ANALYSIS OF RODENT MITOCHONDRIAL AND NUCLEAR DNA

by
Xiaobang Chen

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillement of the Requirements
for the Degree
Master of Science

McMaster University

MASTER OF SCIENCE (1998)
(Biology)

McMaster University
Hamilton, Ontario

TITLE: Analysis of rodent mitochondrial and nuclear DNA

AUTHOR: Xiaobang Chen, B.Sc. (University)

SUPERVISOR: Dr. G. Brian Golding

NUMBER OF PAGES: [xiii], [198]

ABSTRACT

Rodent mitochondrial cytochrome c oxidase subunit II (COII) and nuclear encoded adenine phosphoriboslytransferase (APRT) were PCR amplified and sequenced with special emphasis on the genus *Mus.* The analyses of these COII and APRT sequences along with the sequences retrieved from GenBank and 12S rRNA sequences from GenBank, were carried out using neighbor joining, maximum parsimony and likelihood algorithms. The COII phylogeny and APRT phylogeny agree with the species phylogeny tree previously established by morphological and molecular data. The analysis of COII sequences supports the view that *Mus spretus* is a sister species of *Mus macedonicus*, *Mus spicilegus*, *Mus musculus* and diverged prior to the divergence of *Mus spicilegus* and *Mus musculus.* The COII sequence data also provided evidence that *M. spicilegus* and *M. macedonicus* are sister species within the palearctic group, while *M. caroli* and *M. cervicolor* group together. The 12S rRNA gene sequence analysis showed that the subgenus *Pyromys* species (*M. platythrix* and *M. saxicola*) group together and the subgenus *Coelomys* species (*M. crociduroides* and *M. pahari*) group together. The APRT sequence analysis found two large inserts with homology to the rodent short interspersed repetitive elements B1 and B2 present in *Mus saxicola* and *Mus crociduroides* & *Mus pahari*, respectively. Unusually, these inserts occurred at identical sites in intron 2 of the APRT gene and were not shared by any other *Mus* species examined in this analysis. This finding indicates that 1) *Mus crociduroides* and *Mus pahari* are closely related subgenus species and the mobilization of the B2 element may have occurred before the divergence of the *Coelomys* subgenus. 2) the B1 and B2 elements may favour the same target sequence. The analysis of the APRT sequences provides evidence that *Mus caroli* and *Mus cookii* are distinct from the *Mus musculus* group. The rate variation of the sequences for the APRT genes and COII genes between different closely related species of rodents were examined by a likelihood ratio test of the molecular clock. The test for the COII gene showed that a molecular clock cannot be rejected for the species of the *Mus* genus. For the APRT gene the test revealed that a molecular clock hypothesis must be rejected for these closely related *Mus* species. The existence of different evolutionary patterns among the APRT genes and COII genes in the *Mus* species might indicate that these genes have been under different selective constraints during their evolution but these rate variations are unlikely to be due to any change of generation length.

ACKNOWLEGEMENTS

I would like to thank my supervisor Dr. G. Brian Golding for his encouragement, endless support and patience throughout my two years at McMaster. I would also like to thank Dr. T.M. Finan, my former employer and a member of my supervisory committee for his encouragement to pursue a graduate degree and his concern during difficult times of study.

Many thanks go to Fariborz Yazdani for his willingness to assist in experiments and computer analysis. I am grateful to Monique Muller for her constant encouragement, help and friendship.

Finally, I am most grateful to my parents and my brother for their faith in my ability. Thanks to my mother and my mother-in-law for their time taking care of my son. A very special thanks goes to my husband, Dongya for his advice and encouragement and my son Joshua for his respect and love.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1.1 Classification, Taxonomy, Systematics and Phylogeny

There is an amazing diversity of life in the world, both extant and extinct. To describe and to talk about these many organisms, there must be a classification of these organisms into groups. The very first recorded classifications were done by Aristotle. He built an elaborate system for classifying the species of the animal kingdom which began by dividing animals into groups. He used simple characteristics such as organisms having red blood, corresponding roughly to present day vertebrates, and those lacking it, roughly corresponding to invertebrates, to differentiate and classify these groups. He used further subdivisions to define subgroups, e.g. according to the way in which progeny are produced, whether viviparous, in eggs, and so on (Dunn and Everitt 1982).

Linnaeus (1758) essentially established the science of systematics and established a framework for describing and categorizing biological diversity. He formalized a hierarchical system of nomenclature with a binomial nomenclature for genera and species. This hierarchical system was initially independent of evolutionary theory, but later workers (e.g., Darwin 1859; reviewed by Mayr 1983) developed the notion that classification should be based on phylogenetic relationships.

1.1.1 Terms and concepts

Before proceeding further it is useful to clarify a number of terms that will be used in this thesis. Definitions for the terms character, taxon, taxonomic characters, homoplasies, classification, phylogenetic classification, taxonomy, systematics, phylogenetic systematics, phenetics, and cladistics follow.

Character: A character may be most any feature of an organism that may be described, figured, measured, weighed, counted, scored, or otherwise communicated by one biologist to other biologists.

Taxon: A taxon is a group of organisms of any hierarchical rank that is sufficiently distinct to be worthy of being assigned to a definite category (Mayr 1969). This definition implies that the delimitation of a taxon against other taxa of the same rank is virtually always subject to the judgment of the taxonomist (Dunn and Everitt 1982). Wiley (1981) suggests that a natural taxon is a group of species that exists in nature as a result of a unique history of descent with modification. Natural taxa exist whether or not there are any systematists around to perceive or name them.

Taxonomic characters: A taxonomic character is some biological character that is useful in the taxonomic description of organisms. Mayr (1969, page 121) defined a taxonomic character as an attribute of a taxon by which it differs or may differ from members of a different taxon. But the problem with this definition is that similar characters are used for taxonomic groupings at high levels (Wiley 1981, page 116). Hennig's concept of a "character" (Hennig 1966) is closely akin to Mayr (1969)'s "taxonomic character". Wiley (1981, page 116) defined a character as a feature of an organism which is the product of previously existing features, or a feature of a previously existing parental organism(s). A character in this sense is one product of the evolutionary process.

Homoplasies: Homoplasies are characters that display structural similarities but are thought to have originated independently of each other. Either from two different preexisting characters or from a single preexisting character at two different times or in two different species (Wiley 1981).

Classification: A classification is the ordering of organisms into taxa on the basis of their similarity and relationship as determined by or inferred from their taxonomic characters (Mayr 1969). Dunn and Everitt (1982) described a classification as the ordering of organisms into groups on the basis of their relationships. The relationships may be genetic, evolutionary (phylogenetic) or may simply refer to similarities of phenotype (phenetic). A classification is distinct from an identification of organisms in that a modern classification embodies a hierarchical ordering of taxa.

Phylogenetic classification: A phylogenetic classification is a classification that attempts to represent the historical (or phylogenetic) relationship hypothesized to exist among a given array of organisms (Wiley 1981).

Taxonomy: Mayr (1969) defined taxonomy as the theory and practice of classifying organisms. Wiley (1981) defined taxonomy as the theory and practice of describing the diversity of organisms and ordering this diversity into a system that conveys information concerning the kind of relationship between organisms that the investigator thinks is relevant. In other words, taxonomy comprises description, classification, as well as the theory and history of classification.

Systematics: Simpson (1961) states that "systematics is the scientific study of the kinds and diversity of organisms and of any and all relationships among them". Similarly Wiley (1981) suggests that systematics is the study of organismic diversity as that diversity is relevant to some specified kind of relationship thought to exist among populations, species, or higher taxa. Systematics is an attempt to understand the evolutionary interrelationships of living things, trying to interpret the way in which life has diversified and changed over time. All comparative data are potential systematic data. Molecular systematics is based on molecular data such as nucleic acids (DNA and RNA), proteins, and chromosomes.

Phylogenetic systematics: Phylogenetic systematics is one approach to systematics that attempts to recover the phylogenetic (or historical) relationships among groups of organisms and produces classifications that exactly reflect these phylogenetic relationships (Wiley 1981).

Phenetics: Phenetics is a system of classification based upon phenotypic similarities or differences without regard to phylogenetic relationships.

Cladistics: Cladistics is a method of classification that attempts to reconstruct phylogenies utilizing only characters that are believed to be synapomorphic (these are shared, derived characters).

1.1.2 Classifications; biological and phylogenetic

Classifications are systems of an ordering of things and involve grouping entities or phenomena and giving names to the resulting groups. There are three major types of classifications (Wiley 1981):

1. Classification of natural classes. A natural classes is one that contains individual entities whose origins and behavior are governed by natural processes. The entities placed in a natural class may be quite unrelated in a historical sense.
2. Classification of historical groups and individuals. These classifications are based upon inferred historical connections between the entities classified. Phy-

logenetic classifications are an example of such a historical classification.

3. Classification of convenience classes. These classifications have no particular basis in terms of process or history. They are a by-product of our desire to bring order into the study of nature. Many historical classifications have grouped phylogenetically unrelated classes such as those herbs of medicinal value.

Biological classifications are systems of words which are used to organize the diversity of life and/or to reflect man's estimate of nature's own organization. Biological classifications may be combinations of Wiley's three types. Biological classifications may also be hierarchical or non-hierarchical though all modern taxonomic biological classifications are hierarchical (Wiley 1981). Biological classifications can give a relative idea about the relationships by indicating group membership.

Phylogenetic classifications are ones which attempt to accurately reflect hypotheses concerning the phylogenetic descent of organisms and are usually done at the species level or above (Wiley 1981). The components of phylogenetic classifications are taxa. Any classification that reflects the phylogenetic relationships of the taxa classified is a phylogenetic classification.

1.1.3 Taxonomy

To summarize and make sense of the diversity of organisms the taxonomist customarily constructs a taxonomic hierarchy in which a taxon occupies a position in a nested scheme such as kingdom, phylum, class, order, family, genus and species. The hierarchy is intended to illustrate that different species within a given genus are more similar to one another than to species of other genera. A Linnaean hierarchy is one of several different systems for translating a phylogenetic hypothesis into a phylogenetic classification.

Phenetic relationships are based on the observed similarities in the properties of organisms. The most important phylogenetic relationship is the relative branching order of the taxa and this is called a cladistic relationship. Based on the similarity of DNA (or RNA) sequences a phenetic relationship between organisms also can be defined. One can use the methods of numerical taxonomy to classify organisms on the basis of their phenetic relationships, or a cladistic method can be applied. The latter can be constructed from nucleotide or amino acid sequence data.

1.1.4 Phylogenetic Systematics

Modern phylogenetic systematics was begun by the German entomologist Willi Hennig who published 'Grundzuge einer Theorie der Phylogenetischen Systematik' in 1950. In this book he proposed that classification should be solely based on the historical branching pattern of the organisms.

Wiley (1981) described some of his ideas as ...

1. The relationships lending to the cohesion of living and extinct organisms are genealogical ("blood") relationships.
2. Such relationships exist for individuals within populations, between populations, and between species.
3. All other types of relationship (i.e., phenotypic and genetic) are phenomena correlated with genealogical descent and thus are best understood within the context of genealogical descent with modification (quite literally "evolution")
4. The genealogical relationships among populations and species may be recovered (discovered) by searching for particular characters which document these relationships.

5. The best general classification of organisms is one that exactly reflects the genealogical relationships among these organisms.

Hennig's ideas were first discussed in an American journal by Kiriakoff (1959). Wide discussion of phylogenetic methods in English language journals came after the publication of Hennig's revised book 'Phylogenetic Systematics' in 1966. This book marks a turning point in the history of systematic biology. In a series of publications Hennig (1950, 1965, 1966, 1969, 1975) formalized the practice of systematics and established a series of principles in order to obtain uniform treatments and better classifications. Some of these ideas are considered basic to the discipline (e.g., monophyly) while others have since been discarded (e.g., hierarchical ranks determined by absolute age). Never the less, his ideas formed an important focus for further progress and helped to eliminate subjectivity and arbitrariness in systematics (Wiley 1981).

1.2 DNA taxonomic methods

Before the 1960's systematics was largely based on analysis of morphological and behavioral variation. With the ability to rapidly determine the structure of chromosomes and the sequence of nucleic acids (DNA and RNA) and of proteins, molecular characteristics are becoming more commonly used. Molecular characteristics are now commonly used in studies of population structure (e.g., geographic variation, mating systems, heterozygosity, and individual relatedness), identification of species boundaries, and estimation of phylogenies (Hillis and Moritz 1990).

1.2.1 The application of molecular data in systematics

Early applications of molecular data to systematic problems mainly dealt with proteins. The most useful approaches developed for proteins were isozyme electrophoresis and immunological techniques. In isozyme electrophoresis changes in the amino acid sequence of proteins can sometimes be detected by corresponding changes in the mobility of the protein in an electric field. A substitution which causes an acidic amino acid to be replaced by a basic amino acid will be reflected in the rate in which the protein migrates within an electrophoretic gel. Other amino acid replacements that might not directly affect charge may never-the-less affect the electric charge potential of the protein and can also be detected. Immunological techniques are based on the assumption that a stronger immunological reaction will be observed between more dissimilar proteins. Immunological approaches were pioneered by Nuttall in 1904 using precipitin tests. Precipitin tests rely on mixing antibody and antigen in solution and measuring the resultant antibody-antigen precipitate. The greater the amount of precipitate, the more closely related the two species. With the development of this method, some prominent studies of the relationships and times of divergence between hominoids have been achieved (Goodman 1961; Goodman 1963; Sarich and Wilson 1966; Sarich and Wilson 1967). Relationships among rodents have also been studied using this method (Sarich 1985). Electrophoresis was used to demonstrate a wealth of protein variation within and among species (Hubby and Throckmorton 1965; Hubby and Lewontin 1966; Lewontin and Hubby 1966). Comparisons of amino acid sequences were the first indication of the possibility of a molecular clock (Zuckerkandl and Pauling 1962). Amino acid sequences have been used extensively to estimate phylogeny (Goodman, Miyamoto and Czelusniak 1987). Similarly, the isozyme electrophoresis approach has generated a massive comparative data base and is a widely used approach in molecular systematics (Avise 1974; Avise and Aquadro 1982; Buth 1984).

The use of electrophoresis is generally considered better than immunological methods because the later methods require relatively large amounts of antiserum and antigen for assays, which make it difficult to study small species or species from which only small samples can be obtained. In general electrophoretic methods can also more readily detect single amino acid changes.

Another important application of molecular data in systematics makes use of the number of chromosomes and the patterns among chromosomes. Chromosomes can be identified on the basis of their morphology including their relative size, centromere position, and secondary constrictions. Variation in chromosome structure and number provides a valuable source of genetic markers within and between species. Therefore, the number, size, and shape of chromosomes have also been used to establish phylogenetic relationships among related species (White 1973). The development of chromosome banding techniques permit chromosomes to be identified by their induced banding patterns and then the phylogenetic relationship among species can be inferred from these data. *Drosophila* posses giant polytene chromosomes in the larval salivary glands and are widely studied by this method. The phylogeny of 101 Hawaiian *Drosophila* is a remarkable example of the use of banding pattern in polytene chromosomes to determine phylogenetic history (Carson and Kaneshiro 1976).

Another application of molecular cytogenetics is based on *in situ* hybridization of nucleic acids. Using *in situ* hybridization, specific genes and other kinds of sequences can be localized to particular chromosomes and/or to regions of chromosomes (Gall and Pardue 1969; John, Birnsteil and Jones 1969). Many of the most successful applications of *in situ* hybridization have involved the use of repetitive sequences such as those coding for ribosomal RNA, tRNAs, or histones, as well as highly repeated satellite DNA sequences (MacGregor and Sessions 1986; Cremisi *et al.* 1988; Deininger and Daniels 1986). Other current efforts are using *in situ* hybridization to map the chromosomal location of large fragments cloned in bacteriphage P1 between different *Drosophila* species (Lozovskaya, Petrov and Hartl 1993). A better use of molecular cytogenetic information for phylogenetic analyzes requires a better understanding of the molecular basis of chromosome structure and function.

A third source of molecular data is nucleic acids. The development of rapid methods to sequence DNA has lead to rapidly increasing use of these data in systematics. Amino acid sequence data has been largely replaced by nucleic acid sequence data for most systematic applications mainly due to the fact that amino acid sequencing is technically much more difficult and time-consuming than is nucleic acid sequencing and that amino acid sequences can be easily deduced from nucleic acid sequences. There are three major applications of comparative nucleic acid sequencing in systematic studies (Hillis and Moritz, 1990): (1) construction of molecular phylogenies to evaluate the evolutionary diversification of particular genes or gene families and their RNA or protein products, (2) the tracing of organismal genealogies within species, and (3) construction of species phylogenies to evaluate macroevolutionary patterns and process. In addition to DNA sequence data, nucleic acids are also used to reveal DNA hybridization and RFLP data.

1.2.2 Properties of molecular data

As noted above molecular data in recent decades have been widely used in the studies of systematics and evolution. What is the advantage of these molecules? First of all, molecules are genetically inherited and can be superior taxonomic characters to morphological traits. A morphological change may represent many or only a few genetic changes. In addition, morphological changes in one character might be caused by changes in multiple genetic loci. In contrast, there is generally less ambiguity in deciding whether or not molecular characters are homologous (Patterson 1987). Most molecular data have a clear genetic basis and the total data set is limited only by the genome size. Second, molecules are universal and repeatable. In the whole genome of an organism there are many loci which can provide useful phylogenetic information. Third, molecules are more precise than phenotype. Molecules as biological characters avoid any potentially arbitrary judgments of the investigator. Fourth, some systematic problems can only be addressed with molecular data (Hillis 1987) such as the evolutionary relationships among bacteria and unicellular eukaryotes (Sogin, Elwood and Gunderson 1986; Sogin *et al.* 1989).

In general using both molecular and morphological data will provide much better descriptions and interpretations of biological diversity. Finally, a molecular approach is potentially powerful for generating and testing competing phylogenetic hypotheses. Therefore, studies about the history of an organism using molecules are feasible. However, there are several controversies in molecular systematics, such as the relative value of molecular versus morphological data, the constancy of rates of molecular evolution, the neutrality of molecular variants, the types of data that should be collected, the various philosophical approaches to analyzing data, and the ability to detect "homology" between molecular characters (Hillis and Moritz 1990).

There has been debate over whether molecular or morphological features are inherently better sources of information for estimating phylogeny (Patterson 1987). The real concern should be whether the characters chosen exhibit a level of variation that is appropriate to the question posed, whether the characters have a clear and independent genetic basis, and whether with these data if it is possible to compare and combine phylogenetic hypotheses (Hillis and Moritz 1990).

It is necessary to understand the possible meanings of homology in molecular biology. Comparisons of homologous sequences can be divided into three types. A comparison of homologous sequences that diverged after speciation is termed orthologous (Fitch 1970). A comparison of homologous sequences formed after a gene duplication is referred to as paralogous (Fitch 1970). A comparison of homologous sequences that arose via lateral gene transfer is called xenologous (Gray and Fitch 1983). These distinctions are important because only orthologous sequences can be used to infer the phylogeny of species (Hillis and Moritz 1990).

1.2.3 The advantages and disadvantages of DNA data

We have discussed the properties of molecules as phylogenetic markers. What we are concerned with in this thesis is DNA sequence data. As the primary data in phylogenetic estimates, DNA sequence data can be obtained from nucleus, the mitochondrion, and the chloroplast. Approaches used include analysis of hybridization and dissociation of DNA, use of restriction endonucleases to detect base substitutions and rearrangements, and comparisons of the primary sequences (DNA hybridization, RFLP analysis, and sequencing). The DNA-DNA hybridization approach makes use of the homology between complementary strands of DNA. Typically complementary strands of DNA for each of two species are separated at 100^0C. Their DNA is mixed and is then quickly cooled. DNA absorbs ultraviolet light at 260 nanometers but double stranded DNA absorbs far less ultraviolet light than does single stranded DNA. In this way the proportions of denatured and reassociated DNA can be measured with a spectrophotometer. Heteroduplex formation can occur between well-matched sequences from each species. Never the less, the double-stranded heteroduplex will contain base pair mismatches due to any evolutionary divergence from a common ancestor. Since the extent of mismatch affects the melting temperature, the depression of melting temperature in a heteroduplex relative to a homoduplex serves as an index of divergence between the sequences compared. DNA-DNA hybridization thus provides an estimate of the amount of sequence divergence between genomes but cannot provide discrete character data.

A second approach to study divergence and polymorphism is RFLP analysis. This assay for sequence variation compares the number and size of fragments produced by digestion of the DNA with restriction endonucleases. Base substitutions or insertions/deletions that destroy or create restriction sites will cause changes in the number and size of fragments. These variations are called restriction fragment length polymorphisms (RFLPs). In contrast to DNA-DNA hybridization, RFLP analysis is more laborious and expensive, but provides greater information on the nature of differences between two DNA sequences. RFLP analysis is simpler and cheaper than primary sequence determination, but offers comparatively less information about the evolution of the sequence itself. In comparison the primary sequence offers extremely high resolution and yields character data that can be used to provide a variety of measures including divergence.

DNA is a stable molecule and is far less chemically reactive than proteins. DNA isolated from fresh tissue samples can be stored by freeze-thaw processing or alcohol preservation with only minimal shearing. It is valuable for hybridization, analysis of restriction fragment variants, gene cloning, and sequencing. Unfortunately, DNA is not always recoverable from museum specimens. Depending on the age of the specimens, the recovery of DNA can be markedly reduced. In general this DNA is often single stranded and degraded to between 200 and 300 base pairs, which is too small for restriction fragment analysis (Hillis and Moritz 1990). In these cases mitochondrial DNA is preferred because of the multiple copies present per cell.

The taxonomic characters used in sequence comparisons are individual nucleotide positions. The nucleotides are the basic units of information encoded in an organism's DNA. The potential size of informative data sets are immense: some species contain more than 10^{11} nucleotide pairs per haploid genome. To use DNA sequence positions in phylogenetic studies, orthologous sequences must be aligned. The number and size of orthologous sequences that can be aligned will differ depending on the level of comparison, but for most studies potentially informative variation is essentially inexhaustible.

1.2.4 Methods of DNA data analysis

Molecular phylogenetic reconstruction is an attempt to estimate phylogenetic information from molecular data. Here we are concerned with DNA sequence data. The methods of DNA data analysis discussed here include two parts: techniques for collecting DNA sequencing data and methods for analyzing DNA sequence data.

Techniques for collecting DNA sequence data

There are a number of strategies for obtaining DNA sequence data for use in systematics, all methods have four basic steps (Hillis and Moritz 1990). First, a particular target sequence must be identified that contains an appropriate amount of variation across species or individuals for the problem that is to be addressed. Second, large numbers of copies of the target sequence must be isolated and purified from each individual to be examined. Third, the purified DNA must be sequenced. Finally homologous sequences must be aligned.

Choice of sequence: The general principles to identify a proper target sequence can be classified into several parts: (1) sequences should display sufficient variation to enable phylogenetic analysis, but not so much that there is substantial homoplasy of characters. For example, the control region of mtDNA contains "hotspots" for base substitution that may change so rapidly as to produce considerable homoplasy (Aquadro and Greenberg 1983; Greenberg, Newbold and Sugino 1983). (2) choices should be made according the phylogenetic spectrum. For a broad phylogenetic spectrum, notably nuclear and mitochondrial genes encoding ribosomal RNA can be used because they encompass both highly variable and conserved regions providing information across a broad phylogenetic spectrum. The different regions of the ribosomal RNA sequence also demonstrate a wide range of evolutionary rates and therefore demonstrate a wide potential for phylogenetic resolution across a large time scale (Hillis and Moritz 1990). For a small phylogenetic spectrum, such as among closely related interspecific diversity or intraspecific diversity, mtDNA has proven one of the most effective molecules because it demonstrates a strictly maternal pattern of inheritance, it is easily accessible for analysis and without recombination it is inherited as a single evolutionary unit (Wilson *et al.* 1985, chapter 8). (3) Evolutionary rate and mode of inheritance should be considered. Non-coding sequences typically have high evolutionary rates (Li and Wu 1985) as do most regions of animal mtDNA (Brown 1985). (4) The rate of sequence rearrangement should also be a factor in selecting a sequence for analysis. Rearrangements can lead to gross overestimates of sequence diversity. However, once identified, the rearrangements are themselves potential sources of phylogenetic information (Hillis and Moritz 1990).

Isolation of DNA: Two methods are commonly used to isolate DNA for sequencing. One is *in vivo* cloning which involves the preparation and isolation of viral and/or bacterial vectors that contain copies of the sequence of interest. Another one is *in vitro* amplification which was the method used in this thesis. With the development of the polymerase chain reaction (PCR) technique (Kleppe *et al.* 1971; Mullis and Faloona 1987; Innis *et al.* 1990) direct sequencing from complex genomic DNA has been made possible. The target sequence can be amplified to microgram quantities, which can be sequenced directly using automated sequencing machines. The procedure of PCR is not complex. The important requirement for PCR is that the sequences of the regions flanking the target sequence should be known so that primers to these regions can be constructed. A method called inverted PCR can also be used to amplify outside of a known region (Ochman, Gerber and Hartl 1988). Therefore, this methodology makes it possible to sequence DNA isolated from a wide variety of sources, including preserved museum specimens (Kocher and White 1989) and sub-fossil tissue (Paabo, Gifford and Wilson 1988).

Once sequence data has been obtained it is common practice to submit it to a publically accessible database such as GenBank or EMBL. Many journals now make this a requirement of publication. As a result, these databases have grown enormously and another important source of sequence data has become the public databases themselves.

Method of alignment Once the sequences have been obtained from cloning, PCR processes, or from a database they must be aligned to be of use in systematics. Davison (1985) reviewed the various algorithms that have been developed for nucleotide sequence alignment, and divided the procedures into four general methods: dot matrix plots, global alignments, local alignments and visual inspection.

Dot matrix comparisons are useful for quick determination of major regions of similarity and to provide a visual portrayal of these similarities. Global alignments seek to align two entire regions, whether or not they are homologous, by minimizing some score based on mismatched characters and gaps. Local alignment algorithms are similar but find all subsequence matches above a certain defined threshold, and do not (necessarily) attempt to match all characters in the sequences. A multiple local alignment is usually the best way to align orthologous sequences. There are many computer algorithms to aid alignment.

Methods for analyzing DNA sequencing data

There are three commonly used approaches to construct molecular phylogeny: distance matrix methods (such as Neighbor Joining), parsimony methods, and likelihood methods (Felsenstein 1988).

Distance matrix methods use pairwise measures of differences between species to construct phylogenies (Li and Graur 1991). Evolutionary distances are usually measured in numbers of estimated nucleotide substitutions. One of the simplest distance methods is UPGMA (unweighted pair group method using arithmetic averages). This method starts by constructing a pairwise matrix of distances and noting the two most similar taxa. These are joined together in the phylogeny and a new distance matrix is constructed with these two taxa missing but replaced by the average distance to these two taxa. The pair with the next lowest distance is found and the process iterates until only one taxa remains. This method makes the implicit assumption that the rates of change within each lineage are equal. This may not always be a valid assumption. A commonly used distance method that attempts to correct for this assumption is the Neighbor Joining method (Saitou and Nei 1987). Again this method makes use of pairwise distances but at each stage of the clustering it attempts to adjust the matrix (and the inferred branch lengths) to account for potential differences in evolutionary rate of change between taxa.

Maximum parsimony methods have been the most widely used method to infer phylogenies (Hillis and Moritz 1990). This method ideally evaluates all possible trees and gives each a score that is used to choose between different trees. The score is calculated as the number of evolutionary changes required for a particular topology. The most parsimonious tree is the one with the minimum number of evolutionary changes.

Maximum likelihood methods reconstruct a phylogeny based on a model of evolutionary change that converts any one sequence into another. The model of evolution must be clearly specified. Some of the commonly used models are Jukes-Cantor, the Kimura two-parameter model, and a generalized two-parameter model. The Jukes-Cantor model (Jukes and Cantor 1969) assumes that the rates of change from one nucleotide state to another state are independent of the states considered and depend only on a single parameter measuring the rate of substitution. The Kimura two-parameter model (Kimura 1980) assumes that transitions occur more frequently than transversions however, within each class substitutions occur at an equal rate. The generalized two-parameter model also permits different transition and transversion rates but also takes into account different nucleotide frequencies. The Kimura two-parameter model will eventually lead to equal frequencies of the four nucleotides but these are parameters that must be estimated in the generalized model. The maximum likelihood method determines the likelihood that the given evolutionary model would yield the observed sequences. The phylogenies with the highest likelihood are chosen (Felsenstein 1981c; Felsenstein 1981a; Felsenstein 1981b). However, even with simple models of evolutionary change the computational task is enormous, therefore this method is the slowest of all methods.

1.3 Rodent Phylogeny

Much effort has been made to resolve rodent relationships using both morphological and molecular approaches. In this section we will review the phylogenies of rodents from classical and molecular viewpoints. The following major questions will be considered — what are rodents?; what was the origin of rodents?; are the phylogenies constructed from morphological and molecular approaches supported by each other?.

1.3.1 What are rodents?

Rodents are small, mostly terrestrial mammals which represent about 40% of all living mammal species (Wilson and Reeder 1993). They are a very ancient and highly diversified order. The oldest recognized family of rodents is Ctenodactylidae from the lower Eocene, about 55 million years ago (Hartenberger 1985). The considerable diversity of rodent taxa is exhibited in morphology, habitat, behavior, and life history (Nedbal, Allard and Honeycutt 1994). The size of most rodents is between 80 to 350 mm in length. The distribution of rodents is worldwide and rodents colonize almost all terrestrial environments, including cities (Carleton 1984; Hartenberger 1985).

1.3.2 Classification of rodents

The order Rodentia (kingdom Animalia, phylum Chordata, subphylum Vertebrata and class Mammalia) has existed as a separate group since early classifications of Mammals. These classifications were mainly deduced from characters of the facial skull, more precisely from the masticatory apparatus. Size and shape of the infraorbital foramen, attachments and development of the masseter muscle, and sciurognathy or hystricognathy of the jaw and developmental features (Young 1950) were the main characters used in the elaboration of rodent classification. The Sciuromorpha jaw structure is thought to be the more primitive type of jaw muscles and are those found in squirrels. The Hystricomorpha jaw structure is found in guinea pigs, porcupines, and their relatives (Young 1950). In the former, the masseter is attached to the zygomatic arch and is divided into a more lateral portion with simple up and down action, and a medial part that pulls the jaw forward. In the latter, the lateral portion remains simple and the medial proceeds to a large insertion on the face below (but not through) the infraorbital canal. The lower jaw often carries a large flange which can be the attachment of the masseter muscle. These major taxonomic characters support the monophyly of the group.

However, what the order contains and how it is organized have changed many times in the last 100 years. This is mostly due to the commonness of parallel and convergent characters within the group and the lack of suitably well defined synapomorphic characters. For instance, rabbits were at one time, grouped within the rodents. This was changed with the finding that rabbits have a reduced second pair of incisors hidden behind the large pair of upper incisors. True rodents have only one pair of incisors in each jaw. Rabbits are therefore classified into a different order, Lagomorpha, rather than Rodentia (Willee, Walker and Barnes 1979).

The classification of suborders within Rodentia have also been variable. Brandt (1855) first proposed three suborders, however two suborders were proposed by Tullberg (1899) and Chaline and Mein (1979), four suborders by Wood (1985), five suborders by Simpson (1945) and even ten suborders by Tahler (1966). Some investigators find the task to be so daunting that they prefer simply use the suprafamily and family classifications (Hartenberger 1985).

The most commonly used classification of rodents divides the order into three extant suborders — the Sciuromorpha (squirrel-like rodents), the Myomorpha (rat-like rodents) and the Hystricomorpha (porcupine-like rodents) with a total of 30-33 extant families (Hartenberger 1985; Wilson and Reeder 1993). Among these families, there are about 330-400 genera and 1800-2300 species. Twelve of the families are monogeneric. The four most specious families; Muridae, Cricetidae (including arvicolines and gerbillines), Sciuridae, and Echimyidae; have more than 10 genera each and represent 75% of all rodent genera and 80% of all rodent species (Hartenberger 1985). The family Muridae contains 100 genera and 860-1000 species (Corbet and Hill 1991) including the familiar rodents; the mice, rats, voles, gerbils, and hamsters. Within this family is the subfamily Murinae, containing over 300 species of Old World mice and rats, and within this subfamily is the genus *Mus*.

Another commonly used classification divided order Rodentia into two suborders: Sciurognathi and Hystricognathi. The Sciuromorpha and Myomorpha are two of five infraorders within Sciurognathi. The Myomorpha is grouped into different superfamilies, such as Muroidea, Dipodoidea, Geomyoidea, and Gliroidea. Caviomorpha (guinea-pig-like rodents) is one of three infraorders in the Hystricognathi and the superfamilies within Caviomorpha are Cavioidea, Octodontoidea, Erethizontoidea, and Chinchilloidea (Chaline and Mein 1979). A classification of rodents into two suborders was first suggested in 1899. Based on the sciurognathy or hystrycognathy of the jaw and developmental features, Tullberg (1899) classified rodents into two groups, Sciurognathi and Hystricognathi. The sciurognathi group was split into two infraorders, Sciuromorphi and Myomorphi. Hystricognathi was split into Hystricomorphi and Bathyergomorphi. The classification into three suborders described above can be traced back to Brandt in 1855 who first classified rodents into the Sciuromorpha, Myomorpha and Hystricomorpha suborders (Hartenberger 1985).

Part of the reason for the uncertainty of the classification rests with the types of characters suitable for classification purposes within the rodents. "The presence of identical features in two lines of rodents does not necessarily mean that these features were inherited from a common ancestor" (Wood 1985). In studies of classification, a major problem is distinguishing the characters which evolved by convergence or parallelism from those characters which were inherited from common ancestors. Convergence means that the evolution of unrelated species (perhaps occupying similar adaptive zones), resulted in structures bearing a superficial resemblance. Parallel evolution means that the occurrence of the same or similar features independently evolved in two or more lineages (Lovtrup 1977). Among rodents, parallelism has been exceedingly common in dental, osteological and morphological features. The reasons for the frequent parallelism in rodents is perhaps due to the common basic adaptations of all rodents and the large number of separate but closely related lineages (Wood 1985).

1.3.3 Classical and molecular phylogeny of rodents

The order Rodentia has been accepted as monophyletic since its first classification in the early 19th century (Luckett and Hartenberger 1993). This hypothesis has been supported by traditional systematics due to the large number of shared, derived characteristics. A molecular approach has challenged this view (Graur, Hide and Li 1991). In an analysis of 15 proteins from four taxa, using a maximum parsimony method (Table 1.1), the reconstructed phylogenetic trees suggested rodent paraphyly. From genetic distance data it was suggested that the Cavioidea (guinea pig) diverged before the separation of the primates and the artiodactyls from the myomorph rodents (Muroidea: rat and mouse). Graur, Hide and Li (1991) postulated that the order Rodentia is not monophyletic and that the guinea pig (and relatives) should be placed in a separate order, the Caviomorpha (or the Hystricomorpha).

However, as Graur, Hide and Li mentioned in their paper, many genes in the guinea pig have faster evolutionary rates than those in Muroidea. For example, Shoshani *et al.* (1985) showed a 2 times faster rate in guinea pig than in mice for the α- and β- hemoglobin chains. Since unequal evolutionary rates can mislead parsimony inferences (Felsenstein 1978) this questions their conclusions. Thus a more efficient method should be used to avoid the effect of disparate evolutionary rates. Another problem with this study was that only limited taxa (primates, Myomorphs (mouse or rat), guinea-pig and an outgroup) were examined. More sequence data from more taxa needed to be examined to support this rodent paraphyly hypothesis.

The Graur, Hide and Li paper generated great controversy. Classical morphological data was reanalyzed to demonstrate that Hystricognaths and Sciurognaths share several, presumably derived, morphological characters (synapomorphies) and support the order Rodentia as a natural clade (Luckett and Hartenberger 1993). However, morphological systematics is unable to decipher the relationships among the 30 living families currently recognized within Rodentia (Wilson and Reeder 1993).

Since the initial study by Graur, Hide and Li, many workers in molecular systematics have contributed to the debate by analyzing various mitochondrial and nuclear data sets. Some of these studies have claimed to support rodent paraphyly (Graur *et al.* 1992; Li *et al.* 1992; Ma *et al.* 1993; Wolf *et al.* 1993; D'Erchia *et al.* 1996; Janke, Xu and Arnason 1997; Reyes, Pesole and Saccone 1998) while other molecular analyzes tend to support the traditional classification of rodent monophyly (Hasegawa *et al.* 1992; Martignetti and Brosius 1993; Cao *et al.* 1994; Kuma and Miyata 1994; Frye and Hedges 1995; Porter, Goodman and Stanhope 1996).

In order to avoid the sensitivity of the maximum parsimony method with unequal evolutionary rates, Hasegawa *et al.* (1992) reanalyzed the data used by Graur, Hide and Li with a maximum-likelihood method. Their results did not justify a paraphyletic hypothesis for the rodents. However, they admitted that there may be justification for a separate ordinal status for the Caviomorpha because the phylogenetic distance between Caviomorpha and Myomorpha is so large that the maximum likelihood method could not generate a significant preference for or against monophyly. Li, Hide and Graur (1992) pointed out some limitations which exist in the model-dependent maximum likelihood method and claimed that Hasegawa *et al.*'s results support rodent polyphyly. More protein and nucleotide sequences have been analyzed by Li *et al.* (1992), Graur *et al.* (1992) and Ma *et al.* (1993) (Table 1.1) and provide evidence supporting rodent polyphyly, although not statistically significant.

Wolf *et al.* (1993) analyzed the amino acid sequences of copper-zinc superoxide dismutase and the results showed some support for polyphyly of rodents but no confidence tests were done.

Nuclear data also provide evidence conflicting with rodent polyphyly but supporting the monophyly of rodents. Martignetti and Brosius (1993) detected a molecular marker encoded by a retroposon, BC1 RNA, in the guinea-pig and Sciurognath rodents, but not in the other mammalian orders taxa, and claimed that their result support the monophyly of Rodentia. However, Frye and Hedges (1995) pointed out that a transposable element may not be an ideal indicator of evolutionary history. Cao *et al.* (1994) and Kuma and Miyata (1994) carried out extensive maximum likelihood analyzes, mainly on nuclear encoded proteins, and found support for the traditional monophyletic tree. Another analysis of a nuclear gene, exon 28 from the von Willebrand factor, with a larger representation of several eutherian orders gave further support for Rodentia monophyly (Porter, Goodman and Stanhope 1996).

Mitochondrial DNA sequence data provide evidence which both supports and refutes the monophyly of rodents. The analysis of mtDNA genes such as 12S rRNA, 16S rRNA, transfer RNA-Valine, and cytochrome b from mainly muridae taxa provide evidence supporting rodent monophyly (Frye and Hedges 1995; Cao, Okada and Hasegawa 1997). However, some other mtDNA sequences from complete mitochondrial genomes and longer DNA sequences provide strong evidence to support rodent polyphyly (Ma *et al.* 1993; D'Erchia *et al.* 1996; Janke, Xu and Arnason 1997; Reyes, Pesole and Saccone 1998).

1.4 Mus genus

1.4.1 The phylogeny of the Mus genus

The *Mus* genus (subfamily Murinae, family Muridae, suborder Myomorpha) contains all true Old World mice. It has been divided into four subgenera: *Coelomys*, Ethiopian *Nannomys*, Oriental *Pyromys*, and *Mus* (Table 1.2). *Coelomys* includes the species *pahari*, *crociduroides* and *mayori*; *Nannomys* includes *minutoides* and *setulosus*; *Pyromys* includes *platythrix, shortridge* and *saxicola*; and *Mus* subgenus includes *macedonicus* (equivalent to *spretoides*), *spicilegus*, *spretus*, *caroli*, *cookii*, *cervicolor*, and *musculus*, the so called house mouse. There are three groups within the *Mus* subgenus. The Asian group includes *caroli*, *cervicolor* and *cookii*; the Musculus group includes *domesticus*, *bactrianus*, *castaneus* and *musculus*; and the Palearctic group contains *spicilegus*, *macedonicus* and *spretus* (Bonhomme 1986; Bonhomme *et al.* 1984; She *et al.* 1990; Silver 1995).

The Muridae family of rodents appears to have originated in the area around present-day India and Southeast Asia (Silver 1995). Phylogenetic and paleontological data suggest that mice and rats diverged from a common ancestor 10-15 myr BP. The *Mus* genus itself was established 6 myr BP (Jaeger, Tong and Denys 1986; Silver 1995). Since that time the *Mus* genus has diverged into a variety of species across the Indian subcontinent and neighbouring lands.

Mice have played an important role in genetics studies and modern biomedical research. Nearly all human genes have counterparts in the mouse genome (Silver 1995). To understand human genes one often studies the homologous genes of mice to extrapolate back to an understanding of the human gene. However, the phylogeny of the *Mus* genus is still not very clear. Actually, some even include species from other families as "common mice" (Silver 1995). Therefore, establishing a reliable phylogeny of *Mus* is very important.

1.4.2 Molecular phylogeny of Mus genus

She *et al.* (1990) analyzed the molecular data obtained by three major techniques: protein electrophoresis, scnDNA hybridization and mtDNA RFLP and found that the *Mus* genus can be divided into four major levels of divergence by all methods.

Node 1 groups the four subspecies of *M. musculus* (*domesticus, bactrianus, castaneus* and *musculus*). Node 2 groups the four species: *M. musculus, M. spretus, M. spicilegus* and *M. spretoides.* Node 3 groups the stem of node 2 and the three strictly Oriental species (*M. caroli, M. cervicolor* and *M. cookii*). Node 4 groups the previous lineages with Oriental *Pyromys* and *Coelomys* and the Ethiopian *Nannomys.*

However, She *et al.* (1990) are not able to resolve each level into a series of clear dichotomies. Several questions regarding the phylogeny of the *Mus* genus are currently being examined with molecular techniques. First, the relationship of the four subgenera. Second, the relationships among the *musculus* group (the commensal taxa), the relationship among the Palearctic group (thought to be the aboriginal taxa) and the relationships of the Asian species. Third, the position of *M. spretus.* Fourth, the relative genetic diversity within *M. musculus* and within *M. domesticus.*

1. **Relationship of the four subgenera:** The consensus is that *Pyromys, Coelomys* and *Nannomys* are subgenera of the *Mus* genus and distinct from the *Mus* subgenus, but the relationship of the three subgenera has yet to be resolved (Ferris *et al.* 1983; Bonhomme *et al.* 1984; She *et al.* 1990; Sourrouille *et al.* 1995). Nishioka (1989) compared four repetitive sequences from a B1 element, the MIF (mouse interspersed fragment), centromeric DNA, and Y-chromosomal fragments from *Mus caroli*, *M.cookii*, *M.hortulanus* (equivalent to *M.spicilegus*), *M.musculus*, *M.pahari*, *M.saxicola*, and *M.spretus.* Except for B1, these sequences used as probes showed species-specific hybridization patterns. From the analysis of MIF sequence, they found that *M.pahari* (in the *Coelomys* subgenus)and *M.saxicola* (in the *Pyromys* subgenus) group together due to a shared 1.3-kb EcoR1 band. But no *Nannomys* species

were examined.

2. **Relationships among the musculus, Palearctic and Asian taxa:** It is generally accepted that the musculus group is a monophyletic group clearly separated from the Palearctic group. In the *musculus* group, *M. bactrianus* has been less studied. *M. musculus* and *M. castaneus* are associated according to the studies of electrophoresis of proteins encoded by autosomal; genes and analyzes of mtDNA, the Y chromosome and a pseudogene locus (Prager, Tichy and Sage 1996). However, there is an unresolved trifurcation between *M. musculus-M. castaneus, M. domesticus* and *M. bactrianus.*

These studies suggest that *M. spicilegus* and *M. macedonicus* are grouped together as a clade (Boursot *et al.* 1993; Sage, Atchley and Capanna 1993). However, other analyzes from mitochondrial genomes, differ from analyzes of nuclear sequences and do not or weakly support this association (Fort *et al.* 1985; Bonhomme 1986; She *et al.* 1990). According to electrophoresis and scnDNA data, *M. spicilegus* and *M. spretoides* are clearly sister lineages (She *et al.* 1990). But mtDNA RFLP data weakly support *M. spicilegus* and *M. macedonicus* as sister taxa. Fort *et al.* (1985) showed that the *M. spicilegus* and *M. macedonicus* sequences for 110 bp of the mitochondrial 16s rRNA gene provided insufficient evidence to cluster these species (consistent with the preliminary restriction studies of Ferris *et al.* 1983). A recent mtDNA sequence data showed that *M. spicilegus* and *M. macedonicus* are sister taxa (Prager, Tichy and Sage 1996).

For the Asian species, *M. caroli*, *M. cervicolor* and *M. cookii*, the results from scnDNA and mtDNA (She *et al.* 1990) suggested that these species are monophyletic. And their electrophoretic data suggests that the three Asian species seem to have speciated almost simultaneously (Ferris *et al.* 1983; Bonhomme *et al.* 1984; Silver 1995).

3. **Position of M. spretus**: According to protein electrophoresis data, *M. spretus*

is associated with *M. spicilegus* and *M. macedonicus* in an aboriginal group (Sage 1981; Sage, Atchley and Capanna 1993). Other studies by protein electrophoresis (Bonhomme *et al.* 1984), satellite DNA amplification (Bonhomme 1986; Dod *et al.* 1989), restriction of nuclear rDNA spacer regions (Suzuki and Kurihara 1994), and mtDNA RFLP (She *et al.* 1990) all suggested that *M. spretus* could be an outgroup to all the other commensal taxa. As a result, a trichotomy between the *musculus* group (commensals), *spretus* and *macedonicus-spicilegus* is the current consensus (She *et al.* 1990; Agulnik *et al.* 1993; Boursot *et al.* 1993). But the results of centromeric sequencing (Nishioka 1989) clearly demonstrated that *M. spretus* is distant from *M. spicilegus* and *M. musculus.* A parsimony and neighbor-joining analysis of 10 mtDNA sequences from six species of mice (Prager, Tichy and Sage 1996) show that *M. spretus* diverged early from the *musculus* and *spicilegus-macedonicus* groups.

4. **Relative genetic diversity of M. musculus and M. domesticus:** *M. musculus* and *M. domesticus* are commensal taxa and considered by some to be subspecies within *M. musculus.* Some evidence from t-haplotypes and nuclear DNA RFLPs indicated that *M. musculus* is genetically more uniform (Figueroa *et al.* 1987; Klien, H.Tichy and Figueroa 1987; Klien *et al.* 1988; Ruvinsky *et al.* 1991). Other studies of heterozygosity (Sage, Atchley and Capanna 1993; Dallas *et al.* 1995) implied little or no difference in variability between the two species. An analysis of mtDNA RFLP showed a mtDNA tree with equal depths for the *M. musculus* and *M. domesticus* (Ferris *et al.* 1983) and similar intraspecific pairwise difference. Prager, Tichy and Sage (1996) using parsimony tree depths also showed that *M. musculus* is genetically less diverse and younger than *M. domesticus.*

1.5 Mitochondria

Mitochondria are complex organelles with their own DNA encoded genomes and are the principal sites of cellular energy production. The circular mitochondrial genome is autonomously replicated and transcribed within the organelle. Mitochondrial DNA (mtDNA) differs from nuclear DNA not only by being outside the nucleus but also by existing in thousands of copies per cell. The high copy number of mtDNA makes it much easier to extract the DNA, to carry out PCR amplification and subsequently sequence the mtDNA.

The mitochondrial genome of the mouse and rat have been completely sequenced (Bibb *et al.* 1981; Gadaleta *et al.* 1989). The complete sequence of a rodent mitochondrial genome contains about 16,300 bp and codes for two ribosomal RNA genes (12s and 16s), 22 tRNA genes, 13 proteins including cytochrome c oxidase subunits I, II and III; ATPase subunit 6; cytochrome b which form parts of the respiratory chain (additional respiratory chain polypeptides are encoded by nuclear genes; Bibb *et al.* 1981. All genes are encoded on one strand (the Heavy strand) except for the genes which encode the NADH dehydrogenase subunit 6 (ND6) and eight tRNAs (Janke *et al.* 1994).

The non-coding regions of the mitochondrial genome consist of two control regions – the 879 bp displacement-loop (d-loop) containing the origin of heavy-strand replication and the 32 bp origin of light-strand replication (Bibb *et al.* 1981). The d-loop region is responsible for replicational and transcriptional control and contains the most rapidly changing sequences in mtDNA (Walberg and Clayton 1981; Saccone, Pesole and Sbisa 1991).

1.5.1 Advantages of mtDNA in phylogenetic studies

Mitochondrial DNAs have been widely used in the studies of molecular phylogeny, because of two features that makes it an attractive source of phylogenetically informative characters. First, mtDNA has a maternal inheritance pattern. This haploid inheritance means that mtDNA is more sensitive than nuclear DNA to severe reductions in the number of individuals in a population with an effective population size a quarter of that for a nuclear gene. mtDNA does not have recombination which makes it easy to trace the history of genes divergence (Cann 1987). The mitochondrial genes are inherited as a single linkage group and thus do not provide independent estimates of the species tree. A species tree is best inferred from several independent gene trees (Pamilo and Nei 1988). However, the smaller the effective population size, the higher the probability that a given gene tree will track the species tree. Hence, Moore (1995) argued that the mitochondrial-gene tree might have a better chance of tracking the species tree than any single nuclear-gene tree.

The second advantageous feature of mtDNA is its higher mutation rate (up to ten times; Brown *et al.* 1982) than that found in nuclear DNA (T. *et al.* 1982; Hasegawa, Yano and Kishino 1984). Mutation rates also vary in different functional regions of mtDNA. The variability of different regions in the mouse mitochondria genome is as follows. From least to most variable are ribosomal RNAs, transfer RNAs, protein coding regions and the displacement loop (Ferris *et al.* 1983).

Why is the mitochondrial DNA mutation rate faster than that of nuclear DNA? There are several possible reasons: 1. A lower fidelity of the DNA replication process in mitochondria. 2. Inefficient DNA repair mechanisms or the absence of repair mechanisms in mitochondria. 3. Higher concentrations of mutagens as a by-product of mitochondrial metabolism (Li and Graur 1991, pp. 84-90). 4. Nuclear gene products involved in mtDNA replication or repair may have alleles that accelerate mutations (Paabo 1996).

Table 1.1: Genes used to question the classification of the Caviomorpha.

Paraphyly — Graur, Hide and Li (1991) — Graur *et al.* (1992) Monophyly — Hasegawa *et al.* (1992)	
β-globin	α-crystallin A chain
α-globin	α-lactalbumin
glucagon	pancreatic ribonuclease
"big"gastrin	adrenocorticotrophin
lipocortin	pancreatic polypeptide
proinsulin	lipoprotein lipase
β-nerve growth factor	vasoactive intestinal peptide
vasopressin-neurophysin precursor	
Paraphyly — Li, Hide and Graur (1992)	
α-crystallin A	α-globin
β-globin	α-lactalbumin
pancreatic ribonuclease	lipoprotein lipase
lipocortin	insulin
nerve growth factor-β	factor IX
Paraphyly — Li *et al.* (1992)	
α-crystallin A	α-globin
β-globin	α-lactalbumin
pancreatic ribonuclease	lipoprotein lipase
lipocortin	insulin
myoglobin	"big" gastrin
glucagon	adrenocorticotrophin
pancreatic polypeptide	vasoactive intestinal peptide

β-nerve growth factor	myelin basic protein
vasopressin-neurophysin precursor	N-ras protooncogene
Paraphyly — Wolf *et al.* (1993)	
Cu/Zn superoxide dismutase	
Monophyly — Martignetti and Brosius (1993)	
BC1 RNA, a retroposon element	
Paraphyly — Ma *et al.* (1993)	
cytochrome b gene	
Monophyly — Cao *et al.* (1994)	
α-crystallin A	α-globin
β -globin	lipoprotein lipase
lipocortin I	α-lactalbumin
β-nerve growth factor	factor IX
pancreatic ribonuclease	proinsulin
myoglobin	cytochrome b
COII	gi3 protein a
NADPH-cytochrome p450	integrin β1
myelin	glucagon
Monophyly — Kuma and Miyata (1994)	
α-1 antitrypsin	α-crystallin A chain
α-globin	β-globin
α-lactalbumin	Pancreatic RNase
Complement C3	Transglutaminase
Lipoprotein lipase	Lipocortin
NGF β chain	VP-NP precursor
Phospholipase A_2	Glucagon
IGF1	Insulin

H^+-, K^+-ATPase Corticotrophin	C-reactive protein
Monophyly — Frye and Hedges (1995) 12SrRNA, 16SrRNA, RNA-Valine	
Monophyly — Porter, Goodman and Stanhope (1996) Exon 28 von Willebrand factor	
Paraphyly — D'Erchia *et al.* (1996) Monophyly — Cao, Okada and Hasegawa (1997) 16 complete mtDNA	
Paraphyly — Janke, Xu and Arnason (1997) 12 mt proteins (excluding NADH6)	
Paraphyly — Reyes, Pesole and Saccone (1998) 23 complete mammalian mtDNA sequences	

Table 1.2: Some species of the *Mus* genus.

Genus	Subgenus	Species	Subspecies
Mus	Oriental *Pyromys*	*platythrix*	
		shortridgei	
		saxicola	
	Coelomys	*pahari*	
		crociduroides	
		mayori	
	Ethiopian *Nannomys*	*minutoides*	
		setulosus	
		mattheyi	
	Mus		
	- Musculus group:	*musculus*	*musculus*
			bactrianus
			castaneus
			domesticus
	- Palearctic group:	*macedonicus*[†]	
		spicilegus[*]	
		spretus	
	- Asian species:	*caroli*	
		cervicolor	
		cookii	

Nomenclature and taxonomy follow She *et al.* (1990) and Silver (1995).
[†] *macedonicus* is also known as *spretoides* or *abbotti* and [*] *spicilegus* is also known as *hortulanus*.

Chapter 2

The phylogenetic history of the genus *Mus* inferred from mitochondrial genes

2.1 Abstract

The analysis of COII sequences by neighbor-joining and parsimony approaches support the view that *M. spretus* is a sister species of *macedonicus*, *spicilegus* and *musculus*. Both the neighbor-joining tree and the parsimony tree show a bifurcation between *M. spretus*, and *M. musculus*, *M. spicilegus-M. macedonicus* suggesting that *M. spretus* diverged earlier than the others followed by *M. musculus*. The COII sequence data provided evidence to support that *M. spicilegus* and *M. macedonicus* are sister species within the Palearctic group while *M. caroli* and *M. cervicolor* are grouped together (66 BP). The 12S rRNA gene sequences analyses suggested that the subgenus *Pyromys* species (*M. platythrix* and *M. saxicola*) grouped together and the subgenus *Coelomys* species (*M. crociduroides* and *M. pahari*) grouped together. The

branching pattern among the four subgenera could not be resolved.

2.2 Introduction

The order Rodentia (kingdom Animalia, phylum Chordata, subphylum Vertebrata and class Mammalia) is very old and highly differentiated with 30-33 separate families, numerous genera, and over 1,500 well-defined species (Corbet and Hill 1991). The family Muridae contains over 1,000 species including mice, rats, voles, gerbils, and hamsters. Within this family is the subfamily Murinae, which contains over 300 species of Old World mice and rats, and within this subfamily is the genus *Mus*, the main focus of this project. The *Mus* genus has been divided into four subgenera, *Coelomys* (*pahari, crociduroides* and *mayori*), *Nannomys* (*minutoides, setulosus*), *Pyromys* (*platythrix, shortridgei, saxicola*) and *Mus* (*musculus, macedonicus, spicilegus, spretus, caroli, cookii*, and *cervicolor*).

There is a great deal of uncertainty in the field of rodent systematics about their phylogenetic relationships. Systematics attempts to detect, describe and explain diversity in the biological world. Linnaeus (1758) established a framework for describing and categorizing biological diversity. The notion that classification should be based on phylogenetic relationships was developed by later workers (e.g. Darwin 1859, reviewed by Mayr 1983). Classical taxonomy is dependent on the demonstration of distinct morphological differences. But unfortunately, many small rodent species have gross morphological characteristics that are convergent with those present in other relatively distant species. Thus, traditional taxonomy can fail to provide an accurate systematic description of rodents. Since the 1960s, molecular data have been used to help determine rodent systematics and phylogenetic relationships. DNA sequence comparisons have proven to be a highly effective means of inferring the evolutionary relationships that exist among taxa.

Recently, the phylogenetic locations of several taxa have been challenged by molecular analysis. For instance, the genus *Acomys*, African spiny mouse, was traditionally placed in the Murinae because it shares a tooth pattern with the Murinae which is the presence of two additional lingual cusps on the first upper molar (M1/). However, spiny mice also have an unusual structure of their third upper molar (M3/), and this pattern is not present in most extinct or extant murines. Immunological studies (Wilson, Ochman and Prager 1987) and DNA-DNA hybridization data (Chevret *et al.* 1993) indicated that the spiny mice (*Acomys*) are more closely related to gerbils (Gerbillinae) than to mice and rats (Murinae). Another issue is the phylogenetic location of guinea pig. A molecular analysis suggested that the guinea pig is not a rodent at all, contrary to long-held beliefs (Graur, Hide and Li 1991). A third uncertain taxonomic issue is whether the gundis (family Ctenodactylidae) belong to the order Rodentia (Hartenberger 1985; Graur *et al.* 1992). Molecular data has suggested that gundis branched off before the divergence between myomorphs, primates, and artiodactyls (Graur *et al.* 1992). Lastly, the family Gliridae might be one of the most primitive rodents along with the Ctenodactylidae according to paleontological data (such as a glirid fossil having an enlarged infraorbital foramen as in the Hystricomorpha). However, some genera have a pseudomyomorph pattern which led one study to include Gliridae within the group Myomorpha (Reyes, Pesole and Saccone 1998).

Mice are an economically important group of organisms. They cause major agriculture damage and spread disease. As model organisms, they have played a major role in genetic study and modern biomedical research. Yet we know little of their phylogenetic relationships. Within the genus *Mus*, the phylogenetic relationships of the four subgenera of the *Mus* genus have not been solved by DNA-DNA hybridization, enzyme electrophoresis or 12S rRNA analysis (She *et al.* 1990; Sourrouille *et al.* 1995). More informative sites from additional genes are needed to sort out the phylogenetic relationships of genus *Mus*.

The goal of this project is to reconstruct a reliable phylogeny of the *Mus* genus based on DNA sequences data. The identification of a useful sequence for systematic and phylogenetic analysis is the first step. The sequence should have sufficient variation to enable the analysis, but not so much that homoplasy could be produced. Homoplasy (parallel, convergent substitutions) can lead to gross under estimates of sequence diversity and generate a false phylogenetic tree. DNA recombination and rearrangement can also affect the resolution of the phylogenetic tree.

We chose rodent mitochondrial DNA because 1) it has maternal inheritance and avoids recombination. 2) it has a relatively rapid evolutionary rate which can provide sufficient variation. 3) its genes are not members of a multigene family. The 12S rRNA gene is well conserved in evolutionary history and it has been sequenced from many species. It is used to indicate the phylogenetic history of distantly related species, but is not as efficient to reveal the history of closely related species because of its relatively low evolutionary rate. The cytochrome c oxidase subunit II gene with its relatively high evolutionary rate (Ferris *et al.* 1983) is suitable for estimating a phylogeny for closely related species.

The second goal of this project involves molecular evolutionary rate. The molecular clock hypothesis was proposed by Zuckerkandl and Pauling in 1965. This hypothesis postulates that the rate of evolution in any given protein or DNA sequence is approximately constant over time in all evolutionary lineages. However, since the rate of molecular evolution has become slower in hominoids after their separation from the Old World monkeys (Goodman 1961), which has been known as the hominoid rate-slowdown, a generation-time effect hypothesis was proposed (Li *et al.* 1996). This hypothesis assumes that an important source of mutation is DNA replication errors during germ cell division. Although the molecular clock is known to be false at several levels, within the muroid rodents it was found to be true based on the analysis of mouse, rat, hamster and human sequences (O'hUigin and Li 1992). However the rat APRT sequence analysis revealed that the evolutionary rate varied among exons in rodents and that there was a recent and rapid burst of substitutions within the mouse lineage (Fieldhouse and Golding 1993). A five fold difference of evolutionary rate between two closely related species *Mus spicilegus* and *Mus musculus* was found by Fieldhouse and Golding (1996). Generation time is not a factor leading to the change of rate. The causes of the change of rate among *Mus* species or between these two closely related species should be investigated. But first we should determine if the different rates in rodents (Wu and Li 1985; Fieldhouse and Golding 1993) are consistent across all rodent species, across all genes and even across all parts of a single gene. We hope to investigate whether the molecular clock holds true for all rodents.

2.3 Material and methods

2.3.1 12S rRNA sequence collection and analysis

The 12S rRNA gene is located on the mitochondrial genome. Partial or complete 12S rRNA gene sequences were collected from DNA databases from 107 mammals species including 96 Rodentia species, nine Carnivora, one Edentata and one Lagomorpha. The alignment of sequences was performed using the CLUSTALW program (Thompson, Higgins and Gibson 1994). The initial alignment was manually inspected and adjusted by hand. The aligned sequences were bootstrapped through the SEQBOOT program (Felsenstein 1989) with 100 replicates. Distance matrices were built by DNADIST. Using NEIGHBOR, DNAPARS and CONSENSE programs (Felsenstein 1989), a consensus neighbor-joining tree and a consensus parsimony tree were reconstructed. An Edentata species, *Dasypus novemcinctus*, was used to root the trees.

2.3.2 COII sequence preparation and analysis

Cytochrome oxidase is one of the three enzyme complexes of the respiratory chain. It catalyzes the transfer of electrons from reduced cytochrome c to molecular oxygen. This reaction is carried out by a complex of at least eight subunits, of which three (called subunits I, II, and III) are encoded by the mitochondrion's own genome. The oxidation-reduction units of cytochrome oxidase known as cytochromes a and a3 are located in subunits I and II. The cytochrome c oxidase subunit II gene (COII) is our target gene.

Fifteen rodents species were used in COII PCR amplification and sequencing. The DNA samples of six species are from our lab. These species are *Mus spicilegus*, *Dipldomys merriami*, *Mesocricetus aruatus*, *Peromyscus mexicanus, Stochomys longicudatus, Rattus everti*, and *Rattus exulans.* DNA from another nine species was extracted from soft tissues preserved in 95% ethanol as kindly supplied by Dr. François Catzeflis. They are all *Mus* species: *caroli, cervicolor, cookii, crociduroides, macedonicus, minutoides, platythrix, setulosus, spretoides*, and *spretus.*

The DNA extraction followed the procedure suggested by Dr. Catzeflis; - Rinse ethanol-fixed tissues with cold ddH_2O, mince into particles and dry in a decanter for 18 hours; Add 500 μl of 0.08 M NaCl, 0.01 M Na2-EDTA, pH 8.0; - Shake gently at room temperature for 10 min, centrifuge 3000 rpm for 10 min and discard supernatant; - Resuspend with 400 μl of 0.15 M NaCl, 0.01 M Na2-EDTA, pH 8.0 and shake for 5 min at room temperature; - Add 20 μl 20% SDS (final vol. 1%), shake gently for 10 min at 37^0C and add 6 μl Proteinase K (25mg/mL) up to 0.5 mg/mL for 18 hours at 37^0C; - Centrifuge 10,000 rpm for 10 min, collect supernatant and add equal volume of isopropanol; - Shake gently for 2-3 min and pick up the DNA thread to a new tube; - Dry DNA at room temperature for 10 min, redissolve in 200-300 μl TE buffer (pH 8.0) and store at -20^0C.

The primers were chosen with the aid of Primer3 program (Rozen and Skaletsky

1997). Degenerate primers were used to make this set of primers more suitable for other species. The primers were

pb1 (5′ CACC(A/T)TA(T/C)CACACATTCGAG 3′)

and

pb2 (5′ GGTTAACGCTCT(T/A)AGCTTCA 3′).

The PCR reaction was performed using a DNA thermal cycler (Perkin Elmer Cetus), in 100 μl containing 50 mM KCl, 10 mM Tris, 2 mM MgC12, 0.2 mM each of dNTP, 10 pM each of the primers, 1 μl of DNA template (0.03-0.4 μg), and 2.5 units of AmpliTaq Gold DNA polymerase (Perkin Elmer). Reaction conditions were 12 min at 95^0C for preheating, 1 min at 95^0C, 1 min 30 seconds at $52\text{-}55^0$C (different species at different temperature), and 2 min at 72^0C for 30-50 cycles (different species for different cycles). The PCR amplified products were purified using the QIAquick PCR Purification Kit (QIAGEN Inc., Chatsworth CA). The purified PCR products were sequenced in both directions using primers pb1 and pb2.

The COII sequences of 10 additional rodents species were retrieved from GenBank. The alignment of sequences was performed manually. The aligned sequences were analyzed by the distance and parsimony algorithms described in the section on 12S rRNA analysis. Sequences from a seal (*Phoca vitulina*) and whale (*Balaenoptera musculus*) were used to root the trees. An 813 bp consensus sequence was constructed for COII (appendix A).

2.3.3 12S rRNA and COII sequence analysis

To reconstruct a phylogeny maximizing the information about Muridae species, the aligned sequences of 12S rRNA and COII were concatenated. A parsimony tree was reconstructed. The distance method was not used since it was not capable of estimating accurate distances due to the large number of unknown characters in the sequences.

2.4 Results

2.4.1 12S rRNA phylogeny

In order to determine the phylogenetic tree of rodents, 12S rRNA gene sequences have been obtained in 107 species from DNA databases. The consensus neighbor-joining tree (Fig. 2.1) and parsimony tree (Fig. 2.2) have been reconstructed.

From the neighbor-joining (NJ) tree we found the following results:

1. All Muridae species were clearly grouped together as a monophyletic clad with 97 BP (Bootstrap Percentages). The NJ tree also showed that mice and rats with three other genera *Hylomyscus, Leopoldamys* and *Mastomys* constitute a monophyletic clade (99 BP).
2. The spiny mouse (*Acomys*) clustered within Muridae (97 BP) but its position within or outside of the Murinae remains uncertain.

   There is insufficient evidence (27 BP) that *Acomys* and two African genera of Murinae, *Uranomys* and *Lophuromys* constitute a monophyletic clade by themselves. The monophyly of *Acomys, Uranomys* and *Lophuromys* was suggested on the basis of dental characters and supported by single-copy nuclear DNA-DNA hybridization data (Chevret *et al.* 1993).
3. The *Mus* species were grouped together, but not with statistical significance. The subgenera *Pyromys* species (*M. platythrix* and *M. saxicola*) were grouped together by 100 BP in both NJ and PM trees (Fig. 2.1, Fig. 2.2). The subgenus

Coelomys species (*M. crociduroides* and *M. pahari*) were clustered by 96 BP (Fig. 2.1) and 94 BP (Fig. 2.2) which is identical with Sourrouille et al's analysis (1995). The branching pattern among the four subgenera could not be resolved because of the low BP (less than 80). This is due to the limited information provided by the well conserved 12S rRNA gene.

The neighbor-joining tree and parsimony tree displayed different branching patterns for the genus *Mus*. In the parsimony tree *M. cookii* grouped with *M. musculus*. But in the neighbor-joining tree *M. cookii* clustered with *M. setulosus* and *M. mattheyl* which is contrary to previous analyzes (Silver 1995; Sourrouille *et al.* 1995).

4. Some clades other than Muridae also show high BP (Fig. 2.1) such as the Caviidae clade (guinea-pig) and the Gliridae clade (dormice). *Cavia guianae* and *Cavia porcellus* as one clade (100 BP) grouped with two other Caviidae species *Dolichotis patagonum* and *Hydrochaeris hydochaeris* (81 BP), which indicates that Caviidae is monophyletic. Unfortunately, in the parsimony tree these four Caviidae species were not clustered together but rather scattered throughout the tree. The Gliridae clade (100 BP) includes the well clustered Muscardinus clade (100 BP) and Glirulus clade (100 BP). Graphiurus and Glis are an outgroup of Muridae, Caviidae and Ctenodactylidae that was weakly supported with 59 BP .

5. The neighbor-joining tree showed that the gundis (*Ctenodactylus gundi*) diverged earlier than the guinea-pig but with relatively low bootstrap support (17 BP). This is contrary to the view that the gundis branched off after the divergence of the guinea-pig (Graur *et al.* 1992).

From the 12S rRNA parsimony tree, some clusters are not in accord with the NJ tree. The more reliable clusters in the parsimony tree are among the non-rodents species, Carnivora species. The Caviidae species (guinea-pigs and relatives) were not clustered together but scattered throughout the parsimony tree. Another difference is that the Muridae were not clustered together as a monophyletic clade but separated among almost all rodents species and non-rodents species. Since this family is well delimited by classical methods, this result indicates that 12S rRNA parsimony analysis is not a proper approach to address rodent phylogeny and suggests that the conserved 12S rRNA does not have the potential to solve the relationships among these species.

2.4.2 COII phylogeny

The 12S rRNA phylogenetic tree can only provide a broad view of rodents phylogeny, and it failed to resolve the phylogenetic relationships among the *Mus* species. To determine the *Mus* genus phylogenetic history, we have reconstructed the consensus neighbor-joining trees (Fig. 2.4) and parsimony trees (Fig. 2.3) based on COII sequences. In both trees the *Mus* species grouped together (63 BP in the NJ tree and 24 BP in the PM tree).

The subgenera *Mus* species, including the Palearctic and Asian groups, were well clustered in the neighbor-joining tree by 96 BP (Fig. 2.4). But they were not grouped well in parsimony tree (BP less than 80). The branching pattern within the *Mus* subgenera agreed entirely with the view that *M. spicilegus* and *M. macedonicus*, as well as *M. musculus*, are sister taxa by 100 BP (Fig. 2.4) and 91 BP (Fig. 2.3) (a feature which is still controversial in the literature (Prager, Tichy and Sage 1996)). The pattern also strongly supports that *Mus spretus* diverged earlier than *M. spicilegus* and *M. musculus* (99 BP, 98 BP in Fig. 2.4, Fig. 2.3).

The branching pattern of the four subgenera showed that they all diverged at the same time period. Three *Rattus* species are clustered together by 100 BP in both trees. The *Mus platythrix* and *Mus crociduroides* species were clustered together (62 BP) suggesting that the subgenera *Pyromys* and *Coelomys* may be more closely related than the subgenera Mus and *Nannomys*.

2.4.3 12S rRNA and COII phylogeny

The concatenated 12S rRNA and COII parsimony tree (Fig. 2.5) displayed a different branching pattern for the genus *Mus*. In this tree the subgenus Mus is monophyletically split off from the other three subgenera (69.7 BP). Among the three subgenera, *Pyromys* and *Nannomys* diverged recently and *Coelomys* diverged earlier. This pattern is different from the patterns revealed by either the 12S rRNA and COII neighbor-joining or parsimony trees. The analysis from concatenated 12S rRNA and COII sequences confirmed that *M. spicilegus* and *M. macedonicus* are sister taxa and that *M. spretus* diverged earlier.

2.5 Discussion

2.5.1 Phylogeny inferred from 12S rRNA

The monophyly of *Acomys* and all other Murinae has been challenged by immunological studies (Sarich 1985) that have suggested that *Acomys* is as distantly related to mice as are some other subfamilies of the muroid rodents (e.g., hamsters: Cricetinae, Chevret *et al.* 1993). Chevret *et al.*'s analysis from DNA-DNA hybridization data indicate that the spiny mice (*Acomys*) are more closely related to gerbils (Gerbillinae) than to the true mice and rats (Murinae) with which they have been classified. But their results also indicated that the spiny mice *Acomys*, *Uranomys* and *Lophuromys* constitute a monophyletic clade. They suggested that the murine morphology of spiny mice must have evolved by convergence (Chevret *et al.* 1993). Our data from 12S rRNA indicate a polyphyletic origin of *Acomys* with other Murinae genera suggesting that they are distantly related to Murinae. But the 12S rRNA neighbor-joining tree did not show that spiny mice are more closely related to the Gerbils.

The family Ctenodactylidae is considered to be one of the oldest rodent families and older than the Caviomorphs (Hartenberger 1985). Graur *et al.* suggested that like the guinea-pig, the gundi might also have branched off earlier than the myomorphs, the primates and the artiodactyls, albeit after the divergence of the guinea-pig. But the 12S rRNA neighbor-joining analysis, showed that *Ctenodactylus gundi* diverged earlier than guinea-pig though with relatively low support (58 BP).

The mtDNA analysis of Reyes, Pesole and Saccone (1998), the isochore patterns in the nuclear genome (Sabeur *et al.* 1993) and the B2 and B2-like retroposons (Serdobova and Kramerov 1993) suggest that there is a close relationship between dormice *Glis glis* (Gliridae) and guinea pigs (Caviidae) rather than between dormice and rats and mice (Muridae). However, our neighbor-joining analysis of 12S rRNA seems to indicate that Gliridae might have diverged earlier than Caviidae and Muridae (59 BP).

2.5.2 Phylogeny inferred from COII

The studies from protein electrophoresis (Sage 1981; Sage, Atchley and Capanna 1993; Bonhomme *et al.* 1984), satellite DNA amplification (Bonhomme 1986; Dod *et al.* 1989), restriction of nuclear rDNA spacer regions (Suzuki and Kurihara 1994), and mtDNA RFLP (She *et al.* 1990) all suggest that *M. spretus* are an outgroup of the commensal taxa, but they could not exclude a trichotomy between the *musculus group*, *spretus* and *macedonicus-spicilegus.* A study of Y-chromosomal probes confirmed that *Mus spicilegus*, *Mus musculus* and *Mus spretus* are closely related species and also provide a evidence that *M. spretus* is more distant from *Mus spicilegus* and *Mus musculus* (Nishioka 1989). Our analysis of COII sequences by neighbor-joining and parsimony approaches support this view (99 BP, 98 BP).

Previous mtDNA data did not support that *M. spicilegus* and *M. macedonicus* are sister species. An analysis of 110 bp of mitochondrial 16S rRNA found that the sequences from *M. spicilegus* and *M. macedonicus* were so different that they could not be clustered in a phylogenetic tree (Fort *et al.* 1985). This is consistent with preliminary restriction studies of Ferris *et al.* (1983). A recent study of mtDNA sequences from the Palearctic group of species showed that *M. spicilegus* and *M. macedonicus* are sister taxa (Prager, Tichy and Sage 1996). Our COII sequence data involving all of the *Mus* genus species and covering each subgenus provides more reliable evidence to support that *M. spicilegus* and *M. macedonicus* are sister species (66 BP in the NJ tree and 46 BP in the PM tree).

We found that the Asian species, *M. caroli*, *M. cervicolor* and *M. cookii* grouped with the Mus subgenera species forming a monophyletic clade (96 BP in the NJ tree but less than 80 BP in the PM tree). The relative associations within this group (e.g. of *M. caroli* and *M. cervicolor*) are less certain (66 BP in the NJ tree).

One view is that the subgenera *Pyromys*, *Coelomys* and *Nannomys* form a trichotomy separated from the Mus subgenus (Ferris *et al.* 1983; Bonhomme *et al.* 1984; She *et al.* 1990; Sourrouille *et al.* 1995). Another view based on repetitive sequences B1, MIF (mouse interspersed fragment), centromeric, and Y-chromosomal sequences (Nishioka 1989) among *M. caroli*, *M. cookii*, *M. hortulanus* (synonymous with *Mus spicilegus*), *M. musculus*, *M. pahari*, *M. saxicola*, and *M. spretus* suggests that *M. pahari* (*Coelomys* subgenus) and *M. saxicola* (*Pyromys* subgenus) are grouped together. The results of our COII sequence neighbor-joining analysis support the second view that subgenera *Coelomys* (*M. crociduroides*) and *Pyromys* (*M. platythrix*) are more closely related (62 BP).

Figure 2.1: A neighbor-joining phylogeny of rodents, with nine carnivores and an Edentata representative as outgroups based on 12S rRNA gene. The branch lengths in this phylogeny were calculated using the FITCH algorithm according to the consensus of 100 bootstrapped neighbor-joining trees. Branch points with a bootstrap value above 90% are labelled.

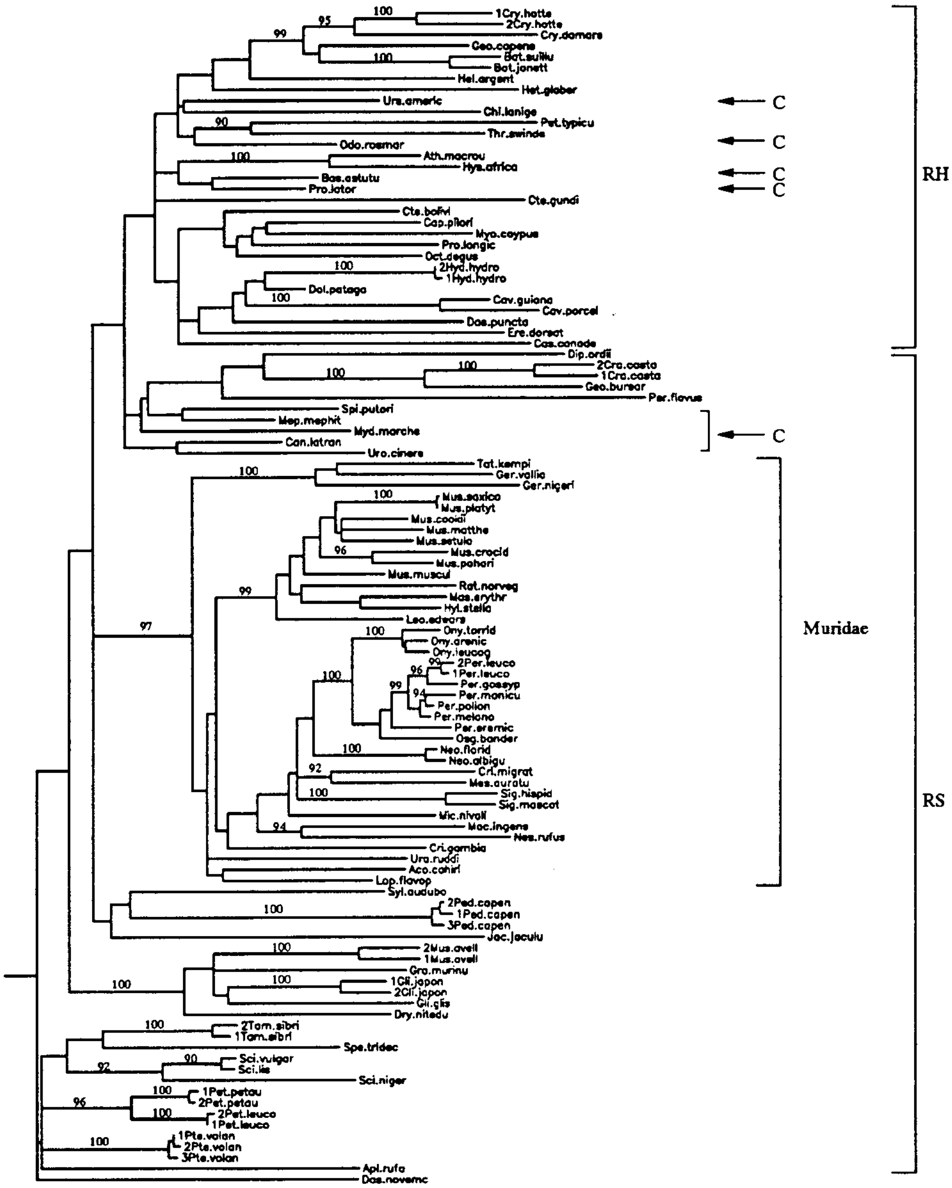

100
95
99
1Cry.hotte
2Cry.hotte
Cry.damara
Geo.capens
100
Bat.suillu
Bat.janetti
Hel.argent
Het.glaber
Urs.americ
Chi.lanige
90
Pet.typicu
Thr.swinde
Odo.rosmar
100
Ath.macrou
Hys.africa
Bas.astutu
Pro.lotor
Cte.gundi
Cte.bolivi
Cap.pilori
Myo.coypus
Pro.longic
Oct.degus
100
2Hyd.hydro
1Hyd.hydro
Dol.patago
100
Cav.guiana
Cav.porcel
Das.puncta
Ere.dorsat
Cas.canade
Dip.ordii
100
100
2Cra.casta
1Cra.casta
Geo.bursar
Per.flavus
Spi.putori
Mep.mephit
Myd.marche
Can.latran
Uro.cinere
C
C
C
C
C
RH
100
Tat.kempi
Ger.vallic
Ger.nigeri
100
Mus.saxico
Mus.platyt
Mus.cooki
Mus.matthe
Mus.setulo
96
Mus.crocid
Mus.pahari
Mus.muscul
99
Rat.norveg
Mas.erythr
Hyl.stella
Leo.edwars
97
100
Ony.torrid
Ony.arenic
Ony.leucog
99
2Per.leuco
96
1Per.leuco
100
99
Per.gossyp
94
Per.manicu
Per.polion
Per.melano
Per.eremic
Osg.bander
100
Neo.florid
Neo.albigu
92
Cri.migrat
Mes.auratu
100
Sig.hispid
Sig.mascot
Mic.nivali
94
Mac.ingens
Nes.rufus
Cri.gambia
Ura.ruddi
Aco.cahiri
Lop.flavop
Syl.audubo
Muridae
100
2Ped.capen
1Ped.capen
3Ped.capen
Jac.jaculu
100
2Mus.avell
1Mus.avell
Gra.murinu
100
100
1Gli.japon
2Gli.japon
Gli.glis
Dry.nitedu
100
2Tam.sibri
1Tam.sibri
Spe.tridec
90
Sci.vulgar
92
Sci.lis
Sci.niger
100
1Pet.petau
96
2Pet.petau
100
2Pet.leuco
1Pet.leuco
100
1Pte.volan
2Pte.volan
3Pte.volan
Apl.rufa
Das.novemc
RS

Figure 2.2: A 12S rRNA parsimony phylogeny of rodents. The branch lengths in this phylogeny were calculated using the FITCH algorithm according to the consensus of 100 bootstrapped parsimony trees. Branch points with a bootstrap value above 80% are labelled.

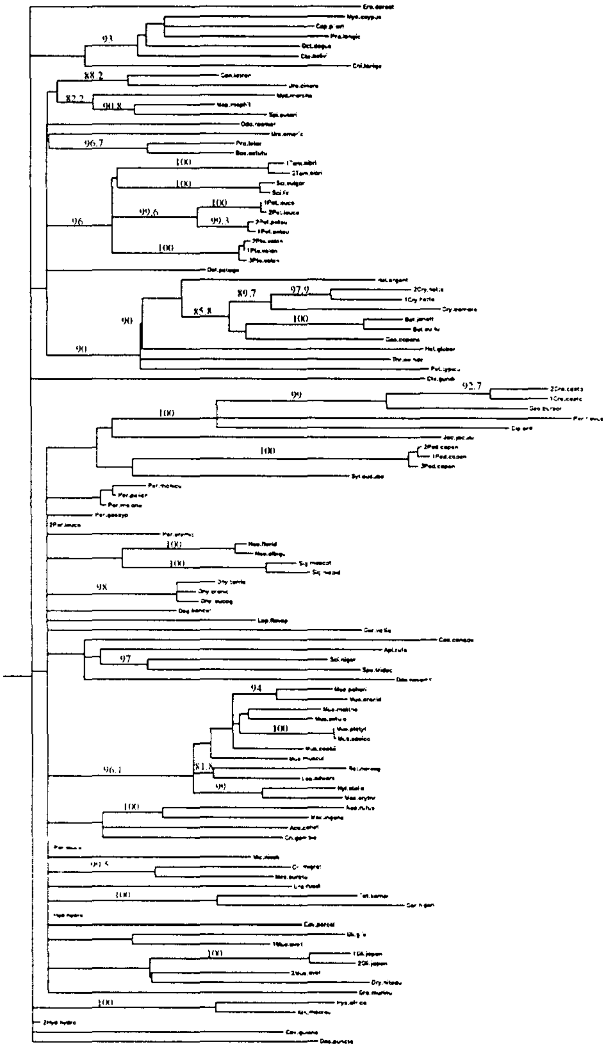

Ere.dorsat
Myo.coypus
Pro.longic
Oct.degus
Chi.lanige
Can.latran
Uro.cinere
Mep.mephit
Spi.putori
Odo.rosmar
Urs.americ
Pro.lotor
Bas.astutu
1Tam.mini
2Tam.mini
Sci.vulgar
1Pet.leuco
2Pet.leuco
2Pte.volan
1Pte.volan
3Pte.volan
Dol.patago
Hel.argent
2Cry.hotte
1Cry.hotte
Cry.damare
Bat.janett
Geo.capens
Het.glaber
Cte.gundi
2Cra.casta
1Cra.casta
Geo.bursar
Jac.jaculu
2Ped.capen
1Ped.capen
3Ped.capen
Per.manicu
Per.gossyp
2Per.leuco
Per.eremic
Neo.florid
Neo.albigu
Sig.hispid
Ony.torrid
Ony.arenic
Ony.leucog
Cas.canade
Apl.rufa
Sci.niger
Spe.tridec
Mus.pahari
Mus.caroli
Mus.platyt
Mus.cookii
Mus.musculu
Rat.norveg
Mes.erythr
Aco.cahiri
Mes.auratu
Cav.porcel
1GK.japon
2GK.japon
Cav.guiana
Das.puncta
93
88.2
82.2
90.8
96.7
100
100
100
99.6
99.3
96
100
89.7
97.9
85.8
100
90
90
92.7
99
100
100
100
100
98
97
94
100
96.1
81.8
99
100
91.5
100
100
100

Figure 2.3: A COII parsimony phylogeny of rodents. The taxa *Balaenoptera musculus* was used as an outgroup.

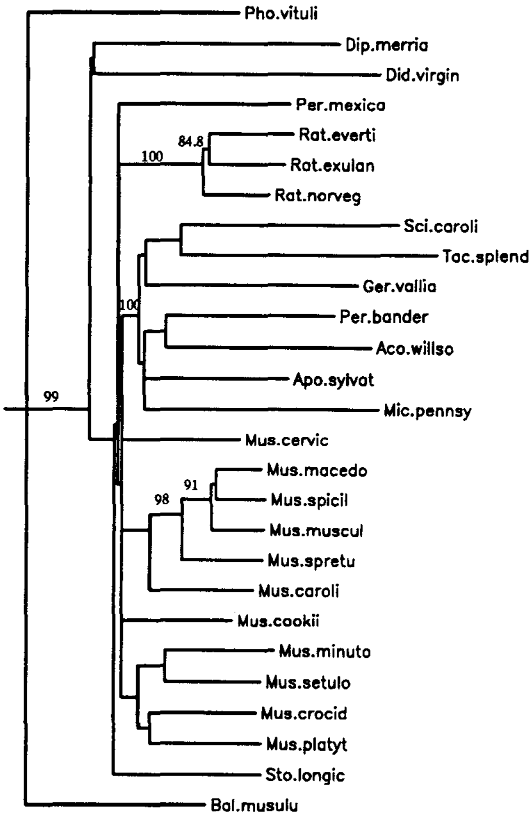

Pho.vituli
Dip.merria
Did.virgin
Per.mexica
84.8
Rat.everti
100
Rat.exulan
Rat.norveg
Sci.caroli
Tac.splend
Ger.vallia
100
Per.bander
Aco.willso
Apo.sylvat
99
Mic.pennsy
Mus.cervic
Mus.macedo
91
98
Mus.spicil
Mus.muscul
Mus.spretu
Mus.caroli
Mus.cookii
Mus.minuto
Mus.setulo
Mus.crocid
Mus.platyt
Sto.longic
Bal.musulu

Figure 2.4: A COII neighbor-joining phylogeny of rodents, with *Balaenoptera musculus* as outgroup. The branch lengths in this phylogeny were calculated using the FITCH algorithm according to the consensus of 100 bootstrapped neighbor-joining trees. Branch points with a bootstrap value above 80% are labelled.

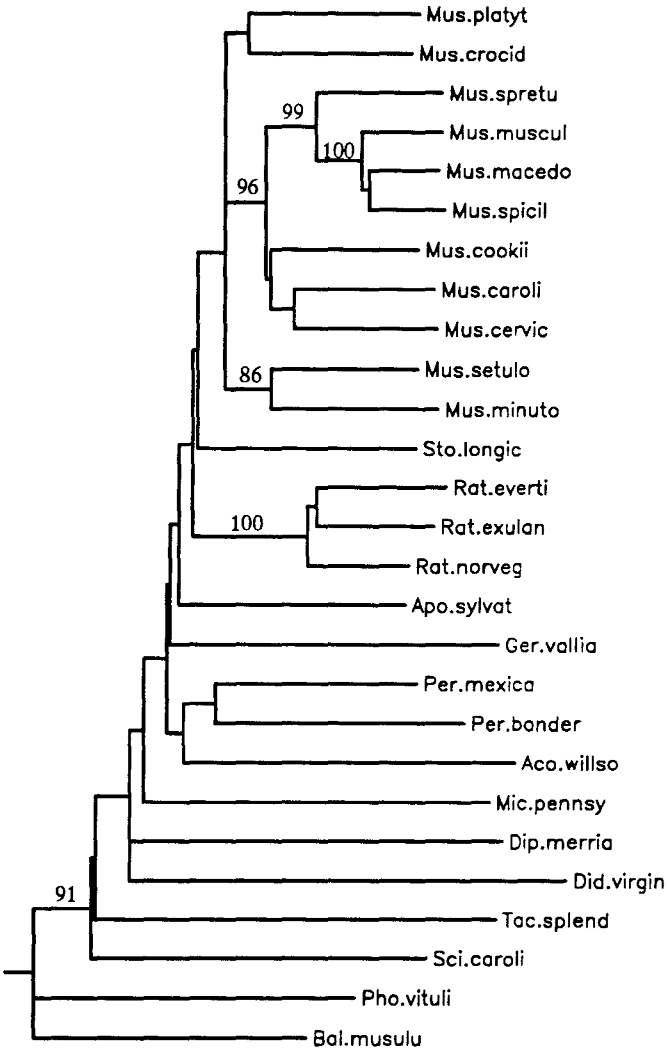
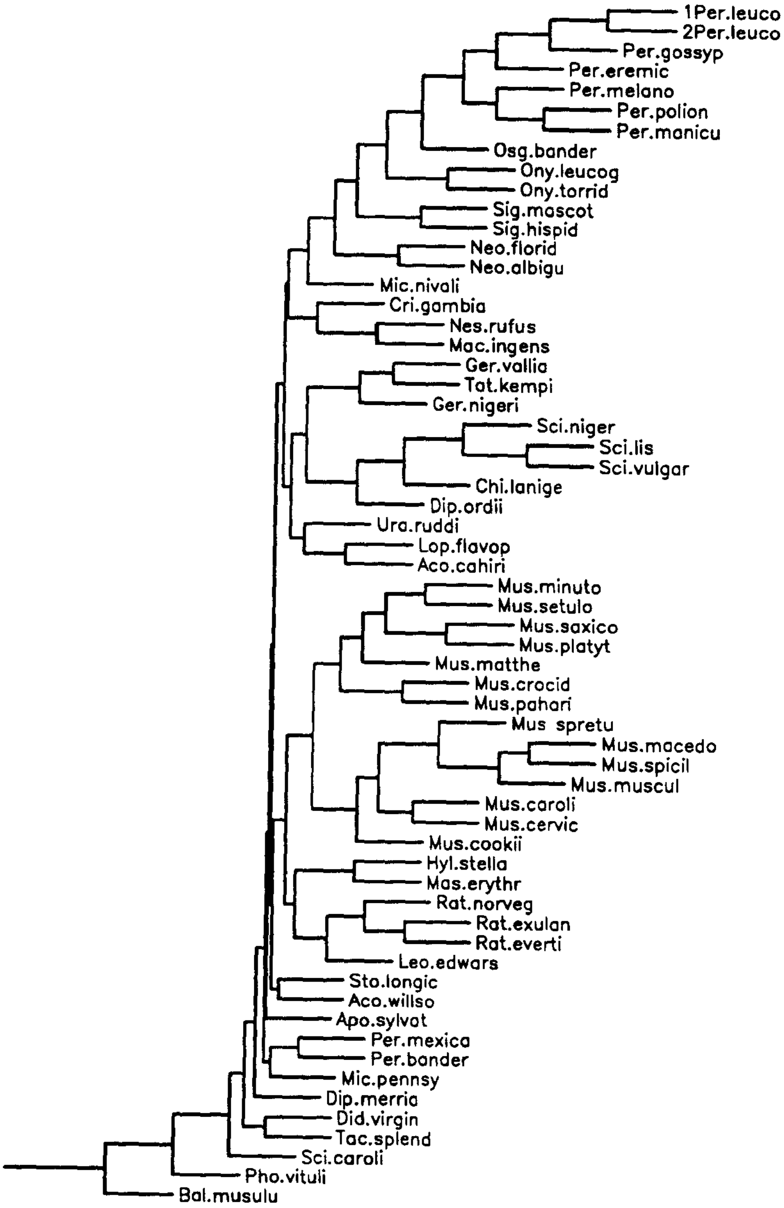

Mus.platyt
Mus.crocid
Mus.spretu
Mus.muscul
Mus.macedo
Mus.spicil
Mus.cookii
Mus.caroli
Mus.cervic
Mus.setulo
Mus.minuto
Sto.longic
Rat.everti
Rat.exulan
Rat.norveg
Apo.sylvat
Ger.vallia
Per.mexica
Per.bonder
Aco.willso
Mic.pennsy
Dip.merria
Did.virgin
Tac.splend
Sci.caroli
Pho.vituli
Bal.musulu
99
100
96
86
100
91

Figure 2.5: A parsimony phylogeny of rodents based on concatenated 12S rRNA and COII genes. The branch lengths in this phylogeny show only the relative degree to which a particular branch might be significant. The length of the branch leading to *Balaenoptera musculus* is a bootstrap of 100%. This taxa was also used an outgroup.


1Per.leuco
2Per.leuco
Per.gossyp
Per.eremic
Per.melano
Per.polion
Per.manicu
Osg.bander
Ony.leucog
Ony.torrid
Sig.mascot
Sig.hispid
Neo.florid
Neo.albigu
Mic.nivali
Cri.gambia
Nes.rufus
Mac.ingens
Ger.vallia
Tat.kempi
Ger.nigeri
Sci.niger
Sci.lis
Sci.vulgar
Chi.lanige
Dip.ordii
Ura.ruddi
Lop.flavop
Aco.cahiri
Mus.minuto
Mus.setulo
Mus.saxico
Mus.platyt
Mus.matthe
Mus.crocid
Mus.pahari
Mus spretu
Mus.macedo
Mus.spicil
Mus.muscul
Mus.caroli
Mus.cervic
Mus.cookii
Hyl.stella
Mas.erythr
Rat.norveg
Rat.exulan
Rat.everti
Leo.edwars
Sto.longic
Aco.willso
Apo.sylvat
Per.mexica
Per.bander
Mic.pennsy
Dip.merria
Did.virgin
Tac.splend
Sci.caroli
Pho.vituli
Bal.musulu

Chapter 3

The phylogenetic history of the genus *Mus* inferred from APRT gene sequences

3.1 Abstract

The DNA sequence for approximately 650 bp from the APRT gene of five *Mus* species was obtained. A region of the APRT gene that covered one complete exon and a complete intron of the gene were amplified by PCR. This sequence was used to help determine the phylogeny of rodent species. It was found that a large insert with homology to a rodent B2 repetitive element was present in *Mus pahari* and *Mus crociduroides* and a B1 element was inserted into *Mus saxicola.* Both B1 and B2 elements are inserted at the same site in intron 2 of the APRT gene and are not shared by any other *Mus* species. This is strong evidence indicating a phylogenetic relationship between *Mus pahari* and *Mus crociduroides* and indicates that the mobilization of the B2 element may have occurred before the divergence of the *Coelomys* subgenus. This sequence contains sufficient phylogenetic information to separate *M. caroli* and

M. cookii as a group from the *Mus musculus* group but it is not sufficient to reliably place the *Mus saxicola* species.

3.2 Introduction

Adenine phosphoribosyltransferase (APRT) is a salvage pathway enzyme in purine biosynthesis. It recycles free adenine bases available from degraded nucleic acids and nucleotides (Stryer 1988). APRT catalyzes the direct synthesis of adenosine-5′-monophosphate (AMP) from adenine and 5-phosphoribosyl-1-pyrophosphate. The AMP can then be further utilized in DNA synthesis and other metabolic reactions that require purine (Fieldhouse and Golding 1996).

The APRT gene is a single copy gene about two kilobase (kb) in length. The gene has five exons and four introns that compose 25% and 75% of the sequence respectively (with one large 1 kb intron). The intron/exon locations have been determined from a mouse cDNA sequence (Dush *et al.* 1985). The rat APRT gene organization is typical of other rodent APRTs with five exons, one large intron of 993 bp and three smaller introns averaging 145 bp. It was also found that there is 194 base insertion in intron III of *Mastomys hildebrandtii* and a 175 base insertion in intron II of *Mus pahari* (Fieldhouse and Golding 1996). The 194 base insertion has a high degree of similarity to B1 repetitive elements found within many mouse genes. The *M. pahari* 175 base insertion has a high degree of similarity to mouse and rat B2 repetitive elements. There are no other known repetitive elements present within the rodent APRT gene (Fieldhouse and Golding 1996). B1 repetitive elements are specific to rodent species, they are not present in the closely related Lagomorphs. B2 elements are not related to B1 elements, they are found in the Muridae and Cricetidae but not in the Gliridae and Caviidae. These repetitive elements may be involved with DNA replication, pre-mRNA processing, and genome reorganization (Fieldhouse and Golding 1996).

The only well characterized functional site of the APRT gene is the purine / pyrimidine phosphoribosyl transferase signature sequence which is located in exon IV. The functions of the remaining coding regions are not well understood (Fieldhouse and Golding 1996).

The *Mus* genus includes four subgenera. The species examined in this project cover three subgenera, *M. crociduroides* in subgenus *Coelomys*, *M. saxicola* in *Pyromys*, and *M. caroli, M. cookii*, and *M. spretus* in subgenus Mus. The Mus subgenus species, *M. cookii* and *M. caroli* belong to the Asian group of species which was suggested as a monophyly (She *et al.* 1990). *M. spretus* is in the Palearctic group. *Mus spretus* is associated with *M. spicilegus* and *M. macedonicus* (Sage 1981; Sage, Atchley and Capanna 1993), could be an outgroup to all the other commensal taxa (She *et al.* 1990; Suzuki and Kurihara 1994) and diverged early from the *musculus* and *spicilegus-macedonicus* association (Prager, Tichy and Sage 1996). The consensus is that *Pyromys*, *Coelomys* and *Nannomys* are subgenera of the *Mus* genus and distinct from the Mus subgenus, but the relationship of the three subgenera has yet to be resolved (Ferris *et al.* 1983; She *et al.* 1990; Sourrouille *et al.* 1995). Nishioka (1989) found that the *Coelomys* and *Pyromys* subgenera group together.

3.3 Material and methods

3.3.1 APRT partial sequences

Genomic DNA from *Mus crociduroides*, *M. caroli*, *M. cookii*, *M. saxicola*, and *M. spretus* was extracted from soft tissue. The procedures are the same as those described in Chapter **2**. The parts of the APRT gene sequence that included 309 bp from intron 2, 133 bp from exon 3, 187 bp from intron 3 and 57 bp from exon 4 were PCR amplified. The primers used were 3int1: TCACTTCCTATTGGT (a 15'mer) and Ix4: CAGAGAGTGGTCATTGTGG (a 19'mer). To ensure specific PCR products, we performed nested PCR in some species using the additional primers X5: AGCTGACCTCGCTGAAGGGC (a 20'mer) and 2intl: TGGTAGCTTCAGGGGC (a 16'mer).

PCR reactions were performed in 100 μl containing 50 mM KCl, 10 mM Tris, 2 mM MgCl2, 0.2 mM of each dNTP, 10 pM each of the primers, 2 μl of DNA template (0.03-0.4 μg), and 2.5 units of AmpliTaq Gold DNA polymerase (Perkin Elmer). Reaction conditions were 12 min at 95^{0}C for preheating, 1 min at 95^{0}C, 2 min at 42- 45^{0}C (since different species amplify best at different temperatures), and 2 min at 72^{0}C for 30-50 cycles. The nested PCR reaction conditions were 12 min at 95^{0}C for preheating, 1 min at 95^{0}C, 1 min 30 sec at 40^{0}C (for the 3Intl/x5 primers set) or 45^{0}C (for the 2intl/Ix4 primer set), and 2 min at 72^{0}C for 30-50 cycles. Then the nested PCR products were used to perform another PCR reaction with primers 3Intl and Ix4.

The PCR products were purified using the QIAquick PCR Purification Kit. The purified PCR products were sequenced in both directions using primers 3intl and Ix4. From these sequences a consensus was constructed (547-707 bp) for four of the *Mus* species but for *M. saxicola* only one direction was useful and four separate sequences from the 3intl to Ix4 direction were used to construct the consensus. Additional APRT sequences from eight other rodent species were retrieved from GenBank. These are *Mus musculus*, *Mus spicilegus*, *Mus pahari*, *Mastomys hildebrandtii*, *Stochomys longicaudatus*, *Rattus norvegicus*, *Gerbillus campestris*, and *Mesocricetus auratus*.

3.3.2 APRT sequences analysis

The APRT sequences were aligned using the CLUSTALW program (Thompson, Higgins and Gibson 1994). Dot plots were used to visually compare the insertions and deletions between different species. The alignment was then manually inspected and adjusted by hand incorporating all this information. The aligned sequences were bootstrapped using the SEQBOOT program (Felsenstein 1989) with 100 replicates. Distance matrices were built by DNADIST and the phylogeny inferred using the NEIGHBOR implementation of the neighbor joining algorithm (Felsenstein 1989). Maximum likelihood trees based on quartet puzzling (Strimmer and von Haeseler 1996) were also used to reconstruct the phylogenies of these genes.

3.4 Results

The alignment of the APRT gene of 13 rodent sequences showed that insertions and deletions have occurred in the intron regions. These are most easily observed by dot plots. Seven *Mus* sequences are compared to the *Mus musculus* APRT sequence in Figure 3.1 and 3.2. This shows that *Mus spicilegus*, *Mus spretus*, *Mus caroli* and the shorter sequence of *Mus saxicola* are colinear with *Mus musculus*. In *Mus cookii* there is a deletion of about 100 bp in intron 2 relative to *Mus musculus*. The most dramatic difference however is an insertion of 175 bp in *Mus pahari* and *Mus crociduroides* that is not present in *Mus musculus*. A dot plot of these two sequences together (Figure 3.3) indicates that they share the same insert (with a small difference of approximately 40 bp indel at the junction of this larger insert).

At sites 245-264 there is a 19 bp deletion which appears to have occurred in the ancestor of both *Mus spicilegus* and *Mus musculus*, but not in *Mus spretus* or the other Asian group species. This suggests that *Mus spretus* is an outgroup of *Mus spicilegus* and *Mus musculus*.

Mus saxicola has an unusual sequence in intron 2 of the APRT gene. It differs from any other sequences of the *Mus* genus in that it contains approximately 20 repeated A's inserted in intron 2 of the APRT gene. We could not sequence through this repeated region. Sequence from the 3int1 primer into intron 2 for *Mus saxicola* yielded sequence that showed homology to that of the other *Mus* species beginning at 233 and extending to 362 but beyond this position there appears to be little homology to any of the other *Mus* species. It is extremely interesting that this homology should break down at the exact same position where the B2 element is inserted into *Mus pahari* and *Mus crociduroides.* Although there are similar though less repetitive stretches of A's in intron 2 of *Mus pahari* and *Mus crociduroides* this sequence does not appear to show homology to these B2 elements. Instead it appears to have strong homology from this point onward to a B1 repetitive element with 90% homology over positions 363 to 497 (a length of 134 bp, $P < 2.7 \times 10^{-42}$) to another B1 element from *Mus musculus* (one of several in Acc. AC005259). Only the sequence from the Ix4 primer sequence in the opposite direction is included in the further analysis.

An alignment of the first part of intron 2 from *Mus saxicola* shows substantial homology with sequences from the remainder of the genus (Figure 3.4). The site where the B1 and B2 elements inserted is GAAAGGTGGCAAGAG. This site is duplicated into a direct repeat and borders the B2 inserts in *Mus pahari* and *Mus crociduroides.* The B1 element in *Mus saxicola* begins immediately after this sequence and appears to end with poly $(A)_{22}$. This region is shown in detail in Figure 3.4.

The phylogenies constructed by neighbor-joining (NJ) and maximum likelihood (ML) methods both showed that the Mus subgenus is well clustered together with high bootstrap (BP) and quartet (QT) values (97.4 in NJ tree and 92.3 in ML tree). The Palearctic species group clustered together with 86.0 BP in the NJ tree and 88.3 QT in the ML tree. The *Mus pahari* and *Mus crociduroides* were grouped with high BP (98.9) in the NJ tree but with less support in the quartet puzzling tree. Finally, all *Mus* species were well grouped together in both trees with high BP 99.6 in NJ tree and QT 94.5 in ML tree.

3.5 Discussion

The short interspersed repetative elements B1 and B2 share the characteristic features of oligo (A)-rich tails at their 3′ ends and direct terminal repeats. This suggests that they are reverse transcripts that have been integrated into the genome (Rogers 1985). B1 and B2 have spread throughout rodents via an RNA-mediated process of retroposition (Moshier, Deutch and Huang 1987). This process requires an RNA transcript of the DNA element to be generated and the RNA transcript is then copied by a reverse transcriptase and subsequently integrated into the chromosome at a new location.

The B1 element is known to be homologous to the human Alu repeat but the B2 element does not contain any apparent homologies to the Alu sequence. The B2 element contains regions of homology to the RNA polymerase III split promoter and 4.5S snRNAI (Krayev *et al.* 1982). The B1 and Alu elements are evolutionarily derived from the terminal segments of a 7SL RNA sequence (Ullu and Tschudi 1984). The 7SL RNA is an abundant cytoplasmic RNA which is essential in the processing of signal sequence of secreted proteins (Li and Graur 1991). The transcription of 7SL RNA is controlled by RNA polymerase III. The major difference between B1 and Alu is that Alu is a 300-bp heterodimer (left and right half differ) arranged head-to-tail and connected by a 19-nucleotide A-rich linker while B1 is a 130-bp monomeric unit which most closely resembles the right Alu monomer.

The presence of a B2 element in both *Mus pahari* and *Mus crociduroides* provides strong evidence as a synapomorphic character phylogenetically linking these two species to the exclusion of all other rodent APRT sequences. As shown in Appendix D, no other Mus subgenus species (Palearctic group nor Asian group) have the 174 bp B2 insertion. Nor do other closely related rodent groups. Hence, this insertion most likely occurred in the last common ancestor of *Mus pahari* and *Mus crociduroides.* This phylogenetic placement is similar to those of previous studies

(She *et al.* 1990; Silver 1995). This raises the interesting possibility that the evolution of the subgenus may have been coincident with genomic rearrangements due to the mobilization of repetitive elements as suggested by Kido *et al.* (1991).

That a B1 element would be inserted in the same location in another species is very unusual. Both elements may favour the same target sequence. The sequence 3′ of the B1 element has not been obtained. It could be the direct terminal repeat just like that of the B2 element or there may be another B1 or B2 element linked by the oligo (A). An oligo (A) has been found as a linker of two B1 elements joined as a dimer (Moshier, Deutch and Huang 1987) and as a linker between two Alu dimers (Zietkiewicz *et al.* 1998). Our electrophoresis results (not shown here) suggest that the PCR product of *Mus saxicola* was about 100bp larger than that of *Mus crociduroides.* This may be due to the existance of another element. In order to confirm this, the sequence of this region in *Mus saxicola* needs to be determined.

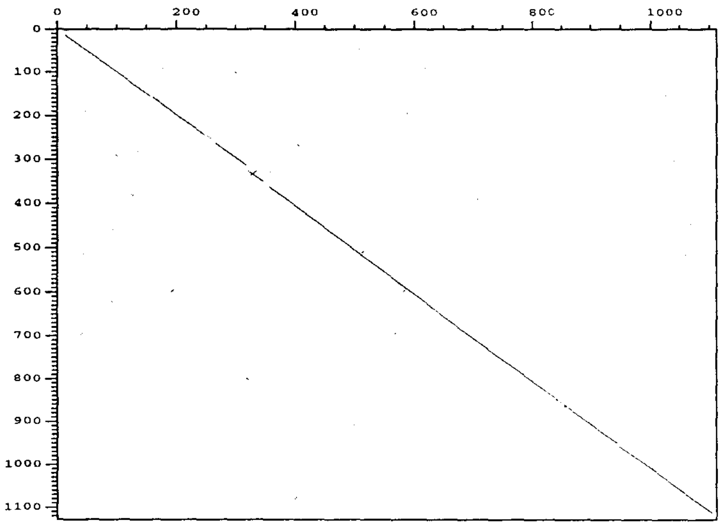
The largest cluster of SINEs within the mouse genome is a cluster of three repetitive elements (one a B1 and the others Alu-related sequences) within a 700 bp span (Kalb *et al.* 1983). The mechanisms of their mobilization are unknown but one possible explanation for this cluster is that a second element may easily insert into the oligo A tract of another B1 element. The significance of these dimers and clusters is unclear.

A 19 bp deletion occurred in both *Mus spicilegus* and *Mus musculus* but not in *Mus spretus* or any other rodents species. This suggests that *Mus spretus* is an outgroup to *Mus spicilegus* and *Mus musculus.* This is the same relationship that has been indicated from studies by Nishioka (1989) and Prager, Tichy and Sage (1996). On the other hand the results in Figure 3.5 and Figure 3.6 indicate the opposite. However, the quartet support for this arrangement is only 73% (not significant) while the bootstrap support for this arrangement via neighbor joining is only 42.4%. Hence the available APRT sequence data are too short on its own to significantly resolve this relationship.

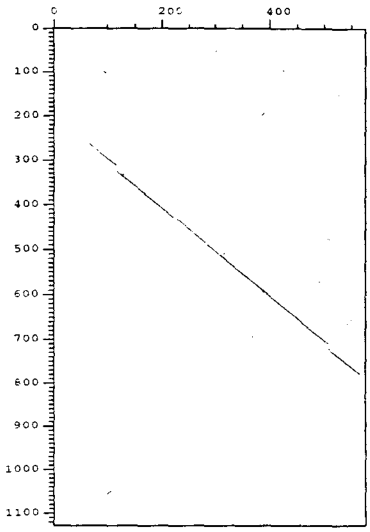
Both the neighbor joining tree and maximum likelihood tree clearly cluster the *Mus* species together and recover the subgeneric classification mentioned previously. These results are comparable with the results from the COII mitochondrial gene. Unfortunately, the APRT gene from species of the subgenus *Nannomys* have not been sequenced, something that needs to be done in the future.

Figure 3.1: A dot plot of four *Mus* species APRT sequences (horizontal) versus the *Mus musculus* APRT sequence (vertical).

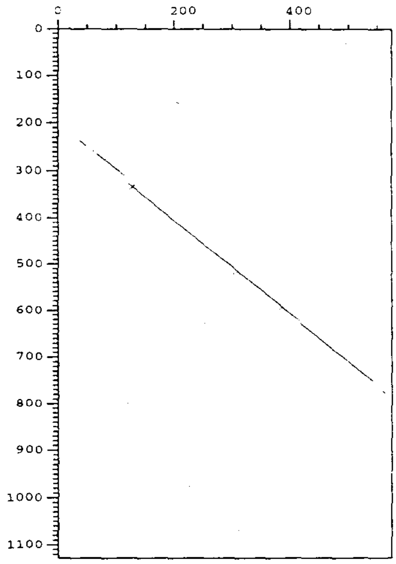
M. spicilegus (horizontal) vs. M. musculus (vertical)



M. spretus (horizontal) vs. M. musculus (vertical)



M. caroli (horizontal) vs. M. musculus (vertical)



M. cookii (horizontal) vs. M. musculus (vertical)

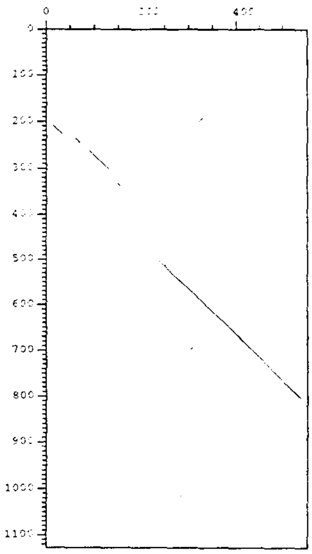
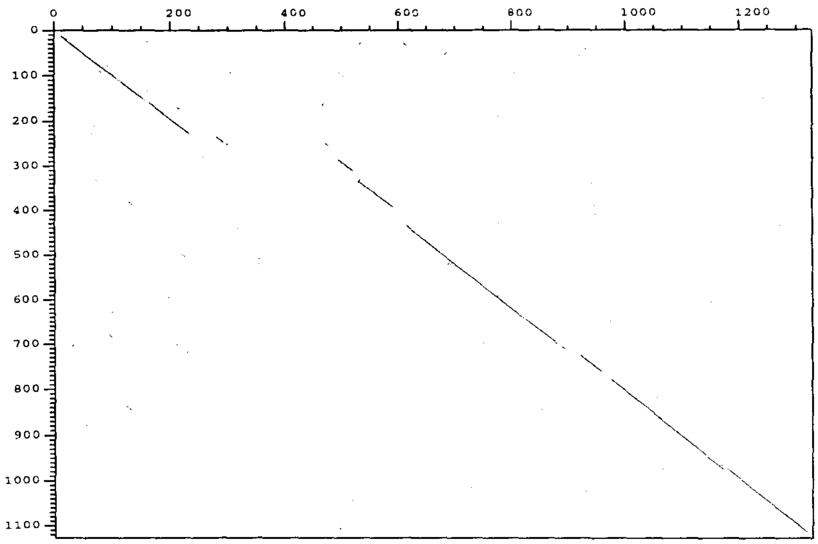
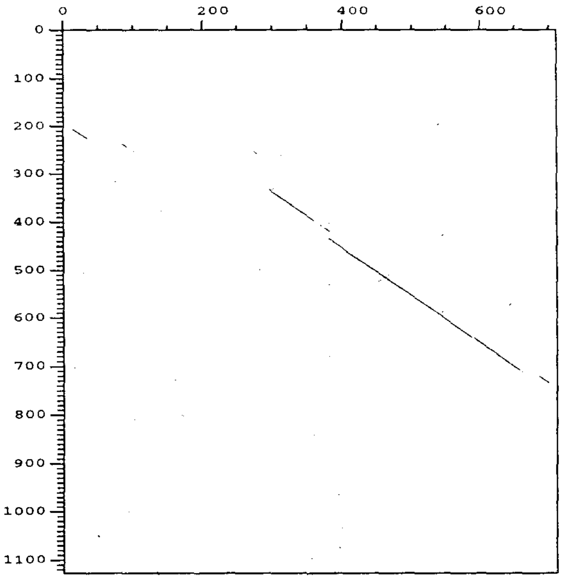


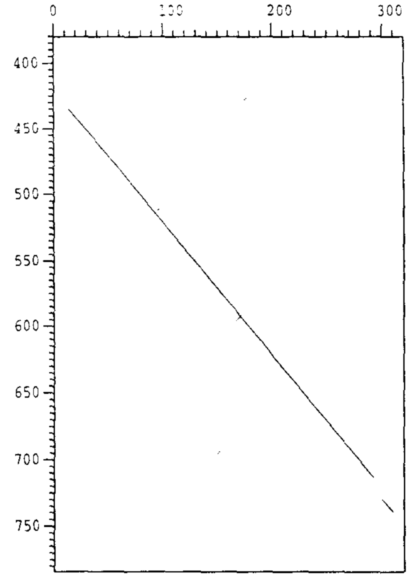
Figure 3.2: A dot plot of three *Mus* species APRT sequences (horizontal) versus the *Mus musculus* APRT sequence (vertical).



M. pahari (horizontal) vs. M. musculus (vertical)



M. crociduroides (horizontal) vs. M. musculus (vertical)



M. saxicola (horizontal) vs. M. musculus (vertical)

Figure 3.3: A dot plot the *M. crociduroides* APRT sequences (horizontal) versus the *M. pahari* APRT sequence (vertical) shows that the *M. pahari* insert is present in both species.

M. crociduroides (horizontal) vs. M. pahari (vertical)

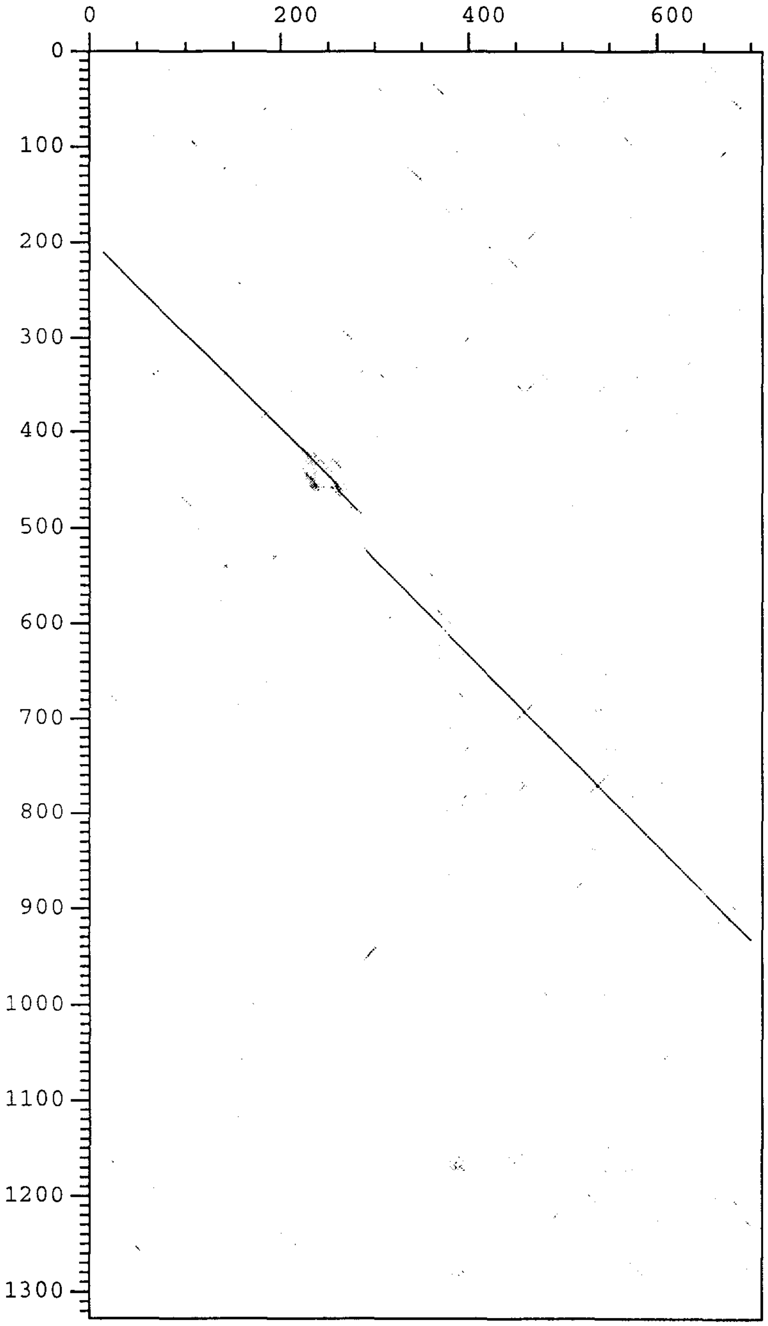


Figure 3.4: A section of the APRT sequence alignment from intron 2 of the *Mus* species showing the identical sites of insertion for B1 and B2 repetitive elements. Numbering is discontinuous and follows appendix D. The direct repeat generated by the SINEs is shaded.

```
          240       250       260       270       280       340       350       360       370
            |         |         |         |         |         |         |         |         |
     CCACGACAGCTCGGGCCGCCTGGAACTGACCTGTAGACAGTGCTCCTGGGTAGATGCTGCATTTGAAAGGTGGCAAGAGGGCTGGTGAG  M.pahari
     CCACGACAGCTCGGGCCGCCTGGAACTGACCTGTAGACAGTGCTCCTGAGTAGATGCTGCATTCGAAAGGTGGCAAGAGGGCTGGTGAG  M.crocid
     CCACGACAGCTCAGGCCTCCTGGAACTGGGT--------------------AGATGCTGCATTTGAAAGGTGGCAAGAG----------  M.caroli
     CCACGACAGCTCAGGCCTCCTGGAACTGGAT--------------------AGATGCTGCATTTGAAAGGTGGCAAGAG----------  M.cookii
.... CCACGACAGCTCAGGCCTCCTGRAACTGGGT--------------------AGATGTTGCATTTGAAAGGTGGCAAGAG----------  M.spretu
     CCACGACAGCTCAGGCCTCCTGGAACTGGGT--------------------AGATGCTGCATTTGAAAGGTGGCAAGAGCCGGGTGTGG  M.saxico
     CCACGACAGCT----------------------------------------AGATGCTGCATTTGAAAGGTGGCAAGAG----------  M.spicil
     CCACGACAGCT----------------------------------------AGATGCTGCATTTGAA-GGTGGCAAGA-----------  M.mus
     CCACGACAGCT----------------------------------------AGATGCTGCATTTGAAA-GTGGCAAGA-----------  M.musA
     CCACGACAGCT----------------------------------------AGATGCTGCATTTGAAAGGTGGCAAGAG----------  M.musB

           380                       500       510       520       530            610
             |                         |         |         |         |              |
     ATGGCTCAGCGGTTAG   B2     TAAATAAACAAATCTTAAAAAAAAAAAAAAGAAAGAAAGGTGGCAAGAGCCACCATAG  M.pahari
     ATGGCTCAGCGGTTAG          TAAATAAATAAATCTTTAAAAATAAA----GAAAGAAAGGTGGCAAGAGCCACC----  M.crocid
     ----------------          -------------------------------------------------CCACCATAG  M.caroli
     ----------------  ......  -------------------------------------------------CCACCATAG  M.cookii
     ----------------          -------------------------------------------------CCACCATAG  M.spretu
     TGGCGCACGCCTGTAA   B1     CTGTCTCGAAAAAACCAAAAAAAAAAAAAAAAAAAAAAnnnnnnnnnnnnnnnnnnnn  M.saxico
     ----------------          -------------------------------------------------CCACCATAG  M.spicil
     ----------------          -------------------------------------------------CCACCATAG  M.mus
     ----------------          -------------------------------------------------CCACCATAG  M.musA
     ----------------          -------------------------------------------------CCA--ATAG  M.musB

             620     700       710       720       730
               |       |         |         |         |
     --TGGA------------GAAGGCAGGTAGGAT-CCCCAAGGCTAAGA               M.pahari
     ------------------------------------------TAAGA               M.crocid
     CGTTGCCTGAGCTGTACAGAAGGCAGGTAGGATTCCCCAAGGCTAAGA               M.caroli
     CGTGGCCTGAGCTGTACAGAAGGCAGGTAGGAT-CCCCAAGGCTAAGA               M.cookii
     CGTGGCCTGAGCTGTTCAGAAGGGAGGTAGGAT CCCCAAGGCTGAGA    ........   M.spretu
     nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn               M.saxico
     CGTGGCCTGAGCTGTTCAGAAGGCAGGTAGGAT-CCCCAAGGCTGAGA               M.spicil
     CGTGGCCTGAGCTGTTCAGAAGGCAGGTAGGAT-CCCCAAGGCTGAGA               M.mus
     CGTGGCCTGAGCTGTTCAGAAGGCAGGTAGGAT-CCCCAAGGCTGAGA               M.musA
     C--GG--TGAGCTGTTCAGAAGGCAGGTAGGAT-CCCCAAGGCTGAGA               M.musB
```

Figure 3.5: A phylogeny of the rodent APRT sequences constructed with the neighbor joining algorithm. All branch points with more than 80% bootstrap support are labeled. All three *Mus musculus* sequences lie along the branch shown.

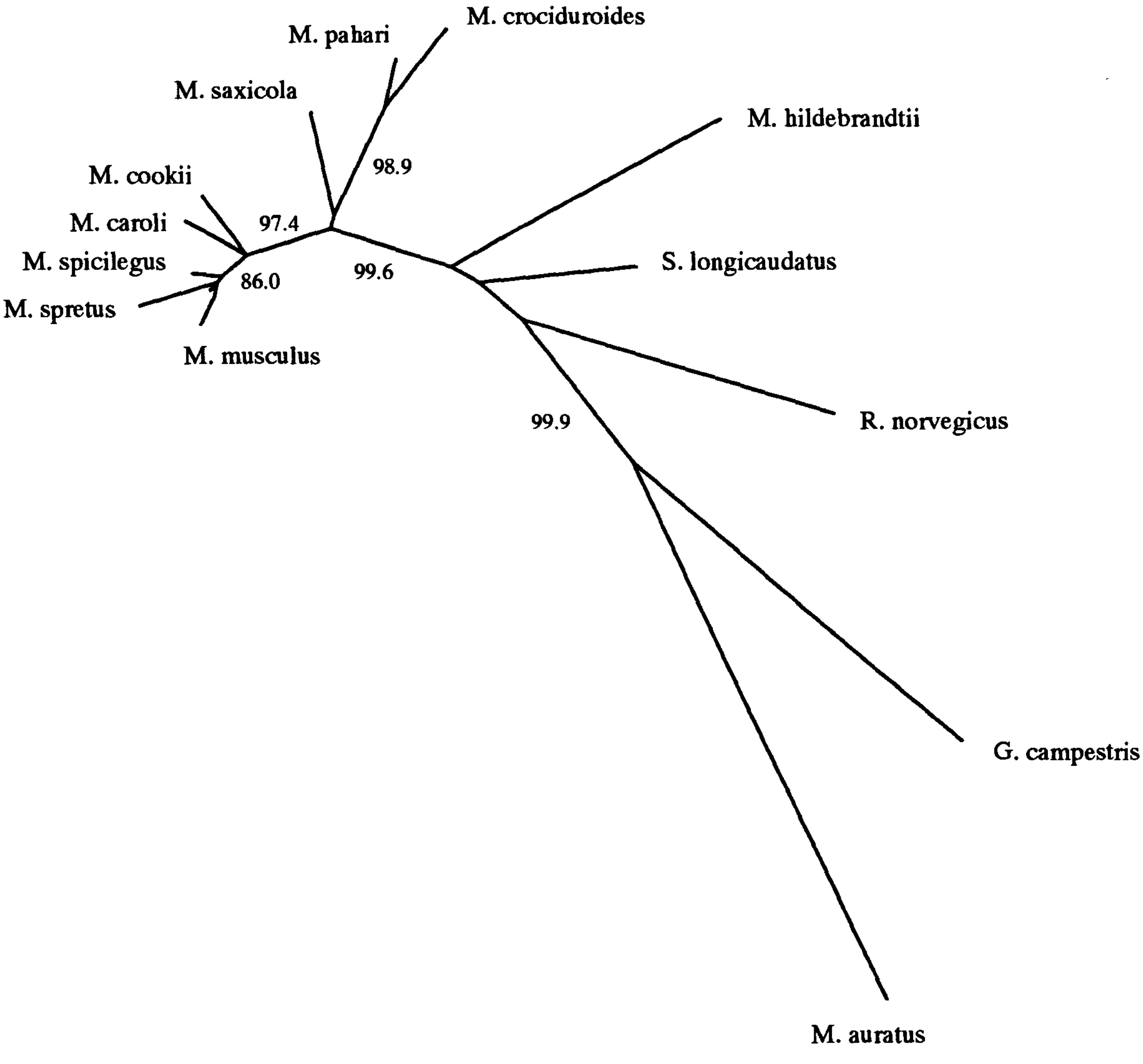
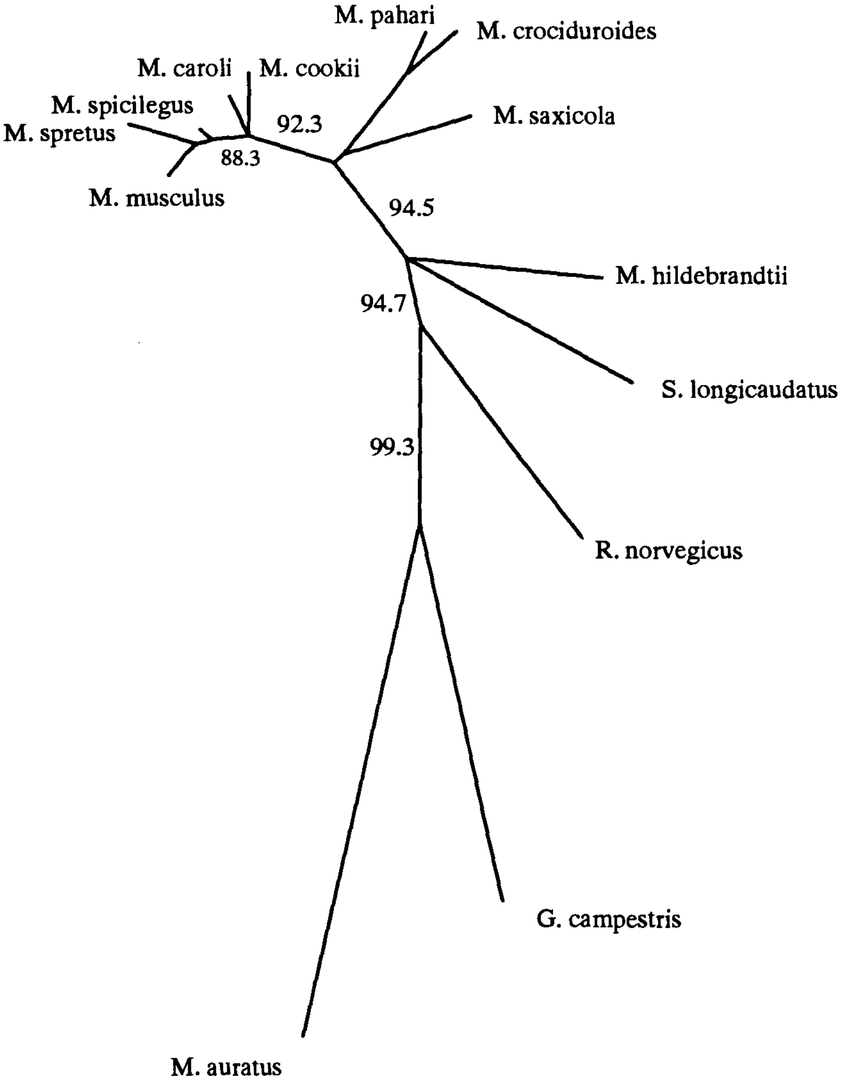

M. crociduroides
M. pahari
M. saxicola
98.9
M. hildebrandtii
M. cookii
M. caroli
97.4
M. spicilegus
86.0
99.6
S. longicaudatus
M. spretus
M. musculus
99.9
R. norvegicus
G. campestris
M. auratus

Figure 3.6: A phylogeny of the rodent APRT sequences constructed with maximum likelihood quartet puzzling. All branch points with more than 80% quartet support are labeled. All three *Mus musculus* sequences lie along the branch shown.


M. pahari
M. crociduroides
M. caroli
M. cookii
M. spicilegus
M. spretus
92.3
M. saxicola
88.3
M. musculus
94.5
M. hildebrandtii
94.7
S. longicaudatus
99.3
R. norvegicus
G. campestris
M. auratus

Chapter 4

An examination of lineage rate variation in the *Mus* genus

4.1 Abstract

To further examine rate variation between different closely related species of rodents the sequences of the COII gene and the APRT gene have been sequenced and analyzed. A likelihood ratio test of the molecular clock hypothesis for the mitochondrial COII gene shows that a molecular clock cannot be rejected for the species of the *Mus* genus. However, the likelihood ratio test for the APRT gene revealed that a molecular clock hypothesis must be rejected for these closely related *Mus* species. The existence of different evolutionary patterns among the APRT gene and mitochondrial COII gene in the *Mus* species might indicate that these genes have been under different selective constraints during their evolution. We suggest that these rate variations can not be due to any change of generation length (a commonly suggested cause of evolutionary rate variation) as these sequences exist within the same species or within species with similar generation length.

4.2 Introduction

The molecular clock hypothesis was proposed by Zuckerkandl and Pauling in 1965. This hypothesis postulates that the rate of evolution in any given protein or DNA sequence is approximately constant over time in all evolutionary lineages. A simplistic neutral theory of molecular evolution suggests that a molecular clock would exist if the mutation rate to neutral alleles per year were equal among all organisms at all times. Any deviation from the equality of neutral mutation rate per year makes the molecular clock less exact (Kimura 1987). Two potential causes leading to the deviation (Kimura 1987) are a change of mutation rate per year (e.g. due to a change of generation span) and an alteration of the selective constraint on each molecule (e.g. due to a change of environment).

There is good evidence that the rate of molecular evolution has become slower in hominoids after their separation from the Old World monkeys (Goodman 1961). This has become known as the hominoid rate-slowdown and a generation-time effect hypothesis was proposed as the cause (Li *et al.* 1996). Higher evolutionary rates have also been observed in monkeys than in humans and higher rates in rodents than in primates (Li, Tanimura and Sharp 1987). These have been explained by a generation-time effect (Li and Graur 1991) since when organisms with similar generation times such as mice and rats are compared, rate constancy holds fairly well (Li, Tanimura and Sharp 1987). However, lagomorphs have about the same generation time as rodents but do not show the same degree of evolutionary rate difference (Li *et al.* 1990). The generation-time hypothesis assumes that an important source of mutation is DNA replication errors coincident with germ cell division. These cell divisions are more strongly correlated with generation length than with chronological time. Although the molecular clock is known to be false at several levels, within the muroid rodents it was found to be true based on the analysis of sequences from the mouse, rat, hamster and human (O'hUigin and Li 1992). However the rat APRT sequence analysis revealed that the evolutionary rate varied among exons and that there was a recent and rapid burst of substitutions within the mouse lineage (Fieldhouse and Golding 1993). A five fold difference of evolutionary rate between two closely related species *Mus spicilegus* and *Mus musculus* was found by Fieldhouse and Golding (1996). Generation time can not be a factor leading to this change of rate. The causes of the change of rate among *Mus* species and between closely related species requires further investigation. As a first step toward that goal we should determine if the different rates in rodents (Wu and Li 1985; Fieldhouse and Golding 1993) are consistent across all rodent species and across all genes. To this end we have analyzed mitochondrial COII and autosomal APRT gene sequences from 27 and 13 species of rodents, respectively.

4.3 Material and methods

In this study we have focused on species within the *Mus* genus, especially those closely related species within the *Mus* subgenus which contains *Mus spicilegus* and *Mus musculus* as these have been found to have a five fold difference in evolutionary rate. The mitochondrial gene used is cytochrome C oxidase subunit II. This is compared with a nuclear gene encoding adenine phosphoribosyltransferase (only partial sequence from intron 2, exon 3, intron 3 and exon 4 of this gene was used). The sequencing and collection of these sequences has been described in Chapters 2 and 3, respectively.

To test for the existence of a molecular clock in the COII or the APRT gene, a maximum likelihood method is employed. The likelihood of the best phylogeny is calculated assuming that rates are free to vary along each branch. This implies that $2n - 3$ rate parameters (one for each subbranch; where n is the number of taxa) must be estimated (Felsenstein 1981b; Felsenstein 1989). The likelihood of the best tree is then calculated assuming that a molecular clock is true. This implies that $n - 1$ branch points must be estimated. If these two phylogenies are the same then a likelihood ratio test can be performed. The ratio of two likelihoods based on the same topology will asymtoptically have a χ^2 distribution. This distribution will have its degrees of freedom determined by the difference in the number of parameters estimated by each likelihood model. In this case $(2n-3)-(n-1) = n-2$ degrees of freedom. Therefore, running these two programs on the same data makes it possible to statistically test the molecular clock hypothesis. If the difference in likelihoods is not excessively large (as determined in comparison to a χ^2 distribution) then the added constraint of a molecular clock does not significantly alter the likelihood of observing these data. A molecular clock hypothesis can not therefore be rejected. If the difference in likelihoods is excessively large then the added constraint does reduce the likelihood and it suggests that it is unlikely to observe these data when a molecular clock is present. The hypothesis of a molecular clock is rejected.

If the two trees support different topologies then the ratio of likelihoods will not properly follow a χ^2 distribution. In this case, we have tested the topology favoured assuming a molecular clock against this same topology without the assumption of a molecular clock. It should be noted however, that this is not the corresponding maximum likelihood tree but it does give an indication of how well the likelihood might be improved.

4.4 Results

The rate tests for the rodent COII sequences demonstrate that a molecular clock does not hold for this gene when all species are considered. The assumption of a molecular clock forces differences in the topology (Figure 4.1). Within the *Mus* genus for example, (i) *Mus cookii* was forced from being an outgroup of all the *Mus* subgenera species to being an outgroup of only the Asian group species *Mus cervicolor* and *Mus caroli* and (ii) *Mus crociduroides* was also changed from an outgroup of *Mus platythrix*, *Mus setulosus* and *Mus minutoides* to a branch clustering with *Mus platythrix* and sharing a common ancestor with *Mus setulosus* and *Mus minutoides*. When a tree with the identical topology to that of the tree with a molecular clock is used, there is still a significant though small difference in the rates of substitutions among species (Table 4.1).

If only the *Mus* species are considered the assumption of a molecular clock again forces a change in the topology of the phylogeny (Figure 4.2). In this case the slowly evolving *Mus cookii* lineage is forced to branch more recently within the *Mus spretus*-*Mus musculus* group. However, this change does not really alter the size of the likelihood and the difference between a phylogeny with or without a molecular clock is not significant. Therefore, there is no reason to reject the hypothesis of a molecular clock for the COII gene within the *Mus* genus.

There is no reason to expect or not to expect that any lineage rate variation demonstrated for mitochondrial genes would or would not be reflected in nuclear encoded genes. Besides their subcellular location a major difference between these two is that the APRT sequence includes non-coding sequence.

The rate tests for the rodent APRT sequences demonstrate that a molecular clock does not hold for the region of the gene tested. The results of likelihood ratio tests (Table 4.2) indicate that a phylogeny that assumes a molecular clock is significantly less likely ($P < 0.01$) than a phylogeny that does not assume a molecular clock (Figure 4.3). If only the sequences of the *Mus* genus are considered the same result is found (Table 4.2; Figure 4.4).

There does not however appear to be any obvious pattern to the rate variation. One of the major differences appears to be the relative rates of *Mus spretus* versus the rest of the *Mus* sequences. Indeed there is such a large difference in the rates that when the phylogeny is forced to follow a molecular clock the *M. spretus* sequence is forced to branch more deeply and the phylogeny differs between the two trees (Figure 4.4). This invalidates the statistical accuracy of the likelihood ratio test because the two trees are different and hence the same likelihood space is not explored by both methods.

To determine if this species was an important factor, the sequence for *Mus spretus* was deleted and the tests were redone. Although *M. spretus* has a large variation in its substitution rate relative to the remainder of the genus there is still sufficient variation that the likelihood of a phylogeny with a molecular clock is significantly less likely ($P < 0.01$) than a phylogeny without (Figure 4.5).

4.5 Discussion

There are at least eight hypotheses about the causes of the variation rate of evolution. They are differences in generation time (or more accurately germ cell division) (Li, Tanimura and Sharp 1987), in DNA repair efficiency (Britten 1986), in metabolic rate, in life history/environmental factors, in base composition/base context, in chromosomal location of the gene, in population size and in selection. The five fold difference of evolutionary rate between *Mus spicilegus* and *Mus musculus* observed by Fieldhouse and Golding (1996) suggests that generation time is not a factor leading to this change of rate. Rather, it was suggested that the cause might be the different levels of association with man that these species exhibit. However, whether the difference of evolutionary rate exists among other closely related *Mus* species or whether the rate variation of APRT genes is consistent across all rodent species requires more evidence.

Closely related rodents species were chosen to attempt to eliminate or reduce differences in the above factors. A comparison between nuclear DNA and mitochondrial DNA indicates if variation in evolutionary rates occurred in both the nuclear genome and the mitochondrial genome. This comparison does however cause a dramatic change in effective population size since mitochondrial genes will have an effective size one quarter that of nuclear genes.

Our results indicate that the pattern of evolutionary rates are different for nuclear genes and mitochondrial genes within the same species. Likelihood ratio test of the mitochondrial COII gene suggest that a molecular clock can not be rejected within the *Mus* genus. However, the same test for the APRT gene revealed that evolutionary rates are not constant within the *Mus* genus and that a molecular clock probably does not hold true.

Possible causes of this difference are that (i) these nuclear genes and mitochondrial genes underwent different selective constraints; (ii) the APRT selective constraint varied in different species; (iii) the difference is a reflection of differences in DNA repair system, metabolic rate and effective population size between nuclear and mitochondrial genes.

The causes leading to the variation of evolutionary rates do not strongly affect mitochondrial DNA within the *Mus* genus but do affect mtDNA among more distant species. Thus, the clock appears to work only on a limited level. The causes of the rate variation in the APRT gene are sufficient to affect the rates even among closely related species.

If the generation-time hypothesis is true, the evolutionary rate should be tied to the mutation rate, and thus the evolutionary rate should be more closely related to the number of generations per unit time than to time itself (Li, Tanimura and Sharp 1987). However, our results show that the variation of evolutionary rate is not correlated to generation time in these *Mus* species and other rodent species. Therefore, if our data prove general it would provide more evidence against the hypothesis of a generation time effect. Our analysis of the variation of evolutionary rate in APRT in the *Mus* genus minimally indicates that the evolutionary forces influencing a molecular clock act totally different between mitochondria and autosomes or that selection has had a strong influence on the evolutionary pattern of these APRT genes.

Our data suggest variation of evolutionary rate among species of the *Mus* genus, but did not show a large burst of evolutionary changes between *Mus spicilegus* and *Mus musculus* as was found previously (Fieldhouse and Golding 1996). Even with the same data as was previously analyzed there is no indication of a rapid burst in *Mus musculus*. The most likely reason for this is the incomplete APRT sequence examined. There are insufficient numbers of substitutions in this small region of the gene to determine if *Mus musculus* and *Mus spretus* (the two species most commensal with man) are evolving at a faster rate. Complete APRT gene sequences in these species are needed to further address this problem.

Table 4.1: Tests of the molecular clock for rodent COII sequences.

	Without a molecular clock log($LIKL$)	With a molecular clock log($LIKL$)	-2log(LR)	degrees of freedom	$\chi^2_{5\%}$
27 Rodent Species[†]	-9263.1	-9296.5	66.8	25	37.7
27 Rodent Species[*]	-9276.9	-9296.5	39.2	25	37.7
11 *Mus* Species[†]	-3519.6	-3529.0	18.8	9	16.9
11 *Mus* Species[*]	-3524.4	-3529.0	9.2	9	16.9

[†] Not a valid test due to the differences in the phylogenies with and without a molecular clock (see text). [*] Using the phylogeny favoured by a tree which follows a molecular clock.

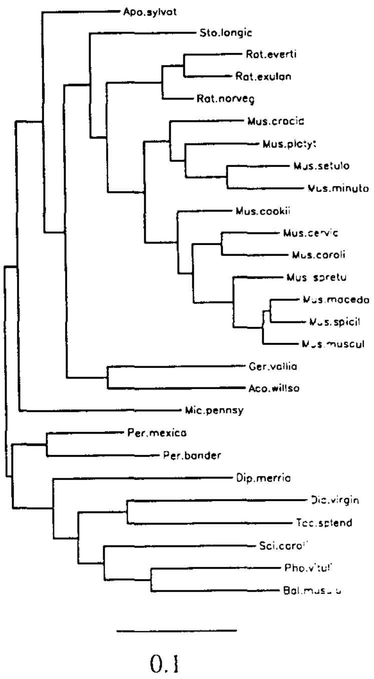
Table 4.2: Tests of the molecular clock for rodent APRT sequences.

	Without a molecular clock log($LIKL$)	With a molecular clock log($LIKL$)	-2log(LR)	degrees of freedom	$\chi^2_{5\%}$
15 Rodent Species†	-5349.5	-5368.0	37.0	13	22.4
15 Rodent Species*	-5351.0	-5368.0	34.0	13	22.4
10 *Mus* Species†	-2681.8	-2693.5	23.4	8	15.5
10 *Mus* Species*	-2683.4	-2693.5	20.3	8	15.5
9 *Mus* Species	-2618.7	-2628.1	18.7	7	14.1

† Not a valid test due to the differences in the phylogenies with and without a molecular clock (see text). * Using the phylogeny favored by a tree which follows a molecular clock.

Figure 4.1: Maximum likelihood trees of the COII sequences without (A) and with (B) a molecular clock. The scale bar denotes 0.1 substitutions per site. There are a minimum of 1879 and 1902 substitutions required for each tree, respectively.

A


Apo.sylvat
Sto.longic
Rat.everti
Rat.exulan
Rat.norveg
Mus.cracic
Mus.platyt
Mus.setulo
Mus.minuto
Mus.cookii
Mus.cervic
Mus.caroli
Mus.spretu
Mus.macedo
Mus.spicil
Mus.muscul
Ger.vallia
Aco.willso
Mic.pennsy
Per.mexica
Per.bander
Dip.merria
Dic.virgin
Toc.splend
Sci.caroli
Pho.vitul
Bal.muscul
0.1

B

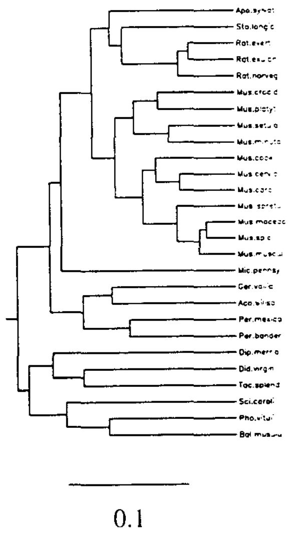
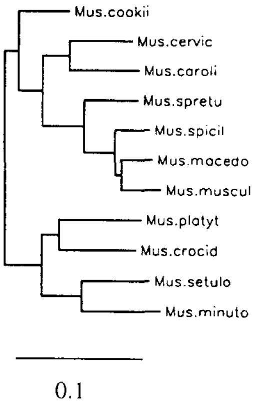

Apo.sylvat
Sto.longic
Rat.everti
Rat.exulan
Rat.norveg
Mus.cracic
Mus.platyt
Mus.setulo
Mus.minuto
Mus.cookii
Mus.cervic
Mus.caroli
Mus.spretu
Mus.macedo
Mus.spicil
Mus.muscul
Mic.pennsy
Ger.vallia
Aco.willso
Per.mexica
Per.bander
Dip.merria
Did.virgin
Toc.splend
Sci.caroli
Pho.vitul
Bal.muscul
0.1

Figure 4.2: Maximum likelihood trees of the *Mus* COII sequences without (A) and with (B) a molecular clock. The scale bar denotes 0.1 substitutions per site. There are a minimum of 522 and 525 substitutions required for each tree, respectively.

A


Mus.cookii
Mus.cervic
Mus.caroli
Mus.spretu
Mus.spicil
Mus.macedo
Mus.muscul
Mus.platyt
Mus.crocid
Mus.setulo
Mus.minuto
0.1

B

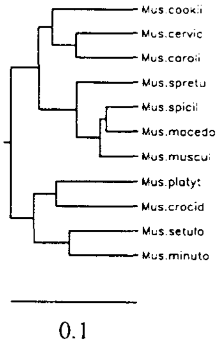
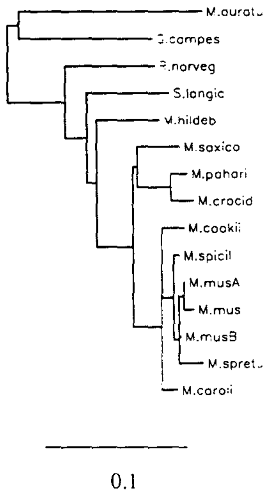

Mus.cookii
Mus.cervic
Mus.caroli
Mus.spretu
Mus.spicil
Mus.macedo
Mus.muscul
Mus.platyt
Mus.crocid
Mus.setulo
Mus.minuto
0.1

Figure 4.3: A maximum likelihood tree of the APRT sequences without (A) and with (B) a molecular clock. The scale bar denotes 0.1 substitutions per site. There are a minimum of 2542 substitutions required for this tree.

A


M.auratu
G.campes
R.norveg
S.longic
M.hildeb
M.saxico
M.pahari
M.crocid
M.cookii
M.spicil
M.musA
M.mus
M.musB
M.spretu
M.caroli
0.1

B

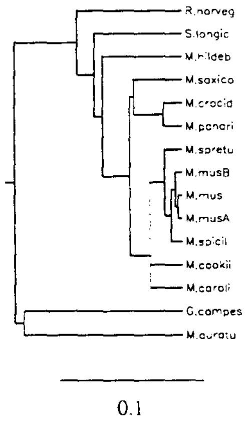
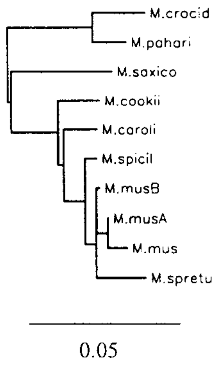

R.norveg
S.longic
M.hildeb
M.saxico
M.crocid
M.pahari
M.spretu
M.musB
M.mus
M.musA
M.spicil
M.cookii
M.caroli
G.campes
M.auratu
0.1

Figure 4.4: A maximum likelihood tree of the *Mus* APRT sequences without (A) and with (B) a molecular clock. The scale bar denotes 0.05 substitutions per site. There are a minimum of 651 and 620 substitutions required for each tree, respectively.

A


M.crocid
M.pahari
M.saxico
M.cookii
M.caroli
M.spicil
M.musB
M.musA
M.mus
M.spretu
0.05

B

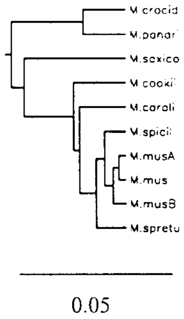
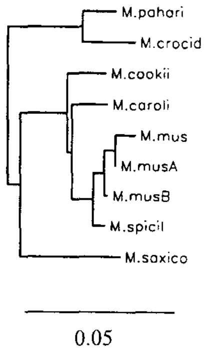
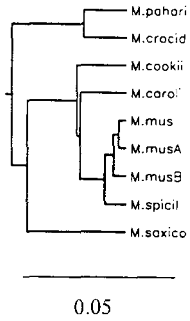

M.crocid
M.pahari
M.saxico
M.cookii
M.caroli
M.spicil
M.musA
M.mus
M.musB
M.spretu
0.05

Figure 4.5: A maximum likelihood tree of the *Mus* APRT sequences (excluding *Mus spretus*) without (A) and with (B) a molecular clock. The scale bar denotes 0.05 substitutions per site. There are a minimum of 596 substitutions required for this tree.

A


M.pahari
M.crocid
M.cookii
M.caroli
M.mus
M.musA
M.musB
M.spicil
M.saxico
0.05

B


M.pahari
M.crocid
M.cookii
M.caroli
M.mus
M.musA
M.musB
M.spicil
M.saxico
0.05

Chapter 5

Summary

5.1 Phylogenetic analysis of mitochondrial genes

Cytochrome c oxidase subunit II (COII) gene sequences were PCR amplified and sequenced from fifteen rodents species; *Dipldomys merriami, Peromyscus mexicanus, Stochomys longicudatus, Rattus everti, Rattus exulans, Mus spicilegus, Mus macedonicus, Mus cervicolor, Mus caroli, Mus minutoides, Mus setulosus, Mus cookii, Mus platythrix, Mus crociduroides,* and *Mus spretus.* The COII sequences of 10 additional rodents species were retrieved from GenBank. Sequences from the seal and whale were used to root the trees. The alignment of all 27 COII sequences was performed using CLUSTALW (Thompson, Higgins and Gibson 1994) and then adjusted manually. The aligned sequences were analyzed by neighbor joining distance and parsimony algorithms. Bootstrap programs were used to give a statistical evaluation.

Partial or complete 12S rRNA gene sequences were collected from DNA databases for 107 mammals species including 96 Rodentia species, 9 Carnivora, one Edentata and one Lagomorpha species. The alignment of sequences was performed using the CLUSTALW program. The initial alignment was manually inspected and adjusted by hand. The aligned sequences were bootstrapped and analyzed by consensus neighbor-joining and consensus parsimony algorithms. An Edentata species (*Dasypus novemcinctus*) was used to root the trees.

The analysis of COII sequence by neighbor-joining and parsimony approaches support the view that *Mus spretus* is a sister species of *Mus macedonicus*, *Mus spicilegus* and *Mus musculus*. Both the neighbor-joining tree and the parsimony tree show a bifurcation between *Mus spretus* and *Mus musculus-Mus spicilegus-Mus macedonicus* suggesting that *Mus spretus* diverged earlier than the others followed by *Mus musculus*. The COII sequence data provide evidence to support a sister species relationship between *Mus spicilegus* and *Mus macedonicus* within the palearctic group. In addition, *Mus caroli* and *Mus cervicolor* are grouped together. Although the overall phylogeny of the 12S rRNA gene sequences give variable results, these analyzes do suggest that the subgenus *Pyromys* species (*Mus platythrix* and *Mus saxicola*) group together and the subgenus *Coelomys* species (*Mus crociduroides* and *Mus pahari*) group together.

5.2 Phylogenetic analysis of the APRT nuclear encoded gene

The DNA sequences for approximately 650 bp covering the region from intron 2 to exon 4 of the APRT gene of five *Mus* species (*Mus crociduroides*, *Mus caroli*, *Mus cookii*, *Mus saxicola* and *Mus spretus*) were obtained by PCR amplification and sequencing. An additional 10 APRT sequences from other eight rodents species were retrieved from GenBank. Dot plots, neighbor joining and maximum likelihood algorithms were used to analyze the sequences. Blast searches were used to identify two large inserts. It was found that a large insert with homology to a rodent B2 repetitive element was present in *Mus pahari* and *Mus crociduroides* and a B1 element was inserted into *Mus saxicola.* Both B1 and B2 elements inserted at the same site in intron 2 of the APRT gene and were not shared by any other *Mus* species. This supports the conjecture that *Mus pahari* and *Mus crociduroides* are closely related subgenus species and indicates that the mobilization of the B2 element may have occurred before the divergence of the *Coelomys* subgenus. The B1 element that inserted at the same site as B2 indicates that B1 and B2 may favour the same target sequence. The analysis of these APRT sequences provides evidence that *Mus caroli* and *Mus cookii* are distinct from the *Mus musculus* group.

5.3 Molecular clocks in mitochondrial genes and nuclear genes

To further examine rate variation between different closely related species of rodents, the sequences of both the COII gene and the APRT gene have been analyzed and a likelihood ratio test of the molecular clock hypothesis has been carried out. For the mitochondrial COII gene this test showed that a molecular clock cannot be rejected for the species of the *Mus* genus. However, the likelihood ratio test for the APRT gene revealed that a molecular clock hypothesis must be rejected for these closely related *Mus* species. The existence of different evolutionary patterns among the APRT gene and mitochondrial COII gene in the *Mus* species might indicate that these genes have been under different selective constraints during their evolution. We suggest that these rate variations can not be due to any change of generation length as these sequences exist within the same species or within species with similar generation length.

Appendix A

Aligned COII sequence data

```
    1          11         21         31         41         50
    |          |          |          |          |          |
1   TCGAACCCCC TAAAATTGGT TTCAAGCCAA TTTCATA-TC CTATATGTCT   49 Mus.plat
1   TCGAACCCCC TAAAACTGGT TTCAAGCCAA TCTCATA-TC CTATATGTCT   49 Mus.croc
1   TCGAACCCCC TAAAACTGGT TTCAAGCCAG CTCCATA-TC CTATATGTCT   49 Mus.setu
1   TCGAACCCCC TGAAATTGGT TTCAAGCCAA CCCCATA-TC CTATATGTCT   49 Mus.caro
1   ---------- ---------- ---------- TCTCATA-TC CTATATGTCT   19 Mus.cerv
1   TCGAACCCCC TAAAACTGGT TTCAAGCCAA TCTCATA-TC CTATATCTCT   49 Mus.cook
1   TCGAACCCCC TAAAACTGGT TTCAAGCCAG TCCCATA-TC CTATATGACT   49 Mus.minu
1   TCGAACCCCC TAGAATTGGT TTCAAGCCAA CCTCATAATC CTATATGTCT   50 Mus.mace
1   TCGAACCCCC TTAAATTGGT TTCAAGCCAA CCTCATA-TC CTATATGTCT   49 Mus.spre
1   TCGAACCCCC TACAACTGGT TTCAAGCCAA TCTCATA-AC CACTATGTCT   49 Rat.ever
1   TCGAACCCCC TACAACTGGT TTCAAGCCAA TTTCATA-AC CACTATGTCT   49 Rat.exul
1   TCGAACCCCC TAAAATTGGT TTCAAGCCAA CTCCATA-AC CTCTATGTCT   49 Sto.long
1   TCGAACCCCC CAAAATTGGT TTCAAGCCAA CCTCATA-TC CTATATGTCT   49 Mus.spic
1   TCGAACCCCC CAAAACTGAT TTCAAGTCAG CCCCATA-TC CAATATGTCT   49 Dip.merr
1   TCGAACCCCC TTAAACTGGT TTCAAGCCAA TCTCATA-AC CTCTATGTCT   49 Per.mexi
1   TCGAACCCCC TAAAATTGGT TTCAAGCCAA TCTCATA-TC CTATATGTCT   49 Mus.musc
1   TCGA-CCCCC TACAACTGGT TTCAAGCCAA TTTCATA-AC CATTATGTCT   48 Rat.norv
1   TCGAACCTTC TCCCATTGGT TTCAAGCCAA TATCATA-AC CACTATGTCT   49 Bal.musu
1   TCGAACCCTC TGAAACTGGT TTCAAGCCAA CACCATA-AC CCTTATGTCT   49 Pho.vitu
1   TTGAACCCCC TAAGATTAAT TTCAAGTCAA TCCCATA-AC CCTTATGACT   49 Did.virg
1   ---------- ---------- ---------- ---------- ----------    0 Ger.vall
1   ---------- ---------- ---------- ---------- ----------    0 Mic.penn
1   ---------- ---------- ---------- ---------- ----------    0 Per.band
1   ---------- ---------- ---------- ---------- ----------    0 Tac.sple
```

```
 1  ---------- ---------- ---------- ---------- ----------    0 Apo.sylv
 1  ---------- ---------- ---------- ---------- ----------    0 Aco.will
 1  ---------- ---------- ---------- ---------- ----------    0 Sci.caro

    51         61         71         81         91       100
    |          |          |          |          |        |
50  TTCTC---AA -TAAGATATT AGTAAAA-TA ATTACATAAC TTTGTCAAA-   93 Mus.plat
50  TTCTC---AA -TAAGATATT AGTAAAA-TA ATTACATAAC TTTGTCAAA-   93 Mus.croc
50  TTCTC---AA -TAAGATATT AGTAAAA-TT ATTACATAAC TTTGTCAAA-   93 Mus.setu
50  TTCTC---AA -TAAGATATT AGTAAAA-TA ATTACATAAC TTTGTCAAA-   93 Mus.caro
20  TTCTC---AA -TAAGATATT AGTAAAA-CA ATTACATAAC TTTGTCAAA-   63 Mus.cerv
50  TTCTC---AA -TAAGATATT AGTAAAA-CA ATTACATAAC TTTGTCAAA-   93 Mus.cook
50  TTCTC---AA -TAAGATATT AGTAAAA-TA ATTACATAAC TTTGTCAAA-   93 Mus.minu
51  TTCTC---AA -TAAGATATT AGTAAAA-CA ATTACATAAC TTTGTCAAA-   94 Mus.mace
50  TTCTC---AA -TAAGATATT AGTAAAA-CA ATTACATAAC TTTGTCGAA-   93 Mus.spre
50  TTCTC---AA -TGAGATATT AGTAAAA-TA ATTACATAAC CTTGTCAAGG   94 Rat.ever
50  TTCTC---AA -TGAGATATT AGTAAAA-TA ATTACATAAC CTTGTCAAGG   94 Rat.exul
50  TTCTC---AA -TGAGATATT AGTAAAAATT ATTACATAAC TTTGTCAAA-   94 Sto.long
50  TTCTC---AA -TAAGATATT AGTAAAA-CA ATTACATAAC TTTGTCAAA-   93 Mus.spic
50  TTCTC---AA ATAAGATATT AGTAAAGTCA ATTACATAAC TTTGTCAAA-   95 Dip.merr
50  TTCTC---AA -TGAGGTATT AGTAAAA-CA ATTACATAAC TTTGTCAAA-   93 Per.mexi
50  TTCTC---AA -TAAGATATT AGTAAAATCA ATTACATAAC TTTGTCAAA-   94 Mus.musc
49  TTCTC---AA -TGAGATATT AGTAAAA-TA ATTACATAAC CTTGTCAAG-   92 Rat.norv
50  TTCTTTATAA ATGAGATATT AGTAAAACC- -TTATATAAC TTTGTCAAA-   96 Bal.musu
50  TTCTCAATTA G-GAGGCATT AGTAAAAA-- -TTACATAAC TTTGTCAAA-   94 Pho.vitu
50  TTCTC---AA A-AAGATATT AGTAAAATTC ATTACATAAC TTTGCCATA-   94 Did.virg
 1  ---------- ---------- ---------- ---------- ----------    0 Ger.vall
 1  ---------- ---------- ---------- ---------- ----------    0 Mic.penn
 1  ---------- ---------- ---------- ---------- ----------    0 Per.band
 1  ---------- ---------- ---------- ---------- ----------    0 Tac.sple
 1  ---------- ---------- ---------- ---------- ----------    0 Apo.sylv
 1  ---------- ---------- ---------- ---------- ----------    0 Aco.will
 1  ---------- ---------- ---------- ---------- ----------    0 Sci.caro

    101        111        121        131        141      150
    |          |          |          |          |        |
94  GTTAAATTAT AGAATAAAA- TC--TATATG TCTT-ATATG GCTTACCCAT  139 Mus.plat
94  GTTAAATTAT AGATTAATAT TC--TATATA TCTT-ATATG GCTTATCCAC  140 Mus.croc
94  GTTAAATTAT AGAGTAAAA- TC--TATATA TCTT-ATATG GCGTATCCAT  139 Mus.setu
94  GTTAAATTAT AGACCAAAAA TC--TATATA TCTT-ACATG GCTTACCCAT  140 Mus.caro
64  GTTAAATTAT AGACCAAAAA TC--TATATA TCTT-ACATG GCTTACCCAT  110 Mus.cerv
94  GTTAAGTTAT AGACCAAAAA TC--TATATA TCTT-ACATG GCTTATCCAT  140 Mus.cook
94  GTTAAATTAT AGAGTAAAA- TCTCTATATA TCTT-TTATG GCATACCCAT  141 Mus.minu
95  GTTAAATTAT AGACCAATAA TC--TATATA TCTT-ATATG GCCTACCCAT  141 Mus.mace
94  GTTAAATTAT AGACTAAAGA TC--TATATA TCTT-ACATG GCCTACCCAT  140 Mus.spre
```

```
 95  GTTAAATTAT AGACTTAAA- TC--TATATA TCTC-ACATG GCTTACCCAT  140 Rat.ever
 95  GTTAAGTTAT AGATTTAAA- TC--TATATA TCTT-ACATG GCTTACCCAT  140 Rat.exul
 95  GTTAAATCAT AGACCAAAA- TC--TATATA TCTT-ATATG GCCTACCCAT  140 Sto.long
 94  GTTAAATTAT AGACCAATAA TC--TATATA TCTT-ATATG GCCTACCCAT  140 Mus.spic
 96  GTTAAATCAT AAGCTAAACA CT--TATATA TCTT--TATG GCTTACCCGC  141 Dip.merr
 94  GTTAAATTAT AGAGTTACA- TC--TATATA TCTC-ATATG GCTTACCCAT  139 Per.mexi
 95  GTTAAATTAT AGATCAATAA TC--TATATA TCTT-ATATG GCCTACCCAT  141 Mus.musc
 93  GTTAAGTTAT AGACCTAA-- TCC-TATATA TCTT-ACATG GCTTACCCAT  138 Rat.norv
 97  GTTAAGTTAC AAGTGAAAA- TCC-TGTATA TCTC--CATG GCATATCCAT  142 Bal.musu
 95  GTTAAATTAT AGGTGGAAA- CCC-TTTATG CCTC--CATG GCATACCCCC  140 Pho.vitu
 95  GTTAAATTAT AGGTTTAA-- CTCCTATATA TCTTAATATG CCCTATCCAA  142 Did.virg
  1  ---------- ---------- ---------- -------ATG GCTTACCCTT   13 Ger.vall
  1  ---------- ---------- ---------- -------ATG GCTTATGGTT   13 Mic.penn
  1  ---------- ---------- ---------- -------ATG GCTTACCCTC   13 Per.band
  1  ---------- ---------- ---------- -------ATG GCCTATCCGT   13 Tac.sple
  1  ---------- ---------- ---------- -------ATG GCTTACCCTT   13 Apo.sylv
  1  ---------- ---------- ---------- -------ATG GCTTACCCCT   13 Aco.will
  1  ---------- ---------- ---------- -------ATG GCATACCCTT   13 Sci.caro

     151        161        171        181        191       200
     |          |          |          |          |         |
140  TTCAATTAGG CTTACAAGAT GCCACTTCTC CTATTATAGA AGAACTAATA  189 Mus.plat
141  TTCAATTAGG CTTACAAGAC GCCACATCTC CTATCATAGA AGAATTAATA  190 Mus.croc
140  TTCAACTAGG CCTTCAAGAT GCAACATCCC CAATCATAGA AGAACTCATA  189 Mus.setu
141  TCCAACTAGG CTTACAAGAC GCCACATCCC CCATTATAGA AGAACTAATA  190 Mus.caro
111  TCCAACTAGG CTTACAAGAC GCCACATCCC CCATCATAGA AGAATTAATA  160 Mus.cerv
141  TCCAACTAGG CTTACAAGAC GCCACATCCC CTATCATAGA AGAATTAATA  190 Mus.cook
142  TTCAATTAGG TCTACAAGAC GCAACATCCC CAATCATGGA AGAACTTATA  191 Mus.minu
142  TCCAACTTGG CTTACAAGAC GCCACATCCC CTATTATAGA AGAATTAATA  191 Mus.mace
141  TCCAACTAGG TTTACAAGAC GCCACATCCC CCATTATAGA AGAACTAACA  190 Mus.spre
141  TTCAACTTGG CTTACAAGAC GCCACATCAC CTATCATAGA GGAACTTACA  190 Rat.ever
141  TTCAACTTGG CTTACAAGAC GCTACATCAC CTATCATAGA AGAGCTTACA  190 Rat.exul
141  TTCAACTAGG ACTACAAGAT GCCACATCAC CTATTATAGA AGAGTTAACA  190 Sto.long
141  TCCAACTTGG CTTACAAGAC GCCACATCCC CTATTATGGA AGAATTAATA  190 Mus.spic
142  TTCAAACAGG ACTGCAAGAC GCCTCATCCC CTATTATAGA AGAACTAACC  191 Dip.merr
140  TCCAACTAGG CTTACAAGAC GCCACATCAC CTATTATAGA AGAACTAACA  189 Per.mexi
142  TCCAACTTGG TCTACAAGAC GCCACATCCC CTATTATAGA AGAGCTAATA  191 Mus.musc
139  TTCAACTTGG CTTACAAGAC GCTACATCAC CTATCATAGA AGAACTTACA  188 Rat.norv
143  TCCAACTAGG CTTCCAAGAT GCAACATCAC CTATCATAGA AGAACTCCTA  192 Bal.musu
141  TACAAATAGG CCTACAAGAT GCAACCTCTC CCATTATAGA GGAGTTACTA  190 Pho.vitu
143  TACAACTAGG TTTCCAAGAC GCTACATCTC CTATTATAGA AGAACTTATA  192 Did.virg
 14  TTCAATTAGG TTTACAAGAC GCAACATCTC CTATTATGGA AGAACTTACA   63 Ger.vall
 14  TCCAACTAGG CTTACAAGAT GCATCTTCAC CCATTATAGA AGAACTAATA   63 Mic.penn
 14  TCCAACTAGG CTTACAAGAT GCTACATCCC CAATTATAGA AGAATTAACA   63 Per.band
 14  TACAATATGG TCTTCAAGAC GCCACTTCAC CTATTATAGA AGAACTTTTA   63 Tac.sple
```

```
 14  TTCAATTAGG CTTACAAGAC GCCACATCTC CTATTATAGA AGAACTAATA   63 Apo.sylv
 14  TCCAACTAGG ACTACAAGAC GCCACATCCC CCATCATAGA AGAGCTTACA   63 Aco.will
 14  TCGAATTAGG ATTTCAAGAT GCTACTTCCC CAATTATAGA AGAACTATTA   63 Sci.caro

     201        211        221        231        241       250
     |          |          |          |          |         |
190  AATTTTCACG ACCATACACT AATAATCGTG TTCCTAATTA GTTCACTAGT  239 Mus.plat
191  AACTTTCATG ACCATACATT AATAATTGTA TTTTTAATCA GCTCTCTAGT  240 Mus.croc
190  AACTTCCATG ATCATACACT AATAATCGTG TTTTTAATTA GCTCCTTAGT  239 Mus.setu
191  AATTTCCACG ATCACACACT AATAATTGTT TTCCTAATTA GCTCACTAGT  240 Mus.caro
161  AATTTCCATG ACCACACATT AATAATTGTA TTCCTAATTA GTTCTTTAGT  210 Mus.cerv
191  AACTTTCATG ACCACACATT AATAATTGTT TTTTTAATCA GCTCCCTAGT  240 Mus.cook
192  AACTTCCATG ACCATACACT GATAATCGTG TTTTTAATTA GCTCCCTAGT  241 Mus.minu
192  AATTTCCACG ATCATACACT AATAATTGTT TTCCTAATTA GCTCCTTAGT  241 Mus.mace
191  AATTTCCACG ACCACACATT AATAATTGTC TTCCTAATTA GCTCCTTAGT  240 Mus.spre
191  AATTTTCATG ACCACACCCT AATAATTGTC TTCCTTATTA GCTCATTAGT  240 Rat.ever
191  AACTTTCATG ACCACACCCT AATAATTGTA TTTCTCATTA GCTCCCTAGT  240 Rat.exul
191  AATTTCCATG ACCACACATT AATAATTGTA TTCCTAATTA GTTCTCTAGT  240 Sto.long
191  AACTTCCACG ACCACACACT AATAATTGTT TTCCTAATTA GCTCCTTAGT  240 Mus.spic
192  AGTTTCCATG ACCACACCCT TATAATCGTT TTCCTAATCA GCACATTAGT  241 Dip.merr
190  AACTTTCATG ACCACACTCT AATAATCGTT TTCCTAATTA GCTCATTAGT  239 Per.mexi
192  AATTTCCATG ATCACACACT AATAATTGTT TTCCTAATTA GCTCCTTAGT  241 Mus.musc
189  AACTTTCATG ACCACACCCT AATAATTGTA TTCCTCATCA GCTCCCTAGT  238 Rat.norv
193  CACTTTCACG ACCACACATT AATAATTGTA TTCCTAATTA GCTCCTTAGT  242 Bal.musu
191  CACTTCCATG ACCACACATT AATAATTGTG TTCCTAATTA GCTCATTAGT  240 Pho.vitu
193  TACTTTCATG ATCATACATT AATAATTGTA TTTCTGATCA GTTCACTAGT  242 Did.virg
 64  AACTTCCATG ATCATACATT AATAATTGTA TTTCTTATTA GCTCGCTTGT  113 Ger.vall
 64  AACTTTCACG ACCATACACT TATAATCGTA TTCCTAATTA GCTCCTTAGT  113 Mic.penn
 64  AACTTTCACG ACCATACTTT AATAATTGTC TTCCTAATCA GCTCACTAGT  113 Per.band
 64  CATTTTCACG ACCATACACT TATAATTGTG TTTCTTATTA GCTTTCTAGT  113 Tac.sple
 64  AACTTTCACG ACCACACACT AATAATTGTC TTCCTAATTA GCTCATTAGT  113 Apo.sylv
 64  AACTTTCACG ACCATACACT AATAATTGTT TTCCTAATTA GCTCCTTAGT  113 Aco.will
 64  CACTTCCATG ATCATACCTT AATAATTGTA TTCCTAATTA GCTCTTTAGT  113 Sci.caro

     251        261        271        281        291       300
     |          |          |          |          |         |
240  TCTCTACATT ATTTCACTAA TGCTAACAAC AAAACTTACA CATACAAGCA  289 Mus.plat
241  ACTCTATATT ATTTCACTTA TATTAACAAC AAAACTTACA CATACAAGCA  290 Mus.croc
240  ACTATACATT ATTTCACTAA TATTAACAAC AAAACTAACT CATACAAGCA  289 Mus.setu
241  TCTTTACATC ATTTCACTAA TATTAACAAC CAAACTAACA CATACAAGCA  290 Mus.caro
211  CCTTTATATT ATTTCACTTA TATTAACAAC AAAACTCACA CACACAAGTA  260 Mus.cerv
241  CCTTTATATT ATTTCACTTA TATTGACAAC AAAACTAACA CATACAAGTA  290 Mus.cook
242  ACTATATATT ATTTCTCTAA TATTAACAAC AAAACTAACA CATACAAGCA  291 Mus.minu
242  CCTCTATATT ATCTCACTAA TGTTAACAAC AAAACTAACA CATACAAGCA  291 Mus.mace
241  CCTCTACATT ATTTCACTAA TATTAACAAC AAAACTAACA CATACAAGCA  290 Mus.spre
```

```
241  ACTATATATC ATCTCACTAA TACTAACAAC AAAACTAACA CACACAAGCA  290 Rat.ever
241  ACTATATATC ATTTCACTAA TACTAACAAC AAAATTAACA CATACAAGCA  290 Rat.exul
241  ACTTTACATT ATTTCACTCA TACTAACAAC AAAACTAACC CACACAAGTA  290 Sto.long
241  CCTCTATATC ATCTCACTAA TATTAACAAC AAAACTAACA CATACGAACA  290 Mus.spic
242  ATTATACATC ATTTCACTTA TGCTAACAAC CAAACTCACA CACACTACTA  291 Dip.merr
240  ACTATACATC ATCACATTAA TATTAACTAC AAAACTAACT CACACAAACA  289 Per.mexi
242  CCTCTATATC ATCTCGCTAA TATTAACAAC AAAACTAACA CATACAAGCA  291 Mus.musc
239  ACTTTATATT ATTTCACTAA TACTAACAAC AAAACTAACA CACACAAGCA  288 Rat.norv
243  TCTCTACATT ATCACCCTAA TACTTACAAC CAAACTAACA CATACTAGTA  292 Bal.musu
241  ACTCTACATT ATCTCACTTA TACTAACCAC GAAACTCACC CACACAAGTA  290 Pho.vitu
243  ATTATATATT ATTATTCTTA TACTTACTAC AAAACTTACT CACACAAGCA  292 Did.virg
114  ACTTTACTTA ATCTCATTAA TATTAACAAC AAAACTGATT CATACTAACA  163 Ger.vall
114  TCTTTACATC ATCACTCTCA TACTCACAAC AAAATTAACT CATACTAGCA  163 Mic.penn
114  ATTATATATT ATTACACTAA TATTAACTCA AAAACTAACT CACACAAGCA  163 Per.band
114  TCTTTACGTA ATTTTCGCTA TACTAAGTAC AAAACTTACA CACACAAGTA  163 Tac.sple
114  ACTTTACATT ATTTCACTTA TGCTAACCAC AAAACTAACT CACACAAGCA  163 Apo.sylv
114  ACTCTATATT ATTTCATCCA TATTAGCTAC CAAAATAACC CACACTAGTA  163 Aco.will
114  ATTATATATC ATTTCACTAA TGCTAACAAC TAAACTAACC CATACCAGTA  163 Sci.caro

     301        311        321        331        341       350
     |          |          |          |          |         |
290  CAATAGATGC ACAAGAAGTA GAAACTATCT GAACTATTTT ACCAGCCGTT  339 Mus.plat
291  CTATAGACGC CCAAGAAGTT GAAACTATCT GAACAATCCT ACCAGCTGTT  340 Mus.croc
290  CAATAGATGC ACAAGAAGTT GAAACTATCT GAACTATTTT ACCAGCTGTT  339 Mus.setu
291  CAATAGACGC TCAAGAGGTT GAGACCATTT GAACAATTTT ACCAGCTGTT  340 Mus.caro
261  CAATAGATGC CCAAGAAGTT GAAACCATCT GAACCATTTT ACCAGCTGTT  310 Mus.cerv
291  CAATAGATGC ACAAGAAGTT GAAACCATCT GAACCATTCT ACCAGCTGTT  340 Mus.cook
292  CTATAGACGC ACAAGAAGTT GAAACCATTT GAACCATCTT ACCAGCCGTA  341 Mus.minu
292  CAATAGATGC ACAAGAAGTT GAAACCATTT GAACTATTCT ACCAGCCGTA  341 Mus.mace
291  CAATAGATGC ACAAGAAGTT GAAACCATTT GAACCATTCT GCCAGCCGTA  340 Mus.spre
291  CAATAGACGC CCAAGAAGTA GAAACAATTT GAACAATTCT CCCAGCTGTT  340 Rat.ever
291  CAATAGACGC CCAAGAAGTA GAAACAATCT GAACAATTCT CCCAGCCGTT  340 Rat.exul
291  CAATAGACGC TCAAGAAGTA GAAACAATCT GAACTATTCT ACCAGCCGTA  340 Sto.long
291  CAATAGATGC ACAAGAAGTT GAGACCATTT GAACTATTCT ACCAGCCGTA  340 Mus.spic
292  CTATAGACGC CCAAGAAGTA GAAACTATCT GAACTATCCT CCCAGCTATA  341 Dip.merr
290  CAATAGACGC TCAAGAAGTA GAGACAATTT GGACAATCCT TCCGGCTGTT  339 Per.mexi
292  CAATAGATGC ACAAGAAGTT GAAACCATTT GAACTATTCT ACCAGCTGTA  341 Mus.musc
289  CAATAGACGC CCAAGAAGTA GAAACAATTT GAACAATTCT CCCAGCTGTC  338 Rat.norv
293  CAATAGACGC CCAAGAAGTA GAAACTGTCT GAACTATCCT CCCAGCCATC  342 Bal.musu
291  CAATAGACGC ACAAGAAGTG GAAACGGTGT GAACGATCCT ACCCGCTATC  340 Pho.vitu
293  CTATAGATGC CCAAGAAGTG GAAACAATTT GAACAATTTT ACCAGCCGTA  342 Did.virg
164  CAATAGATGC CCAAGAAGTA GAAACAGTTT GAACTATTTT ACCAGCCATT  213 Ger.vall
164  CCATAGACGC TCAAGAAGTA GAGACTATCT GAACTATCTT ACCCGCTGTT  213 Mic.penn
164  CAATAGATGC TCAAGAAGTT GAAACAATTT GAACCATCCT ACCAGCAGTA  213 Per.band
164  CTATAGACGC TCAAGAAGTA GAGACAATTT GAACTATTTT ACCTGCTGTA  213 Tac.sple
```

```
164   CAATAGACGC CCAAGAAGTA GAAACTATTT GAACCATCTT ACCAGCCGTT  213 Apo.sylv
164   CCATAGACGC CCAAAGCATA GAAACAATCT GAACTATTTT ACCAGCTGTA  213 Aco.will
164   CAATAGATGC CCAGGAAGTT GAGACTATTT GAACTATCCT ACCAGCTATT  213 Sci.caro

      351        361        371        381        391       400
      |          |          |          |          |         |
340   ATTCTTATTT TAATTGCCTT ACCATCTCTA CGTATTCTAT ATATAATAGA  389 Mus.plat
341   ATTCTTATTC TAATTGCCCT ACCATCTCTA CGCATTTTAT ATATAATAGA  390 Mus.croc
340   ATTCTTATCT TAATTGCCCT CCCATCACTT CGTATTCTTT ACATAATGGA  389 Mus.setu
341   ATCCTAATCA TAATTGCCCT GCCCTCCCTT CGCATTCTTT ACATGATAGA  390 Mus.caro
311   ATTCTCATTA TAATTGCTTT ACCTTCCCTT CGCATTCTAT ACATAATAGA  360 Mus.cerv
341   ATCCTTATCA TAATTGCCCT ACCTTCCCTT CGCATTCTAT ATATAATAGA  390 Mus.cook
342   ATTCTCATTT TAATTGCTCT CCCCTCCCTG CGTATTCTCT ACATAATAGA  391 Mus.minu
342   ATCCTTATCA TAATCGCCCT ACCCTCTTTA CGCATTCTGT ATATAATAGA  391 Mus.mace
341   ATCCTTATTA TAATTGCTCT CCCTTCTCTA CGCATTTTAT ATATGATAGA  390 Mus.spre
341   ATCCTTATTC TAATCGCCCT CCCTTCTCTA CGAATTTTAT ATATAATAGA  390 Rat.ever
341   ATCCTTGTTC TAATTGCCCT CCCCTCCCTA CGAATTTTAT ATATAATAGA  390 Rat.exul
341   ATTCTTATTT TAATTGCCCT GCCCTCTCTT CGTATTCTCT ACATAATAGA  390 Sto.long
341   ATCCTTATCA TAATTGCTCT ACCCTCTTTA CGCATTCTTT ACATAATAGA  390 Mus.spic
342   ATCCTAATTA TAATTGCTCT CCCTTCCCTA CGCATCCTCT ATATAATAGA  391 Dip.merr
340   ATCTTAATCC TCATTGCACT ACCATCACTC CGAATTCTTT ACATAATAGA  389 Per.mexi
342   ATCCTTATCA TAATTGCTCT CCCCTCTCTA CGCATTCTAT ATATAATAGA  391 Mus.musc
339   ATTCTTATTC TAATTGCCCT TCCCTCCCTA CGAATTCTAT ACATAATAGA  388 Rat.norv
343   ATCTTAATTT TAATCGCCCT ACCTTCCCTA CGAATCCTCT ATATAATAGA  392 Bal.musu
341   ATTTTAATTC TCATTGCCCT ACCATCATTA CGAATCCTCT ACATAATGGA  390 Pho.vitu
343   ATTCTTATCC TTATTGCCCT TCCTTCCTTA CGAATTCTTT ACATAATAGA  392 Did.virg
214   ATTCTTATTA TAATTGCACT TCCTTCCCTT CGCATTCTTT ATCTAATAGA  263 Ger.vall
214   ATTCTTATCC TAATTGCTCT TCCCTCCTTA CGAATCTTGT ACATGATAGA  263 Mic.penn
214   ATTTTAATTC TTATTGCGCT CCCATCTCTT CGAATCCTCT ACATAATAGA  263 Per.band
214   ATTTTAGTTC TAATCGCTCT ACCTTCACTT CGAATTCTAT ACATAATAGA  263 Tac.sple
214   ATTCTTATCT TAATTGCACT TCCATCCCTA CGAATCCTAT ATATAATAGA  263 Apo.sylv
214   ATCCTAGTTC TTATTGCACT CCCGTCCCTT CGAATCCTAT ATATAATAGA  263 Aco.will
214   ATTTTAATCT TAATCGCCTT ACCTTCACTA CGAGTTCTCT ATATGATAGA  263 Sci.caro

      401        411        421        431        441       450
      |          |          |          |          |         |
390   CGAAATTAAT AACCCAGTAT TAACAGTGAA AACAATAGGA CACCAATGAT  439 Mus.plat
391   TGAAATTAAT AATCCCGTAT TAACAGTAAA GACTATAGGC CATCAATGAT  440 Mus.croc
390   TGAAATCAAC AACCCTGTGT TAACAGTAAA AACCATAGGT CACCAATGAT  439 Mus.setu
391   TGAAATTAAT AATCCTGTCT TAACCGTAAA AACCATAGGA CATCAATGGT  440 Mus.caro
361   TGAAATTAAT AATCCTGTTT TAACCGTTAA AACTATGGGT CACCAATGAT  410 Mus.cerv
391   CGAAATTAAC AACCCCGTTC TTACCGTTAA AACTATAGGC CACCAATGAT  440 Mus.cook
392   CGAAATTAAT AACCCCTTAT TAACAGTAAA AACTATAGGT CACCAATGAT  441 Mus.minu
392   CGAAATTAAT AACCCCGTAT TAACCGTGAA AACCATAGGT CACCAATGAT  441 Mus.mace
391   CGAAATTAAT AACCCCGTCT TAACCGTGAA GACCATAGGA CATCAGTGAT  440 Mus.spre
```

```
391  CGAGATTAAT AACCCAGCTT TAACAGTAAA AACTATGGGA CACCAATGAT  440 Rat.ever
391  TGAGATTAAC AATCCAGTAT TAACAGTAAA AACAATAGGA CACCAATGAT  440 Rat.exul
391  CGAAATTAAT AACCCCACAT TAACAGTAAA GACCATAGGG CACCAATGAT  440 Sto.long
391  CGAAATTAAC AACCCTGTCT TAACCGTGAA AACTATAGGA CACCAATGAT  440 Mus.spic
392  TGAAATCAAC GACCCAGCAC TAACAGTAAA AACTATAGGA CATCAATGAT  441 Dip.merr
390  TGAGATCAAC AACCCTGTAC TAACAGTGAA AACTATAGGC CATCAATGAT  439 Per.mexi
392  CGAAATCAAC AACCCCGTAT TAACCGTTAA AACCATAGGG CACCAATGAT  441 Mus.musc
389  CGAGATTAAT AACCCAGTTC TAACAGTAAA AACTATAGGA CACCAATGAT  438 Rat.norv
393  CGAAGTCAAC AACCCCTCCC TCACTGTAAA AACAATAGGC CACCAATGAT  442 Bal.musu
391  CGAGATCAAT AACCCTTCCT TGACCGTAAA AACTATAGGA CATCAGTGAT  440 Pho.vitu
393  TGAAATCTAT AATCCTTATC TAACAGTTAA AGCAATGGGT CATCAATGAT  442 Did.virg
264  TGAAATTAAT AATCCTGTAC TAACAGTAAA AACCATAGGA CATCAATGAT  313 Ger.vall
264  CGAGATTAAC AACCCAGCCC TTACAGTAAA AACAATAGGC CACCAATGGT  313 Mic.penn
264  TGAAATTAAT AACCCTGCTT TAACAGTAAA AACCATAGGA CACCAATGAT  313 Per.band
264  CGAAATTTAC AATCCAGCAT TAACAGTAAA AACCATAGGC CATCAATGAT  313 Tac.sple
264  TGAAATTAAT AACCCAGTAC TTACAGTAAA AACCATGGGC CACCAATGAT  313 Apo.sylv
264  TGAAATTAAC AACCCAGTAT TAACCGTTAA AACAATAGGA CACCAATGAT  313 Aco.will
264  CGAAATCAAT GATCCTTCAT TAACTGTAAA AACTATAGGT CATCAATGAT  313 Sci.caro

     451        461        471        481        491       500
     |          |          |          |          |         |
440  ACTGAAGTTA TGAATATACC GACTACGAAG ACCTATGTTT TGACTCGTAC  489 Mus.plat
441  ACTGAAGCTA TGAATATACT GACTATGAAG ATCTATGCTT TGACTCTTAT  490 Mus.croc
440  ACTGAAGCTA CGAATATACC GATTATGAAG ACCTATGCTT CGATTCTTAC  489 Mus.setu
441  ACTGAAGTTA TGAGTATACT GACTATGAAG ACTTATGCTT CGATTCATAT  490 Mus.caro
411  ACTGAAGTTA TGAGTATACA GACTATGAAG ATTTATGCTT CGACTCATAT  460 Mus.cerv
441  ACTGAAGCTA TGAGTATACT GACTATGAAG ACCTATGCTT CGACTCATAC  490 Mus.cook
442  ATTGAAGCTA TGAGTATACC GATTATGAAG ATTTATGCTT TGACTCATAT  491 Mus.minu
442  ACTGAAGCTA CGAATACACT GACTATGAAG ACTTATGCTT TGATTCGTAT  491 Mus.mace
441  ACTGAAGCTA CGAGTATACT GACTATGAGG ACCTATGCTT TGACTCATAC  490 Mus.spre
441  ACTGAAGTTA CGAGTACACC GACTATGAGG ACCTGTGCTT TGACTCCTAC  490 Rat.ever
441  ACTGAAGTTA CGAATATACC GACTATGAAG ATTTATGCTT TGATTCCTAC  490 Rat.exul
441  ATTGAAGCTA CGAATACACC GATTACGAAG ACCTATGCTT TGACTCCTAT  490 Sto.long
441  ACTGAAGCTA CGAATACACC GACTATGAAG ACTTATGCTT TGATTCATAT  490 Mus.spic
442  ATTGAAGCTA TGAATACACC GACTTCAACG ATCTCATATT CGACTCCTAC  491 Dip.merr
440  ACTGAAGCTA TGAGTATACA GATTATGAAG ACCTATGTTT TGACTCATAT  489 Per.mexi
442  ACTGAAGCTA CGAATATACT GACTATGAAG ACCTATGCTT TGATTCATAT  491 Mus.musc
439  ACTGAAGCTA TGAATATACT GACTATGAAG ACCTATGCTT TGACTCCTAC  488 Rat.norv
443  ATTGAAGCTA TGAGTATACT GATTACGAAG ACCTAAGCTT TGACTCTTAC  492 Bal.musu
441  ACTGAAGCTA TGAGTACACA GACTACGAAG ACCTGAACTT TGACTCATAT  490 Pho.vitu
443  ATTGAAGCTA TGAGTTCACA GACTATGAAA ATTTAATATT CGACTCATAC  492 Did.virg
314  ACTGAAGCTA TGAGTATACC GACTATGAGG ATCTCTGTTT TGACTCGTAT  363 Ger.vall
314  ATTGAAGCTA CGAGTATACA GACTATGAAG ATCTCTGCTT CGACTCGTAC  363 Mic.penn
314  ATTGAAGCTA TGAGTATACA GACTATGAAG ACTTATGTTT CGACTCATAC  363 Per.band
314  ACTGAAGTTA CGAATACACA GATTACGAAG ACCTTAACTT TGACTCATAT  363 Tac.sple
```

```
314  ACTGAAGCTA TGAGTACACT GACTACGAAG ACTTATGCTT TGACTCCTAT  363 Apo.sylv
314  ACTGAAGCTA CGAGTATACT GATTATGAAG ATCTATGCTT TGACTCTTAT  363 Aco.will
314  ATTGAAGCTA TGAGTACACA GATTATGAAG ACCTGAACTT TGACTCCTAC  363 Sci.caro

     501        511        521        531        541       550
     |          |          |          |          |         |
490  ATGATTCCAA CAAATGACTT AAAACCAGGT GAACTTCGAT TACTAGAAGT  539 Mus.plat
491  ATAATTCCAA CAAGTGACTT AAAACCAGGC GAACTCCGAC TTCTAGAAGT  540 Mus.croc
490  ATAATTCCAA CAAATGACCT AAAACCAGGT GAACTTCGAC TTTTAGAAGT  539 Mus.setu
491  ATAATCCCAA CAAATGACTT AAAACCTGGT GAACTTCGAC TTTTAGAAGT  540 Mus.caro
461  ATGATCCCAA CAAATGACTT GAAACCAGGC GAACTTCGCC TATTAGAAGT  510 Mus.cerv
491  ATAATTCCAA CAACTGACTT AAAACCAGGT GAGCTTCGAC TACTAGAAGT  540 Mus.cook
492  ATAATTCCAA CAAATGAACT CAAACCGGGT GAACTACGAC TGTTAGAAGT  541 Mus.minu
492  ATAATCCCAA CAAGCGACTT AAAACCTGGT GAACTTCGAC TGCTAGAAGT  541 Mus.mace
491  ATAATTCCAA CAAATGACTT GAAACCTGGC GAACTACGAC TACTAGAAGT  540 Mus.spre
491  ATAATCCCAA CCAATGACCT AAAACCAGGC GAACTTCGCC TATTAGAAGT  540 Rat.ever
491  ATAATCCCAA CCAATGATCT GAAACCAGGT GAGCTTCGCC TCTTAGAAGT  540 Rat.exul
491  ATGATCCCAA CAACCGACCT AAAACCAGGG GAACTACGTC TCCTAGAAGT  540 Sto.long
491  ATAATCCCAA CAAACGACTT GAAACCCGGC GAACTACGAC TGCTAGAAGT  540 Mus.spic
492  ATAATTGCTA CATCAGACTT AAAACCTGGC GAACTACGTC TGCTAGAAGT  541 Dip.merr
490  ATGATCCCAA CAAATGACTT AAAACCAGGA GGTCTACGTC TACTAGAAGT  539 Per.mexi
492  ATAATCCCAA CAAACGACCT AAAACCTGGT GAACTACGAC TGCTAGAAGT  541 Mus.musc
489  ATAATCCCAA CCAATGACCT AAAACCAGGT GAACTTCGTC TATTAGAAGT  538 Rat.norv
493  ATAATCCCAA CGTCAGACCT AAAGCCAGGA GAATTACGAC TACTAGAAGT  542 Bal.musu
491  ATGATCCCCA CACAAGAACT AAAGCCCGGA GAACTACGAC TGCTAGAAGT  540 Pho.vitu
493  ATAATCCCAA CCAAAGACCT TAGTCCTGGG CAACTTCGTT TACTAGAAGT  542 Did.virg
364  ATAACTCCAA CAAACGAACT AAAACCAGGA GAACTTCGCT TATTAGAAGT  413 Ger.vall
364  ATGATTCCAA CCAATGACTT AAAACCCGGA GAACTTCGCC TCCTAGAAGT  413 Mic.penn
364  ATAATTCCAA CAAATGATTT AAAACCGGGT GAACTACGCC TATTAGAAGT  413 Per.band
364  ATAATCCCAA CAACCGAGCT AAAGCCAGGA GAACTACGAC TTCTAGAAGT  413 Tac.sple
364  ATAATTCCAA CAAGTGACCT TAAACCAGGT GAATTACGAC TACTAGAAGT  413 Apo.sylv
364  ATAGTACCAA CAAACGACTT AAAACCAGGA GAACTTCGTC TATTAGAAGT  413 Aco.will
364  ATAATTCCAA CATCTGAACT AAAGCCAGGT GAGCTGCGTC TTCTAGAAGT  413 Sci.caro

     551        561        571        581        591       600
     |          |          |          |          |         |
540  AGATAATCGA GTTGTTCTAC CCATAGAACT TCCAATCCGT ATACTAATTT  589 Mus.plat
541  TGATAATCGA GTTGTTCTAC CAATAGAACT TCCAATCCGC ATACTAATCT  590 Mus.croc
540  AGATAACCGA GTAGTTCTAC CAATAGAACT TCCAATCCGT ATATTAATTT  589 Mus.setu
541  TGACAATCGA GTCGTCCTAC CAATGGAACT CCCAATCCGT ATATTAATTT  590 Mus.caro
511  TGATAATCGA GTCGTCCTAC CTATAGAACT TCCAATCCGT ATATTAATTT  560 Mus.cerv
541  TGACAATCGA GTTGTTCTAC CAATGGAACT TCCAATTCGT ATATTAATCT  590 Mus.cook
542  TGACAACCGA GTTGTCTTAC CAATAGAGCT TCCAATTCGT ATATTAATCT  591 Mus.minu
542  TGATAACCGA GTCGTCCTAC CAATAGAACT CCCAATCCGT ATGCTAATCT  591 Mus.mace
541  TGACAACCGA GTCGTCCTAC CAATAGAACT TCCAATCCGT ATGCTAATCT  590 Mus.spre
```

```
541   TGATAACCGA GTAGTCTTAC CGATAGAACT TCCAATTCGT ATATTAATTT  590 Rat.ever
541   CGACAACCGA GTAGTCCTGC CAATAGAACT TCCAATTCGT ATATTAATCT  590 Rat.exul
541   CGACAATCGA GTAGTTCTAC CCATAGAACT TCCAATTCGC ATATTAATCT  590 Sto.long
541   TGACAACCGA GTTGTCCTAC CAATAGAACT TCCAATCCGT ATATTAATTT  590 Mus.spic
542   GGATAATCGA GTTGTTCTTC CTATAGAACT ACCTATTCGA ATATTAATCT  591 Dip.merr
540   AGATAACCGA GTTGTCCTAC CAATAGAATT ACCAATCCGC ATACTAATTT  589 Per.mexi
542   TGATAACCGA GTCGTTCTGC CAATAGAACT TCCAATCCGT ATATTAATTT  591 Mus.musc
539   TGATAATCGG GTAGTCTTAC CAATAGAACT TCCAATTCGT ATACTAATCT  588 Rat.norv
543   AGACAACCGA GTTGTCCTAC CCATAGAAAT AACAATCCGA ATGCTAGTCT  592 Bal.musu
541   AGACAATCGA GTAGTCCTCC CAATAGAAAT AACAATCCGC ATACTAATCT  590 Pho.vitu
543   TGATAACCGA ATTGTTCTCC CAATAGAACT ACCAATTCGC ATGCTAATTT  592 Did.virg
414   AGACAACCGA GTAGTTCTCC CAATAGAACT TCCTATTCGT ATATTAATTT  463 Ger.vall
414   AGATAACCGA GTAGTCCTAC CCATAGAATT ACCAATCCGC ATACTAATCT  463 Mic.penn
414   AGATAACCGA GTTGTATTAC CAATAGAACT ACCAATTCGT ATATTAATTT  463 Per.band
414   AGATAATCGC GTAGTTCTAC CAATAGAACT ACCAATTCGC ATACTAGTCT  463 Tac.sple
414   CGACAACCGA GTAGTCTTAC CAATAGAACT CCCAATTCGC ATGCTTATCT  463 Apo.sylv
414   AGATAACCGA GTAGTATTAC CCATAGAACT CCCTATTCGT ATATTAATTT  463 Aco.will
414   TGATAATCGA GTCGTTCTCC CCATAGAGCT ACCTATCCGC ATGCTAATCT  463 Sci.caro

      601        611        621        631        641       650
      |          |          |          |          |         |
590   CATCCGAAGA TGTACTTCAC TCATGAGCGG TTCCTTCATT AGGATTAAAA  639 Mus.plat
591   CATCTGAAGA TGTTCTTCAC TCATGAGCCG TCCCTTCTCT AGGATTAAAA  640 Mus.croc
590   CATCAGAAGA CGTACTTCAC TCATGAGCCG TTCCTTCATT AGGACTAAAA  639 Mus.setu
591   CATCTGAAGA TGTACTTCAC TCATGAGCCG TCCCTTCACT AGGATTAAAA  640 Mus.caro
561   CATCTGAAGA TGTTCTTCAT TCATGAGCTG TTCCTTCACT TGGACTAAAA  610 Mus.cerv
591   CATCCGAAGA CGTACTTCAC TCATGGGCTG TCCCATCATT AGGACTAAAA  640 Mus.cook
592   CATCTGAAGA TGTCCTCCAC TCATGAGCCG TCCCTTCATT AGGACTAAAA  641 Mus.minu
592   CATCTGAAGA CGTCCTTCAC TCATGAGCCG TACCCTCACT AGGACTTAAA  641 Mus.mace
591   CATCCGAAGA TGTCCTCCAT TCATGAGCTG TGCCCTCACT AGGACTAAAA  640 Mus.spre
591   CATCCGAAGA CGTATTACAC TCATGAGCTG TCCCTTCACT AGGATTAAAG  640 Rat.ever
591   CATCCGAAGA CGTCCTACAC TCATGAGCTG TTCCTTCACT AGGATTAAAA  640 Rat.exul
591   CATCCGAAGA TGTACTTCAC TCATGGGCTG TTCCCTCATT AGGATTAAAA  640 Sto.long
591   CACCTGAAGA CGTCCTCCAC TCATGAGCCG TCCCCTCACT AGGACTTAAA  640 Mus.spic
592   CTTCAGAAGA CGTACTACAC TCATGAGCAG TACCTTCATT AGGAGTTAAA  641 Dip.merr
590   CATCAGAAGA CGCACTTCAC TCATGAGCCG TTCCATCCCT AGGACTAAAA  639 Per.mexi
592   CATCTGAAGA CGTCCTCCAC TCATGAGCAG TCCCCTCCCT AGGACTTAAA  641 Mus.musc
589   CATCCGAAGA CGTCCTGCAC TCATGAGCCA TCCCTTCACT AGGGTTAAAA  638 Rat.norv
593   CATCAGAAGA CGTACTCCAC TCATGAGCCG TACCCTCTCT AGGCCTAAAA  642 Bal.musu
591   CATCAGAAGA TGTACTCCAC TCATGAGCCG TACCGTCCCT AGGACTAAAA  640 Pho.vitu
593   CATCAGAAGA CGTTCTCCAT GCATGAACAA TGCCATCATT AGGCTTAAAA  642 Did.virg
464   CATCAGAAGA TGTACTACAT TCATGAGCTG TTCCCTCCTT AGGATTAAAA  513 Ger.vall
464   CATCCGAAGA CGTACTTCAC TCATGAGCCG TTCCCTCATT AGGACTAAAA  513 Mic.penn
464   CATCAGAAGA CGTACTTCAC TCATGAGCCG TACCCTCTCT AGGACTAAAA  513 Per.band
464   CCTCAGAAGA TGTTCTACAC TCCTGAGCAG TTCCATCTCT TGGCTTAAAA  513 Tac.sple
```

```
464  CATCTGAAGA TGTCCTCCAC TCATGAGCCG TACCTTCACT AGGACTAAAA  513 Apo.sylv
464  CATCCGAAGA TGTTCTACAC TCATGAGCTG TACCTTCATT AGGGTTAAAA  513 Aco.will
464  CATCTGAAGA CGTATTACAC TCTTGGGCTG TTCCATCACT TGGCCTAAAA  513 Sci.caro

     651        661        671        681        691      700
     |          |          |          |          |        |
640  ACTGATGCAA TTCCAGGACG ATTGAACCAA GCAACAGTTA CCTCTAACCG  689 Mus.plat
641  ACTGACGCAA TTCCAGGACG ATTAAACCAA GCAACAGTAA CATCAAATCG  690 Mus.croc
640  ACTGACGCCA TCCCAGGACG ATTAAATCAA GCAACAGTTA CATCAAACCG  689 Mus.setu
641  ACTGACGCTA TTCCAGGTCG ATTAAACCAA GCAACAGTAA CATCAAACCG  690 Mus.caro
611  ACCGACGCCA TCCCAGGACG ACTAAATCAA GCAACAGTAA CATCAAACCG  660 Mus.cerv
641  ACCGACGCCA TCCCAGGTCG ACTAAACCAA GCAACAGTAA CATCAAACCG  690 Mus.cook
642  ACCGATGCAA TTCCAGGACG ATTAAACCAA GCAACAGTGA CATCAAACCG  691 Mus.minu
642  ACTGACGCCA TCCCAGGCCG ACTAAATCAA GCAACAGTAA CATCAAACCG  691 Mus.mace
641  ACCGACGCCA TCCCAGGCCG ACTGAATCAA GCAACAGTAA CATCAAACCG  690 Mus.spre
641  ACCGATGCAA TCCCAGGCCG CCTAAATCAA GCCACAGTAA CATCAAACCG  690 Rat.ever
641  ACCGACGCAA TCCCAGGTCG CCTAAACCAA GCTACAGTAA CATCAAACCG  690 Rat.exul
641  ACAGATGCAA TCCCAGGACG CTTAAACCAA GCCACAGTAA CATCAAATCG  690 Sto.long
641  ACTGACGCTA TCCCAGGCCG ACTAAATCAA GCAACAGTAA CATCAAACCG  690 Mus.spic
642  ACTGACGCAA TCCCAGGACG ACTTAACCAA GCAACCCTAT CATCCTCACG  691 Dip.merr
640  ACAGATGCTA TCCCAGGACG CCTAAATCAG GCCACAATCT CATCCAACCG  689 Per.mexi
642  ACTGATGCCA TCCCAGGCCG ACTAAATCAA GCAACAGTAA CATCAAACCG  691 Mus.musc
639  ACCGACGCAA TCCCCGGCCG CCTAAACCAA GCTACAGTCA CATCAAACCG  688 Rat.norv
643  ACGGACGCAA TCCCAGGACG TCTAAACCAA ACAACCTTAA TATCAACACG  692 Bal.musu
641  ACTGATGCTA TCCCAGGACG ACTAAACCAA ACAACCCTAA TAACCATACG  690 Pho.vitu
643  GCAGATGCTA TTCCAGGGCG ATTAAACCAA ATTACCTTAA CATCATCCCG  692 Did.virg
514  ACAGATGCCA TCCCTGGACG TCTAAATCAA GCCACAATTA CATCAAATCG  563 Ger.vall
514  ACAGACGCTA TTCCTGGGCG ACTAAACCAA GCAACCATTT CATCCAACCG  563 Mic.penn
514  ACAGATGCTA TTCCAGGCCG TCTAAATCAG GCCACAATCT CATCTAATCG  563 Per.band
514  ACAGATACTA TCCCTGGACG GCTAAATCAA GCAACTTTAT CATCTACTCG  563 Tac.sple
514  ACCGACGCAA TTCCAGGACG CCTAAATCAA GCCACAGTCA CATCGAACCG  563 Apo.sylv
514  ACAGATGCTA TTCCCGGACG CCTCAATCAA GCTACTGTCT CGTCCAACCG  563 Aco.will
514  ACCGATGCTA TTCCTGGCCG ATTAAATCAA GCAACACTAA CATCTACCCG  563 Sci.caro

     701        711        721        731        741      750
     |          |          |          |          |        |
690  ACCAGGACTA TTTTATGGCC AATGTTCTGA AATTTGCGGC TCTAACCATA  739 Mus.plat
691  ACCAGGACTA TTTTACGGCC AATGCTCTGA AATTTGCGGG TCAAATCACA  740 Mus.croc
690  ACCAGGACTA TTTTATGGTC AATGCTCCGA AATTTGTGGC TCTAATCACA  739 Mus.setu
691  ACCAGGTTTA TTTTATGGTC AATGCTCCGA AATTTGCGGC TCCAATCATA  740 Mus.caro
661  ACCAGGGTTA TTCTATGGTC AATGCTCTGA AATTTGCGGC TCCAATCATA  710 Mus.cerv
691  ACCAGGATTA TTCTACGGCC AATGCTCTGA AATTTGCGGC TCTAACCATA  740 Mus.cook
692  ACCAGGATTG TTTTATGGTC AATGCTCTGA AATCTGTGGC TCTAATCACA  741 Mus.minu
692  ACCAGGATTA TTCTATGGCC AATGCTCTGA AATTTGTGGA TCTAACCATA  741 Mus.mace
691  ACCAGGATTA TTCTACGGCC AATGCTCTGA AATTTGCGGA TCTAACCATA  740 Mus.spre
```

```
691  ACCTGGTCTA TTCTACGGAC AGTGCTCCGA AATCTGCGGC TCAAATCACA  740 Rat.ever
691  ACCTGGCTTA TTCTATGGCC AATGCTCTGA AATCTGCGGC TCAAATCACA  740 Rat.exul
691  ACCAGGCTTA TTCTATGGTC AATGCTCTGA GATTTGTGGC TCTAACCACA  740 Sto.long
691  ACCAGGATTA TTCTATGGCC AATGCTCCGA AATTTGTGGA TCTAACCATA  740 Mus.spic
692  ACCAGGACTG TTCTATGGAC AATGCTCTGA AATCTGCGGC TCCAATCACA  741 Dip.merr
690  ACCAGGGTTA TTTTACGGTC AATGCTCTGA AATTTGTGGA TCTAACCACA  739 Per.mexi
692  ACCAGGGTTA TTCTATGGCC AATGCTCTGA AATTTGTGGA TCTAACCATA  741 Mus.musc
689  ACCAGGTCTA TTCTATGGCC AATGCTCTGA AATTTGCGGC TCAAATCACA  738 Rat.norv
693  ACCAGGCCTA TTCTACGGAC AATGCTCAGA AATCTGTGGC TCAAACCATA  742 Bal.musu
691  ACCAGGACTG TACTACGGTC AATGCTCAGA AATCTGTGGT TCAAACCACA  740 Pho.vitu
693  ACCAGGGGTG TTTTATGGTC AATGTTCAGA AATCTGTGGT TCAAACCACA  742 Did.virg
564  CCCAGGAGTG TTTTATGGGC AATGCTCAGA AATCTGTGGG TCAAATCATA  613 Ger.vall
564  TCCCGGATTG TTCTACGGCC AATGTTCAGA AATCTGTGGG TCTAACCATA  613 Mic.penn
564  ACCTGGACTA TACTACGGAC AATGTTCTGA AATCTGTGGC TCAAATCACA  613 Per.band
564  CCCAGGGTTA TTCTACGGAC AATGCTCAGA AATCTGCGGG TCAAATCATA  613 Tac.sple
564  TCCAGGCTTA TTCTATGGAC AATGCTCTGA AATCTGTGGC TCTAACCACA  613 Apo.sylv
564  TCCAGGCCTG TTCTACGGAC AATGCTCAGA AATTTGTGGC TCAAATCACA  613 Aco.will
564  ACCCGGCTTA TATTACGGTC AATGCTCCGA AATCTGTGGC TCTAACCACA  613 Sci.caro

     751        761        771        781        791       800
     |          |          |          |          |         |
740  GTTTTATACC TATTGTACTT GAAATAGTTC CACTAAAATA TTTTGAAAAC  789 Mus.plat
741  GTTTCATACC CATTGTTCTT GAAATAGTAC CTCTAAAATA TTTTGAAAAT  790 Mus.croc
740  GCTTTATACC TATTGTTCTT GAAATAGTCC CATTAAAACA TTTTGAAA    787 Mus.setu
741  GTTTCATGCC TATCGTCCTT GAAATAGTTC CACTAAAATA TTTCGAAAA   789 Mus.caro
711  GTTTCATGCC TATTGTCCTT GAAATAGTCC CACTAAAATA TTTCGAAAA   759 Mus.cerv
741  GCTTTATACC TATCGTCCTT GAAATAGTCC CACTAAAGTA CTTCGAAAA   789 Mus.cook
742  GCTTCATGCC TATTGTTCTT GAAATAGTTC CAYTAAAACA TTTTGAAAA   790 Mus.minu
742  GCTTTATGCC CATTGTCCTT GAAATGGTTC CACTAAAGTA TTTCGAAAA   790 Mus.mace
741  GCTTTATACC CATTGTCCTT GAAATGGTTC CACTAAAATA TTTTGAAAA   789 Mus.spre
741  GCTTCATGCC TATCGTATTA GAAATAGTAC CCCTAAAATA CTTCGAAAA   789 Rat.ever
741  GCTTTATACC CATCGTACTA GAAATAGTAC CCCTAAAACA CTTCGAAAA   789 Rat.exul
741  GTTTTATACC AATTGTCCTT GAAATAGTAC CCCTAAAATA CTTTGAAAA   789 Sto.long
741  GCTTTATGCC CATTGTCCTT GAAATAGTTC CACTAAAATA TTTCGAAAA   789 Mus.spic
742  GCTTTATACC CATTGTCCTT GAAATAGTAC CACTCAAATA TTTCGAA     788 Dip.merr
740  GCTTCATACC CATTGTACTT GAAATAGTAC CACTAAAACA TTTTGAAAA   788 Per.mexi
742  GCTTTATGCC CATTGTCCTA GAAATGGTTC CACTAAAATA TTTCGAAAAC  791 Mus.musc
739  GCTTCATACC CATTGTACTA GAAATAGTGC CTCTAAAATA TTTCGAAAAC  788 Rat.norv
743  GTTTTATACC AATCGTCCTA GAACTAGTGC CCTTAGAATT CTTTGAAAAA  792 Bal.musu
741  GCTTCATACC TATTGTCCTC GAATTGGTCC CACTATCCCA CTTCGAGAAA  790 Pho.vitu
743  GCTTTATGCC TATTGTCCTA GAAATAGCCT CACTAAAATA TTTCGAGAAA  792 Did.virg
614  GCTTTATACC TATTGTCCTT GAAATAATTC CACTAAAACT TTTTGAAAAT  663 Ger.vall
614  GCTTTATGCC TATTGTACTT GAAATAGTTC CCCTAAAAAA CTTCGAGGAC  663 Mic.penn
614  GTTTTATACC TATCGTACTT GAAATAGTGC CACTAAAACA CTTTGAAAAC  663 Per.band
614  GCTTTATACC AATTGTCTTA GAAATAGTCC CACTAAAACA ATTTGAAAGC  663 Tac.sple
```

```
614  GCTTTATACC TATTGTCCTC GAAATAGTAC CACTAAAACA CTTCGAAAAT  663 Apo.sylv
614  GTTTCATACC AATCGTACTT GAAATAGTAC CATTAAAATT CTTCGAAAAC  663 Aco.will
614  GCTTTATGCC TATTGTCCTT GAAATAGTTC CACTAAAACA TTTCGAAAAC  663 Sci.caro

791  TGATCAGCTT CAATAAT                                      807 Mus.croc
792  TGATCTGCTT CAATAATTTA A                                 812 Mus.musc
789  TGATCAGCTT CTATAATTTA A                                 809 Rat.norv
793  TGATCTGCAT CAATACTATA A                                 813 Bal.musu
791  TGATCTACCT CAATGCTTTA A                                 811 Pho.vitu
793  TGATCTTCTA TAATGCAATC A                                 813 Did.virg
664  TGATCAACAT CAATAATTTA A                                 684 Ger.vall
664  TGATCTCTAT CAATAATCTA A                                 684 Mic.penn
664  TGATCTTCAT CAATAATTCA A                                 684 Per.band
664  TGATCAGCTA CAATAATCTA A                                 684 Tac.sple
664  TGATCAGCCT CAATAATCTA A                                 684 Apo.sylv
664  TGATCAGCTT CTATAATCTA A                                 684 Aco.will
664  TGATCTTCAT CAATATTATA A                                 684 Sci.caro
```

Appendix B

Taxa used for 12S rRNA sequences

Taxa used and their ten character code

1Cra.casta	Cratogeomys castanops
1Cry.hotte	Cryptomys hottentotus
1Gli.japon	Glirulus japonicus
1Hyd.hydro	Hydrochaeris hydrochaeris
1Mus.avell	Muscardinus avellanarius
1Ped.capen	Pedetes capensis
1Per.leuco	Peromyscus leucopus
1Pet.leuco	Petaurista leucogenys
1Pet.petau	Petaurista petaurista
1Pte.volan	Pteromys volans
1Tam.sibri	Tamias sibricus
2Cra.casta	Cratogeomys castanops
2Cry.hotte	Cryptomys hottentotus
2Gli.japon	Glirulus japonicus
2Hyd.hydro	Hydrochaeris hydrochaeris
2Mus.avell	Muscardinus avellanarius
2Ped.capen	Pedetes capensis

2Per.leuco	Peromyscus leucopus
2Pet.leuco	Petaurista leucogenys
2Pet.petau	Petaurista petaurista
2Pte.volan	Pteromys volans
2Tam.sibri	Tamias sibricus
3Ped.capen	Pedetes capensis
3Pte.volan	Pteromys volans
Aco.cahiri	Acomys cahirinus
Apl.rufa	Aplodontia rufa
Ath.macrou	Atherurus macrourus
Bas.astutu	Bassaricus astutus
Bat.janett	Bathyergus janetta
Bat.suillu	Bathyergus suillus
Can.latran	Canis latrans
Cap.pilori	Capromys pilorides
Cas.canade	Castor canadensis
Cav.guiana	Cavia guianae
Cav.porcel	Cavia porcellus
Chi.lanige	Chinchilla laniger
Cri.gambia	Cricetomys gambianus
Cri.migrat	Cricetulus migratorius
Cry.damare	Cryptomys damarensis
Cte.bolivi	Ctenomys boliviensis
Cte.gundi	Ctenodactylus gundi
Das.novemc	Dasypus novemcinctus
Das.puncta	Dasyprocta punctata
Dip.ordii	Dipodomys ordii
Dol.patago	Dolichotis patagonum
Dry.nitedu	Dryomys nitedula
Ere.dorsat	Erethizon dorsatum
Geo.bursar	Geomys bursarius
Geo.capens	Georychus capensis
Ger.nigeri	Gerbillus nigeriae
Ger.vallia	Gerbillurus vallianus
Gli.glis	Glis glis

Gra.murinu	Graphiurus murinus
Hel.argent	Heliophobius argenteocinereus
Het.glaber	Heterocephalus glaber
Hyl.stella	Hylomyscus stella
Hys.africa	Hystrix africaeaustralis
Jac.jaculu	Jaculus jaculus
Leo.edwars	Leopoldamys edwarsi
Lop.flavop	Lophuromys flavopunctatus
Mac.ingens	Macrotarsomys ingens
Mas.erythr	Mastomys erythroleucus
Mep.mephit	Mephitis mephitis
Mes.auratu	Mesocricetus auratus
Mic.nivali	Microtus nivalis
Mus.cookii	Mus cookii
Mus.crocid	Mus crociduroides
Mus.matthe	Mus mattheyi
Mus.muscul	Mus musculus
Mus.pahari	Mus pahari
Mus.platyt	Mus platythrix
Mus.saxico	Mus saxicola
Mus.setulo	Mus setulosus
Myd.marche	Mydaus marchei
Myo.coypus	Myocastor coypus
Neo.albigu	Neotoma albigula
Neo.florid	Neotoma floridana
Nes.rufus	Nesomys rufus
Oct.degus	Octodon degus
Odo.rosmar	Odobenus rosmarus
Ony.arenic	Onychomys arenicola
Ony.leucog	Onychomys leucogaster
Ony.torrid	Onychomys torridus
Osg.bander	Osgoodomys banderanus
Per.eremic	Peromyscus eremicus
Per.flavus	Perognathus flavus
Per.gossyp	Peromyscus gossypinus

Per.manicu	Peromyscus maniculatus
Per.melano	Peromyscus melanotis
Per.polion	Peromyscus polionotus
Pet.typicu	Petromus typicus
Pro.longic	Proechimys longicaudatus
Pro.lotor	Procyon lotor
Rat.norveg	Rattus norvegicus
Sci.lis	Sciurus lis
Sci.niger	Sciurus niger
Sci.vulgar	Sciurus vulgaris
Sig.hispid	Sigmodon hispidus
Sig.mascot	Sigmodon mascotensis
Spe.tridec	Spermophilus tridecemlineatus
Spi.putori	Spilogale putorius
Syl.audubo	Sylvilagus audubonii
Tat.kempi	Tatera gkempi ambiana
Thr.swinde	Thryonomys swinderianus
Ura.ruddi	Uranomys ruddi
Uro.cinere	Urocyon cinereoargenteus
Urs.americ	Ursus americanus

Appendix C

Aligned 12S rRNA sequence data

Species names are listed in Appendix B.

```
     1          11         21         31         41        50
     |          |          |          |          |         |
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAAG ATTACACATG    50 Rat.norv
  1  AAAAGGTTTG GTCCTGGCCT TATAATTAGT TGGAGGTAAG ATTACACATG    50 Hyl.stel
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAAA ATTACACATG    50 Mus.croc
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAAA ATTACACATG    50 Mus.cook
  1  AAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAAG ATTACACATG    50 Mas.eryt
  1  TAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAGA ATTACACATG    50 Mus.matt
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAGT TGGAGGTAAA ATTACACATG    50 Mus.plat
  1  TAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAAA ATTACACATG    50 Mus.setu
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAGT TGGAGGTAAA ATTACACATG    50 Mus.saxi
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAGT TGGAAGTAAA ATTACACATG    50 Cri.gamb
  1  AAAAGGTTTG GTCCTGGCCT TATAGTTAGT TAAAGGAAAA ATTACACATG    50 Mac.inge
  1  AAAAGGTTTG GTCCTGGCCT TATAATTAGT TGGAGGCAGA ATTACACATG    50 Mic.niva
  1  CAGAGGTTTG GTCCTAGCCT TATGGTTAAT TAGAGGAAGA ATTACACATG    50 Nes.rufu
  1  AAAAGGTTTG GTCCTGGCCT TATAATTAGT TAGAGGTAAA ATTACACATG    50 1Per.leu
  1  AATAGGTTTG GTCCTAGCCT TATAATTAGC TGAAAGTAAA ATTACACATG    50 Aco.cahi
  1  TAAAGGTTTG GTCCTGGCCT TATAATTAAT TGAAGGTATG GTTACACATG    50 Cri.migr
  1  CAAAGGTTTG GTCCCAGCCT TCTTATTAGT TTATAGCAAG ATTACACATG    50 Gli.glis
  1  TAAAGGTTTG GTCCTGGCCT TGTAGTTATT TCTAGGTAAG TTTACACATG    50 Ger.nige
  1  CAAAGGTTTG GTCCTGGCCT TATAATTAGT TGGAGGTAAG ATTACACATG    50 Leo.edwa
  1  AAAAGGTTTG GTCCTAGCCT TATAGTTAGT TAGAGGTAGA GTTACACATG    50 Mes.aura
```

```
1  TAAAGGTTTG GTCCCAGCCT TCTTATTAAT TTATAGCAGG ATTACACATG   50 1Mus.ave
1  CAAAGGTTTG GTCCTGGCCT TATAATTAAT TAGAGGTAAA ATTACACATG   50 Mus.musc
1  CAAAGGTTTG GTCCTGGCCT TATAATTAAT TGGAGGTAAA ATTACACATG   50 Mus.paha
1  CAAAGGTTTG GTCCTGGCCT TATAGTTAAT TGAAGGTGAG ATTACACATG   50 Tat.kemp
1  TAAAGGTTTG GTCCTGGCCT TATAGTTAAT TAGGGGTAGA ATTACACATG   50 Ura.rudd
1  CAAAGGTTTG GTCCTGGCCT TCTTATTAGT CTATAACAAG ATTACACATG   50 Dry.nite
1  TAAAGGTTTG GTCCCAGCCT TCTTATTAAT TTATAACAAG ATTACACATG   50 2Mus.ave
1  TAAAGGTTTG GTCCTAGCCT TCTTATTAAT TTATAGTAAG ATTACACATG   50 1Gli.jap
1  TAAAGGTTTG GGCCTGGCCT TCTTATTGAT TAATAGTAAA ATTACACATG   50 2Gli.jap
1  AACAGGTTTG GTCCTGGCCT TTTTATTAGT TGCTTACAGA ATTATACATG   50 1Hyd.hyd
1  ---------- ----CGGCCN TTTTATTAGC TGTCTGCAGG ATTATACATG   36 Cav.porc
1  ---------- ----TAGCTT TTTTATTAGT TATAAGCAGA ATTACACATG   36 Bat.suil
1  ---------- ----TAGCTT TTTTATTAAT TGTAAGCAGA ATTACACATG   36 Bat.jane
1  ---------- ----TAGCTT TTTTATTAGT TGTATGCAAA TTTACACATG   36 1Cry.hot
1  ---------- ----TAGCTT TTTTATTAGT TGTATGCAAA TTTACACATG   36 2Cry.hot
1  ---------- ----TGGCTT TTTTATTAGT TGTAAGCAGA ATTACACATG   36 Cry.dama
1  ---------- ----TGGCTT TTTTATTAAT TGTAAGCAAA ATTACACATG   36 Geo.cape
1  ---------- ----TAGCTT TTTTATTAGT TTAATGCGAA ATTACACATG   36 Hel.arge
1  ---------- ----TGGCTT TTTTATTAGT TTAGCGCAAA ATTATACATG   36 Het.glab
1  ---------- ----TGGCTT TTTTATTAGT TGTTAGTAGC GTTATACATG   36 Thr.swin
1  ---------- ---------- ---------- -----GTAAA ATTACACATG   15 2Cra.cas
1  ---------- ---------- ---------- -----GTAAA TTTACACATG   15 Dip.ordi
1  ---------- ---------- ---------- -----GCAAA ACTACACATG   15 1Ped.cap
1  ---------- ---------- ---------- -----GCAAA ACTACACATG   15 2Ped.cap
1  ---------- ---------- ---------- ---------- ---------G    1 Ath.macr
1  ---------- ---------- ---------- ---------- ---------G    1 Cte.boli
1  ---------- ---------- ---------- ---------- ---------G    1 Cav.guia
1  ---------- ---------- ---------- ---------- ---------G    1 Chi.lani
1  ---------- ---------- ---------- ---------- ---------G    1 Cap.pilo
1  ---------- ---------- ---------- ---------- ---------G    1 Das.punc
1  ---------- ---------- ---------- ---------- ---------G    1 Ere.dors
1  ---------- ---------- ---------- ---------- ---------G    1 Hys.afri
1  ---------- ---------- ---------- ---------- ---------G    1 Myo.coyp
1  ---------- ---------- ---------- ---------- ---------G    1 Oct.degu
1  ---------- ---------- ---------- ---------- ---------G    1 Pro.long
1  ---------- ---------- ---------- ---------- ---------G    1 Apl.rufa
1  ---------- ---------- ---------- ---------- ---------G    1 Cas.cana
1  ---------- ---------- ---------- ---------- ---------G    1 1Cra.cas
1  ---------- ---------- ---------- ---------- ---------G    1 Cte.gund
1  ---------- ---------- ---------- ---------- ---------G    1 Das.nove
1  ---------- ---------- ---------- ---------- ---------G    1 Geo.burs
1  ---------- ---------- ---------- ---------- ---------G    1 Ger.vall
1  ---------- ---------- ---------- ---------- ---------G    1 Jac.jacu
1  ---------- ---------- ---------- ---------- ---------G    1 Lop.flav
1  ---------- ---------- ---------- ---------- ---------G    1 Osg.band
```

```
1 ---------- ---------- ---------- ---------- ---------G   1 3Ped.cap
1 ---------- ---------- ---------- ---------- ---------G   1 Per.flav
1 ---------- ---------- ---------- ---------- ---------G   1 Syl.audu
1 ---------- ---------- ---------- ---------- ---------G   1 Sci.nige
1 ---------- ---------- ---------- ---------- ---------G   1 Spe.trid
1 ---------- ---------- ---------- ---------- ---------G   1 Gra.muri
1 ---------- ---------- ---------- ---------- ---------G   1 2Hyd.hyd
1 ---------- ---------- ---------- ---------- ---------G   1 Pet.typi
1 ---------- ---------- ---------- ---------- -ATACACGTG   9 Ony.leuc
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Per.poli
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Neo.albi
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Neo.flor
1 ---------- ---------- ---------- ---------- -ATATATATG   9 Ony.torr
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Per.erem
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Per.goss
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Per.mani
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Per.mela
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Sig.hisp
1 ---------- ---------- ---------- ---------- -TTACACATG   9 Sig.masc
1 ---------- ---------- ---------- ---------- -ATACACATG   9 Ony.aren
1 ---------- ---------- ---------- ---------- -TTACACATG   9 2Per.leu
1 ---------- ---------- ---------- ---------- ----------   0 Dol.pata
1 ---------- ---------- ---------- ---------- ----------   0 1Pet.leu
1 ---------- ---------- ---------- ---------- ----------   0 2Pet.leu
1 ---------- ---------- ---------- ---------- ----------   0 1Pet.pet
1 ---------- ---------- ---------- ---------- ----------   0 2Pet.pet
1 ---------- ---------- ---------- ---------- ----------   0 1Pte.vol
1 ---------- ---------- ---------- ---------- ----------   0 2Pte.vol
1 ---------- ---------- ---------- ---------- ----------   0 3Pte.vol
1 ---------- ---------- ---------- ---------- ----------   0  Sci.lis
1 ---------- ---------- ---------- ---------- ----------   0 Sci.vulg
1 ---------- ---------- ---------- ---------- ----------   0 1Tam.sib
1 ---------- ---------- ---------- ---------- ----------   0 2Tam.sib
1 ---------- ---------- ---------- ---------- ----------   0 Bas.astu
1 ---------- ---------- ---------- ---------- ----------   0 Spi.puto
1 ---------- ---------- ---------- ---------- ----------   0 Pro.loto
1 ---------- ---------- ---------- ---------- ----------   0 Mep.meph
1 ---------- ---------- ---------- ---------- ----------   0 Odo.rosm
1 ---------- ---------- ---------- ---------- ----------   0 Can.latr
1 ---------- ---------- ---------- ---------- ----------   0 Uro.cine
1 ---------- ---------- ---------- ---------- ----------   0 Urs.amer
1 ---------- ---------- ---------- ---------- ----------   0 Myd.marc

  51         61         71         81         91       100
  |          |          |          |          |         |
```

```
51  CAAACATCCA TAAACCGGTG TAAAATCCCT TAAAGATTTG CCT----AAA   96 Rat.norv
51  CAAATATCCA TAAACCGGTG TAAAATCCCT TAAAGATAT- -CC---AAAA   95 Hyl.stel
51  CAAATCTCTC TTAACCAGTG TAAAATCCCT TAAACGTTT- AACA---AAA   96 Mus.croc
51  CAAATCTCTA TAAACCGGTG TAAAATCCCT TAAATATTT- ACCT---AAA   96 Mus.cook
51  CAAACATCCA TAAACCGGTG TAGAATCCCT TAAAGAACT- --AT--TAAA   95 Mas.eryt
51  CAAATCTCTG TAAACCGGTG TAAAATCCCT TAAACATTT- AATT---AAA   96 Mus.matt
51  CAAATATCCG CAAACCGGTG TAAAATCCCT TAAAAGTTT- -AAT---AAA   95 Mus.plat
51  CAAATCTCTA TATACCGGTG TAAAATCCCT TAAATATTTT AGCT---AAA   97 Mus.setu
51  CAAATATCCG CAAACCGGTG TAAAATCCCT TAAAAGTTT- -AAT---AAA   95 Mus.saxi
51  CAAATATCCT TACACCAGTG TCAAATCCCT TAAAATTTTA ATCT--TAAA   98 Cri.gamb
51  CAAATCTCCT TAAACCAGTG CCTAATCCCT TATAAATTTA CTCA--GAAA   98 Mac.inge
51  CAAACGTCCA TAAACCAGTG TAGAATCCCA CAGAGTTCAG CC-----TAA   95 Mic.niva
51  CAAAGATCCC TAAACCAGTG TCAAATCCCT TATAATTCTA CC----TAGA   96 Nes.rufu
51  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA TT----CAAA   96 1Per.leu
51  CAAGACTCCC CACACCAGTG TCAAATCCCT TAGAACTAGA CAGCT----A   96 Aco.cahi
51  CAAAACTCCA TGAGCCAGTG TAAAATCCCT AGGG-TTTTG TAT---TAAA   96 Cri.migr
51  CAAGCATCTA CGACCCTGTG AGAATGCCCT CTACGTCAAA ------CAGA   94 Gli.glis
51  CAAACATCCA TGAACCTGTG CTTAATTCCT TACAGTTAGT CATCC-TCAA   99 Ger.nige
51  CAAATATCCA TAAACCGGTG TAAAATCCCT TAAACATTTG TCCC--AAAA   98 Leo.edwa
51  CAAATCTCTA TAAACCAGTG TCAAATCCCT AGG--TTTTA CTT---TAAA   95 Mes.aura
51  CAAGCATCTA CGACCCTGTG AGAATGCCCT CTATGTTAAT ACATGTTTAA  100 1Mus.ave
51  CAAACCTCCA TAGACCGGTG TAAAATCCCT TAAACATTT- ACTT---AAA   96 Mus.musc
51  CAAATCTCTT TTAACCAGTG TAAAATCCCT TAAACGTTT- AATA---GAA   96 Mus.paha
51  CAAATTTCCA TGAACCTGTG CGAATTCCCT TATAATTTAT A-----TTAA   95 Tat.kemp
51  CAAATATCCA TATACCAGTG TCAGATCCCC TTAAAATTTA ATC---TTAA   97 Ura.rudd
51  CAAGTATCCG CACACCTGTG AGAATGCCCT CTACGTTTAT TCA---TTAA   97 Dry.nite
51  CAAGCATCTA CGACCCTGTG AAAATGCCCT CTATGTTAAC AAATATTTAA  100 2Mus.ave
51  CAAGCATCCA CGATCCTGTG AGAATACCCT CTTTGTCACG ------CTGA   94 1Gli.jap
51  CAAGCATCCA CGATCCTGTG AGAATGCCCT CTATGTCACA ------CTGA   94 2Gli.jap
51  CAAGAGTCAT CGCCCCGGTG AAAATGCCCT CTAAACCACA TA-----CGG   95 1Hyd.hyd
37  CAAAAATCCC TACACCGGTG AGAATGCCCT ATGTACCACA CT----CAGG   82 Cav.porc
37  CAAATATCCT TAAGCCAGTG AAAATGCCCT TTAGATCATT ------CCGA   80 Bat.suil
37  CAAATATCCT TAAGCCAGTG AAAATGCCCT TTAGATCATT ------CCGA   80 Bat.jane
37  CAAGAATCCC CAAGCCGGTG AGAATGCCCT TAAAATCAAC ------ATGA   80 1Cry.hot
37  CGAGCATCCC CAAGCCAGTG AGAATGCCCT TAAAATCAAC ------ATGA   80 2Cry.hot
37  CAAGTATCCC CAAGCCAGTG AGGATGCCCT TTATAATCGA TT----CCGA   82 Cry.dama
37  CAAAAATCCC TGAGCCAGTG AGAATGCCCT TTAAATCATC ------CTGA   80 Geo.cape
37  CAAGAATCCC CAAACCAGTG AAAATACCCT TAAAATCATA ------TTGA   80 Hel.arge
37  CAAGTCTCCC CC-GCCAGTG AGAATGCCCT TAAAATCTTA AA----CTGA   81 Het.glab
37  CA-GCATCCC T--CCCAGTG AGCATGCCCT TTAAATCATA ACTAA-ATGA   82 Thr.swin
16  CAAATATCCG TACCCCAGTG AATTTTGCCC TAA-TTTCCG CAAAT-TAGA   63 2Cra.cas
16  CAAATATCCG TAAGCCAGTG AGTTTAGCCC TATGTTCTCT AA-----AGA   60 Dip.ordi
16  CAAGACTCCC CATACCTGTG AGAATACCCT TAGTATCTTA AA----CTGA   61 1Ped.cap
16  CAAGACTCCT CATACCTGTG AGAATACCCT TAGTATCTTA AA----CTGA   61 2Ped.cap
 2  CGAGGATCCC CTAACCGGTG AGAATACCCT TTAAATCCCA CAT---TGGA   48 Ath.macr
```

```
 2  CAAGAATCAT CTGTCCAGTG AAAATGCCCT TTAAATTAAC AA-----A-A   45 Cte.boli
 2  ACAAAATCCC TACACCGGTG AGAATGCCCT CTGTACCACA CTT----AGG   47 Cav.guia
 2  CAAGGCTCCC AATCC-AGTG AGAATGTCCT TAAGGTCTCT CA-----AG-   44 Chi.lani
 2  CAAGACTCAT CCTTCCAGTG AGAATGCCCT CCATGTCAAC AA------GA   45 Cap.pilo
 2  CAAGACTCCT CTCCCCGGTG AAAATGCCCT TTTAACCACA A------AGG   45 Das.punc
 2  CAAACATCCC CGCACCAGTG AGAACGCCCT TTAAATCTTA TA----AAGA   47 Ere.dors
 2  CGAGGAATCC CTTACCGGTG AGA-TACCCT TTTAATCCAA TAT---AGGA   47 Hys.afri
 2  CAAGAGTCAT CATTCCATGT GAGATGCCCA TCAAATCAAC A-------GA   44 Myo.coyp
 2  CAAAACTCAT CAAACCTGTG AGAATGCCCT CCAAATCGCA TC-----GGA   46 Oct.degu
 2  CAAGAGTCAT CACTCCTGTG AGAATGCCCT ACAAATCTTA TA------GA   45 Pro.long
 2  CAAGATTCCT CGCCCCAGTG AAAATGCCCT CCTCACCTAG CC----TAGA   47 Apl.rufa
 2  CAAATATCCG TGCCCCAGTG AGTACGCCCT CTAAGTCTAA ------CAGA   45 Cas.cana
 2  CAAATATCCG TACCCCAGTG AATTTTGCCC TAAATTTCCG CAAATTAGGA   51 1Cra.cas
 2  CAAGCATCCA CAAGCCCGTG AAAATGCCCT CAACATCCTC AA----GAGA   47 Cte.gund
 2  CAAGTATCAG CACACCAGTG AGAATGCCCT CTAACTCTTA ------TAGA   45 Das.nove
 2  CAAATATCCG TACCCCGGTG AGTTTTGCCC TGAATTTCCG CAAGCTGGGA   51 Geo.burs
 2  CAAATTTCCA TGTACCTGTG AGAATTCCCT TAAAGTTCAT T-----TAAA   46 Ger.vall
 2  CAAGCCTCCC CATACCAGTG AGAATACCCT TAAGTATCTA AA----ATGA   47 Jac.jacu
 2  CAAATCTCCA TATTCCGGTG TCAAATCCCA TAAAACTAGC TTCAA-CCTA   50 Lop.flav
 2  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAACTTCA AC----CTAA   47 Osg.band
 2  CAAGACTCCC CATACCTGTG AGAATACCCT TAGTATCTTA AA----CTGA   47 3Ped.cap
 2  CAAATTTCCG TAATCCGGTG AGTACAACCC TGAAGTTCTA TGAGAACAGA   51 Per.flav
 2  CAAGCATCCC CACACCGGTG AGAATGCCCT TTACATCA-A TTT----AGA   46 Syl.audu
 2  CAAGTATCCT CGCCCCAGTG AGAATGCCCT CTATCTCCAC AA-----GGA   46 Sci.nige
 2  CAAGCATCCC CGCCCCAGTG AGAATGCCCT CTATATCTAT AA----CTGA   47 Spe.trid
 2  CAAGCATCTA CGTCCCTGTG AGAATGCCCT CTATGTCAAT A-----ACGA   46 Gra.muri
 2  CAAGAGTCAT CGCCCCGGTG AAAATGCCCT CTAAACCACA TA-----CGG   46 2Hyd.hyd
 2  CAAGGATCCC CATGCCAGTG AGTACGCCCT CTAGATCATT T-----ATGA   46 Pet.typi
10  CAAAAGTCCA TAAACCAGTG TCAAATCCCC TAGAACTTTT TT----CTAA   55 Ony.leuc
10  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA TT----CAAA   55 Per.poli
10  CGAAACTCC- TAAGCCAGTG TCAAATCCCT TAGAACTTTT TT----ATAA   54 Neo.albi
10  CGAAACTCC- TAAGCCAGTG TCAAATCCCT TAGAACTTTT TT----ATAA   54 Neo.flor
10  CAAAAGTCCA TAAACCAGTG TCAAATCCCC TAGAACTTTT TT----CTAA   55 Ony.torr
10  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA AAA---TAAA   56 Per.erem
10  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA TT----CAAA   55 Per.goss
10  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA TT----CAAA   55 Per.mani
10  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA TT----CAAA   55 Per.mela
10  CAAAACTCCC TAAGCCAGTG TCAAATCCCC CAAAATTCTA CA-----GAA   54 Sig.hisp
10  CAAAGCTCCC TGTACCAGTG TCAAATCCCC CAAAATTCTA CA-----GAA   54 Sig.masc
10  CAAAAGTCCA TAAACCAGTG TCAAATCCCC TAGAACTTTT TT----CTAA   55 Ony.aren
10  CAAAAATCCG TAAACCAGTG TCAAATCCCC TAAAGCTTTA TT----CAAA   55 2Per.leu
 1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
 1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
 1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
 1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
```

```
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     101        111        121        131        141      150
     |          |          |          |          |        |
 97  ACTTAAGGAG AGGGCATCAA GCACATAATA ---------- ---TAGCTCA  133 Rat.norv
 96  ATTTAAGGAG AGGGTATCAA GCACATAATA ---------- ---TAGCTAA  132 Hyl.stel
 97  CTTCAAGGAG AGGGTATCAA GCACATTCTT A--------- ---TAGCTTA  134 Mus.croc
 97  ATTTAAGGAG AGGGTATCAA GCACATAAAA A--------- ---TAGCTTA  134 Mus.cook
 96  ATTTAAGGAG AGGGTATCAA GCACATAAAA ---------- ---TAGCTAA  132 Mas.eryt
 97  CTTTAAGGAG AGGGTATCAA GCACATTTAA A--------- ---TAGCTTA  134 Mus.matt
 96  CTTCAAGGAG AGGGTATCAA GCACATTTTA AA-------- ---TAGCTTA  134 Mus.plat
 98  CTTTATGGAG AGGGTATCAA GCACATTATA A--------- ---TAGCTCA  135 Mus.setu
 96  CTTCAAGGAG AGGGTATCAA GCACATTTTA AA-------- ---TAGCTTA  134 Mus.saxi
 99  -TTTAAGGAG AGGGTATCAA GTTCATTATC TA-------- ---TAGCTTA  136 Cri.gamb
 99  TTTCAAGGAG AGGATATCAA GCACATGTAT ATATTAATTT ACATAGCTCA  148 Mac.inge
 96  CTCTTAGGAG AGGGTATCAA GTACATACAA ---------- ---TAGCTAA  132 Mic.niva
 97  ATTTAAGGAG AGGGTATCAA GTACATTCTT A--------- ---TAGCTAA  134 Nes.rufu
 97  ACCTAGGGAG AGGGCATCAA GTTCATACAG T--------- --ATAGCTAA  135 1Per.leu
 97  ACCTAAGGAG AGGATATCAA GCACATTAAT A--------- ---TAGCTCA  134 Aco.cahi
 97  CTCTAAGGAG AGGGTATCAA GCACATTATA A--------- ---TAGCTAA  134 Cri.migr
 95  CTTAAAGGAG CAGGTATCAA GCACACATCA ---------C CGTAGCTCA-  134 Gli.glis
100  ACTTAAGGGA ACAGTATCAA GCACATAAAT A--------- ---TAGCTTA  137 Ger.nige
 99  ATTTAAGGAG AGGGTATCAA GCACATAAAA ---------- ---TAGCTCA  135 Leo.edwa
 96  CCCTAAGGAG AGGGTATCAA GCACATACAC A--------T ATATAGCTAA  137 Mes.aura
101  CCTAAAGGAG CAGGTATCAA GCACACTATA A--------- TAGTAGCTCA  141 1Mus.ave
 97  ATTTAAGGAG AGGGTATCAA GCACATTAAA A--------- ---TAGCTTA  134 Mus.musc
 97  CTTCAAGGAG AGGGTATCAA GCACATTTTT A--------- ---TAGCTTA  134 Mus.paha
 96  ACCTAAGGAG AAGGTATCAA GCACATAAAC A--------- ---TAGCTTC  133 Tat.kemp
 98  ATATAAGGAG GAGATATCAA GCACATATCT ---------- GTATAGCTTA  137 Ura.rudd
```

```
 98  TCCAAAGGAG CAGGTATCAA GCACACTTTA ---------- AAGTAGCTCA  137 Dry.nite
101  CCTAAAGGTG CAGGTATCAA GCACACTAT- ---------- TAGTAGCTCA  139 2Mus.ave
 95  CAAAAAGGAG TAGGTATCAA GTTCACTTTA ---------- CAGTTGCTCA  134 1Gli.jap
 95  CAAAAAGGAG TAGGTATCAA GTTCACTTTA ---------- TAGTAGCTCA  134 2Gli.jap
 96  ATAAAAGGAG CGGGTATCAG GCGCACACCA CA-------- AAGTAGCCCA  137 1Hyd.hyd
 83  TTTAAAGGAG CAGACATCAA GC--ACACTG CT-------- AAGTAGCTCA  122 Cav.porc
 81  TATAAAGGAG CGGGTATCAA GCACACATAT A--------- -AGCAGCTCA  120 Bat.suil
 81  TATAAAGGAG CGGGTATCAA GCACACATAT A--------- -AGTAGCTCA  120 Bat.jane
 81  TAAGAAGGAG CGGGTATCAA GTGCACACAA A--------- -AGTAGCTCA  120 1Cry.hot
 81  TAAGAAGGAG CTGGTATCAA GTACACACAA A--------- -AGTAGCTCA  120 2Cry.hot
 83  TAAAAAGGAG CAGGTATCAA GTTCGCACAA A--------- -AGTAGCTCA  122 Cry.dama
 81  TATAATGGAG CGGGTATCAA GTACACATTA A--------- -AGTAGCTCA  120 Geo.cape
 81  TTCGAAGGAG TTGGTATCAA GTACACACCA AT-------- -AGTAGCTCA  121 Hel.arge
 82  CAAAAAGGAG CGGGTATCAA GTACACTACC A--------- -AGTAGCTCA  121 Het.glab
 83  TCAAAAGGAG CGGGCATCAA GCACGCTTAA ATC------- AAGCAGCTCA  125 Thr.swin
 64  AAAAAAGGAG CAGGTATAAA GCACACTAAA CGAAAT---- TAGTAGCTTA  109 2Cra.cas
 61  CTACAAGGAG CAAGTATAAA GCACGCGTGA TAACTA---C ACGCAGCTAA  107 Dip.ordi
 62  TATAAAGGAG TGGGTATCAA GCACACTATA T--------- -AGTAGCTAA  101 1Ped.cap
 62  TATAAAGGAG TGGGTATCAA GCACACTATA T--------- -AGTAGCTAA  101 2Ped.cap
 49  TTAAAAGGAG TGGGTATCAA GCACAGCTTG CG-------- AAGTAGCTCA   90 Ath.macr
 46  TCAAAAGGAG CTGATATCAA GCACACCAA- T--------- -GGTAGCTCA   84 Cte.boli
 48  TTTAAAGGAG CGGACATCAA GCACACTGCT A--------- -AGTAGCTCA   87 Cav.guia
 45  -TCAAAGAA- CAGGTATCAG GTGCACTCAC C--------- -AGTAGCCCA   82 Chi.lani
 46  CCCAGAGGAG CGGGTATCAA GCACACCTT- C--------- -AGTAGCTCA   84 Cap.pilo
 46  ATGAAAGGAG CGGGTATCAA GCACACTAAT T--------- -AGTAGCTCA   85 Das.punc
 48  CAAGAAGGAG CAGGTATCAA GCACGC-CAA C--------- -AGCAGCTCA   86 Ere.dors
 48  TTAAAAGGAG TAGGTATCAA GCACACTTAT A--------- -AGTAGCTCA   87 Hys.afri
 45  TCTAAAGGAG CTGGTATCAA GCACACATAA C--------- -AGTAGCTCA   84 Myo.coyp
 47  CCTAGAGGAG CTGGTATCAA GCACACTAC- ---------- -CGTAGCTCA   84 Oct.degu
 46  TCCAAAGGAG CAGGTATCAA GCACACCCA- C--------- -GGTAGCTCA   84 Pro.long
 48  T-AAAAGGAG CAGGTATCAA GCACGCTTAA CTC------- TAGCAGCTCA   89 Apl.rufa
 46  CCAAAAGGAG CAAGCATCAA GCACACAAAA C--------- TTGTAGCTAA   86 Cas.cana
 52  AAAAAAGGAG CAGGTATAAA GCACACTAAA CGAAAT---T AGTAGCTTAA   98 1Cra.cas
 48  TATAAAGGAG CCGGTATCAA GTACACAAT- ---------- -TGTAGCTCA   85 Cte.gund
 46  TCAAAAGGAG CAAGCATCAA GTACACACAG CCCTTA---- CAGTAGCTCA   91 Das.nove
 52  AAAAAAGGAG CAGGTATAAA GCACACTAAA ACTAAA---T CAGTAGCTAA   98 Geo.burs
 47  GTTTAAGGAG TTGGTATCAA GCACATTCTT ---------- ---TAGCTCA   83 Ger.vall
 48  TTGTAAGGAG TAGGTATCAA GCGCACTAG- ---------- -CGTTGCTCA   85 Jac.jacu
 51  GCCAAAGGAG AGGATATCAA GCACATTTAA ---------C CTATAGCTCA   91 Lop.flav
 48  GCTTAGGGAG AGGGCATCAA GTACATACAA A--------- ---TAGCTAC   85 Osg.band
 48  TATAAAGGAG TGGGTATCAA GCACACTATA T--------- -AGTAGCTAA   87 3Ped.cap
 52  GGGAGTAGGC CTTA-AGCTC TCTACATAAA TAAATTA--- CAGATGCTCA   97 Per.flav
 47  TCAAGAGGAG CAGGCATTAA GCACACTATT T--------- -AGTAGCTCA   86 Syl.audu
 47  TCAAAAGGTG CAGGTATTAA GTACGCTCTT TCTT------ GAGCAGCTCA   90 Sci.nige
 48  TCAAAAGGTG CAGGCATCAA GTTCACTACT CCT------- -AGTAGCTCA   89 Spe.trid
```

```
 47  CCAAAAGGAG CAGGTATCAA GCACACACCT A--------- ATGTAGCTCA   87 Gra.muri
 47  ATAAAAGGAG CGGGTATCAG GCGCACACCA CA-------- AAGTAGCCCA   88 2Hyd.hyd
 47  TCAGATGGAG TGGGTATCAA GTACACTTAA AC-------- AAGTAGCTCA   88 Pet.typi
 56  GCCTAAGGAG AGGACATCAA GCACATAA-- ACA------- ---TAGCTAA   93 Ony.leuc
 56  ACCTAGGGAG AGGGCATCAA GTACATAC-A ATA------- ---TAGCTAA   94 Per.poli
 55  ACCTAAGGAG AGGGCATCAA GCACATTAAA TAA------- ---TAGCTCA   94 Neo.albi
 55  ACCTAAGGAG AGGGCATCAA GCACATTAAA TAA------- ---TAGCTCA   94 Neo.flor
 56  GCCTAAGGAG AAGACATCAA GCACATAA-- ATA------- ---TAGCTAA   93 Ony.torr
 57  ACCTAGGGAG AGGGCATCAA GCACATATTT CAA------- ---TAGCTAA   96 Per.erem
 56  ACCTAGGGAG AGGGCATCAA GTTCATAC-A ATA------- ---TAGCTAA   94 Per.goss
 56  ACCTAGGGAG AGGGCATCAA GTACATACAA ATA------- ---TAGCTAA   95 Per.mani
 56  ACCTAGGGAG AGGGCATCAA GTACATATTA GTA------- ---TAGCTAA   95 Per.mela
 55  TTTGAAGGAG AGGACATCAA GTTCATTC-- TAA------- ---TAGCTAA   92 Sig.hisp
 55  TTTGAAGNAG AGGACATCAA GTGCATTC-- TCA------- ---TAGCTAA   92 Sig.masc
 56  GCCTAGGGAG AGGACATCAA GCACATAA-- ACA------- ---TAGCTAA   93 Ony.aren
 56  ACCTAGGGAG AGGGCATCAA GTTCATAC-A NTA------- ---TAGCTAA   94 2Per.leu
  1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     151        161        171        181        191      200
     |          |          |          |          |        |
134  AGACGCCTTG CCT-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAATA  181 Rat.norv
133  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAATA  180 Hyl.stel
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-TCAGCA GTGATAAATA  182 Mus.croc
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-CCAGCA GTGATAAATA  182 Mus.cook
133  AAACACCTTG CCC-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAATA  180 Mas.eryt
```

```
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-CCAGCA GTGATAAATA  182 Mus.matt
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-TCAGCA GTGATAAATA  182 Mus.plat
136  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-ACAGCA GTGATAAATA  183 Mus.setu
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-TCAGCA GTGATAAATA  182 Mus.saxi
137  AGACACCTTG CCA-AGCCAC ACCCCCACGG GAC-GCAGCA GTGATAAAAA  184 Cri.gamb
149  AGACATCTTG CCA-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAAAA  196 Mac.inge
133  AGACACCTTG CCT-AGCCAC GCCCCCACGG GAC-TCAGCA GTGATAAAAA  180 Mic.niva
135  AAACACCTTG CCT-AGCCAC ACCCCCACGG GAC-CCAGCA GTGATAAATA  182 Nes.rufu
136  AGACGCCTTG CCT-AGCCAC GCCCCCACGG GAT-TCAGCA GTGATAAAAA  183 1Per.leu
135  AGACATCTCG CCA-AGCCAC ACCCCCACGG GAA-ACAGCA GTGATAAAAA  182 Aco.cahi
135  AGACACCTTG CCT-AGCCAC GCCCCCACGG GAC-TCAGCA GTGATAAAAA  182 Cri.migr
135  AAACACCTTG CTA-AACCAC ACCCCCACGG GTT-ACAGCA GTGATTAAAA  182 Gli.glis
138  TGACGCTTAG CCT-AGCCAC ACCCCCACGG GATTTCAGCA GTGATAAAAA  186 Ger.nige
136  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-CCAGCA GTGATAAATA  183 Leo.edwa
138  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAAAA  185 Mes.aura
142  AAACACCTTG CTT-AACCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  189 1Mus.ave
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAATA  182 Mus.musc
135  AGACACCTTG CCT-AGCCAC ACCCCCACGG GAT-TCAGCA GTGATAAATA  182 Mus.paha
134  TGACATCTAG CCT-AGCCAC TCCCCCACGG GAT-TCAGCA GTGATAAAAA  181 Tat.kemp
138  AGACATCTTG CCT-AGCCAC ACCCCCACGG GTA-ACAGCA GTGATAAAAA  185 Ura.rudd
138  AGACACCTTG CTT-AACCAC ACCCCCACGG GAT-ACAGCA GTGATTAAGA  185 Dry.nite
140  AAACACCTTG CTT-AACCAC ACCCCCACGG GAA-ACAGCA GTGATTAAAA  187 2Mus.ave
135  AGACACCTTG CTA-AACCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  182 1Gli.jap
135  AAACACCTTG CTA-AACCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  182 2Gli.jap
138  TAACACCTTG CTT-CGCCAC ACCCCTACGG GAG-ACAGCA GTAACAAAAA  185 1Hyd.hyd
123  CGACGTCTTG CTT-TGCCAC ACCCCCACGG GAA-ACAGCA GTAATAAAAA  170 Cav.porc
121  CGACACCTTG CTT-AGCCAC GCCCCCACGG GAA-ACAGCA GTGATAAATA  168 Bat.suil
121  CGACACCTTG CTT-AGCCAC GCCCCCACGG GAA-ACAGCA GTGATAAAAA  168 Bat.jane
121  ATACACCTTG TAC-AACCAC ACCCCCACGG GAA-ACAGCA GTGATAAAAA  168 1Cry.hot
121  ATACACCTTG TAC-AACCAC ACCCCCACGG GAA-ACAGCA GTGATAAAAA  168 2Cry.hot
123  CAACACCTTG TAA-AACCAC ACCCCCACGG GTA-ACAGCA GTGATAAAAA  170 Cry.dama
121  TAACACCTTG CTT-AACCAC ACCCCCACGG GAA-ACAGCA GTGATAAAAA  168 Geo.cape
122  TTACACCTTG CCA-AGCCAC ACCCCCACGG GAT-ACAGCA GTGACAAGAA  169 Hel.arge
122  AGACACCTTG CTT-AGCCAC ACCCCCACGG GAA-ACAGCA GTAATTAATA  169 Het.glab
126  TAACGCCTTG CTC-AGCCAC ACCCCCACGG GAT-ACAGCA GTAATAAAAA  173 Thr.swin
110  AAACACCAGG TCC-AACCAC GCCCCCACGG GAT-ACAGCA GTGATAAAAA  157 2Cra.cas
108  AAACACTAAG TTT-AACCAC ACCCCCACGG GAT-ACAGCA GTGATAAAAT  155 Dip.ordi
102  AGACATCTTG CCC-AGCCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  149 1Ped.cap
102  AGACATCTTG CCC-AGCCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  149 2Ped.cap
 91  TTACACCTTG CCT-AGCCAC ATCCCCACGG AAACACAGCA GTGATAAAAA  139 Ath.macr
 85  TAACATCTTG CCC-AGCCAC ACCCCCACGG GAT-ACAGCA GTAATTAAAA  132 Cte.boli
 88  CGACGTCTTG CTT-TGCCAC ACCCCCACGG GAG-ACAGCA GTAATAAAAA  135 Cav.guia
 83  CAACACCTTG TAA-AACCAC ACCC-CACGG GAC-ACAGCA GTAATTAAAC  129 Chi.lani
 85  TTACACCTTG CTC-AGCCAC ACCCCCACGG GAT-ACAGCA GTAATTAAAA  132 Cap.pilo
 86  CAACGCCTTG CTT-TGCCAC ACCCCCATGG -AA-ACAGCA GTAATAAAAA  132 Das.punc
```

```
87  CGACATCTCG CTT-TGCCAC GCCCTCACGG GAA-ACAGCA GTAATAAAAA  134 Ere.dors
88  CTACACCTTG CTT-AGCCAC ATCCCCACGG AAATACAGCA GTGATAAAAA  136 Hys.afri
85  CAACACCTTG CTT-AGCCAC ACCCCCACGG GAT-ACAGCA GTAATCAAAA  132 Myo.coyp
85  TGACACCCCG CTT-AGCCAC ACCCCCACGG GAT-ACAGCA GTAATTAAAA  132 Oct.degu
85  TAACACCTTG CTT-TGCCAC ACCCCCACGG GAC-ACAGCA GTAACTAAAA  132 Pro.long
90  TGACACCTTG CAC-AGCCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  137 Apl.rufa
87  TGACGCTTTG CCT-TGCCAC ACCCCCACGG GAT-ACAGCA GTAATAAAAA  134 Cas.cana
99  AAACACCAAG TCC-AACCAC GCCCCCACGG GAT-ACAGCA GTGATAAAAA  146 1Cra.cas
86  AAACACCTTG CTA-AGCCAC ACCCCCACGG GAA-ACAGCA GTGATAAAAC  133 Cte.gund
92  TAACGCTTTG CTC-AACCAC ACCCCCACGG GAT-ACAGCA GTGATAAAAA  139 Das.nove
99  AAACATCAAG TCT-GACCAC ACCCCCACGG GAT-ACAGCA GTGATAAAAA  146 Geo.burs
84  CGACACCTAG CTC-AGCCAC TCCCCCACGG -AT-ACAGCA GTGATAAAAA  130 Ger.vall
86  AGACACCTTG CTT-AGCCAC ACCCCCACGG GAA-ACAGCA GTGATAAAAA  133 Jac.jacu
92  AGACATCTTG CCCCAGCCAC ACCCCCACGG -AT-ACAGCA GTGATAGAAA  139 Lop.flav
86  AGACGCCTTG CCT-AGCCAC GCCCCCACGG GAC-CCAGCA GTGATAAAAA  133 Osg.band
88  AGACATCTTG CCC-AGCCAC ACCCCCACGG GAT-ACAGCA GTGATTAAAA  135 3Ped.cap
98  CAACGCCAAG TCT-AACCAC ACCCCCACGG GTT-ACAGCA GTAATAAAAC  145 Per.flav
87  AGATGCCTTG CTT-AACCAC ACCCCCAAGG GAA-ACAGCA GTGATAGAAA  134 Syl.audu
91  TGACACCTTG CTA-TACCAC ACCCCCACGG GAC-ACAGCA GTGATTAAAA  138 Sci.nige
90  TAACGCCTTG CTC-CACCAC ACCCCCACGG GAC-ACAGCA GTAATTAACA  137 Spe.trid
88  AAACACCTTG CTTTCGCCAC ACCCCCACGG GAA-ACAGCA GTGATCAAAA  136 Gra.muri
89  TAACACCTTG CTT-CGCCAC ACCCCTACGG GAG-ACAGCA GTAACAAAAA  136 2Hyd.hyd
89  AGACACCCAG CCA-AGCCAC GCCCCCACGG GAC-ACAGCA GTAATAAAAA  136 Pet.typi
94  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AC-CCAGCA GTGATAAAAA  140 Ony.leuc
95  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AC-TCAGCA GTGATAAAAA  141 Per.poli
95  AGACGCCTTG CCT-AGCCAC GCCCCCACGG GAC-CCAGCA GTGATAAAAA  142 Neo.albi
95  AGACGCCTTG CCT-AGCCAC GCCCCCACGG GAC-CCAGCA GTGATAAAAA  142 Neo.flor
94  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AC-CCAGCA GTGATAAAAA  140 Ony.torr
97  AGACGCCTTG CCA-AGCCAC GCCCCCACGG -AT-TCAGCA GTGATAAAAA  143 Per.erem
95  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AT-TCAGCA GTGATAAAAA  141 Per.goss
96  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AC-TCAGCA GTGATAAAAA  142 Per.mani
96  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AC-TCAGCA GTGATAAAAA  142 Per.mela
93  AGACGTCTTG CCT-AGCCAC ACCCCCACGG GAC-CCAGCA GTGATAAAAA  140 Sig.hisp
93  AGACGTCTTG CCT-AGCCAC ACCCCCACGG GAC-TCAGCA GTGATAAAAA  140 Sig.masc
94  AGACGTCTTG CCT-AGCCAC GCCCCCACGG -AC-TCAGCA GTGATAAAAA  140 Ony.aren
95  AGACGCCTTG CCT-AGCCAC NCCCCCACGG -AT-TCAGCA GTGATAAAAA  141 2Per.leu
 1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
 1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
 1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
 1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
 1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
 1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
 1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
 1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
 1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
```

```
  1 ---------- ---------- ---------- ---------- ----------      0 Sci.vulg
  1 ---------- ---------- ---------- ---------- ----------      0 1Tam.sib
  1 ---------- ---------- ---------- ---------- ----------      0 2Tam.sib
  1 ---------- ---------- ---------- ---------- ----------      0 Bas.astu
  1 ---------- ---------- ---------- ---------- ----------      0 Spi.puto
  1 ---------- ---------- ---------- ---------- ----------      0 Pro.loto
  1 ---------- ---------- ---------- ---------- ----------      0 Mep.meph
  1 ---------- ---------- ---------- ---------- ----------      0 Odo.rosm
  1 ---------- ---------- ---------- ---------- ----------      0 Can.latr
  1 ---------- ---------- ---------- ---------- ----------      0 Uro.cine
  1 ---------- ---------- ---------- ---------- ----------      0 Urs.amer
  1 ---------- ---------- ---------- ---------- ----------      0 Myd.marc

    201        211        221        231        241       250
    |          |          |          |          |         |
182 TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CTAGTACCT CTC-------  221 Rat.norv
181 TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CTA-TACCT CTA-------  219 Hyl.stel
183 TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  221 Mus.croc
183 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  221 Mus.cook
181 TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CTA-TACCT CTT-------  219 Mas.eryt
183 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TCA-TACCT CTT-------  221 Mus.matt
183 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  221 Mus.plat
184 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  222 Mus.setu
183 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  221 Mus.saxi
185 TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACTT CCA-------  223 Cri.gamb
197 TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TTCCT TCA-------  235 Mac.inge
181 TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CCA-TGCCT CTCT-A----  221 Mic.niva
183 TTTAGC-AAT AAATGAAAGT TTGA-CTAAG -TCA-TTCCT CTT-------  221 Nes.rufu
184 TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTCA------  223 1Per.leu
183 TTAAGC-AAT GAACGAAAGT TTGA-CTTAG -TTA-TACTT CCA-------  221 Aco.cahi
183 TTAAGC-AAT GAACGAAAGT TTGA-CTTAG -TAA-TACCT TATACA----  224 Cri.migr
183 TTAAAA-TAT AAACGAAAGT TTGA-CTAAG -CTA-TGCTA TAATCT----  224 Gli.glis
187 TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -CTA-TACCT TAAAG-----  227 Ger.nige
184 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -CTA-TACCT CTT-------  222 Leo.edwa
186 TTAAGC-CAT AAACGAAAGT TTGA-CTTAG -TCA-TACCT CATC------  225 Mes.aura
190 TTAAGT-TAT AAACGAAAGT TTGA-CTAAG -CCA-TGTTA TAT-------  228 1Mus.ave
183 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  221 Mus.musc
183 TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  221 Mus.paha
182 TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -CTA-CACCT TAG-------  220 Tat.kemp
186 TTAAGC-AAT GAACGAAAGT TCGA-CTAAG GTCA-TACCC CAA-------  225 Ura.rudd
186 TTAAGC-TAT GAACGAAAGT TTGA-CTAAG -CTA-TGTTA TAGT-T----  226 Dry.nite
188 TTAAGA-CAT AAACGAAAGT TTGA-CTAAG -CTA-TGTTA TAAA------  227 2Mus.ave
183 TTAAGT-TAT AAACGAAAGT TTGA-CTAAG -CTA-TACTA TTTAAT----  224 1Gli.jap
183 TTAAGT-TAT AAACGAAAGT TTGA-CTAAG -CTA-TACTA TTAA------  222 2Gli.jap
186 TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TCA-TGTAG CTA-------  224 1Hyd.hyd
```

```
171  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TGCAG CAC-------  209 Cav.porc
169  TTAAGC-CAT GAACGAAAGT TTGA-CTTAG -TCA-TGCAA TATAAT----  210 Bat.suil
169  TTAAGC-TAT GAACGAAAGT TTGA-CTTAG -TCA-TGCAA TATAAT----  210 Bat.jane
169  TTAAGC-TAT GAACGAAAGT TCGA-CTTAG -TTA-TGCAA TTTACC----  210 1Cry.hot
169  TTAAGC-TAT AAACGAAAGT TTGA-CTTAG -TTA-TGCAA TT-ATT----  209 2Cry.hot
171  TTAAGC-TAT GAACGAAAGT TCGA-CTTAG -TCA-TGCAA TACAGT----  212 Cry.dama
169  TTAAGC-TAT GAACGAAAGT TTGA-CTTAG -TTA-TGCAA TACAAT----  210 Geo.cape
170  TTAAGA-AAT AAACGAAAGT TTGA-CTTAG -TTA-CGCAA TACAAT----  211 Hel.arge
170  TTAAGC-AAT GAACGAAAGT TTGA-CTTAA -TTA-TGCGA TCAGAC----  211 Het.glab
174  TTAAGC-AAT AAACGAAAGT TTGAACTTAG -TCA-TACTA CT--------  212 Thr.swin
158  TTAGGC-AAT GAACGAAAGT TTGA-CTAAG -CAA-TACTA ATTC------  197 2Cra.cas
156  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACTG CAAAAA----  197 Dip.ordi
150  TTAAGC-CAT GAACGAAAGT TTGA-CTAAG -TTA-TGCTA TT--------  187 1Ped.cap
150  TTAAGC-CAT GAACGAAAGT TTGA-CTAAG -TTA-TGCTA TT--------  187 2Ped.cap
140  TTAAGC-AAT GAACGAAAGT TTGA-CTTAG -TTA-TGCTA ATT-------  178 Ath.macr
133  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TATAA CAAA----TA  174 Cte.boli
136  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TCA-TGCAG CAA-------  174 Cav.guia
130  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-CGCAA CACAC-----  170 Chi.lani
133  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TGCAA CATTA---TA  175 Cap.pilo
133  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TACAC TAGCA----A  174 Das.punc
135  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TGCTT CAACTCTTTA  180 Ere.dors
137  TTAAGC-AAT AAACGAAAGT TTGA-CTTAG -TTA-TGCTA ATA-------  175 Hys.afri
133  TTAAGC-TAT AAACGAAAGT TCGA-CTAAG -TCA-TGCAA TTTCT----T  174 Myo.coyp
133  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TTA-TATAA CAACT-----  173 Oct.degu
133  TTAAGC-TAT AAACGAAAGT TTGA-CTAAG -TTA-TGCCA TTCTTATT--  176 Pro.long
138  TTAAGC-TAT AAATGAAAAT TTGA-CTAAG -CTA-TGCTA CTCT------  177 Apl.rufa
135  TTAAGC-AAT AAACGAAAGT TTGA-CTTAG -TTA-CACTG ACCA------  174 Cas.cana
147  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CAA-TACTA ATT-------  185 1Cra.cas
134  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TCA-TGCAT AA--------  171 Cte.gund
140  TTAAGC-AAT AAGCGAAAGC TTGA-CTAAG -TTA-TGTTA TTAT------  179 Das.nove
147  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TAA-TACTA ATC-------  185 Geo.burs
131  TTTAGC-AAT AAACGAAAGT TTGA-CTAAG -CTA-CACCT TAT-------  169 Ger.vall
134  TTAGGC-AAT GAACGAAAGT TTGA-CCTAG -TTA-TACTA TTC-------  172 Jac.jacu
140  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACTC TCG-------  178 Lop.flav
134  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTC-------  172 Osg.band
136  TTAAGC-CAT GAACGAAAGT TTGA-CTAAG -TTA-TGCTA TT--------  173 3Ped.cap
146  TTAAGC-CAT GAACGAAAGT TTGA-CTAAG -TTA-TACTG ACT-------  184 Per.flav
135  TTTAGC-AAT GAACGCAAGT TTGA-CTAAG -TTA-TGCTA CTC-------  173 Syl.audu
139  TTGAGA-TAT GAACGAAAGT TTGA-CTGAG -TCA-TACTA ATAC------  178 Sci.nige
138  TTAAGCCTAT AAACGAAAGT TTGA-CTAAG -TTA-AGCTA AAT-------  177 Spe.trid
137  TTAATT-TAT AAACGAAAGT TTGA-CTAAG -TTA-TGCTA TTA-------  175 Gra.muri
137  TTAAGC-AAT AAACGAAAGT TTGA-CTAAG -TCA-TGTAG CTAT------  176 2Hyd.hyd
137  TTAAGC-AAT GAACGAAAGT TCGA-CTTAG -TCA-TACAA TC--------  174 Pet.typi
141  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTC-------  179 Ony.leuc
142  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTC-------  180 Per.poli
```

```
143  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CTA-TATCT CCA-------  181 Neo.albi
143  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -CTA-TATCT CCA-------  181 Neo.flor
141  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  179 Ony.torr
144  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTC-------  182 Per.erem
142  TTATGC-AAT GAACGAAAGT TTGA-CTAAG -CTA-TATCC TTC-------  180 Per.goss
143  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTT-------  181 Per.mani
143  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTC-------  181 Per.mela
141  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT C-C-------  178 Sig.hisp
141  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT C-C-------  178 Sig.masc
141  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TACCT CTC-------  179 Ony.aren
142  TTAAGC-AAT GAACGAAAGT TTGA-CTAAG -TTA-TNCCT CTCA------  181 2Per.leu
  1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     251        261        271        281        291      300
     |          |          |          |          |        |
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Rat.norv
220  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  263 Hyl.stel
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.croc
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.cook
220  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  263 Mas.eryt
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.matt
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.plat
223  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  266 Mus.setu
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.saxi
224  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  267 Cri.gamb
```

```
236  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  279 Mac.inge
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mic.niva
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Nes.rufu
224  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  267 1Per.leu
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Aco.cahi
225  -----GGGTT GGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  268 Cri.migr
225  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  268 Gli.glis
228  -----AGTTT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTGA  271 Ger.nige
223  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  266 Leo.edwa
226  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  269 Mes.aura
229  ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  273 1Mus.ave
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.musc
222  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  265 Mus.paha
221  -----AGAGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  264 Tat.kemp
226  -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  269 Ura.rudd
227  ----AAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  271 Dry.nite
228  ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  272 2Mus.ave
225  ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  269 1Gli.jap
223  ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  267 2Gli.jap
225  --TTAAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  271 1Hyd.hyd
210  ---TTAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTGA  255 Cav.porc
211  ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-G-TC ATACGATTAA  254 Bat.suil
211  ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  255 Bat.jane
211  ----TAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-G-TT ATACGATTGA  254 1Cry.hot
210  ----TAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  254 2Cry.hot
213  ----TAGGGT TGGTAAATCT CGTGCCAGCC GCCGC-G-TC ATACGATTAA  256 Cry.dama
211  ----TAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  255 Geo.cape
212  ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  256 Hel.arge
212  ----TAGGGT TGGTAAATAT CGTGCCAGCC ACCGC-GGTT ATACGATTAA  256 Het.glab
213  ----CAGGGT TGGTCAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  257 Thr.swin
198  ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  242 2Cra.cas
198  ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  242 Dip.ordi
188  ----TAGAGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTT ATTCGATTAA  232 1Ped.cap
188  ----TAGAGT TGGTAAATTT CGTGCCAGCC GCCGC-GGTT ATTCGATTAA  232 2Ped.cap
179  ----CAGAGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  223 Ath.macr
175  A---AAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  220 Cte.boli
175  ---TCAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTGA  220 Cav.guia
171  ----CAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  215 Chi.lani
176  ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  220 Cap.pilo
175  A---AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  220 Das.punc
181  ATCTTAGGGT TGGTAAATAT CGTGCCAGCC ACCGC-GGTT ATACGATTAA  229 Ere.dors
176  ----TAGAGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  220 Hys.afri
175  A---TAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  220 Myo.coyp
174  ----AAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  218 Oct.degu
177  -CT-AAGGGT TGGTAAATCT CGTGCCAGCC ACCGC-G-TC ATACGATTAA  222 Pro.long
```

```
178 ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  222 Apl.rufa
175 ----CAGGGT TGGTAAATAT CGTGCCAGCC ACCGC-GGTC ATACGATAGA  219 Cas.cana
186 ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  230 1Cra.cas
172 ----TAGAGC TGGTCAACTT CGTGCCAGCC ACCGC-GGTC ATACGATGAG  216 Cte.gund
180 ----AAGGGT TGGTAAATTT CGTGCCAGCA ACCGC-GGTC ATACGATTAA  224 Das.nove
186 ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  230 Geo.burs
170 ----CAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  214 Ger.vall
173 -----AGGGT TGGTCAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  216 Jac.jacu
179 -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC AT-CGATTAA  221 Lop.flav
173 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  217 Osg.band
174 ----TAGAGT TGGTAAATTT CGTGCCAGCC GCCGC-GGTT ATACGATTAA  218 3Ped.cap
185 ----CAGGGT TGGTAAACTT CGTGCCAGCC ACCGC-GGTC ATACGAATGA  229 Per.flav
174 ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  218 Syl.audu
179 ---TTAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  224 Sci.nige
178 ---TTAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  223 Spe.trid
176 ---TTAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  221 Gra.muri
177 ---TAAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  222 2Hyd.hyd
175 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-G-TC ATACGATTAA  218 Pet.typi
180 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  224 Ony.leuc
181 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  225 Per.poli
182 ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  226 Neo.albi
182 ----TAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  226 Neo.flor
180 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  224 Ony.torr
183 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  227 Per.erem
181 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  225 Per.goss
182 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATAGCATTAA  226 Per.mani
182 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  226 Per.mela
179 ----AAGGGT TGGTAAATCT CGTGCCAGCC ACCGCCGGTC ATACGATTAA  224 Sig.hisp
179 ----AAGGGT TGGTAAATCT CGTGCCAGCC ACCGCCGGTC ATACGATTAA  224 Sig.masc
180 ----AAGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  224 Ony.aren
182 -----AGGGT TGGTAAATTT CGTGCCAGCC ACCGC-GGTC ATACGATTAA  225 2Per.leu
  1 ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1 ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1 ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1 ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1 ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1 ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1 ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1 ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1 ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1 ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1 ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1 ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1 ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1 ---------- ---------- ---------- ---------- ----------    0 Spi.puto
```

```
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     301        311        321        331        341      350
     |          |          |          |          |        |
266  CCC--AAACT AATTATTT-T CGGCGTAAAA CGTGCCAACT ATAAATCT--  310 Rat.norv
264  CCC--TAACT AATTACTC-T CGGCGTAAAA CGTGTTAACT GAACAAAC--  308 Hyl.stel
266  CCC--AAACT AATTATTT-C CAGCGTAAAA CGTGTTAACT ATTAACTA--  310 Mus.croc
266  CCC--AAACT AATTATTT-T CGGCGTAAAA TGTGTTAACT ATAGATAA--  310 Mus.cook
264  CCC--TAATT AATTATTC-T CGGCGTAAAA CGTGTTAATT GAGCATGC--  308 Mas.eryt
266  CCC--AAACT AATTATTC-T CGGCGTAAAA TGTGTTATCT ATAAATAA--  310 Mus.matt
266  CCC--AAACT AATTATTC-T CGGCGTAAAA TGTGTTAACT ACAAATAA--  310 Mus.plat
267  CCC--AAACT AATTATTC-T CGGCGTAAAA TGTGTTAACT ATAAATTA--  311 Mus.setu
266  CCC--AAACT AATTATTC-T CGGCGTAAAA TGTGTTAACT ACAAATAA--  310 Mus.saxi
268  CCC--AAACT AATTAAT--T CGGCGTAAAA AGTG---ACT -----TAC--  303 Cri.gamb
280  CCC--GAATT AACTACTT-C CGGCGTAAAA CGTGTTTATG GATACAATAA  326 Mac.inge
266  CCC--AAACT AATTATTA-C CGGCGTAAAA CGTGTCACTG GAAGC-----  307 Mic.niva
266  CCC--AAACT AACCAT-AGA CGGCGTAAAA CGTGAATATG GAAAC-----  307 Nes.rufu
268  CCC--AAATT AATTATTCTA CGGCGTAAAA CGTGTCCATA GGAA------  309 1Per.leu
266  CCC--AAACT AATTATTTTA CGGCGTAAAA GGTGTTATAT CT--------  305 Aco.cahi
269  CCC--AAACT AATTATTT-C CGGCGTAAAA CGTGTTATAG ACTT------  309 Cri.migr
269  CCC--AAATT AATAAGTT-T CGGCGTAAAG GGTGTTTAGA T---TATT-C  311 Gli.glis
272  CTC--AAATT AACTACTTTT CGGCGTAAAA CGTGTCCCTG CCTTA-----  314 Ger.nige
267  CCC--AAACT AATTATTT-C CGGCGTAAAA CGTGTTAACT ATAATCAA--  311 Leo.edwa
270  CCC--AAACT AACTATTCTC CGGCGTAAAA TGTGTTTTTA TTACG-----  312 Mes.aura
274  CCC--AAATT AATAAGTC-A CGGCGTAAAG AGTGTTTTAG A---TTTT-A  316 1Mus.ave
266  CCC--AAACT AATTATCT-T CGGCGTAAAA CGTGTCAACT ATAAATAA--  310 Mus.musc
266  CCC--AAACT AATTATTT-C CGGCGTAAAA CGTGTTAACT ATAAGCACTT  312 Mus.paha
265  CTC--AAACT AACTAATT-T CGGCGTAAAA CGTGTTATAA ACTTA-----  306 Tat.kemp
270  CCC--AAGTT AATTATATTT CGGCGTAAAA CGTGCTATAT TAATA-----  312 Ura.rudd
272  CCC--AAATT AATAAGCC-C CGGCGTAAAG AGTGTTTTAG A---TAAT-C  314 Dry.nite
273  CCC--AAATT AATAAGTC-A CGGCGTAAAG AGTGTTTTAG A---TTAT-T  315 2Mus.ave
270  CCC--AAATT AATAAGCT-T CGGCGTAAAG AGTGTTTTAG A---GTTA-A  312 1Gli.jap
268  CCC--AAATT AATAAGCT-T CGGCGTAAAG AGTGTTTTAG A---GTTA-A  310 2Gli.jap
272  CCC--AAACT AATAAACTTC CGGCGTAAAA AATGTTTTAG A---GA----  312 1Hyd.hyd
256  CCC--TAGTT AATAAANC-C CGGCGTAAAA AGTGTTTTGG A---AA----  295 Cav.porc
255  CCC--GAATT AATAAACCAC CGGCGTAAAG AGTGTTTAAA ----AATA-T  297 Bat.suil
256  CCC--AAATT AATAAACCGC CGGCGTAAAG AGTGTTTAAG ----AGTA-T  298 Bat.jane
255  CCC--GAATT AATAAACA-C CGGCGTAAAG AATGTTTAGG ----AATA-T  296 1Cry.hot
255  CCC--AAATT AATAAACA-C CGGCGTAAAG AATGTTTAAG ----AATA-T  296 2Cry.hot
```

```
257  CCC--GAATT AATAAACAAT CGGCGTAAAG AATGTTTAAG ----ATTA-T  299 Cry.dama
256  CCC--AAATT AATAAACA-T CGGCGTAAAG AGTGTTTAGG ----ATTA-A  297 Geo.cape
257  CCC--AAATT AATAAAAA-C CGGCGTAAAG AGTGTTTAAG ----TTAT-A  298 Hel.arge
257  CCC--AAATT AATAAAAC-C CGGCGTAAAG AGTGTTTAGG G---AAAA-C  299 Het.glab
258  CCC--AAACT AATAAACTAA CGGCGTAAAG GGTGTTTAGG A---TTAA-A  301 Thr.swin
243  CCC--TAGTT AATATCAAAA CGGCGTAAAG AGTG--CT-- --AAAAAC-T  283 2Cra.cas
243  CCC--AAGTT AATATTTCTA CGGCGTAAAG AGTG--CT-- --AAAATA-T  283 Dip.ordi
233  CTC--AAGTT AATA-GCATT CGGCGTAAAG CGTG-TTATA TTTATAA---  275 1Ped.cap
233  CTC--AAGTT AATA-GCATT CGGCGTAAAG CGTG-TTATA TTTATAA---  275 2Ped.cap
224  CTC--AAATT AATGGGAAAC CGGCGTAAAA CGTGTTTTAA A---GA---T  265 Ath.macr
221  CCC--TAATT AATAAAACC- CGGCGTAAAG AGTGTTAAAA T---TAAT-A  263 Cte.boli
221  CCC--TAGTT AATAAA-TCC CGGCGTAAAA AGTGTTTTGG A---AC----  260 Cav.guia
216  CCC--GAACT AATAAACTC- CGGCGTATAG AGTGTTTTAG A---AC---A  256 Chi.lani
221  CCC--TAATT AATAAAGCC- CGGCGTAAAG AGTGTTAAAA C---TA---C  261 Cap.pilo
221  CCC--AAACT AATAAA-ACC CGGCGTAAAG AGTGTTTTAG A---AA----  260 Das.punc
230  CCC--AAATT AATAAACACA CGGCGTAAAA GTGTTTTAGG G---AA---G  271 Ere.dors
221  CTC--AAATT AATGGAAAAC CGGCGTAAAG CGTGTTTTAG A---AA---T  262 Hys.afri
221  CCC--TGATT AATAAACCC- CGGCGTAAAG AGTATTAAAG A---TA---C  261 Myo.coyp
219  CCC--TAATT AATAAACCC- CGGCGTAAAG AGTGTTAAAG A---CA---A  259 Oct.degu
223  CCC--TAATT AATAAATTT- CGGCGTAAAG AGTGTTAAAA A---C----T  262 Pro.long
223  CCC--TAGTT AATAATTTTA CGGCGTAAAG CGTGTTTAAG A---TCAC-C  266 Apl.rufa
220  CCC--AAACT AATAAACCC- CGGCGTAAAG CGTGTTTAAG A---ATAT-C  262 Cas.cana
231  CCC--TAGTT AATA-CAAAA CGGCGTAAAG AGTG--CT-- --AAAAAC-T  270 1Cra.cas
217  CTC--TAATT AATAAACTA- CGGCGTAAAG AGTGTTTAAG A---AATA-C  259 Cte.gund
225  CCC--AAATT AATAGTTAT- CGGCGTAAAG CGTGTTTAAG A---CACC-T  267 Das.nove
231  CCC--TAGTT AATA-CAAAA CGGCGTAAAG AGTG--CT-- --AATACC-C  270 Geo.burs
215  CTC--AAACT AACTAATCTT CGGCGTAAAA CGTGTTATAA CATAA-----  257 Ger.vall
217  CCC--ACATT AATGAATTCA CGGCGTAAAG CGTG--TTTA GAGAACCC-C  261 Jac.jacu
222  CCC--AAACT AATTATTAAA CGGCGTAAAA AGTGTTACAC AAGAC-----  264 Lop.flav
218  CCC--AAATT AATTATTATA CGGCGTAAAA CGTGTTCATA GGAAT-----  260 Osg.band
219  CTC--AAGTT AATAAGCATT CGGCGTAAAG CGTG-TTATA TTTATAA---  262 3Ped.cap
230  CCC--TAGTA AATA-AACCC CGGCGTAAAG AGT------- ---ATATA-T  265 Per.flav
219  CCC--AAACT AATAAATTCC CGGCGTAAAG CGTGATTAGA ATTAAAAA-C  265 Syl.audu
225  CCC--TAGTT AATAAAGCA- CGGCGTAAAG CGTGATAAAG GAATAATC-T  270 Sci.nige
224  CCC--TAGTT AATGAAATA- CGGCGTAAAG CGTGATTAAG AGACTAAT-T  269 Spe.trid
222  CCC--AAATT AATAAGCTT- CGGCGTAAAG AGTGTTTAAG A---ATAA-T  264 Gra.muri
223  CCC--AAACT AATAAACTTC CGGCGTAAAA AATGTTTTAG A---GA----  263 2Hyd.hyd
219  CCC--AAGTT AATAAACCTA CGGCGTAAAG GGTGTTTAGG A---ATAA-C  262 Pet.typi
225  CCC--AAATT AATTATTCTC CGGCGTAAAA CGTGTTAATA GATAAC----  268 Ony.leuc
226  CCC--AAATT AATTATTATA CGGCGTAAAA CGTGTCCATA GGAA------  267 Per.poli
227  CCC--GAATT AATTATCC-A CGGCGTAAAA CGTGTCCATA GAAA------  267 Neo.albi
227  CCC--GAATT AATTATCC-A CGGCGTAAAA CGTGTCCATA GAAA------  267 Neo.flor
225  CCC--AAATT AATTATTCTC CGGCGTAAAA CGTGTTAATA GATAAT----  268 Ony.torr
228  CCC--AAATT AATTATTAAA CGGCGTAAAA CGTGTCCATA GGAA------  269 Per.erem
226  CCC--AAATT AATTATTATA CGGCGTAAAA CGTGTCCATA GGAA------  267 Per.goss
```

```
227  CCC--AAATT AATTATTATA CGGCGTAAAA CGTGTCCACA GGAA------  268 Per.mani
227  CCC--AAATT AATTATTATA CGGCGTAAAA CGTGTCCATA GGAA------  268 Per.mela
225  CCCCAAAATT AACTATCACC CGGCGTAAAA CGTGTCCCTA GAGT------  268 Sig.hisp
225  CCCCAAAATT AACTATTTCC CGGCGTAAAA CGTGTCCCTA GANT------  268 Sig.masc
225  CCC--GAATT AATTATTCTC CGGCGTAAAA CGTGTTAATA GACAAC----  268 Ony.aren
226  CCC--AAATT AATTATTCTA CGGCGTAAAA CGTGTCCATA GGAA------  267 2Per.leu
  1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     351        361        371        381        391      400
     |          |          |          |          |        |
311  ----CATAAT AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  356 Rat.norv
309  ---AATAAAT AGAATTAAAA TTCAACTAAT ATGTGAAAAT TCATTGTTAG  355 Hyl.stel
311  ----ATTAAT AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  356 Mus.croc
311  ----ATAAAT AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  356 Mus.cook
309  ---ATTAA-T AGAGTTAAAA TTCAACTAAT ATGTAAAAAT TCATTGTTAG  354 Mas.eryt
311  ATGAATAAAT AGAATTAAAA TCCAACTAAT ATGTGAAAAT TCATTGTTAG  360 Mus.matt
311  AATAAT---- AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  356 Mus.plat
312  ATTAAT---- AGAATTAAAA TCCAACTAAT ATGTGAAAAT TCATTGTTAG  357 Mus.setu
311  AATAAT---- AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  356 Mus.saxi
304  -AT-ACAAAT AGAATTAAAA ATCAACTAAT ATGTGAAAAT TCATCGTTGA  351 Cri.gamb
327  TAACACTAAT AGAATTAAAA TCCAATTTAT ATGTGAAAAT TCATTGTTGG  376 Mac.inge
308  -ACTAAAAAT AGAATAAAAA TCCATCCAAT ATGTGAAAAT TCATCGCTGG  356 Mic.niva
308  CAATATTAAT AGAATTAAAA CTCAGCTAAT AAGTAAAAAT TCATCGCTGA  357 Nes.rufu
310  -ACAACTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  358 1Per.leu
306  -AAAACAAAT AGAATTAAAA CCTGACTTAT ATGTGAAAAT TCATTGTCAG  354 Aco.cahi
```

```
310  --AACATAAT AGAATTAAAA CCCAACTAAT ATGTGAAAAT TCATAGTTGG  357 Cri.migr
312  CT---ACAAT AAGGTTAAAT TCTGTCTAAG CTGTAAAAAG CCTTTGCCAG  358 Gli.glis
315  ---CCATAAT AGAATTGAAA ACCAACAAAT GCGTAGAAGT TTATCGTTCG  361 Ger.nige
312  ----AACAAT AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  357 Leo.edwa
313  --AACATAAT AGAATTAAAA CCCAACTAAT ATGTGAAAAT TCATTGTTGG  360 Mes.aura
317  ACT--ATAAT AAAGTTAAAC TTTAACTAAG CCGTAAAAAG CC-CTAGTTA  363 1Mus.ave
311  ----ATAAAT AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  356 Mus.musc
313  ATACATAAAT AGAATTAAAA TCCAACTTAT ATGTGAAAAT TCATTGTTAG  362 Mus.paha
307  ---CCCTAAT AGAATTGAAA CCCAACTAAT ACGTAGAAGT TCATCGTTCG  353 Tat.kemp
313  ---TATAAAT AGAATTAAAA CTCAACTTAT ATGTGAAAAT TCATTGTTAG  359 Ura.rudd
315  CACA-AAAAT AAAGTTAAAT CTTGTCTAAG CCGTAAAAAG CAA-TAGACA  362 Dry.nite
316  ACC--TTAAT AAAGTTAAAC TTTAACTAAG CTGTAAAAAG CCC-TAGTCA  362 2Mus.ave
313  CT---ACAAT AAGGTTAAAT TTTATCTAAG CCGTAAAAAG CCC-TAGCTA  358 1Gli.jap
311  CTT--ATAAT AAGGTTAAAT TTTATCTAAG CCGTAAAAAG CCC-TAGCTA  357 2Gli.jap
313  TATAAAAAAT AAGATTAAAC TCTATCTAAG TTGTAAAAAA CAC-CAGATA  361 1Hyd.hyd
296  TATAAAAAAT AAGGCTAATC TTTGTCTAAG TTGTAGAAAA CTC-TAGACA  344 Cav.porc
298  AACAAAAC-T AAGATTAAAT TCTGTCTATG TTGTAAAAAA CCA-CAGACA  345 Bat.suil
299  AACAAAAC-T AAGATTAAAT TCTGTCTCTG TTGTAAAAAA CCA-CAGACA  346 Bat.jane
297  GATCAAAC-T AAGATTAAAA TTTATCTATA TCGTAAAAAA TTC-TAGACA  344 1Cry.hot
297  AACCAAAC-T AAGATTAAAT CTTGTCTATA TCGTAAAAAA TTT-TAGACG  344 2Cry.hot
300  AACAAAAC-T AAGATTAAAT CTTGTCTATG TCGTAAAAAA TTA-TAGATA  347 Cry.dama
298  AATTTAAC-T AAGATTAAAT TCTATTTATG TCGTAAAAAA CCA-CAAATA  345 Geo.cape
299  AATTGAAC-T AAGATTAAAT CTTATCTAAG TTGTAAAAAA CTAATAGATA  347 Hel.arge
300  AATCAGAC-T AAGATTAAAT TCTATCTAGG TCGTAAAAAA CTTCAGATAA  348 Het.glab
302  AACAACAAAT AAAATTAAGA TCTACCTGGG TTGTAAAATG CTT-CAGGTA  350 Thr.swin
284  TT-ACCAAAT AAAACTAAGC CTAATTGAAG CTGTAGAAAG CACTAATTAG  332 2Cra.cas
284  TT-ATGAAAT AGAGCTAAGA CTTGTCTTGG CTGTAAAAAG CCAAAGCCAA  332 Dip.ordi
276  -GATAAGATT AAAGTTAAGC TATAACTAAG CCGTAAAAAG CCATAGTTA-  323 1Ped.cap
276  -GATAAGATT AAAGTTAAGC TATAACTAAG CCGTAAAAAG CCATAGTTA-  323 2Ped.cap
266  AAAAACTAAT AAGATTAAGT TCTATTTAGG TCGTAAAAAA CTT-CAAATA  314 Ath.macr
264  AATCAAAA-T AAGATTAAAT TTCATCTAAG TTGTAAAAAA CTA-TAGACA  311 Cte.boli
261  TATAAAAAAT AAGACTAATC CTTGTCTAAG TTGTAGAAAA CTC-TAGACA  309 Cav.guia
257  ATAAACAA-T AAGATTAAAC TTTATTTAAG TTGTAAAAGA CTC-TAAATA  304 Chi.lani
262  AAAAACAA-T AAGATTAAAT TTTACCTAAG TCGTAAAAAA CAA-CAGATA  309 Cap.pilo
261  AACATAAAAT AAGACTAAAA TTTATCTAAG TCGTAGAAAA CAC-CAGATA  309 Das.punc
272  ACCGAAAA-T AAGACTAAAC TTCATCTAAG TCGTAAAAAA CTT-TAGATA  319 Ere.dors
263  AAAAATTAAT GAGATTAAGT TCTATCTAGG TCGTAAAAAA CTC-CAGATT  311 Hys.afri
262  AATAAAAA-T AAGATTAAAT TTCATCTGGG TCGTAAAAAA CTA-TAGATA  309 Myo.coyp
260  AACTAAAA-T AAGATTAAAG TTTACCCAAG TCGTAAAAAA CTA-CAGGTA  307 Oct.degu
263  AAACAAAA-T AAGACTAAAT TTTACCCAAG TCGTAAAAAA CAT-CAGGTA  310 Pro.long
267  ACC--ATAAT AAGGTTAAGA CCAAACTAAG CTGTAAAAAG CCTCTAGCTT  314 Apl.rufa
263  CACCCTCAAT AAGGATAAAC TTGACCTAAG CTGTAAAAAG CTA-TAGTTG  311 Cas.cana
271  TTTACCAAGT AAAACTAAAG CCTAATTGAG CTGTAGAAAG CACTAATTAG  320 1Cra.cas
260  CAAA-TAATT AAGAGAAAAC TTAGTTTAAG TCGTAAAAAA CC--TTAACT  306 Cte.gund
268  AG---ACAAT AGAGTTAAAC CCTTACTACG CTGTAAAAAG CCTTAGTAGG  314 Das.nove
```

```
271  TTAATCAGGT AAAACTAAAG TCTAATTAAG CTGTAGAAAG CACCAATTAG  320 Geo.burs
258  ---TTAAAAT AGAATTAAAA CCCAACCAAT ACGTTTAAGT TCATCGTTAG  304 Ger.vall
262  CCCCCCCAAT AAGGCTAAAC TCCGACTAAT ATGTAGAAAT TCATAGTCTG  311 Jac.jacu
265  ---AAAAAAT AGAATTAAAA TTCAACTTAT ATGTGAAAAT TCATTGTTAG  311 Lop.flav
261  -AAAACTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  309 Osg.band
263  -GATAAGATT AAAGTTAAGC TATAACTAAG CCGTAAAAAG CCATAGTTA-  310 3Ped.cap
266  TCCAATTAGT AAAGCTAAAA TTAGGCTAAT CTGTAAAAAG TCTTAGCCCA  315 Per.flav
266  AGAAAAAAAT AAAATCAAAT AACAACTAAG CTGTAAAAAG TAATAGTTAC  315 Syl.audu
271  AA----AGAT AAGATTAAGC TCTAACTAAG CTGTAAAAAG CCTTAGTTAA  316 Sci.nige
270  ------AGAT AAGATTAAAA TTATACTAAA CTGTAAAAAG TCTTGGTGTT  313 Spe.trid
265  ATGACAAAAT AAGGTTAAAC TCTAACTAAA CTGTAAAAAG TCA-TAGACA  313 Gra.muri
264  TATAAAAAAT AAGATTAAAC TCTATCTAAG TTGTAAAAAA CAC-CAGATA  312 2Hyd.hyd
263  CATCATAA-T AAGATTAAAT TTTATCTAAG TTGTAGAACA CTA-CAGATA  310 Pet.typi
269  -AAATTTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  317 Ony.leuc
268  -ACAACCAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  316 Per.poli
268  -CAAAAAAAT AGAATTAAAA ATCAACCAAT ATGTGAAAAT TCATCGCTGA  316 Neo.albi
268  -CAAAAAAAT AGAATTAAAA ATCAACCAAT ATGTGAAAAT TCATCGCTGA  316 Neo.flor
269  -AAATTTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  317 Ony.torr
270  -AAAACAAAT AGAATTAAAA ACCAGCTAAT ATGTGAAAAT TCATAGATGG  318 Per.erem
268  -ACAACTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  316 Per.goss
269  -ACAATCAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  317 Per.mani
269  -ACAACCAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  317 Per.mela
269  -ACCCCCAAT AGAATTAAAA TTCAACCAAT ATGTGAAAAT TCATCGTTGA  317 Sig.hisp
269  -TTCCCTAAT AGAATTAAAA TCCAACCAAT ATGTGAAAAT TCATCGTTGA  317 Sig.masc
269  -AAATTTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  317 Ony.aren
268  -ACAACTAAT AGAATTAAAA ACCAACCAAT ATGTGAAAAT TCATCGTTGG  316 2Per.leu
  1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
```

```
  1   ---------- ---------- ---------- ---------- ----------     0 Urs.amer
  1   ---------- ---------- ---------- ---------- ----------     0 Myd.marc

      401        411        421        431        441      450
      |          |          |          |          |        |
357   GACCTAAGCC CAATAACGAA ---------- AGTAATTCTA ATCATTTAT-   395 Rat.norv
356   AACCTAAACA CAGTAACGAA ---------- GGTAATTCTA GATATTCTCA   395 Hyl.stel
357   GACCTAAACA CAATAACGAA ---------- AGTAATTCTA ATTATTT---   393 Mus.croc
357   GACCTAAACT CAATAACGAA ---------- AGTAATTCTA ATAATTT---   393 Mus.cook
355   AACTTAAACA CGACAACGAA ---------- AGTAATTCTA AACATTTC--   392 Mas.eryt
361   GACATAAATT CAATAACGAA ---------- AGTAATTCTA ATAGTTTT--   398 Mus.matt
357   GACCTAAATT CAATTACGAA ---------- AGTAATTCTA ATAATTG---   393 Mus.plat
358   GACCTAAATT CAATTACGAA ---------- AGTAATTCTA ATCATTT---   394 Mus.setu
357   GACCTAAATT CAATTACGAA ---------- AGTAATTCTA ATAATTG---   393 Mus.saxi
352   ATCTAAAATC ACTAACGAAA ---------- -GTAATTCTA ATCAATTT--   388 Cri.gamb
377   ACTTAAAATC AATAACGAAA ---------- -GTAATTCTA -GAAGCCC--   412 Mac.inge
357   ACCTAAGACC AATGACGAAA ---------- -GTAATTCTA ATTAACCTT-   394 Mic.niva
358   CTATAAAATC AATAACGAAA ---------- -GTAATTCTA -CCAATCC--   393 Nes.rufu
359   GATTAAACTC AATAACGAAA ---------- -GTAATTCTA ATTATCTTA-   396 1Per.leu
355   ACAAAAACAC AACCACGAAA ---------- -GTGATTCTA ATAAACC---   390 Aco.cahi
358   TATTAAACTC AATGACGAAA ---------- -GTAATTCTA ATATTTTAAT   396 Cri.migr
359   AG-AAAAAAT AACATACGAA ---------- AGTCACCTTA CTAAAA--TC   395 Gli.glis
362   GGCCTAAAAT CACCCACGAA ---------- AGTAATTCTA ATATATAC--   399 Ger.nige
358   GACCTAAGCC CAATAACGAA ---------- AGTAATTCTA ACCTCTAAT-   396 Leo.edwa
361   ACTTAAAATC AATAACGAAA ---------- -GTAATTCTA ATTATATTA-   398 Mes.aura
364   AAAGTAAAAT AACGAACGAA ---------- AGTCACTTTA ATATTC--TC   401 1Mus.ave
357   GACCTAAACT CAATAACGAA ---------- AGTAATTCTA GTCATTT---   393 Mus.musc
363   GACTTAAACA CAATAACGAA ---------- AGTAATTCTA ATTGTTT---   399 Mus.paha
354   GATCTAAAAT CACTAACGAA ---------- AGTAATTCTA GTAAATCT--   391 Tat.kemp
360   AATTTAAAAT CAATAACGAA ---------- GGTAATTCTA ATAT--AT--   395 Ura.rudd
363   AAGATAAGAC AAAGCACGGA ---------- AGTCACTTTA ATATTA-CTC   401 Dry.nite
363   AAAGTAAAGT AATGAACGAA ---------- AGTCACTTTA GTATTT--TC   400 2Mus.ave
359   AAGGAAAAAT AATATACGAA ---------- AGTCACCTTA TAACAA--TC   396 1Gli.jap
358   AAGGAAAAAT AATACACGAA ---------- AGTCACCTTA TAAAAA--TC   395 2Gli.jap
362   AAATGTAAGT CAATAACGAA ---------- AGTAATCTTA ATACCT---C   398 1Hyd.hyd
345   TGGTAGAAAT CATAAACGAA ---------- AGTAGCTTTA ATAAGT---C   381 Cav.porc
346   GATGTAAACC CAATAAACGC ---------A AGTAATCTTA TTGTAT---C   383 Bat.suil
347   GATGTAAACC CAATAAACGC ---------A AGTAATCTTA TTATAT---C   384 Bat.jane
345   AATCAAAACT CAATAAACGC ---------A AGTAATCTTA TAATAT---C   382 1Cry.hot
345   AATTATAACT CAATAAACGT T--------A AGTAATCTTA TAAAAT---C   383 2Cry.hot
348   AGACGAAACT CAGAATAAAC GC-------A GGTAATCTTA TAATAT---C   387 Cry.dama
346   GACTCAAACT CAATAAATGT ---------A AGTAATCTTA TTATAT---C   383 Geo.cape
348   TTTAAAATTT CAATAAACGA ---------A AGTAATCTTA TTAAAC---C   385 Hel.arge
349   GACAAACCCA CTACTATCTA CG-------A AGTAATCTTA ACAAAC---C   388 Het.glab
351   AAATAAAAAT CAACACTCCC TCCAACGA-A AGTGATTTTA GCCTTT---C   396 Thr.swin
```

```
333  AACGAAAAAT ACATCACGAA ---------- AGTAGTTTTA TATA-----A  367 2Cra.cas
333  AGTAAAAA-C CAATAACGAA ---------- AGTAACTCTA ATTT-----T  366 Dip.ordi
324  GTATAAAAAT AAACAACGAA ---------- GGTGACTTTA TAAGTC---T  360 1Ped.cap
324  GTATAAAAAT AAACAACGAA ---------- AGTGACTTTA TAAGTC---T  360 2Ped.cap
315  GTTACAAAAC CAATACGCGA ---------A AGCAATCTCA ATAACT---T  352 Ath.macr
312  AAAATAAAAC CATGCACGAA ---------- AGTAATCTTA TTATAA---T  348 Cte.boli
310  CGGTAGAAAC CATAAACGAA ---------- AGTAGTTTTA ATAAGT---C  346 Cav.guia
305  AAAGTAAACC CAACAAACGA ---------- AGTAATCTTA ACCCAT---C  341 Chi.lani
310  AAAATAAAAT CATAAACGAA ---------- AGTAATCTTA ATAAAT---C  346 Cap.pilo
310  AAATACATAC TG-AAACGAA ---------- AGTAGTCTTA CTATAT---C  345 Das.punc
320  AATTTAAAAT CACAAACGAA ---------- AGTAACCTTA ATTTAC---C  356 Ere.dors
312  GATACAAAAC CAGTACGCGA ---------A AGCAATCTCA ATAATT---C  349 Hys.afri
310  AAAATAAAAT CGATAACGAA ---------- GGTAATCTTA ATATAT---C  346 Myo.coyp
308  AAAAATANNT CCACATACGA ---------- AGTTATCTTA GTACAC---C  344 Oct.degu
311  AAAACAAGCC CGTTAACGAA ---------- AGTAGTCTTA ATAATC---A  347 Pro.long
315  CAGTAAAACC CCACCACGAA ---------- AGTAACCTTA GTATCC----  350 Apl.rufa
312  AAGAAAAAAT AACCCACGAA ---------- AGTGATCTTA AAATAAAT-C  350 Cas.cana
321  AACGAAAA-T ACATCACGAA ---------- AGTAGTTTTA TATA-----A  354 1Cra.cas
307  AAAGTAAGCC AGCCAACGAA ---------- AGTACTCTTA TATAGT---C  343 Cte.gund
315  ACCATAAACC CTTCAACGAA ---------- AGTGACTCTA ATTTAT---C  351 Das.nove
321  AGCGAAAA-T AAACCACGAA ---------- GGTGATTTTA TCAAC----A  355 Geo.burs
305  GGCCTAAACC CACAAACGAA ---------- AGTAATTCTA ATTTAT---C  341 Ger.vall
312  AAGTAAAA-T AAACTACAAA ---------- AGTGGTCTTA AAATTC---T  347 Jac.jacu
312  AACCTAAACA CAATAACGAA ---------- AGTAATTCTA ATATCT---A  348 Lop.flav
310  GATTAAATTC AATAACGAAA ---------- -GTAATTCTA ATCAA----C  344 Osg.band
311  GTATAAAAAT AAACAACGAA ---------- AGTGACTTTA TAAGTC---T  347 3Ped.cap
316  AATGAAAA-A CTACCTCGAA ---------- AGAGGCTTTA TTAA-----A  349 Per.flav
316  AAACAAAAAT AAACAACGAA ---------- AGTGATTTTA TAGCCT---T  352 Syl.audu
317  AATAAAAATA TAAAACGAAA ---------- -GTGGTCTTA TAACTC---C  352 Sci.nige
314  AATGAAAATC AAATACGAAA ---------- -GTAATCTTA AATTTT---C  349 Spe.trid
314  GAGAAAAAAT AACATACGAA ---------- AGTAACCTTA CTATATATTC  353 Gra.muri
313  AAATGTAAGT CAATAACGAA ---------- AGTAATCTTA ATACCT---C  349 2Hyd.hyd
311  AAATTAAACT CAGT--TAAC CCCAACGAA- AGTAATCTTA CCACAT---C  354 Pet.typi
318  AATTAAAACC AGTAACGAAA ---------- -GTAATTCTA ATAAA----C  352 Ony.leuc
317  AATTAAAACC AATAACGAAA ---------- -GTAATTCTA ATTAA----C  351 Per.poli
317  ACTTAAAAAC ACTAACGAAA ---------- -GTAATTCTA ACTAA----C  351 Neo.albi
317  ACTTAAAAAC ACTAACGAAA ---------- -GTAATTCTA ACTAA----C  351 Neo.flor
318  AATTAAACCC AGTAACGAAA ---------- -GTAATTCTA ATAAA----C  352 Ony.torr
319  TACTAAAATC AGTAACGAAA ---------- -GTAATTCTA ATTAA----C  353 Per.erem
317  GATTAAACTC AATAACGAAA ---------- -GTAATTCTA ATTAT----C  351 Per.goss
318  GACTAAAACC AATAACGAAA ---------- -GTAATTCTA ATTAG----C  352 Per.mani
318  GACTAAAATC AATAACGAAA ---------- -GTAATTCTA ATTAA----C  352 Per.mela
318  AGTTAAAATC AGAAACGAAA ---------- -GTTATTCTA GTAGA----C  352 Sig.hisp
318  GATTAAAATC AGAAACGAAA ---------- -GTTATTCTA GTAGA----C  352 Sig.masc
318  AACTAAAACC AATAACGAAA ---------- -GTAATTCTA ATAAA----C  352 Ony.aren
```

```
317  GATTAAACTC AATAACGAAA ---------- -GTAATTCTA ATTAT----C  351 2Per.leu
  1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     451        461        471        481        491      500
     |          |          |          |          |        |
396  -ATAATGCAC GATAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  443 Rat.norv
396  C-TAAAACAC GATAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  443 Hyl.stel
394  -ATAACACAC GACAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  441 Mus.croc
394  -ATAAGACAC GACAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  441 Mus.cook
393  -ATAAAACAC GATAGCTAAG -ATCCAAACT GGGATTAGAT ACCCCACTAT  440 Mas.eryt
399  -ATAATACAC GACAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  446 Mus.matt
394  --TAATACAC GACAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  440 Mus.plat
395  -ATAATACAC GACAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  442 Mus.setu
394  --TAATACAC GACAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  440 Mus.saxi
389  --TAAGCCAC GAAAGCTAAG -ACCCAAACT GGGATTAGAT ACCCCACTAT  435 Cri.gamb
413  -TAAAACACG AAAACTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  459 Mac.inge
395  ---ATGCACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  439 Mic.niva
394  -TAATGCACG AAAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  440 Nes.rufu
397  ---ATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  441 1Per.leu
391  ---AAACACG AAAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  435 Aco.cahi
397  ---ATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  441 Cri.migr
396  TGAATACACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  443 Gli.glis
400  ---ATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  444 Ger.nige
397  ACAACACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  444 Leo.edwa
399  ---ATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  443 Mes.aura
```

```
402 TGAGTACACG ATAGCTAAG- -ACACAAACT GGGATTAGAT ACCCCACTAT 449 1Mus.ave
394 ATAATACACG ACAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 441 Mus.musc
400 ATAATACACG ACAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT 447 Mus.paha
392 ---ATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT 436 Tat.kemp
396 -CAATGCACG ACAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT 442 Ura.rudd
402 TGAACACACG ATAGCTAAG- -ATACAAACT GGGATTAGAT ACCCCACTAT 449 Dry.nite
401 TGAACACACG ATAGCTAAG- -ACACAAACT GGGATTAGAT ACCCCACTAT 448 2Mus.ave
397 TGAATACACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 444 1Gli.jap
396 TGAATACACG ATAGCTAAG- -ACCCAAACT GAGATTAGAT ACCTCACTAT 443 2Gli.jap
399 TGAATACATG AAAGCTAAG- -GCTCAAACT GGGATTAGAT GCCCCACTAT 446 1Hyd.hyd
382 CGA-CACACG AAAGCTAAG- -GCCCAAACT GGGATTAGAT ACCCCACTAT 428 Cav.porc
384 TGAACACACG ACAGCTAAG- -ACCCAAA-- ---------- ---------- 409 Bat.suil
385 TGAACACACG ACAGCTGAG- -ACCCAAA-- ---------- ---------- 410 Bat.jane
383 TGAACACATG ACAGCTCAG- -ACCCAAA-- ---------- ---------- 408 1Cry.hot
384 TGAATACATG ACAGCTCAG- -GCTCAAA-- ---------- ---------- 409 2Cry.hot
388 TGAACACATG ACAGCTTGG- -GTACAAA-- ---------- ---------- 413 Cry.dama
384 TGAACACACG AAAGCTAAG- -ATTCAAA-- ---------- ---------- 409 Geo.cape
386 TGAATACCCG ACAGTTAAG- -ACCCAAA-- ---------- ---------- 411 Hel.arge
389 TGAACACACG ACAGCTATG- -GT-CAAA-- ---------- ---------- 413 Het.glab
397 TGAATACACG AAAACTAAG- -GCCCAAA-- ---------- ---------- 422 Thr.swin
368 CTGCAGCACT AAAGCTATAT AGCACAAACT GGGATTAGAT ACCCCACTAT 417 2Cra.cas
367 TTGTAGCACT AAAGCTATT- -ACCTAAACT GGGATTAGAT ACCCCACTAT 414 Dip.ordi
361 AGATTACACG ACAGCTAGA- -TCCCAAACT GAGATTAGAT ACCTCACTAT 408 1Ped.cap
361 ATATTACACG ACAGCTAGA- -TCCCAAACT GAGATTAGAT ACCTCACTAT 408 2Ped.cap
353 TGAACACACG ACATCTAAG- -ACTCAAACT GGGATTAGAT ACCCCACTAT 400 Ath.macr
349 TAAACACACT AAAGCTAAG- -GTACAAACT GGGATTAGAT ACCCCACTAT 396 Cte.boli
347 CGA-CACACG AAAGCTAAG- -GCCCAAACT GGGATTAGAT ACCCCACTAT 393 Cav.guia
342 TGAACACACG AAAGCTACG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 389 Chi.lani
347 TGAATACACT AAAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 394 Cap.pilo
346 TGAATACACG AAAGCTAGA- -ATCCAAACT GGGATTAGAT ACCCCACTAT 393 Das.punc
357 TGAACACACG AAAGCTAGG- -GCACAAACT GGGATTAGAT ACCCCACTAT 404 Ere.dors
350 TGAATACACG ATACCTAAG- -ACTCAAACT GGGATTAGAT ACCCCATTAT 397 Hys.afri
347 AGAATATACT AAAGCTAAG- -ACACAAACT GGGATTAGAT ACCCCACTAT 394 Myo.coyp
345 TGAATGCACT AAAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT 392 Oct.degu
348 TGAAAACACC AAAGCTGAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT 395 Pro.long
351 TGAAAACACG ACAGCTAAG- -ACCTAAACT GGGATTAGAT ACCCCACTAT 398 Apl.rufa
351 TGATAACACG ACAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 398 Cas.cana
355 CTGCAGCACT AAAGCTATA- -GCACAAACT GGGATTAGAT ACCCCACTAT 402 1Cra.cas
344 TGAATACACG ATAGCTAAG- -GCCCAAACT GGGATTAGAT ACCCCACTAT 391 Cte.gund
352 TGACTACACG ATAGCTAGG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 399 Das.nove
356 CTGCAGCACT AAAGCTATA- -GAACAAACT GGGATTAGAT ACCCCACTAT 403 Geo.burs
342 T-AAACACGA ACAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 388 Ger.vall
348 TGAAGACACG ATAGCTAGG- -GGACAAACT GGGATTAGAT ACCCCATTAT 395 Jac.jacu
349 TCCATACACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT 396 Lop.flav
345 TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT 392 Osg.band
```

```
348  ATATTACACG ACAGCTAGA- -TCCCAAACT GAGATTAGAT ACCTCACTAT  395 3Ped.cap
350  CTATTATACT AAAGCTATA- -GCACAAACT GGGATTAGAT ACCCCACTAT  397 Per.flav
353  CGAATTCACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  400 Syl.audu
353  TGACTTCACG ATAGCTAAG- -GCCCAAACT GGGATTAGAT ACCCCACTAT  400 Sci.nige
350  TGAATTCACG ATAGCTAAG- -ACTCAAACT GGGATTAGAT ACCCCACTAT  397 Spe.trid
354  TGAATACACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  401 Gra.muri
350  TGAATACATG AAAGCTAAG- -ACTCAAACT GGGATTAGAT ACCCCACTAT  397 2Hyd.hyd
355  TGAACACACG AAAACTAAG- -GCCCAAA-- ---------- ----------  380 Pet.typi
353  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Ony.leuc
352  TTAACACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  399 Per.poli
352  TTAATACACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  399 Neo.albi
352  TTAATACACG ATAGCTAAG- -ACCCAAACT GGGATTAGAT ACCCCACTAT  399 Neo.flor
353  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Ony.torr
354  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  401 Per.erem
352  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  399 Per.goss
353  TTAACACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Per.mani
353  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Per.mela
353  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Sig.hisp
353  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Sig.masc
353  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  400 Ony.aren
352  TTAATACACG ATAGCTAAG- -ATCCAAACT GGGATTAGAT ACCCCACTAT  399 2Per.leu
  1  ---------- ---------- ---------- ---------- ----------    0 Dol.pata
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.leu
  1  ---------- ---------- ---------- ---------- ----------    0 1Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 2Pet.pet
  1  ---------- ---------- ---------- ---------- ----------    0 1Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 2Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0 3Pte.vol
  1  ---------- ---------- ---------- ---------- ----------    0  Sci.lis
  1  ---------- ---------- ---------- ---------- ----------    0 Sci.vulg
  1  ---------- ---------- ---------- ---------- ----------    0 1Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 2Tam.sib
  1  ---------- ---------- ---------- ---------- ----------    0 Bas.astu
  1  ---------- ---------- ---------- ---------- ----------    0 Spi.puto
  1  ---------- ---------- ---------- ---------- ----------    0 Pro.loto
  1  ---------- ---------- ---------- ---------- ----------    0 Mep.meph
  1  ---------- ---------- ---------- ---------- ----------    0 Odo.rosm
  1  ---------- ---------- ---------- ---------- ----------    0 Can.latr
  1  ---------- ---------- ---------- ---------- ----------    0 Uro.cine
  1  ---------- ---------- ---------- ---------- ----------    0 Urs.amer
  1  ---------- ---------- ---------- ---------- ----------    0 Myd.marc

     501        511        521        531        541     550
     |          |          |          |          |        |
```

```
444  GCTTA-GCCC TAAA-CCTTA ATAATTAAA- --CCTACAAA ATTATTTGCC  488 Rat.norv
444  GCTTA-GCCC TAAA-CCATA ATAATTCCG- ---TAACAAA ATTATTTGCC  487 Hyl.stel
442  GCTTA-GCCA TAAA-CCTAA ATATTTTGA- --CTAACAAA ACTATTTGCC  486 Mus.croc
442  GCTTA-GCCA TAAA-CCTAA ATAACTTAA- --TTAACAAA ATTATTTGCC  486 Mus.cook
441  GCTTA-GCCC TAAA-CTATA ATAATTGTA- ---TAACAAA ATTATTTGCC  484 Mas.eryt
447  GCTTA-GCCA TAAA-CTTAA ATACTTTTA- --TTAACAAA ATTATTTGCC  491 Mus.matt
441  GCTTA-GCCA TAAA-CTTAA ATATTTAAT- --AAAACAAA ACTATTTGCC  485 Mus.plat
443  GCTTA-GCCA TAAA-CCTAA ATAATTTAA- --TTAACAAA ACTATTTGCC  487 Mus.setu
441  GCTTA-GCCA TAAA-CTTAA ATATTTAAT- --AAAACAAA ACTATTTGCC  485 Mus.saxi
436  GCTTA-GCCC TAAA-CTTAA ATAGTTATT- ---TAACAAA ACTATTTGCC  479 Cri.gamb
460  GCTTA-GCCG TAAA-CTTGA ATAATTATT- ---TAACAAA ATTATTTGCC  503 Mac.inge
440  GCTTA-GCCC TAAA-CCTCA ATAATTCAA- ---AAACAAA ACTATTTGCC  483 Mic.niva
441  GCCTA-GCCA TAAA-CTTAA ATAATTAAA- ---TAACGAA ATTATTTGCC  484 Nes.rufu
442  GCTTA-GCCC TAAA-CCTTA AAGATTAAA- ---TAACAAA ATCATTTGCC  485 1Per.leu
436  GCTTA-GCCC TAAA-CTTAA ATAATTAA-- ---TAACAAA ATTATTTGCC  478 Aco.cahi
442  GCTTA-GCCT TAAA-CCTTA AGGGTTAAA- CTATAACAAA ACCCTTTGCC  488 Cri.migr
444  GCTTA-GCCC TAAA-CATAA ACCCTT---- -ACTAACGCA ACCGTTCGCC  486 Gli.glis
445  GCTTA-GCCC TAAA-CATAA ATAATTCCT- ---TAACAAA ATTATTTGCC  488 Ger.nige
445  GCTTA-GCCC TAAA-CCTAA ATAATTCAA- --CCCACAAA AATATTTGCC  489 Leo.edwa
444  GCTTA-GCCC TAAA-CCTAA GTGATTAAA- ---TAACAAA ATCACTTGCC  487 Mes.aura
450  GCTTA-GCCC TAAA-CATAA ACCCTT---- -ATTAACCTA ACTGTTCGCC  492 1Mus.ave
442  GCTTA-GCCA TAAA-CCTAA ATAATTAAA- -TTTAACAAA ACTATTTGCC  487 Mus.musc
448  GCTTA-GCCA TAAA-CCTAA ATAATTTAA- --CAAACAAA ACTATTTGCC  492 Mus.paha
437  GCTTA-GCCC TAAA-CTTAA ATAATTTTT- --ATAACAAA ATTATTTGCC  481 Tat.kemp
443  GCTTA-GCCG TAAA-CTTAA ATAATTTAT- ----AACAAA ATTATTTGCC  485 Ura.rudd
450  GCTTA-GCCC TAAA-CATAA ATTATTC--- -GCAAACACA AATATTCGCC  493 Dry.nite
449  GCTTA-GCCC TAAA-CATAA ACTCTT---- -ATTAACCTA ATTGTTCGCC  491 2Mus.ave
445  GCTTA-GCCA TAAA-CAAAA ATAACT---- -AATAACACA GTTATTCGCC  487 1Gli.jap
444  GCTTA-GCCA TAAA-CAAAA ATAACT---- -AATAACACA GTTATTCGCC  486 2Gli.jap
447  GCTTA-GCCA TAAA-CATAA AAGTTCA--- -CATAACAAG AACTTTCGCC  490 1Hyd.hyd
429  GCTTA-GCCA TAAA-CATAA ANACTTA--- -TACAACAAA AGATTTCGCC  472 Cav.porc
410  GCT-AAGCTA TAAA-CACAG AACATTAAA- -TTAAACAAA ATGTTTCGCC  455 Bat.suil
411  GCT-AAGCTA TAAA-CACAG AACATTAAA- -CGAAACAAA ATGTTTCGCC  456 Bat.jane
409  GCT-GAGCCG TAAA-CATAA ATTATTAAA- -TGAAACAAA ATATTTCGCC  454 1Cry.hot
410  GCT-AAGCCG TAAA-CACAA ATTATTGAG- -CAAAACAAA ATATTTCGCC  455 2Cry.hot
414  GCC-AAGCCG TAAA-CACAA AATATTAAAT -AGAACGGAA AAGTTTCGCC  460 Cry.dama
410  GCT-AAGCTG TAAA-CACAA AATATTAAAT -AGAA-CAAA ACATTTCGCC  455 Geo.cape
412  GCTTAACCTA TAAA-CATAA AATATTATAC TCATAAACAA ATATTTCGCC  460 Hel.arge
414  GCTTA-GCCG TAAAACATAA AAAATATTAT -AACAACAAA TA-TTTCGCC  460 Het.glab
423  GCT-TAGTCG TAAAA-ATAA AATATT---- -CATAACAAA ATATTTCGCC  465 Thr.swin
418  AATTA-GCCA TAAA-CATAA GTAGATAT-- -AGTAACAAA TTTATTCGCC  462 2Cra.cas
415  AATTA-GCCA TAAA-CATAA GTAGTT---- -CATAACATT ACTACTCGCC  457 Dip.ordi
409  GTCTA-GCCG TAAA-CTTAA ACTATTAAA- -TTCAACAAA ATAGTTCGCC  454 1Ped.cap
409  GTCTA-GCCG TAAA-CTTAA ACTATTAAA- -TTCAACAAA ATAGTTCGCC  454 2Ped.cap
401  GCTTA-GACG TAAA-CATAA AAAATT---- -CATAACAAA ATTTTTCGCC  443 Ath.macr
```

```
397  GCTTA-GCCA TAAA-CAAAG ACATTT---- -AAAAACAAA AATGTTCGCC  439 Cte.boli
394  GCTTA-GCCA TAAA-CACAA AAACTTA--- -CATAACAAA AGATTTCGCC  437 Cav.guia
390  GCTTA-GCCA TAAA-CATAA AAACTC---- -CACAACAAA AGTTTTCGCC  432 Chi.lani
395  GCTTA-GCCA TAAA-CAAAG ATATTT---- -AATAACAAA AATTTTCGCC  437 Cap.pilo
394  GTCTA-GCCA TAAA-CATAA AAATTT---- -TATAACAAA AATTCTCGCC  436 Das.punc
405  GCCTA-ACTA TAAA-CATAA AAATTT---- -CCCAACAAA AATTTTTGCC  447 Ere.dors
398  GCTTA-GGCG TAAA-CACAA AAAATT---- -CATAACAAA ATTATTCGCC  440 Hys.afri
395  GCTTA-GTTG TAAA-CACAG ATTCTT---- -AACAACAAA AATATCCGCC  437 Myo.coyp
393  GCTTA-GCCA TAAA-CACAG ACATTT---- -AACAACAAA AATGTTCGCC  435 Oct.degu
396  GCTTA-GCCA TAAA-CACAA ATATTT---- -TATAACAAA AATATTCGCC  438 Pro.long
399  ACTTA-GCCC TAAA-CATAA ATTTTCC--- -TATAACAAG AAAATTCGCC  442 Apl.rufa
399  GCTTA-GCCG TAAA-CTTAG ATATGTAAA- -CCTAACAAA CATATTCGCC  444 Cas.cana
403  GTATA-GCCA TAAA-CATAA ATAGATAT-- -AGTAACAAA TTTATTCGCC  447 1Cra.cas
392  GCTTA-GCCG TAAA-CAATA ACATTTA--- -TACAACAAA AATCTTCGCC  435 Cte.gund
400  GCCTA-GCCC TAAA-CTAAA ACAGTTC--- -ACAAACAAA ACTGTTCGCC  443 Das.nove
404  GTATA-GCCG TAAA-CTTAA ATAGATTA-- -AATAACAAA TCTATTCGCC  448 Geo.burs
389  GCTTA-GCCG TAAA-CTTAA ATAATTTC-- -AATAACAAA ATTATTTGCC  433 Ger.vall
396  GCCTA-GCCC TAAA-CAGAG ATAGCCA--- -ACTAATAAG GCTACCCGCC  439 Jac.jacu
397  GCTTA-GCCC TAAA-CCTCA ATAATTG--- --CTAACAAA ATTATTTGCC  439 Lop.flav
393  GCTTA-GCCT TAAA-CTTCA ATGATTA--- -AATAACAAA ATCATTTGCC  436 Osg.band
396  GTCTA-GCCG TAAA-CTTAA ACTATTAA-- -TTCAACAAA ATAGTTCGCC  440 3Ped.cap
398  GTATA-GCCA TAAA-CATAA GCAGATC--- -CAAAACAAA TCTGCTCGCC  441 Per.flav
401  GCTTA-GCCC TAAA-CTTAA ATAATTC--- -CATAACAAA ATTACTCGCC  444 Syl.audu
401  GCTTA-GCCC TAAA-CATAA ATATTCA--- -ACTAACAAG AATATTCGCC  444 Sci.nige
398  GCTTA-GCCC TAAA-CATAA ATATTCA--- -ACAAACAAG AATATTCGCC  441 Spe.trid
402  GCTTA-GCCT TAAA-CATAA GTTTTTT--- -ATCAACACA ACAATTCGCC  445 Gra.muri
398  GCTTA-GCCA TAAA-CATAA AAGTTCA--- -CATAACAAG AACTTTCGCC  441 2Hyd.hyd
381  GCTTA-GTTG TAAA-CATAG GATATTCA-- -AATAACAAA -TATCCCGCC  424 Pet.typi
401  GCTTA-GCCC TAAA-CCTCA AAGATTG--- -AATAACAAA ATCATTTGCC  444 Ony.leuc
400  GCTTA-GCCC TAAA-CCTTA AAGATTA--- -AATAACAAA ATCATTTGCC  443 Per.poli
400  GCTTA-GCCT TAAA-CCCTA ATGATTT--- -AATAACAAA ATTATTTGCC  443 Neo.albi
400  GCTTA-GCCC TAAA-CCCTA ATAATTC--- -AATAACAAA ATTATTTGCC  443 Neo.flor
401  GCTTA-GCCC TAAA-CCTTA AAGATTA--- -AATAACAAA ATCATTTGCC  444 Ony.torr
402  GCTTA-GCCC TAAA-CTTTA ACGATTA--- -AATAACAAA ATCATTTGCC  445 Per.erem
400  GCTTA-GCCC TAAA-CCTCA AAGATTA--- -AATAACAAA ATCATTTGCC  443 Per.goss
401  GCTTA-GCCC TAAA-CCTTA AAGATTA--- -AATAACAAA ATAATTTGCC  444 Per.mani
401  GCTTA-GCCC TAAA-CCTTA AAGATTA--- -AATAACAAA ATCATTTGCC  444 Per.mela
401  GCTTA-GCCC TAAA-CCACA ATAACTT--- -AAAAACAAA GTTATTTGCC  444 Sig.hisp
401  GCTTA-GCCC TAAA-CCATA ATAACTT--- -AATAACNAA GTTATTTGCC  444 Sig.masc
401  GCTTA-GCCC TAAA-CCTCA AAGATTA--- -AATAACAAA ATCATTTGCC  444 Ony.aren
400  GCTTA-GCCC TAAA-CCTTA AAGATTA--- -AATAACAAA ATCATTTGCC  443 2Per.leu
  1  GCTTA-GCCA TAAA-CTTTA AAATTCA--- -TACAACAAG AATCTTCGCC   44 Dol.pata
  1  GCTTA-GCCC TAAA-CACAA ATACTTA--- -ACTAACAAA ACTATTCGCC   44 1Pet.leu
  1  GCTTA-GCCC TAAA-CACAA ATGCTTA--- -ACTAACAAA ACTATTCGCC   44 2Pet.leu
  1  GCTTA-GCCC TAAA-CATAA ATACTTA--- -ATTAACAAA ACTATTCGCC   44 1Pet.pet
```

```
  1  GCTTA-GCCC TAAA-CATAA ATATTTA--- -ATTAACAAA ACTATTCGCC   44 2Pet.pet
  1  GCTTA-GCCT TAAA-CATAA ATTTTTC--- -AATAACAAA ATTATTCGCC   44 1Pte.vol
  1  GCTTA-GCCT TAAA-CATAA ATTTTTC--- -AATAACAAA ACTATTCGCC   44 2Pte.vol
  1  GCTTA-GCCT TAAA-CATAA ATTCTTC--- -AATAACAAA ATTATTCGCC   44 3Pte.vol
  1  GCTTA-GCCC TAAA-CATAG ACACTCA--- -ACTAACAAG ACTGTTCGCC   44  Sci.lis
  1  GCTTA-GCCC TAAA-CATAG ACACTCA--- -ATTAACAAG ACTGTTCGCC   44 Sci.vulg
  1  GCTTA-GCCC TAAA-CATAA ATACTCA--- -ATAAACAAG G-TATTCGCC   43 1Tam.sib
  1  GCTTA-GCCC TAAA-CATAA ATACTCA--- -ATAAACAAG G-TATTCGCC   43 2Tam.sib
  1  ---------- ---------- -AATTAA--- -CCTAACAAA ATTGTCTGCC   25 Bas.astu
  1  ---------- ---------- -AATTAA--- -TATAACAAA ACTGTTCGCC   25 Spi.puto
  1  ---------- ---------- -AATTAA--- -CATAACAAA ATTATTTGCC   25 Pro.loto
  1  ---------- ---------- -AATTAA--- -TGTAACAAA ATTGTTCGCC   25 Mep.meph
  1  ---------- ---------- -AATTTA--- -CACAACAAA ATTACTCGCC   25 Odo.rosm
  1  ---------- ---------- -AATTT---- -TACAACAAA ATAATTCGCC   24 Can.latr
  1  ---------- ---------- -AGTTCT--- -ATAAACAAA ATAGTTCGCC   25 Uro.cine
  1  ---------- ---------- -AATTTA--- -TTAAACAAA ATTATTCGCC   25 Urs.amer
  1  ---------- ---------- --GATTC--- -AATAACAAA ATCATTCGCC   24 Myd.marc

     551        561        571        581        591       600
     |          |          |          |          |         |
489  AGAGAACTAC TAGCTACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  537 Rat.norv
488  AGAGAACTAC TAGCCATAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  536 Hyl.stel
487  AGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  535 Mus.croc
487  AGAGAACTAC TAGCCATAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  535 Mus.cook
485  AGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  533 Mas.eryt
492  AGAGAACTAC TAGCTACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  540 Mus.matt
486  AGAGAACTAC TAGCTATAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  534 Mus.plat
488  AGAGAACTAC TAGCTACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  536 Mus.setu
486  AGAGAACTAC TAGCTATAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  534 Mus.saxi
480  AGAGAACTAC TAGCCACAGC CTAAAACTCA AAGGACTTGG -CGGTACTTT  528 Cri.gamb
504  AGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  552 Mac.inge
484  TGAGAACTAC TGGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  532 Mic.niva
485  AGAGAACTAC TAGCTACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  533 Nes.rufu
486  TGAGAACTAC TGGCTACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  534 1Per.leu
479  AGAGAACTAC TAGCCATAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  527 Aco.cahi
489  TGAGAACTAC TAGCCACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  537 Cri.migr
487  AGAGTACTAC AAGCCACAGC TCAAAACTCA AAGGACTTGG -CGGTGCTTT  535 Gli.glis
489  AGAGAACTAC TAGCCAACGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  537 Ger.nige
490  AGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  538 Leo.edwa
488  TGAGAACTAC TGGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  536 Mes.aura
493  AGAGAACTAC AAGCCAAAGC TAAAAACTCA AAGGACTTGG -CGGTGCTTT  541 1Mus.ave
488  AGAGAACTAC TAGCCATAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  536 Mus.musc
493  AGAGAACTAC TAGCTACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  541 Mus.paha
482  AGAGAACTAC GAGCCACCGC TTAAAACTCA AAGGACTTGG -CGGCGCTTT  530 Tat.kemp
486  AGAGAACTAC GAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  534 Ura.rudd
```

```
494  AGAGTACTAC AAGCCACAGC TCAAAACTCA AAGGACTTGG -CGGTGCTTC  542 Dry.nite
492  AGAGAACTAC AAGCCAAAGC TAAAAACTCA AAGGACTTGG -CGGTGCTTT  540 2Mus.ave
488  AGAGTACTAC AAGCCACAGC TCAAAACTCA AAGGACTTGG -CGGTGCTTT  536 1Gli.jap
487  AGAGAACTAC AAGCCACAGC TCAAAACTCA AAGGACTTGG -CGGTGCTTT  535 2Gli.jap
491  CGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  539 1Hyd.hyd
473  ANANAACTAC TAGCAATAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  521 Cav.porc
456  AGAGAACTAC TAGCCACAGC TAAAAACTCA AAGGACTTGG -CGGTGCTTC  504 Bat.suil
457  AGAGAACTAC TAGCCACGGC TAAAAACTCA AAGGACTTGG -CGGTGCTTC  505 Bat.jane
455  AGAGGACTAC CAGCAACAGC TAAAAACTCA AAGGACTTGG -CGGTGCTTC  503 1Cry.hot
456  AGAGAACTAC CAGCAACGGC TAAAAACTCA AAGGACTTGG -CGGTGCTTC  504 2Cry.hot
461  AGAGAACTAC TAGCCACGGC TAAAAACTCA AAGGACTTGG -CGGTGCTTC  509 Cry.dama
456  AGAGGACTAC TAGCCATAGC TAAAAACTCA AAGGACTTGG -CGGTGCTTC  504 Geo.cape
461  AGAGGACTAC TAGCCATGGC TAAAAACTCA AAGGACTTGG -CGGTGCTTT  509 Hel.arge
461  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGA -CGGTGCTTT  509 Het.glab
466  CGAGAACTAC GAGCCACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  514 Thr.swin
463  TGAGAACTAC AAGCAACCGC TAGAAACTCA AAGGACTTGG -CGGTGCTTT  511 2Cra.cas
458  AGAGAACTAC AAGCCACCGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  506 Dip.ordi
455  CGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  503 1Ped.cap
455  CGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  503 2Ped.cap
444  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTC  492 Ath.macr
440  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  488 Cte.boli
438  AGAGAACTAC TAGCAATAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  486 Cav.guia
433  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  481 Chi.lani
438  AGAGTACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  486 Cap.pilo
437  CGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  485 Das.punc
448  AGAGGACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGA -CGGTGCTT-  495 Ere.dors
441  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTC  489 Hys.afri
438  AGAGAAGTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  486 Myo.coyp
436  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  484 Oct.degu
439  AGAGAACTAC TAGCTACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  487 Pro.long
443  AGAGTACTAC TAGCCAAAGC TAAAAACTCA AAGAACTTGG -CGGTGCTTT  491 Apl.rufa
445  AGAGAACTAC TAGCAATTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  493 Cas.cana
448  TGAGAACTAC AAGCAACCGC TAGAAACTCA AAGGACTTGG -CGGTGCTTT  496 1Cra.cas
436  AGAGAACTAC TAGCCACAGC TAAAAACTCA AAGGACTTGG -CGGCACTTC  484 Cte.gund
444  AGAGTACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  492 Das.nove
449  TGAGGACTAC AAGCAACTGC TTGAAACTCA AAGGACTTGG -CGGTGCTTT  497 Geo.burs
434  AGAGAACTAC GAGCCATTGC TTAAAACTCA AAGGACTTGG -CGGCGCTTT  482 Ger.vall
440  AGAGAACTAC TAGCAACTGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  488 Jac.jacu
440  AGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  488 Lop.flav
437  TGAGAACTAC TGGCTATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  485 Osg.band
441  CGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  489 3Ped.cap
442  CGAGAACTAC TAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  490 Per.flav
445  AGAGAACTAC AAGCCAGAGC TTGAAACTCA AAGGACTTGG -CGGTGCTTT  493 Syl.audu
445  AGAGAACTAC TAGCCACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  493 Sci.nige
442  AGAGTACTAC TAGCAATGGC CTAAAACTCA AAGGACTTGG -CGGTGCTTT  490 Spe.trid
```

```
446  AGAAAACTAT AGGCAACAGC CAAAAACTCA AAGGACTTGG -CGGTGCTTT  494 Gra.muri
442  CGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG ACGGTGCTTT  491 2Hyd.hyd
425  AGAGAACTAC AAGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT  473 Pet.typi
445  TGAGAACTAC TGGCCATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Ony.leuc
444  TGAGAACTAC TGGCCATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  492 Per.poli
444  CGAGGACTAC TGGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  492 Neo.albi
444  CGAGAACTAC TGGCCACAGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  492 Neo.flor
445  TGAGAACTAC TGGCCATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Ony.torr
446  TGAGAACTAC TGGCCACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  494 Per.erem
444  TGAGAACTAC TGGCCACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  492 Per.goss
445  TGAGAACTAC TGGCCATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Per.mani
445  TGAGAACTAC TGGCCATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Per.mela
445  TGAGAACTAC TGGCCACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Sig.hisp
445  TGAGAACTAC TGGCCACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Sig.masc
445  TGAGAACTAC TGGCCATCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  493 Ony.aren
444  TGAGAACTAC TGGCTACCGC TTAAAACTCA AAGGACTTGG -CGGTACTTT  492 2Per.leu
 45  CGAGGACTAC TAGCAATAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 Dol.pata
 45  AGAGTACTAC TAGCCATTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 1Pet.leu
 45  AGAGTACTAC TAGCCATTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 2Pet.leu
 45  AGAGTACTAC TAGCCATTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 1Pet.pet
 45  AGAGTACTAC TAGCCATTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 2Pet.pet
 45  AGAGTACTAC TAGCAACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 1Pte.vol
 45  AGAGTACTAC TAGCAACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 2Pte.vol
 45  AGAGTACTAC TAGCAACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 3Pte.vol
 45  AGAGAACTAC TAGCTACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93  Sci.lis
 45  AGAGGACTAC TAGCTACTGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   93 Sci.vulg
 44  AGAGTACTAC TAGCAACAGC CTAAAACTCA AAGGACTTGG -CGGTGCTTT   92 1Tam.sib
 44  AGAGTACTAC TAGCAACAGC CTAAAACTCA AAGGACTTGG -CGGTGCTTT   92 2Tam.sib
 26  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   74 Bas.astu
 26  AGAGAACTAC TAGCAATAGC CTGAAACTCA AAGGACTTGG -CGGTGCTTT   74 Spi.puto
 26  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   74 Pro.loto
 26  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   74 Mep.meph
 26  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTC   74 Odo.rosm
 25  AGAGGACTAC TAGCAATAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   73 Can.latr
 26  AGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   74 Uro.cine
 26  GGAGAACTAC TAGCAACAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   74 Urs.amer
 25  AGAGAACTAC TAGCAATAGC TTAAAACTCA AAGGACTTGG -CGGTGCTTT   73 Myd.marc

     601        611        621        631        641      650
     |          |          |          |          |        |
538  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTCTACCT  587 Rat.norv
537  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATACACCC CGCTTTACCT  586 Hyl.stel
536  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATATACCC CGCTCCACCT  585 Mus.croc
536  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTCTACCT  585 Mus.cook
534  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATACACCC CGCTTTACCT  583 Mas.eryt
```

```
541  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTCTACCT  590 Mus.matt
535  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTCCACCT  584 Mus.plat
537  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTATACCT  586 Mus.setu
535  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTCCACCT  584 Mus.saxi
529  ACATCCATCT AGAGGAGCCT GTTCTATAAT CGATATACCC CGTTCTACCT  578 Cri.gamb
553  ACATCCGTCT AGAGGAGCCT GTTCTATAAT CGATACACCC CGCTTTACCT  602 Mac.inge
533  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  582 Mic.niva
534  ATATCCACCT AGAGGAGCCT GTTCTATAAT CGATATACCC CGCTACACCT  583 Nes.rufu
535  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  584 1Per.leu
528  ATATCCACCT AGAGGAGCCT GTTCTATAAT CGATATACCC CGTTACACCT  577 Aco.cahi
538  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTATACCT  587 Cri.migr
536  ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  585 Gli.glis
538  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTCCACCT  587 Ger.nige
539  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTCTACCT  588 Leo.edwa
537  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  586 Mes.aura
542  ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATATACCT  591 1Mus.ave
537  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTCTACCT  586 Mus.musc
542  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATATACCC CGCTCCACCT  591 Mus.paha
531  ATATCCGTCT AGAGGAGCCT GTTTTATAAT CGATAATCCC CGTTCTACCT  580 Tat.kemp
535  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTATACCT  584 Ura.rudd
543  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATCAACCT  592 Dry.nite
541  ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATATACCT  590 2Mus.ave
537  ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATATACCT  586 1Gli.jap
536  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  585 2Gli.jap
540  ATATCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  589 1Hyd.hyd
522  ATACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  571 Cav.porc
505  ACACCCGCCT AGAGGAGCCT GTTCTATAAT CGATATTCCC CGATAAACCT  554 Bat.suil
506  ACACCCACCT AGAGGAGCCT GTTCTATAAT CGATATACCC CGATAAACCT  555 Bat.jane
504  ACACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  553 1Cry.hot
505  ACACCCGCCT AGAGGAGCCT GTTCTGTAAT CGATAAACCC CGATAAACCT  554 2Cry.hot
510  ACACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  559 Cry.dama
505  ACACCCATCT AGAGGAGCCT GTTCTGTAAT CGATACACCC CGATAAACCT  554 Geo.cape
510  ATATCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  559 Hel.arge
510  AAACCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTAAACCT  559 Het.glab
515  ATACCCACCT AGAGGAGCCT GTTTTATAAT CGATAAACCC CGATACACCT  564 Thr.swin
512  ACATCCGTCT AGAGGAGCCT GTTCCGAAAC CGATAAACCC CGATATACCT  561 2Cra.cas
507  ATATCCACCT AGAGGAGCCT GTTCCATAAC CGATGATCCC CGATTTACCT  556 Dip.ordi
504  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTTTACCT  553 1Ped.cap
504  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTTTACCT  553 2Ped.cap
493  ACAACCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  542 Ath.macr
489  AAACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTACACCT  538 Cte.boli
487  ATACCCGCCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  536 Cav.guia
482  ATATCCACCT AGAGGGGCCT GTTCTATAAT CGATAAACCC CGATCAACCT  531 Chi.lani
487  AAACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATATACCT  536 Cap.pilo
486  ACACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  535 Das.punc
```

```
496  ACACCCATCT AGAGGAGCCT GTTCTGTAAT CGATAAACCC CGATTTACCT  545 Ere.dors
490  ACAACCATCT AGAGGAGCCT GTCCTATAAT CGATAATCCC CGATAAACCT  539 Hys.afri
487  AAACCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  536 Myo.coyp
485  AAACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  534 Oct.degu
488  AAACCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  537 Pro.long
492  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  541 Apl.rufa
494  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATTCACCT  543 Cas.cana
497  ACATCCGTCT AGAGGAGCCT GTTCCGAAAC CGATAAACCC CGATATACCT  546 1Cra.cas
485  ACATCCCTCT AGAGGAGCTT GTTCTATAAT CGATAAACCC CGATAAACCT  534 Cte.gund
493  ACATCCTTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATATACCT  542 Das.nove
498  ATACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATCTACCT  547 Geo.burs
483  ATATCCGTCT AGAGGAGCCT GTTCTTTAAT CGATAAACCC CGTTCCACCT  532 Ger.vall
489  ATATCCACCT AGAGGAGCCT GTTCTATAAC CGATACACCC CGTTCTACCT  538 Jac.jacu
489  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATACACCC CGTTATACCT  538 Lop.flav
486  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATACACCC CGTTATACCT  535 Osg.band
490  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTTTACCT  539 3Ped.cap
491  ATATCCATCT AGAGGAGCCT GTTCTGTAAT CGATAAACCC CGATATACCT  540 Per.flav
494  ATACCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  543 Syl.audu
494  ATATCCCCCT AGAGGAGCCT GTTCTGTAAT CGATAAACCC CGTTAAACCT  543 Sci.nige
491  ATACCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  540 Spe.trid
495  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  544 Gra.muri
492  ATATCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATACACCT  541 2Hyd.hyd
474  ATACCCACCT AGAGGAGCCT GTTTCATAAT CGATAAACCC CGATAAACCT  523 Pet.typi
494  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC -GTTATACCT  542 Ony.leuc
493  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  542 Per.poli
493  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  542 Neo.albi
493  ACATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  542 Neo.flor
494  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC -GTTATACCT  542 Ony.torr
495  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  544 Per.erem
493  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  542 Per.goss
494  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  543 Per.mani
494  ATATCCGTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  543 Per.mela
494  ATATCCACCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTATACCT  543 Sig.hisp
494  ATATCCANCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTATACCT  543 Sig.masc
494  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC -GTTATACCT  542 Ony.aren
493  ATATCCATCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTATACCT  542 2Per.leu
 94  ATACCCACCT AGAGGAGCCT GTTCTGTAAT CGATAAACCC CGATATACCT  143 Dol.pata
 94  ATACCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTCTACCT  143 1Pet.leu
 94  ATACCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTCTACCT  143 2Pet.leu
 94  ACATCCCCCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTTTACCT  143 1Pet.pet
 94  ACATCCCTCT AGAGGAGCCT GTTCTATA-T CGATAAACCC CGTTTTACCT  142 2Pet.pet
 94  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  143 1Pte.vol
 94  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  143 2Pte.vol
 94  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  143 3Pte.vol
 94  ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTAAACCT  143  Sci.lis
```

```
94   ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGTTAAACCT  143 Sci.vulg
93   ACACCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTTAACCT  142 1Tam.sib
93   ACACCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGCTTAACCT  142 2Tam.sib
75   ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  124 Bas.astu
75   ATATCCCTCT AGAGGAGCCT GTCCTATAAT CGATAAACCC CGATAAACCT  124 Spi.puto
75   ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  124 Pro.loto
75   ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAGACCT  124 Mep.meph
75   ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  124 Odo.rosm
74   ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  123 Can.latr
75   ACATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAGACCT  124 Uro.cine
75   AAACCCCCCT AGAGGAGGCT GTTCTATAAT CGATAAACCC CGATAGACCT  124 Urs.amer
74   ATATCCCTCT AGAGGAGCCT GTTCTATAAT CGATAAACCC CGATAAACCT  123 Myd.marc

     651        661        671        681        691       700
     |          |          |          |          |         |
588  TACCCCTTCT CGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  636 Rat.norv
587  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAATCCC  635 Hyl.stel
586  CACCATCCCT TGCTAATT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  634 Mus.croc
586  CACCATCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  634 Mus.cook
584  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  632 Mas.eryt
591  CACCATCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  639 Mus.matt
585  CACCATCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  633 Mus.plat
587  CACCATCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  635 Mus.setu
585  CACCATCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  633 Mus.saxi
579  CACCATCTCT TGCTAATC-C AGTCTATATA CCGCCATCTT CAGCAAACCT  627 Cri.gamb
603  CACCGTCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  651 Mac.inge
583  CACCACCCCT TGCTAATT-C AGCCTATTTT CCGCCATCTT CAGCAAACCC  631 Mic.niva
584  CACCGTCTCT TGCTAATT-C AGCCTATATA CCGCCATCTC CAGCAAACCC  632 Nes.rufu
585  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  633 1Per.leu
578  CACCACCCCT TGCTAATC-C AGCCTATATA CCGCCATCTT CAGCAAACCC  626 Aco.cahi
588  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  636 Cri.migr
586  CACCACTTCT TGC-TACCAC AGCTTATATA CCGCCATCTT CAGCAAACCT  634 Gli.glis
588  TACCAACTCT TGCTATTT-C AGCTTATATA CCGCCATCTT AAGTAAACCC  636 Ger.nige
589  CACCACCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  637 Leo.edwa
587  TACCACCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  635 Mes.aura
592  CACCATTTCT TGT-CACCAC AGCCTATATA CCGCCATCTT CAGCAAACCT  640 1Mus.ave
587  CACCATCTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  635 Mus.musc
592  CACCATCCCT TGCTAATT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  640 Mus.paha
581  TACCAACTCT TGCTTTTT-C AGCTTATATA CCGCCATCTT CAGTAAACCC  629 Tat.kemp
585  TACCACCTTT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  633 Ura.rudd
593  CACCACTCCT TGTTCATCAC AGCCTATATA CCGCCATCTT CAGCAAACCT  642 Dry.nite
591  CACCATTTCT TGT-CACCAC AGCCTATATA CCGCCATCTT CAGCAAACCT  639 2Mus.ave
587  CACCTCTTCT TGC-CACCAC AGCTTATATA CCGCCATCTT CAGCAAACCT  635 1Gli.jap
586  CACCTCTTCT TGC-CACCAC AGCTTATATA CCGCCATCTT CAGCAAACCT  634 2Gli.jap
590  CACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  638 1Hyd.hyd
```

```
572  CACCTCTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCCAACCC  620 Cav.porc
555  TACCACTTTT TGCTAATA-C AGTCTATATA CCGCCATCTT CAGCAAACCT  603 Bat.suil
556  TACCACTTTT TGCTAATA-C AGTCTATATA CCGCCATCTT CAGCAAACCT  604 Bat.jane
554  CACCACTATT TGCTAATA-C AGCTTATATA CCGCCATCTT CAGCGAACCT  602 1Cry.hot
555  CACCACTATT TGCTAATA-C AGTTTATATA CCGCCATCTT CAGCAAACCT  603 2Cry.hot
560  CACCACTTTT TGCTAATA-C AGCCTATATA CCGCCATCTT CAGCGAACCT  608 Cry.dama
555  CACCACTTTT TGCCAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  603 Geo.cape
560  CACCACTTTT TGCTAATC-C AGTCTATATA CCGCCATCTT CAGCTAACCT  608 Hel.arge
560  CACCATTTCT GGCTACTT-C AGTCTATATA CCGCCATCTT CAGCGAACCT  608 Het.glab
565  CACCATCTCT TGCTTATAAC AGTCTATATA CCGCCATCTT CAGCGAACCC  614 Thr.swin
562  TACCACTTCT TGTTAATT-C AGCCTGTATA CCGCCATCTT CAGCAAACCC  610 2Cra.cas
557  CACCACTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCTAACCC  605 Dip.ordi
554  TACCACATCT TGCTAATT-C AGCCTACATA CCGCCATCTT CAGCAAACCT  602 1Ped.cap
554  TACCACATCT TGCTAATT-C AGCCTACATA CCGCCATCTT CAGCAAACCT  602 2Ped.cap
543  CACCACTTCT TGCTAATC-C AGCTTATATA CCGCCATCTT CAGCAAACCC  591 Ath.macr
539  CACCACCCCT TGCTAATC-C AGCCTATATA CCGCCATCTT CAGCAAACCC  587 Cte.boli
537  CACCTCTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCCAACCC  585 Cav.guia
532  TACCACTTCT TGCTAATA-C AGCCTATATA CCGCCATCTT CAGCAAACCC  580 Chi.lani
537  TACCACTTCT TGCTAATC-C AGTCTATATA CCGCCATCTT CAGCAAACCC  585 Cap.pilo
536  CACCACCCCT TGCTAATT-C AGCTTATATA CCGCCATCTT CAGCAAACCC  584 Das.punc
546  CACCGTTTCT TGCTAATT-C AGTTTATATA CCGCCATCTT TAGCAAACCC  594 Ere.dors
540  TACCACTCCT TGCTAATC-C AGCTTATATA CCGCCATCTC CAGCAAACCC  588 Hys.afri
537  CACCACTTCT CGCTAATT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  585 Myo.coyp
535  CACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  583 Oct.degu
538  CACCACTTCT CGCTAATT-C AGTTTATATA CCGCCATCTT CAGCAAACCC  586 Pro.long
542  CACCACCTTT TGC-TAAATC AGCCTATATA CCGCCATCTT CAGCAAACCC  590 Apl.rufa
544  CACCATCTCT TGC-TAAGTC AGCCTATATA CCGCCATCTT CAGCAAACCC  592 Cas.cana
547  TACCACTTCT TGTTAATT-C AGCCTGTATA CCGCCATCTT CAGCTAACCC  595 1Cra.cas
535  AACCACCTTT TGC-TAAATC AGCCTATATA CCGCCATCTT CAGCAAACCT  583 Cte.gund
543  CACCACCCCT TGC-TAATAC AGCCTATATA CCGCCATCTT CAGCAGACCC  591 Das.nove
548  TACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCCT CAGCTAACCC  596 Geo.burs
533  TACCAACTCT TGCTTTT--C AGCTTATATA CCGCCATCTT CAGTAAACCC  580 Ger.vall
539  CACCACTTCT TGCTAATT-C AGCCTGTATA CCGCCATCTT CAGCGAACCC  587 Jac.jacu
539  CACCACCCCT TGCTAATA-C AGCCTATATA CCGCCATCTT CAGCAAACCC  587 Lop.flav
536  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  584 Osg.band
540  TACCACATCT TGCTAATT-C AGCCTACATA CCGCCATCCT CAGCAAACCT  588 3Ped.cap
541  TACCATTTCT TGCTAATA-C CGTCTATATA CCGCCATCTT CAGCAAACCC  589 Per.flav
544  TACCACTCTT TGCCAACT-C AGCCTATATA CCGCCATCTT CAGCGAACCC  592 Syl.audu
544  CACCACTCTT TGCAA-TGTC AGCCTATATA CCGCCATCTT CAGCAAACCC  592 Sci.nige
541  CACCACCTTT AGCAAATATC AGCCTATATA CCGCCATCTG CAGCAAACCC  590 Spe.trid
545  CACCACTTCT TGC-CACCTC AGCTTATATA CCGCCATCTT CAGCAAACCT  593 Gra.muri
542  CACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCC  590 2Hyd.hyd
524  TACCGCCCCT CGCTAATC-C AGTCTATATA CCGCCATCTT CAGCAAACCC  572 Pet.typi
543  CACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Ony.leuc
543  CACCACCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Per.poli
```

```
543  TACCACTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Neo.albi
543  TACCACTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Neo.flor
543  CACCACTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Ony.torr
545  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  593 Per.erem
543  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Per.goss
544  CACCACCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  592 Per.mani
544  CACCACCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  592 Per.mela
544  CACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  592 Sig.hisp
544  CACCACTTCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  592 Sig.masc
543  CACCACTCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 Ony.aren
543  CACCATCCCT TGCTAATT-C AGCCTATATA CCGCCATCTT CAGCAAACCT  591 2Per.leu
144  CACCACTTCT TGCTAATA-C AGCCTATATA CCGCCATCTT CAGCAAACCC  192 Dol.pata
144  CACCACCTTT TGCAA-TATC AGCCTATATA CCGCCATCTT CAGCAAACCC  192 1Pet.leu
144  CACCACCTTT TGCAA-TATC AGCCTATATA CCGCCATCTT CAGCAAACCC  192 2Pet.leu
144  CACCACCTTT TGCAA-TGTC AGCCTATATA CCGCCATCTT CAGCAAACCC  192 1Pet.pet
143  CACCACCTTT TGCAA-TGTC AGCCTATAT- CCGCCATCTT CAGCAAACCC  190 2Pet.pet
144  CACCACCCTT TGCAACTATC AGCCTATATA CCGCCATCTT CAGCAAACCC  193 1Pte.vol
144  CACCACCCTT TGCAACTATC AGCCTATATA CCGCCATCTT CAGCAAACCC  193 2Pte.vol
144  CACCACCCTT TGCAACTATC AGCCTATATA CCGCCATCTT CAGCAAACCC  193 3Pte.vol
144  CACCACTCTT TGCAA-TATC AGCCTATATA CCGCCATCTT CAGCAAACCC  192  Sci.lis
144  CACCACTCTT TGCAA-TATC AGCCTATATA CCGCCATCTT CAGCAAACCC  192 Sci.vulg
143  TACCACCTCT TGCAA-CATC AGCCTATATA CCGCCATCTT CAGCAAACCC  191 1Tam.sib
143  TACCACCCCT TGCAA-CATC AGCCTATATA CCGCCATCTT CAGCAAACCC  191 2Tam.sib
125  CACCACCTCT AGCTAAAC-C AGTCTATATA CCGCCATCTT CAGCAAACCC  173 Bas.astu
125  CACCGTTCCT AGCTAATT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  173 Spi.puto
125  CACCATCTCT AGCTAAAT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  173 Pro.loto
125  CACCGTTTCT AGCTAATC-C AGTCTATATA CCGCCATCTT CAGCAAACCC  173 Mep.meph
125  CACCATCCCT TGCTAATT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  173 Odo.rosm
124  CACCACCCTT TGCTAATT-C AGTCTATATA CCGCCATCTT CAGCAAACCC  172 Can.latr
125  CACCATTCCT TGCTAATA-C AGTCTATATA CCGCCATCTT CAGTAAACCC  173 Uro.cine
125  CACCACCTCT TGCTAATC-C AGTCTATATA CCGCCATCTT CAGCTAACCC  173 Urs.amer
124  CACCGTTCCT AGCTAATC-C AGTCTATATA CCGCCATCTT CAGCAAACCC  172 Myd.marc

     701        711        721        731        741      750
     |          |          |          |          |        |
637  TAA--AAAGG CACTAAAGTA AGCACAAGAA CA----AA-C ATAAAAACGT  679 Rat.norv
636  TTA--AAAGG TCTAAAAGTA GGCAAGAGAA TC----AAAC ATAAAAACGT  679 Hyl.stel
635  TAA--AAAGG TAAAAAAGTA AGCACAATAA CA----AA-C ATAAAAACGT  677 Mus.croc
635  TAA--AAAGG AATTAAAGTA AGCACAAGAA TC----AAAC ATAAAAACGT  678 Mus.cook
633  TTAA-AAAGG AATAAGAGTA AGCAAGAGAA CC----AATC ATAAAAACGT  677 Mas.eryt
640  TAA--AAAGG AAGTATAGTA AGCACAAGAA CA----TC-C GTAAAAACGT  682 Mus.matt
634  TAA--AAAGG TATCAGAGTA AGCACAAGAA TT----TTAC ATAAAAACGT  677 Mus.plat
636  TAA--AAAGG AAAAAAAGTA AGCACAAGAA TA----CT-C ATAAAAACGT  678 Mus.setu
634  TAA--AAAGG TATCAGAGTA AGCACAAGAA TT----TTAC ATAAAAACGT  677 Mus.saxi
628  TAA--AAAAG AAACACAGTA AGCAAGATAA TA----A-CC ATAAAAACGT  670 Cri.gamb
```

```
652  TAA--AAAGG AATGACAGTA AGCAAAAGAA TA----A-CC ATAAAAACGT  694 Mac.inge
632  TAA--AAAGG AGTAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  674 Mic.niva
633  TAA--AAAGG AACCACAGTA AGCTCAAGAA TG----A-CC ATAAAGACGT  675 Nes.rufu
634  TAA--AAAGG AGTAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  676 1Per.leu
627  TCA--AAAGG ACACACAGTA AGCAAGAGCA CA----A-AC ATAAAAAAGT  669 Aco.cahi
637  TTA--AAAAG AATAAAAGTA AGCAAGAGCA TA----CC-C GTAAAAACGT  679 Cri.migr
635  TAA--CAAGG AATCAAAGTA AGCCCAAGTA TT----A-AC ATAAAAACGT  677 Gli.glis
637  TCA--AAAGG ATACATAGTG AGCGAAAAAA CA----AA-C ATAAAGACGT  679 Ger.nige
638  TAA--AAAGG CATAGAAGTA AGCGCAAGAA CA----AG-C ATAAAAACGT  680 Leo.edwa
636  TAA--AAAAG AACAAGAGTA AGCAAGAGAA TA----CC-C ATAAAAACGT  678 Mes.aura
641  TAA--TAAGG AGATAAAGTA AGCTCAAGCA CT----G-AC ATAAAAACGT  683 1Mus.ave
636  TAA--AAAGG TATTAAAGTA AGCAAAAGAA TC----AAAC ATAAAAACGT  679 Mus.musc
641  TAAT-AAAGG TAGTAAAGTA AGCACAAGAA CA----AA-C ATAAAAACGT  684 Mus.paha
630  TAA--AAAGG AAGTAAAGTA AACGAAAGAA CA----AA-C ATGAAGACGT  672 Tat.kemp
634  TAA--AAAAG AGAAAAAGTA AGCAAAAGAA TT----AAAC ATAAAAACGT  677 Ura.rudd
643  TAA--CAAAG AATAAAAGTA AGCACAAGTA TA----A-TC ATAAAAACGT  685 Dry.nite
640  TAA--CAAGG AGTAAAAGTA AGCTCAAGCA CT----G-AC ATAAAAACGT  682 2Mus.ave
636  TAA--CAAGG CACCCAAGTA AGCACAAATA TT----A-AC ATAAAAACGT  678 1Gli.jap
635  TCA--CAAGG TACCCAAGTA AGCACAAATA TC----A-AC ATAAAAACGT  677 2Gli.jap
639  AAT--CATGG CAACAAAGTA AGCACAAACA TA----TTAC ATAAAAACGT  682 1Hyd.hyd
621  CAT--TATGG AAACAAAGTG AGCGCAAGTA CA----CTAC ATAAAAACGT  664 Cav.porc
604  TAA--CAAAG TATAAAAGTA AGCGCAATAA T-----AATC GTAAAAACGT  646 Bat.suil
605  TTA--TAAAG AATAAAAGTA AGCGCAATAA T-----AATC GTAAAAACGT  647 Bat.jane
603  TTA--TAAAG AAAAAAAGTA AGCACAATCA C-----AAGC ATAAAGACGT  645 1Cry.hot
604  TAA--TAAAG AAAAAAAGTG AGCACAATCA C-----AAGC ATAAAGACGT  646 2Cry.hot
609  TAA--CAAAG TAAAAAAGTG AGCACAATCA C-----AAAC ATAAAGACGT  651 Cry.dama
604  TAA--CAAAG AGAAAAAGTA AGCACAATCA C-----AATC ATAAAAACGT  646 Geo.cape
609  TTA--TAAAG AACAAAAGTA AGCACAATAA T-----AATC ATAAAAACGT  651 Hel.arge
609  CTA--CAGGG GACAAAAGTA AGCTCAAGTA C-----TGTC ATAAAAACGT  651 Het.glab
615  TAA--AAAGG CAAAAAAGTA AGCATAATCA TCT---AATC ATGAAAACGT  659 Thr.swin
611  TAA--CAAGG AACATTAGTA AGCAAAAGCT TTATAAACTA ATAAAAAAGT  658 2Cra.cas
606  TAA---AAGG AAACAAAGTA AGCCAAATGA GTA---ATTA CTAAAGAAGT  649 Dip.ordi
603  TAA--CAAAG AAAAAAAGTA AGCACAAGCA TTT---A-CC ATTAAAACGT  646 1Ped.cap
603  TAA--CAAAG AAAAAAAGTA AGCACAAGCA TTT---A-CC ATTAAAACGT  646 2Ped.cap
592  TAA--CAAGG AAAAAAAGTA AGCACAAGTA C-----ACAC ATAAAAACGT  634 Ath.macr
588  TAA--CAAGG ATAAAAAGTA AGCATAATCA T-----ATAC ATAAAAACGT  630 Cte.boli
586  TAT--TATGG AAACAAAGTG AGCGCAAGTA CA----CTAC ATAAAAACGT  629 Cav.guia
581  CTAC-ATGGA ACAAAAAGTA AGCACAATAA T-----CCCC ATAAAAACGT  624 Chi.lani
586  CAA--CAGGG ATTAAAAGTA AGCACAACTA T-----ACTC GTAAAAACGT  628 Cap.pilo
585  TAT--TATGG AACAACAGTA AGCTTAACTA TC----AACA ATAAAAACGT  628 Das.punc
595  AAT--TATGG AATTAAAGTA AGCACAAGTA T-----TAAC ATAAAAACGT  637 Ere.dors
589  TAA--CAAGG AACAAAAGTA AGCGCAAGTA T-----GCAC GTAAATACGT  631 Hys.afri
586  CAA--CAGGG ATTAAAAGTA AGCACAACGA T-----CATC ATAAAAN-GT  627 Myo.coyp
584  TAA--TAAGG AAAAAAAGTA AGCACAAATA T-----ACTC ATAAAAACGT  626 Oct.degu
587  CCA--CAGGG AATTAAAGTA AGCACAATTA T-----TATC ATAAAAACGT  629 Pro.long
```

```
591  TCA--CAAGG AATAACAGTA AGCAGGAGCA TT----TTAC ATTAAAACGT  634 Apl.rufa
593  TAATTAAGGC ACTCTAAGTA AGCCAAAACA TA----CAAC ATAAAAACGT  638 Cas.cana
596  TAA--AAAGG AACATTAGTA AGCAAAAGCA TTATAAACTA ATAAAAAAGT  643 1Cra.cas
584  TC---CAAAG AACAGAAGTA AGCAAGAGAA TC----TACC GTTAAAAAGT  626 Cte.gund
592  TAG--TAAGG CACCACAGTG AGCACAATAA CA----T-AC ATAAAGACGT  634 Das.nove
597  TAA--AAAGG AACATTAGTA AGCACAATCA TTATAATCTA ATAAAAAAGT  644 Geo.burs
581  TAA--AAAGG AAAAGAAGTA AACGAGAGAA CA----AA-C ATAAAGACGT  623 Ger.vall
588  TTA--TAGGG AAAAAAAGTA AGCTCAATAA TTA----CCC ATAAAAACGT  631 Jac.jacu
588  TAA--AAAGG AACACAAGTA AGCAAGAGAA C-----AAAC ATAAAAACGT  630 Lop.flav
585  TAA--AAAGG AATAAAAGTA AGCAAGAGAA TT----AC-C ATAAAAACGT  627 Osg.band
589  TAA--CAAAG AAAAAAAGTA AGCACAAGCA TTT---A-CC ATAAAAACGT  632 3Ped.cap
590  TTA--AAAGG CTGACAAGTA AGCTCAATAA TAA----TAT ATTAAAAAGT  633 Per.flav
593  TAA--AAAGG AACAAAAGTA AGCTCAATTA TCC-----CC ATAAAAACGT  635 Syl.audu
593  TAA--CAAGG CATCATAGTA AGCATAATAA TT----CTAC ATAAAAACGT  636 Sci.nige
591  TAA--AAAGG TCCTATAGTA AGCAAGAAAA TT----CTAC ATTAGTACGT  634 Spe.trid
594  CAA--TAGAG AGTTAAAGTA AGCTCGAGCA TT----AGAC ATAAAAACGT  637 Gra.muri
591  AAT--CATGG CAACAAAGTA AGCACAAACA TA----TTAC ATAAAAACGT  634 2Hyd.hyd
573  TAA--AAAGG ATGAACAGTG AGCATAATCA TGA---AACC ATAAAAACGT  617 Pet.typi
592  TAA--AAAGG AATAAAAGTA AGCAAGAGAA TT----AC-C ATAAAAACGT  634 Ony.leuc
592  TAA--AAAGG AATAAAAGTA AGCAAGAGAA TC----TC-C ATAAAAACGT  634 Per.poli
592  TAA--AAAG- AATAACAGTA AGCAGGAGAA TC----AC-C ATAAAAACGT  633 Neo.albi
592  TAA--AAAG- AACAACAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  633 Neo.flor
592  TAA--AAAGG AACAAAAGTA AGCAAGATAA TA----AC-C ATAAAAACGT  634 Ony.torr
594  TAA--AAAGG AATAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  636 Per.erem
592  TAA--AAAGG AGTAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  634 Per.goss
593  TAA--AAGGG AATAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  635 Per.mani
593  TAA--AAAGG AATAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  635 Per.mela
593  TTTA-AAAAG AACCATAGTA AGCAAGAGAA CA----AT-C ATAAAAACGT  636 Sig.hisp
593  TTTA-AAAAG AATTATAGTA AGCAAGAGAA CA----AT-C ATAAAAACGT  636 Sig.masc
592  TAA--AAAGG AAAAAAAGTA AGCAAGAGAA TA----AC-C ATAAAAACGT  634 Ony.aren
592  NAA--AAAGG ANTAAAAGTA AGCAAGAGAA TC----AC-C ATAAAAACGT  634 2Per.leu
193  AAT--CATGG AACCAAAGTA AGCAAAAGTA TC----TTGC ATAAAAACGT  236 Dol.pata
193  TAA--TAAGG CAGCATAGTA AGCATAAAAA TT----CTAC ATAAAAACGT  236 1Pet.leu
193  TAA--TAAGG CAGCATAGTA AGCATAAAAA TT----CTAC ATAAAAACGT  236 2Pet.leu
193  TAA--TAAGG CAATAAAGTA AGCACAAAAA TG----TCAC ATAAAAACGT  236 1Pet.pet
191  TAA--TAAGG CAATAAAGTA AGCACAAAAA TG----TCAC ATAAAAACGT  234 2Pet.pet
194  TAA--CAAGG CACTAAAGTA AGCATAATAA TAC---TTAC ATAAAAACGT  238 1Pte.vol
194  TAA--CAAGG CACTAAAGTA AGCATAATAA TAC---TTAC ATAAAAACGT  238 2Pte.vol
194  TAA--CAAGG CACTAAAGTA AGCATAATAA TAC---TTAC ATAAAAACGT  238 3Pte.vol
193  TAA--CAAGG CACCATAGTA AGCACAATAA TC----TTAC ATAAAAACGT  236  Sci.lis
193  TAA--CAAGG CACTATAGTA AGCACAATAA TT----TTAC ATAAAAACGT  236 Sci.vulg
192  TAT--TAAGG CATTATAGTA AGCACAATAA TT----TAAC ACTAACACGT  235 1Tam.sib
192  TAT--TAAGG TAGTATAGTA AGCACAATAA TT----TAAC ACTAATACGT  235 2Tam.sib
174  TTA--AAAGG AAGAATAGTA AGCACAATAA TA----ATAC ATAAAAAAGT  217 Bas.astu
174  TTA--AAAGG AGTACAAGTA AGCACAATCA TA----AGAC ATAAAAAAGT  217 Spi.puto
```

```
174  TTA--AAAGG AAGAGCAGTA AGCACAATAA TA----ATAC ATAAAAAAGT  217 Pro.loto
174  TTA--AAAGG CACACAAGTA AGCACAATCA TA----CGAC ATAAAAAAGT  217 Mep.meph
174  TTA--AAAGG AAAGAAAGTA AGCATAATCA TC----ACAC GTAAAAAAGT  217 Odo.rosm
173  TCA--AAAGG TAGAGCAGTA AGCACAATCA TT----TTAC ATAAAAAAGT  216 Can.latr
174  TTA--AAAGG TAAAGTAGTA AGCACGATCA CT----TCAC GCAAAAAAGT  217 Uro.cine
174  TTA--AAAGG AGTAAAAGTA AGCACAATCA TC----CCAC ATAAAAAAGT  217 Urs.amer
173  TTA--AAAGG ACAAGCAGTA AGCATGATCA TA----GGAC ATAAAAAAGT  216 Myd.marc

     751        761        771        781        791      800
     |          |          |          |          |        |
680  TAGGTCAAGG TGTAGCCAAT GAAGCGGAAA GAAATGGGCT ACATTTTCTT  729 Rat.norv
680  TAGGTCAAGG TGTAGCCAAT GGGATGGGAA GCAATGGGCT ACATTTTCTT  729 Hyl.stel
678  TAGGTCAAGG TGTAACTAAT GGGATGGGAA GAAATGGGCT ACATTTTCTT  727 Mus.croc
679  TAGGTCAAGG TGTAGCCAAT GAAATGGGAA GAAATGGGCT ACATTTTCTT  728 Mus.cook
678  TAGGTCAAGG TGTAGCCAAT GGGATGGAAA GCAATGGGCT ACATTTTCTA  727 Mas.eryt
683  TAGGTCAAGG TGTAGCCAAT GAGATGGGAA GAAATGGGCT ACATTTTCTT  732 Mus.matt
678  TAGGTCAAGG TGTAGCCAAT GAGATGGAAA GAAATGGGCT ACATTTTCTT  727 Mus.plat
679  TAGGTCAAGG TGTAGCCAAT GAGATGGGAA GAAATGGGCT ACATTTTCTT  728 Mus.setu
678  TAGGTCAAGG TGTAGCCAAT GAGATGGAAA GAAATGGGCT ACATTTTCTT  727 Mus.saxi
671  TAGGTCAAGG TGTAACCCAT GAGATGGAAA GCAATGGGCT ACATTTTCTC  720 Cri.gamb
695  TAGGTCAAGG TGTAGCCTAT GAGACGGAAA GTAATGGGCT ACATTTTCTT  744 Mac.inge
675  TAGGTCAAGG TGTAGCCAAT GAGGTGGGAA GCAATGGGCT ACATTTTCTT  724 Mic.niva
676  TAGGTCAAGG TGTAGCCTAT GAGACGGAAA GCAATGGGCT ACATTTTCTT  725 Nes.rufu
677  TAGGTCAAGG TGTAGCTTAT GAGATGGGAA GCAATGGGCT ACATTTTCTT  726 1Per.leu
670  TAGGTCAAGG TGTAGCCCAT GAGGTGGGAA GTAATGGGCT ACATTTTCTT  719 Aco.cahi
680  TAGGTCAAGG TGTAGCCTAT GGGATGGGAA GTAATGGGCT ACATTTTCTT  729 Cri.migr
678  TAGGTCAAGG TGTAGCCTAT GAAGTGGAAA GAAATGGGCT ACATTTCCTT  727 Gli.glis
680  TAGGTCAAGG TGTAGCCTAT GAGTTGGAAA GCAATGGGCT ACATTTTCTT  729 Ger.nige
681  TAGGTCAAGG TGTAGCCAAT GAGGTGGGAA GAAATGGGCT ACATTTTCTT  730 Leo.edwa
679  TAGGTCAAGG TGTAGCCTAT GGGCTGGGAA GTAATGGGCT ACATTTTCTT  728 Mes.aura
684  TAGGTCAAGG TGTAGCTTAT GAAATGGGAA GAAATGGGCT ACATTTACTT  733 1Mus.ave
680  TAGGTCAAGG TGTAGCCAAT GAAATGGGAA GAAATGGGCT ACATTTTCTT  729 Mus.musc
685  TAGGTCAAGG TGTAGCTAAT GGGATGGGAA GAAATGGGCT ACATTTTCTT  734 Mus.paha
673  TAGGTCAAGG TGTAGCCTAT GAGTTGGAAA GTAATGGGCT ACATTTTCTT  722 Tat.kemp
678  TAGGTCAAGG TGCAGCCTAT AAGGTGGGAA GTAATGGGCT ACATTTTCTC  727 Ura.rudd
686  TAGGTCAAGG TGTAGCCTAT GGAGTGGGAA GAAATGGGCT ACATTTACTT  735 Dry.nite
683  TAGGTCAAGG TGTAGCTTAT GAAATGGAAA GAAATGGGCT ACATTTATTT  732 2Mus.ave
679  TAGGTCAAGG TGTAGCCTAT GAAGAGGAAA GAAATGGGCT ACATTTCCTT  728 1Gli.jap
678  TAGGTCAAGG TGTAGCCTAT GAAGAGGAAA GAAATGGGCT ACATTTCCTT  727 2Gli.jap
683  TAGGTCAAGG TGTAGCTAAT GAAGTGGGAA GAAATGGGCT ACATTTTCTT  732 1Hyd.hyd
665  TAGGTCAAGG TGTAGCCAAT GGAGTGGGAA GAAATGGGCT ACATTTTCTT  714 Cav.porc
647  TAGGTCAAGG TGTAGCTAAT AAAGTGGAGA GAAATGGGCT ACATTTTCTT  696 Bat.suil
648  TAGGTCAAGG TGTAGCTAAT AAAGTGGAGA GAAATGGGCT ACATTTTCTT  697 Bat.jane
646  TAGGTCAAGG TGTAGCCAAT AAAGTGGGAA GAAATGGGCT ACATTTTCTT  695 1Cry.hot
647  TAGGTCAAGG TGTAGCCAAT AAAGTGGGAA GAAATGGGCT ACATTTTCTT  696 2Cry.hot
```

```
652  TAGGTCAAGG TGTAGCCAAT AAAGTGGAAA GTAATGGGCT ACATTTTCTT  701 Cry.dama
647  TAGGTCAAGG TGTAGCCAAT AAAGTGGGAA GAAATGGGCT ACATTTTCTT  696 Geo.cape
652  TAGGTCAAGG TGTAGCCAAT GAAGTGGGAA GAAATGGGCT ACATTTTCTT  701 Hel.arge
652  TAGGTCAAGG TGTAACCAAT GAAATGGGAA GAAATGGGCT ACATTTTCTT  701 Het.glab
660  TAGGTCAAGG TGTAACTTAT GAGATGGAAA -AAATGGGCT ACATTTTCTT  708 Thr.swin
659  TAGGTCAAGG TGTAGCCAAT GAAGTGGAGT GCAATGGGCT ACATTTTCTA  708 2Cra.cas
650  TAGGTCAAGG TGTAGCCTAT GGAGTGGAGC GAAATGGGCT ACATTTTCTA  699 Dip.ordi
647  TAGGTCAAGG CGTAGCTTAT GAGGTGGAAA GTAATGGGCT ACATTTTCTG  696 1Ped.cap
647  TAGGTCAAGG CGTAGCTTAT GAGGTGGAAA GTAATGGGCT ACATTTTCTG  696 2Ped.cap
635  TAGGTCAAGG TGTAGCCCAT GAAGTGGGAA GTAATGGGCT ACATTTTCT-  683 Ath.macr
631  TAGGTCAAGG TGTAGCCAAT GGAGTGGAAA GAAATGGGCT ACATTTTCTT  680 Cte.boli
630  TAGGTCAAGG TGTAGCCAAT GGAGAGGGAA GAAATGGGCT ACATTTTCTT  679 Cav.guia
625  TAGGTCAAGG TGTAGCCCAT GAAGTGGAAA GAAATGGGCC ACATTTTCTT  674 Chi.lani
629  TAGGTCAAGG TGTAACTAAT GAAGTGGAAA GCAATGGGCT ACATTTTCTT  678 Cap.pilo
629  TAGGTCAAGG TGTAGCCTAT GGGGTGGGAA GAAATGGGCT ACATTTTCTT  678 Das.punc
638  TAGGTCAAGG TATAGCCTAT GAAACGGGAA GCAATGGGCT ACATTTTCTT  687 Ere.dors
632  TAGGTCAAGG TGTAGCCTAT GGAGTGGGAA GAGATGGGCT ACATTTTCTA  681 Hys.afri
628  TAGGTCAAGG TGTAGCCAAA NAAGTGGAAA GAAATGGGCT ACATTTTCTT  677 Myo.coyp
627  TAGGTCAAGG TGTAGCTAAT GAAGTGGTAA GAAATGGGCT ACATTTTCTT  676 Oct.degu
630  TAGGTCAAGG TGTAACCAAT GAAGTGGGAA GAAATGGGCT ACATTTTCTT  679 Pro.long
635  TAGGTCAAGG TGTAGCCTAT AAGGTGGAAA GAAATGGGCT ACATTTTCTA  684 Apl.rufa
639  TAGGTCAAGG TGTAGCCAAT GAGGTGGGAA GAAATGGGCT ACATTTTCTT  688 Cas.cana
644  TAGGTCAAGG TGTAGCCAAT GAAGTGGAGT GCAATGGGCT ACATTTTCTA  693 1Cra.cas
627  TAGGTCAAGG TGTAGCCAAT AAAGTGGTAA GAAATGAGCT ACATTTCTAA  676 Cte.gund
635  TAGGTCAAGG TGTAGCTTAT GGGGTGGGAA GAAATGGGCT ACATTTTCTA  684 Das.nove
645  TAGGTTAAGG TGTAGCCAAT GAAGTGGAGT GAAATGGGCT ACATTTTCTG  694 Geo.burs
624  TAGGTCAAGG TGTAGCCTAT GAGTTGGAAA GCAATGGGCT ACATTTTCTT  673 Ger.vall
632  TAGGTCAAGG TGTAGCTTAT GAAGTGGAAA GCAATGGGCT ACATTTTCTC  681 Jac.jacu
631  TAGGTCAAGG TGTAGCCTAT GAGGTGGGAA GAAATGGGCT ACATTTTCTT  680 Lop.flav
628  TAGGTCAAGG TGTAGCTTAT GGAGTGGGAA GTAATGAGCT ACATTTTCTT  677 Osg.band
633  TAGGTCAAGG CGTAGCTTAT GAGGTGGAAA GTAATGGGCT ACATTTTCTG  682 3Ped.cap
634  TAGGTCAAGG TGTAACTTAT GAAATGGAAT GAAATGGGCT ACATTTTCTG  683 Per.flav
636  TAGGTCAAGG TGTAGCCTAT AGAGTGGAAA GTAATGGGCT ACATTTTCTA  685 Syl.audu
637  TAGGTCAAGG TGTAGCCTAT AGAGTGGGAA GAAATGGGCT ACATTTTCTA  686 Sci.nige
635  TAGGTCAAGG TGTAGCCTAT AAGGTGGGAA GAAATGGGCT ACATTTTCTA  684 Spe.trid
638  TAGGTCAAGG TGTAGCTTAT GAAGTGGAAA GAAATGGGCT ACATTTCCTT  687 Gra.muri
635  TAGGTCAAGG TGTAGCTAAT GAAGTGGGAA GAAATGGGCT ACATTTTCTT  684 2Hyd.hyd
618  TAGGTCAAGG TGTAACCCAT GGGGCGGGGA -AAATGGGCT ACATTTTCTT  666 Pet.typi
635  TAGGTCAAGG TGTAGCCTAT GAAGTGGGAA GTAATGGGCT ACATTTTCTT  684 Ony.leuc
635  TAGGTCAAGG TGTAGCTTAT GAGGTGGGAA GCAATGGGCT ACATTTTCTT  684 Per.poli
634  TAGGTCAAGG TGTAGCCTAT GGAGTGGGAA GTAATGGGCT ACATTTTCTC  683 Neo.albi
634  TAGGTCAAGG TGTAGCCTAT GGAGTGGGAA GTAATGGGCT ACATTTTCTA  683 Neo.flor
635  TAGGTCAAGG TGTAGCCCAT GGAGTGGGAA GTAATGGGCT ACATTTTCTT  684 Ony.torr
637  TAGGTCAAGG TGTAGCTTAT GAGATGGGAA GCAATGGGCT ACATTTTCTT  686 Per.erem
635  TAGGTCAAGG TGTAGCTTAT GAGATGGGAA GCAATGGGCT ACATTTTCTT  684 Per.goss
```

```
636  TAGGTCAAGG TGTAGCTTAT GAGGTGGGAA GCAATGGGCT ACATTTTCTT  685 Per.mani
636  TAGGTCAAGG TGTAGCTTAT GAGGTGGGAA GCAATGGGCT ACATTTTCTT  685 Per.mela
637  TAGGTCAAGG TGTAGCCTAT GAAGTGGGAA GCAATGGGCT ACATTTTCTC  686 Sig.hisp
637  TAGGTCAAGG TGTAGCCTAT GAAGTGGAAA GCAATGGGCT ACATTTTCTT  686 Sig.masc
635  TAGGTCAAGG TGTAGCCTAT GGAGTGGGAA GTAATGGGCT ACATTTTCTT  684 Ony.aren
635  TAGGTCAAGG TGTAGCTTAT GAGATGGGAA GCAATGGGCT ACATTTTCTT  684 2Per.leu
237  TAGGTCAAGG TGTAGCCAAT GAAGTGGAAA GAAATGGGCT ACATTTTCTT  286 Dol.pata
237  TAGGTCAAGG TGTAGCCTAT AAGGTGGAAA GAAATGGGCT ACATTTTCTA  286 1Pet.leu
237  TAGGTCAAGG TGTAGCCTAT AAGGTGGAAA GAAATGGGCT ACATTTTCTA  286 2Pet.leu
237  TAGGTCAAGG TGTAGCCAAT AAGGTGGAAA GAAATGGGCT ACATTTTCTA  286 1Pet.pet
235  TAGGTCAAGG TGTAGCCAAT AAGGTGGAAA GAAATGGGCT ACATTTTCTA  284 2Pet.pet
239  TAGGTCAAGG TGTAGCCTAT AGGGTGGAAA GAAATGGGCT ACATTTTCTA  288 1Pte.vol
239  TAGGTCAAGG TGTAGCCTAT AGGGTGGAAA GAAATGGGCT ACATTTTCTA  288 2Pte.vol
239  TAGGTCAAGG TGTAGCCTAT AGGGTGGAAA GAAATGGGCT ACATTTTCTA  288 3Pte.vol
237  TAGGTCAAGG TGTAGCCTAT AGAGTGGAAA GAAATGGGCT ACATTTTCTA  286  Sci.lis
237  TAGGTCAAGG TGTAGCCTAT AGAGTGGAAA GAAATGGGCT ACATTTTCTA  286 Sci.vulg
236  TAGGTCAAGG TGTAGCCTAT GAGGTGGGAA GAAATGGGCT ACATTTTCTA  285 1Tam.sib
236  TAGGTCAAGG TGTAGCCTAT GAGGTGGGAA GAAATGGGCT ACATTTTCTA  285 2Tam.sib
218  TAGGTCAAGG TGTAACCCAT GAGGTGGAAA GAAATGGGCT ACATTTTCTA  267 Bas.astu
218  TAGGTCAAGG TGTAACCCAT GGAACGGGAA GAAATGGGCT ACATTTTCTT  267 Spi.puto
218  TAGGTCAAGG TGTAACCTAT GAGGTGGGAA GAAATGGGCT ACATTTTCTA  267 Pro.loto
218  TAGGTCAAGG TGTAGCCTAT GAAACGGGAA GAAATGGGCT ACATTTTCTT  267 Mep.meph
218  TAGGTCAAGG TGTAACCCAT GGGATGGAAA GAAATGGGCT ACATTTTCTA  267 Odo.rosm
217  TAGGTCAAGG TGTAACTTAT GAGGTGGGAA GAAATGGGCT ACATTTTCTA  266 Can.latr
218  TAGGTCAAGG TGTAACCCAT GGGATGGGAA GAAATGGGCT ACATTTTCTA  267 Uro.cine
218  TAGGTCAAGG TGTAACCCAT GGGGTGGGAA GAAATGGGCT ACATTTTCTA  267 Urs.amer
217  TAGGTCAAGG TGTAACCAAT GGAACGGAAA GAAATGGGCT ACATTTTCTT  266 Myd.marc

     801        811        821        831        841      850
     |          |          |          |          |        |
730  ----TTCCCA GAGAACATT- ---ACGAA-A CC-TTTATGA AAC---TAAA  766 Rat.norv
730  -----TTTAA GAACAT---- ---ACGAT-A CCCTTTATGA AAC---TAAA  763 Hyl.stel
728  ----TTTTAA GAACATT--- ---ACTTT-A CCCTTTATGA AAC---TAAA  763 Mus.croc
729  ----TTAAAA GAACATT--- ---ACTAT-A CCCTTTATGA AAC---TAAA  764 Mus.cook
728  -----TCTAA GAACATC--- ---ACGATCA CCCTTTATGA AAC---TAAA  763 Mas.eryt
733  ----TTAAAA GAACAAT--- ---ACGAT-A TCCTTTATGA AAC---TAAA  768 Mus.matt
728  ----TTAAAA GAACATC--- ---ACCAC-A CCCTTTATGA AAC---GAAA  763 Mus.plat
729  ----TTCTAA GAACATT--- ---ACGAT-A TCCTTTATGA AAC---TAAA  764 Mus.setu
728  ----TTAAAA GAACATC--- ---ACCAC-A CCCTTTATGA AAC---GAAA  763 Mus.saxi
721  -----TCAGA GAACATTC-- ---ACGCT-A CCCTTTATGA AACC--TAAA  757 Cri.gamb
745  -----AAAAA GAACATTT-- ---ACGTT-A CCCTCCATGA AAT---GGGA  780 Mac.inge
725  -----TTCAA GAACATT--- ---ACGAT-A CCCTTTATGA AAC---TAAA  759 Mic.niva
726  ----ACTCAA GAACACTC-- ---ACGCT-A CCCTCTATGA AATA--GGGA  763 Nes.rufu
727  -----AAAAA GAACATT--- ---ACGAT-A CCCTTATTGA AACA--TAA-  761 1Per.leu
720  -----AAAAA GAACAC---- ---ACGAA-A CCCTTTATGA AAT---TAAA  753 Aco.cahi
```

```
730  -----GACAA GAACATTC-- ---ACGAT-A CCCTTTATGA AAT---TAAA  765 Cri.migr
728  ----TCCCAA GTACATTC-- ---ACGTC-A ACTTTTATGA AACC--TAAA  765 Gli.glis
730  -----AAAAA GAACATTA-- ---ACGAT-G TCTTTCATGA AATT--TTAA  766 Ger.nige
731  ----GCCCAA GAACATT--- ---ACGAA-A TCCTTTATGA AAC---TAAA  766 Leo.edwa
729  -----CTAAA GAACAGTT-- ---ACGCT-A TCCTCTATGA AACT--TAGA  765 Mes.aura
734  ----CACCAA GTATATCC-- ---ACGTT-A ACTCTTATGA AATC--TGAG  771 1Mus.ave
730  ----ATAAAA GAACATT--- ---ACTAT-A CCCTTTATGA AAC---TAAA  765 Mus.musc
735  ----TTTTAA GAACATT--- ---ACTAT-A CCCTTTATGA AAT---TAAA  770 Mus.paha
723  -----AAAAA GAACATT--- ---ACGAT-G TCCTTTATGA AATC--TAAA  758 Tat.kemp
728  -----ATAAA GAATAT---- ---ACGAA-A CCCTTTATGA AAC---TAAA  761 Ura.rudd
736  ----AGCCAA GTATATATT- ---ACGAT-A CCTTTCATGA AAC---TGAA  773 Dry.nite
733  ----CACCAA ATACATTT-- ---ACGTT-A ACTCTTATGA AACC--TGAG  770 2Mus.ave
729  -----TATAA GTATATC--- ---ACGCA-A ACTCTTATGA AAC---TAAA  763 1Gli.jap
728  -----TACAA GTACATC--- ---ACGTA-A ACTCTTATGA AAT---TAAA  762 2Gli.jap
733  ----ACCAAA GAACAT---- -AAACGTA-A ATCTTTATGA AACC--TAAA  770 1Hyd.hyd
715  ----ACCCAA GAACATT--- -AAACGCA-A ATCTTTATGA AATTC-AAGA  754 Cav.porc
697  ----TACAAA GAACACTT-- -AAACGGA-A ACTATTATGA AATC--TAAT  736 Bat.suil
698  ----TATAAA GAACACTT-- -AAACGGA-A ACTATTATGA AATT--TGAT  737 Bat.jane
696  ----AACAAA GAACATCT-- -AAACGAA-A GCCATTATGA AATT--TAAT  735 1Cry.hot
697  ----AACGAA GAACATCT-- -AAACGAG-A ACTATTATGA AATT--TAGT  736 2Cry.hot
702  ----GACGAA GAACATTT-- -AAACGAA-A GCTATTATGA AATT--CAAT  741 Cry.dama
697  ----CGTTAA GAACACT--- -AAACGAA-A ACTATTGTGA AAAT--TAAT  735 Geo.cape
702  ---TTTTCAA GAACATTA-- -AAACGAA-A GTCACTATGA AAATAATATT  744 Hel.arge
702  ----ATACAA GAACAT---- -AAACGAT-A GTTACCATGA AATC--TGAT  739 Het.glab
709  ------AACA GAATATCC-- ---ACGAT-A ATCATTATGA AAC---TAAT  743 Thr.swin
709  ----CCAGTA GAACATC--- ---ACGAA-T TTCTCTCTAA AAA---CAGA  744 2Cra.cas
700  ----TCACTA GAACATT--- ---ACGAA-A CCCTACGTGA AATC--CATA  736 Dip.ordi
697  ----TTTACA GAACAAATTT -ATACGCT-A TCTCTTATGA AACC--TAAG  738 1Ped.cap
697  ----TTTACA GAACAAATTT -ATACGTT-A TCTCTTATGA AACC--TAAG  738 2Ped.cap
684  ----AACCAA GAACATCC-- ---ACGAC-A ATCTTTATGA AAAC--TGAA  721 Ath.macr
681  CATTACAAAA GAATATT--- --AACGGT-G ATCTCTATGT AACC--TAAA  722 Cte.boli
680  ----ACCCAA GAACATT--- -AAACGCA-A ATCTTTATGA AATT--CAAA  718 Cav.guia
675  -----ACCAA GAAAATTAAT T-AACGAC-A GTCACTATGA AACT--TAGA  715 Chi.lani
679  ---CTTCCAA GAAAACAAC- --AACAAGTA ATCTTTATGA CAT---TAAA  719 Cap.pilo
679  ----ATCAAA GAACATTCT- ---ACGCA-A ATCCTCATGA AACT---GAG  716 Das.punc
688  -----TATAA GAAAATTATT --TACGGC-A ACCCTTATGA AACT--TCG-  726 Ere.dors
682  -----ACCAA GAATATAT-- ---ACGAC-A ATCTTTATGA AAAT--TAAA  718 Hys.afri
678  -----TTTAA GAACATCT-- --AACAGT-A ATCCTTATGA AAC---TAAG  714 Myo.coyp
677  ---ACCTTAA GAATATT--- --AACAGT-A ATCTTTATGA AATC--TAAA  715 Oct.degu
680  ----TTACAA GAACACATAC --AACAGT-T ACCCTTATGA AAAT--TAAG  720 Pro.long
685  ----ATACTA GAACACCCCT ---ACAAC-A GTTTTTATGA AAC---TAAA  723 Apl.rufa
689  ---GCCCTAA GAAAACT--- --AACGGA-A ACTCTTATGA AACT--GAAG  727 Cas.cana
694  ----CCAGTA GAACATC--- ---ACGAA-T TTCTCTCTGA AAA--ACAGA  730 1Cra.cas
677  ----AATTTA GAAAAT---- ---ACGAA-G TCCATTATGA AAAC--TAAA  712 Cte.gund
685  -----ATAAA GAGCAAAT-- ---ACAAA-A AACTTAATGA AAC-AATTTA  722 Das.nove
```

```
695  ----GTAACA GAACACT--- ---ACGAA-G TTCTCTCTGA AAC--ACAGA  731 Geo.burs
674  -----TGAAA GAACATT--- ---ACGAT-G CCTCTTATGA AATT--TAAG  709 Ger.vall
682  ----TCTCGA GAACATA--- ---ACGAT-A CCCTTCTTGA AAC---AGAA  717 Jac.jacu
681  ----TAACAA GAACAT---- ---ACGAA-A ACCTTTATGA AAC---TAAA  715 Lop.flav
678  -----AAAAA GAACATT--- ---ACGAT-A CCCTTATTGA AAC---ATAA  712 Osg.band
683  -----TTTCA GAATAAATTA T--ACGTT-A TCTCTTATGA AAC--CTAAG  722 3Ped.cap
684  -----TTTAA GAATAC---- ---ACGAA-A CCCTGCCTGA AAT--AGGCA  718 Per.flav
686  -----CTTCA GAACAT---- ---ACGAA-A GCCCTTATGA AACT-CTAAG  721 Syl.audu
687  -TTTCTCCTA GAACACTTC- ---ACGAC-A GCAGTTATGA AACA-TGACT  729 Sci.nige
685  --TTTTCCTA GAACAAATAT ---ACAAT-A GCTTTTATGA AATT---CAG  725 Spe.trid
688  ----TTTCAA GTACATTC-- ---ACGTA-A ACTTTTATGA AACC--TAAA  725 Gra.muri
685  ----ACCAAA GAACATAA-- ---ACGTA-A ATCTTTATGA AACCT--AAA  722 2Hyd.hyd
667  -----GTTAA GAACAGTTCT ---ACGAT-T ACCATTATGA AATC--TAAT  705 Pet.typi
685  -----AAAAA GAACATT--- ---ACGAT-A CCCTTTTTGA AACA--AAA-  719 Ony.leuc
685  -----AAAAA GAACATT--- ---ACGAT-A CCCTTATTGA AACA--TAA-  719 Per.poli
684  -----AACAA GAACATT--- ---ACGAA-A CCCTTTATGA AATC--TAAA  719 Neo.albi
684  -----AATAA GAACATT--- ---ACGAA-A CCCTTTATGA AATC--TAAA  719 Neo.flor
685  -----AAAAA GAACATT--- ---ACGTT-A CCCTTTCTGA AACA---AAA  719 Ony.torr
687  -----AAAAA GAACATT--- ---ACGAT-A TCCTTATTGA AACA--TAA-  721 Per.erem
685  -----AAAAA GAACATT--- ---ACGAA-A CCCTTATTGA AATA--TAA-  719 Per.goss
686  ---AAAAAAA GAACATT--- ---ACGAT-A CCCTTATTGA AACA--TAA-  722 Per.mani
686  -----AAAAA GAACATT--- ---ACGAT-A CCCTTATTGA AACA--TAA-  720 Per.mela
687  -----CACAA GAACATT--- --AACGAC-A CCCTCTATGA AACT--AGA-  722 Sig.hisp
687  -----CACAA GAACATT--- --AACGAT-A CCCTCTATGA AACT--AGA-  722 Sig.masc
685  -----AAAAA GAACATT--- ---ACGAT-A CCCTTTCTGA AACA--AAA-  719 Ony.aren
685  -----AAAAA GAACATT--- ---ACGAT-A CCCTTATTGA AACA--TAA-  719 2Per.leu
287  ----ACCAAA GAACAT---- ---ACGCA-A ATCTTTATGA AAAT--TAAA  322 Dol.pata
287  --GCCTCCTA GAATAATAAA -CAACGAC-A GCTTCCATGA AAC---TTCA  329 1Pet.leu
287  --GCCTCCTA GAATAATAAA -CAACGAC-A GCTTCCATGA AAC---TTCA  329 2Pet.leu
287  --GCCCCCTA GAACAACAAA TCAACGAC-A GCTTCCATGA AAC---TTCA  330 1Pet.pet
285  --GCCCCCTA GAACAACAAA TCAACGAC-A GCTTCCATGA AAC---TTCA  328 2Pet.pet
289  -GCATTCATA GAACAACACA ---ACGAT-A ACTTATATGA AACA--TATA  331 1Pte.vol
289  -GCATTCATA GAACAACACA ---ACGAT-A ACTTATATGA AACA--TATA  331 2Pte.vol
289  -GTATTCATA GAACAACACA ---ACGAT-A ACTTATATGA AACA--TATA  331 3Pte.vol
287  --ATTTCCTA GAATATA--- --CACGAT-A GCAATCATGA AATATGATTA  328  Sci.lis
287  --ATTTCCTA GAATATA--- --CACGAT-A GCAATCATGA AATATGATTA  328 Sci.vulg
286  --CTTTACTA GAACAACACT --CACGAT-A ACTTTCATGA AACAT-TGAA  329 1Tam.sib
286  --CTCTACTA GAACAACACT --CACGAT-A ACTTTCATGA AACAC-TGAA  329 2Tam.sib
268  -----AATAA GAACACACC- --CACGGA-A GTTTTTATGA AAC---TAAA  305 Bas.astu
268  ----TAGCAA GAATATCCT- --CACGAA-A GACTTTATGA AAT---TAAA  306 Spi.puto
268  --ATAAATAA GAATATATAC --CACGGA-A ATTTTTATGA AAC---TAAA  309 Pro.loto
268  ----TAACAA GAACATTCT- ---ACGAA-A GACTTTATGA AAT---TAAA  305 Mep.meph
268  -----AGTAA GAACAGCCAT ---ACAAA-A GTTTTTATGA AAT---TAAT  305 Odo.rosm
267  ----CTCCAA GAACATTTC- ---ACGAA-T GTTTTTATGA AAT---TAAA  304 Can.latr
268  ----CTCCAA GAATACTTC- ---ACGGA-A GTTTTTATGA AAC---TAAA  305 Uro.cine
```

```
268  -----TTCAA GAACAACCT- ---ACGAA-A GTTTTTATGA AAC---TAAA  304 Urs.amer
267  -----AAAAA GAATACCCA- ---ACGGA-A GTTCCTATGA AAA---CAAG  303 Myd.marc

     851        861        871        881        891      900
     |          |          |          |          |        |
767  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  816 Rat.norv
764  GGACGAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  813 Hyl.stel
764  GGACTAGGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  813 Mus.croc
765  GGATTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  814 Mus.cook
764  GGACAAAGGA GGATTTAGTA GTAAATTAAG AGTAGAGAGC TTAATTGAAT  813 Mas.eryt
769  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  818 Mus.matt
764  GGATAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAGTTGAAT  813 Mus.plat
765  GGACTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  814 Mus.setu
764  GGATAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAGTTGAAT  813 Mus.saxi
758  GGACCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  807 Cri.gamb
781  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  830 Mac.inge
760  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  809 Mic.niva
764  GGGCCAAGGA GGATTTAGTA GTAAATTAAG AAT-GAGAGC TTAATTGAAC  812 Nes.rufu
762  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  811 1Per.leu
754  GGATAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  803 Aco.cahi
766  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  815 Cri.migr
766  AGTAAAAGGC GGATTTAGTA GTAAGCTAAG AATAGAGAGC CTAGCTGAAT  815 Gli.glis
767  AGATCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGTGC TTAATTGAAT  816 Ger.nige
767  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  816 Leo.edwa
766  GGACAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  815 Mes.aura
772  AGTAGAAGGA GGATTTAGTA GTAAGCTAGG AATAGAGAGC CTAACTGAAT  821 1Mus.ave
766  GGACTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  815 Mus.musc
771  GGACTAGGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  820 Mus.paha
759  GGATCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGTGC TTAATTGAAT  808 Tat.kemp
762  GGCTAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGTGC TTAATTGAAT  811 Ura.rudd
774  AGAAAAAGGA GGATTTAGTA GTAAGCTAAG AATAGAGTGC TTAGCTGAAT  823 Dry.nite
771  AGTAGAAGGA GGATTTAGTA GTAAGCTAGG AGTAGAGAGC CTAACTGAAT  820 2Mus.ave
764  AGTACAAGGA GGATTTAGTA GTAAGCTATG AATAGAGAGC CTAGCTGAAT  813 1Gli.jap
763  AGTATAAGGA GGATTTAGCA GTAAGCTATG AATAGAGAGC CTAGCTGAAT  812 2Gli.jap
771  GATAGAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAC  820 1Hyd.hyd
755  TCTAAGGGAG GGATTTAGTA GTAA-TTAAG AATAGAGAGC TTGATTGAAC  803 Cav.porc
737  GGTGTAAGGA GGATTTAGTA GTAAACCAAG AATAGAGAGC TTGGTTG     783 Bat.suil
738  AGTATAAGGA GGATTTAGTA GTAAATCAAG AATAGAGAGC TTGGTTG     784 Bat.jane
736  AGTGCAAGGA GGATTTAGTA GTAAATTGAG AATAGAGAGC TTAATTG     782 1Cry.hot
737  AGTGTAAGGA GGATTTAGTA GTAAATTGAG AATAGAGAGC TTAATTG     783 2Cry.hot
742  AGCACAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG     788 Cry.dama
736  AGTCTAAGGA GGATTTAGCA GTAAATTAAG AATAGAGTGC TTAATTG     782 Geo.cape
745  GACCTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAGTTG     791 Hel.arge
740  AACACAAGGT GGATTTAGTA GTAAAACAAG AATAGAGAGC TTGTTTG     786 Het.glab
744  GATAGAAGGC GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG     790 Thr.swin
```

```
745  GAATAAAGGA GGATTTAGTA GTAAATTAAG AG--GAGAGC TTAAT        787 2Cra.cas
737  GGGTAAAGGA GGATTTAGTA GTAAATTAAA ACTAGAGAGC TTAAT        781 Dip.ordi
739  AGACAAAGGC GGATTTAGTA ---AATTAAG ATTAGAGAGC TTAAT        780 1Ped.cap
739  AGACAAAGGC GGATTTAGTA GTAAATTAAG ATTAGAGAGC TTAAT        783 2Ped.cap
722  GATCCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGCGC TTAATTG      768 Ath.macr
723  GATTTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG      769 Cte.boli
719  GATCTAAGGA GGATTTAGTA GTAAATCAAG AATAGAGAGC TTGATTG      765 Cav.guia
716  GATCTAAGGA GGATTTAGCA GTAAATTAAG AATAGAGAGC TTAGTTG      762 Chi.lani
720  GATTTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG      766 Cap.pilo
717  GATATAAGGA GGATTTAGTA GTAAATTAAG AACAGAGAGC TTAATTG      763 Das.punc
727  GGTCTAAGGA GGATTTAGCA GTAAATTAAG AATAGAGAGC TTAATTG      773 Ere.dors
719  GATCCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGCGC TTAATTG      765 Hys.afri
715  GATTTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTGATTG      761 Myo.coyp
716  GATTTAAGGA GGATTTAGCA GTAAATTAAG AATAGAGAGC TTATTG       761 Oct.degu
721  GGTTTAAGGA GGATTTAGTA GTAAATTAAG AACAGAGAGC TTAATTG      767 Pro.long
724  AGCTCAAGGC GGATTTAGTA GTAAGCTGAA AATAGAGTGT TTAGCTG      770 Apl.rufa
728  AAGATAAGGA GGATTTAGCA GTAAATTGAA CTTAGAAAGC TCAATTG      774 Cas.cana
731  GAATGAAGGA GGATTTAGTA GTAAGTTAAA AATAGAGAGT TTAACTG      777 1Cra.cas
713  GGATGAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG      759 Cte.gund
723  AGACTAAGGT GGATTTAGTA GTAAGCTAAA AATAGAGAGT TTAGCTG      769 Das.nove
732  GAATGAAGGA GGATTTAGAA GTAAGTTAAA AATAGAGTGT TTAGCTG      778 Geo.burs
710  AGGTCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGTGC TTAATTG      756 Ger.vall
718  GGAATAAGGA GGATTTAGTA GTAAATTAAG AGTAGAGTGC TTAATTG      764 Jac.jacu
716  GGATTAAGGA GGATTTAGCA GTAAATTAAG AATAGAGAGC TTAATTG      762 Lop.flav
713  GGGCAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG      759 Osg.band
723  AGACAAAGGC G-ATTTAGTA GTAAATTAAG ATTAGAGAGC TTAATTG      768 3Ped.cap
719  GGGTGAAGGA GGATTTAGCA GTAAACTAAA AATAGAGTGT TTAGTTG      765 Per.flav
722  AGCCAAAGGA GGATTTAGTA GTAAATTAAG AATAGAGTGC TTAATTG      768 Syl.audu
730  AGCCTAAGGC GGATTTAGTA GTAAGTTGGG AATAGAGTGC CCAACTG      776 Sci.nige
726  AGCATAAGGC GGATTTAGTA GTAAGTTAAG AATAGAGAGC TTAACTG      772 Spe.trid
726  AGTATAAGGA GGATTTAGTA GTAAGCTAAG AATAGAGAGC TTAGCTG      772 Gra.muri
723  GATAGAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTG      769 2Hyd.hyd
706  GGTCTAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC ATTATTG      752 Pet.typi
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Ony.leuc
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Per.poli
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Neo.albi
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Neo.flor
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Ony.torr
722  GGACTAAGGA GGATTTAGTA GTAAATTAAG AA                      753 Per.erem
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Per.goss
723  GGACAAAGGA GGATTTAGTA GTAAATTAA                          751 Per.mani
721  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      752 Per.mela
723  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      754 Sig.hisp
723  GGACCAAGGA GGATTTAGTA GTAAATTAAG AA                      754 Sig.masc
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                      751 Ony.aren
```

```
720  GGACAAAGGA GGATTTAGTA GTAAATTAAG AA                    751 2Per.leu
323  GACTCAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAC  372 Dol.pata
330  AGCTCAAGGT GGATTTAGTA GTAAGCTAAG AATAGAGAGC TTAGCTGAAT  379 1Pet.leu
330  AGCTCAAGGT GGATTTAGTA GTAAGCTAAG AATAGAGAGC TTAGCTGAAT  379 2Pet.leu
331  AGCCCAAGGT GGATTTAGTA GTAAGCTAAG AATAGAGAGC TTAGCTGAAT  380 1Pet.pet
329  AGCCCAAGGT GGATTTAGTA GTAAGCTAAG AATAGAGAGC TTAGCTGAAT  378 2Pet.pet
332  AGTCCAAGGC GGATTTAGTA GTAAGCCAAG AATAGAGAGC TTGACTGAAT  381 1Pte.vol
332  AGTCCAAGGC GGATTTAGTA GTAAGCCAAG AATAGAGAGC TTGACTGAAT  381 2Pte.vol
332  AGTCCAAGGC GGATTTAGTA GTAAGCCAAG AATAGAGAGC TTGACTGAAT  381 3Pte.vol
329  AGCCCAAGGC GGATTTAGTA GTAAGTTAGG AATAGAGTGC CTAACTGAAC  378  Sci.lis
329  AGCCCAAGGC GGATTTAGTA GTAAGTTAGG AATAGAGTGC CTAACTGAAC  378 Sci.vulg
330  AGTCAAAGGC GGATTTAGTA GTAAGTTAGG AATAGAGAGC CTAACTGAAT  379 1Tam.sib
330  AGTCAAAGGC GGATTTAGTA GTAAGTTAGG AATAGAGAGC CTAACTGAAT  379 2Tam.sib
306  AACTGAAGGC GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  355 Bas.astu
307  GACCGAAGGA GGATTTAGCA GTAAATTAAG AATAGAGAGC TTAATTGAAT  356 Spi.puto
310  AATCAAAGGT GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAGTTGGAT  359 Pro.loto
306  GGCCGAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  355 Mep.meph
306  AACTGAAGGT GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAC  355 Odo.rosm
305  AACTGAAGGA GGATTTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAT  354 Can.latr
306  AACTGAAGGA GGATTTAGTA GTAAGTTAAG AATAGAGAGC TTAACTGAAT  355 Uro.cine
305  AACTAAAGGT GGATTTAGCA GTAAACCAAG AATAGAGAGC TTGGTTGAAT  354 Urs.amer
304  AACCGAAGGA GGATCTAGTA GTAAATTAAG AATAGAGAGC TTAATTGAAC  353 Myd.marc

     901        911        921        931        941       950
     |          |          |          |          |         |
817  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAG  863 Rat.norv
814  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  860 Hyl.stel
814  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  860 Mus.croc
815  CGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  861 Mus.cook
814  AGA--GCAAT GAAGTACGTA CACACCGCCC GTCACCCTCC T-CAAGTTAA  860 Mas.eryt
819  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  865 Mus.matt
814  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  860 Mus.plat
815  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  861 Mus.setu
814  CGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  860 Mus.saxi
808  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAACTAG  854 Cri.gamb
831  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAG  877 Mac.inge
810  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAACTAA  856 Mic.niva
813  AGA--GCAAT GAAGTACGTA CACACCGCCC GTCACCCTCC T-CAAATTAG  859 Nes.rufu
812  AGC--GCAAT GAAGTACGTA CACACCGCCC GTCACCCTCC T-CAAATTAA  858 1Per.leu
804  AGA--GCAAT GAAGTGCGCA CACACCGCCC GTCACCCTCC T-CAAACTAA  850 Aco.cahi
816  AGA--GCAAT GAAGTACGTA CACACCGCCC GTCACCCTCC T-CAAATTAG  862 Cri.migr
816  AGG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGCATG  862 Gli.glis
817  TGA--GCCAT GAAGCACGTA CACACCGCTC GTCACCCTCC TTCGAATAAC  864 Ger.nige
817  AGC--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  863 Leo.edwa
816  AGA--GCAAT GAAGTACGTA CACACCGCCC GTCACCCTCC T-CAAATTAA  862 Mes.aura
```

```
822  AGG--GCCAT TAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGCATT  868 1Mus.ave
816  TGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  862 Mus.musc
821  AGA--GCAAT GAAGTACGCA CACACCGCCC GTCACCCTCC T-CAAATTAA  867 Mus.paha
809  AGAAAGCTCT GAAGCGCGCA CACACCGCCC GTCACCCTCC T-CAAATAAA  857 Tat.kemp
812  AGA--GCAAT GAAGTGCGCA CACATCGCCC GTCACCCTCC T-CAAACTAA  858 Ura.rudd
824  AAG--GCCAT GAAG                                         835 Dry.nite
821  AGG--GCCAT TAAG                                         832 2Mus.ave
814  AGG--GCCAT GAAG                                         825 1Gli.jap
813  AGG--GCCAT GAAG                                         824 2Gli.jap
821  TAG--GCAAT GAAGCACGTA CACACCGCCC GTCACCCTCC T-CAAATA--  865 1Hyd.hyd
804  TAG--GCCAT GAAGCACGTA CACACCGCCC GTCACCCTCC T-CAAGTA--  848 Cav.porc
373  TAG--GCCAT GAAGCACGT                                    389 Dol.pata
380  CGG--GCAAT AAAGC                                        392 1Pet.leu
380  CGG--GCAAT AAAGC                                        392 2Pet.leu
381  CGG--GCAAT AAAGC                                        393 1Pet.pet
379  CGG--GCAAT AAAGC                                        391 2Pet.pet
382  TGG--GCAAT AAAGC                                        394 1Pte.vol
382  TGG--GCAAT AAAGC                                        394 2Pte.vol
382  TGG--GCAAT AAAGC                                        394 3Pte.vol
379  TGG--GCAAT AAAGC                                        391  Sci.lis
379  TGG--GCAAT AAAGC                                        391 Sci.vulg
380  AGG--GCAAT AAAGC                                        392 1Tam.sib
380  AGG--GCAAT AAAGC                                        392 2Tam.sib
356  CGG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGCAGT  402 Bas.astu
357  GGG--GCCAT GAAGCGCGCA CACACCGCCC GTCACCCTCC T-CAAGTAAC  403 Spi.puto
360  TGG--GCCAT AAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGCAGT  406 Pro.loto
356  AGG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGTAAC  402 Mep.meph
356  TGG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAACAAC  402 Odo.rosm
355  AGG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGTAAT  401 Can.latr
356  AGG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGTAAT  402 Uro.cine
355  TAG--GGCAT GGAGCATGCA CACACCGCCC GTCACCCTCC T-CAAGTGGC  401 Urs.amer
354  -GG--GCCAT GAAGCACGCA CACACCGCCC GTCACCCTCC T-CAAGTAAA  399 Myd.marc

     951        961        971        981        991       1000
     |          |          |          |          |         |
864  ATTGACATT- -CACATATAC ATAATTTCAC TAACAAA--- ----------  898 Rat.norv
861  ATTCACTAA- ---ACTATTA ATAATTCCTA GTAATAAA-- ----------  894 Hyl.stel
861  ATTTATCAC- ---TATATAA ATAATTATAA CCCACTTA-- ----------  894 Mus.croc
862  ATACAACTT- ---ATTATAA ATAATTCCTA AGTAACTAA- ----------  896 Mus.cook
861  ATACACTTA- ---ATTATAC ATAATAACAA GTAATAAAA- ----------  895 Mas.eryt
866  ATAGCAATA- ATAAACATAG ATAATTATAA ATCA-CTTA- ----------  902 Mus.matt
861  ATTTCACCT- ----TAATAA ATATTCATAT ACTATCTAAT ATTTTATCAA  905 Mus.plat
862  ATATTTCTT- ---AGTATAA ATAATAATGA TAAATTATAT A---------  898 Mus.setu
861  ATTTCACCT- ----TAATAA ATATTCATAT ACTATCTAAT ATTTTATCAA  905 Mus.saxi
855  ACAAGTCCA- ---TAAAAAT ATATACATAA TACAACAGAC T-------AA  893 Cri.gamb
```

```
878  CCAAATGAA- --AATCACAC TTACATAAGC CCAATCAAAG GC--------  916 Mac.inge
857  ATAAACGAA- --AACTATAT ATAA-TAAAA TCAAACT--- ----------  889 Mic.niva
860  ACTAGCGAC- ----GCACCA ATAAATAATA TAAGTCAAAG CC--------  896 Nes.rufu
859  ACGATTATT- ---AAACCCT ATACATAATT ACATATACAA G---------  895 1Per.leu
851  TCTTCAAGT- ---AAAATAA ATAAAATTAA ACTAACTAGA A---------  887 Aco.cahi
863  ACTGCCGAT- ---ATTATAC CTAATACTTA TTAATAAGCC ----------  898 Cri.migr
863  AATATAATT- -TATCTATAA TTAAATCTCT GCAC------ ----------  894 Gli.glis
865  CCCGTACA-- -CCTAAATAT ATAATTTTAC GTACAT---- -------CAA  900 Ger.nige
864  ATTGATA--- -CAACCATAG ATAATAGCGC TAAAAAA--- ----------  896 Leo.edwa
863  GCTAACTGT- -TAACTATAC CTAATATTGA CTCATAAAC- ----------  899 Mes.aura
869  TACTATTAA- ---TATTAGT TTATATTACC TTACATACA- ----------  903 1Mus.ave
863  ATTAAACTT- ---AACATAA TTAATTTCTA GACATCCG-- ----------  896 Mus.musc
868  ATTTAATAT- ---AATATAA ATAATAACAA TTAACTTA-- ----------  901 Mus.paha
858  GAAAAAAAA- ------ACAA ATACATAATT CTCAAAAT-- -------AAA  891 Tat.kemp
859  ATACTAATA- ---CTATTAA ATAAATTACA TTAGAA---- -------AAA  893 Ura.rudd
866  -----CCCAA AAATTTAAAA ATCAAAAATT AACAA----- ----------  895 1Hyd.hyd
849  -----TCCAA GGGTTTTG-- -TACAAA-CT AATAA----- ----------  874 Cav.porc
403  ACTAATTCAA CTATAACAT- ATTAACGAAC AACAC----- ----------  436 Bas.astu
404  TACAT--CAA AATATAACAT ATTAACATTT ATTCA----- ----------  436 Spi.puto
407  AGTAATTCAA CCACAATATT ATTAACGGAT AAATTC---- ----------  442 Pro.loto
403  TCTAT--TTA GACACAACAT ATTACTCTGC CCTAA----- ----------  435 Mep.meph
403  ACCTC-AAAC ATACATAAAT ATGTACAAAA TA-------- ----------  433 Odo.rosm
402  AAGAC-ACAA CCATAACCAT ATTAACTTAA CT-AAA---- ----------  435 Can.latr
403  AACCCTACAA CAATAATCAT ATTAATCAAA CCCAAA---- ----------  438 Uro.cine
402  ACAGT-CAAA ATTTACACCT ATTGAAATTT AAATAA---- ----------  436 Urs.amer
400  TATAT--CAT ATTATAACAT ATCAAACAAT ATCGAG---- ----------  433 Myd.marc

     1001       1011       1021       1031       1041     1050
     |          |          |          |          |        |
899  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  943 Rat.norv
895  ----TTTATC AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  939 Hyl.stel
895  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  939 Mus.croc
897  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  941 Mus.cook
896  ----TTTATT AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  940 Mas.eryt
903  ----TTTATG AGAGGAGACA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  947 Mus.matt
906  CACATTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  954 Mus.plat
899  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  943 Mus.setu
906  CACATTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  954 Mus.saxi
894  CC----TACG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  938 Cri.gamb
917  ------TATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  959 Mac.inge
890  ----TTTACG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  934 Mic.niva
897  ------TCTG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  939 Nes.rufu
896  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  940 1Per.leu
888  ------TACG AGAGGAGACA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  930 Aco.cahi
899  ------TATG AGAGGAGACA GGTCG-TAAC AAGGTAAGCA TACTGGAAAG  941 Cri.migr
```

```
895  ----TTTGCA AGAGGAGATA AGTCG-TAAC ATGGTAAGCA TACTGGAAAG  939 Gli.glis
901  CC---TTACG AGAAGAGGTA AGTCG-TAAC AAGGTAAGCG TACTGGAAAG  946 Ger.nige
897  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  941 Leo.edwa
900  -----TTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  943 Mes.aura
904  ------TGCA AGAGGAGATA AGTCG-TAAC ATGGTAAGCA TACTGGAAAG  946 1Mus.ave
897  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  941 Mus.musc
902  ----TTTATG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  946 Mus.paha
892  CTA---TATG AGAGGAGGTA AGTCG-TAAC ATGGTAAGCA TACTGGAAAG  937 Tat.kemp
894  ----TTTACG AGAGGAGATA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  938 Ura.rudd
896  -----ATATG AGAGGAGACA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  939 1Hyd.hyd
875  -----ATATT AGAGGAGACA AGTCG-TAAC AAGGTAAG               906 Cav.porc
437  -----ATGCA AGAGGAGACA AGTCGCTAAC AAGGTAAGCA TACTGGAAAG  481 Bas.astu
437  ------CACA AGAGGAGACA AGTCGT-AAC AAGGTAAGCG TACTGGAAAG  479 Spi.puto
443  ----AATGTA AGAGGAGACA AGTCG-TAAC ACGGTAAGCA TACTGGAAGG  487 Pro.loto
436  ------TACA AGAGGAGACA AGTCGTCAAC AAGGTAAGCG TACTGGAAGG  479 Mep.meph
434  -----TATCT AGAGGAGATA AGTCGCTAAC AAGGTAAGCG TACTGGAAAG  478 Odo.rosm
436  -----ACACA AGAGGAGACA AGTCGTCAAC AAGGTAAGCA TACCGGAAGG  480 Can.latr
439  -----GCACA AGAGGAGACA AGTCGCTAAC AAGGTAAGCG TACCGGAAGG  483 Uro.cine
437  ----AACGCA AGAGGAGACA AGTCG-TAAC AAGGTAAGCA TACTGGAAAG  481 Urs.amer
434  ------TACA AGAGGAGATA AGTCGTCAAC AAGGTAAGCC TACTGGAAAG  477 Myd.marc

     1051       1061
     |          |
944  TGTGCTTGGA ATA                                          956 Rat.norv
940  TGTGCTTGGA ATAAT                                        954 Hyl.stel
940  TGTGCTTGGA                                              949 Mus.croc
942  TGTGCTTGGA ATAAT                                        956 Mus.cook
941  TGTGCTTGGA ATAAT                                        955 Mas.eryt
948  TGTGCTTGGA                                              957 Mus.matt
955  TGTGCTTGGA AAACC                                        969 Mus.plat
944  TGTGCTTGGA                                              953 Mus.setu
955  TGTGCTTGGA AAAAG                                        969 Mus.saxi
939  TGTGCTTGGA                                              948 Cri.gamb
960  TGTGCTTGGA                                              969 Mac.inge
935  TGTGCTTGGA                                              944 Mic.niva
940  TGTGCTTGGA                                              949 Nes.rufu
941  TGTGCTTGGA                                              950 1Per.leu
931  TGTGCTTGGA                                              940 Aco.cahi
942  TGTGCTTGGA                                              951 Cri.migr
940  TGTGCTTGGA                                              949 Gli.glis
947  TGTGCTTGGA                                              956 Ger.nige
942  TGTGCTTGGA ATAAT                                        956 Leo.edwa
944  TGTGCTTGGA                                              953 Mes.aura
947  TGTGCTTGGA                                              956 1Mus.ave
942  TGTGCTTGGA ATAAT                                        956 Mus.musc
```

```
947  TGTGCTTGGA ATAAT                                      961 Mus.paha
938  TGTGCTTGGA                                            947 Tat.kemp
939  TGTGCTTGGA                                            948 Ura.rudd
940  TGTGCTTGGA AA                                         951 1Hyd.hyd
482  TGTGCTTGGA TTAAT                                      496 Bas.astu
480  TGCGCTTGGC TTAAC                                      494 Spi.puto
488  TGTGCTTGGA TTAAT                                      502 Pro.loto
480  TGCGCTTGGA TTAAC                                      494 Mep.meph
479  TGTGCTTGGA TTGAC                                      493 Odo.rosm
481  TGTGCTTGGA TTAAT                                      495 Can.latr
484  TGCGCTTGGA TTCAC                                      498 Uro.cine
482  TGTGCTTGGA TAAAC                                      496 Urs.amer
478  TGTGCTTGGA AAC                                        490 Myd.marc
```

Appendix D

Aligned APRT sequence data

The sequences used for the APRT analysis are shown here. The alignments were done by hand with aid from CLUSTALW. The species listed are *Mus crociduroides* (M.crocid), *Mus caroli* (M.caroli), *Mus cookii* (M.cookii), *Mus spretus* (M.spretu), *Mus saxicola* (M.saxico) (all sequenced for this thesis), *Mus pahari* (Acc. U28721) (M.pahari), *Mus spicilegus* (Acc. U28720) (M.spicil), *Mus musculus domesticus* (Acc. M11310) (M.mus), *Mus musculus domesticus* (Acc. M86439) (M.musA), *Mus musculus domesticus* (Acc. M86440) (M.musB), *Mastomys hildebrandtii* (Acc. U28722) (M.hildeb), *Stochomys longicaudatus* (Acc. U28723) (S.longic), *Rattus norvegicus* (Acc. L04970) (R.norveg), *Gerbillus campestris* (Acc. U28961) (G.campes), and *Mesocricetus auratus* (Acc. X0360) (M.auratu). The sequence presented begins in the middle of the second intron and terminates with the stop codon signalling the end of the APRT coding sequence. Exon 3 begins at position 897 and ends at 1030, intron 3 is located at 1031-1462, exon 4 is located at 1463-1541, intron 4 is located at 1542-1662 and exon 5 is located at 1663-1805. The sequencing primers used are located at 174-188 in intron 2 for primer 3int1 while primer Ix4 is located at 1502-1520 in exon 4.

```
     1          11         21         31         41       50
     |          |          |          |          |        |
  1  AA-GCTTGTG CTAAACAACT GCTGTATACC AGGCTCCATG CTTGAGCTTC    49 M.pahari
  1  ---------- ---------- ---------- ---------- ----------     0 M.crocid
  1  ---------- ---------- ---------- ---------- ----------     0 M.caroli
  1  ---------- ---------- ---------- ---------- ----------     0 M.cookii
  1  ---------- ---------- ---------- ---------- ----------     0 M.spretu
  1  ---------- ---------- ---------- ---------- ----------     0 M.saxico
  1  AA-GCTTGTG TTAAACAACT GCTGTCTACC TGGCTCCATG CCTGAGCTTC    49 M.spicil
  1  AA-GCTTGTG TTAAACAACT GCTGTCTACT TGGCTCCATG CCTGAGCTTG    49    M.mus
  1  AA-GCTTGTG TTAAACAACT GCTGTCTACC TGGCTCCATG CCTGAGCTTC    49   M.musA
  1  AA-GCTTGTG TTAAACAACT GCTGTCTACC TGGCTCCATG CCTGAGCTTC    49   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
  1  ---------- ----ACAACT GCTCTATACC AGGCTCCATG CTTGAGCTTC    36 S.longic
  1  AA-GCTTGTG CTAAACA--- TGTGCACACC AGGCTCTGTG ACTGAGATTC    46 R.norveg
  1  AA-GCT---- ---------- --------CC AGGCTCCATG CGTGAGTTTC    27 G.campes
  1  AATTCTTGTG CTAAATAACT TTCACTTACC AGTG-CCAAG CACGGGCTTC    49 M.auratu

     51         61         71         81         91       100
     |          |          |          |          |        |
 50  AGAAACACCC TAGGGCAGCT GAATGTCCAC CAGGAGTGTC CAGA------    93 M.pahari
  1  ---------- ---------- ---------- ---------- ----------     0 M.crocid
  1  ---------- ---------- ---------- ---------- ----------     0 M.caroli
  1  ---------- ---------- ---------- ---------- ----------     0 M.cookii
  1  ---------- ---------- ---------- ---------- ----------     0 M.spretu
  1  ---------- ---------- ---------- ---------- ----------     0 M.saxico
 50  AGAAACACCC TAGGGCAGCT GAATGTCCAC CAGGAGTGTC CAGA------    93 M.spicil
 50  AGAAACACCC TAGGACAGCT GAATGTCCAC CAGGAGTGTC CAGA------    93    M.mus
 50  AGAAACACCC TAGGACAGCT GAATGTCCAC CAGGAGTGTC CAGA------    93   M.musA
 50  AGAAACACCC TAGGGCAGCT GAATGTCCAC CAGGAGTGTC CAGA------    93   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
 37  AGAAACACCC TAGGGTAGCT AAATGTCCAC CAGGAGTGTC CAGA------    80 S.longic
 47  AGAAACACCC TGGGGTAGCT GAATGTCCAC CAGGAGTGTC CAGA------    90 R.norveg
 28  TGAAACACGC TAGGGTAGCT GAATGTCCAC CAGGGGAGGC CAGA------    71 G.campes
 50  AGAAACACCC TAGGGTCGCT GAATGTCCAC CAGGGGAGTC -AGACATGTC    98 M.auratu

     101        111        121        131        141      150
     |          |          |          |          |        |
 94  GGGAGGGTGA GCACCCCAAG AGAACAGAGT GGCCCTAGTA AATGCTCAGG   143 M.pahari
  1  ---------- ---------- ---------- ---------- ----------     0 M.crocid
  1  ---------- ---------- ---------- ---------- ----------     0 M.caroli
  1  ---------- ---------- ---------- ---------- ----------     0 M.cookii
  1  ---------- ---------- ---------- ---------- ----------     0 M.spretu
  1  ---------- ---------- ---------- ---------- ----------     0 M.saxico
 94  GGGAGGGTGG GCACCCCA-G AGAACAGAGT GGCCCTGGTA AGTGCTCAGG   142 M.spicil
```

```
 94  GGGAGGGTGG GCACCCCA-G AGAACAGAGT GGCCTTGGTA AGTGCTCGGG  142    M.mus
 94  GGGAGGGTGG GCACCCCA-G AGAACAGAGT G-CCCTGGTA AGTGCTCAGG  141   M.musA
 94  GGGAGGGTGG GCACCCCA-G AGAACAGAGT G-CCCTGGTA AGTGCTCAGG  141   M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
 81  GGGAGGGTGA GCACCCCA-G AGAACAGAGT GGCCCTGGCA AGTGCTCAGG  129 S.longic
 91  GGGAGG-TGA ACACCCCA-G AGAACAGAGT GGCCCTCACA AGTGCTCAGG  138 R.norveg
 72  GGGAGGGTGG GCACCCCA-- ------GGGT GGCCCTGGGA AATGCTCAGG  113 G.campes
 99  CAGAGGGTGA GAACCCCA-G AGAATTCGGT AGCCCTGACA TGTGCT----  143 M.auratu

     151        161        171        181        191      200
     |          |          |          |          |        |
144  GACCACAGA- ACTTTTGCCC ACTCCACTTC CTATTGGTAC CC-CCGGCCA  191 M.pahari
  1  ---------- ---------- ---------- ---------- ----------    0 M.crocid
  1  ---------- ---------- ---------- ---------- ----------    0 M.caroli
  1  ---------- ---------- ---------- ---------- ----------    0 M.cookii
  1  ---------- ---------- ---------- ---------- ----------    0 M.spretu
  1  ---------- ---------- ---------- ---------- ----------    0 M.saxico
143  GACCACAG-- ACTTTTGCCT ACTTCACTTC CTATTGGTAC CC-CTGGCCA  189 M.spicil
143  GACCACAG-- ACTTT-GCC- ACTTCACTTC CTATTGGTAC CC-TTGGCCA  187    M.mus
142  GACCACAG-- ACTTTTGCC- ACTTCACTTC CTATTGGTAC CC-CTGGCCA  187   M.musA
142  GACCACAG-- ACTTTTGCCT ACTTCACTTC CTATTGGTAC CC-CTGGCCA  188   M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
130  GGCCACAGT- ACTTTTGTCC ACTTCCCTTC CTGTTGGTAC CCCCTGGTAG  178 S.longic
139  GACCACAGT- CCTTTTGCCC ACTTCACTTC CTATTGGTAC CCCCTGACCA  187 R.norveg
114  GGCCAGAGT- ACTCGTGCCC ACTTGACTTC CTGTTGGAAC CCCCTGGCCA  162 G.campes
144  ----ACAATT ACTGATGCCC ACTT-----C CTACTGGTTC CTCCTGGCCA  184 M.auratu

     201        211        221        231        241      250
     |          |          |          |          |        |
192  TGCCCCAGAA ATCAGGGCAT GTTTGTACCC TCCCCACGAC AGCTCGGGCC  241 M.pahari
  1  ------AGAA ATCAGGGCAT GTTTGTACCC TCCCCACGAC AGCTCGGGCC   44 M.crocid
  1  ---------- ---------- ---------- --CCCACGAC AGCTCAGGCC   18 M.caroli
  1  --------AA ATTAGGGCAT GTATGTATCC TTCCCACGAC AGCTCAGGCC   42 M.cookii
  1  ---------- ---------- ---------- --CCCACGAC AGCTCAGGCC   18 M.spretu
  1  ---------- ---------- --------CC TTCCCACGAC AGCTCAGGCC   22 M.saxico
190  TGCTCCAGAA ATTAGGGCAT GTATGTATCC TTCCCACGAC AGCT------  233 M.spicil
188  TGCTCCAGAA ATTAGGGCAT GTATGTATCC TTCCCACGAC AGCT------  231    M.mus
188  TGCTCCAGAA ATTAGGGCAT GTATGTATCC TTCCCACGAC AGCT------  231   M.musA
189  TGCTCCAGAA ATTAGGGCAT GTATGTATCC TTCCCACGAC AGCT------  232   M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
179  TGCTCCTGAA ATGAGGGCAT GTGTGTATCC TCTC------ --CTCAGGCC  220 S.longic
188  TGCTGTAGAA ATTAGGG--- ---------- ---------- --TCCAGGCT  212 R.norveg
163  TGCTCCAGAA ATGAGGGTAT GTATGCATCT TT-CCACTTT AGCACAGGT-  210 G.campes
185  TACCTCAGGA ATTAGGGCAT GCTTTCTGCC TG-CTACAGT AGCTCATCCT  233 M.auratu
```

```
     251        261        271        281        291       300
     |          |          |          |          |         |
242  GCCTGGAACT GACCTGTAGA CAGTGCTCCT GGGT------ ----------  275 M.pahari
 45  GCCTGGAACT GACCTGTAGA CAGTGCTCCT GAGT------ ----------   78 M.crocid
 19  TCCTGGAACT GGGT------ ---------- ---------- ----------   32 M.caroli
 43  TCCTGGAACT GGAT------ ---------- ---------- ----------   56 M.cookii
 19  TCCTGRAACT GGGT------ ---------- ---------- ----------   32 M.spretu
 23  TCCTGGAACT GGGT------ ---------- ---------- ----------   36 M.saxico
234  ---------- ---------- ---------- ---------- ----------  233 M.spicil
232  ---------- ---------- ---------- ---------- ----------  231    M.mus
232  ---------- ---------- ---------- ---------- ----------  231   M.musA
233  ---------- ---------- ---------- ---------- ----------  232   M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
221  TCCTGGAAAT GACC------ ---------- ---------- ----------  234 S.longic
213  TCCTGGAAAT GGCC------ ---------- ---------- ----------  226 R.norveg
211  --CTGGAAGT GACC------ ---------- ---------- TGGAGTCAAA  232 G.campes
234  CCCTGGAAGT GACCCCAGAC ATATACCCTG AACTGTAACC ----------  273 M.auratu

     301        311        321        331        341       350
     |          |          |          |          |         |
276  ---------- ---------- ---------- ----AGATGC TGCATTTGAA  291 M.pahari
 79  ---------- ---------- ---------- ----AGATGC TGCATTCGAA   94 M.crocid
 33  ---------- ---------- ---------- ----AGATGC TGCATTTGAA   48 M.caroli
 57  ---------- ---------- ---------- ----AGATGC TGCATTTGAA   72 M.cookii
 33  ---------- ---------- ---------- ----AGATGT TGCATTTGAA   48 M.spretu
 37  ---------- ---------- ---------- ----AGATGC TGCATTTGAA   52 M.saxico
234  ---------- ---------- ---------- ----AGATGC TGCATTTGAA  249 M.spicil
232  ---------- ---------- ---------- ----AGATGC TGCATTTGAA  247    M.mus
232  ---------- ---------- ---------- ----AGATGC TGCATTTGAA  247   M.musA
233  ---------- ---------- ---------- ----AGATGC TGCATTTGAA  248   M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
235  ---------- TGT-AGATAA AGTGCTCCTG GG-CAGATGC TGCATTTGAA  272 S.longic
227  ---------- TGTTAGAGAA AGTGCTCATG GG-CAGATGC TGCATTTGAC  265 R.norveg
233  GGAACCCAAG -----GTGAA ACTACACCAG GGGC------ ----CTTGA-  266 G.campes
274  ---------- -----GATAA AGTGCGCCTG GG-CAGATGT ---ATTTGAG  304 M.auratu

     351        361        371        381        391       400
     |          |          |          |          |         |
292  AGGTGGCAAG AGGGCTGGTG AGATGGCTCA GCGGTTAGGA GCACTGACTG  341 M.pahari
 95  AGGTGGCAAG AGGGCTGGTG AGATGGCTCA GCGGTTAGGA ACACTGACTG  144 M.crocid
 49  AGGTGGCAAG AG-------- ---------- ---------- ----------   60 M.caroli
 73  AGGTGGCAAG AG-------- ---------- ---------- ----------   84 M.cookii
 49  AGGTGGCAAG AG-------- ---------- ---------- ----------   60 M.spretu
 53  AGGTGGCAAG AGCCGGGTGT GGTGGCGCAC GCCTGTAATC CCAGCACTCG  102 M.saxico
250  AGGTGGCAAG AG-------- ---------- ---------- ----------  261 M.spicil
```

```
248  -GGTGGCAAG A--------- ---------- ---------- ----------   257    M.mus
248  A-GTGGCAAG A--------- ---------- ---------- ----------   257   M.musA
249  AGGTGGCAAG AG-------- ---------- ---------- ----------   260   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
273  AGGTGGC--- ---------- ---------- ---------- ----------   279 S.longic
266  AGATGGCAAG AG-------- ---------- ---------- ----------   277 R.norveg
267  -----GCAAG A--------- ---------- ---------- ----------   272 G.campes
305  AGGTGGCAAA AG-------- ---------- ---------- ----------   316 M.auratu

     401        411        421        431        441      450
     |          |          |          |          |        |
342  CTCTTCCAAA GGTCCTGAGT TCAAATCCCA GCAACCACAT GGTGGCTCAC   391 M.pahari
145  CTCTTCCAAG GGTCCTGAGT TCAAATCCCA GCAACCACAT GGTGGCTCAC   194 M.crocid
 61  ---------- ---------- ---------- ---------- ----------    60 M.caroli
 85  ---------- ---------- ---------- ---------- ----------    84 M.cookii
 61  ---------- ---------- ---------- ---------- ----------    60 M.spretu
103  GGAGGCAGAG GCAGGCGGAT TTCTGAGTTC AAGGCCAGCC TGGTCTACAG   152 M.saxico
262  ---------- ---------- ---------- ---------- ----------   261 M.spicil
258  ---------- ---------- ---------- ---------- ----------   257    M.mus
258  ---------- ---------- ---------- ---------- ----------   257   M.musA
261  ---------- ---------- ---------- ---------- ----------   260   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
280  ---------- ---------- ---------- ---------- ----------   279 S.longic
278  ---------- ---------- ---------- ---------- ----------   277 R.norveg
273  ---------- ---------- ---------- ---------- ----------   272 G.campes
317  ---------- ---------- ---------- ---------- ----------   316 M.auratu

     451        461        471        481        491      500
     |          |          |          |          |        |
392  AACCACCTAC AGCTACAGTG TACACACATA TAATAAAATA AATAAACAAA   441 M.pahari
195  AACCATCTAC AGCTACAGTG TACACACATA TAATAAAATA AATAAATAAA   244 M.crocid
 61  ---------- ---------- ---------- ---------- ----------    60 M.caroli
 85  ---------- ---------- ---------- ---------- ----------    84 M.cookii
 61  ---------- ---------- ---------- ---------- ----------    60 M.spretu
153  AGTGAGTTCC AGGACAGCCA GGGCTACACA GAGAAACCCT GTCTCGAAAA   202 M.saxico
262  ---------- ---------- ---------- ---------- ----------   261 M.spicil
258  ---------- ---------- ---------- ---------- ----------   257    M.mus
258  ---------- ---------- ---------- ---------- ----------   257   M.musA
261  ---------- ---------- ---------- ---------- ----------   260   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
280  ---------- ---------- ---------- ---------- ----------   279 S.longic
278  ---------- ---------- ---------- ---------- ----------   277 R.norveg
273  ---------- ---------- ---------- ---------- ----------   272 G.campes
317  ---------- ---------- ---------- ---------- ----------   316 M.auratu
```

```
      501        511        521        531        541       550
      |          |          |          |          |        |
 442  TCTTAAAAAA AAAAAAAAGA AAGAAAGGTG GCAAGAG--- ----------   478 M.pahari
 245  TCTTTAAAAA TAAA----GA AAGAAAGGTG GCAAGAG--- ----------   277 M.crocid
  61  ---------- ---------- ---------- ---------- ----------    60 M.caroli
  85  ---------- ---------- ---------- ---------- ----------    84 M.cookii
  61  ---------- ---------- ---------- ---------- ----------    60 M.spretu
 203  AACCAAAAAA AAAAAAAAAA AAAAAA---- ---------- ----------   229 M.saxico
 262  ---------- ---------- ---------- ---------- ----------   261 M.spicil
 258  ---------- ---------- ---------- ---------- ----------   257    M.mus
 258  ---------- ---------- ---------- ---------- ----------   257   M.musA
 261  ---------- ---------- ---------- ---------- ----------   260   M.musB
   1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
 280  ---------- ---------- ---------- ---------- ----------   279 S.longic
 278  ---------- ---------- ---------- ---------- ----------   277 R.norveg
 273  ---------- ---------- ---------- -----CCTGC GCCGGCCCTC   287 G.campes
 317  ---------- ---------- ---------- ---------- ----------   316 M.auratu

      551        561        571        581        591       600
      |          |          |          |          |        |
 479  ---------- ---------- ---------- ---------- ----------   478 M.pahari
 278  ---------- ---------- ---------- ---------- ----------   277 M.crocid
  61  ---------- ---------- ---------- ---------- ----------    60 M.caroli
  85  ---------- ---------- ---------- ---------- ----------    84 M.cookii
  61  ---------- ---------- ---------- ---------- ----------    60 M.spretu
   1  ---------- ---------- ---------- ---------- ----------     0 M.saxico
 262  ---------- ---------- ---------- ---------- ----------   261 M.spicil
 258  ---------- ---------- ---------- ---------- ----------   257    M.mus
 258  ---------- ---------- ---------- ---------- ----------   257   M.musA
 261  ---------- ---------- ---------- ---------- ----------   260   M.musB
   1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
 280  ---------- ---------- ---------- -----CCTGA GCTGTACA--   292 S.longic
 278  ---------- ---------- ---------- ---------- ----------   277 R.norveg
 288  CTGGTCCAGA AGGCTGGCCC CGCAGCCTGG GATGC----- ----------   322 G.campes
 317  ---------- ---------- ---------- ---------- --------TA   318 M.auratu

      601        611        621        631        641       650
      |          |          |          |          |        |
 479  --CCACCATA G-TGGA---- ---------- ---------- ----------   491 M.pahari
 278  --CCACC--- ---------- ---------- ---------- ----------   282 M.crocid
  61  --CCACCATA GGTTGCCCTG -AGCTG---- ---------- ----------    83 M.caroli
  85  --CCACCATA GGTGGCCCTG -AGCTG---- ---------- ----------   107 M.cookii
  61  --CCACCATA GGTGGCCCTG -AGCTG---- ---------- ----------    83 M.spretu
   1  ---------- ---------- ---------- ---------- ----------     0 M.saxico
 262  --CCACCATA GGTGGCCCTG -AGCTG---- ---------- ----------   284 M.spicil
```

```
258  --CCACCATA GGTGGCCCTG -AGCTG---- ---------- ----------  280   M.mus
258  --CCACCATA GGTGGCCCTG -AGCTG---- ---------- ----------  280  M.musA
261  --CCA--ATA G--GG--CTG -AGCTG---- ---------- ----------  277  M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
293  ---------- ---------- ---------- ---------- ----------  292 S.longic
278  --CCACCATA GGTGACCCTG -AGCTG---- ---------- ----------  300 R.norveg
323  ---------- ----ACCCTG TAGCTGAAAT GCTCCGGGTA GATGCTGCAT  358 G.campes
319  AA---CCATA GGTGTCCCCG -AGCT----- ---------- ----------  339 M.auratu

     651       661       671       681       691      700
     |         |         |         |         |        |
492  ---------- ---------- ---------- ---------- ----------  491 M.pahari
283  ---------- ---------- ---------- ---------- ----------  282 M.crocid
 84  ---------- ---------- ---------- ---------- -------TAC   86 M.caroli
108  ---------- ---------- ---------- ---------- -------TAC  110 M.cookii
 84  ---------- ---------- ---------- ---------- -------TTC   86 M.spretu
  1  ---------- ---------- ---------- ---------- ----------    0 M.saxico
285  ---------- ---------- ---------- ---------- -------TTC  287 M.spicil
281  ---------- ---------- ---------- ---------- -------TTC  283   M.mus
281  ---------- ---------- ---------- ---------- -------TTC  283  M.musA
278  ---------- ---------- ---------- ---------- -------TTC  280  M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
293  ---------- ---------- ---------- ---------- ----------  292 S.longic
301  ---------- ---------- ---------- ---------- -------TAA  303 R.norveg
359  TTGAGCGTGG CCAGAGCCAT CTCTATTATT GTCCTCTCTA CCTC------  402 G.campes
340  ---------- ---------- ---------- ---------- ----AGATAC  345 M.auratu

     701       711       721       731       741      750
     |         |         |         |         |        |
492  -GAAGGCAGG TAGGAT-CCC CAAGGCTAAG AT-------- ----GCTACC  527 M.pahari
283  ---------- ---------- ------TAAG AT-------- ----GCTAAC  294 M.crocid
 87  AGAAGGCAGG TAGGATTCCC CAAGGCTAAG AT-------- ---GGCTACC  125 M.caroli
111  AGAAGGCAGG TAGGAT-CCC CAAGGCTAAG AT-------- ---GGCTACC  148 M.cookii
 87  AGAAGGGAGG TAGGAT-CCC CAAGGCTGAG AT-------- ---GGCCACC  124 M.spretu
  1  ---------- ---------- ---------- ---------- ----------    0 M.saxico
288  AGAAGGCAGG TAGGAT-CCC CAAGGCTGAG AT-------- ---GGCCACC  325 M.spicil
284  AGAAGGCAGG TAGGAT-CCC CAAGGCTGAG ATGATGAGTT GATGGCTACC  332   M.mus
284  AGAAGGCAGG TAGGAT-CCC CAAGGCTGAG ATGATGAGTT GATGGCTACC  332  M.musA
281  AGAAGGCAGG TAGGAT-CCC CAAGGCTGAG ATGATGAGTT GATGGCTACC  329  M.musB
  1  ---------- ---------- ---------- ---------- ----------    0 M.hildeb
293  -GAAGGCAGG GAAGAT-CCC CAAGGCTGAG AT-------- ----GCTACC  328 S.longic
304  AGAAGGCAGG TAAGAT-CCC CACGGCTAAG AT-------- ----ACTACC  340 R.norveg
403  ---------- ---------- ---------- ---------- ----------  402 G.campes
346  AGAAGGCAGA TAACAT-CCC CAAGGCTAAG CT-------- ----GCTGCC  382 M.auratu
```

```
     751        761        771        781        791      800
     |          |          |          |          |        |
528  G-AGTAACCA TCAGTGTTCT TCTAGCCA-T AGTGGGCAAG ACCTAGTGTT   575 M.pahari
295  C-AGTAGCCA TCAGTGTTCT TCTAGCCA-T AGTCAGCAAG ACCTAGTGTT   342 M.crocid
126  C-AGTAGCCA TCAACGTTCT TCTAGCCG-T AGTCAGCAAG ACCTAGTGTT   173 M.caroli
149  C-AG------ ---------- ---------- ---------- ----------   151 M.cookii
125  C-AGTAGCCA TCAACGTTCT TCTAGCCG-T AGTCAGCAAG ACCTAGTGTT   172 M.spretu
  1  ---------- ---------- ---------- ---------- ----------     0 M.saxico
326  C-AGTAGCCA TCAACGTTCT TCTAGCTGGT AGTCAGCAAG ACCTAGTGTT   374 M.spicil
333  C-AGTAGCCA TCAACGTTCT TCTAACCG-T AGTCAGCAAG ACCTAGTGTT   380    M.mus
333  C-AGTAGCCA TCAACGTTCT TCTAGCCG-T AGTCAGCAAG ACCTAGTGTT   380   M.musA
330  C-AGTAGCCA TCAACGTTCT TCTAGCCG-T AGTCAGCAAG ACCTAGTGTT   377   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
329  C-AGTAGCCA TCTGTGTTCT ---ACCTA-T AGTCAGCAAG ACTTAGTGTT   373 S.longic
341  C-AGTAGCCA TCAACATTCT ---AGCTA-T AGTCAGCAAG GCCTAGTGTT   385 R.norveg
403  ---------- ---------- ---------- ---------- ----------   402 G.campes
383  CCAATAGCCA TCAGC---CT TCTAGTTA-T AGCTAGTAAG ACCTAGTATT   428 M.auratu

     801        811        821        831        841      850
     |          |          |          |          |        |
576  CCTAGTCAAT GTTGACCTCT CC-ATACTT- ---------- -------GCC   606 M.pahari
343  CCTAGTCAAT GTTGACCTC- TCCATACTTG ACCTCT---- ----------   377 M.crocid
174  CCTAGCCAGT GTTGACCTCG CTCATACTTG GCCTCTAGAT TCCCAT-GCC   222 M.caroli
152  ---------- ---------- ---------- ---------T GCCCATTGCC   162 M.cookii
173  CCTAGCAAGT GTTGACCTCG CCCATACTTG GCCTCTAGAT TCCCAT-GCC   221 M.spretu
  1  ---------- ---------- ---------- ---------- -CCCA-CGCC     8 M.saxico
375  CCTAGCAAGT GTTGACCTCG CCCATACTTG GCCTTTAGAT TCCCAT-GCC   423 M.spicil
381  CCTAGCAAGT GTTGACCTCG CCCATACTTG GCCTCTAGAT TCCCAT-GCC   429    M.mus
381  CCTAGCAAGT GTTGACCTCG CCCATACTTG GCCTCTAGAT TCCCAT-GCC   429   M.musA
378  CCTAGCAAGT GTTGACCTCG CCCATACTTG GCCTCTAGAT TCCCAT-GCC   426   M.musB
  1  ---------- ---------- ---------- ---------- ----------     0 M.hildeb
374  CCCAGCCTA- -TTGACCTCT CC-ATCCTTG GCCTGTAGAT TCCCAT-GCC   419 S.longic
386  CCTAGCCAAT GTTGACCTCT CC-ATCCCTG GCCTGGAGAT CCCTAT-GCC   433 R.norveg
403  ---------- ---------- --------TG GCCTGTAGG- --CTAT-CAC   420 G.campes
429  CCTGGTCAAT ACTATTCACT CA-ATCCTT- ---------- ----A--CAC   460 M.auratu

     851        861        871        881        891      900
     |          |          |          |          |        |
607  TCTCG----G CTCC-ATCCC ACACCCTTCC CTCCTTACCC TAACAGGTCT   651 M.pahari
378  ---CG----G CTCCC-TCCC ACACCCTTCC CTCCTTACCC TAACAGGTCT   419 M.crocid
223  CCTCG----G CTCC-ATCCC ACACCCTTCC CTCCTTACCC TAACAGGTCT   267 M.caroli
163  CCTCG----G CTCC-ATCCC ACACCCTTCC CTCCTTACCC TAACAGGTCT   207 M.cookii
222  CCTCCCTCAG CTCC-ATCCC ACACCCTTCC CTCCTTACCC TAACAGGTCT   270 M.spretu
  9  CTTCG----G CTCC-ATCCC ACACCCTTCC CTCCTTACCC TCACAGGTCT    53 M.saxico
424  CCTCA----G CTCC-ATCCC ACAACCTTCC CTCCTTACCC TAACAGGTCT   468 M.spicil
```

```
430   CCTCA----G CTCC-ATCCC ACAACCTTCC CTCCTTACCC TAACAGGTCT   474     M.mus
430   CCTCA----G CTCC-ATCCC ACAACCTTCC CTCCTTACCC TAACAGGTCT   474    M.musA
427   CCTCA----G CTCC-ATCCC ACAACCTTCC CTCCTTACCC TAACAGGTCT   471    M.musB
  1   ---------- ---------- ---------- ---------- ------GTCT     4 M.hildeb
420   CCTCG----G CTCC-ATCTA ATACCTTTCC CTCCTTACCC AAACAGGTCT   464 S.longic
434   TCTGG----G CTCCCCTCCA ACACCCTTCT CTCCTTACCC TAACAGGTCT   479 R.norveg
421   CC--A----G CTCC-ACCCA ACACCTCTCC TCCCTTACCC TGACAGGCCT   463 G.campes
461   C-TCA----G C-----CCTA ACACGCCCCC TCTCTCATCC TAACAGGCCT   500 M.auratu

      901        911        921        931        941       950
      |          |          |          |          |         |
652   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   701 M.pahari
420   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTTGGCG   469 M.crocid
268   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   317 M.caroli
208   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGATCAG GAGCTGGGCG   257 M.cookii
271   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   320 M.spretu
 54   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   103 M.saxico
469   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   518 M.spicil
475   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   524     M.mus
475   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   524    M.musA
472   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   521    M.musB
  5   AGACTCCAGG GGTTTCCTGT TTGGCCCTTC ACTAGCTCAG GAGCTGGGCG    54 M.hildeb
465   AGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTAGCTCAG GAGCTGGGCG   514 S.longic
480   GGACTCCAGG GGCTTCCTGT TTGGCCCTTC CCTTGCTCAG GAGCTGGGGG   529 R.norveg
464   GGACTCCAGG GGCTTCCTGT TTGGCCCCTC CCTAGCTCAG GAGCTGGGCC   513 G.campes
501   AGACTCCAGG GGATTCTTGT TTGGCCCCTC CCTAGCTCAG GAGCTGGGCC   550 M.auratu

      951        961        971        981        991       1000
      |          |          |          |          |         |
702   TGGGCTGCGT GCTCATCCGG AAGCAGGGGA AGCTGCCGGG CCCCACTATA   751 M.pahari
470   TGGGCTGCGT GCTCATCCGG AAGCAGGGGA AGCTGCCGGG CCCCACTATA   519 M.crocid
318   TGGGCTGTGT GCTCATCCGG AAACAGGGGA AGCTGCCGGG CCCCACTGTG   367 M.caroli
258   TGGGCTGTGT GCTCATCCGG AAACAGGGGA AGCTGCCGGG CCCCACTGTG   307 M.cookii
321   TGGGCTGTGT GCTCATCCGA AAACAGGGGA AGCTGCCGGG CCCCACTGTG   370 M.spretu
104   TGGGCTGTGT GCTCATCCGG AAGCAGGGGA AGCTGCCGGG CCCCACTGTG   153 M.saxico
519   TGGGCTGTGT GCTCATCCGG AAACAGGGGA AGCTGCCGGG CCCCACTGTG   568 M.spicil
525   TGGGCTGTGT GCTCATCCGG AAACAGGGGA AGCTGCCGGG CCCCACTGTG   574     M.mus
525   TGGGCTGTGT GCTCATCCGG AAACAGGGGA AGCTGCCGGG CCCCACTGTG   574    M.musA
522   TGGGCTGTGT GCTCATCCGG AAACAGGGGA AGCTGCCGGG CCCCACTGTG   571    M.musB
 55   TGGGCTGCGT GCTCATCCGG AAACGAGGGA AGCTTCCCGG CCCCACTCTG   104 M.hildeb
515   TGGGCTGTGT GCTCATCCGG AAGCGCGGGA AGCTGCCAGG CCCCACTTTG   564 S.longic
530   TGGGCTGTGT GCTCATCCGT AAGCGCGGGA AGCTACCAGG CCCCACTGTA   579 R.norveg
514   TGGGCTGTGT GCTCATCCGA AAGCGTGGGA AGCTGCCGGG TCCCACTGTG   563 G.campes
551   TGGGCTGTGT GCTCATCCGG AAGCGAGGGA AGCTGCCAGG CCCCACAGTG   600 M.auratu
```

```
      1001       1011       1021       1031       1041     1050
      |          |          |          |          |        |
752   TCAGCCTCCT ATGCTCTGGA GTATGGGAAG GTAAGGGAGC TGTGGGTAGA    801 M.pahari
520   TCAGCCTCCT ATGCTCTGGA GTATGGGAAG GTAAGGRAGC TGTGGGTAGA    569 M.crocid
368   TCAGCCTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTG---A    414 M.caroli
308   TCAGCCTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTGTAGA    357 M.cookii
371   TCARCGTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTGTAGA    420 M.spretu
154   TCAGCCTCCT ATGCTCTGGA GTATGGGAAG GTAAGCGAGC TCTGGGGAGA    203 M.saxico
569   TCAGCCTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTGTAGA    618 M.spicil
575   TCAGCCTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTGTAGA    624    M.mus
575   TCAGCCTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTGTAGA    624   M.musA
572   TCAGCCTCCT ATTCTCTGGA GTATGGGAAG GTAAGCGAGC TGTGTGTAGA    621   M.musB
105   TCAGCCTCCT ATGCCCTGGA GTATGGGAAG GTAAGCGAGC TGGGGGTAGA    154 M.hildeb
565   TCAGCCTCCT ATGCTCTGGA GTATGGAAAG GTAAGTGAGC TGTGGGTGGA    614 S.longic
580   TCAGCCTCCT ATTCTCTAGA GTATGGGAAG GTAAGCAAGC CGTGGGTGGA    629 R.norveg
564   TCAGCCTCCT ACGCCCTAGA GTACGGGAAG GTAAGCGAGC TTTGGGTAGA    613 G.campes
601   TCAGCCTCCT ATGCTCTCGA GTATGGCAAG GTAAGCAGGC AGTGGGTAGC    650 M.auratu

      1051       1061       1071       1081       1091     1100
      |          |          |          |          |        |
802   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    819 M.pahari
570   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    587 M.crocid
415   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    432 M.caroli
358   GGAAGAT-CA GGGTCTTAT- ---------- ---------- ----------    375 M.cookii
421   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    438 M.spretu
204   GGAAGGA-CA GGGTCTTAC- ---------- ---------- ----------    221 M.saxico
619   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    636 M.spicil
625   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    642    M.mus
625   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    642   M.musA
622   GGAAGGG-CA GGGTCTTAT- ---------- ---------- ----------    639   M.musB
155   GGAAGGG-CA GGGTCTT--T CGTCTTTGTT GTGGTGGTGG TTTTTCTTTC    201 M.hildeb
615   GCAAGGGGCA GGGTCTA--- ---------- ---------- ----------    631 S.longic
630   GGAAGGG-CA GGGTCAA--- ---------- ---------- ----------    645 R.norveg
614   GGAAGCA-CA GGGTCAG--- ---------- ---------- ----------    629 G.campes
651   ---------- ---------- ---------- ---------- ----------    650 M.auratu

      1101       1111       1121       1131       1141     1150
      |          |          |          |          |        |
820   ---------- ---------- ---------- ---------- ----------    819 M.pahari
588   ---------- ---------- ---------- ---------- ----------    587 M.crocid
433   ---------- ---------- ---------- ---------- ----------    432 M.caroli
376   ---------- ---------- ---------- ---------- ----------    375 M.cookii
439   ---------- ---------- ---------- ---------- ----------    438 M.spretu
222   ---------- ---------- ---------- ---------- ----------    221 M.saxico
637   ---------- ---------- ---------- ---------- ----------    636 M.spicil
```

```
643  ---------- ---------- ---------- ---------- ----------  642    M.mus
643  ---------- ---------- ---------- ---------- ----------  642   M.musA
640  ---------- ---------- ---------- ---------- ----------  639   M.musB
202  TTTGGGGTTT TTTTGTTGTT TTTCGAGACA GGGTTTCTCT GTGTAGCCCT  251 M.hildeb
632  ---------- ---------- ---------- ---------- ----------  631 S.longic
646  ---------- ---------- ---------- ---------- ----------  645 R.norveg
630  ---------- ---------- ---------- ---------- ----------  629 G.campes
651  ---------- ---------- ---------- ---------- ----------  650 M.auratu

     1151       1161       1171       1181       1191      1200
     |          |          |          |          |         |
820  ---------- ---------- ---------- ---------- ----------  819 M.pahari
588  ---------- ---------- ---------- ---------- ----------  587 M.crocid
433  ---------- ---------- ---------- ---------- ----------  432 M.caroli
376  ---------- ---------- ---------- ---------- ----------  375 M.cookii
439  ---------- ---------- ---------- ---------- ----------  438 M.spretu
222  ---------- ---------- ---------- ---------- ----------  221 M.saxico
637  ---------- ---------- ---------- ---------- ----------  636 M.spicil
643  ---------- ---------- ---------- ---------- ----------  642    M.mus
643  ---------- ---------- ---------- ---------- ----------  642   M.musA
640  ---------- ---------- ---------- ---------- ----------  639   M.musB
252  GGCTGTCCTG TAGACCAGGC TGGTCTCGAA CTCAGAAATC TGCCTGCCTC  301 M.hildeb
632  ---------- ---------- ---------- ---------- ----------  631 S.longic
646  ---------- ---------- ---------- ---------- ----------  645 R.norveg
630  ---------- ---------- ---------- ---------- ----------  629 G.campes
651  ---------- ---------- ---------- ---------- ----------  650 M.auratu

     1201       1211       1221       1231       1241      1250
     |          |          |          |          |         |
820  ---------- ---------- ---------- ---------- ----------  819 M.pahari
588  ---------- ---------- ---------- ---------- ----------  587 M.crocid
433  ---------- ---------- ---------- ---------- ----------  432 M.caroli
376  ---------- ---------- ---------- ---------- ----------  375 M.cookii
439  ---------- ---------- ---------- ---------- ----------  438 M.spretu
222  ---------- ---------- ---------- ---------- ----------  221 M.saxico
637  ---------- ---------- ---------- ---------- ----------  636 M.spicil
643  ---------- ---------- ---------- ---------- ----------  642    M.mus
643  ---------- ---------- ---------- ---------- ----------  642   M.musA
640  ---------- ---------- ---------- ---------- ----------  639   M.musB
302  AGCCTCCCAA GTGCTGGGAT TAAAGGTGCG TGCCACCACT GCCTGGCCAA  351 M.hildeb
632  ---------- ---------- ---------- ---------- ----------  631 S.longic
646  ---------- ---------- ---------- ---------- ----------  645 R.norveg
630  ---------- ---------- ---------- ---------- ----------  629 G.campes
651  ---------- ---------- ---------- ---------- ----------  650 M.auratu
```

```
     1251       1261       1271       1281       1291       1300
     |          |          |          |          |         |
820  ---------- ---TACCACG GCTACCAGTG CCTAGGAGTA AATGTGGGTG   856 M.pahari
588  ---------- ---TACCACG GCTACCAGTG CCTAGGAGTA AATGTGGGTG   624 M.crocid
433  ---------- ------CATG GCTACCAGTG TCTAGGAGTA AATGTGGGTG   466 M.caroli
376  ---------- ------CACG GCTACCAGTG TCTTGGAGTA AATGTGGGTG   409 M.cookii
439  ---------- ------CACG GCTACCAGTG TCTAGGAGTA AATGTGGGTG   472 M.spretu
222  ---------- ------CACG GCTACCAGTG TCTAGGAGTA AATGTGGGTG   255 M.saxico
637  ---------- ------CACG GCTACCAGTG TCTAGGAGTA AATGTGGGTG   670 M.spicil
643  ---------- ------CACG GCTACCAGTG TCTAGGAGTA AATGTGGGTG   676    M.mus
643  ---------- ------CACG GCTACCAGTG TCTAGGAGTA AATGTGGGTG   676   M.musA
640  ---------- ------CACA GCTACCAGTG TCTAGGAGTA AATGTGGGTG   673   M.musB
352  AGGGCAGGGT CATAACCACT GCTAGCAGTG TCTAGGAGTA AATGTAGGTG   401 M.hildeb
632  ---------- ----ACCACT GCCACCCGCG TCTAGGAGTA AATGTGGGTG   667 S.longic
646  ---------- ----ACCACT GCTGCCAGTG TCTAGGAGTA AAAGTGGGTG   681 R.norveg
630  ---------- ----ACCACT GCCTCGGATG TTTAGGAGTA AATGTTGGGG   665 G.campes
651  ---------- ---------- --------TG TCTAGGAGTA AATGTGGGGG   672 M.auratu

     1301       1311       1321       1331       1341       1350
     |          |          |          |          |         |
857  CTCAGAGAG- ---GTTGAGA CATTGGGGTG AGGTTTACAA CTCCTGAAAT   902 M.pahari
625  CTCAGAGAG- ---GTTGAGA CATCGGG-TC GGGTTTACAA CTCCTGAAAT   669 M.crocid
467  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTTACAC CACC------   505 M.caroli
410  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTTACAC CACC------   448 M.cookii
473  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTT---- ----------   503 M.spretu
256  CTCAGAGAG- ---GTCGAGG CATCGGG-TC GGGTTTACAA CACC------   294 M.saxico
671  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTTACAC CACC------   709 M.spicil
677  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTTACAC CACC------   715    M.mus
677  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTTACAC CACC------   715   M.musA
674  CTCAGAGAG- ---GTTGAGA CATTGGG-TC AGGTTTACAC CACC------   712   M.musB
402  CTCAGAGAGA GAGGTTGAGA CATCTGG--- ---------- --CC------   430 M.hildeb
668  CTCAGAGAG- ---GTTGAGA CATCGGGGGC AGGTTTACAC CACC------   707 S.longic
682  CCCGGAGAG- ---GTTGGGA CATCGGG-CC AGGTTTACAC CACC------   720 R.norveg
666  CTCAGAGAG- ---GTTGAGA CACCGGG-CC AGGCTTGCAC CACC------   704 G.campes
673  CTCAGAGAG- ---GTTAAGT CATCAGG-CC AGGTTTATAC CACC------   711 M.auratu

     1351       1361       1371       1381       1391       1400
     |          |          |          |          |         |
903  GCTCAGCCTC AGAAATGCT- CCAGGCTAGG GAG------- ----------   934 M.pahari
670  GCTCAGCCTC AGAAATGCT- CCAGGCTAGG GAG------- ----------   701 M.crocid
506  ---------C AGAAACGCT- CGAGGCTAGG GAG------- ----------   528 M.caroli
449  ---------C AGAAATGCT- CGAGGCTAGG GAG------- ----------   471 M.cookii
504  ---------- --AAACGCT- CGAGGCTACG GAG------- ----------   523 M.spretu
295  ---------- --AAA----- -GAGGCTAGG GAG------- ----------   309 M.saxico
710  ---------C AGAAACGCT- CGAGGCTAGG GAG------- ----------   732 M.spicil
```

```
716  ---------C AGAAACGCT- CGAGCCTAGG GAG------- ----------   738    M.mus
716  ---------C AGAAACGCT- CGAGGCTAGG GAG------- ----------   738   M.musA
713  ---------C AGAAACGCT- CGAGGCTAGG GAG------- ----------   735   M.musB
431  ---------A GGAAACGCT- TGAGGCTAGG GAG------- ----------   453 M.hildeb
708  ---------A GGAAATGCT- CGAGGCTAAG GAG------- ----------   730 S.longic
721  ---------A GGAAACACT- AGAGGCTAGG GAG------- ----------   743 R.norveg
705  ---------A GGAAACACGG GGAGGCCAGC AGTGGAAAGC AGGCCAGGCA   745 G.campes
712  ---------A GGAAACATGG AGAAGCTAGG GGTG------ ----------   736 M.auratu

     1401       1411       1421       1431       1441     1450
     |          |          |          |          |        |
935  --------GT GGCCACTTGT TAGCATCTAG ACTCTC--TT AACGC-TACT   973 M.pahari
702  --------GT GACC                                          707 M.crocid
529  --------GT GGCCACTTGT TCCCACCT-- -CCCCC--TT A-CAC-TACT   563 M.caroli
472  --------GT GGCCACTTGT TCGCACCTAG ACTCTC--TT A-CAC-TGCT   509 M.cookii
524  --------GT GGCCACTTGT TTGCACCTAG ACTCTGTCTT A-CAC-TACT   563 M.spretu
310  --------GT GGCCACTTGT TC                                 323 M.saxico
733  --------GT GGCCACTTGT TCGCATCTAG ACTCTCTCTT A-CAC-TACT   772 M.spicil
739  --------GT GGCCACTTGT TCGCGCCTAG ACTCTGTCTT A-CAC-TACT   778    M.mus
739  --------GT GGCCACTTGT TCGCGCCTAG ACTCTGTCTT A-CAC-TACT   778   M.musA
736  --------GT GGCCACTTGT TCGCACCTAG ACTCTGTCTT A-CAC-TACT   775   M.musB
454  --------GT GGCCACCTGT TTGCACCAAG ACTCTCTCTT AACAC-TATT   494 M.hildeb
731  --------GT GGCTACTTGT CCGCACCTGG ACTCTC--TT A-CGT-TACT   768 S.longic
744  -AGGTCCAGT AGCCACTTGT CCGCACCTAG ACTCTCTCTT A-CAC-TACT   790 R.norveg
746  A-------GT GGACACTTGT TAGCTCCTAG ACTCTC---- A-CACATGCT   783 G.campes
737  --------GT GGTCACTTGT TAGCTACTAG ACTCTC---- A-CTC-TACT   772 M.auratu

     1451       1461       1471       1481       1491     1500
     |          |          |          |          |        |
974  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TAGAACCCGG  1023 M.pahari
564  TCCTGTCTGC AG                                            575 M.caroli
510  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGC                547 M.cookii
564  TCCTGTCTGC AG                                            575 M.spretu
773  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TAGAACCCGG   822 M.spicil
779  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TGGAACCCGG   828    M.mus
779  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TGGAACCCGG   828   M.musA
776  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TGGAACCCGG   825   M.musB
495  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TAGAACCGGG   544 M.hildeb
769  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TAGAACCTGG   818 S.longic
791  TCCTGTCTGC AGGCTGAGCT GGAAATCCAG AAAGATGCCT TAGAACCGGG   840 R.norveg
784  TCCTGTCTTC AGGCTGAGCT GGAAATCCAG AAAGATGCTT TAGAACCCGG   833 G.campes
773  TCCTGTCTGC AGGCTGAACT AGAAATCCAG AAAGACGCCT TAGAACCTGG   822 M.auratu

     1501       1511       1521       1531       1541     1550
     |          |          |          |          |        |
```

```
1024  GCAGAGAGTG GTCATTGTGG ATGACCTCCT GGCCACTGGA GGTAAAGAAC  1073 M.pahari
 823  GCAGAGAGTG GTCATTGTGG ATGACCTCCT GGCCACAGGA GGTAAAGAAC   872 M.spicil
 829  GCAGAGAGTG GTCATTGTGG ATGACCTCCT GGCCACAGGA GGTAAAGAAC   878    M.mus
 829  GCAGAGAGTG GTCATTGTGG ATGACCTCCT GGCCACAGGA GGTAAAGAAC   878   M.musA
 826  GCAGAGAGTG GTCATTGTGG ATGACCTCCT GGCCACAGGA GGTAAAGAAC   875   M.musB
 545  GCAGAGAGTG GTCATTGTGG ATGATCTCCT GGCCACTGGA GGTGA-GAAC   593 M.hildeb
 819  GCAGAGAGTG GTCATTGTGG ATGATCTTCT GGCCACTGGA GGTAA-GATC   867 S.longic
 841  GCAGAAAGTG GTCATTGTGG ATGATCTCCT GGCCACTGGA GGTAA-GGAC   889 R.norveg
 834  GCAGAAAGTG GTCATTGTAG ATGACCTCCT GGCCACTGGA GGTAA-GAGC   882 G.campes
 823  CCAGAAAGTG GTTGTTGTAG ATGATCTCCT GGCCACTGGA GGTAA-GAGC   871 M.auratu

      1551       1561       1571       1581       1591     1600
      |          |          |          |          |        |
1074  CAGCCCAAGA C--AAACAGG CTTCAAAGGG CCAGGCCCTG TCTGGGGTGC  1121 M.pahari
 873  CAACCCAAGA C--AAACAGA CTTCAAAGGG CCAGGCCCTG TCCTGGGTGC   920 M.spicil
 879  CAACCCAAGA C--AAACAGA CTTCAAAGGG CCAGACCCTG TCCTGGGTGC   926    M.mus
 879  CAACCCAAGA C--AAACAGA CTTCAAAGGG CCAGACCCTG TCCTGGGTGC   926   M.musA
 876  CAACCCAAGA C--AAACAGA CTTCAAAGGG CCAGGCCCTG TCCTGGGTGC   923   M.musB
 594  CACCCCAAGA CATAAACAGG CTTCAAAGGG CCTGTCCCTG TCCGGGGTGC   643 M.hildeb
 868  CACCCCAAGA CA-CAACAGG CTCTAAA--- ---GACCCTG TGCGGGGTGC   910 S.longic
 890  TACCCCAAGA CGTAAACTGG TTTCAAAGGG TCAGGCCCTG TCCGGGGTGC   939 R.norveg
 883  CACCCGCAC- --TTTACAGG TTACGAAGGG CCAGGCCCTG TCCAGG-TGC   928 G.campes
 872  CACTCTGTAG CATAAAGAGG TTTCAAAGGG ATAAGCCCTA TCCGGGGTGC   921 M.auratu

      1601       1611       1621       1631       1641     1650
      |          |          |          |          |        |
1122  TGACTAAACA AAGCGCTTGA ATACCT-TCT CTTTCTCTGT CCCTTCCCCC  1170 M.pahari
 921  TGACTAAGCA AAGAGCTTGA ATACCT---- CTTTCTCTGT CCCTTCCCCC   966 M.spicil
 927  TGACTAAGCA AAGAGCTTGA ACACCT-CCT CTTTCTCTGT CCCTTCCCCC   975    M.mus
 927  TGACTAAGCA AAGAGCTTGA ACACCT-CCT CTTTCTCTGT CCCTTCCCCC   975   M.musA
 924  TGACTAAGCA AAGAGCTTGA ACACCT-CCT CTTTCTCTGT CCCTTCCCCC   972   M.musB
 644  TGACTAAGCA ACGCTCTTTA ATACCT-TCT CTTTC--TGT CCCTTCTCCC   690 M.hildeb
 911  TGACTCAATA AAGAGCTTTA ATACCT-TCT CTTTTCTTGT CCCTCCGCCC   959 S.longic
 940  TGACTCAGCA AAGCGCTTTA ACACCT-TCT CCTTCCTTGT CCCCTTGCCC   988 R.norveg
 929  TGGCTTAGCA AAGAACCTTA CCAC-T-TCT CTTC-----T CCCTTCACAC   971 G.campes
 922  TGACTAAGCA A-GAGCCTTA CTACCTGTGT CTTTCCTCGT CCCTTCACCC   970 M.auratu

      1651       1661       1671       1681       1691     1700
      |          |          |          |          |        |
1171  CCCCCCCCCC AGGAACCATG TTTGCAGCCT GTGATCTGCT GCACCAGCTA  1220 M.pahari
 967  C--------- AGGAACCATG TTTGCGGCCT GTGATCTGCT GCACCAGCTC  1007 M.spicil
 976  C--------- AGGAACCATG TTTGCGGCCT GTGACCTGCT GCACCAGCTC  1016    M.mus
 976  C--------- AGGAACCATG TTTGCGGCCT GTGACCTGCT GCACCAGCTC  1016   M.musA
 973  C--------- AGGAACCATG TTTGCGGCCT GTGACCTGCT GCACCAGCTC  1013   M.musB
 691  C--------- AGGAACCATG TGTGCAGCCT GTGAGCTGCT GAACCAGCTA   731 M.hildeb
```

```
 960  C--------- AGGAACTATG TGTGCGGCCT GTGAGCTGCT GAACCAGCTA  1000 S.longic
 989  C--------- AGGAACCATG TGTGCAGCCT GTGAGCTGCT GAGCCAGCTG  1029 R.norveg
 972  C--------- AGGAACCATG TGTGCCGCCT GTCAGCTGCT GGGCCAGCTC  1012 G.campes
 971  C--------- AGGAACCATG TGCGCTGCCT GTGAGCTGCT GGGCCAGCTA  1011 M.auratu

      1701       1711       1721       1731       1741     1750
      |          |          |          |          |        |
1221  CGGGCTGAGG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCGCTGAA  1270 M.pahari
1008  CGGGCTGAAG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCGCTGAA  1057 M.spicil
1017  CGGGCTGAAG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCGCTGAA  1066    M.mus
1017  CGGGCTGAAG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCGCTGAA  1066   M.musA
1014  CGGGCTGAAG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCGCTGAA  1063   M.musB
 732  CGGGCTGAGG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCCCTGAA   781 M.hildeb
1001  CGGGCTGAGG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCCCTGAA  1050 S.longic
1030  CGGGCTGAGG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCCCTGAA  1079 R.norveg
1013  CGGGCCGAGG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCGCTGAA  1062 G.campes
1012  CAGGCTGAGG TGGTGGAGTG TGTGAGCCTG GTGGAGCTGA CCTCACTTAA  1061 M.auratu

      1751       1761       1771       1781       1791     1800
      |          |          |          |          |        |
1271  GGGCAGGGAG AGGCTAGGAC CTATACCATT CTTCTCTCTC CTCCAGTATG  1320 M.pahari
1058  GGGCAGGGAG AGGCTAGGAC CTATACCGTT CTTCTCTCTC CTCCAGTATG  1107 M.spicil
1067  GGGCAGGGAG AGGCTAGGAC CTATACCATT CTTCTCTCTC CTCCAGTATG  1116    M.mus
1067  GGGCAGGGAG AGGCTAGGAC CTATACCATT CTTCTCTCTC CTCCAGTATG  1116   M.musA
1064  GGGCAGGGAG AGGCTAGGAC CTATACCGTT CTTCTCTCTC CTCCAGTATG  1113   M.musB
 782  GGGCAGGGAG AGGCTAGGAC CTATACCCTT CTTCTCTCTC CTGCAGTATG   831 M.hildeb
1051  GGGCAGGGAG AGGCTAGGAC CTATACCATA CTTCTCTCTC CTACAGTATG  1100 S.longic
1080  GGGCAGGGAG AAGCTGGGAC CTGTGCCATT CTTCTCTCTC CTGCAGTATG  1129 R.norveg
1063  GGGCAGGGAG AAGCTGGGAC CTGTACCGTT CTTCTCTCTC CTGCAGTACG  1112 G.campes
1062  GGGCAGAGAG AAGCTAGGAT CAGTACCATT CTTCTCTCTC CTGCAATATG  1111 M.auratu

      1801
      |
1321  ACTGA                                                   1325 M.pahari
1108  ACTGA                                                   1112 M.spicil
1117  ACTGA                                                   1121    M.mus
1117  ACTGA                                                   1121   M.musA
1114  ACTGA                                                   1118   M.musB
 832  ACTGA                                                    836 M.hildeb
1101  AATGA                                                   1105 S.longic
1130  AGTGA                                                   1134 R.norveg
1113  AGTGA                                                   1117 G.campes
1112  AGTGA                                                   1116 M.auratu
```

Bibliography

Agulnik, S., C. Plass, W. Traut, and H. Winking (1993). Evolution of a long-range repeat family in chromosome 1 of the genus *Mus*. *Mammalian Genome 4*, 704–710.

Aquadro, C. and B. Greenberg (1983). Human mitochondrial DNA variation and evolution: Analysis of nucleotide sequences from seven individuals. *Genetics 103*, 287–312.

Avise, J. (1974). Systematic value of electrophoretic data. *Syst. Zool. 23*, 465–481.

Avise, J. and C. Aquadro (1982). A comparative summary of genetic distances in the vertebrates. *Evol. Biol. 15*, 151–158.

Bibb, M. J., R. A. Van Etten, W. C. T., M. Walberg, and D. A. Clayton (1981). Sequence and gene organization of the mouse mitochondrial DNA. *Cell 26*, 167–180.

Bonhomme, F. (1986). Evolutionary relationships in the genus *Mus*. *Curr Top Microbiol Immunol 127*, 19–34.

Bonhomme, F., J. Catalan, J. Britton-Davidian, V. M. Chapman, K. Moriwaki, E. Nevo, and L. Thaler (1984). Biochemical diversity and evolution in the genus *Mus*. *Biochem. Genet. 22*, 275–303.

Boursot, P., J.-C. Auffray, J. Britton-Davidian, and F. Bonhomme (1993). The evolution of house mice. *Ann. Rev. Ecol. Syst. 24*, 119–152.

Britten, R. (1986). Rates of DNA sequence evolution differ between taxonomic group. *Science 231*, 1393–1398.

Brown, W. (1985). The mitochondrial genome of animals. In R. MacIntyre (Ed.), *Molecular Evolutionary Genetics.*, pp. 95–130. New York: Plenum.

Brown, W., E. M. Prager, A. Wang, and A. C. Wilson (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol. 18*, 225–239.

Buth, D. (1984). The application of electrophoretic data in systematic studies. *Annu. Rev. Ecol. Syst. 15*, 501–522.

Cann, R. L. (1987). Mitochondrial DNA and human evolution. *Nature 325*, 31–36.

Cao, Y., J. Adachi, T. Yano, and M. Hasegawa (1994). Phyolgenetic place of guinea pigs: no support of rodent-polyphyly hypothesis from maximum-likelihood analyses of multiple protein sequences. . *Mol. Biol. Evol. 11*, 593–604.

Cao, Y., N. Okada, and M. Hasegawa (1997). Phyogenetic position of guinea pigs revisited. *Mol. Biol. Evol. 14*, 461–464.

Carleton, M. (1984). Introduction to rodents. In S. Anderson and J. Jones (Eds.), *Orders and families of recent mammals*, pp. 255–265. New York: John Wiley and Sons.

Carson, H. and K. Kaneshiro (1976). Drosophila of Hawaii: Systematics and ecological genetics. *Annu. Rev. Ecol. Syst. 7*, 311–345.

Chaline, J. and P. Mein (1979). *Les Rongeurs et l'Ecolution.* Paris: Doin.

Chevret, P., C. Denys, J.-J. Jaeger, and J. Michaux (1993). Molecular evidence that the spiny mouse (Acomys) is more closely related to gerbils (Gerbillinae) than to true mice (Murinae). *Proc. Natl. Acad. Sci. USA 90*, 3433–3436.

Corbet, G. B. and J. E. Hill (1991). *A world list of Mammalian species* (3rd ed.). New York: Oxford University Press.

Cremisi, F., R.Vignali, R.Batistoni, and G. Barsacchi (1988). Heterochromatic DNA in *Triturus* (Amphibia, Urodela) II. A centromeric satellite DNA. *Chromosoma 97*, 204–211.

Dallas, J., B. Dod, P. Boursot, E. Prager, and F. Bonhomme (1995). Population subdivision and gene flow in Danish house mice. *Mol. Ecol. 4*, 311–320.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection.* London: Murray.

Davison, D. (1985). Sequence similarity ('homology') searching for molecular biologists. *Bull. Math. Biol. 47*, 437–474.

Deininger, P. and G. Daniels (1986). The recent evolution of mammalian repetitive DNA elements. *Trends Genet. 2*, 76–80.

D'Erchia, A., C. Gissi, G. Pesole, C. Saccone, and U. Arnason (1996). The guinea-pig is not a rodent. *Nature 381*, 597–600.

Dod, B., E. Mottez, E. Desmarais, F. Bonhomme, and G. Roizes (1989). Concerted evolution of light satellite DNA in genus *Mus* implies amplification and homogenization of large blocks of repeats. *Mol. Biol. Evol. 6*, 478–491.

Dunn, G. and B. Everitt (1982). *An Introduction to Mathematical Taxonomy.* New York: Cambridge University Press.

Dush, M., J. Sikela, S. Khan, J. Tischfield, and P. Stambrook (1985). Nucleotide Sequence and Organization of the Mouse Adenine Phosphoribosyltransferase Gene: Presence of a Coding Region Common to Animal and Bacterial Phosphoribosyltransferases that has a Variable Intron/Exon Arrangement. *Proc. Natl. Acad. Sci. USA 82*, 2731–2735.

Felsenstein, J. (1978). Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool. 27*, 401–410.

Felsenstein, J. (1981a). A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *J. Linn. Soc. 16*, 183–196.

Felsenstein, J. (1981b). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol. 17*, 368–376.

Felsenstein, J. (1981c). Evolutionary trees from gene frequencies and quantitative characters: Finding maximum likelihood estimates. *Evolution 35*, 1229–1242.

Felsenstein, J. (1988). Phylogenies from molecular sequences: Inference and reliability. *Annu. Rev. Genet. 22*, 521–565.

Felsenstein, J. (1989). Phylip-phylogeny inference package, version3.4. *Cladistics 5*, 164–166.

Ferris, S., R. Sage, E. Prager, U. Ritte, and A. Wilson (1983). Mitochondrial DNA evolution in mice. *Genetics 105*, 681–721.

Fieldhouse, D. and G. B. Golding (1993). The rat adenine phosphoribosyltransferase sequence shows evolutionary rate variation among exons in rodents. *Genome 36*, 1107- 1110.

Fieldhouse, D. and G. B. Golding (1996). Evolutionary rate variation within *Mus* APRT. *Genome 39*, 909–913.

Figueroa, F., M. Kasahara, H. Tichy, E. Neufeld, U. Ritte, and J.Klein (1987). Polymorphism of unique noncoding DNA sequences in wild and laboratory mice. *Genetics 117*, 101–108.

Fitch, W. (1970). Distinguishing homologous from analogous proteins. *Syst. Zool. 19*, 99–113.

Fort, P., F. Bonhomme, P. Darlu, M. Piechaczyk, P. Jeanteur, and L. Thaler (1985). Clonal divergence of mitochondrial DNA versus populational evolution of nuclear genome. *Evol. Theor. 7*, 81–90.

Frye, M. and S. Hedges (1995). Monophyly of the order Rodentia inferred from mitochondrial DNA sequences of the genes for 12SrRNA, 16SrRNA, and tRNA-valine. *Mol. Biol. Evol. 12*, 168–176.

Gadaleta, G., G. Pepe, G. D. Candid, C. Quagliariello, E. Sbisa, and C. Saccone (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.

Gall, J. and M. Pardue (1969). Formationand detection of RNA-DNA hybrid molecules in cytological preparations. *Proc. Natl. Acad. Sci. USA 63*, 378–383.

Goodman, M. (1961). The role of immunochemical differences in the phyletic development of human behavior. *Hum. Biol. 33*, 131–162.

Goodman, M. (1963). Serological analysis of the systematics of recent hominoids. *Hum. Biol. 35*, 377–424.

Goodman, M., M. Miyamoto, and J. Czelusniak (1987). Pattern and process in vertebrate phylogeny revealed by coevolution of molecules and morphologies. In C. Patterson (Ed.), *Molecules and Morphology in Evolution: Conflict or Compromise?*, pp. 141–176. Cambridge, England: Cambridge Univ. Press.

Graur, D., A. Hide, A. Zarkikh, and W.-H. Li (1992). The biochemical phylogeny of guinea pigs and gundies and the paraphyly of the order Rodentia. *Comp. Biochem. Physiol. B 101*, 495–498.

Graur, D., W. Hide, and W.-H. Li (1991). Is the guinea-pig a rodent? *Nature 351*, 649–652.

Gray, G. and W. Fitch (1983). Evolution of antibiotic resistance genes: The DNA sequence of a kanamycin resistance gene from *Staphylococcus aureua. Mol. Biol. Evol. 1*, 57–66.

Greenberg, B., J. Newbold, and A. Sugino (1983). Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene 21*, 33–49.

Hartenberger, J.-L. (1985). The order rodentia: Major questions on their evolutionary origin, relationships, and suprafamilial systematics. In W. Luckett and J.-L. Hartenberger (Eds.), *Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis*, pp. 1–33. New York: Plenum.

Hasegawa, M., Y. Cao, J. Adachi, and T. Yano (1992). Rodent polyphyly? *Nature 355*, 595.

Hasegawa, M., T. Yano, and H. Kishino (1984). A new molecular clock of mtDNA and the evolution of hominoids. *Proceedings of the Japan. Academy 60*, 95–98.

Hennig, W. (1950). *Grundzuge einer Theorie der Phylogenetischen Systematik.* Berlin: Deutscher Zentralverlag.

Hennig, W. (1965). Phylogenetic systematics. *Annu. Rev. Entomol. 10*, 97–116.

Hennig, W. (1966). *Phylogenetic Systematics.* Urbana: University of Illinois Press.

Hennig, W. (1969). *Die Stammesgeschichte der Insekten.* Frankfurt am Main: Senchenberg Naturf. Ges.

Hennig, W. (1975). Cladistic analysis or cladistic classification?: A reply to Ernst Mayr. *Syst. Zool. 24*, 244–256.

Hillis, D. (1987). Molecular versus morphological approaches to systematics. *Annu. Rev. Ecol. Syst. 18*, 23–42.

Hillis, D. and C. Moritz (1990). *Molecular systematics.* Sunderland, Massachusetts: Sinauer Associates, Inc.

Hubby, J. and R. Lewontin (1966). A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura. Genetics 54*, 577–594.

Hubby, J. and L. H. Throckmorton (1965). Protein differences in Drosophila. II. Comparative species genetics an evolutionary problem. *Genetics 52*, 203–215.

Innis, M., D. Gelfand, J. Sninsky, and T. White. (1990). *PCR Protocols: A Guide to Methods and Applications.* San Diego: Academic Press.

Jaeger, J.-J., H. Tong, and C. Denys (1986). The age of *Mus-Rattus* divergence: paleontological data compared with the molecular clock. *C. R. Acad. Sci. Paris, Ser. II 302*, 917–922.

Janke, A., G. Feldmaier-Fuchs, W. Thomas, A. von Haeseler, and S. Paabo (1994). The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics 137*, 243–256.

Janke, A., X. Xu, and U. Arnason (1997). The complete mitochondrial genome of the wallaroo (Macropus robustus) and the phylogenetic relationship among Monotremata, Marsupialia and Eutheria. *Proc. Natl. Acad. Sci. USA 94*, 1276–1281.

John, H., M. L. Birnsteil, and K. W. Jones (1969). RNA-DNA hybrids at cytological levels. *Nature 223*, 582–587.

Jukes, T. and C. Cantor (1969). Evolution of protein molecules. In H. Munro (Ed.), *Mammalian Protein Metabolism*, pp. 21–132. New York: Academic Press.

Kalb, V., S. Glasser, D. King, and J. Lingrel (1983). A cluster of repetitive elements within a 700 base pair region in the mouse genome. *Nucleic Acid Res. 11*, 2177–2184.

Kido, Y., M. Aono, T. Yamaki, K. Matsumoto, S. Murata, M. Saneyoshi, and N. Okada (1991). Shaping and reshaping of salmonid genomes by amplification of tRNA- derived retroposons during evolution. *Proc. Natl. Acad. Sci. USA 88*, 2326–2330.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol. 16*, 111–120.

Kimura, M. (1987). Molecular evolutionary clock and the neutral theory. *J. Mol. Evol. 26*, 24–33.

Kiriakoff, S. (1959). Phylogenetic systematics versus typology. *Syst. Zool. 8*, 117–118.

Kleppe, K., E. Ohtsuka, R. Kleppe, I. Molineux, and H. Khorana (1971). Studies on polynucleotides xcvi. repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *J. Mol. Biol. 56*, 341–361.

Klien, J., H.Tichy, and F. Figueroa (1987). Complete mitochondrial DNA Sequence of the fat dormouse, *Glis glis*: further evidence of rodent paraphyly. *Mol. Biol. Evol. 15*, 499–505.

Klien, J., V. Vincekm, M. Kasahara, and F. Figueroa (1988). Probing mouse origins with random DNA probes. *Curr. Top. Microbiol. Immun. 137*, 55–63.

Kocher, T. and T. White (1989). Evolutionary analysis via PCR. In H. Erlich (Ed.), *PCR Technology: Principles and Applications for DNA Amplification.* New York: Stockton Press.

Krayev, A., T. Markusheva, D. Kramerov, A. Ryskov, K.G.Skryabin, A. Bayev, and G. Georgiev (1982). Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing. *Nucleic Acid Res. 10*, 7461–7475.

Kuma, K. I. and T. Miyata (1994). Mammalian phylogeny inferred from multiple protein data. *Jpn. J. Genet. 69*, 555–566.

Lewontin, R. and J. Hubby (1966). A molecular approach to the study of genic heterozygosity in natural populations. II. Amounts of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura. Genetics 54*, 595–609.

Li, W. and D. Graur (1991). *Fundamentals os Molecular Evolution.* Sunderland, Massachusetts: Sinauer Associates, Inc.

Li, W.-H., D. L. Ellsworth, J. Krushkal, H. J. Chang, and D. Hewett-Emmett (1996). Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Mol. Phylogenet. Evol. 5*, 182–187.

Li, W.-H., M. Gouy, P. Sharp, C. O'hUigin, and Y. Yang (1990). Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. *Proc. Natl. Acad. Sci. USA 87*, 6703–6707.

Li, W.-H., W. Hide, and D. Graur (1992). Origin of roders and guinea-pigs. *Nature 359*, 277–278.

Li, W.-H., W. Hide, A. Zharkika, D. Ma, and D. Graur (1992). The molecular taxonomy and evolution of the guinea pig. *J. Hered. 83*, 174–181.

Li, W.-H., M. Tanimura, and P. Sharp (1987). An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *J. Mol. Evol. 25*, 330–342.

Li, W.-H. and C.-I. Wu (1985). Evolution of DNA sequence. In R. MacTntyre (Ed.), *Molecular Evolutionary Genetics*, pp. 1–130. New York: Plenum.

Linnaeus, C. (1758). *Systema Naturae* (10th ed.). Stockholm: Stockholm.

Lovtrup, S. (1977). *The phylogeny of vertebrata.* London: Wiley Interscience.

Lozovskaya, E., D. Petrov, and D. Hartl (1993). A combined molecular and cytogenetic approach to genome evolution in *Drosophila* using large-fragment DNA cloning. *Chromosoma 102*, 253–266.

Luckett, W. P. and J.-L. Hartenberger (1993). Monophyly or polyphyly of the order Rodentia: possible conflict between morphological and molecular interpretations. *J. Mammal. Evol. 1*, 127–147.

Ma, D.-P., A. Zharkika, D. Graur, J. L. Vanderberg, and H.-W. Li (1993). Structure and evolution of opposum, guinea pig, and porcupine cytochrome b genes. *J. Mol. Evol. 36*, 327–334.

MacGregor, H. and S. Sessions (1986). The biological significance of variation in satellite DNA and heterochromatin in news of the genus *Triturus*: An evolutinary perspective. *Roy. Soc. London Ser. B 312*, 243–259.

Martignetti, J. and J. Brosius (1993). Neural BC1 RNA as an evolutionary marker: guinea pig remains a rodent. *Proc. Natl. Acad. Sci. USA 90*, 9698–9702.

Mayr, E. (1969). *Principles of Systematic Zoology.* New York: McGraw-Hill.

Mayr, E. (1983). *The growth of biological thought: diversity, evolution, and inheritance.* Cambridge, MA: Harvard Univ. Press.

Moore, W. (1995). Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution 49*, 718–726.

Moshier, J., A. Deutch, and R. Huang (1987). Structure and *in vitro* transcription of a mouse B1 cluster containing a unique B1 dimer. *Gene 58*, 19–27.

Mullis, K. and F. Faloona (1987). Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods Enzymol. 155*, 335–350.

Nedbal, M., M. Allard, and R. Honeycutt (1994). Molecular systematics of hystricognath rodents: Evidence from the mitochondrial 12SrRNA gene. *Mol. Phylo. Evol. 3*, 206–220.

Nishioka, Y. (1989). Genome comparison in the genus *Mus*: a study with B1, MIF (mouse interspersed fragment), centromeric, and Y- chromosomal repetitive sequences. *Cytogenet. Cell Genet. 50*, 195–200.

Nuttall, G. (1904). *Blood immunity and Blood Relationship.* Cambridge, England: Cambridge Univ. Press.

Ochman, H., A. S. Gerber, and D. Hartl (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics 120*, 621–623.

O'hUigin, C. and W.-H. Li (1992). The molecular clock ticks regularly in muroid rodents and hamsters. *J. Mol. Evol. 35*, 377–384.

Paabo, S. (1996). Mutation hot spots in the mitochondrial microcosm. *Am. J. Hum. Genet. 59*, 493–496.

Paabo, S., J. Gifford, and A. Wilson (1988). Mitochondrial DNA sequences from a 7000-year-old brain. *Nucl. Acid Res. 16*, 9775- 9787.

Pamilo, P. and M. Nei (1988). Relationships between gene trees and species trees. *Mol. Biol. Evol. 5*, 568–583.

Patterson, C. (1987). *Molecules and morphology in evolution: Conflict or Compromise?* Cambridge, England: Cambridge Univ. Press.

Porter, C. A., M. Goodman, and M. Stanhope (1996). Evidence on mammalian phylogeny from sequences of exon 28 of the von Willebrand factor gene. *Mol. Phylo. Evol. 5*, 89–101.

Prager, E., H. Tichy, and R. D. Sage (1996). Mitochondrial DNA sequence variation in the eastern house mouse, *Mus musculus*: comparison with other house mice and report of a 75-bp tandem repeat. *Genetics 143*, 427–446.

Reyes, A., G. Pesole, and C. Saccone (1998). Complete mitochondrial DNA Sequence of the fat dormouse, *Glis glis*: further evidence of rodent paraphyly. *Mol. Biol. Evol. 15*, 499–505.

Rogers, J. (1985). The origin and evolution of retroposones. *Int. Rev. Cytol. 93*, 187–279.

Rozen, S. and H. J. Skaletsky (1997). Primer 3. *http://www-genome.wi.mit.edu/ genome_software/ other/ primer3.html.*

Ruvinsky, A., A. Polyakov, A. Agulnik, H. Tichy, and F. Figueroa (1991). Low diversity of t-haplotypes in the estern form of the house mouse, *Mus musculus* L. *Genetics 127*, 161–168.

Sabeur, G., G. Macaya, F. Kadi, and G. Bernardi (1993). The isochore patterns of mammalian genomes and their phylogenetic implications. *J. Mol. Evol. 37*, 93–108.

Saccone, C., G. Pesole, and E. Sbisa (1991). The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *J. Mol. Evol. 33*, 83–91.

Sage, R. (1981). Wild mice. In H. Foster, J. Small, and J. Fox (Eds.), *The Mouse in Biomedical Research, Vol. 1*, pp. 39–90. New York: Academic Press.

Sage, R., W. Atchley, and E. Capanna (1993). House mice as models in systematic biology. *Syst. Biol. 42*, 523–561.

Saitou, N. and M. Nei (1987). The neighbor-joining method: A new method for reconstruction phylogenetic trees. *Mol. Biol. Evol. 4*, 406–425.

Sarich, V. (1985). Rodent macromolecular systematics. In W. Luckett and J.-L. Hartenberger (Eds.), *Evolutionary relationships among rodents*, pp. 423–452. New York: Plenum.

Sarich, V. and A. Wilson (1966). Quantitative immunochemistry and the evolution of primate albumins: Micro-complement fixation. *Science 154*, 1563–1566.

Sarich, V. and A. Wilson (1967). Immunological time scale for hominid evolution. *Science 158*, 1200–1203.

Serdobova, I. and D. Kramerov (1993). Use of the short retroposon B2 in the study of phylogenetic relationship in rodents. *Genetika 29*, 1969–1981.

She, J., F. Bonhomme, T. Boursot, L. Thaler, and F. M. Catzeflis (1990). Molecular phylogenies in the genus *Mus*: comparative analysis of electrophoretic scnDNA hybridization, and mtDNA RFLP data. *Biological Journal of the Linnean Society 41*, 83–103.

Shoshani, J., M. Goodman, J. Czelusniak, and G. Braunitzer (1985). A phylogeny of Rodentia and other eutherian orders: Parsimony Analysis utilizing amino acid sequences of alpha and beta hemoglobin chains. In W. Luckett and J.-L. Hartenberger (Eds.), *Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis*, pp. 192–210. New York: Plenum.

Silver, L. M. (1995). *Mouse genetics: Concepts and Applications.* New York: Oxford Univ. Press.

Simpson, G. (1945). The principles of classification and a classification of mammals. *Bull. Amer. Mus. Nat. Hist. 85*, 1–350.

Simpson, G. (1961). *Principles of Animal Taxonomy.* New York: Columbia University Press.

Sogin, M., H. Elwood, and J. Gunderson (1986). Evolutionary diversity of eukaryotic small-subunit rRNA genes. *Proc. Natl. Acad. Sci. USA 83*, 1383–1387.

Sogin, M., J. Gunderson, H. Elwood, R. Alonso, and D. Peattie (1989). Phylogenetic meaning of the kingdom concept: An unusual ribosomal RNA from *Giardia lamblia. Science 243*, 75–77.

Sourrouille, P., C. Hanni, M. Ruedi, and F. M. Catzeflis (1995). Molecular systematics of *Mus crociduroides*, an endemic mouse of Sumatra (Muridae: Rodentia).

Mammalia 59, 91–102.

Strimmer, K. and A. von Haeseler (1996). Quartet Puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol. 13*, 964–969.

Stryer, L. (1988). *Biochemistry* (3rd ed.). New York: W. H. Freeman and Company.

Suzuki, H. and Y. Kurihara (1994). Genetic variation of ribosomal rna in the house mouse, *mus musculus*. In K. Moriwaki, T. Shiroishi, and H. Yonekawa (Eds.), *Genetics in Wild Mice*, pp. 107–119. Tokyo: Jap. Sci. Soc. Press.

T., M., H. Hayashida, R. Kikuno, M. Hasegawa, M. Kobayashi, and K. Koike (1982). Molecular clock of silent substitution: at least six-fold preponderance of silent changes in mitochondrial genes over those in nuclear genes. *J. Mol. Evol. 19*, 18–35.

Tahler, L. (1966). Les rongeurs fossiles du Bas-Languedoc dans leurs rapports avec 1'histoire des faunes et la stratigraphie du tertiare d'Europe. *Mem. Mus. Nat. Hist. Paris, series C 17*, 1–297.

Thompson, J., D. G. Higgins, and T. J. Gibson (1994). ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, postions-specific gap penalties and weight matrix choice. *Nucleic Acids Research 22*, 4673–4680.

Tullberg, T. (1899). Ueber das System der nagetiere: eine phylogenetische studie. . *Nova Acta Reg. Soc. Scient. Upsala 18*, 1–514.

Ullu, E. and C. Tschudi (1984). Alu sequences are processed 7 SL RNA genes. *Nature 312*, 171–172.

Walberg, M. and D. Clayton (1981). Sequence and properties of the human KB cell and mouse D-loop regions of mitochondrial DNA. *Nucl. Acids Res. 9*, 5411–5421.

White, M. (1973). *Animal Cytology and Evolution.* Cambridge, England: Cambridge Univ. Press.

Wiley, E. (1981). *Phylogenetics: The Theory and Practice of Phylogenetic Systematics.* New York: John Wiley and Sons.

Willee, C. A., W. Walker, and R. Barnes (1979). *Introduction to Animal Biology.* London: W.B. Saunders.

Wilson, A., R. Cann, S. Carr, M. G. Jr., U. Gyllenstin, K. Helm-Bychowski, R. Higuchi, S. Palumbi, E. Prager, R. Sage, and M. Stoneking (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc. 26*, 375–400.

Wilson, A., H. Ochman, and E. M. Prager (1987). Molecular scale for evolution. *Trends Genet. 3*, 241–247.

Wilson, D. and D. Reeder (1993). *Mammal species of the world. A taxonomic and geographic reference.* Washington: Smithsonian Institution Press.

Wolf, B., K. Reinecke, K. Aumann, R. Brigelius-Flohe, and L. Flohe (1993). Taxonomical classification of the guinea pig based on its Cu/Zn superoxide dismutase sequence. *Biol. Chem. Hoppe Seyler 374*, 641–649.

Wood, A. (1985). The relationship, origin and dispersal of the hystricognathous rodents. In W. Luckett and J.-L. Hartenberger (Eds.), *Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis*, pp. 475–513. New York: Plenum.

Wu, C.-I. and W.-H. Li (1985). Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA 82*, 1741–1745.

Young, J. (1950). *The life of vertebrates.* London: Oxford University Press.

Zietkiewicz, E., C. Richer, D. Sinnett, and D. Labuda (1998). Monophyletic origin of alu elements in primates. *J. Mol. Evol. 47*, 172–182.

Zuckerkandl, E. and L. Pauling (1962). Molecular disease, evolution and genic heterogeneity. In M. Kasha and B. Pullman (Eds.), *Horizons in Biochemistry.*, pp. 189–255. New York: Academic Press.

Zuckerkandl, E. and L. Pauling (1965). Evolutionary divergence and convergence in proteins. In V.Bryson and H. J. Vogel (Eds.), *Evolving Genes and Proteins*, pp. 97–166. New York: Academic Press.