A. agrarius* and *A. chejuensis* might had originated in eastern China, where it split into 2 lineages; one lineage could moved to the western islet of the Korean Peninsula, and the other could moved to Jeju Island. Arrows with dotted lines indicate dispersion of *A. agrarius*. (D) Modified eustatic sea level history (Zhong *et al.*, 2004): Gray bar indicates divergence time (0.3 Mya) of *A. agrarius* and *A. chejuensis*.

LITERATURE CITED

- Adkins R.M., Gelke E.L. Rowe D. and Honeycutt R.L. 2001. Molecular phylogeny and divergence time estimates for major rodent groups: Evidence from multiple genes. *Mol. Biol. Evol.* 18:777-791.
- Akaike H. 1973. Information theory as an extension of the maximum likelihood principle, in: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest, Hungary.
- Anderson S., Bankier A.T., Barrell B.G., De Bruijn M.H.L., Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A., Sanger F., Schreier P.H., Smith A.J.H., Staden R. and Young I.G. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465.
- Arnason U., Adegoke J.A., Gullberg A., Harley E.H., Janke A. and Kullberg M. 2008. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. *Gene* 421:37-51.
- Baker R.J. and Bradley R.D. 2006. Speciation in mammals and the genetic species concept. *J. Mamm.* 87:643-662.
- Baker R.J., Solari S. and Hoffmann F.G. 2002. A new Central American species from the *Carollia brevicauda* complex. *Occasional Papers, Museum of Texas Tech University* 217:1-12.
- Bayona-Bafaluy M.P., Acin-Perez R., Mullikin J.C., Park J.S., Moreno-Loshuertos R., Hu P., Perez-Martos A., Fernandez-Silva P., Bai Y. and Enriquez J.A. 2003. Revisiting the mouse mitochondrial DNA sequence. *Nucleic Acids Res.* 31:5349-5355.
- Bensch S. and Härlid A. 2000. Mitochondrial genomic rearrangements in songbirds. *Mol. Biol. Evol.* 17:107-113.
- Blanga-Kanfi S., Miranda H., Penn O., Pupko T., DeBry R.W. and Huchon D.

2009. Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evol. Biol.* 9:71.
- Boore J.L. 1999. Animal mitochondrial genomes. *Nucleic Acid Res.* 27:1767-1780.
- Boore J.L., Collins T.M., Stanton D., Daehler L.L. and Brown W.M. 1995. Deducing arthropod phylogeny from mitochondrial DNA rearrangements. *Nature* 376:163-165.
- Boore J.L., Medina M. and Rosenberg L.A. 2004. Complete sequences of two highly rearranged molluscan mitochondrial genomes, those of the scaphopod *Graptacme eborea* and of the bivalve *Mytilus edulis*. *Mol. Biol. Evol.* 21:1492-1503.
- Brand M.D. 1997. Regulation analysis of energy metabolism. *J. Exp. Biol.* 200:193-202.
- Brennicke A. and Clayton D.A. 1981. Nucleotide assignment of alkali-sensitive sites in mouse mitochondrial DNA. *J. Biol. Chem.* 256:10613-10617.
- Bridge D., Cunningham C.W., Schierwater B., Desalle R. and Buss L.W. 1992. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. *Proc. Natl. Acad. Sci. USA.* 89:8750-8753.
- Broughton R.E., Milam J.E. and Roe B.A. 2001. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. *Genome Res.* 11:1958-1967.
- Brown W.M., George M. and Wilson A.C. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA.* 76:1967-1971.
- Cantatore P. and Saccone C. 1987. Organization, structure, and evolution of mammalian mitochondrial genes. *Int. Rev. Cytol.* 108:149-208.
- Cantatore P., Roberti M., Morisco P., Rainaldi G., Gadaleta M. and Saccone C. 1987. A novel gene order in the *Paracentrotus lividus* mitochondrial genome. *Gene* 53:41-54.
- Cao Y., Adachi J. and Hasegawa M. 1994. Eutherian phylogeny as inferred

- from mitochondrial DNA sequence data. Jap. J. Gen. 69:455-472.
- Cao Y., Okada N. and Hasegawa M. 1997. Phylogenetic position of guinea pigs revisited. Mol. Biol. Evol. 14:461-464.
- Chu K.H., Li C.P. and Ho H.Y. 2001. The first internal transcribed spacer (ITS-1) of ribosomal DNA as a molecular marker for phylogenetic and population analyses in Crustacea. Mar. Biotechnol. 3:355-361.
- Clary D.O. and Wolstenholme D.R. 1985. The mtDNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization and genetic code. J. Mol. Evol. 22:252-271.
- Clayton D.A. 1982. Replication of animal mitochondrial DNA. Cell 28:693-705.
- Clayton D.A. 1991. Nuclear gadgets in mitochondrial DNA replication and transcription. Trends Biochem. Sci. 16:107-111.
- Clayton D.A. 1992. Structure and function of the mitochondrial genome. J. Inherited Metab. Dis. 15:439-447.
- Conroy C.J. and Cook J.A. 2000. Molecular systematics of a Holarctic rodent (*Microtus*: Muridae). J. Mammal. 81:344-359.
- Corbet G.B. and Hill J.E. 1980. A world list of mammalian species. British Museum (Natural History), London, 226 pp.
- Corneli P.S. 2002. Complete mitochondrial genomes and eutherian evolution. J. Mamm. Evol. 9:281-305.
- Covacin C., Shao R., Cameron S. and Barker S.C. 2006 Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). Insect Mol. Biol. 15:63-68.
- Crease T.J. 1999. The complete sequence of the mitochondrial genome of *Daphnia pulex* (Cladocera: Crustacea). Gene 233:89-99.
- D'Erchia A.M., Gissi C., Pesole G., Saccone C. and Arnason U. 1996. The guinea-pig is not a rodent. Nature 381:597-600.
- Doda J.N., Wright C.T. and Clayton D.A. 1981. Elongation of displacement loop strands in human and mouse mitochondrial DNA is arrested near

- specific template sequences. Proc. Natl. Acad. Sci. USA. 78:6116-6120.
- Dong S. and Kumazawa, Y. 2005. Complete mitochondrial DNA sequences of six snakes: phylogenetic relationships and molecular evolution of genomic features. J. Mol. Evol. 61:12-22.
- Dubey S., Michaux J., Brunner H., Hutterer R. and Vogel P. 2009. False phylogenies on wood mice due to cryptic cytochrome-*b* pseudogene. Mol. Phylogenet. Evol. 50:633-641.
- Excoffier L. and Lischer H.E. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564-567.
- Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22:521-565.
- Filippucci M.G., Macholán M. and Michaux J.R. 2002. Genetic variation and evolution in the genus *Apodemus* (Muridae: Rodentia). Biol. J. Linn. Soc. 75:395-419.
- Frye M.S. and Hedges S.B. 1995. Monophyly of the order Rodentia inferred from mitochondrial DNA sequences of the genes for 12S rRNA, 16S rRNA, and tRNA-Valine. Mol. Biol. Evol. 12:168-176.
- Fu Y.X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925.
- Gadaleta G., Pepe G., De Candia G., Quagliariello C., Sbisà E. and Saccone C. 1989. The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J. Mol. Evol. 28:497-516.
- Gouy M. and Gautier C. 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. 10:7055-074.
- Graeber M.B. and Muller U. 1998. Recent developments in the molecular genetics of mitochondrial disorders. J. Neurol. Sci. 153:251-263.
- Grant S.W. and Bowen B.W. 1998. Shallow population histories in deep

- evolutionary lineages of marine fishes: insight from sardines and anchovies and lessons for conservation. *J. Hered.* 89:415-426.
- Graur D., Hide W.A. and Li W.H. 1991. Is the guinea-pig a rodent? *Nature* 351:649-651.
- Graur D., Hide W.A. Zharkikh A. and Li W.H. 1992. The biochemical phylogeny of guinea-pigs and gundis, and the paraphyly of the order Rodentia. *Comp. Biochem. Physiol.* 101B:495-498.
- Graybeal A. 1993. The phylogenetic utility of cytochrome *b*: lessons from bufonid frogs. *Mol. Phylogenet. Evol.* 2:256-269.
- Gupta P.K., Roy J.K. and Prasad M. 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Curr. Sci.* 4:524-535.
- Gutiérrez G., Márquez L. and Marín A. 1996. Preference for guanosine at first codon position in highly expressed *Escherichia coli* genes. A relationship with translational efficiency. *Nucleic Acid Res.* 24:2525-2527.
- Han S.H., Wakana S., Suzuki H., Hirai Y. and Tsuchiya K. 1996. Variation of the mitochondrial DNA and the nuclear ribosomal DNA in the striped field mouse *Apodemus agrarius* on the mainland and offshore islands of South Korea. *Mammal Study* 21:125-136.
- Hanada T., Suzuki T., Yokogawa T., Takemoto-Hori C., Sprinz M. and Watanabe K. 2001. Translation ability of mitochondrial tRNAs^{Ser} with unusual secondary structures in an in vitro translation system of bovine mitochondria. *Genes Cells* 6:1019-1030.
- Harlid A. and Arnason U. 1999. Analyses of mitochondrial DNA nest ratite birds within the Neognathae-supporting a neotenus origin of ratite morphological characters. *Proc. R. Soc. Lond. B Biol. Sci.* 266:305-309.
- Helfenbein K.G., Fourcade H.M., Vanjani R.G. and Boore J.L. 2004. The mitochondrial genome of *Paraspadella gotoi* is highly reduced and reveals that chaetognaths are a sister-group to protostomes. *Proc. Natl.*

- Acad. Sci. USA. 101:10639-10643.
- Hixson J.E., Wong T.W. and Clayton D.A. 1986. Both the conserved stem-loop and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA replication. *J. Biol. Chem.* 261:2384-2390.
- Hodgkinson V.H., Birungi J., Quintana M., Deitze R. and Munstermann L.E. 2003. Mitochondrial cytochrome b variation in populations of the visceral leishmaniasis vector *Lutzomyia longipalpis* across eastern Brazil. *Am. J. Trop. Med. Hyg.* 69:386-392
- Hoffmann R.J., Boore J.L. and Brown W.M. 1992. A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. *Genetics* 131:397-412.
- Horner D.S., Lefkimiatis K., Reyes A., Gissi C., Saccone C. and Pesole G. 2007. Phylogenetic analyses of complete mitochondrial genome sequences suggest a basal divergence of the enigmatic rodent *Anomalurus*. *BMC Evol. Biol.* 7:16.
- Huchon D., Catzeflis F.M. and Douzery E.J.P. 1999. Molecular evolution of the nuclear von Willebrand factor gene in mammals and the phylogeny of rodents. *Mol. Biol. Evol.* 16:577-589.
- Huchon D., Catzeflis F.M. and Douzery E.J.P. 2000. Variance of molecular datings, evolution of rodents and the phylogenetic affinities between Ctenodactylidae and Hystricognathi. *Proc. R. Soc. Lond. B Biol. Sci.* 267:393-402.
- Ikemura T. 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J. Mol. Biol.* 146:1-21.
- Ingman M., Kaessmann H., Pääbo S. and Gyllensten U. 2001. Mitochondrial genome variation and the origin of modern humans. *Nature* 408:708-713.
- Inoue J.G., Miya M., Tsukamoto K. and Nishida M. 2000. Complete

- mitochondrial DNA sequence of the Japanese sardine *Sardinops melanostictus*. Fisheries Sci. 66:924-932.
- Irwin D.M., Kocher T.D. and Wilson A.C. 1991. Evolution of cytochrome *b* gene in mammals. J. Mol. Biol. Evol. 2:13-34.
- Iwasa M.A., Utsumi Y., Nakata K., Kartavtseva I.V., Nevedomskaya I.A., Kondoh N. and Suzuki H. 2000. Geographic patterns of cytochrome *b* and *Sry* gene lineages in the gray red-backed vole *Clethrionomys rufocanus* from far east Asia including Sakhalin and Hokkaido. Zool. Sci. 17:477-484.
- Jacobs H., Elliot D., Math V. and Farguharson A. 1988. Nucleotide sequence and gene organization of sea urchin mtDNA. J. Mol. Biol. 201:185-217.
- Jia W. and Higgs P.G. 2008. Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection. Mol. Biol. Evol. 25:339-351.
- Jones J.K. and Johnson D.H. 1965. Synopsis of the lagomorphs and rodents of Korea. University of Kansas Publications, Museum of Natural History 16:357-407.
- Jordan K. 1896. On mechanical selection and other problems. Novit. Zool. 3:426-525.
- Kang S.W. 1971. Ecological studies of the striped field mouse. Korean J. Zool. 14:57-74.
- Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111-120.
- Knight R.D., Landweber L.F. and Yarus M. 2001. How mitochondria redefine the code. J. Mol. Evol. 53:299-313.
- Koh H.S. and Yoo B.S. 1992. Variation of mitochondrial DNA in two subspecies of striped field mice, *Apodemus agrarius coreae* and *Apodemus agrarius chejuensis*, from Korea. Korean J. Zool. 35:332-338.

- Koh H.S., Chun T.Y., Yoo H.S., Zhang Y., Wang J., Zhang M. and Wu C. 2001. Mitochondrial cytochrome b gene sequence diversity in the Korean hare, *Lepus coreanus* Thomas (Mammalia, Lariiformes). *Biochem. Genet.* 39:417-429.
- Koh H.S., Lee W.J. and Kocher T.D. 2000. The genetic relationships of two subspecies of striped field mice, *Apodemus agrarius coreae* and *Apodemus agrarius chejuensis*. *Heredity* 85:30-36.
- Kondo R., Satta Y., Matsuura E.T., Ishiwa H., Takahata N. and Chigusa S.I. 1990. Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics* 126:657-663.
- Kroemer G., Dallaporta B. and Resche-Rigon M. 1998. The mitochondrial death/life regulator in apoptosis and necrosis. *Annu. Rev. Physiol.* 60:619-642.
- Kumazawa Y., Ota H., Nishida M. and Ozawa T. 1998. The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. *Genetics* 150:313-329.
- Laslett D. and Canbäck B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* 24:172-5.
- Le T.H., Blair D., Agatsuma T., Humair P.F., Campbell N.J., Iwagami M., Littlewood D.T., Peacock B., Johnston D.A., Bartley J., Rollinson D., Herniou E.A., Zarlenga D.S. and McManus D.P. 2000. Phylogenies inferred from mitochondrial gene orders—a cautionary tale from the parasitic flatworms. *Mol. Biol. Evol.* 17:1123-1125.
- Lee H.I., Lee H.S., Lee I.Y. and Yoshida Y. 1991. Epidemiological studies on host animals of tsutsugamushi disease in Korea. *Kisaengchunghak-chapchi* 29:181-188.
- Lee H.W. and Lee P.W. 1976. Korean hemorrhagic fever I. Demonstration of causative antigen and antibodies. *Korean J. Int. Med.* 19:371-383.
- Li W.-H., Hide W.A. and Graur D. 1992. Origin of rodents and guinea-pigs.

- Nature 359:277-278.
- Librado P. and Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
- Lin Y.H., Waddell P.J. and Penny D. 2002. Pika and vole mitochondrial genomes increase support for both rodent monophyly and glires. *Gene* 294:119-129.
- Liu X., Wei F., Li M., Jiang X., Feng Z. and Hu J. 2004. Molecular phylogeny and taxonomy of wood mice (genus *Apodemus* Kaup, 1829) based on complete mtDNA cytochrome *b* sequences, with emphasis on Chinese species. *Mol. Phylogenet. Evol.* 33:1-15.
- Lovejoy N.R. and de Araújo M.L.G. 2000. Molecular systematics, biogeography, and population structure of Neotropical freshwater needlefishes of the genus *Potamorhaphis*. *Mol. Ecol.* 9:259-268.
- Lowe T.M. and Eddy S.R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955-964.
- Luckett W.P. and Hartenberger J.-L. 1993. Monophyly or polyphyly of the order Rodentia: Possible conflict between morphological and molecular interpretations. *J. Mammal. Evol.* 1:127-147.
- Mabuchi K., Miya M., Satoh T.P., Westneat M.W. and Nishida M. 2004. Gene rearrangements and evolution of tRNA pseudogenes in the mitochondrial genome of the parrotfish (Teleostei: Perciformes: Scaridae). *J. Mol. Evol.* 59:287-297.
- Macey J.R., Larson A., Ananjeva N.B., Fang Z. and Papenfuss T.J. 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Mol. Biol. Evol.* 14:91-104.
- Machado T.M.M., Chakir M. and Lauvergne J.J. 2000. Genetic distances and taxonomic trees between goats of Ceará State (Brazil) and goats of the

- Mediterranean region (Europe and Africa). *Genet. Mol. Biol.* 23:121–125.
- Mackay S.L.D., Olivo P.D., Laipis P.J. and Hauswirth W.W. 1986. Template-directed arrest of mammalian mitochondrial DNA synthesis. *Mol. Cell. Biol.* 6:1261–1267.
- Martin Y., Gerlach G., Schlotterer C. and Meyer A. 2000. Molecular phylogeny of European muroid rodents based on complete cytochrome *b* sequences. *Mol. Phylogenet. Evol.* 16:37–47.
- Matocq M.D. 2002. Morphological and molecular analysis of a contact zone in the *Neotoma fuscipes* species complex. *J. Mamm.* 83:866–883.
- Mayr E. 1969. Principles of systematic zoology. McGraw-Hill, New York.
- Mayr E. 1982. The Growth of Biological Thought: Diversity, Evolution and Inheritance. Harvard University Press, Cambridge.
- Mercer J.M. and Roth V.L. 2003. The effects of Cenozoic global change on squirrel phylogeny. *Science* 299:1568–1572.
- Meyer A. 1994. Shortcomings of the cytochrome *b* gene as a molecular marker. *Trends Ecol. Evol.* 9:278–280.
- Michaels G.S., Hauswirth W.W. and Laipis P.J. 1982. Mitochondrial DNA copy number in bovine oocytes and somatic cells. *Dev. Biol.* 94:246–251.
- Michaux J., Reyes A. and Catzeflis F. 2001. Evolutionary history of the most speciose mammals: molecular phylogeny of muroid rodents. *Mol. Biol. Evol.* 18:2017–2031.
- Michaux J.R., Chevret P., Filippucci M.G. and Macholán M. 2002. Phylogeny of the genus *Apodemus* with a special emphasis to the subgenus *Sylvaemus* using the nuclear IRBP gene and two mitochondrial markers: cytochrome *b* and 12S rRNA. *Mol. Phylogenet. Evol.* 23:123–136.
- Milbury C.A. and Gaffney P.M. 2005. Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*. *Mar. Biotechnol.* 7:697–712.
- Miya M., Takeshima H., Endo H., Ishiguro N.B., Inoue J.G., Mukai T., Satoh T.P., Yamaguchi M., Kawaguchi A., Mabuchi K., Shirai S.M. and

- Nishida M. 2003. Major patterns of higher teleostean phylogenies; a new perspective based on 100 complete mitochondrial DNA sequences. *Mol. Phylogent. Evol.* 26:121-138.
- Mouchaty S.K., Catzeflis F., Janke A. and Arnason U. 2001. Molecular evidence of an African Phiomorpha-South American Caviomorpha clade and support for Hystricognathi based on the complete mitochondrial genome of the cane rat (*Thryonomys swinderianus*). *Mol. Phylogenet. Evol.* 18:127-135.
- Mueller R.L. and Boore J.L. 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in Plethodontid salamanders. *Mol. Biol. Evol.* 22:2104-2112.
- Musser G.G. and Carleton M.D. 2005. Order Rodentia. In: Wilson D.E., Reeder D.M. (eds) *Mammal Species of the World, a Taxonomic and Geographic Reference*, 3rded. Johns Hopkins University Press.
- Nedbal M.A., Honeycutt R.L. and Schlitter D.A. 1996. Higher-level systematics of rodents (Mammalia, Rodentia): evidence from the mitochondrial 12S rRNA gene. *J. Mamm. Evol.* 3:201-237.
- Nilsson M.A., Harlid A., Kullberg M. and Janke A. 2010. The impact of fossil calibrations, codon positions and relaxed clocks on the divergence time estimates of the native Australian rodents (Conilurini). *Gene* 455:22-31.
- Noack K., Zardoya R. and Meyer A. 1996. The complete mitochondrial DNA sequence of the bichir (*Polypterus ornatipinnis*), a basal ray-finned fish: ancient establishment of the consensus vertebrate gene order. *Genetics* 144:1165-1180.
- Oh D.J. and Jung Y.H. 2008. The mitochondrial genome of the threespot wrass *Halichoeres trimaculatus* (Perciformes, Labridae). *Genes Genom.* 30:113-120.
- Oh D.J., Chang M.H., Oh H.S. and Jung Y.H. 2007a. The complete mitochondrial DNA sequence of the Jeju salamander, *Hynobius*

- quelpaertensis*, and the phylogenetic relationships among the Hynobiidae. Korean J. Genetics 29:331-341.
- Oh D.J., Kim J.Y., Lee J.A., Yoon W.J., Park S.Y. and Jung Y.H. 2007b. Complete mitochondrial genome of the rock bream *Oplegnathus fasciatus* (Perciformes, Oplegnathidae) with phylogenetic considerations. Gene 392:174-180.
- Oh D.J., Oh B.S., Jung M.M. and Jung Y.H. 2010. Complete mitochondrial genome of three *Branchiostegus* (Perciformes, Malacanthidae) species: genome description and phylogenetic considerations. Mitochondrial DNA 21:151-159.
- Oh H.S. and Mori T. 1998b. Taxonomic re-examination of the striped field mouse, *Apodemus agrarius coreae* and *A. a. chejuensis*: evidence from crossbreeding experiments (Mammalia: Rodentia). J. Fac. Agr. 43:143-151.
- Oh H.S. and Mori T. 1998a. Reproduction, growth and development of the striped field mouse, *Apoemus agrarius chejuensis*. Honyurui Kagaku (Mammalian Science), 38:23-37.
- Ojala D., Montoya J. and Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470-474.
- Okimoto R., Chamberlin H.M., MacFarlane J.L. and Wolstenholme D.R. 1991. Repeated sequence sets in mitochondrial DNA molecules of root knot nematodes (Meloidogyne): nucleotide sequences, genome location and potential for host race identification. Nucleic Acids Res. 19:1619-1626.
- Okimoto R., MacFarlane J.L., Clary D.O. and Wolstenholme D.R. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. Genetics 130:471-498.
- Pääbo S., Thomas W.K., Whitfield K.M., Kumazawa Y. and Wilson A.C. 1991. Rearrangements of mitochondrial transfer RNA genes in marsupials. J. Mol. Evol. 33:426-430.
- Park Y.C., Maekawa K., Matsumoto T., Santoni R. and Choe J.C. 2004.

- Molecular phylogeny and biogeography of the Korean woodroaches *Cryptocercus* spp. *Mol. Phylogenet. Evol.* 30:450-464.
- Parsons T.J. and Coble M.D. 2001. Increasing the forensic discrimination of mitochondrial DNA testing through the analysis of the entire mitochondrial DNA genome. *Croatian Med. J.* 42:304-309.
- Partridge M.A., Davidson M.M. and Hei T.K. 2007. The complete nucleotide sequence of Chinese hamster (*Cricetulus griseus*) mitochondrial DNA. *DNA Seq.* 18: 341-346.
- Piaggio A.J., Valdez E.W., Bogan M.A. and Spicer G.S. 2002. Systematics of *Myotis occultus* (Chiroptera: Vespertilionidae) inferred from sequences of two mitochondrial genes. *J. Mamm.* 83:386-395.
- Posada D. 2008. jModeltest: phylogenetic model averaging. *Mol. Biol. Evol.* 25: 1253-1256.
- Rand D.M. 1993. Endotherms, ectotherms, and mitochondrial genome-size variation. *J. Mol. Evol.* 37:281-295
- Reyes A., Pesole G. and Saccone C. 1998. Complete mitochondrial DNA sequence of the fat dormouse, *Glis glis*: further evidence of rodent paraphyly. *Mol. Biol. Evol.* 15:499-505.
- Robin E.D. and Wong R. 1988. Mitochondrial DNA molecules and virtual number of mitochondrial per cell in mammalian cells. *J. Cell. Physiol.* 136:507-513.
- Robins J.H., McLenachan P.A., Phillips M.J., Craig L., Ross H.A. and Matixxo-Smith E. 2008. Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. *Mol. Phylogenet. Evol.* 49:460-466.
- Robins J.H., McLenachan P.A., Phillips M.J., McComish B.J., Matisso-Smith E. and Ross H.A. 2010. Evolutionary relationships and divergence times among the native rats of Australia. *BMC Evol. Biol.* 10:375.
- Robinson-Rechavi M., Ponger L. and Mouchiroud D. 2000. Nuclear gene

- LCAT supports rodent monophyly. *Mol. Biol. Evol.* 17:1410–1412.
- Roca A.L., Georgiadis N., Pecon-Slattery J. and O'Brien S.J. 2001. Genetic evidence for two species of elephant in Africa. *Science* 293:1473–1477.
- Serizawa K., Suzuki H. and Tsuchiya K. 2000. A phylogenetic view on species radiation in *Apodemus* inferred from variation of nuclear and mitochondrial genes. *Biochem. Genet.* 38:27–41.
- Shadel G.S. and Clayton D.A. 1997. Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.* 66:409–435.
- Shim J.H. and Park B.S. 1998. Vertebrate fauna, speciation and geological history in the Cheju Island. *Kor. J. Env. Eco.* 12:42–57. (in Korean)
- Sinniger F., Chevaldonné P. and Pawlowski J. 2007. Mitochondrial genome of *Savalia savaglia* (Cnidaria, Hexacorallia) and early metazoan phylogeny. *J. Mol. Evol.* 64:196–203.
- Steppan S.J., Adkins R.M. and Anderson J. 2004. Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst. Biol.* 53:533–553.
- Steppan S.J., Adkins R.M., Spinks P.Q. and Hale C. 2005. Multigene phylogeny of the Old World mice, Murinae, reveals distinct geographic lineages and the declining utility of mitochondrial genes compared to nuclear genes. *Mol. Phylogenet. Evol.* 37:370–388.
- Suzuki H., Filippucci M.G., Chelomina G.N., Sato J.J., Serizawa K. and Nevo E. 2008. A biogeographic view of *Apodemus* in Asia and Europe inferred from nuclear and mitochondrial gene sequences. *Biochem. Genet.* 46:329–346.
- Suzuki H., Sato J.J., Tsuchiya K., Luo J., Zhang Y.-P., Wang Y.-X. and Jiang X.-L. 2003. Molecular phylogeny of wood mice (*Apodemus*, Muridae) in East Asia. *Biol. J. Linn. Soc.* 80:469–481.
- Suzuki H., Tsuchiya K. and Takezaki N. 2000. A molecular phylogenetic framework for the Ryukyu endemic rodents *Tokudaia osimensis* and

- Diplothrix legata*. 15:15-24.
- Tamura K. and Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10:512-526.
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M. and Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.
- Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F. and Higgins D.G. 1997. The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876-4882.
- Touriol C., Bornes S., Bonnal S., Audigier S., Prats H., Prats A.C. and Vagner S. 2003. Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol. Cell* 95:169-178.
- Triant D.A. and DeWoody A.J. 2007. The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. *J. Mammal.* 88:908-920.
- Valverde J.R., Marco R. and Garesse R. 1994. A conserved heptamer motif for ribosomal RNA transcription termination in animal mitochondria. *Proc. Natl. Acad. Sci. USA.* 91:5368-5371.
- van Embden J.D., Cave M.D., Crawford J.T., Dale J.W., Eisenach K.D., Gicquel B., Hermans P., Martin C., McAdam R. and Shinnick T.M. 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J. Clin. Microbiol.* 31:406-409.
- Voris H.K. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J. Biogeogr.* 27:1153-1167.
- Waddell P.J. and Shelley S. 2003. Evaluating placental inter-ordinal

- phylogenies with novel sequences including RAG1, gamma-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. *Mol. Phylogenet. Evol.* 28:197–224.
- Wei Y.H. 1998. Oxidative stress and mitochondrial DNA mutations in human aging. *Proc. Soc. Exp. Biol. Med.* 217:53–63.
- Wilson E.O. 1998. *Consilience the unity of knowledge*. Alfred A. Knopf, Inc., New York.
- Won C.M. and Smith K.G. 1999. History and current status of mammals of the Korean peninsula. *Mammal Rev.* 29:3–33.
- Won H. and Renner S.S. 2006. Dating dispersal and radiation in the Gymnosperm *Gnetum* (Gnetales)–clock calibration when outgroup relationships are uncertain. *Syst. Biol.* 55:610–622.
- Wong T.W. and Clayton D.A. 1985. *In vitro* replication of human mitochondrial DNA: Accurate initiation at the origin of light-strand synthesis. *Cell* 42:951–958.
- Wyman S. and Boore J. 2003. Annotating animal mitochondrial tRNAs: an experimental evaluation of four methods. In *Proceedings of European Conference on Computer. Biology, Local Proceedings, Self Published, Paris, France, pp. 44–46.*
- Xia W.P. 1984. A study on Chinese *Apodemus* with a discussion of its relations to Japanese species. *Acta Theriol. Sin.* 4:93–98. (in Chinese)
- Yamazaki N., Ueshima R., Terrett J.A., Yokobori S.-I., Kaifu M., Segawa R., Kobayashi T., Numachi K.-I., Ueda T., Nishikawa K., Watanabe K. and Thomas R.H. 1997. Evolution of Pulmonate Gastropod mitochondrial genomes: comparisons of gene order organizations of *Euhadra*, *Cepaea* and *Albinaria* and implications of unusual tRNA secondary structures. *Genetics* 145:749–758.
- Yan J., Li H. and Zhou K. 2008. Evolution of the mitochondrial genome in snakes: gene rearrangements and phylogenetic relationships. *BMC*

- Genomics 9:269.
- Yang S.Y., Jung H.S., Lim D.I. and Li C.X. 2003. A review on the provenance discrimination of sediments in the Yellow Sea. *Earth-Sci. Rev.* 63:93-120.
- Yasuda N., Hamaguchi, M., Sasaki, M., Nagai, S., Saba, M. and Nadaoka, K. 2006. Complete mitochondrial genome sequences for Crown-of-thorns starfish *Acanthaster planci* and *Acanthaster brevispinus*. *BMC Genomics* 7:17.
- Yoon M.H. Jung S.J. and Oh H.S. 1997. Studies on population structure and reproductive pattern of the Korean striped field mice, *Apodemus agrarius*. *Korean J. Biol. Sci.* 1:53-61.
- Yoon M.H., Han S.H., Oh H.S. and Kim J.G. 2004a. The mammals of Korea. Dong-bang Media. Seoul. (in Korean)
- Yoon M.H., Phillips C.J., Kim I. and Oh H.S. 2004b. Biogeography, genetic structure, and speciation in the striped field mouse, *Apodemus agrarius*, in Southern Korea. *Korean J. Genet.* 26:15-28.
- Yoon S. 1997. Miocene-Pleistocene volcanism and tectonics in southern Korea and their relationship to the opening of the Japan Sea. *Tectonophysics* 281:53-70.
- Yoon S., Koh G.W. and Kim I.T. 1995. Development history of Cheju Island. *Chejudo (Bull. Cheju-do Prov. Gov.)* 98:148-164. (in Korean)
- Zardoya R. and Doadrio I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *J. Mol. Evol.* 49:227-237
- Zhang D.X. and Hewitt G.M. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol. Evol.* 11:247-251.
- Zhang D.X. and Hewitt G.M. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Mol. Ecol.* 12:563-584.
- Zhang P., Chen Y.Q., Zhou H., Liu Y.F., Wang X.L., Papenfuss T.J., Wake D.B. and Qu L.H. 2006. Phylogeny, evolution, and biogeography of

Asiatic salamanders (Hynobiidae). Proc. Natl. Acad. Sci. USA. 103:7360-7365.

Zhang Y.Z. 2002. Geological events and mammalian distribution in China. Acta Zool. Sin. 48:141-153.

Zhong G., Geng J., Wong H.K., Ma Z. and Wu N. 2004. A semi-quantitative method for the reconstruction of eustatic sea level history from seismic profiles and its application to the southern South China Sea. Earth Planet Sc. Lett. 223:443-459.

국문 초록

제 1장. 한국산 붉은쥐속의 미토콘드리아 유전체 비교 분석

붉은쥐속(*Apodemus*)을 포함하는 쥐목(Rodentia)의 계통학적 분석을 위한 기초정보를 제공하기 위해 한국산 붉은쥐속인 등줄쥐(*Apodemus agrarius*)와 제주 등줄쥐(*Apodemus chejuensis*) 그리고 흰넓적다리붉은쥐(*Apodemus peninsulae*)의 미토콘드리아 유전체의 전체 염기서열을 분석하였다. 분석된 미토콘드리아 유전체 서열은 등줄쥐, 제주등줄쥐, 그리고 흰넓적다리붉은쥐에서 각각 16,260, 16,261, 그리고 16,266 bp로 확인되었다. 다른 척추동물들과 마찬가지로 한국산 붉은쥐속 3종의 미토콘드리아 유전체에는 13 개의 단백질 유전자, 22 개의 tRNAs 유전자, 2 개의 rDNAs, 그리고 비암호화 영역인 조절부위(Control region)를 포함하고 있었다. 등줄쥐와 제주등줄쥐의 미토콘드리아 유전체에서는 단백질 암호화 유전자에서 ATG 코돈이 개시코돈으로 가장 많이 사용되었고, GTG와 ATA 개시코돈도 확인되었다. 한편, 흰넓적다리붉은쥐에서는 추가로 ATC 개시코돈도 ND2 유전자에서 사용되고 있음을 확인할 수 있었다. 13개 단백질 암호화 유전자에서의 뉴클레오티드의 출현 빈도는 한국산 붉은쥐속 3종 모두 비슷하게 나타났다. 코돈내 뉴클레오티드 출현 양상은 등줄쥐와 제주등줄쥐가 매우 유사하였고, 흰넓적다리붉은쥐는 다소 차이를 보였다. 12S rDNA는 등줄쥐와 제주등줄쥐가 955 bp로 동일하였고, 흰넓적다리붉은쥐는 956 bp로 확인되었다. 반면에 16S rDNA에서는 등줄쥐와 제주등줄쥐, 그리고 흰넓적다리붉은쥐가 각각 1,572, 1,573, 그리고 1,571 bp이었다. tRNA 유전자들 중에서는 tRNA-Leu(UUR과 CUN)과 tRNA-Ser(UCN과 AGY)이 두 형태로 존재함을 확인하였고, 전형적인 설치류에서처럼 3종 모두에서 3 종류 tRNA cluster(IQM, WANCY, HSL)들이 모두 보존되어 있었다. 일반적으로 포유류의 미토콘드리아 게놈은 2개의 비암호화 영역을 포함하고 있다. 그 중 하나는 light strand의 복제 기점(O_L)이며 나머지는 D-loop으로도 불리는 조절부위(CR)이다. 특히 조절부위는 tRNA-Pro와 tRNA-Phe 유전자의 중간에 위치하고 있었으며, 이들의 크기는

각각 등줄쥐에서 854 bp, 제주등줄쥐에서 857 bp, 그리고 흰넓적다리붉은쥐에서 866 bp이었다. 미토콘드리아 13 개 단백질 암호화 유전자들을 이용한 쥐목의 계통학적 분석에서 청설모(*Sciurus vulgaris*)와 큰동면쥐(*Myoxus glis*)의 분기가 쥐목 중에서도 가장 초기에 분화된 종임을 보여주고 있으며, 집쥐속(*Rattus*)은 두 개의 그룹으로 명확히 구분되고, 한국산 붉은쥐속은 분석된 종들 중에서 생쥐속(*Mus*)과 가장 근연종으로 확인되었다. 현재까지 붉은쥐속 종들간의 계통학적인 유연관계에 대한 여러 연구보고들이 있지만, 아직까지도 일부 종에 대해서는 분포범위가 매우 광범위하고, 아종으로써의 취급 문제 및 단편적인 정보들로 인하여 논쟁거리로 남아있다. 따라서 본 연구에서 얻은 미토콘드리아 계통 정보는 향후 쥐목 및 붉은쥐속의 종들간의 계통학적 관계를 밝히는데 기초자료로 활용될 것으로 기대된다.

제 2장. 제주등줄쥐의 역사적 이주 경로 추정

등줄쥐(*Apodemus agrarius*)와 제주등줄쥐(*Apodemus chejuensis*) 집단에서 미토콘드리아 *CYTB* 유전자의 haplotype의 상관관계를 분석함으로써 두 종의 진화적 상관관계와 분포적 특성을 추정하였다. 등줄쥐에서 19 개, 제주등줄쥐에서는 24 개의 *CYTB* haplotype들이 확인되었다. 가장 빈도가 높은 haplotype는 Hap27로 제주등줄쥐 6 개체에서 관찰되었고, 등줄쥐 집단에서는 Hap15가 가장 많은 3 개체에서 관찰되었다. 등줄쥐와 제주등줄쥐 집단의 확장 상태를 확인하기 위해 mismatch distribution 분석과 Fu's F_s 분석을 수행하였고, 그 결과는 COR-II와 CHE 집단이 확장되었음을 보여주었다. 반면, COR-I 집단은 확장이 멈춘 집단임을 추정할 수 있었다. 또한, median-joining network 분석에서 제주등줄쥐는 등줄쥐와 확연히 구분되는 종임을 알 수 있었고, 제주등줄쥐는 다시 4 개의 subgroup으로 구분되었으나, subgroup 간의 지역적 연관성을 나타내지는 않았다. 흥미로운 점들 중 하나는 제주등줄쥐 Hap42가 등줄쥐 Hap9와 가장 가까운 거리지수를 보인다는 점이다. 이에, 등줄쥐와 제주등줄쥐 두 종의 분기연대를 추정하기 위한 분자시계를 계산한 결과, 이 두 종은 약 30만 년 전에 분화된 것으로 추정되었다. 이상의 분석 결과들을 바탕으로 제주등줄쥐와 등줄쥐의 이주 경로에 대한 3 가지 가설을 제시할 수 있었다. 첫째는 등줄쥐의 조상계보가 중국 북동부 지역을 거쳐 북한 지역을 통해 한반도로 유입되었고, 빙하기를 거치면서 해수면이 현재보다 낮아짐으로 인해 황해지역이 육지가 되었을 때 등줄쥐 집단 중 일부가 황해를 거쳐 제주도로 유입되었을 가능성이 있다. 둘째는 중국 동부에 서식하던 제주등줄쥐의 조상계보가 황해를 건너 제주도로 유입되면서, 그 중 일부가 COR-I 지역까지 도달했을 가능성이 있다는 점이다. 마지막 가설은 중국 동부 지역에 서식하던 등줄쥐 집단 중 일부가 동쪽으로 서식범위를 넓히는 과정에서 집단 중 일부는 COR-I 지역으로, 다른 일부는 제주도로 거의 동시에 유입되었을 가능성도 배제할 수 없다는 점이다. 그리고 세 가설 모두에서 등줄쥐와 제주등줄쥐의 분기 시점이 약 30만 년 전이라고 추정된다. 현재 등줄쥐와 제주등줄쥐는 지리적으로 서로 완벽히 분리되어 있을 뿐만 아니라, 생식적으로도 두 종은 완전

히 격리되었음을 감안하면 제주등줄쥐와 등줄쥐는 완전히 별개의 종으로 취급해야 할 것이다. 본 연구 결과는 제주등줄쥐가 종으로써의 지위를 가져야 한다는 점을 지지한다. 그러나 등줄쥐와 제주등줄쥐의 정확한 유입 경로를 확정하기 위해서는 북한 및 중국 지역의 등줄쥐 개체들을 이용한 추가적인 연구가 필요하다.

APPENDICES

Appendix 1. The sequence of mitochondrial genome of *Apodemus agrarius*

LOCUS HM034866 16260 bp DNA circular ROD 30-DEC-2011
 DEFINITION *Apodemus agrarius* mitochondrion, complete genome.
 ACCESSION HM034866
 VERSION HM034866.1 GI:296100340
 KEYWORDS .
 SOURCE mitochondrion *Apodemus agrarius* (Eurasian field mouse)
 ORGANISM *Apodemus agrarius*
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; *Apodemus*.
 REFERENCE 1 (bases 1 to 16260)
 AUTHORS Oh, D. -J., Jung, Y. -H., Kim, T. -W., Chang, M. -H., Oh, H. -S., Han, S. -H.
 and Kim, S. -J.
 TITLE Biogeography and speciation of Korean striped field mice *Apodemus*
agrarius and *A. chejuensis* inferred from mitochondrial DNA
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 16260)
 AUTHORS Oh, D. -J., Jung, Y. -H., Kim, T. -W., Chang, M. -H., Oh, H. -S., Han, S. -H.
 and Kim, S. -J.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2010) Department of Biology, Jeju National
 University, 66 Jejudaehakno, Jeju 690-756, Republic of Korea
 FEATURES Location/Qualifiers
 source 1..16260
 /organism="Apodemus agrarius"
 /organelle="mitochondrion"
 /mol_type="genomic DNA"
 /specimen_voucher="JBRI-Mam-001"
 /db_xref="taxon:39030"
 tRNA 1..67
 /product="tRNA-Phe"
 rRNA 68..1022
 /product="12S ribosomal RNA"
 tRNA 1023..1091
 /product="tRNA-Val"
 rRNA 1092..2663
 /product="16S ribosomal RNA"
 tRNA 2664..2738
 /product="tRNA-Leu"
 /note="codons recognized: UUR"
 gene 2739..3693
 /gene="ND1"
 CDS 2739..3693
 /gene="ND1"
 /note="TAA stop codon is completed by the addition of 3' A

```

residues to the mRNA"
/codon_start=1
/transl_except=(pos:3693, aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 1"
/protein_id="ADG95699.1"
/db_xref="GI:296100341"
/translation="MYFINTLMLLVPILIAMAFLLVERKILGYMQLRKGPNVVGPYGV
VLQPFADAMKLFIKPMRPLTTSISLFI|APTLSLTLALSLWIPLPMPHPLINLNLGL
LFILATSSLSVYSILWSGWASNSKYSLFGALRAVAQTSYEVTMAIILLSVLLMSGFS
SLQMLIYQEHMWLILPAWPMAMMWYISTLAETNRAPFDLTEGESELVSGFNVEYAAG
PFALFFMAEYTNILMNALTSIVFLGPIHMINYPELYSINFMTETLLSTFTLWIRAS
YPRFRYDQLMHLLWKNFLPLTLALCTWHISLP|FMASVPPYI"
tRNA 3694..3762
/product="tRNA-Ile"
tRNA complement(3760..3834)
/product="tRNA-Gln"
tRNA 3835..3903
/product="tRNA-Met"
gene 3904..4939
/gene="ND2"
CDS 3904..4939
/gene="ND2"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:4939, aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="ADG95700.1"
/db_xref="GI:296100342"
/translation="MNPITIIIIYFTILIGPVI|TMSSNLLLMWVGLLELSLLAI|PLL
INKNRPRSTEAATKYFIQTATSMI|LLAIVLNFKQLGLWTFQQQNTSLLNITLIAL
SMKLG LAPFHFWLPEVTQGIPLHMGL|LLTWQKIAPLSILFQFYQLLNPTIIMILAIS
SILIGAWGGLNQTQMRKIMAYSSIAHMGWMLA|LPYNPTMTLLNLLIYI|LTVPMFLM
LMLNASTTINSISLLWNAKPTTLVLIPL|LLSLGGLPPLTGFLPKWAI|TELLKNCL
I|ITMMAI|MALLNLFYTRLIYSTSLTMFPTNNNSKMLAHHLNYKNNI|LPSLT|IST
L|LPLSPQLII"
tRNA 4940..5007
/product="tRNA-Trp"
tRNA complement(5009..5077)
/product="tRNA-Ala"
tRNA complement(5079..5149)
/product="tRNA-Asn"
rep_origin complement(5150..5182)
/note="L-strand replication origin"
tRNA complement(5181..5248)
/product="tRNA-Cys"
tRNA complement(5249..5315)
/product="tRNA-Tyr"
gene 5317..6861
/gene="CO1"
CDS 5317..6861

```

```

/gene="C01"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit 1"
/protein_id="ADG95701.1"
/db_xref="GI:296100343"
/translation="MFINRWLFSTNHKDIGTLYLLFGAWAGMVGITALSILIRAE LGQP
GALLGDDQIYNVIVTAHAFVMIFFMVMPMMIGGFGNWLVP LMI GAPDMAFPRMNMSF
WLLPPSFLLLLASSMVEAGAGTGWTVYPPLAGNLAHAGASVDLTI FSLHLAGVSSILG
AINFITTIINMKPPAMTQYQTPLFVWSVLI TAVLLLLSLPALAAGITMLLTDRLNNTT
FFDPAGGGDPILYQHLEWFFGHPEVYILILPGFGIISHVVYYS GKKEPFGYMGMVWA
MMSIGFLGFI VWAHMFVGLDVDTRAYF TSATMIIA IPTGVKVF SWLATLHGGINIKW
SPAMLWALGFI FLFTVGGLTGIVLSNSSLDIVLHDTYVVVAHFHYVLSMGAVFAIMAG
FVHWFPLFTGYTLDDMWAKTHFAIMFVGVNMTFFPQHFLGLSGMPRRYSYDPA YTTW
NTVSSMGSFISLTAVLLMIFMIWEAFASKREVLTVSYSSNTLEWLHGCPPPYHTFEEP
SYVKVK"
tRNA complement(6859..6927)
/product="tRNA-Ser"
/note="codons recognized: UCN"
tRNA 6931..6999
/product="tRNA-Asp"
gene 7000..7683
/gene="C02"
CDS 7000..7683
/gene="C02"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit 2"
/protein_id="ADG95702.1"
/db_xref="GI:296100344"
/translation="MAYPFQLGLQDATSPIMEELMNFHDHTLMIVFLISTLVLYIISL
MLTTELHTSTMDAQEVETVWTILPAVILILIALPSLRILYMMDEINNPVLTVKTMGH
QWYWSYEYTDYEDLCFDSYMIPTDLKPGDLRLLEVDNRV LPMGLPIRMLISSEDVL
HSWAVPSLW LKTDAPGRLNQATVTSNRPGLFYGCSEICGSNHSFMPIVLEMVPLKH
FENWSASMI"
tRNA 7687..7750
/product="tRNA-Lys"
gene 7752..7955
/gene="ATP8"
CDS 7752..7955
/gene="ATP8"
/codon_start=1
/transl_table=2
/product="ATP synthase F0 subunit 8"
/protein_id="ADG95703.1"
/db_xref="GI:296100345"
/translation="MPQLDTSTWFITIISSMATLFI LFQLKISSQSFPTPPSPKFTFA
QETKTPWESKWKIYLPLLSLPQ"
gene 7913..8593
/gene="ATP6"
CDS 7913..8593
/gene="ATP6"
/codon_start=1

```



```

/db_xref="GI:296100349"
/translation="MSSAFINLTLAFMLSLLGTLTFRSHLMSTLLCLEGMMLSLFIMT
SLASLNSNSMASMPIPITILVFAACEAAVGLALLVKVSNTYGTDYVQNLNLLQC"
gene      10154..11531
          /gene="ND4"
CDS      10154..11531
          /gene="ND4"
          /note="TAA stop codon is completed by the addition of 3' A
          residues to the mRNA"
          /codon_start=1
          /transl_except=(pos:11531,aa:TERM)
          /transl_table=2
          /product="NADH dehydrogenase subunit 4"
          /protein_id="ADG95708.1"
          /db_xref="GI:296100350"
          /translation="MLKIIFPSLMLLPLTWLSKPKKTWMNVTSYSFMISLISLSILWQ
          TDGSPLNFSSTMFSSDPLSSPLIILTLLPLMLMASQNHKKENFTYQKLYISMLVSL
          QILLVMTFSATELIMFYILFEATLIPTLIITRWGNQTERLNAGLYFLFYTLIGSIPL
          LIALIFIQNSMGTLNFTMLSLTTNPLNPLWSNNILWLACMMAFMIKMPLYGVHLWLPK
          AHVEAPIAGSMILAAILLKLGSGMMRIAIIIDPLTKYMAYPFI LLSLWGMIMTSSIC
          LRQTDLKSLEYSSVSHMALVIASIMIQTPWSFMGATMLMVAHGLTSSLLFCLANSNY
          ERHSRTMIMARGLQMI FPLMATCHLWLVASLANLALPPSINLMGELFITMSLFSWSNFS
          IILMGVNIITIGMYSMYMIITQRGKLTSHMNNLQPSHTRELTMLIHIIPVLLTIN
          PKLITGLTM"
tRNA     11532..11599
          /product="tRNA-His"
tRNA     11600..11658
          /product="tRNA-Ser"
          /note="codons recognized: AGY"
tRNA     11658..11729
          /product="tRNA-Leu"
          /note="codons recognized: CUN"
gene     11729..13556
          /gene="ND5"
CDS     11729..13556
          /gene="ND5"
          /note="TAA stop codon is completed by the addition of 3' A
          residues to the mRNA"
          /codon_start=1
          /transl_except=(pos:13556,aa:TERM)
          /transl_table=2
          /product="NADH dehydrogenase subunit 5"
          /protein_id="ADG95709.1"
          /db_xref="GI:296100351"
          /translation="MNIITSFIMAI FLILMSPILISSTSLFKHINFPMYATLSIKLSF
          FL SLLPLMMLFYHNT EYMITTWHWITINSIELTMSFKMDYFSTMFLSVALFVTSIMQ
          FSSWYMHSDPHINRFIKYLLFLITMLILTSANNLFQLFIGWEGVIGMSFLLIGWWSG
          RADANTAALQAILYNRIGDIGFILAMTWFCNMNSWELQIILLTNNNNLIPLLGLLIA
          ATGKSAQFGLHPWLP SAMEGPTPVSALLHSSTMVVGIFLLVRFHPMTSNNPSILTMM
          LCLGALTTLFTAICALTQNDIKKIVAFSTSSQLGLMMVTLGINQPYLAFLHICTHAFF
          KAMLFMCSGSIHNLNDEQDIRKMGNNMKTMPFTSSCLTIGSLALTGMPFLTGFYSKD
          LIIEAINTCNTNAWALLITL IATSMTAIYSMRIYFVVMTKPRFPPLISINENPNLIT
          NPIKRLAFGSI LAGFFITYNIPPTNIQVLTMPWYKTTALLISILGFLLALELNNLTL

```

```

NLMSKKNYSYSSFSTSLGYFPSI I HRI IPNKTLNLSFKTSLNLLDLYWLEKSI PKSTS
TMHSYMSKLLTNQKGLVKLYFMSFLLSILLTTTLFI I NLEWFQ"
gene      complement (13536..14054)
          /gene="ND6"
CDS       complement (13536..14054)
          /gene="ND6"
          /codon_start=1
          /transl_table=2
          /product="NADH dehydrogenase subunit 6"
          /protein_id="ADG95710.1"
          /db_xref="GI:296100352"
          /translation="MTNYIFVLSLLFLTGCLGLALKPSP IYGGFGLIMSGFVGLMVL
GFGGSFLGLMVFLIYLGMMVVFVGYTTAMATEEYPETWGSNWFVFGFLIAGVLMEMFF
VYFLDYNGEEVDFDSLGDWLMYEIDDVGVMLEGGI GVAAMYSCATWMMVAVGWSLF
AGIFIIIETRD"
tRNA      complement (14055..14123)
          /product="tRNA-Glu"
gene      14127..15270
          /gene="CytB"
CDS       14127..15270
          /gene="CytB"
          /note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
          /codon_start=1
          /transl_except=(pos:15270,aa:TERM)
          /transl_table=2
          /product="cytochrome b"
          /protein_id="ADG95711.1"
          /db_xref="GI:296100353"
          /translation="MTNIRKTHPLFKI INHSFIDL PAPSNISSWWNFGSLLGLCLV I Q
ILTGLFLAMHYTSGTMTAFSSVTHICRDVNYGWLIRYMHANGASMF I CLFLHVGRGM
YYGSYAFMETWNI GVVLLFAVMATAFMGYVLPWGQMSFWGATVI TNLLSAIPYIGTTL
VEW I WGGFSVDKATL TRFFAFHFILPFI I AALVIVHLLFLHETGSNNPTGLNSDADKI
PFHPYYTIKDI LGIFIMIFLMTLVLFDPDLLGDPDNYTPANPLNTPPHIKPEWYFLF
AYA ILRSIPNKLGGV LALVLSILILALLPLLHTSKQRSLMFRPI TQMLYWI LVANLLV
LTWIGGQPVEYPFVI IGQLASISYFSIILIFMPISGI IENMLKWNL"
tRNA      15271..15339
          /product="tRNA-Thr"
tRNA      complement (15340..15406)
          /product="tRNA-Pro"
D-loop   15407..16260
ORIGIN
  1 gttaatgtag cttaaaacaa agcaaagcac tgaaatgct tagatggatt ttttaatccc
  61 ataaacacaa aggtttggtc ctggccttat aattagttgg aggtaagatt acacatgcaa
 121 acatccataa accggtgtaa aatccotaa atatttgtat aaaatttaag gagagggcat
 181 caagcacata ccatagctta agacgccttg cctggccaca cccccacggg actcagcagt
 241 gataaatatt aagcaatgaa cgaaagtttg actaagctat acctcttagg ggtggtaaat
 301 ttcgtgccag ccaccgcggt catacgatta acccaaaacta attacctctc ggcgtaaaac
 361 gtgttaacta taaaccata aatagaatta aaatccaact tataatgtgaa aattcattgt
 421 taggacctaa atccaataac gaaagtaatt ctaatattgt tatgttacac gatagctaag
 481 acccaagctg ggattagata cccactatg cttagcccta aacctcaata atttgacaac
 541 aaaattatth gccagagaac tactagctgt agcttaaac tcaaaggact tggcgtact
 601 ttatatccac ctagaggagc ctgttctata atcgataaac cccgctctac ctaccatct

```


661 cttgctaatt cagcctatat accgccatct tcagcaaacc ctaaaaagga actaaagtaa
721 gcacaagaat ttagcataaa aacgttaggt caaggtgtag ccaatgagat gggaagaaat
781 gggctacatt ttcttattaa agaacattta cgaaacctt tatgaaatta aaggataaag
841 gaggatttag tagtaaatta agaatagaga gcttagttga attgagcaat gaagtacgca
901 cacaccgcc gtcaccctcc tcaactaaa taaactaac tatacataat tcataaacia
961 aatttacgag aggagataag tcgtaacaag gtaagcatac tggaaagtgt gcttggata
1021 atcacagtgt agcttaacta caaagcatct ggcctacacc cagaagaatt cataataaat
1081 gaacactttg aactaatcct agocctcaac tcttaataat ataactatit cacactataa
1141 attaaacat ttaatctcgt aaaagtattg gagaaagaaa ttatttcata ggagctatag
1201 agtaagtacc gcaagggaaa gatgaaagaa taattaaaag taacaataag caagataaaa
1261 acctgtacc ttttgcataa tgagttaact agaataaatt taactaaaag cattagagct
1321 aataccocg aaaccaaag agctaccta aaacaattta aacgaatcaa cccgtctatg
1381 ttgcaaaata gtgggaagat ttttaggtag aggtgaaaag octaacgagc ttggtgatag
1441 ctggttacc aaataatgaa ttaagtta acttaaaatt taccaaaaga acatttaac
1501 ctcatgtaaa ttttaaatat agccaaaaga gggacagctc tttaggaaa gaaaaaac
1561 ttaagtagtg aataaactca ataaataacc aaaccattgt aggcccaaaa gcagccacca
1621 ataagaag cgttcaagct caacattagc atttaataa ttaacaaat aataataaat
1681 tcctacatta aaaattgggt taatctatta atgtatagat gagatactgt taacatgagt
1741 aacaagaact ttattctcca agcacaagcg tataacaact cggataacca ttgttaatta
1801 tcaacctata gataaaaact acaaatat agatctaaa acaacttgtt aatccaacac
1861 aggagtcta taaggaaaga ttaaaaaaa taaaaggaa cggcaaac aaaacccgc
1921 ctgtttacca aaaacatcac ctctagcata ttaagtatta gaggcattgc ctgcccagtg
1981 actaaagtt aacggccgag gtatcctgac cgtgcaaagg tagcataatc actgttctc
2041 taattagga ctagcatgaa cggctagac agggttgac tctctttat tttgatcag
2101 tgaattgac ctttcagtga agaggctgaa atactataat aagacgagaa gaccctatgg
2161 agctttaatt tatcagttta atcattctaa taaaattacc taatgggata accacaatga
2221 tataaactca taatttcggt tgggtgacc tgggaaaa gaaaatctc cgaatgattt
2281 taacatagat ccacaaatca aagtaagcaa tctatcttat tgacccaaaa atctttttg
2341 atcaacggac caagttacc tagggataac agcgaatcc tattaagag ttcataatga
2401 caattaggt ttacgacct gatgttgat caggacatcc caatggtgca gaagctatta
2461 atggttcgt tttcaacga ttaaagctc acgtgatctg agttcagacc ggagtaatcc
2521 agctcgttt ctatctatta acaattctc ccagtacgaa aggcaagag aatggagcc
2581 acctataaa taagcgcoc caacttaata tatgaattaa tctaaatata gtaagtatg
2641 taactcttg octagacaag gctattagg tggcagacc aggaaattgc gtaagactta
2701 aaacctgtt occagaggt caaatcctc tcctaagat gtattttatt aatacctta
2761 tattattagt cccgatcct attgccatag ctttttaac actagtagaa cgaaaaattt
2821 taggctacat acaactacga aaaggccta acgtttagg tccatacga gtcctcaac
2881 cattcgcaga cgctataaaa ctctttatta aagaacctat gcgcccccta actacctca
2941 tttactatt tattattgca ccaaccttt cactcacact agccctaagc ctatgaattc
3001 cactaccaat accacacca ctaatcaatc taaacctagg tctttattc attttagcta
3061 catcaagtct ttcagtatat tctatcttat ggtcaggatg agcatcaaac tcaaaatct
3121 cgtatattg tgcttacga gcagtagcac aaacaattc atatgaagt acaatagcta
3181 ttattctct atcagtttta ttaataagcg gctcattctc attacaaata cttatttata
3241 ccaagaaca catatgacta attttacct cctgaccaat agctataata tgatacatc
3301 caacctagc agaaacaaat cgagctccat tcgacctaac agaaggagaa tcagaactag
3361 tttccgggt caacgttgag taogctgcag gtcccttgc actattttt atagcggaat
3421 atacaaatat tattcttata aagccctaa catccattgt cttctgggc ccaatccata
3481 taattaacta cccagaacta tactcaatta atttataac agaaacta ttacttcaa
3541 caacattct atgaattcga gcatcatatc cccgtttcg ttatgacaa ctatacatc
3601 tcctatggaa aaattttct cccctacac tagcactatg tacatgacat atctccttac
3661 ctatctttat agcaagtgt ccaccataca tctagaaca tgtctgaca aagagttact
3721 ttgatagagt aaaatataga gtttaagcc ctctgtttc taggataata gaattgaac
3781 ctattcttaa gaattcaaaa ttctcgtgc taccataca octaatccta acatagtaag

3841 gtcagctaat taagctatcg ggcccatacc ccgaaaacgt tggtttaaac ccttcccgta
3901 ctaataaatc ctatcaccat catcattatt tacttcaacta tcctcattgg tcccgtcatc
3961 acaatatcaa gctccaacct tctactcata tgagtaggat tagaactaag ttactagca
4021 attattcccc tattaattaa caaaaaaat ccccgatcaa ctgaagcagc aacaaaatac
4081 tttattacac aagcaacagc ctcaataatt atccttttag ctattgtcct aaattttaaa
4141 caactcggac tatgaacatt tcaacaacaa acaaacagcc tcctactcaa cattacacta
4201 attgcactat caatgaaact agggctagcc ccattccact tttgactccc agaagtaaca
4261 caaggaatcc cactacacat aggactaatc ctctcacat gacaaaaaat tgcccacta
4321 tcaatcctat ttcaatttta tcagctactt aatccaacca tcattatgat cctagcaatt
4381 tcattccattc ttatcggago atgaggtgga cttaaccaaa cacaaatgag aaaaatcata
4441 gcatactcat caatcgctca cataggatgg atattagcta tcctgoccta caaccact
4501 ataacacttc ttaacttact aatttatatt atcctcactg ttcccatatt cctcactt
4561 atacttaatg catcaacaac tatcaactca atctcactcc tatgaataa agctccaaca
4621 acccttgtat taattccatt aatcctcctc tccttaggag gtctaccccc actaacagga
4681 tttttaccaa aatgagctat tatcacagaa ctactaaaaa acaactgcct aattattacc
4741 acaataatag ctattatagc cctattaaac ctatttttct acacacgtct gatctactct
4801 acttctctca ccataatttc aactaataat aactcaaaaa tattagccca ccatttaa
4861 tataaaaaa atattatcct tccatcactt actatcatca gcacattaat tctcccctca
4921 tcaccacaac taatcattta gaagtttagg atatatgcag tccaagagcc ttcaaagccc
4981 taagaaaaa aacaagtta acttctgata aggactgtaa gactacttcc tacatctatt
5041 gaatgcaaat caatcgcttt aattaagcta agaccttact agattggtag gaattaaacc
5101 tacgaaactt tagttaacag ctaaataccc tacttctggc tccaatctac ttctcccgcc
5161 tatcagaaaa ggggcgggag aagccttagt agatataatg ctctacccct tcgaatttgc
5221 aattcgacat gactaacacc ttaaggcttg gtaaaaagag gacttaaacc tctgtattta
5281 gatttacagt ctaatgctta ctacagccatt ttacctatgt tcattaatog ttgattattc
5341 tcaactaacc acaagacat tggcacccta tacctcctat tcggagcttg agccggaata
5401 gtagggactg cattgagcat tctgatccga gctgaactcg gacaaccagg tgcacttta
5461 ggggacgacc agatctataa tgcattgttt accgcccag cgtttgttat aatcttcttc
5521 atagtaatac caataataat tgggggcttt ggaactgac tcgtaccact tatgattgga
5581 gctcccagata tggcattccc acgaataaat aatataagct tctgacttct accaccatct
5641 ttctactccc tattagcacc atoccatagtt gaagcaggag ccggtacagg atgaacagta
5701 ttccaccctt tagctggaac cttagcccat gctggagcat cagtagacct gacaatttcc
5761 tctcttact tagccggtgt gtctcaatc ctaggagcta ttaactttat taccactatt
5821 attaatataa aaccaccagc tataaccocaa tatcaaacac cactgtttgt atgatctgta
5881 ctaattacag ctgttctact cctattatcc ctcccagcac tagcagcagg aattacaata
5941 ctattaactg accgaaacct aaacacaact ttctttgacc ctgcaggagg tggagacca
6001 attctatatac aacatctatt ctgattcttt ggccaccocg aagtttatat ccttatcctg
6061 cctgggttcg gaattatctc tcatgtagtt acttattact ccggtaaaaa agaaccattc
6121 gggtagatgg gaatagtttg agctataata tcaattggct tcctaggttt tattgtatga
6181 gcacatcaca tgtttacagt aggccttagat gtagacacac gagcctactt tacatctgcc
6241 acaataatta tcgctatccc aacaggcgtta aaagtattta gctgacttgc aacactacac
6301 ggaggtaata tcaaatgac tccagctata ctttgagccc taggctttat ctctctattt
6361 acagtaggtg gattaacagg aattgtcctc tcaaacctct cactcgacat tgtacttcat
6421 gacacatatt atgtagtagc ccatttccat tatgtactgt ctataggagc agtatttct
6481 attatagctg gttttgtcca ttgattcccc ctattcacag ggtataccct agacgatatg
6541 tgagcaaaaa cacactttgc cattatattt gtaggagtta acataacatt ctccctcaa
6601 cttttctag gactctcggg cataccacga cgatactcag actatocaga tgcattatcc
6661 acatgaaata cagtatcaco tataggatca tttatttctc ttacagccgt cctactaata
6721 atttttatga tctgagaago attcgttca aaacgagaag ttctaacagt atottattct
6781 tcaaccaact tagagtact acatggctgc ccacctcctt atcacacatt cgaagagcct
6841 tcttacgtaa aagtaaaata agaaaggaag gactcgaacc ccctaaaact gtttcaagc
6901 caacctcata gcctctatgt ctttctcaat gagatattag taaaacaatt acataacct
6961 gtcagggtta aattatagac ttaaatctat atatcttaca tggcttatcc attccaatta

7021 ggtctacaag atgccacatc ccctattata gaagaattaa taaacttcca tgaccacaca
7081 ctaataattg ttttccttat tagcacotta gtactttaca ttatttctact gatattgaca
7141 acagaactta cacacacaag cacaatagat gcccaagaag tagagactgt ttgaactatc
7201 ttaccagctg ttattccttat ttttaattgca cttccctcac tacgaatttt atatatgata
7261 gacgaaatta ataaccccggt actaacagta aaaactatag gccaccaatg atactgaagc
7321 tatgaatata ctgactatga agacctatgc tttgattcctt acataattcc cacaactgac
7381 ttaaaaccag gtgacctagc actattagaa gttgacaacc gagtagtttt accaatagga
7441 cttcccatcc gaatgottat ctcacagaa gacgttctcc actcatgagc tgttccctca
7501 ctatgactaa aaactgacgc tatcccagga cgctaaatc aagcaacagt aacctctaac
7561 cgtccaggac tattttacgg tcaatgctct gaaatctgag ggtctaatca cagcttcaata
7621 cctattgttc tagaaatagt gccactaaaa cacttcgaaa attgatcagc ctcaataatt
7681 taatctcact gtgaagctta gagcgttaac cttttaagtt aaagttagag accttaaatc
7741 tccatagtaga tatgcccaca ctagatacat ccacatgatt tattaccatc atctcatcta
7801 tggctacact atttatttta ttcaactaa aaatttcctc ccaatctttt cccactcctc
7861 catcccaaaa aaccttttaca gcacaagaaa caaaaactcc ttgagaatca aatgaacga
7921 aaatctattht gcctctttta tcaactcctc agtaatagga ctcccaattg taattactat
7981 tatcatattht ccatctatcc tattcccac c atcagaacgt ttaattagta atogacttca
8041 ctcatthtcaa cactgactaa tcaaaacttat tatcaaaaca ataacttca ttcacacacc
8101 aaaaggacgg acttgaactc taataattgt ctcccttatt atattcatcg gotcaactaa
8161 ccttctagcc cttcttccac acacatttac accgactact caactatcaa taatcttag
8221 tatagcaatc ccactatgag caggagcagc cctattaggg tttcgcata aattaaag
8281 ctcccttagct catttctctc cacaaggcac ccctatttca cttattccta tacttattat
8341 tatcgaaaca atcagtctat ttatccaacc tatagctcta gcagtagcat taacagccaa
8401 tattaccgca ggacatctac ttatacatct aattggagga gcaaccctag tcttataaaa
8461 tattagccca cccaccgcta ctattacatt cattatctta ctcttctca cagtactaga
8521 gtttgcgtga gccctaattc aagcatacgt attcactctt ctagtaagcc tatactaca
8581 tgataacaca taatgacca ccaaacacat gcctatcata tagttaacc aagtcacatga
8641 ccacttacag gagcatttct agoccttcta cttacatctg gattagtaat atgatttcc
8701 tatagctcaa ccattctctc gtctctagga ctactagcca atacaactac tatatatcag
8761 tgatgatgag atattattctc tgaaggaacc tatcaaggcc atcataccct tattgtaca
8821 aaaggctctc gatacggatc aatcttgctc attgtctctg aagtattttt ctttgcctgc
8881 ttcttctgag ctttctacca ttcaagccta gtaccaacc cagatctcgg aggttgcctga
8941 cctccaacag gaattactcc tcttaactct ctagaagtc cactattaa tacatcagtg
9001 ctactacatc caggtgttct aatcacatga gccaccata gtctcataga aggcaaacga
9061 aatcacataa atcaagccct actaatcact attctattag gactttactt caccactctt
9121 caagcctcag aatattttga aactccattc tctatttcag acggaattta cggctccacc
9181 ttctttatag caacaggatt ccacgggctc catgtaatta ttggaactac attcctcgtg
9241 gtttgcctcc tacgacaatt aaaatttccac ttacatcaa agcaccattt tggatttga
9301 gcagcagcat gatattgaca cttttagtag gtatgtctgac tcttctata tttatccatc
9361 tattgatgag gatcttactc ttttagtata attaatataa ctgacttcca attagtagat
9421 tctgaacaac atcagaagag agtaataaat cttataatta tcatatttat caatactatc
9481 ctatcattaa tccctaattc agttgcattc tgactcccc aaataaactt atcacagaa
9541 aaagccaatc catatgaatg tggatttgac ccatcaagct ctgctcgttt accattttcg
9601 ataaaatttt tcttagtagc tattacattc ctattatttg atctagagat tgcctacta
9661 cttctctctc catgagcaat ccaaacaact aataccaatg taatagtaac ctgacttca
9721 attttagtca ccattctatc attaggccta acctatgaat gaactcaaaa aggactcga
9781 tgaacagaat aactggtaat tagtttaaat aaaattaatg atttgcactc attagattat
9841 gatcttattc ataattacca aaaatgtcat ctgcctttat taatcttaca ctgacttca
9901 tactttctact tctaggaaca ttaacgttctc gttctcactt aatgtctact cttctatgct
9961 tagaaggaaat aatgttatcc ttatttatta taacatccct agcttccctg aatttcaact
10021 ctatagcatc tataccatt ccaattacta tcttagtatt tgcagcatgc gaagcagcag
10081 tagggctagc cctattagt t aaagtgtcca acacctatgg taccgattac gttcaaac
10141 tcaatttact tcaatgctaa aaattatctt tccatcactc atacttctgc cattaacatg

10201 attatcaaaa cggaaaaaaa cctgaataaa cgtcacttcc tacagcttca tgattagcct
10261 tattagcctt tcaatcctat gacaaactga cggaagccct cttaactttt caactatatt
10321 ctctcagat cctctatcat ccccacttat tattttaaca acttgactat tacccttat
10381 acttatagct agtcaaaatc acctaaaaaa agaaaacttc acctaccaaa aactctatat
10441 ttcaatacta gttagcctac aaattcttct agtaataact ttttcagcaa ctgagcta
10501 catattttat attctatttg aagctaccct tattctaca ttaattatta ttacacgatg
10561 aggtaatcaa acagaacgat taaacgcagg actttacttc ttattttaca ccctcatcgg
10621 ctctatccca ctcttaatcg cccttatctt tatccaaaac tccataggtta cactcaactt
10681 taccatacta tcactaaca ctaaccccct gaacccotta tgatccaaca acatcttatg
10741 actagcatgc ataatagcat ttataattaa aatacctcta tatggagtto acctatgact
10801 cccaaaagcc catgttgaag caoccatgagc aggatccata atcctagcgg caattctctt
10861 aaaactagga agttacggaa taatacgaat tgccatcatt ctggaccacac taacaaaata
10921 catagcctac cctttcattt tactatocct atggggcata atcataacaa gotcaatttg
10981 cctacgcaa acagacctaa aatcccta atcgttattcc tcagttagcc atatagcttt
11041 agtaattgca tcaattataa ttcaactcct ttggagcttt ataggagcta caatattaat
11101 agttgctcac ggcctaactt catcacttct attctgctta gcaactcca actatgaacg
11161 aattcacagc ogaactataa ttatagoccg aggacttcaa ataacttttc cactcatagc
11221 aacttgctgg ttagtagcaa gcctagctaa tcttgctota ccaccatcaa tcaatctcat
11281 aggggaacta tttattacta tatcactatt ttcttgatcc aacttttcaa ttatccta
11341 aggagtcaac attattatta caggcatata ctcaatata ataattatca caaccacagc
11401 aggaaaacta acaagtcaca taacaactct ccaaccttcc cacacccgag aactaacact
11461 tataattctc cacatcatcc ccctagtcct attaactatt aacccaaac taattacagc
11521 gttgacaata tgtaaatata gtttcaaaa acattagact gtgaatctaa aaacaggagt
11581 tcaacacttc ttatttacca agaaagcacg caagaactgc taactcatgc ccccatgttt
11641 aaccacatgg ctttcttact tttataggat agaagtaact cattggtctt aggagccaaa
11701 aaccttggtg caattccaaa taaaagtaat aaatcattt acatcattta ttatagcaat
11761 ttttctcatt cttatctcac caattttaat ctctcactt agtctattca aacatattaa
11821 ttttccaata tatgcaaac tatcaatcaa gctctcattc tttctaagcc tcttacctct
11881 cataatatta ttttatcaca acacagaata cataattact acctgacatt gaatcacaat
11941 caattctatt gaactcaca taagctttaa aatagattat ttctcaacta tattttatc
12001 cgtagcccta ttctgtacat ggtctatcat acaattttcc tcatgatata tacactcaga
12061 tccctcacatt aatcgattta ttaaactctt actattattt ttaactacta tactaatttt
12121 aacttcagct aacaatctat ttcaactatt tatcgggtgg gaaggagtag gtattatc
12181 tttcctctc atcggatgat gatctggacg agcagacgca aataccgagc ctctcaagc
12241 aattttatat aaccgaatcg gtgatcgg ctttatctta gctataacat gattttgtct
12301 aaacataaac tcctgagaac ttcaacaaat tctattaacc aataataata atctaattcc
12361 actgctaggt ttactaattg cagccacagg aaaatcagcc caattcggcc ttcattcatg
12421 acttccatca gccatagaag gacccacacc agtttcagcc ctactacact caagcacaat
12481 agtagtagca ggaatttttc tactagtctg gtttcacctt ataactcca acaaccatc
12541 aatcctaact ataatactat gtctaggagc actaactacc ctatttaccg ctatctgtgc
12601 cctcacccaa aacgatatta aaaaaattgt agcattttcc acatcaagtc agctaggcct
12661 aataatagtt acattaggaa ttaaccaacc atatctagca ttcttacata tctgcacaca
12721 tgcattcttc aaagccatc tttctatg ctccgatca attattcaca atctaaccga
12781 tgaacaagac attcgaaaa taggaaacat aataaaaact atgccattca catcatcctg
12841 cctaacaatt ggtagtctg ccctaacagg aatacattc ctaacaggtt tttactcaaa
12901 agacctaatc attgaagcta tcaatacctg caataccaac gctgagccc tactaattac
12961 actaatgcc acatcaata cagctatcta cagcatacga attatttact ttgtagcat
13021 aacaaaacca cgattccac cattaatctc aatcaatgaa aataoccaa atcttactaa
13081 ccctattaaa cgcctagcat ttggaagtat cctagctggg tttttcatta catataatat
13141 cccaccaact aacatccaag ttcttactat accatgatac ttaaaaacta cagccctgct
13201 aatctctatc ctaggctttc ttctagccct agaactaat aacctacat taatctatc
13261 cataagtaaa aaaaactcat atctatcatt ttccacatca ctaggctact tccatcaat
13321 tatccatoga attatcccaa aaaaacact taacctagc ttcaaacat cacttaacct

13381 tctagatctg tattgactag aaaaatcaat cccaaaatct acttcaacca tacactotta
13441 tatatcaaaa ctactaacta accaaaaagg actagtaaaa ctatatTTTA tatcattttt
13501 actttctatt ctactaacca ctacctctt tatcattaat ctcgagtgat ttcaataata
13561 ataaaaattc cagcaaaaa agaccatcca gcaactacta ttatccaagt agcacaacta
13621 tacatcgag caaccocaat accacctctt aacattacac caacatcctc aatctcatic
13681 atcaatcaat cacctagact atcaaaaatca actacctctt cccattata ataatcaaga
13741 aaatacacia aaaacatttc cattaaaacc cctgcaatta aaaaaccaa aacaaaacca
13801 ttagaccctc aggtctcagg atactctca gtagctatag ctgtggata cccaaact
13861 actattattc ctcccaata aattaaaaat accattagac ctaaaaagga tccaccaac
13921 cctaaaacta ttagacaacc caaaaacca ctcataatta aaccaaacc cccataata
13981 ggagaaggct ttaaagctaa ccaagacia ccagtcaaaa ataataaact taaaacaaa
14041 atataattag tcattatttc tacacagcat tcaactgtga ctaatgacat gaaaaatcat
14101 cgttgtaatt caactacaga aacctaatga caaacatccg aaaaactcac ccctattta
14161 aatcattaa ccactcttc atogacctcc ctgccccatc taacatctca tctgtatgaa
14221 actttggctc cctcctaggt ctatgcctcg taattcaaat tcttacaggc ttattcctag
14281 ccatacacta cacatcaggc acaataacag cattctcctc agttacacat atttgcogag
14341 acgtaacta tggatgacta attcgtatata tacacgctaa tggagcttca atatttttta
14401 tctgcttatt tctccatgta ggacgaggaa tgtactacgg atcctatgca ttatagaaa
14461 catgaaatat oggagtagtc ctattatttg cagtcatagc tacagcattc ataggctatg
14521 tacttccgtg aggacaaata tcctctgag gagcaacagt aattacaaat ctctctcag
14581 ccattccata catcggcact accctggtag aatggatttg aggaggattc tcagtatata
14641 aagccacttt aacacgtttc ttgcattcc attttattct cccattcatt atcgcagccc
14701 tggtaatcgt ccattctccta tttctccacg aaacaggctc aaacaacca acaggtttta
14761 actcagacgc cgataaaaac ccatttcacc catactacac aattaagat atcctaggca
14821 ttttcattat aattatattc ctaataaacc tggctctatt ctccccggac ctacttggag
14881 acccagataa ttacacacca gcaaatccac ttaatacacc accacatctc aaaccagaat
14941 ggtactttct atttgcatat gcaattctac gctctattcc aaacaacta ggaggagtcc
15001 tagctctagt cctatccatc cttatcctag cttactgcc acttcttcc acttcaaac
15061 aacgaagcct aatattcctg cctattactc aaatactcta ctgaatccta gtagtaacc
15121 ttctcgtctc tacctgaatc ggoggacaac cagttgaata cccattccta atcatcggcc
15181 aactcgcctc catcagctac ttctcaatta tcctaattct tatgccaatc tcaggcatca
15241 ttgaaaataa tatactaaa tgaaacctat gccctgatag tatagaacca ttactctggt
15301 cttgtaaac aaaaacgaag aacttctctt ctccaggcat caagaagaag gaaatcctc
15361 ccaccatcag caccocaaagc tgatattcta cttaaactac ttcttgagta cataaaatta
15421 catagtacat tcatacatat atgtatagc tacattaaat tatttaccoc tagcatataa
15481 gcaagtacat ttaatcaatg atcctagaca tcaatggta accaaatata tttctcctc
15541 acatgactat tcaagactta cattttaatt aatgatctta agacatctc gcgttatctt
15601 acatacacca ttcagtcata aactcttctc ttccatagc ctatccctct ccacatttgg
15661 tctatatatc taccatcctc cgtgaaacca acaaccgccc cacctatgcc cctctctcag
15721 ctccgggccc ataaaacttg gggtagctc acctgaaact ttatcagaca tctggttctt
15781 acttcagggc catcaaatgc gtatctgccc atactgtccc cttaaaatag acatctogat
15841 ggtacgggtc taatcagcca agaccgtcat aactgtggtc tcggcaggtt ggtattttt
15901 tatttttagg atgctgtgac tcaccatagc cgtcaaggca tgaaggctcag cccaccatga
15961 agctggactt acggtgaagg atcatttctc cacataaac accaccgaga caatttatga
16021 atgcttgta gacataactc tattactact tataatttta acttaacca acccccttc
16081 cccctttgat gtcaaacccc aaaacctca aagactaac ttaaatttca taagttttat
16141 tctattctag tagttcacia aattaaactt atattacagt attaggcaaa atttataaa
16201 aaattaattt tgaactatca aaaactcacc ttattcaaat tccctaataa atttatataa

Appendix 2. The sequence of mitochondrial genome of *Apodemus chejuensis*

LOCUS HM034867 16261 bp DNA circular ROD 30-DEC-2011
 DEFINITION *Apodemus chejuensis* mitochondrion, complete genome.
 ACCESSION HM034867
 VERSION HM034867.1 GI:296100354
 KEYWORDS .
 SOURCE mitochondrion *Apodemus chejuensis* (Jeju striped field mouse)
 ORGANISM *Apodemus chejuensis*
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; *Apodemus*.
 REFERENCE 1 (bases 1 to 16261)
 AUTHORS Oh, D.-J., Jung, Y.-H., Kim, T.-W., Chang, M.-H., Han, S.-H., Oh, H.-S.
 and Kim, S.-J.
 TITLE Biogeography and speciation of Korean striped field mice *Apodemus*
agrarius and *A. chejuensis* inferred from mitochondrial DNA
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 16261)
 AUTHORS Oh, D.-J., Jung, Y.-H., Kim, T.-W., Chang, M.-H., Han, S.-H., Oh, H.-S.
 and Kim, S.-J.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2010) Department of Biology, Jeju National
 University, 66 Jejudaehakno, Jeju 690-756, Republic of Korea
 FEATURES Location/Qualifiers
 source 1..16261
 /organism="Apodemus chejuensis"
 /organelle="mitochondrion"
 /mol_type="genomic DNA"
 /specimen_voucher="JBRI-Mam-002"
 /db_xref="taxon:754351"
 tRNA 1..67
 /product="tRNA-Phe"
 rRNA 68..1022
 /product="12S ribosomal RNA"
 tRNA 1023..1091
 /product="tRNA-Val"
 rRNA 1091..2663
 /product="16S ribosomal RNA"
 tRNA 2664..2738
 /product="tRNA-Leu"
 /note="codons recognized: UUR"
 gene 2739..3693
 /gene="ND1"
 CDS 2739..3693
 /gene="ND1"
 /note="TAA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /codon_start=1
 /transl_except=(pos:3693, aa:TERM)
 /transl_table=2
 /product="NADH dehydrogenase subunit 1"

```

/protein_id="ADG95712.1"
/db_xref="GI:296100355"
/translation="MYFINILMLLVPIIIAMAFLLVERKILGYMQLRKGPNVVGPYGV
VLQPFADAMKLFIKEPMRPLTTSISLFI IAPTLSLTLALSLWIPLPMPHPLINLNLGV
LFI LATSSLSVYSILWSGWASNSKYSLFGALRAVAQTSYEVTMAI ILLSVLLMSGSF
SLQMLIYTQEHMWLILPAWPMAMMWYI STLAETNRAPFDL TEGESELVSGFNVEYAAG
PFALFFMAEYTN IILMNALTSIVFLGPIHMINYPELYSINFMETELLLSTTFLWIRAS
YPRFRYDQLMHLLWKNFLPLTLALCTWHI SLP IFMASIPPYI"
tRNA 3694..3762
/product="tRNA-Ile"
tRNA complement(3760..3830)
/product="tRNA-Gln"
tRNA 3835..3903
/product="tRNA-Met"
gene 3904..4939
/gene="ND2"
CDS 3904..4939
/gene="ND2"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:4939,aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="ADG95713.1"
/db_xref="GI:296100356"
/translation="MNPITIIIIYFTILIGPVI TMSSSNLLLMWVGLLELSLLAI IPLL
INKNPRSTEAATKYFI TQATSMI ILLAI IILNFKQLGLWTFQQQTNSLLLNITLIAL
SMKLG LAPFHFWLPEVTQGI PLHMGL ILLTWQKI APLSILFQFYQLLNPTI IMILAIS
SILIGAWGGLNQTQMRKI VAYSSI AHMGWMLA ILPYNPTMTLLNLLIYI ILTVPMFLM
LMLNASTTINSISLLWNKAPTTLVLTPL ILLSLGGLPPLTGFLPKWAI I TELLKNNCL
IITTMMAIMALLNLFYTRLI YSTSLTMFPTNNNSKMLTHHLNKNKI I LPSLTIIST
LILPLSPQLII"
tRNA 4940..5007
/product="tRNA-Trp"
tRNA complement(5009..5077)
/product="tRNA-Ala"
tRNA complement(5079..5149)
/product="tRNA-Asn"
rep_origin 5150..5180
/note="L-strand replication origin"
/direction=LEFT
tRNA complement(5181..5248)
/product="tRNA-Cys"
tRNA 5249..5315
/product="tRNA-Tyr"
gene 5317..6861
/gene="C01"
CDS 5317..6861
/gene="C01"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit 1"

```

```

/protein_id="ADG95714.1"
/db_xref="GI:296100357"
/translation="MFINRWLFSTNHKDIGTLYLLFGAWAGMVGTAISILRAELGKP
GALLGDDQIYNVIVTAHAFVMIFFMVMPMMIGFGNWLVPMLI GAPDMAFPRMNNMSF
WLLPPSFLLLLASSMVEAGAGTGWTVYPPLAGNLAHAGASVDLTI FSLHLAGVPSILG
AINFITTI INMKPPAMTQYQTPLFVWSVLI TAVLLLLSLPVAAGITMLLTDRLNNTT
FFDPAGGGDPI LYQHFWFFGHPEVYILILPGFGI ISHVVTYYSGKKEPFGYMGMVWA
MMSIGFLGFIVWAHMFVGLDVDTRAYFTSATMIIA IPTGVKVF SWLATLHGNI KW
SPAMLWALGFI FLFTVGGTGI VLSNSSLDIVLHDTYYVVAHFHYVLSMGAVFAIMAG
FVHWFPFLTGYTDDMMAKTHFAIMFVGVNMTFFPQHFLGLSGMPRRYSYDYPDAYTTW
NTVSSMGSFISLTAVLLMI FMIWEAFASKREVLTVSYSSTNLEWLHGCPPPYHTFEEP
SYVKVK"
tRNA complement(6859..6927)
/product="tRNA-Ser"
/note="codons recognized: UCN"
tRNA 6931..6998
/product="tRNA-Asp"
gene 7000..7683
/gene="C02"
CDS 7000..7683
/gene="C02"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit 2"
/protein_id="ADG95715.1"
/db_xref="GI:296100358"
/translation="MAYPFQLGLQDATSPIMEELMNFHDHTLMIVFLISTLVLYIISL
MLTTKLTHTSTMDAQEVETVWTILPAVILILIALPSLRILYMMDEINNPVLTVKTMGH
QWYWSYEYTDYEDLCFDSYMIPTTDLKPGDLRLLEVDNRVLPMELPIRMILSSEDVL
HSWAVPSLGLKTDAPGRLNQATVTSNRPGLFYGGQCEICGSNHSFMPIVLEMVPLKH
FENWSASMI"
tRNA 7687..7750
/product="tRNA-Lys"
gene 7752..7955
/gene="ATP8"
CDS 7752..7955
/gene="ATP8"
/codon_start=1
/transl_table=2
/product="ATP synthase FO subunit 8"
/protein_id="ADG95716.1"
/db_xref="GI:296100359"
/translation="MPQLDTSTWFI TIISSMATLFI LFQLKISSQSFPPTPPSKTFTA
QETKTPWESKWTKIYLPLLSLPQ"
gene 7913..8593
/gene="ATP6"
CDS 7913..8593
/gene="ATP6"
/codon_start=1
/transl_table=2
/product="ATP synthase FO subunit 6"
/protein_id="ADG95717.1"
/db_xref="GI:296100360"

```



```

/translacion="MNENLFASFITPSVMGLPIVITIIMFPSILFPSSERLISNRLHS
FQHWLIKLIKQMMLIHTPKGRTWTLMI VSLIMFISSTNLLGLPHTFTPTTQLSMNL
SMAIPLWAGAVLLGFRHKLKSSLAHFLPQGTPI SLIPMLIIETISLFIQPMALAVRL
TANITAGHLLMHLIGGATLVLMNISPPTATITFII LLLLTVLEFAVALIQAYVFTLLV
SLYLHDNT"
gene      8593..9376
          /gene="CO3"
CDS       8593..9376
          /gene="CO3"
          /note="TAA stop codon is completed by the addition of 3' A
          residues to the mRNA"
          /codon_start=1
          /transl_except=(pos:9376, aa:TERM)
          /transl_table=2
          /product="cytochrome c oxidase subunit 3"
          /protein_id="ADG95718.1"
          /db_xref="GI:296100361"
          /translation="MTHQTHAYHMVNPSPWPLTGAFSALLLTSGLVMWFHYSSTILLS
          LGLLANTLTMYQWWRDI IREGTYQGHHTPIVQKGLRYGMILFIVSEVFFFAGFFWAFY
          HSSLVPTHDLGGCWPTGI TPLNPLEVPLLNTSVLLASGVSITWAHHSLEMEGKRNMN
          QALLITILLGLYFTMLQASEYFETPFSISDGIYGSTFFMATGFHGLHVIIGTTFLLVVC
          LLRQLKFHFTSKHHFGFEAAAWYWHFVDVWVFLYVSIYWWGS"
tRNA      9377..9444
          /product="tRNA-Gly"
gene      9445..9792
          /gene="ND3"
CDS       9445..9792
          /gene="ND3"
          /codon_start=1
          /transl_table=2
          /product="NADH dehydrogenase subunit 3"
          /protein_id="ADG95719.1"
          /db_xref="GI:296100362"
          /translation="MNLMIIMFINTILSLILISVAFWLPQMNLYTEKANPYECGFDP
          SSARLPFSMKFFLVAITFLFDLEIALLLPLPWAIQTTNTNVMMSALILVTILSLGL
          TYEWTQKGLEWTE"
tRNA      9794..9861
          /product="tRNA-Arg"
gene      9864..10160
          /gene="ND4L"
CDS       9864..10160
          /gene="ND4L"
          /codon_start=1
          /transl_table=2
          /product="NADH dehydrogenase subunit 4L"
          /protein_id="ADG95720.1"
          /db_xref="GI:296100363"
          /translation="MSSAFINLTLAFMLSLLGTLTFRSHLMSTLLCLEGMMLSLFIMT
          SLASLNSNSMASMPIPIITILVFAACEAAVGLALLVKVSNTYGTDYQNLNLLQC"
gene      10154..11531
          /gene="ND4"
CDS       10154..11531
          /gene="ND4"

```

```

/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:11531,aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 4"
/protein_id="ADG95721.1"
/db_xref="GI:296100364"
/translation="MLKIIFPSLMLLPLTWLSKPKKTWMNVTSYSFMISLISLSILWQ
TDGSPLNFSMTFSSDPLSSPLIILTTWLLPLMLMASQNHKKENFTYQKLYISMLVSL
QILLVMTFSATELIMFYILFEATLIPTLIIITRWGNQTERLNAGLYFLFYTLIGSIPL
LIALIFIQNSMGTLNFTMLSLTTNPLNPLWSNNILWLACMMAFMIKMPLYGVHLLWPK
AHVEAPIAGSMILAAILLKLGSGMMRIAII LDPLTKYMAYPFILLSLWGMIMTSSIC
LRQTDLKSLEYSSVSHMALVIASIMIQTWFSFMGATMLMI AHGLTSSLLFCLANSNY
ERIHSTRTIMARGLQMI FPLMATCWLVASLANLALPPSINLMGELFITMSLFSWSNFS
IILMGVNIITGMYSMYMIITQRGKLTSHMNNLQPSHTRELTMLIHLI IPLVLLTIN
PKLITGLTM"
tRNA 11532..11599
/product="tRNA-His"
tRNA 11600..11658
/product="tRNA-Ser"
/note="codons recognized: AGY"
tRNA 11658..11728
/product="tRNA-Leu"
/note="codons recognized: CUN"
gene 11729..13558
/gene="ND5"
CDS 11729..13558
/gene="ND5"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 5"
/protein_id="ADG95722.1"
/db_xref="GI:296100365"
/translation="MNIITSFIMAIFLILMSPILISSTSLFKHINFPMYATLSIKLSF
FLSLLPLMMLFYHNTEYMITTWHWITINSIELTMSFKMDYFSTMFLSVALFVTSIMQ
FSSWYMHSDPHINRFIKYLLFLITMLILTSANNLFQLFIWEGVGIMSFLLIGWWFG
RADANTAALQAILYNRIGDIGFILAMTWFCNMNSWELQQILLTNNNLIPLGLLIA
ATGKSAQFGLHPWLP SAMEGPTPVSA LQHSSTMV VAGIFLLVRFHPMTSNNPSILTMM
LCLGALTTLFTAICALTQNDIKKIVAFSTSSQLGLMMVTLGINQPYLAFLHICTHAFF
KAMLFMCSGSIHSLNDEQDIRKMGNNMKTMPFTSSCLTIGSLALTGMPFLTGFYSKD
LIEAINTCNTNAWALLITL IATSMTAIYSMRIIFVVMTKPRFPPLISINENPDLT
NPIKRLAFGSI LAGFFITYNIPPTNIQVLTMPWYLKTTALLISILGFLLALELNNLTL
NLSMSKKNYSYSSFSTSLGYFPSIHRIPNKTLNLSFKTSLNLLDLYWLEKSI PKSTS
TMHSYMSKLLTNQKGLVKLYFMSFLLSILLTTTLFIINLEWFQ"
gene complement(13536..14054)
/gene="ND6"
CDS complement(13536..14054)
/gene="ND6"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 6"
/protein_id="ADG95723.1"

```

```

/db_xref="GI:296100366"
/translation="MTNYIFVLSLLFLTGCLGLALKPSPYGGFGLMSGFVGCMLMVL
GFGGSFLGLMVFLIYLGMMVVFYTTAMATEEYPETWGSNWFVFGFLIAGVLMEMFF
VYFLDYNGEEVDFDSLGDWLMYEIDDVGMLEGGI GVAAMYSCATMMM VVAGWSLF
AGIFIIIEITRD"
tRNA      complement(14055..14123)
          /product="tRNA-Glu"
gene      14127..15270
          /gene="CytB"
CDS       14127..15270
          /gene="CytB"
          /note="CytB; TAA stop codon is completed by the addition
of 3' A residues to the mRNA"
          /codon_start=1
          /transl_except=(pos:15270,aa:TERM)
          /transl_table=2
          /product="cytochrome b"
          /protein_id="ADG95724.1"
          /db_xref="GI:296100367"
          /translation="MTNIRKTHPLFKIINHSFIDL PAPSNISSWWNFGSLLGLCLVIQ
ILTGLFLAMHYTSDTMTAFSSVTHICRDVNYGWLIRYMHANGASMFFICLFLHVGRGM
YYGSYAFMETWNI GVILLFAVMATAFMGYVLPWGQMSFWGATVINLLSAIPYIGTTL
VEW IWGGFSVDKATL TRFFAFHFILPFI I AALVIVHLLFLHETGSNNPTGLNSDADKI
PFHPYYTIKDI LGIFIMIMFLMVLVLFPPDLLGDPDNYTPANPLNTPPHIKPEWYFLF
AYA ILRSIPNKLGGVLALVLSILILALLPLLHTSKQRSLMFRPITQMLYWILVANLLV
LTWIGGQPVEYPFVIGQLASISYFSIILIFMPISGI IENMLKWNL"
tRNA      15271..15338
          /product="tRNA-Thr"
tRNA      complement(15339..15404)
          /product="tRNA-Pro"
D-loop    15405..16261
ORIGIN
1  gttaatgtag cttaaacaa agcaaagcac tgaaaatgct tagatggatt ttttaatccc
61  ataaacacaa aggtttgctc ctggccttat aatagttgg aggtaagatt acacatgcaa
121 acatccataa accggtgtaa aatcccttaa atatttatac aaaatttaag gagagggcat
181 caagcacata ccatagctta agacgccttg cctggccaca cccccacggg actcagcagt
241 gataaatatt aagcaatgaa cgaagtttg actaagctat acctottagg gttgtaaat
301 ttcgtgccag ccaccgpgt catacgatta acccaaacta attacctctc ggcgtaaac
361 gtgttaacta taaaccata aatagaatta aaatccaact tataatgtgaa aattcattgt
421 taggacctaa atccaataac gaaagtaatt ctaatattgt tatgctacac gatagctaag
481 acccaaactg ggattagata cccactatg cttagcccta aacctcaata atttgacaac
541 aaaattatit gccagagaac tactagctat agcttaaac tcaaaggact tggcgggtact
601 ttatatccac ctagaggagc ctgttctata atcgataaac cccgctctac ctcaccatct
661 cttgctaatt cagcctatat accgccatct tcagcgaacc ctaaaaagga actaaagtaa
721 gcacaagaat ttagcataaa aacgttaggt caaggtgtag ccaatgagat gggaagaaat
781 gggctacatt ttcttattaa agaacattta cgaaccctt tatgaaatta aaggataaag
841 gaggatttag tagtaaatta agaatagaga gcttaattga attgagcaat gaagtacgca
901 cacaccgcc gtcaccctcc tcaactaaa taaattaac tataaataat tcataaacia
961 aatttacgag aggagataag tcgtaacaag gtaagcatac tggaaagtgt gcttggata
1021 atcacagtgt agcttaatta caagcatct ggctacacc cagaagaatt cataataaat
1081 gaacacttg aactaatcct agccctaac tcttaataat ataactattt cacactataa
1141 attaaacat ttaatctcgt aaaagtattg gagaaagaaa ttattcata ggagctatag
1201 agtaagtacc gcaagggaaa gatgaagaa taattaaag tagcaataag caagataaa

```

1261 accttgtagc ttttgataa tgagttaact agaataaatt taactaaaag cattaagct
1321 aaataccccg aaaccaaagc agctacctaa aaacaattta aacgaatcaa cccgtctatg
1381 ttgcaaaaata gtgggaagat ttttaggtag aggtgaaaag cctaacgagc ttggtgatag
1441 ctggttacc aaataatgaa ttttaagtca acttaaaatt taccaaaaga acatttaac
1501 ctcatgtaaa ttttaaatat agccaaaaga gggacagctc tttaggaaaa ggaaaaaac
1561 ttaagtagtg aataaactca ataaataacc aaaccattgt aggcccaaaa gcagccacca
1621 ataagaagag cgttcaagct caacattagc atttaataa tttacaagt aataataat
1681 tcctacatta aaaattgggt taatctatta atgtatagat gagatactgt taacatgagt
1741 aacaagaatt ttattctcca agcacaagcg tataacaact cggataacca ttgttaatta
1801 tcaacctata gataaaaact ataaattaat agatctaaaa acaacttgtt aatccaacac
1861 aggagtgcata taaggaaaga ttaaaaaaaa taaaaggac tcggcaaac aaacccgc
1921 ctgtttacca aaaacatcac ctctagcata ttaagtatta gaggcattgc ctgcccagtg
1981 actaaagttt aacggccgag gtatcctgac cgtgcaaagg tagcataatc acttgttct
2041 taattagga ctagcatgaa cggctagac agggtttgac tgtctctat tttgatcag
2101 tgaattgac ctttcagtga agaggctgaa atactataat aagacgagaa gaccctatgg
2161 agctttaatt tatcagttta atcactactaa taaagttacc taatggaata accacaatga
2221 tgtaaacca taatttcggt tgggtgacc tcggagaaaa gaaaatcctc cgaatgattt
2281 taacatagat ccacaatca aagtaagcaa tctatcttat tgacccaaaa atctttttg
2341 atcaacggac caagttacc tagggataac agcgaatcc tattaagag ttcatatcga
2401 caattaggt ttacgacct gatgttgat caggacatcc caatgggtgca gaagctatta
2461 atggttcgtt tttcaacga ttaaagtcct acgtgatctg agttcagacc ggagtaatcc
2521 aggtcgggtt ctatctatta acaatttctc ccagtacgaa aggacaagag aatggagcc
2581 acottataaa taagcgcacc caacttaata tatgaattaa tctaaatata gtaagtagt
2641 taactctttg cctagacaag gctattaggg tggcagacc aggaattgc gtaagactta
2701 aaacctgtt cccagaggtt caaatcctct tctaatagt gtattttatt aatctctta
2761 tattattagt cccgatcctt attgcatag ccttttaac actagtagaa cgaaaaattt
2821 tgggtacat acaactacga aaaggcccta acgtttagg tccatagga gtcctcaac
2881 cattcgcaga cgtataaaa ctotttatta aagaacctat ggcctcccta actacctca
2941 tttactatt tattattgca ccaaccttt cactcacact agccctaagc ctatgaattc
3001 cactaccaat accacacca ctaatcaatc taaacctagg tgttttattc attttagcta
3061 catcaagctt ttcagtatat tctatcttat ggtcaggatg agcatcaaac tcaaaatac
3121 cgtatttgg tgccttacga gcagtagcac aaacaattc atatgaggtc acaatagcta
3181 ttattctctt atcagtttta ttaataagcg gctcattctc attcaaaata ctatttata
3241 ccaagaaca catatgacta attctactg cctgaccaat agctataata tgatacatc
3301 caacctagc agaaacaaat cgagctocat tcgacctaac agaaggagaa tcagaactag
3361 tttccgggtt taacgttag tagctgcag gtccctttgc actattttt atagcggat
3421 atacaaacat tattcttata aatgccctaa catcattgt cttcctagc cccatccata
3481 taattaacta cccagaactc tactcaatta atttataac agaaacgcta ttacttcaa
3541 caacatttct atgaattcgg gcatcatacc cccgctttcg ttatgaccaa ctatacatc
3601 tcctatgaaa aaattttctc ccccttacac tagcactatg tacatgacat atctcctac
3661 ctatctttat agcaagtatc ccaccataca tctagaaaca tgtctgacaa aagagttact
3721 ttgatagagt aaaatataga ggtttaagcc ctcttgtttc taggataata ggaattgaac
3781 ctattcttaa gaattcaaaa ttctctgtgc taccataca cctaactccta acatagtaag
3841 gtcagctaat taagctatcg ggccataacc cggaaaacgt tggtttaaac ccttccgta
3901 ctaataaatc ctatcaccat tatcattatt tacttcaact tctcattgg tccctcctc
3961 acaatatcaa gctccaacct tctactcata tgagtaggat tagaactaag tttactagca
4021 atcattcccc tattaattaa taaaaaaat ccccgatcaa ctgaagcagc aacaaaatat
4081 tttattacac aagcaacagc ctcaataatc atccttttag ctattatcct aaattttaaa
4141 caactcggac tatgaacatt tcaacaacaa acaaacagcc ttctactcaa cattactcta
4201 attgactat caatgaaact aggactagcc ccattccact tttgactccc agaagtaaca
4261 caaggaatcc cactacacat aggactaatc ctctcacat gacaaaaaat tgcccacta
4321 tcaatcctat ttcaatttta tcagctactt aatccaacca tcaattatgat ctagcaatt
4381 tcactattc ttatcggagc atgaggtgga ctttaacaaa cacaatgagc aaaaatcgt

4441 gcatattcat caatcgctca cataggatgg atattagcta tcttgccta caaccaact
4501 ataacacttc ttaacttact aatttatatt atcctcactg ttcccatatt cctcatgctt
4561 ataacttaag catcaacaac tatcaactca atctcactcc tatgaataa ggctccaaca
4621 acccttgat taactccatt aatcctcctc tccctaggag gtctaccccc actaacagga
4681 tttttacca aatgagctat tatcacagaa ctactaaaaa acaactgcct aattattacc
4741 acaataatag ctattatagc cctatataac ctatTTTTct acacacgtct gatctactct
4801 acttctctca ccatattccc aactaataat aactcaaaaa tattaacca ccatttaa
4861 tataaaaaa atattatcct tccatcactt actatcatca gcacattaat tctccctcta
4921 tcaccacaac taatcattta gaagtttagg atatatgcag tccaagagcc ttcaaagccc
4981 taagaaaaca aacaagtta acttctgata aggactgtaa gactacttcc tacatctatt
5041 gaatgcaaat caatcgcttt aattaagcta agacctact agattgtag gaattaaacc
5101 tacgaaattt tagttaacag ctaaataacc tacttctggc ttcaatctac ttctccgccc
5161 tatcagaaaa ggggCGggag aagccttagt agatatattg ctctaccctc togaatttgc
5221 aattcgacat gactaacacc ttaaggcttg gtaaaaagag gacttaaaccc tctgtattta
5281 gatttacagt ctaatgctta ctcagccatt ttacctatgt tcattaatcg ttgattattc
5341 tcaactaacc acaaagacat tggcacoccta tatctcctat tccggagcttg agccggaata
5401 gtagggactg cattgagcat tctgatocga gctgaaactg gaaaaccagg tgcactttta
5461 ggggacgacc agatctataa tgcattgtt accgcccattg cgtttgttat aatcttcttc
5521 atagtaatac caataataat tgggggcttt ggaaactgac tctgaccact tatgattgga
5581 gctcccgata tagcattccc acgaataaat aatataagct tctgacttct accaccatct
5641 ttctgtctcc tattagcacc atccatagtt gaagcaggag ccggtacagg atgaacagta
5701 taccacactt tagctggaaa cttagcccat gctggagcat cagtagacct gacaattttc
5761 tctcttact tggcoggtgt accctcaatt ctaggagcta ttaactttat taccactatt
5821 attaataata aaccaccagc tataacccea tatcaaacac cactgtttgt atgatccgta
5881 ctaattacag ccgttctgct cctattatcc ctcccagtac tagcagcagg aattacaata
5941 ctattaactg accgaaacct aaacacaact ttctttgacc ctgacggagg tggagacca
6001 attctatatac aacatctatt ctgattcttt ggccacccc aagtttata ccttctctg
6061 cctgggttcg gaattatctc tcatgtagtt acttattact ccggtaaaaa agaaccattc
6121 gggtacatgg gaatagctg agctataata tcaattggct tcttaggttt tattgtatga
6181 gcacatcaca tgtttacagt agccttagat gtagacacac gagcctactt tacatctgcc
6241 acaataatta tgcctatccc aacagcgcta aaagtattta gctgacttgc aacactacac
6301 ggaggtaata tcaaatgac tccagctata ctttaggcc taggctttat ctctctattt
6361 acagtaggtg ggtaacagg aattgtctc tcaactctc cactcagcat tgtcttcat
6421 gacacatatt atgtagtagc ccatttccat tatgtactgt ctataggagc agtatttgc
6481 attatagctg gttttgtcca ctgattcccc ctattcacag gatataacct agacgatata
6541 tgagcaaaaa cacactttgc cattatattt gtaggagtta acataacatt ctccctcaa
6601 cttttctag gactctcagg cataccacga cgatactcag attatccaga tcttataacc
6661 acatgaaata cagtatcacc tataggatca tttatctctc ttacagccgt cctactaata
6721 atttttataa tctgagaagc attcgttca aaacgagaag ttctaacagt atcttattct
6781 tcaaccaact tagagtgact acatggctgc ccacctcctt atcatacatt cgaagagcct
6841 tcttacgtaa aagtaaaata agaaaggaag gattcgaacc ccctaaaact gtttcaagc
6901 caacctcata gcctctatgt ctttctcaat gagatattag taaaacaatt acataacctt
6961 gtcaagggtta aattatagac ttaaactctat atactttaca tggcttacc attccaatta
7021 ggtctacaag atgccacac ccctattata gaagaattaa taaacttcca tgaccacaca
7081 ctaatgattg ttttcttat tagcacotta gtaactttata ttatttact gatattaaca
7141 acaaaactta cacatacaag cacaatagat gcccaagaag tagagactgt ttgactatc
7201 ttaccagctg ttattottat ttaattgca ctccctcac tacgaatttt atatatgata
7261 gacgaaatta ataacctgt actaacagta aaaactatag gccaccaatg atactgaagc
7321 tatgaatata ctgactatga agatctatgc tttgactctt acataattcc aacaactgac
7381 ttaaaaccag gtgacctag actattagaa gttgacaacc gagtagtttt accaatagaa
7441 ctccaatcc gaatgcttat ctcatcagaa gacgttctcc actcatgagc tgttccctca
7501 ctaggactaa aactgacgc tatcccagga cgcctaaatc aagcaacagt aaccttaac
7561 cgtccaggat tttttacgg tcaatgctct gaaatctgcg ggtctaatca cagcttcaata

7621 cctattgttc tagaaatagt gccactaaaa cacttcgaaa actgatcagc ctcaataatt
7681 taatctcact gtgaagctta gagcgtaac cttttaagtt aaagttagag accctaatac
7741 tccatagtga catgccacaa ctagatacat ccacatgatt tattaccatc atctcatcta
7801 tggctacact atttatttta tttcaactaa aaatttcctc acaatctttt cccactcctc
7861 catcccaaaa aacctttaca gcacaagaaa caaaaactcc ttgagaatca aatgaacga
7921 aaatctatth gcctctttta tcaactcctc agtaatagga ctcccaattg taattactat
7981 tatcatatth ccactctatc tattcccac c atcagaacgt ttaattagta atcggcttca
8041 ctcatthcaa cactgactaa tcaactaat tatcaaaaa ataacttta ttcacacacc
8101 aaaaggacgg acttgaactc taataattgt ctcccttatt atattcatca gctcaactaa
8161 ccttctagcc cttcttccac acacatttac accgactact caactatcaa taaatcttag
8221 tatagcaatc ccactatgag caggagcagt cctattaggg tttcgcata aattaaag
8281 ctcccttagct ctttctctc cacaaggcac ccctatttca cttatoccta tactaattat
8341 tatcgaacaa attagtctat ttattcaacc tatagctcta gcagtacgat tgacagccaa
8401 tattaccgca gggcatctac ttatgcatct aattggagga gcaaccctag tcttataaa
8461 tattagccca cccaccgcta ctattacatt cattatttta ctcccttctca cagtactaga
8521 gtttgctgta gccctaattc aagcatacgt atttactctt ctagtaagcc tatatctaca
8581 cgataacaca taatgaccca ccaaaccat gcctatcata tagttaacc aagtccatga
8641 ccacttacag gagcatttcc agcccttcta cttacatctg gattagtaat atgattcac
8701 tatagctcaa ccattcttct gtctctagga ctactagcca atacacttac tatatatcag
8761 tgatgacgag atattattcg tgaaggaacc tatcaaggcc atcatacccc tattgtacaa
8821 aaaggtcttc gatatggaat aatcttgttc attgtctctg aagtattttt cttgtctggc
8881 ttcttctgag ctttctacca ttcaagccta gtaccaacc acgatctcgg aggttgctga
8941 cctccaacag gaattactcc tcttaactct ctagaagtc cactattaaa tacatcagtg
9001 ctactagcat cagggttttc aattacatga gccaccata gtctcataga aggcaaacga
9061 aatcatataa atcaagccct actaattact attctattag gactttactt caccacttt
9121 caagcctcag aatattttga aactccattc tctatttcag acggaattta cggctccacc
9181 ttctttatag caacaggatt ccacgggctc catgtaatta ttggaactac attcctcgt
9241 gtttgctccc tacgacaact aaaatttcac ttatcatcaa aacaccattt tggatttga
9301 gcagcagcgt gatattgaca ctttgtagac gtagtctgac ttttctata tgtatccatc
9361 tattgatgag gatcttactc ttttagtata attaatatga ctgacttcca attagtagat
9421 tctgaacaac atcagaagag agtaataaat cttataatta tcatatttat taatactatc
9481 ctatcattga tccataattc agttgcattc tgaactcccc aaataaactt atatacagaa
9541 aaagccaatc catatgaatg tgggttgac ccatcaagct ctgctcgttt accatttca
9601 ataaaatttt tcctagtagc tatcacattc ctattatttg atctagagat tgcctacta
9661 cttcctcttc catgagcaat ccaaacact aataccaatg taataataat ctgagcttta
9721 atcttagtca ccattctatc attaggccta acctatgaat gaactcaaaa aggactogaa
9781 tgaacagaat agctggaat tagtttaaat aaaattaatg atttcgactc attagattat
9841 gatcctatcc ataattacca aaaatgtcat ctgcctttat taatctcaca ctagccttca
9901 tactttcact tctaggaaca ttaacgtttc gttctcactt aatatctact cttctatgct
9961 tagaaggaat aatattatcc ttatttatta taacatccct agcttccctg aattctaact
10021 ctatagcatc tatacctatt ccaattacta ttttagtatt tgcagcatgc gaagcagcag
10081 taggactagc cctattagtt aaagtgtcca acacctatgg taccgactac gtccaaaacc
10141 tcaatttact tcaatgctaa aaattatttt tccatcactc atacttctcc cgttaacatg
10201 actatcaaaa ccgaaaaaaa cctgaataaa cgtcacttcc tacagtttca tgattagcct
10261 tattagcctt tcaattctat gacaaactga cggagccct ctttaacttt caactatatt
10321 ctccctagat cctctatcat cccacttat tattttaaca acttgactat tacccttat
10381 acttatagct agtcaaaatc atctaaaaaa agaaaacttc acctaccaaa aactctatat
10441 ttcaactacta gttagcctac aaattcttct agtaataact ttttcagcaa ctgagctaat
10501 catattttat attctatttg aagctaccct tattcctaca ttaattatta ttacacgatg
10561 aggcaatcaa acagaacggt taaacgcagg actttacttc ttattctata cctcatcgg
10621 ctctatccca ctcttaatcg cccttatctt tatccaaaac tccataggtta cactcaactt
10681 taccatacta tcaactacaa ctaccocct gaaccctta tgatccaaca acatcttatg
10741 actagcatgc ataatagcat ttataattaa aataccctta tatggagttc acctatgact

10801 tccaaaagcc catgttgaag cacccattgc aggatccata atcctagcgg caattctctt
10861 aaaactagga agttacggaa taatacgaat tgccatcatt ctggacccac taacaaaata
10921 catagcctat cccttcaccc tactatccct gtgggtata atcataacaa gctcaatttg
10981 cctacgccaa acagacctaa aatcccta atcgttattcc tcagttagcc atatagcttt
11041 agtaattgca tcaattataa ttcaaaactcc ttggagcttt ataggagcta caatattaat
11101 aattgctcac ggcctaactt catcacttct attctgctta gcaaaactcca actatgagcg
11161 gattcacagc cgaactataa ttatagccc aggacttcaa ataactttc cactcatagc
11221 aacttgctga ttagtagcaa gcctagctaa tcttgctcta ccaccatcaa tcaatctcat
11281 aggagaacta tttattacca taccactatt ttcttgatcc aacttttcaa ttatccta
11341 aggagtcaac attattatta caggcatata ctcaatata ataattatca caacccaacg
11401 aggaaaacta acaagtcaca taacaatct ccaaccttcc cacacccgag aactaacact
11461 tataattctc cacattatcc ccctagtcct attaactatt aacccaaaac taattacagg
11521 gttgacaata tgtaaatata gtttcaaaa acattagact gtgaatctaa aaacaggagt
11581 tcaacacttc ttattacca agaaagtacg caagaactgc taactcatgc cccatggtt
11641 aaccacatgg ctttcttact ttataggat agaagtaatc cattggtctt aggaacccaa
11701 aaccttggtg caattccaaa taaaagtaat aaatatcatt acatcattta ttatagcaat
11761 ttttctcatt cttatcacac caattttaat ctctcaact agcctgttca aacatattaa
11821 ttttccaata tatgcaacac tatcaatcaa actctcattc tttctaagcc tcttacctct
11881 cataatatta tttatcaca acacagaata cataattact acctgacatt gaatcacaat
11941 caattctatt gagctcacia taagctttaa aatagattat ttctcaacta tatttctatc
12001 cgtagcccta ttcgttacat ggtctatcat acaattctcc tcatgatata tacactcaga
12061 tctcacatt aatcgattta ttaaactatt actattggtt ttaatcacta tactaatttt
12121 aacttcagct aacaatctat ttcaactatt tatoggatgg gaaggagtag gtattatatac
12181 tttctcctc atcggatgat gatttggacg agcagacgca aataccgcag ctcttcaagc
12241 aattttatat aaccgaatog gtgatcgg cttatctta gctataacat gattttgtct
12301 aaacataaac tctgagaac ttcaacaaat tctattaacc aataataata atctaattcc
12361 actgctcggg ttactaattg cagccacagg aaaatcagcc caattoggcc ttcacatagc
12421 acttccatca gccatagaag gaocccacacc agtttcagcc ctacaacact caagcacaat
12481 agtagtagca ggaatttttc tactagtctg gtttaccct ataacatcca acaaccatc
12541 aatcctaact ataactat gtctaggagc actaactacc ctattaccg ctatctgtgc
12601 cctaccccaa aacgatatta aaaaaattgt agcattttcc acatcaagtc agctaggcct
12661 aataatagtt acattaggaa ttaatcaacc atatctagca ttcttacata tctgacaca
12721 tgcattcttc aaagccatgc tcttcatatg ctccggatca attattcaca gtctaaacga
12781 tgaacaagac attcgaaaaa taggaaacat aataaaaaact atgccattca catcatcctg
12841 cctaacaatt ggtagtcttg ccctaacagg aataccattc ctaacagggt tctactcaaa
12901 agacctaatc attgaagcta tcaatacctg caataccaac gcctgagccc tactaattac
12961 actaatgcc acatcaataa cagctatcta cagcatacga attatttact ttgtatcat
13021 aacaaaacca cgattcccac cattaatctc aatcaatgaa aataaccag atcttactaa
13081 cctattataa cgcctagcat ttggaagcat cctagctgga tttttcatta catataatat
13141 cccaccaact aacatccaag ttctcactat accatgatac ttaaaaaacta cagccctgct
13201 aatctctatc ctaggctttc tactagccct agaactaaat aacctcacta taaacctttc
13261 cataagtaaa aaaaactcgt attcatcatt ttccacatca ctaggctact tccatcaat
13321 catccatcga attattccaa acaaaact taacctagc ttcaaaacat cacttaacct
13381 tctagatctg tattgactag agaaatcaat cccaaaatct acttcaacta tacactotta
13441 tatatcaaaa ctactaacta accaaaaagg attagtaaaa ctatatttta tatcattttt
13501 actttctatt ctactaacca ctacctctt tatcattaat ctcgagtgat ttcaataata
13561 ataaaaatc cagcaacaa agaccatcca gcaactacta ttattcaagt agcacaacta
13621 tacatcggcg caaccccaat accaccctct aatattacac caacatcatc aatctcatic
13681 atcaatcaat cacctagact atcaaaatca actacctct cccattata ataatcaaga
13741 aaatacacia aaaacatttc cattaaaacc cctgcaatta aaaaacccaa aacaaaacca
13801 ttgaccctc aagtctcagg atactctca gtagctatag ctgtctgata cccaaacact
13861 actattatc ctccaaata aattaaaac accattagac ctaaaaagga tccaccaaac
13921 cctaaaacta ttgacaacc cacaaaacca ctcataatta aaccaaaccc ccataataa

13981 ggagaaggct ttaaagctaa cccaagacaa ccagtcacaaa ataataaaact taaaacaaaa
14041 atatagttag tcattatttc tacacagcat tcaactgtga ctaatgacat gaaaaatcat
14101 cgttgtaatt caactacaga aacctaatga caaacatccg aaaaactcac cccctattta
14161 aaatcattaa ccactctttc atcgacctcc ctgccccatc taacatctca tcctgatgaa
14221 actttggctc cctcctaggt ctatgcctcg taattcaaat tcttacaggc ttattcctag
14281 ccatacacta cacatcagac acaataacag cattctcatc agttacacat atttggccgag
14341 acgtaaaacta tggatgacta attcgatata tacacgctaa tggagcttca atatttttta
14401 tctgcttatt tctccatgta ggacgaggaa tgtactacgg atcctacgca tttatagaaa
14461 catgaaatat cggagtaatc ctattatttg cagtaatagc tacagcattc ataggctatg
14521 tacttccatg aggacaaaata tcattttgag gggcaacagt aattacaaat ctctctcag
14581 ccatcccata tatcggaact accctggtag aatgaatttg aggaggattc tcagtagata
14641 aagccacttt aacacgtttc ttgcattcc attttattct cccatttacc atcgcagccc
14701 tggtaaatgt ccatctccta ttctccacg aaacaggctc aaacaacca acaggtctaa
14761 actcagacgc cgataaaaac ccatctcacc catactacac aattaaagat attctaggca
14821 tttcattat aattatattc ctaatagtcc tggctctatt tttccagac ctacttggag
14881 acccagataa ttacacacca gcaaatccac ttaatacacc accacatac aaaccagagt
14941 gatacttctc atttgcataat gcaatctac gctccattcc aaacaaacta ggaggagtcc
15001 tagctctagt cctatccatc cttatcctag ccttactgcc acttcttcac acttcaaaaac
15061 aacgaagcct aatattccgt cctattactc aaatactota ctgaatccta gtagccaacc
15121 ttctcgtcct tacctgaatc ggcggaacac cagttgaata tccattccta atcatcggcc
15181 aactcgctc catcagctac ttctcaatta ttctaactct tatgccaatc tcaggcatca
15241 tcgaaaataa tatactaaaa tgaacctat gccctgatag tatagaacca ttactctggt
15301 cttgtaaac aaaaacgaag aacttctct ctccaggcatc aagaagaagg aatcctcccc
15361 accatcagca cccaaagctg gtattctact taaactactt cttgagtaca taaaattaca
15421 tagtacattc atacatatat gtatatagta cattaaatta tttacccta gcatataagc
15481 aagtacattt aattaatgat cctagacatt aatgattaat caaatatctc tcttctcca
15541 catgaatatt caggacttac attttaatta atgatcttaa gacatatctg cgttatctta
15601 cataccatc tcagtcataa actcttctct tccatatgac tatccctctc cacatttggg
15661 ctatatact accatctcgt gtgaaaccaa caaccggccc acctatgccc ctcttctcgc
15721 tccgggcca taaaacttgg gggtagctaa cctgaaactt tatcagacat ctggttcta
15781 cttcagggcc atcaaatgcg ttatcgcca tacgttccc ttaaataaga catctcgtatg
15841 gtacgggtct aatcagccaa gacgggcata actgtggtct cgcgtagtgt gtattttttt
15901 attttttagga tgctgtgact caccatagcc gtcaggcat gaaggtcagc ccaccatgaa
15961 gctggactta cagtgaagga tcatttatcc acataaaaca accaccgaag acaattatg
16021 aatgcttgtt agacataact ctattattac ttctaatttt aacttaccba aaccocctta
16081 cccoccttga cgtcaaaccc caaaaccatc aaagacttaa cttaaatcc ataagtttta
16141 ttctattcta gtagttcaca aaattaaact tatattacag tattgggcaa aattttataa
16201 aaaattaatt ttgaattatc aaaaactcac cttattcaaa tttcctaata aatttatata
16261 a

Appendix 3. The sequence of mitochondrial genome of *Apodemus peninsulae*

LOCUS HQ660074 16266 bp DNA circular ROD 30-DEC-2011
 DEFINITION *Apodemus peninsulae* mitochondrion, complete genome.
 ACCESSION HQ660074
 VERSION HQ660074.1 GI:316993300
 KEYWORDS .
 SOURCE mitochondrion *Apodemus peninsulae* (Korean field mouse)
 ORGANISM *Apodemus peninsulae*
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; *Apodemus*.
 REFERENCE 1 (bases 1 to 16266)
 AUTHORS Oh, D. J., Kim, T. W., Chang, M. H., Han, S. H., Oh, H. S. and Kim, S. J.
 TITLE The mitochondrial genome of *Apodemus peninsulae* (Rodentia, Muridae)
 JOURNAL Mitochondrial DNA 22 (4), 99-101 (2011)
 PUBMED 22040077
 REFERENCE 2 (bases 1 to 16266)
 AUTHORS Oh, D. -J., Jung, Y. -H., Kim, T. -W., Han, S. -H., Chang, M. -H., Oh, H. -S.
 and Kim, S. J.
 TITLE Direct Submission
 JOURNAL Submitted (29-NOV-2010) Department of Biology, Jeju National
 University, 66 Jejudaehakno, Jeju 690-756, Republic of Korea
 FEATURES Location/Qualifiers
 source 1..16266
 /organism="Apodemus peninsulae"
 /organelle="mitochondrion"
 /mol_type="genomic DNA"
 /db_xref="taxon:105297"
 tRNA 1..67
 /product="tRNA-Phe"
 rRNA 68..1023
 /product="12S ribosomal RNA"
 tRNA 1024..1092
 /product="tRNA-Val"
 rRNA 1093..2663
 /product="16S ribosomal RNA"
 tRNA 2664..2738
 /product="tRNA-Leu"
 /note="codons recognized: UUR"
 gene 2739..3693
 /gene="ND1"
 CDS 2739..3693
 /gene="ND1"
 /note="TAA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /codon_start=1
 /transl_except=(pos:3693, aa:TERM)
 /transl_table=2
 /product="NADH dehydrogenase subunit 1"
 /protein_id="ADU78207.1"
 /db_xref="GI:316993301"

```

/translacion="MYFINIITLLVPILIAMAFLLVERKILGYMQLRKGPNNVGPYGV
LQPFADALKLFIKEPMRPSTTSISLFIAPTLSLTLALSLWIPLPMPHPLINLNLGI
LFILATSSLSVYSILWSGWASNSKYSLFGALRAVAQTSYEVTMAIILSVLLMSGSF
SLQTLIYQEHMWLIVPTWPMAMMWYISTLAETNRAPFDLTEGESELVSGFNVEYAAG
PFVLFMAEYMNILMNALTSIFLGPMHKIYYPEFYSINFMTETLILSTTFLWIRAS
YPRFRYDQLMHLLWKNFLPLTLAFCMWHISLPFIMASIPPYM"
tRNA 3694..3762
/product="tRNA-Ile"
tRNA complement(3760..3830)
/product="tRNA-Gln"
tRNA 3834..3902
/product="tRNA-Met"
gene 3903..4938
/gene="ND2"
CDS 3903..4938
/gene="ND2"
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:4938,aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="ADU78208.1"
/db_xref="GI:316993302"
/translacion="MNPITLIIYLTILMGPVITMSSTNLLLMMVGLLESLLAIPLL
INKKNPRSTEAATKYFVTQATASMIILLAIIILNYKQLGLWVQQQTNPLLLNITLIAL
SMKLG LAPFHFWLPEVTQGIQLHTGLILLTWQIAPLSMLIQNYLLDHTIIMIFSIL
SIFIGAWGGLNQTQMRKIMAYSSIAHMGWMLAILPFNPTMTLLNLMIIYIILTPMFLV
HMFNSSTTMNSISLMWNKTPPTLIMIPLILLSLGGLPPLTGFLPKWAIITELLKNNCL
IMTTMMAMMALLNLFFYMRLIYSTTLTMFPTNNSKMLTHHTNHKYNILPLSLIIST
LMLPLSPQLII"
tRNA 4939..5006
/product="tRNA-Trp"
tRNA complement(5008..5076)
/product="tRNA-Ala"
tRNA complement(5078..5148)
/product="tRNA-Asn"
rep_origin complement(5149..5179)
/note="L-strand replication origin"
/direction=LEFT
tRNA complement(5180..5245)
/product="tRNA-Cys"
tRNA complement(5246..5312)
/product="tRNA-Tyr"
gene 5314..6858
/gene="C01"
CDS 5314..6858
/gene="C01"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit 1"
/protein_id="ADU78209.1"
/db_xref="GI:316993303"

```

```

/translacion="MFINRWLFSTNHKDIGTLYLLFGAWAGMVGTAALSILIRAE LGQP
GALLGDDQIYNVIVTAHAFVMIFFMVMPMMIGFGNWLVPMLIGAPDMAFPRMNMMSF
WLLPPSFLLLLASSVVEAGAGTGWTVPPLAGNLAHAGASVDLTI FSLHLAGVSSILG
AINFITTIINMKPPAMTQYQTPLFVWSVLI TAVLLLLSLPVAAGITMLLTDRNLNTT
FFDPAGGGDPI LYQHFWFFGHPEVYILILPGFGIISHVVTTYSGKKEPFGYMGMVWA
MMSIGFLGFIVWAHMFVGLDVDTRAYFTSATMIIAIP TGVKVFSWLATLHGGINIKW
SPAMLWALGFIFLFTVGLTGIVLSNSPLDIVLHDTYYVVAHFHYVLSMGAVFAIMAG
FVHWFPLFTGYTLDDMMWAKTHFAIMFVGVNMTFFPQHFLGLSGMPRRYSYDPDAYTTW
NTVSSMGSFISLTAVLLMIFMIWEAFASKREVLTVTYSSTNLEWLHGCPPPYHTFEEP
SYVKVK"
tRNA complement (6856..6924)
/product="tRNA-Ser"
/note="codons recognized: UCN"
tRNA 6928..6995
/product="tRNA-Asp"
gene 6997..7680
/gene="C02"
CDS 6997..7680
/gene="C02"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit 2"
/protein_id="ADU78210.1"
/db_xref="GI:316993304"
/translacion="MAYPFQLGLQDATSPIMEELMNFHDHTLMIVFLISSLVLYIISL
MLTTKLTHSTMDAQEVETIWTILPAVILILIALPSLRILYMMDEINNPNLTVKTMGH
QWYWSYDYEDLCFDSYMIPTNDLKPGDLRLLEVDNRVLPME LPIRMLISSEDVL
HSWAVPSLGLKTDIIPGRLNQATVTSNRPGFLFYGQCSEICGSNHSFMPVLEMVPLKH
FENWSASMI"
tRNA 7684..7747
/product="tRNA-Lys"
gene 7749..7952
/gene="ATP8"
CDS 7749..7952
/gene="ATP8"
/codon_start=1
/transl_table=2
/product="ATPase subunit 8"
/protein_id="ADU78211.1"
/db_xref="GI:316993305"
/translacion="MPQLDTSTWFITIISSMATLFLVLFQLKISSQSFPPTPPSPKTFST
QKTKNPWESKWTKIYLP LSSPLQ"
gene 7910..8590
/gene="ATP6"
CDS 7910..8590
/gene="ATP6"
/codon_start=1
/transl_table=2
/product="ATPase subunit 6"
/protein_id="ADU78212.1"
/db_xref="GI:316993306"
/translacion="MNENLFASFITPSVMGLPIVVTIIMFPSILFPSSERLINNRLHS
FQHWLIKLIKQMMLIHTPKGRTWTLMI VSLIMFIGSTNLLG LPHFTPTTQLSMNL

```



```

/codon_start=1
/transl_except=(pos:11528, aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 4"
/protein_id="ADU78216.1"
/db_xref="GI:316993310"
/translation="MLKII FPSLM LPLTWLSKPKMTWNTVTSYSFLISLISLSIMWQ
TNETPLNFS TMFSSDPLSTPLIILTTWLLPLMLMASQNHKKKEKTMHQKLYISLLVSL
QILLVMTFSATELIMFYILFEATLIPTLIIITRWGNQTERLNAGIYFLFYTLIGSIPL
LIALIFMQTSMGTLNFSMLSMPSPDLWSNINLWACMMAMFKMPLYGVHLLWLPK
AHVEAPIAGSMILAAILLKLGSGMMRIAIMLDPLTKFMAYPFILLSLWGMIMTSSIC
LRQTDLKSLIAYSSVSHMALVIASIMIQTPWSFMGATMLMI AHGLTSSLLFCLANSNY
ERIH SR TMIMARGLQMI FPLMATWWLVASLTNLALPPSINLIGELFITMSLFSWSNFS
IILMGINIIITGMYSYMIITQRGKLTSHMNNLQPSHTRELTLMALHIIPLLLLTIS
PKFITGLTM"
tRNA 11529..11595
      /product="tRNA-His"
tRNA 11596..11654
      /product="tRNA-Ser"
      /note="codons recognized: AGY"
tRNA 11654..11724
      /product="tRNA-Leu"
      /note="codons recognized: CUN"
gene 11725..13554
      /gene="ND5"
CDS 11725..13554
      /gene="ND5"
      /codon_start=1
      /transl_table=2
      /product="NADH dehydrogenase subunit 5"
      /protein_id="ADU78217.1"
      /db_xref="GI:316993311"
      /translation="MNTITSSIMAI FIVLLTPILISSTNMLKQINFPLYATMSIKLSF
      FLSLLPLMMFFYYNTEYMITTWHWMTINSIELTMSFKMDYFISMFLSVALFVWTSMQ
      FSSWYMHSDPHINRFIKYLLFLITMLILTSANNLFQLFI GWEGVIMSFLIIGWWFG
      RSDANTAALQAILYNRIGDIGFILAMTWFC LNMNSWELQQILLTENNLIPLGLLIA
      ATGKSAQFGLHPWLP SAMEGPTPVSALLHSSTMV VAGIFLLVRFHPLTSNPKILTMM
      LCLGGLTTLFTAICALTQNDIKKIVAFSTSSQLGMMVTLGINQPYLAFLHICTHAFF
      KAMLFMCSGSIHSLNDEQDIRKMGNNMKTMPFTSSCLTIGSLALTGMPFLTGFYSKD
      LIIEAMNTCNTNAWALMITL IATSMTAIYSMRIIYFVTMTKPRFPPMILINENDPNLM
      NPIKRLALGSI MAGFFISYNI PPTNIQILTMPWYLKTTALFISILGFLMALELNNLTL
      KLSMNKKNLYSSFSTSLGYFPSIHRIPNKTLNLSLKTSLNLLDLYWLENSIPKPI S
      TMHSYASKLLTNQKGLVKLYFLSFFLSILLTVILYIINLEWSQ"
gene complement(13532..14050)
      /gene="ND6"
CDS complement(13532..14050)
      /gene="ND6"
      /codon_start=1
      /transl_table=2
      /product="NADH dehydrogenase subunit 6"
      /protein_id="ADU78218.1"
      /db_xref="GI:316993312"
      /translation="MTNYIFVLRLLFLTGCLGLALKPSP IYGGFGLIMSGFVGCLMVL

```

GFGGSFLGLMVFLIYLGMMVVFYTTAMATEEYPETWGSNWFVFGSLIIGVLMELFF
 MYLIDYYDGVDDVDFDKLGDWLMYEVDVGVMLEGGIGVAAMYS CATMMM VVAGWSLF
 AGIFIIIEITRD"

tRNA complement (14051..14119)
 /product="tRNA-Glu"

gene 14123..15266
 /gene="CytB"

CDS 14123..15266
 /gene="CytB"
 /note="TAA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /codon_start=1
 /transl_except=(pos:15266, aa:TERM)
 /transl_table=2
 /product="cytochrome b"
 /protein_id="ADU78219.1"
 /db_xref="GI:316993313"
 /translation="MTNIRKTHPLLKIINHSFIDL PAPSNISSWWNFGSLLGVCLIIQ
 IVTGLFLAMHYTSDTMTAFSSVTHICRDVNYGWLIRYMHANGASMF IICLFLHVGRGM
 YYGSYFMETWNI GVI LLFAVMATAFMGYVLPWGQMSFWGATVI TNLLPAIPYIGTTL
 VEWIWGGFSVDKATL TRFFAFHFILPFI I AALAVVHLLFLHETGSNNPTGLNSDADKI
 PFHPYYTIKDI LGIFILVGLM TLVLFSPDLLGDPDNYMPANPLNTPPHIKPEWYFLF
 AYA ILRSIPNKLGGVLALILSILILALLPFLHTSKQRSLMFRPITQTLYWILVANLLT
 LTWIGGQPV EHPFVIIGQLASISYFSIILILMPLSGI IEDKILKWNL"

tRNA 15267..15333
 /product="tRNA-Thr"

tRNA complement (15334..15400)
 /product="tRNA-Pro"

D-loop 15401..16266

ORIGIN

1 gttaatgtag ctaaataaa gcaaagcact gaaaatgctt agatggattt ttattatccc
 61 ataacacaaa aggtttggtc ctagccttat aatagttgg aggtaagatt acacatgcaa
 121 atgtccataa accagtgtaa aatcccttaa atatttttac aaaatttaag gagagggtat
 181 caagcacata aatagctta agacaccttg cctagccaca ccccacggg actcagcagt
 241 gataaatatt aagcaatgaa cgaagtttg actaagctat gcctcttagg gttgtaaat
 301 ttcgtgcag ccaccgggt catacgatta accogaacta attactctcg gcgtaaaacg
 361 tgttactat aaacttacta atagaattaa aattcaactt atatgtaaaa attcattggt
 421 agaacttaaa ctcaataacg aaagtaattc taataacatt attttacag atagctaaga
 481 tccaaactgg gattagatac cccactatgc ttagccctaa accttaataa ttagataaca
 541 aaattatttg ccagagaact actagctata gcttaaaact caatggactt ggcggactt
 601 tatatccatc tagaggagcc tgttctataa tcgataaacc cgcctttacc tcaccatctc
 661 ttgctaattc agcctatata ccgccatctt cagcaaacc taaaaaggaa taatagtaag
 721 cacaagaaca aacataaaaa cgttaggtca aggtgtagcc aatgagatgg gaagaaatgg
 781 gctacatctt ctttccaaag aacaataacg aaactcttta tgaaattaa gaacaaagga
 841 ggatttagta gtaaattaag aatagagagc ttaattgaat agagcaatga agtacgcaca
 901 caccgcccgt caccctcctc aactaaata aatttatata tataaataat ttctaaattg
 961 atatttacta gaggagataa gtogtaacaa ggtaagcata ctggaaagtg tgcttggat
 1021 aatcacagtg tagcttaact acaaagcactc tggcctacac ccagaagatt tcatagcaaa
 1081 tgaacacttt gaactaatac tagcccttat atttattaa ttaaattatt ttattatata
 1141 aattaaatca ttaacttat taaaagtatt ggagaaagaa attataagac aggagcaata
 1201 gagtaagtac cgtaaggga agatgaaaga acaattaaaa gtacaaaaaa gcaaagatta
 1261 tacctgttac ctttgcata atgaattaac tagaaaaat ttgactaaaa gaatttaagc
 1321 caaatacccc gaaaccaaac gagctaccta agaacaattt tatgaatcaa ccgctctatg

1381 ttgcaaaata gtgggaagat tottaggtag aggtgaaaag cctaacgagc ttggtgatag
1441 ctggttaccg aaaaaatgaa ttttaagtca actttaaaact taccaaaaga acaccaaatc
1501 cacacgtaag tttaaaatat agccaaaaga gggacagctc tttaggaaaag gaataaacct
1561 ttaatagtga ataaacatta aatatttata aaaccattgt aggcctaaaa gcagccacca
1621 atagagaaag cgttcaagct caacattaat ttttctctaa ttccataaag aataatcaat
1681 tcctattata aaaattgggt taatctattc atatatagat gagatactgt taatatgagt
1741 aacaagaata ttattctcca agcacaagtg tataacaact cggatgtcca ttggttagtta
1801 tcaagtcata gatgaaaact acaaattaaa ctttctaaat actatttgtt agoccaaacc
1861 aggagtgtc taaggaaaga ttaaaagaaa taaaaggaaac tcggcaaaaca agaaccocgc
1921 ctgtttacca aaaacatcac ctotagcata aaaagtatta gaggcattgc ctgcccagtg
1981 actagagttt aacggccgcg gtatcctgac cgtgcaaagg tagcataatc acttgttctc
2041 taattagggc ctagtatgaa tggctaaacg agggttcaac tgtctcttat ttccaatcag
2101 tgaattggc ctttcagtga agagactgaa ataaattaat aagacgagaa gaccctatgg
2161 agctttaatt tgtcagtta attattacta tttattaacc taatggttta ttatacaata
2221 acctaaactt aaaatttcgg ttgggtgac ctccggagaaa aaagaatcct ccgaatgatt
2281 ataacataga cctacaagtc aaagtaaaca acctatctt attgatccaa ttcttttttg
2341 atcaacggac caagttaccg tagggataac agcgcaatcc tattaagag ttcatatcga
2401 caattagggc ttacgacctc gatgttggat caggacatcc caatggtgca gaagctatta
2461 atggttctgt tgttcaacga ttaaagtcct acgtgatctg agttcagacc ggagcaatcc
2521 aggtcggttt ctatctatta acaatttctc ccagtacgaa aggacaagag aatggggcc
2581 acctataaaa taagcgcgcc aactcaatt tatgaaaaaa tctaataaaa gtaagtacgt
2641 atatttatag cctagacaag gcaattaggg tggcagagcc aggaattgc gtaagattta
2701 aaaccttgtt cccagagggt caaatctctc ccctaatagt gtattttatt aatattatca
2761 cactcctagt acctatcctt attgctatag ccttctgac attagtagaa cgaaaaattt
2821 taggctatat acaattacgt aaaggcccta acgtagtagg cccttatggg gtctacaac
2881 catttgcaga tgccttaaaa ctatttatca aagaacctat acgcccctca accacatcca
2941 tctcactctt tattattgca ccaacctat cactcacact agcccttagt ctatgaattc
3001 cactaccaat acctaccct ctaattaacc taaacttagg tattttattt attctagcta
3061 catcaagctt ttctgtttat totattctat gatcaggatg agcatccaac tcaaaaattt
3121 ccctatttgg tgcctacga gcagtagccc aaacaatttc atacgaagta acaatagcta
3181 ttatccttct atcagttctg ttaataagtg gctcattctc actacaaaaca ctcatctata
3241 ctcaagaaca tatatgactt atgttccaa cctgacctat ggctataata tgatatact
3301 caacctagc agaaactaac cgagctccat ttgacctaac agaagggggag tcagaattag
3361 tttcaggctt caatgttgaa tacgctgcag gtcatttgt actattcttc atagccgaat
3421 acataaatat tattcttatg aatgctctta catctattat tttcctaggt ccaatacata
3481 aaatttacta occagaattt tactcaatta actttataac agaaactcta attctatcaa
3541 caacattttt atgaattcga gcatcatatc cacgctttcg ttacgaccaa cttatacatc
3601 ttctatgaaa aaacttctta ccaactcacac tagcattctg catatgacat atctccttgc
3661 ctattttcat agcgagtatc ccaccataca tatagaaata tgtctgataa aagagttact
3721 ttgatagagt aaattataga ggtttaaatc ctcttatttc taggataaca ggaattgaac
3781 caactcctaa gaattcaaaa ttcttctgtc taccataca cttaatccta aatagtaagg
3841 tcagctaatt aagctatcgg gccataacc cgaaaacgtt ggtttagatc cttcccgtac
3901 taatcaaccg tattaccctt attattattt acttaactat tcttataggt ccagtaatta
3961 ccatatctag cactaacctt ctacttataat gagtaggttt agaacttagc cttctagcaa
4021 tcattccatt attaattaac aaaaaaac cccgatcaac tgaagctgca accaaatatt
4081 ttgtaaccga agcaacagct tcaataatta ttttactagc tattattcta aactataaac
4141 aacttggatt atgagtattt caacaacaaa caaaccttt acttcttaac attacactaa
4201 ttgcaactac aataaaatta ggocctggcc cttccactt ctgattacca gaagtaaac
4261 aaggatttca acttcacaca ggactaattc tacttacatg acaaaaaatt gctccactat
4321 caatattaat tcaaaactac tatctccttg atcatacctat cattataatt ttctcaattt
4381 tatcaatctt tattggagca tgaggtggac ttaaccaaac tcaaatacga aaaattatag
4441 catattcatc tattgtctac ataggatgaa tattagctat tttaccattt aaccaacta
4501 taacctcctc aaacttaata atttatatta ttctcacaat tcctatattc ctgtacata

4561 tatttaactc atcaacaaca ataaactcaa tttcacttat atgaacaaa actccaacaa
4621 ctctaattat aattccacta atccttttat ccctaggagg cttaccccca cttacaggct
4681 tcctacccaaa atgagccatt attacagaat tattaaaaaa caattgtcta attataacca
4741 ctataatagc tataatagct cttctaaatc tattctttta tatacgccta atttactcta
4801 ccacacttac catatttccc actaacaata actcaaaaat attaacccat catacaaac
4861 ataaatataa tattatcctt ccatcactta ccattattag cacactgata cttcctctct
4921 caccacaact catcatttag aagtttagga tatttatagt ccaagagcct tcaaagcct
4981 aagaaaacga acaagtttaa cttctgctaa ggactgtaag actacttctc acatctattg
5041 aatgcaaadc aattgottta attagctaa gacctcacta gattgtagg aatcaaacct
5101 acgaaatfff agttaacago taaataccct acttctggct tcaatctact tctcccgcct
5161 atcagaaaag gggcgggaga agccttagta gaatattatc tacaccttgc aatttgcatt
5221 tcgacatgac taacacctta agacttgga aaaagaggac ttaaacctct gtatttagat
5281 ttacagtcta atgcttactc agccatttta cctatgttca ttaaccgttg attattctca
5341 accaatcata aagacattgg cacattatat ctctattcgc gcgcttgagc tggaaatagta
5401 ggaaccgat taagtatttt aatccgagcc gaactaggac aaccggcgc acttttagga
5461 gatgacccaaa tctataatgt cattgtcact gccatgcat ttgttataat tttctttata
5521 gttataccaa taataattgg aggcttogga aactgactag taccactaat aattggagcc
5581 ccagacatag catttccacg aataaacaat ataagctttt gacttttacc accatcattt
5641 ctactcctac tagcatcctc tgtagtagaa gcaggagcag gaactggatg gacagtttac
5701 ccaccactag caggaactt agctcacgct ggagcatcag tagacttaac aattttctct
5761 cttcatttgg ctggggctc tccaatctta ggagctatta actttattac aactattatt
5821 aatataaaac cacctgccat aactcaatat caaacaccat tattcgtctg atcagtgtta
5881 attacagctg tctgtctttt attatcactc ccagtagtag cagcaggaat tacaactactg
5941 ttaaccgacc gcaatcttaa tacaaccttt tttgatcctg caggaggtgg agaccaatt
6001 ctctatcaac acctattctg attccttggc caccctgaag tttatattct tattttacca
6061 ggattcggaa ttatttccca tgtagttact tactactctg gaaaaaaga accatttggc
6121 tatataggaa tagtttagc cataatgtca attggtttcc taggatttat tgtatgagcc
6181 catcatatat ttacagtagg cttagatgta gacacacgag cttacttcaac atctgccaca
6241 ataattattg ctatcccaac aggtgtcaaa gtatttagct gacttgcaac actacacgga
6301 ggtaataatta aatgatcccc agctatatta tgagccctag gctttatctt cttatttaca
6361 gtaggaggac taactggaat tgttctatca aactctccac tagatattgt actccaagac
6421 acatactatg tagtagcaca ttccactat gtcttatcca taggagcagt attcgtatt
6481 atagctggat tcgttcactc atttccacta tttacaggct acactctaga cgatattga
6541 gcaaaaactc actttgccat tatatttcta ggagttaaca taacattttt tctcaacat
6601 ttcttaggac tatcaggaat accacgacga tactcagact acccagatgc ctatactaca
6661 tgaaacacag tgtcatctat aggatcattc atctccctaa cagcgttct tctaataatt
6721 tttataatft gagaagcatt tgcctcaaaa cgagaagtac taacagtaac atactcctca
6781 accaacctag aatgacttca cggctgcccc ccaccttacc atacattcga agaaccttca
6841 tatgttaaag taaaataaga aaggaaggaa tcgaaccccc taaaattggt ttcaagccaa
6901 ccccataacc tctatgtctt tctcaatgag atattagtaa aataattacg taacctgtc
6961 aaggttaaat tatagactaa aatctatata tcttatatgg cttatccatt tcaattaggc
7021 ttacaagatg ccacatcccc tattatagaa gaactgataa actttcatga ccacacacta
7081 ataattgfff tcttatttag ttccctagta ctatatatta tttcacttat attacaaca
7141 aaacttacac acacaagtac catagatgoc caagaagttg aaactatttg aaccatttta
7201 ccagcagtaa ttcttatttt aattgcccct ccctcactac gaattctata tataatagat
7261 gaaattaata accctgtatt aacagtaaag actataggtc atcaatgata ctgaagctat
7321 gagtacactg actatgaaga cctatgcttt gattcctata taatcccaac aaacgatcta
7381 aaaccaggtg acttacgact actagaagta gataatcgag tagtactacc aatagaacta
7441 ccaatccgaa tactaatttc atctgaagat gtacttact catgagctgt tccctcacta
7501 ggactaaaaa ctgacgctat cccaggacga ttaaatcaag caacagtaac atctaatcgc
7561 ccaggcctat tttacggta atgctcagaa atttgcgggt ctaaccacag ctttatacct
7621 attgtcctag agatagtgc acttaacat tttgaaatt gatcagcctc aataatttag
7681 tttcattatg aagcttagag cgttaacctt ttaagttaaa gttagagacc ttaaatctcc

7741 atagtgatat gccacaacta gacacatcca catgatttat tactattatc tcatctatgg
7801 ctacactatt tgtcctatit caacttaaaa tttcttcaca atcattccca acccctcctt
7861 ctcccaagac tttctctact caaaaaaaca aaaacccttg agaatcaaaa tgaacgaaaa
7921 tctatitggc tctttcatca ccccttcagt aataggatta ccaatcgtag tcaccattat
7981 catatitcca tctattctat ttocatcatc agaacgttta attaacaacc gccttcactc
8041 attccaacat tgactaatta aacttattat taaacaaata atattaattc atacacccaaa
8101 aggacgtacc tgaaccttaa taattgtttc ctttaattata tttattggct caacaaacct
8161 actaggtctt ctccctcata cattcacacc tactactcaa ctatcaataa accttagtat
8221 agcaattcca ctatgagcag gagccgtact actaggattt cgtcataaac taaaaagctc
8281 tctagcccac tttcttcac aaggaacacc aatctcacta atcccaatac tcattattat
8341 tgaacaatt agcctattta tccaacctat agcccttga gtacgattaa cagccaacat
8401 caccgcagga cacttactta tgcatacat tggaggagca actttagtcc ttataaacat
8461 tagcccacca actgtacta ttacatttat tattcttctc cttcttacag tactagaatt
8521 tgotgtagcc ttaattcaag catatgtatt tactttacta gtaagccttt atttcaatga
8581 taacacataa tgacacacca aaccatgca tatcatatag ttaaccaag tccatgacca
8641 ctaacaggag ctttttcagc tcttcttcta acatccggat tagtaatatg atttcactac
8701 aactcaacta cactattatc tctaggactc ttagccaata ctcttactat atatcaatga
8761 tgacgagata ttatccgtga aggaacctac caaggccatc acactcctat gtacaaaaa
8821 ggcttacgat atggaataat cttgtttatc gtctccgaag tattcttctt cgcaggcttc
8881 ttctgagcct tttaccactc aagcctagta ccaacacacg atttaggagg ctgctgacca
8941 ccaacaggaa ttaccccact taatccacta gaagttccac tattaataac atcagttcta
9001 ctgacatcag gtgtttcaat tacatgagct catcacagcc taatagaagg aaaacgaaac
9061 catataaatc aagccctatt aattactatt ctgctaggac tttatitfac tattcttcaa
9121 gcctcagaat atitgtaaac tcctttctct atctcagacg gaatttacgg ctccactttc
9181 tttatagcaa caggatttca tggattacat gtaatcattg gaacctcctt tctaattgta
9241 tgctctttac gacaattaaa gtttcatttt acatcaaaac atcacttcgg atttgaagca
9301 gcagcatgat actgacactt ttagaogta gtctgactat tcctttacgt ttocatctat
9361 tgatgaggat cttattctct tagtataatt aatataactg acttccaatt agtagattct
9421 gaaaataatc agaagagaat aataaattta gcaattatit tacttatcaa tactacacta
9481 tcactaatc taatctcagt agcattctta ctcccctaca taaatitata ttcagaaaaa
9541 acaaatccat acgaatgtgg atttgcctc tcaagctctg cccgcctacc atititcaata
9601 aaatititcc tagtggctat tacattctc ttatttgatt tagaaattgc tcttctgctc
9661 ccacttccat gagcaattca aacaaccaac attactacta taataacctc agccttaatc
9721 ttagtaacta tcttatcatt aggactaacc tacgaatgaa ctcaaaaagg acttgaatga
9781 acagaataat tggtaattag ttaaaaaa ataatgatt tgcactcatt agattatgat
9841 ttactcata attacaaaa atgtcatctg cctttattaa tctcacatta gcctttacat
9901 tgcctttatt aggaaccctc acattccgct ctacacatac atccaccctt ctatgcttag
9961 aaggaataat actatcctta tttatcataa cctcaatagc ctctttaaac gctaactcaa
10021 tagcatcaat gcctatccca attaccattc tagtgtttgc agcatgtgaa gcagcagtag
10081 gtttggccct attagttaaa gtatctaaca cctacggcac tgactatgta caaaatctta
10141 atttactaca atgctaaaaa ttatititcc atcattaata cttctccac tgacatgact
10201 ttcaaaacca aaaataacct gaaccaatgt aacttcttac agcttcttaa ttgcttaat
10261 tagtctctca attatatgac aaacaaacga aaccctctc aacttctcaa ctatattctc
10321 ctctgaccct ctatcgacc cactaattat tttacaacc tggcttctac ccttaatatt
10381 aatagcaagt caaatcacc taaaaaaga aaaaactatg caccaaaaac tctatatctc
10441 cctactagtt agcctacaaa ttctactagt aataacattc tcagcaactg aactaattat
10501 attctatata ctatttgaag ccactctaact tctacgcta attattatta cacgatgggg
10561 taatcaaaaca gaacgattaa atgcaggaat ttacttctta ttttatacac ttattggctc
10621 aatccactt ttaattgcct taatititac gaaacctct ataggaacat taaactitc
10681 tatgctatca cttatgcca gtcttttaga ttctttatga tctaacaata tcttatgact
10741 agcatgcata atagcattta taattaaaat accactatac ggtgtocacc tctgactacc
10801 aaaagcccat gttgaagcac caattgcagg ctctataatt cttagcagta ttctactaaa
10861 gttaggaagt tatgtataa tacgaattgc cattatacta gatcccttaa ctaaatitac

10921 agcttatacct ttcattttat tatcattatg aggaataatt ataacaagtt caatctgctt
10981 acgcccacaac gacctaataat cactaatcgc ataactcttca gttagccata tagccctagt
11041 cattgcatca atcataatcc aaaccccctg aagctttata ggagctacaa tactgataat
11101 tgcccattgca ttaacttcat cactcctatt ctgtctggca aactcaaaact atgaacgaat
11161 tcacagccga actataatta tggcccggag tcttcaaata atcttcccac tcatggcaac
11221 ctgatgactt gtagcaagct taaccaacct agctttacca ccatcaatta acctcattgg
11281 agaattatit atcactatat cactattctc ttgatctaat ttttctatta ttcttatggg
11341 aattaacatt atcattacag gaatatactc aatatatata attattacaa cccaacgagg
11401 aaaactaact agccatataa ataaccttca accttcacac acacgagaac taacacttat
11461 agctctacat attatcccgc ttctcctctt aacaattagc ccaaaaattta ttacaggctt
11521 aacaatatgt aaatatagtt tacaaaaaca ttagactgtg aatctaaaaa caggaattaa
11581 tactccttat ttaccaagaa agaatgcaag aactgctaac tcatgcctcc atgtttaaac
11641 acatggcttt cttacttita taggataaaa gtaatccatt ggtcttagga accaaaaaca
11701 ttggtgcaaa tccaataaaa agtaataaac actattacat cttccatcat agcaatitit
11761 attgtttctac tcacgcccat cctaatttct tcaactaata tattaaaaa aatcaatitit
11821 ccgctatacag ctacaatata catcaaaactc tcattttttc tcagctttct cccactaata
11881 atatitititit actataatac agaatatata attacaactt gacactgaat aacaatcaac
11941 tctatcgagc taactatgag cttcaaaaata gattatitit caattatatt tttatcagt
12001 gccctcttcg tgacatgac tattatacaa ttttcttcat gatataata ctcagatcct
12061 cacattaatc gatttattaa gtacctacta ttatitctga ttactatatt aattctaacc
12121 tcagctaaca acttatttca actatitatt gggtgagaag gagtaggaat tatatcctit
12181 cttctaattg gctgatgatt tggacgatca gacgcaaaata cagcagccct gcaagcaatt
12241 ttataaacc gaatcggaga catcggattt atittagcca taacatgatt ctgcctaaat
12301 ataaactctt gagaacttca acaaatitita ttaaccgaaa ataataacct aattccactt
12361 ctaggactac taattgcagc aacaggtaaa tcagctcaat ttggtcttca cccatgacta
12421 ccctccgcta tagaaggacc aacaccagtt tcagctctac tgcattcaag cacaatagta
12481 gtacgaggca tttitctact agtacgattt catcccctca cctcaaaaca ccaaaaatc
12541 cttacaataa tgttatgtct tggaggactc actaccctat ttaccgctat ctgcccctc
12601 acacaaaaatg atattaaaaa aatcgtagca ttctctacat caagccaact aggcctaata
12661 atagtgcac taggtattaa ccaaccatat ttagccttcc tacatatctg cacacagca
12721 ttcttcaaag ccatactitit catgtgctcc ggatcaatca ttcacagcct taatgatgaa
12781 caagacattc gcaaaatagg aaacataata aaaaccatac cattcacac atcctgttta
12841 acaactggaa gccttgctct cacaggaaata ccattcctaa caggctitita ttcaaaagac
12901 ctcatctatg aagccataaa cacctgtaat accaacgctt gagccctaata aattacacta
12961 atcggccat ccataacagc tttttacagc atacgaatca tctacttctg cacaatgaca
13021 aaaccacgat tcccaccaat aatcttaatt aacgaaaacg accccaacct tataaatcct
13081 attaaacgac tagcactagg aagcattata gctggattct ttatctctta caatattcca
13141 ccaaccaata tccaatcct tactatacca tgatatttaa aaaccacagc actattcatt
13201 tctatititg gattcctaata agccttagaa ttaacaatc tcacactcaa actitctata
13261 acaaaaaaaa atctatactc ctcatititca acatcactag gctacttccc atcaattatt
13321 caccgaatta ttccaacaa aacactaaac ctaagtttaa aaacatcatt aatctccta
13381 gacctctatt gattagaaaa ttcaattcca aaacctct caacaatata ttctatgca
13441 tctaaattat taactaacca aaaaggacta gtcaaatat atitctatc tttititctt
13501 tccatcttat taacagttat tctatatatt attaatctcg agtgatctca ataataataa
13561 aaattcccgc aaataaagat caaccagcca ctaccatcat tcaagtagca caactgtata
13621 ttgccgcaac tccaataccg ccctccaata ttacaccaac atcatcaacc tcatacatta
13681 gtcaatcacc caacttatca aatcaacta catctacccc atcataatag tcaattaaat
13741 atataaaaa caactctatt aacactccaa taattaaaga accaaaaaca aatcaattag
13801 accctcaagt ttctggatac tctcagtag ctatagcagt agtataacca aataccacca
13861 tcataccacc taagtaaatt aaaaacacta ttaaacctaa aatgatcca ccaaaccta
13921 acaccataag acaacctaca aacccgctca taattaaacc aaaccacca taataggag
13981 aaggctitaa agccaacccc agacaccag ttaaaaataa taaacgtaa acaaaaatat
14041 aattagctat ttttccaca cagcatttaa ctgtgactaa tgacatgaaa aatcatgctt

14101 gtaattcaac tatagaaact aaatgacaaa catccgaaaa acacatccac tactaaaaat
14161 tattaatcac ttttcatcg acttaccgc tccatctaat atttcatctt gatggaactt
14221 tggctcccta ctaggtgtct gcctaattat tcaaatcgtt acaggcttat tcttagctat
14281 aactacact tcagatacaa taacagcatt ttcacagta acacatattt gtcgagacgt
14341 aaattacgac tgactaatc gatatacaca cgcaaacgga gcctcaatat ttttcatctg
14401 cttgtttctt catgtaggac gaggaatata ctacggatca tatakattta tagaaacatg
14461 aaacattgga gttattttac ttttgcagt aatagccaca gcattcatag gttatgttct
14521 cccatgaggc caaatatcct ttgaggagc aacagtaatt acaaacctcc ttccagcaat
14581 tccctacatc ggaaccacc tagtagaatg aatctgagg ggcttctcag tcgataaggc
14641 tactctaaca cgcttcttcg ctttccactt tattcttcca ttcattattg cagctctagc
14701 agttgttcac ctgcttttcc tccacgaaac aggcctcaaac aatccaacag gcctaaactc
14761 agacgctgat aaaattccat tccaccata ctacacaatt aaagacatcc taggtatttt
14821 cattctagtg ggatttctaa taacctagt tctattctcc ccagatcttc taggagacc
14881 agacaactat ataccagcaa atccactcaa tactccccca cacattaaac cagaatgata
14941 tttctatatt gcctatgcaa ttctacgctc aattccaaac aaactaggag gactactagc
15001 cctaactctc tcaattttaa ttctagctct actaccattc cttcacactt caaaacaacg
15061 aagcctaata tttcgtccta tcaccaaaac actatactga atcctagtag ctaattttct
15121 taccctcacc tgaattggag gccaacaccg tgaacatcca tttgtaatta ttggccaact
15181 agcctcaatt agctatttct ccattattct tatcctcata ccactatcag gtatcattga
15241 agataaaatc ctaaaatgaa atctatgcc cgatagtata aacattactc tggctctgta
15301 aacaaaaaat gaaaaactac ttttctcagg gcacaaagaa gaaggagtgt ctccccacca
15361 tcagcaccoca aagctgatat tctcattaaa ctacttcttg tgtacataaa attacatagt
15421 acattcatal atctatgtat aaagtacatt aaattatttc cccctagcat ataagcaagt
15481 atattaatc aatgtatcaa gacattacaa acaatcgtac attaattata tccctcatgt
15541 ctatcaaaaca catacataaa ataatgtaa taaagacata tctgcttat catacataca
15601 ccattaagtc ataaacctt ctcttcata tgactatccc cttccacatt ttgtctatat
15661 ttctaccatc ctccgtgaaa ccaacaacc gccacatc gccctcttc tcgctccggg
15721 cccatttaac ttggggtag ctacacatga atttatcag acatctggtt cttacttcag
15781 ggccatcaaa tgcgttatcg ccatacgtt ccccttaaat aagacatctc gatggtacgg
15841 gttaatcag ccaagacggt cataactgtg gtctcgagca gttggtattt ttttattttt
15901 aggatgctgt gactcaacat agcgtcaag gcatgaagg cagccatca ttagctgga
15961 ctcaagtta aggatcatt atccacataa accaatcatc taagactatt tattaatgct
16021 ttttagacat aacacattta ataaactta cttctactta atcaaaccc ctcccccaa
16081 aaaatcctc aatgccaaac ccaaaaatc attaaaagat ttaactttaa tttataagt
16141 ataattctat tctagtagtt cataaaacta gactcacatt ttagtatctg tcaattttt
16201 attaaaatta ttttaattg tcaaaaactt acctcaccta aatcttctaa tattaattg
16261 tactaa

감사의 글

박사학위를 받기까지 저를 도와주시고, 응원해주신 많은 분들에게 감사를 드립니다. 우선 살아계셨다면 가장 크게 기뻐하셨을 아버님, 늘 저를 믿어주시고 응원해주신 어머님께 감사드리며, 하나뿐인 동생 종철이에게도 고맙다는 말을 전하고 싶습니다. 그리고, 잦은 야근에도 불구하고 항상 걱정해주고 격려해 준 아내와 많이 놀아주지 못해 항상 미안한 우리 딸들 유진, 유현에게 고맙고 미안한 마음을 전합니다. 또한, 일을 핑계로 집안일에 소홀한 저를 대신해 제 아내와 두 딸을 많이 보살펴주신 장인 장모님께도 감사의 말씀을 전하며, 공부하느라 바쁜 와중에도 누나의 일이라면 항상 적극적으로 도와주는 처남 웅이에게도 고맙다는 말을 전합니다.

석사과정에서부터 부족한 저를 이끌어주시고 박사학위를 받을 수 있도록 항상 저를 응원해주시고 지도해주신 김세재 교수님과 연구를 진행하는데 큰 도움을 주신 오홍식 교수님께 무한한 감사를 드립니다. 뿐만 아니라, 지난 수년간 저에게 많은 것을 가르쳐주시고, 이 자리까지 올 수 있도록 음양으로 적극 지원해주신 정용환 소장님과 평소 일에 치여 지내면서도 어려움이 있을 때마다 항상 도움을 주시는 박수영 부장님, 연구방향 및 결과 분석에 많은 도움을 주신 선배 한상현 박사님께 진심으로 감사를 드립니다.

현재 몸담고 있는 생물종다양성연구소에서 같이 근무하고 있는 바이오기술개발부의 고미희 박사님, 원종이형, 영민이에게도 감사의 마음을 전하고 싶습니다. 그리고, 항상 성실한 자세로 맡은 업무에 최선을 다하는 모습이 믿음직한 전형식형님께도 감사의 말을 전하며, 곤충연구에 연구소 업무에 정신없이 지내는 경식이, 지금은 퇴사했지만 그 동안 많은 고생을 함께 했던 지영이와 정아, 그리고 현재 연구소에서 각 분야 업무에 충실히 매진하고 있는 동료 연구원들에게도 고맙다는 말을 전합니다.

학부시절부터 시간이 날 때면 서로의 넋두리를 들어주고 서로에 대한 조언을 아끼지 않았던 후배 강성일 박사, 석사과정에 있을 때부터 같이 실험실 생활을 해왔던 박사과정 혜선이와 현재 열심히 연구에 매진하고 있는 정환이, 승우, 선

아에게도 감사의 말을 전함과 동시에 앞으로 좋은 연구 성과를 기대합니다. 그리고, 자주 모이지는 못 하지만 오랜 친구들 성하, 효진, 길홍, 종석, 정석, 후영, 세미, 희선, 은정에게도 그 동안 힘이 돼주어 고맙다는 말을 전하고 싶습니다.

마지막으로 지면을 통해 일일이 언급을 하지 못했지만 그 동안 저를 아끼고 사랑해주신 분들께 다시 한 번 진심으로 감사를 드립니다.