A SYSTEMATIC APPROACH FOR IDENTIFYING EVOLUTIONARILY SIGNIFICANT UNITS FOR CONSERVATION: THE DILEMMA OF SUBSPECIES

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

in Candidacy for the Degree of

Doctor of Philosophy

by

George Dale Amato

May 1994

ABSTRACT

A SYSTEMATIC APPROACH FOR IDENTIFYING EVOLUTIONARILY SIGNIFICANT UNITS FOR CONSERVATION: THE DILEMMA OF SUBSPECIES

George Dale Amato

Yale University

1994

Identifying the units of conservation is an essential first step in designing a successful taxon oriented conservation strategy. In this thesis I propose a scientific framework for identification of conservation units employing a cladistic analysis of molecular characters. Two applications of this proposed framework are presented to illustrate useful laboratory techniques, methods of analysis, and important limitations of this methodology. In chapter 1, an analysis of conservation units for the *Caiman crocodilus* (common caiman) complex using a character analysis of 12S and 16S ribosomal mitochondrial DNA sequences demonstrates the usefulness of combining limited direct sequencing with allele specific PCR technology. The distribution of characters supports designating three conservation units corresponding to the trinomials *Caiman crocodilus crocodilus*, *C.c. fuscus*, and *C.c.yacare*. Controversies in higher level crocodilian relationships are reviewed in chapter 2. A total evidence analysis of extant rhinoceros is presented in chapter 3 in order to provide context for an

assessment of conservation units for *Dicerorhinus sumatrensis* (Sumatran rhino). In this study I assess the number of conservation units determined by a cladistic analysis of completely sequenced 12S and 16S ribosomal mitochondrial DNA for the allopatric populations of Sumatran Rhinos. Nucleotide sequence data weakly supports separation of the three units. Subjective assessments of the effects of sample size and inferred degrees of divergence support a policy of treating *Dicerorhinus sumatrensis* as a single conservation unit.

Finally, a summary of the strengths and limitations of this generalized approach to determining conservation is presented.

ACKNOWLEDGEMENTS

First, I would like to thank the members of my dissertation committee for their valuable advice and support. Jeff Powell not only taught me how to be a biologist, but included me in an extended traditional "family" of evolutionary biologists. To a large extent I owe my professional development to Jeff. Rob DeSalle is my closest colleague and friend. Anyone familiar with Rob's research will see his strong influence on all of the research presented here. I cannot overstate Rob's contributions to the successful completion of this dissertation. Alison Richard has provided important professional and personal support. Alison's first hand experiences in doing "on the ground" conservation provided me with unique and important insights. Charles Remington's extraordinary knowledge of organismal biology makes him a treasure in the Yale Department of Biology. Rob Dorit kindly joined my committee after his recent arrival here at Yale and provided many "fresh" insights and suggestions. Elisabeth Vrba is always enthusiastic and challenging in her alternative views of the world. Oliver Ryder has informally acted as an outside committee member. His early encouragement helped shape the following research and his continued collegiality is greatly appreciated.

John Gatesy has worked closely with me on many research projects. His constant challenges to work harder and think more clearly have dramatically improved the quality of my science. His friendship has helped me through many difficult periods.

Special friendships and encouragement play important roles during the completion of a dissertation. My good friend Mark Norell shares many of my interests and curiosity about the world. He is was also the first person to teach me about rigorous systematics. We have been friends for so long that he seems

more like a brother. In the same way, Gisella Caccone seems like a sister. She has always been caring and supportive in a way that seems more like family than friend. Robin Absher and Annie Williams have also been sources of strength and encouragement. Trying to live up to their expectations has made me a better person.

Colleagues in the lab have been valuable for their contributions to my science as well as their personal support. I would especially like to thank Jenny Gleason, Dave Hale, Michel Milinkovitch, and James Gibbs.

This dissertation has been shaped by my work in conservation for NYZS/The Wildlife Conservation Society. Dan Wharton has contributed to every aspect of this research. He has also patiently taught me the zoological park and conservation business. His respect and friendship has sustained me through very difficult times. Other NYZS colleagues have taught and inspired me. Peter Brazaitis knows more about caiman than anyone else in the world. I am determined to continue his legacy. As a laboratory biologist I have relied on my field biologist colleagues to help me view the world from a more complete perspective. Patricia Moehlman is an extraordinarily gifted and dedicated field biologist. I will always appreciate her generous sharing of her valuable time. George Schaller was my childhood hero. Our work together has been intellectually enlightening and spiritually uplifting. And finally, I am grateful to be part of the group of conservation biologists instructed and inspired by William Conway. His extraordinary intellect has largely shaped modern conservation. It is a privilege to work with him to conserve the world's wildlife. Financial support for this research has come from NYZS/WCS and the Institute for Museum Services, the Conservation Endowment Fund of the Association of Zoos and Aquariums, Mrs. L. Emery Katzenbach, and the Charles E. Culpeper Foundation.

Finally, I would like to thank the most important people in my life. I have the most wonderful, loving, supportive, extended family. It is from them that I draw all my strength and joy. Dick and Beverly Waterhouse have been my "second" parents. They have always been supportive and loving. I deeply regret that Beverly passed away just as I was nearing completion of this dissertation. She and Dick always had such pride in me -- I would have loved for her to see me complete it.

George and Marie Amato are perfect parents. I owe them everything that is important to me. There is nothing that can substitute for loving, caring, interesting parents. I have always felt so "lucky" to be their son.

I am blessed with two beautiful children - Mary and Grace. They are constant sources of joy. There are many elements of this dissertation that were completed especially for them. My wife Amanda is extraordinarily intelligent, beautiful, and patient. It is not an overstatement to say that without her I would not have completed this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Conservation biology is a relatively new discipline. It arose out of a perception that the application of rigorous scientific methodologies from more traditional disciplines could enhance the success of attempts to conserve biodiversity in the face of the current extinction crisis. This dissertation explores the utility of applying a framework derived from both systematics and evolutionary biology to determine units of conservation. Assessing units of conservation has proved problematic and been termed the "dilemma of subspecies".

Species Survival Programs (SSPs) are coordinated efforts by accredited zoological parks and aquariums to comprehensively and cooperatively develop conservation plans for selected endangered species. Historically, species-oriented conservation programs have dominated conservation (especially vertebrates) and are likely to continue to do so in the immediate future. The emphasis on species-oriented programs is demonstrated not only by SSPs, but also by the international programs of the International Union for the Conservation of Nature (IUCN) including the Captive Breeding Specialist Group (CBSG) and the U.S. governments Endangered Species Act. Identifying the unit of conservation is critical to the establishment of a species-oriented program. More precisely, the selection of the unit of conservation is complicated by broad disagreements about taxonomy, species concepts, and partitioning of genetic variation at or below the species boundary.

Ryder (1986) summarized these concerns as they relate to *ex situ* conservation by focusing on the subspecies dilemma. Tigers were the target of one of the first SSP programs, and have, ever since, provided an easily understood example of this problem. Currently, tigers (*Panthera tigris*) are

categorized as five extant subspecies. The following question confronts conservation planners: do we need five separate conservation programs for tigers with each named subspecies as our unit of conservation, or can we consider tigers a single taxon existing as a metapopulation? This decision has far-reaching consequences for managers that are both practical, biological, and perhaps political. If the decision is to manage them as separate units, the consequences not only include obtaining a sufficient number of founders for five separate *ex situ* programs, but finding money and space for these programs as well. Additionally, a commensurate number of refuges and national parks in all five range areas will need to be set aside and protected. There is also the possibility that we will artificially prevent gene flow between subdivided populations that historically interacted to varying degrees. Conversely, the concern of treating tigers as a single population is the risk of outbreeding depression (Templeton 1986) and the resulting loss of fitness.

The conservation community recognized early (Ryder 1986) that the existing taxonomy of trinomial names was not useful for determining units of conservation. There are numerous anecdotes of both unjustified splitting and lumping. It is clear that the term "subspecies" constitutes varying degrees of differentiation to different taxonomists. This led to the reluctant introduction of the term evolutionarily significant unit (ESU; Conway pers. com.) as a more accurate description of what should be the unit of conservation. It was suggested that researchers should attempt to identify evolutionary units below the species level by use of "natural history information, morphometrics, range and distribution data, as well as protein electrophoresis, cytogenetic analysis, and restriction mapping of nuclear and mitochondrial DNA" (Ryder 1986). However, it is worth noting that one of the recommendations that was made is that mixing of ESUs is appropriate if the extinction of a small population would

jeopardize the higher taxon.

While the dilemma of subspecies for conservation is obvious, methods for partitioning genetic variation at, or below, the species level for taxonomy has always been problematic. Linnaeus (1751) first included the term *varietas* (variety) to describe individuals that varied from the described species type. His varieties included genetic and nongenetic variants. Also, the ambiguity between intra- and interpopulation variation rendered the term "variety" meaningless for understanding units of evolution. However, it persisted in the literature for a long time (Gloger 1833). Rothschild, Hartert, and Jordan (1894) argued that the term variety should be discarded; and that individual variation be termed "aberration" while geographic races be termed subspecies. Taxonomists continued to develop terminology to deal with the variation below the level of Linnaeus's "typological" species. For example, the term *Formenkreis* was used by Kleinschmidt (1900) to describe groups of geographic variants that belonged to a higher category. *Rassenkreis* was introduced as a better alternative to *Formenkreis* (Rensch 1929) since *Formenkreis* had a different meaning in paleontology. Huxley (1940) was the first to introduce the terms "monotypic species" and "polytypic species" in order to reconcile taxonomy and variation. After this, the concept of polytypic species was demonstrated by reviews of variation in numerous taxa (Mayr 1951; Remington 1951; Voipio 1950). Mayr (1963) provides a detailed discussion of the importance of the role of the polytypic species concept in the synthesis of modern population genetics and evolutionary theory. Unfortunately, he also suggests that it plays an important role in reducing the number of named species (an argument of convenience, not science). Mayr does, however, recognize that many authors included a number of "good species" under a single polytypic species (e.g. Vaurie 1955; Ellerman and Morrison-Scott 1951).

Furthermore, Mayr suggests that the term subspecies be used for taxonomy with an emphasis on geographic "types" (not gradients) identified by a trinomial, while population terms (i.e. deme, race, geographic isolate, cline) be used for evolutionary biology.

Clearly, population genetics deals more comprehensively with partitioning of genetic variation than does taxonomy. Genetic distances have often been used as indicators of genetic discontinuities between and within taxa (Ayala 1976; Selander 1976; Avise and Aquadro 1982; Nei 1987). The implications of these studies for various taxonomic ranks was to suggest certain levels of divergence as markers for taxonomic rank. Although increasing genetic distance is frequently proportional to taxonomic rank, all authors suggest that these levels of divergence are not universal.

For partitioning within-species genetic variation, such distance measures as Fst's, Kst's, and Gst's (DeSalle et al. 1987; Takahata and Palumbi 1986; Hudson et al. 1993) provide models for measuring and inferring biological consequences of gene flow, migration rates and patterns, founder events, etc. Dobzhansky (1951) defines evolution as "a change in the genetic composition of populations". However, using population genetics measures for taxonomy (in this case identifying conservation units) is fraught with problems. Populations in nature reflect varying degrees of subdivision that may reveal hierarchy, but are tokogenetic rather than phylogenetic (Hennig 1966). In naming meaningful conservation units, it is important to remain above the "line of death" (Vrana and Wheeler 1992) that separates tokogeny from phylogeny. This is not to deny that evolutionary events happen at the level best described by population genetics theory, but that subjective assessments of gene frequency differences for taxonomy are not useful. Rather, population genetics theory offers valuable insights into managing units of conservation by allowing for assessments of

levels of heterozygosity, reconstructing natural levels of gene flow between fragmented populations, determining coefficients of inbreeding, etc. However, the naming of a conservation unit is ultimately taxonomy and rests on our view of species as distinct units.

For this reason, the definition of a conservation unit (or ESU) has been regarded as inexorably tied with various definitions of species concepts (O'Brien and Mayr 1991; Amato 1991; Rojas 1992; Vogler and DeSalle in press). Two species concepts have been suggested as important guides for determining conservation units. O'Brien and Mayr (1991) suggested that the biological species concept (BSC) can serve as a useful framework for identifying units of conservation. This suggestion was made in response to the controversial determination that Florida panthers (*Felis concolor*) are a "hybrid" population (O'Brien et al. 1990). Defining the Florida panther population as a hybrid population caused concern among managers since "hybrids" are not protected under the Endangered Species Act. The BSC [defined as "groups of actually or potentially interbreeding populations that are reproductively isolated from other such groups (Mayr 1942)"] was invoked to provide an evolutionary framework for distinguishing between intra and interspecific hybrids. O'Brien and Mayr, further suggest that the BSC provides important guidance for determining units of conservation. A number of authors have since criticized this suggestion for both theoretical and practical reasons (Amato 1991; Amato and Wharton in press, Vogler and DeSalle in press). In most cases the criteria of "potential to interbreed" must be inferred by measures of genetic and morphological similarity in the absence of characterizing this trait genetically. This inference is subjective since it is based on an assumption that reproductive isolation is proportional to these similarity measures. A number of data sets demonstrate a lack of concordance for such inferences (Aquadro and Avise

1982; Vogler and DeSalle in press).

The second species concept that has been suggested as useful for identifying units of conservation is the phylogenetic species concept (PSC) (Cracraft 1991; Flessness and Barowclough in press; Amato and Gatesy in press). Cracraft (1983) defines a phylogenetic species as an "irreducible cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent". The PSC assumes that species are distinct, diagnosable units identified by a character or suite of characters (Nelson and Platnick 1981; Nixon and Wheeler 1990; deQueiroz and Donoghue 1990). The most important advantage of this concept is that it is operationally, as well as theoretically tractable (Amato 1991; Cracraft 1991; Vogler and DeSalle in press). Criticisms of the PSC have centered around the typological nature of its definition and the notion that current fine levels of resolution offered by molecular techniques can provide characters that unite groups of organisms below the species level. While surveys of characters for identifying phylogenetic species can be objective, there may be problems with assumptions of what constitutes a population (Davis and Nixon 1992; Amato and Gatesy in press).

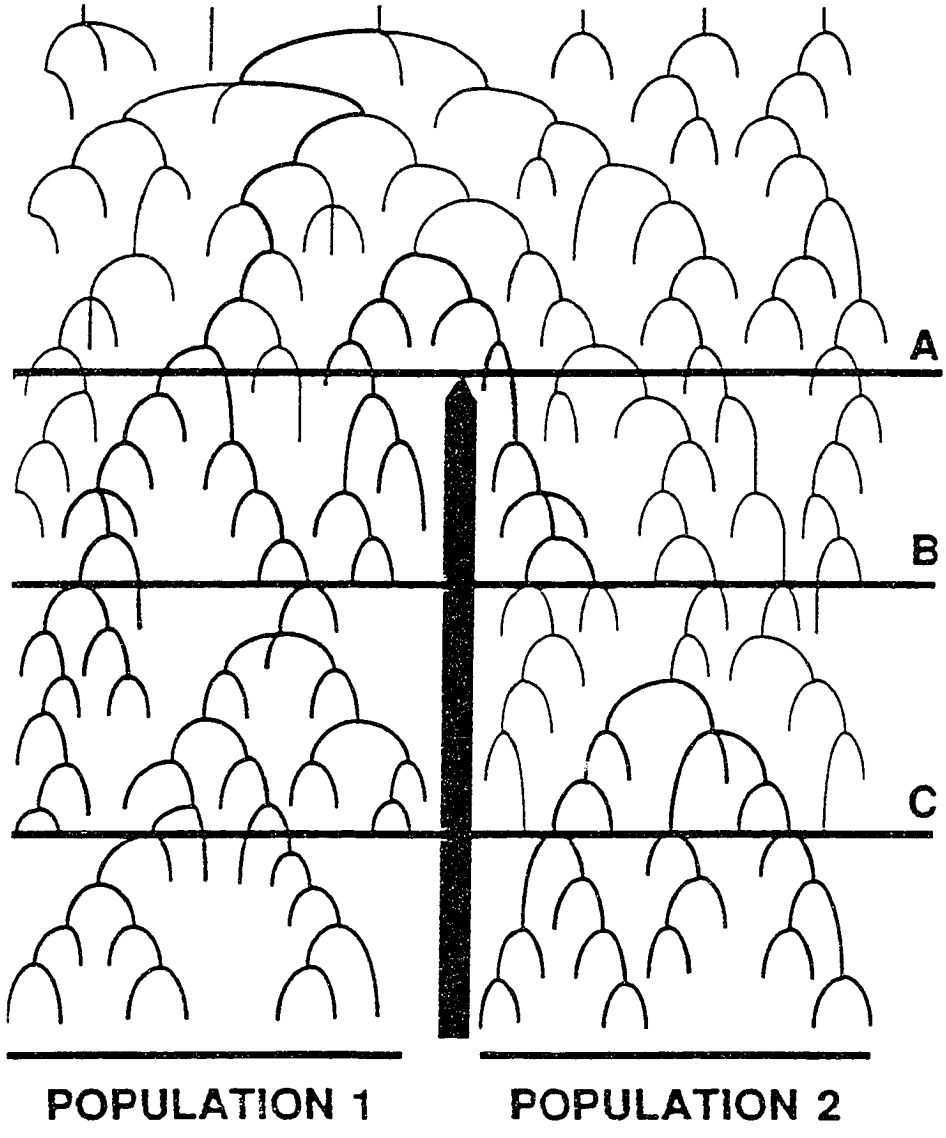
There are many other species concepts in addition to the BSC and PSC. Templeton's (1989) "cohesion" concept is a very broad, unifying concept in which biological and "recognition" (Patterson 1985) species are subsets. However, if we are to use species concepts to guide us in defining conservation units (especially if we choose to add evolutionary significance as in ESU), a phylogenetic, operational, and typological species concept implies evolutionary significance that can be observed by a distribution of characters. Further, it has been suggested that the best approach for identifying units of conservation is to follow a systematics model of character analysis (Amato, 1991; Cracraft, 1991;

Vogler and DeSalle, in press). The use of the phylogenetic species concept has the utility and philosophical logic appropriate for this task. Additionally, there is a large body of literature that uses this framework, along with a parsimony based character analysis, to identify patterns of phylogeny (Cracraft, 1983; Nelson and Platnick, 1981; Nixon and Wheeler, 1990).

Presented in this thesis are two major studies employing this phylogenetic systematics based framework for identifying conservation units. However, successfully identifying evolutionarily diverged taxa by character analysis requires a level of differentiation that approximates Mayr's "polytypic species". Below this level, a strict application of a PSC model may confuse gene trees and species trees (Powell 1992). If cladogenesis is viewed as an area occupied by populations (Fig. 1), rather than branching lines, there is a lag period in evolutionary time from the occurrence of a cladogenetic event and sufficient separation for successful character analysis.

In this way, the outlined framework employs PSC character based analysis at a level that largely overlaps with biological species. Avise and Ball (1992) proposed a biological taxonomy based on genealogical concordance as a bridge between the BSC and the PSC. The number of genes that should be used and how often they need to be concordant is not addressed and is clearly subjective. This study is based on the assumption that the region characterized should be one that varies between closely related species, but not within species. The choice of 12S and 16S ribosomal mitochondrial genes is supported by a large number of studies demonstrating this level of variation in vertebrates (Kocher et al. 1989; Gatesy et al. 1992; Vrana et al. in press). Comparisons of these studies provide insights into the strengths and limitations of this proposed framework. These studies also generated important primary data on the evolutionary history of these taxa.

FIGURE 1. A schematic of cladogenesis (modified from Avise and Ball 1990). At point A the splitting event has occurred but there has been insufficient time for the accumulation of diagnostic characters. Point B indicates greater separation, but it may still be difficult to distinguish gene trees and species trees. After point C, a sufficient time has elapsed in order to identify distinct conservation units.



CHAPTER 1

PCR ASSAYS OF VARIABLE NUCLEOTIDE SITES FOR IDENTIFICATION OF CONSERVATION UNITS IN CAIMAN CROCODILUS

SUMMARY

A number of authors have recently suggested that the best approach for identifying units of conservation is to follow a systematics model of character analysis (Amato 1991; Cracraft 1991; Vogler and DeSalle in press). This approach requires the use of an operational, typological evolutionary species concept. The use of the phylogenetic species concept has the utility and philosophical logic appropriate for this task. Additionally, there is a large body of literature that uses this framework, along with a parsimony based character analysis to identify patterns of phylogeny (Hennig 1966; Cracraft 1983; Nelson and Platnick 1981; Nixon and Wheeler 1990).

While advocating this approach, it is important to recognize that one of its limiting factors is sample size. It is proposed that by selective direct sequencing plus rapid sampling of variable target characters by polymerase chain reaction (PCR) assays of specific sites, sufficiently large numbers of individuals can be accurately, inexpensively, and quickly surveyed for diagnostic characters. This procedure is demonstrated by a survey of variable nucleotide sites in the *Caiman crocodilus* complex.

INTRODUCTION

Both conservation biologists and systematists attempt to characterize biological diversity. Conservation biologists have primarily been concerned with identifying "natural" units which form the basis of both *in situ* and *ex situ* conservation programs. This objective identification of conservation units has been an attempt to conserve naturally occurring biodiversity and avoid the potential problems of outbreeding. Classifying units at this level is problematic since it is near the interface of phylogeny and tokogeny (Hennig 1966; Vrana and Wheeler 1992). In an attempt to overcome this problem I have employed a species level, systematics based approach, rather than an evaluation of population level assessments (i.e. allele frequencies). It has been proposed that the use of overall similarity or genetic distances is ineffective for identifying taxonomic rank or units of conservation (Avise and Aquadro 1982; Vogler and DeSalle in press).

In 1985 a special meeting sponsored by the American Association of Zoological Parks and Aquariums (AAZPA) was held in Philadelphia to address the "subspecies dilemma" for endangered species. Subspecies designations have historically been applied to everything from undifferentiated local populations to biological species. At this meeting the term "evolutionarily significant unit" (ESU) was introduced to describe the "natural" unit that should be the focus of conservation efforts. Since this time it has been suggested that the ESU corresponds to the phylogenetic species (Cracraft 1991; Barrowclough and Flessness in press; Vogler and DeSalle in press) and that to objectively identify units of conservation requires a phylogenetic species and lower level systematics approach. Phylogenetic species are defined as "an irreducible cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent" (Cracraft 1989).

Operationally, phylogenetic species are diagnosably distinct populations (or groups of populations) where every individual shares the diagnostic character (or suite of characters). However, other lower level, systematics based approaches have been proposed. Vrana and Wheeler (1992) have proposed using only individuals as terminal taxa. While this approach is more assumption free than defining populations for a phylogenetic species analysis, it has less application to conservation where populations are the unit of management.

Systematic analyses at this level have been problematic due to a paucity of discrete morphological characters and incomplete geographic sampling. However, the increase in availability of molecular techniques has provided a new source of characters, especially DNA sequences, for use at lower level systematics studies. PCR technology has been the driving force in the greatly accelerated rate of data collection. While all types of characters (morphological, behavioral, karyotypic, and genetic) are useful and important, easily obtained DNA sequence data provides enormous numbers of genetic characters useful for a systematic approach to identifying conservation units. Along with this improved ability to generate large numbers of characters has been an explosion of ideas and algorithms for analyzing molecular character data for phylogenetic study (Swofford 1990; Felsenstein 1990; Farris 1989). The large number of available molecular phylogenetic studies has also provided useful information about character evolution for conservation. While identification of phylogenetic species rests only on demonstrating that a population has diagnostic characters (which are shared by all members of a population but are not found in other groups), the use of higher level phylogeny reconstruction may be important in identifying regions of DNA that are useful for characterization (Amato et al. 1993).

PCR has also allowed the amplification of target DNA sequences from

nontraditional biological samples. Before PCR, biochemical techniques required careful preparation of large quantities of fresh blood or organ tissues (for vertebrates) or entire organisms (for smaller invertebrates). This proved especially problematic for conservation research because samples were often needed from animals that were handled infrequently, existed only in small, isolated populations, and might only be handled by field researchers who had difficulty in obtaining and preserving biological samples. PCR advances allowed the use of such samples as hair, small skin biopsies, shed feathers, dried blood, museum specimens, and others (Amato et al. 1993; Walsh et al. 1991; Garza and Woodruff 1992). With the collection of materials made easier, as well as the generation of large numbers of molecular genetic characters, we are better prepared than ever to tackle questions concerning units of conservation.

While DNA characters offer a wealth of discrete character information for such lower level systematics problems (e.g. Thomas et al. 1990), the expense and time involved in direct sequencing has often resulted in the sampling of small numbers of specimens. In contrast many morphological characters can be scored rapidly and inexpensively. Although PCR technology has accelerated the rate of data collection, most DNA sequencing studies employ only a few exemplars that are used to represent a particular population or subspecies. The use of direct sequencing for broad character surveys is not generally cost effective given the number of phylogenetically relevant characters that are discovered per dollar and hour.

PCR based sampling methods

One option is to use polymerase chain reaction assays of specific sites to increase the sample size of a lower level systematic analysis. The simplest

approach to this assay is the identification of potentially informative (polymorphic) sites by direct sequencing of a subset of available samples. Sites can then be assayed by designing PCR primers that match the polymorphic sites at the 3' end of the primer. Sommer et al. (1992) describe various methods of assuring specificity of assaying for single base changes with any primer:template mismatch. This methodology has been termed PCR amplification of specific alleles (PASA), allele-specific amplification (ASA), allele-specific PCR (ASP), and amplification refractory mutation system (ARMS) (Sommer et al. 1992; Newton et al. 1989; Nichols et al. 1989; Okayama et al. 1989; Wu et al. 1989). In all cases amplification takes place if there is a perfect match between the primer with the correct sequence to complement the polymorphic site and no amplification if there is a mismatch at the variable site. Varying magnesium concentrations, enzyme concentrations, and cycling conditions allows for optimization of this assay (Sommer et al. 1992). Both primers can be used to provide positive and negative controls, thus excluding amplification problems from confusing the results. Additionally, any samples that were assayed as ambiguous can be completely sequenced.

Other techniques are designed to identify nucleotide sequence changes (down to the level of a single base insertion/deletion/substitution) without completely sequencing all samples. Strobeck and Polziehn (submitted 1993) used a variation of this technique that they term primer generated RFLPs (PG-RFLPs) to survey populations of North American bison (*Bison bison*). In this study, allele specific PCR primers were used that also contain restriction site sequences with the variable site as part of the restriction site. The purpose of this approach is to enhance the ability to discriminate between absence of PCR products, apparently due to problems of amplification specificity. In addition to PG-RFLPs, they surveyed for polymorphic sites by simple restriction enzyme

assays where direct sequencing had revealed a polymorphism that fell within a restriction site sequence. However, this does not unambiguously identify the specific change within the restriction site.

Lessa and Applebaum (1993) recently reviewed related techniques for identifying allelic variation. Single-strand conformational polymorphism (SSCP) (Orita et al. 1989) is based on the change in mobility of single-stranded DNA fragments that differ in base substitutions, insertions, and deletions. PCR products are denatured by heating in the presence of formamide and are separated on acrylamide gels. Strands migrating differentially can then be sequenced to identify base changes.

Two other techniques, denaturing gradient gel electrophoresis (DGGE) (Myers et al. 1986, 1989a, 1989b) and temperature gradient gel electrophoresis (TGGE) (Whartell et al. 1990) rely on the separation of denatured or "melted" double-stranded DNA. The physical properties of the DNA fragment when it reaches its melting point in a gradient gel allow for separation based on a single base change. However, direct sequencing is necessary to identify the position of the change. Carefully controlling conditions should allow for the inference of base change identity for identically migrating fragments.

Another related method for assessing allelic variation is coupled amplification and sequencing (CAS) (Ruano and Kidd 1991a, 1991b). This method allows for the simultaneous direct sequencing of the two complementary strands of a PCR product identified as a different allele by a DGGE gel.

In this study I use a variation of the ASA technique in order to increase the number of surveyed individuals for an assessment of conservation units.

The Caiman crocodilus complex

There are five currently recognized species of caiman- *Caiman latirostris*, *Caiman crocodilus*, *Melanosuchus niger*, *Paleosuchus trigonatus*, and *Paleosuchus palpebrosus*. The most wide spread species is *Caiman crocodilus* which ranges from Central America through northern South America including the Amazon and Orinoco river systems south and west to Brazil, Peru, Bolivia, Paraguay, and Argentina. Taxonomic designations for the "*Caiman crocodilus* complex" have historically been problematic (Brazaitis 1973; Frair and Behler 1983). At the present time the complex is most frequently described as three subspecies- *Caiman crocodilus crocodilus* (northern South America, Amazon and Orinoco rivers), *C. c. fuscus* (Central America and northern Columbia and Venezuela), and *C. c. yacare* (southern and western regions of Brazil bordering Bolivia, Argentina and Paraguay including Rio Paraguay and Rio Pilcomayo) (King and Rocca 1987; Medem 1983). At various times *C. c. yacare* has been designated a full species (Daudin 1802; Medem and Marx 1955, Carvalho 1955; Medem 1960; Medem 1983; King and Burk 1989). Full species designations were based on a variety of morphological analyses of distinctive skull morphology, color, pattern, and scalation correlated with geographic distribution. The number of taxonomic units has been further confused by a leather industry booklet naming two additional subspecies (Fuchs 1974). These designations were introduced into the literature by Wermuth and Mertens (1977) in spite of the fact that they were based on unreferenced, incomplete, tannery skins from unknown collectors and localities, and without deposited voucher specimens (Frair and Behler 1983). All of this complicates the fact that the subspecies have had different levels of protection under the Convention for International Trade in Endangered Species (CITES).

Our interest in applying a molecular systematics approach to identifying

phylogenetic species/conservation units for the *C. crocodilus* complex is in part a response to the confusion in the literature. Accurately identifying these units has important implications for designing *in situ* and *ex situ* conservation strategies for this group. Identifying numbers and distributions of caiman taxa will impact the commercial trade in skins and potentially provide important forensic tools for monitoring trade.

METHODS

Samples used in this study are a subset of samples collected by a number of independent field studies (Brazaitis et al. 1988; 1990). These field surveys sampled all major populations and important river systems in the range countries. A variety of protocols were employed to preserve blood samples including desiccating whole blood on sterile cotton surgical sponges with room temperature storage in sealed plastic bags and preserving whole blood in an equal volume of buffer containing 100 mM EDTA, 100 mM Tris, 2%SDS (RT buffer). All samples were obtained without harm to the study animals which were immediately released after blood samples, morphological measurements, and photographs were taken. These sampling procedures were easily carried out in the field since neither refrigeration nor special handling were required. Total genomic DNA was obtained from the dried blood by a method employing a chelating resin (Chelex 100, BioRad) optimized for forensics samples (Walsh et al. 1991). A standard phenol/chloroform DNA isolation procedure (Caccone et al. 1987) was used for the samples stored in RT buffer. All samples yielded microgram quantities of total genomic DNA.

A subsample of DNA sequence characters was used to identify nucleotide sites that vary within the *C. crocodilus* complex. Approximately 1000 nucleotides of mitochondrial (mt) DNA (fragments of 12S and 16S mt ribosomal

DNA and mt cytochrome b - Kocher et al. 1989; Irwin et al. 1991; Gatesy and Amato 1992) were sequenced from nine individuals of *C. crocodilus*. These samples include the three "subspecies", *C.c. crocodilus*, *C.c. yacare*, and *C.c. fuscus*, from across the broad range of the species complex.

An additional 64 samples were surveyed for four polymorphic sites by use of PCR assays. Reactions were carried out using approximately 250 ngs of template DNA, a magnesium concentration of 1.5 mM, primer concentration of 0.1 um, and 0.5 units of Taq polymerase in 50 ul reaction volumes. PCR was performed in a Perkin-Elmer Cetus DNA thermal cycler at 94^{o}C for 1 min., 52^{o}C for 1.5 min., and 72^{o}C for 2 min. for forty cycles. Two seventeen base long primers were constructed for each surveyed site with the most 3' base specific to the two alternate bases identified by direct sequencing. These primers were designed from previously sequenced *Caiman crocodilus*. Base specific primers were paired with universal vertebrate primers (Kocher et al. 1989) amplifying fragments of approximately 150 bases. Each sample was amplified with both primers providing a positive and negative control (Figure 2).

Samples of *C. crocodilus* were assigned to a subspecies according to morphological criteria and geographic position (Brazaitis et al. 1988; 1990). We then surveyed for character states, inferred from direct sequencing and PCR assays of specific site information, that were unique to each "subspecies" (Nixon and Davis 1993). If there were no fixed differences between subspecies, the division of *C. crocodilus* into several phylogenetic species/conservation units would not be warranted.

RESULTS

There were 22 variable nucleotide sites apparent from the alignment of the nine *C. crocodilus* sequences. Only two of these sites are homoplastic within this sample as determined by character tracing on the most parsimonious tree for this data. The six sites that were more extensively sampled by PCR assays of specific sites, showed consistency across the ranges of each "subspecies" (Table 1). Sites 1, 2, and 3 from the 12S mtDNA sequence display a unique state in *C.c. yacare.* Site 4 is unique to *C.c. crocodilus,* while sites 5 and 6 from the 16S mtDNA sequence are unique in *C.c. fuscus* (Table 1). These sites remain monomorphic within "subspecies" despite the increase in sample size from nine to 73 (Table 2; Figure 3).

DISCUSSION

A population aggregation analysis (PAA) (Davis and Nixon 1992) of the three haplotypes identified by PCR assays of specific sites showed them to be diagnostic characters identifying three phylogenetic species (Figure 4). These units correspond to the named subspecies of *Caiman crocodilus* as identified by Brazaitis et al. (1973). Additionally, specific sites were diagnostic for different populations (e.g. sites 5 and 6 are diagnostic for *fuscus*). This concordance of molecular data, morphological characters, and biogeographic distribution support a management plan with three conservation units corresponding to *crocodilus*, *yacare*, and *fuscus*.

The presence of three closely related, but diagnosably distinct units of the *Caiman crocodilus* complex suggest separation of the units in different refugia during glacial periods. Pleistocene vegetation distribution indicates the presence of broad savannas in central South America (Dixon 1979) could possibly have divided the ranges. Also, Simpson (1979) presents a view of the

major geomorphological units of the Andean Cordillera and the principal shields of South America that could explain periods of isolation.

However, crocodilians appear genetically conservative as demonstrated by their ability to hybridize in captivity. There are at least three examples of Asian and New World crocodiles producing viable offspring (Behler pers. com.). Certainly, members of the genus *Caiman* can hybridize (Brazaitis pers. com.). The persistence of diagnosable units today probably reflects a degree of geographical separation that is not obvious in a distribution map of caiman in South America. Actually, *C.c. fuscus* is confined to Central America and the northern parts of South America where river systems drain toward the Caribbean sea and Pacific ocean. *C. c. yacare* populations may effectively be separated from *C.c. crocodilus* populations by the flow of rivers away from the Rio Amazonas basin, as well as the highlands of eastern Brazil and the Andean foothills to the west in Peru. The tributaries and basin of the Amazon itself coincides with the range of *C.c. crocodilus*. Additional higher level studies of caiman are in progress and are presented in chapter 2. By rooting the network of haplotypes within *C. crocodilus*, relationships between phylogenetic species may be apparent, and have relevance to biogeographical hypotheses for South America.

The use of PCR assays of specific sites allows us to increase our survey of attributes and test hypotheses about phylogenetic species. Only continued research using this approach, followed by decision-making and action by managers will ultimately demonstrate its usefulness. Systematics provides us with an important framework to aid in identifying conservation units while population genetics provides models for managing these units. It is the managers in zoological parks, governments, and international conservation organizations that must use these results to implement the important

management actions.

TABLE 1. 676 bases of 12S and 16S mitochondrial DNA. Dashes signify base identitiy with reference sequence. (C.c.c.= *Caiman crocodilus crocodilus*, C.c.y.= *C.c. yacare*, C.c.f.= *C.c. fuscus*)

```
C.c.c.   gacttgacag tacttcaaat ccacctagag gagcctgtcc
C.c.y.   ---------- ---------- ---------- ----------
C.c.f.   ---------- ---------- ---------- ----------

C.c.c.   tataatcgaa agtacacgat tcacctaacc acccttagtt
C.c.y.   ---------- ---------- ---------- ----------
C.c.f.   ---------- ---------- ---------- ----------

C.c.c.   atccagttg cataccgccg tcgcaagct tgtctcgct
C.c.c.   ------c-- t--------- --------- --------
C.c.f.   --------- ---------- --------- --------

C.c.c.   gagagaaaca aaatgagca caatagcccc ccgcta
C.c.y.   ---------- --------- ---------- ------
C.c.f.   ---------- --------- ---------- ------

C.c.c.   aaacgtcag gtcaacgcgc agctaatggg gtgggaagga
C.c.y.   --------- ---------- ---------- ----------
C.c.f.   --------- ---------- ---------- ----------

C.c.c.   tgtgctacat tttctaaca catagaaata cgtgacggaa
C.c.y.   ---------- --------- ---------- ----------
C.c.f.   ---------- --------- ---------- ----------

C.c.c.   gcgtaatcac ttgttctcca aataaggact agtatggaa
C.c.y.   ---------- ---------- ---------- ---------
C.c.f.   ---------- ---------- ---------- ---------

C.c.c.   cggttaaacg agaatctaac tgtctcctgc aagcagccaa
C.c.y.   ---------- ---------- ---------- ----------
C.c.f.   ---------- ---------- ---------- ----------

C.c.c.   tgaaattgat cttcctgttg caaaagcagg aatgacatc
C.c.y.   ---------- ---------- ---------- ---a-----
C.c.f.   ---------- ---------- ---------- ---------

C.c.c.   accagacgag aagaccctgt gaaacttaaa cccactaagt
C.c.y.   ---------- ---------- ---------- ----------
C.c.f.   ---------- ---------- ---------- ----------

C.c.c.   taaaccaaca acattaactg caacacccac gactgttgaa
C.c.y.   ---------- ---------- ---------- ----------
C.c.f.   ---------- ---------- ---------- ----------

C.c.c.   acctgactta acgttttcgg ttggggtgac cctaaaacaa
C.c.y.   ---------- ---------- ---------- -t--------
C.c.f.   ---------- ---------- ---------- -t--------
```

```
C.c.c.    agaaaaactt ttaagacaat tataactaag accaaattat
C.c.v.    ---------- ---------- ---------- ----------
C.c.f.    ---------- ---------- ---------- ----------

C.c.c.    taaccaagac ccactcctca aagtaccttg aatgtaatta
C.c.v.    ---------- ---------- ---------- ----------
C.c.f.    ---------- ---------- ---------- ---a------

C.c.c.    gatccgacaa cgtcgatcaa cggacaaagc tactccaggg
C.c.v.    ---------- ---------- ---------- ----------
C.c.f.    ---------- ---------- ---------- ----------

C.c.c.    ataacagcgc aatccccctc aagagcctca
C.c.v.    ---------- ---------- ----------
C.c.c.    ---------- ---------- -------c--
```

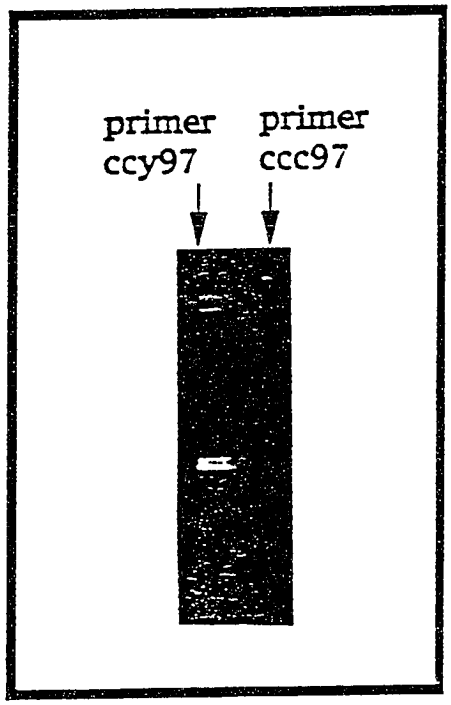
TABLE 2. *Caiman crocodilus* samples included in this study.

IDENTIFICATION NUMBER	COUNTRY
3309	Bolivia
3303	Bolivia
3301	Bolivia
3295	Bolivia
3287	Bolivia
3201	Bolivia
A40159	Bolivia
A40152	Bolivia
A40150	Bolivia
A40137	Bolivia
A40135	Bolivia
A40125	Bolivia
A40120	Bolivia
A40111	Bolivia
A40065	Bolivia
A40050	Bolivia
A40004	Bolivia
C17	Paraguay
C18	Paraguay
C22	Paraguay
C15	Paraguay
C7	Paraguay
C5	Paraguay
C9	Paraguay
C31	Paraguay
C28	Paraguay
C147	Paraguay
C110	Paraguay
C82	Paraguay
C109	Paraguay
C98	Paraguay
C84	Paraguay
C107	Paraguay
C78	Paraguay
C71	Paraguay
C73	Paraguay
C77	Paraguay
C45	Paraguay
C51	Paraguay
C64	Paraguay
C67	Paraguay
C133	Paraguay
C128	Paraguay

46194	Brazil
46191	Brazil
46171	Brazil
46063	Brazil
46140	Brazil
46131	Brazil
46134	Brazil
46027	Brazil
46070	Brazil
46247	Brazil
46210	Brazil
46500	Brazil
46507	Brazil
46506	Brazil
46281	Brazil
46249	Brazil
46269	Brazil
367	Venezuela
54093	Panama
54090	Panama
70469	Costa Rica
70471	Costa Rica

FIGURE 2. 12S ribosomal mitochondrial DNA fragment of *Caiman crocodilus yacare* amplified with base specific primers (ccy 97 and ccc 97).

PCR Amplified Fragment from Caiman crocodilus Sample #C107 Collected in Paraguay



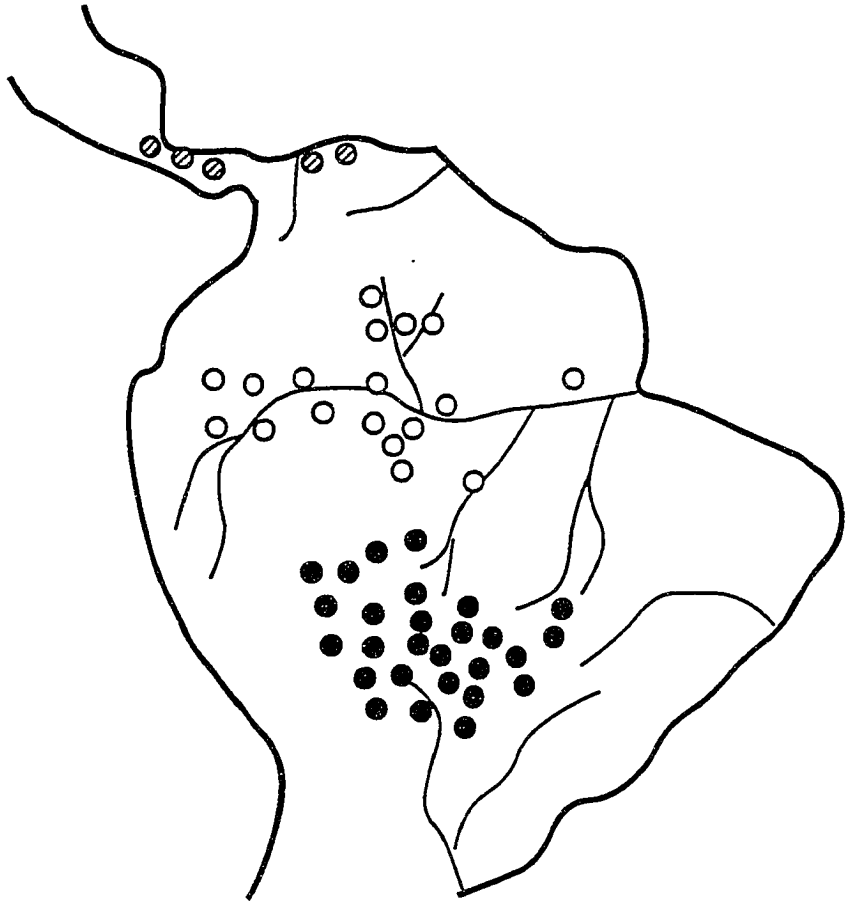
Base Specific Primer For PCR

ccc 97 '5 ACC CTT AGT TAT CCA GT 3'

ccy 97 '5 ACC CTT AGT TAT CCA GC 3'

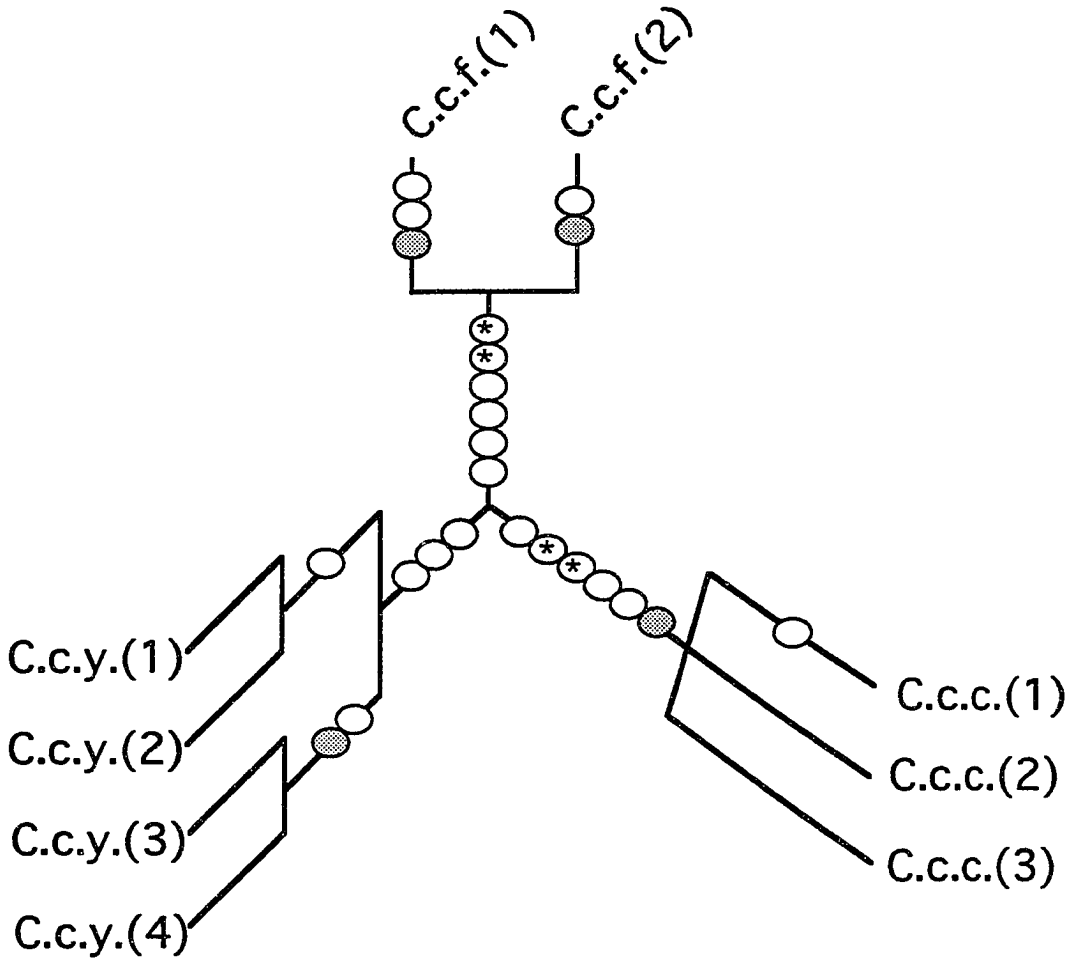
FIGURE 3. Sample localities for *Caiman crocodilus* study employing direct sequencing and PCR assays of specific sites.

SAMPLE LOCALITIES



○ C. c. crocodilus N=18
⊘ C. c. fuscus N = 5
● C. c. yacare N = 50

FIGURE 4. The most parsimonious network (24 steps, consistency index - 0.889) for the *Caiman crocodilus* mtDNA sequences sampled here (Swofford, 1990). Homoplastic changes are shown in gray. Sites that changed once are represented by white circles. C.c.f. = *C.c. fuscus*, C.c.c. = *C.c. crocodilus*, C.c.y = *C.c. yacare*.



CHAPTER 2

ISSUES IN HIGHER LEVEL RELATIONSHIPS OF EXTANT CROCODILIANS

SUMMARY

Two issues in higher level relationships of crocodilians were addressed to assess the usefulness of 12S and 16S mitochondrial sequences for conservation units studies. I first address the phylogenetic relationships of longirostrine crocodilians. This relationship [between the gharial (*Gavialis gangeticus*) and the false gharial (*Tomistoma schlegelii*)] has been the most problematic in studies on higher level relationships of crocodilians. In this study the placement of *Tomistoma* as the sister taxon to *Gavialis* is supported.

Second, higher level relationships of caiman are reviewed to assess monophyly of the genus *Caiman*. Within the monophyletic alligatorid clade (*Alligator mississippiensis*, *A. sinensis*, *Paleosuchus trigonatus*, *P. palpebrosus*, *Melanosuchus niger*, *Caiman crocodilus*, and *C. latirostris*) morphological and biochemical studies fail to resolve this relationship. The addition of DNA sequence character data adds little resolution to this relationship.

INTRODUCTION

Longirostrine crocodilians

Numerous workers have investigated the evolutionary history of the Crocodilia. Nevertheless, the phylogenetic relationships of the longirostrine species, *Tomistoma schlegelii* (the false gharial) and *Gavialis gangeticus* (the gharial) remain controversial. Two favored hypothesis have emerged from past studies (Figure 5). The first, with *Gavialis* as the sister taxon to all other extant

crocodilaians, is the traditional arrangement based on cladistic and evolutionary systematic analyses of morphology (Mook 1934; Norell 1989; Tarsitano et al. 1989). The genus *Tomistoma* is placed as part of a monophyletic group that includes the genera *Crocodylus* and *Osteolaemus*. The alternate hypothesis is primarily based on phenetic biochemical evidence (Densmore 1983; Densmore and Dessauer 1984; Densmore and Owen 1989). Allozyme studies, immunodiffusion experiments, protein fingerprint comparisons (Densmore 1983), and mitochondrial DNA restriction fragment length polymorphism data (Densmore and Owen 1989) revealed that *Gavialis* and *Tomistoma* were most genetically similar, with *Crocodylus* and *Osteolaemus* grouped, and Alligatorinae (*Alligator*, *Caiman*, *Melanosuchus*, and *Paleosuchus*) more distantly related. The molecular and gross anatomical data support radically different trees. Unfortunately, it is not clear whether this discrepancy is due to the methods of data analysis (cladistic versus phenetic) or the quality of data (Densmore 1983; Norell 1989; Tarsitano et al. 1989).

Caimanoids

Caimans comprise a group of alligatorid crocodilians confined to the Neotropics. Morphological data (Norell 1988) and biochemical data (Densmore 1983) generally agree on the relationships of the five extant species. Caimans are monophyletic with the genus *Paleosuchus* (*P. trigonatus* and *P. palpebrosus*) as the sister group to the clade containing *Caiman crocodilus*, *C. latirostris*, and *Melanosuchus niger*. This latter clade, Jacarea (Gray 1844) has been unresolved. Morphological (Norell 1988) and biochemical (Densmore 1983) studies weakly support a paraphyletic *Caiman*.

MATERIALS AND METHODS

In the initial study, 259 bases of 12S mitochondrial DNA were sequenced for *T. shlegelii*, *G. gangeticus*, a crocodile (*Crocodylus rhombifer*) and an alligatorine (*Caiman crocodilus*). Later, all species of caiman and the two species of alligator were sequenced for 676 bases of 12S and 16S mitochondrial DNA. Blood samples were drawn from captive animals, and DNA extracted (Caccone et al. 1987). The 12S and 16S fragments were PCR amplified from each DNA sample using unbalanced proportions of universal vertebrate primers (Kocher et al. 1989). The single stranded product was sequenced directly using the dideoxy protocol (Gyllensten and Erlich 1988). Sequences were initially aligned by eye. In a later study (Gatesy et al. 1993), these sequences and additional crocodilian sequences were aligned with MALIGN (Wheeler and Gladstein 1991) to explore alignment ambiguities and the effect of excluding such areas for systematic study. Maximum parsimony cladograms were constructed using PAUP 3.1 (Swofford 1991).

RESULTS

Longirostrine crocodilians

The orthologous 12S sequences, varying in length from 240 to 256 base pairs were obtained (Table 3). Inspection of 86 variable nucleotide positions revealed that *Gavialis* and *Tomistoma* share 22 unique nucleotide characters, whereas *Tomistoma* and *Crocodylus* share only four. Not surprisingly, the sequence of *Tomistoma* and *Gavialis* are 94% similar whereas the sequences of *Tomistoma* and *Crocodylus* are only 84% similar (Table 4). These data are consistent with all other molecular studies. Additional 16S sequences support the same relationship (Table 5). All maximum parsimony trees derived from

various alignments and weighting schemes support a sister taxa relationship between *Tomistoma* and *Gavialis* (Gatesy et al. 1993; Gatesy and Amato unpublished).

Caimanoids

The mitochondrial sequences, aligned by eye (Table 5) and analyzed using maximum parsimony with unweighted unordered characters, offer weak support for *Caiman* monophyly. 12S data aligned with MALIGN (Wheeler and Gladstein 1991) introduces an alignment ambiguity that results in a paraphyletic Caiman when analyzed by maximum parsimony (Gatesy et al. 1993) (Figure 6). While higher level caiman relationships are unresolved, *Caiman crocodilus* subspecies are diagnosably distinct (Chapter 1, this volume).

DISCUSSION

Longirostrine cocodilians

Both mitochondrial DNA sequence similarity, and a cladistic analysis of 12S and 16S (Gatesy and Amato, in prep) sequences support the monophyly of the longirostrine crocodilians. Apparently, the discrepancy in trees between the gross morphological data and biochemical data is not due to methods of data analysis, but rather to the quality of data. Ongoing higher level studies (Gatesy and Amato unpublished) using Aves as an outgroup will provide estimates of ancestral states, offering greater resolution for higher level crocodilian groupings.

Caimanoids

Additional ongoing studies (Gatesy and Amato unpublished) are necessary to resolve the phylogenetic relationship of *Caiman* and *Melanosuchus*. Also, other data sets may provide important evidence. Norell (1988) noted the largely vicariant nature of the caimanoid's distribution.

Species and subspecies ranges roughly correspond to recognized avian areas of endemism (Cracraft 1988). Concordant results from future systematic and biogeographical studies of extant and fossil taxa may further resolve this problematic relationship.

TABLE 3. 12S mitochondrial DNA sequences for gharial *(Gavialis gangeticus)*, false gharial (*Tomistoma schlegelii*), Caiman (*Caiman crocodilus*), and Cuban crocodile (*Crocodylus rhombifer*).

	10	20	30	40	50
C. crocodilus	gacttgacag	tacttcaaat	ccacctagag	gagcctgtcc	tataatcgaa
C. rhombifer	........g.	..t...g..c	....	....	c
G. gangeticus	........g.	......gc.c	....	....	c
T. schlegelii	........g.	......gc.c	....	....	c
	60	70	80	90	100
C. crocodilus	agtacacgat	tcacctaacc	accettagtt	atc-------	---cagtttg
C. rhombifer	.a........	c....c....	...t..t.cc	c.aag-----	---...cc..
G. gangeticus	.....t....	at....t...	.a.t..t.cc	t.aaacgtc-	taa...cc..
T. schlegelii	.ac..t....	a....ct...	.a.t..t.cc	t.aaactaca	taa...cc..
	110	120	130	140	150
C. crocodilus	catacegccg	tcgcaagctt	gtctcgctga	gagaaacaaa	a-tgagcaca
C. rhombifer	t.........	....	ag.c.-a...	.g..c.ag..	c-ct......
G. gangeticus	t.........	......a..a	.c.c.-....	.g..c.a.c.	.t.t..t...
T. schlegelii	t.........	......a..a	ac.c.-....	.g..cga.c.	gt.a..tg..
	160	170	180	190	200
C. crocodilus	atagcccccc	-----gctaa	aacgtcaggt	caacgcgcag	ctaatggggt
C. rhombifer	...a.t.a.t	tctga....g	t.........	...g.t....	.c...aa.t.
G. gangeticus	.....ttatt	t--ga.....	t.........	...g.t....	.c....a.t.
T. schlegelii	.c...t.att	t--ga.....	t.........	...g.t....	.c...aa.t.
	210	220	230	240	250
C. crocodilus	gggaaggatg	tgctacattt	tctaacacat	agaaatacgt	gacggaacgt
C. rhombifer	..t.ga....	g.........	....cac...	......tg..	c.....ga.g
G. gangeticus	..a.ga....	g.........	....c.....	.......t..	c.....ga.c
T. schlegelii	..a.ga....	g.........	....c.t...	.......t..	c.....ga.c
	259				
C. crocodilus	cccgtgaaa				
C. rhombifer	..t......				
G. gangeticus	..t......				
T. schlegelii	..t......				

TABLE 4. Similarity of longirostrine 12S DNA sequences (*from* Gatesy and Amato, 1992).

TABLE 2. PERCENTAGE NUCLEOTIDE SEQUENCE SIMILARITY (NUMBER OF BASE POSITIONS SHARED/ TOTAL NUMBER OF COMPARABLE BASE POSITIONS × 100) OF THE FOUR SAMPLED CROCODILIAN SPECIES (*Caiman crocodilus, Crocodylus rhombifer, Tomistoma schlegelii,* AND *Gavialis gangeticus*) FOR APPROXIMATELY 250 BASE PAIRS OF 12S MITOCHONDRIAL DNA.

	Caiman crocodilus	*Crocodylus rhombifer*	*Tomistoma schlegelii*
Crocodylus rhombifer	75	X	X
Tomistoma schlegelii	71	84	X
Gavialis gangeticus	75	84	94

TABLE 5. Crocodilian 12S and 16S sequences. (C.C. = *Caiman crocodilus*, C.L. = *Caiman latirostris*, A.M. = *Alligator mississippiensis*, A.S. = *Alligator sinensis*, P.P. = *Paleosuchus palpebrosus*, P.T. = *Paleosuchus trigonatus*, M.N. = *Melanosuchus niger*, C.R. = *Crocodylus rhombifer*, G.G. = *Gavialis gangeticus*, T.S. = *Tomistoma schlegelii*)

12S

C.C.
GACTTGACAG TACTTCAAAT CCACCTAGAG GAGCCTGTCC TATAATCGAA AGTACACGAT
TCACCTAACC ACCCTTAGTT ATCCAGTTTG CATACCGCCG TCGCAAGCTT GTCTCGCTGA
GAGAAACAAA ATGAGCACAA TAGCCCCCCG CTAAAACGTC AGGTCAACGC GCAGCTAATG
GGGTGGGAAG GATGTGCTAC ATTTTCTAAC ACATAGAAAT ACGTGACGGA ACGTCCCGTG
AAA

C.L.
GACTTGACAG CACTTCAAAT CCACCTAGAG GAGCCTGTCC TATAATCGAA AGTACACGAT
TCACCTAACC ACCCCTAGTC TACCCAGTTT GTATACCGCC GTCGCAAGCC TGTCTCGCTG
AGAGAAACAA AATGGGCACA ACAGCCTCCT AGGCTAAAAC GTCAGGTCAA CGTGCAACTA
ATGGGGTGGG AAGGATGTGC TACATTTTCT AACACATAGA AATACGTGAC GGAACGTCCC
GTGAAA

A.M.
GACTTGACGG CACTTTAAAC CCCCCTAGAG GAGCCTGTCC TATAATCGAC AGTACACGTT
ACACCCGACC ACCTTTAGCC TACTCAGTCT GTATACCGCC GTCGCAAGCC CGTCCCATTT
GAGGGAAACA AAACGCGCGC AACAGCTCAA CCGAGCTAAC ACGTCAGGTC AAGGTGCAGC
CAACAAGGTG GAAGAGATGG GCTACATTTT CTCAACATGT AGAAATATTC AACGGAGAGC
CCTATGAAA

A.S.
GACTTGACGG CGCTTCGAAC CCACCTAGAG GAGCCTGTCC TATAATCGAC GGTACACGAT
TCACCCGACC ACCTCTAGCC CCTCAGCCTG TATACCGCCG TCGCCAAGCC CGTCCCCCTG
AGGGAGACAA AACGAGCACA ATAGCCTCCC AGGCTAGCAC GTCAGGTCAA GGTGCAGCCA
ATGAGGTGGA AGAGATGAGC TACATTTTCT AACACATAGA AAATGCAACG GAGAGCCCTG
TGAAA

P.P.
GACTTGACGG TACTTCGAAC CCACCTAGAG GAGCCTGTCC TATAATTGAA GATACACGAT
TCACCTAACC CCTCCTTGCC TCTCAGTCTG TATACCGCCG TCGCCAAACT CGTCCCACTG
AGGGAAACAA AACAAGTGCA ACAGCCTCCC AGGCTAATAC GTCAGGTCAA GGTGCAGCTA
ATGGAGCGGA AGAGATGTGC TACATTTTCT AAAATAGAAA TACGTAACAG AACGCCCTCT
GAAA

P.T.
GACTTGACGG TACTTTAAAC CCACCTAGAG GAGCCTGTCC TATAATTGAA GATACACGAT
CCACCTAACC CCTCCTAGCC TCTCAGTCTG TATACCGCCG TCGCAAACTT GTCCCACTGA
AGGAAACAAA ACGAGTACAA CAGCCTCCCA GGCTAATACG TCAGGTCAAG GTGCAACTAA
CGGAGCGGAA GAGATGTGCT ACATTTTCTA AAATAGAAAT ACGTAACAGA ACGCCCTATG
AAA

M.N.
GACTTGACAG CACTTCAAAA CCACCTAGAG GAGCCTGTCC TATAATCGAA AGTACACGAT
CCACCTGACC ACCCCTGGCC CTCCAGTCTG TATACCGCCG TCGCAAGCTT GTCTCGCTGA
GAGAAACAAA ATAAGCACAA CAGCCTCCCA GGCTAAAACG TCAGGTCAAC GTGCAGCCAA
TGGGGTGGGA AGGATGTGCT ACATTTTCTA ACACATAGAA ATAGGTAACG GAGCGTCCCA
TGAAA

C.R.
GACTTGACGG TATTTCGAAC CCACCTAGAG GAGCCTGTCC TATAATCGAC AATACACGAT
CCACCCAACC ACCTTTTGCC CTAAGCAGCC TGTATACCGC CGTCGCAAGC TTAGCCCATG
AGGGACAAGA ACCTAGCACA ATAACTCACT TCTGAGCTAG TACGTCAGGT CAAGGTGCAG
CCAATAAGTT GGTAGAGATG GGCTACATTT TCTACACCAT AGAAATTGGT CACGGAGAGG
CCTGTGAAA

G.G.
GACTTGACGG TACTTCGCAC CCACCTAGAG GAGCCTGTCC TATAATCGAC AGTACTCGAT
ATACCTTACC AACTTTTGCC TTAAACGTCT AACAGCCTGT ATACCGCCGT CGCAAACTAG
CCCCCTGAGG GACAAACAAT TTAGTACAAT AGCTTATTTG AGCTAATACG TCAGGTCAAG
GTGCAGCCAA TGAGTTGGAA GAGATGGGCT ACATTTTCTA CCACATAGAA ATATGTCACG
GAGAGCCCTG TGAAA

T.S.
GACTTGACGG TACTTCGCAC CCACCTAGAG GAGCCTGTCC TATAATCGAA AACACTCGAT
ACACCCTACC AACTTTTGCC TTAAACTACA TAACAGCCTG TATACCGCCG TCGCAAACTA
ACCCCCTGAG GGACGAACAG TTAAGTGCAA CAGCTCATTT GAGCTAATAC GTCAGGTCAA
GGTGCAGCCA ATAAGTTGGA AGAGATGGGC TACATTTTCT ACCTCATAGA AATATGTCAC
GGAGAGCCCT GTGAAA

16S

C.C.
GCGTAATCAC TTGTTCTCCA AATAAGGACT AGTATGAACG GTTAAACGAG AATCTAACTG
TCTCCTGCAA GCAGCCAATG AAATTGATCT TCCTGTTGCA AAAGCAGGAA TGACATCACC
AGACGAGAAG ACCCTGTGAA ACTTAAACCC ACTAAGTTAA ACCAACAACA TTAACTGCAA
CACCCACGAC TGTTGAAACC TGACTTAACG TTTTCGGTTG GGGTGACCCT AAAACAAAGA
AAAACTTTTA AGACAATTAT AACTAAGACC AAATTATTAA CCAAGACCCA CTCCTCAAAG
TACTTGAATG TAATTAGATC CGACAACGTC GATCAACGGA CAAAGCTACT CCAGGGATAA
CAGCGCAATC CCCCTCAAGA GCTCATATC

C.L.
GCGTAATCAC TTGTTCTCTA AATAAGGACC AGTATGAATG GTTAAACGAG AATCCATCTG
TCTCTTGCAG GCAGCCAATG AAATTGATCT CCCTGTGCAA AAGCAGGGAT TCATACATTA
GACGAGAAGA CCCTGTGAAA CTTTAAACCC CTAGGCCACA ACAAATGTAA CCTAAACCCA
CATAGGCCCA CTATCATTAG ACCCCTTGAC CTAGTGTTTT CGGTTGGGGC GACCCCAAAA
TAAAAAAAAC TTTCCCGGAA AACAGTAACA TGACTACTAC TAACTAAGAC CTACACCCCA
AAGTGCTTAA ATGTAATCAG ATCCGCAAAG CCGATCTATG AACCAAGCTA CTCCAGGGAT
AACAGCGCAA TCCCCCTCAA GAGCCCTTAT C

A.M.
GCGTAATCAC TTGTTCTCCA AATAAGGACC AGTATGAACG GCTAAACGAG AATCTAACTG
TCTCCTGCAA ACAGCCAATG AAATTGATCT TCCTGTGCAA AAGCAGGAAT AACATCACCA
GACGAAAAGA CCCTGTGAAA CTTAAACCAC CTAAGTTAAA CCAACAATAC AACTGTAACC
CCTACAACCG TTTACACCTA ACTTAGCGTT TTCGGTTGGG GTGACCTTAA AACAAAAAAA
ACTTTTAAGA CAACTATAAC TAAGACCAAA TTATTAACCA AGACCCACAC CTCAAGGTAC
TTAAATGTAA TTAGATCCGA CAACGTCGAT CAACGAACAA AGCTACTCCA GGGATAACAG
CGCAATCCCC CTCAAGAGCC CATATC

A.S.
GCGTAATCAC TTGTTCTTTA AATAAGGACC AGTATGAACG GCTAAACGAG AATCTAACTG
TCTCCTGCAA GCAACCAATG AAATTGATCT CCCTGTGCAA AAGCAGGAAT GACCCCACCA
GACGAGAAGA CCCTGTGAAA CTTTAACCGA CTAAGTCACA CACTAGGAAC AACACAACCC
ACAACCACTT AAACCCATGA CTTAGCGCTT TTGGTTGGGG TGACCCTAAA ACAAAAAAAA
ACTTTTAAGA CAATCATAAC AAAATTAGAC TATTAACTAA GACCCACACC TCAAAGTACT
TAACTGTAAT TAGATCCGAC AATGTCGATC CACGAACTAA GCTACTCCAG GGATAACAGC
GCAATCCCCT TCAAGAGCCC CTATC

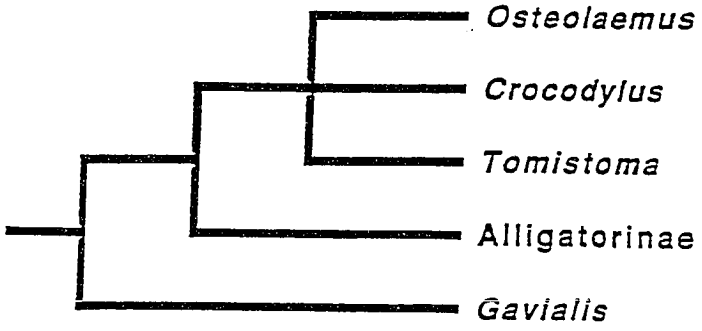
M.N.
GCGTAANCAC TTGTTCTCTA AATAAGGACC GGTATGAATG GTCAAACGAG AGTCTAACTG
TCTCCTGCAA GCAGCCAATG AAATTGATCT TCCTGTGCAA AAGCAGGAAT AGCCCCACCA
GACGAGAAGA CCCTGTGAAA CTTTAATCGG CTAAGTCATA CACACAACTA ATAATCACCC
ATAATTACCT GGACCGTGAC TTAGCGTTTT TGGTTGGGGT GACCTTGAAA CAAAGAAAAA
CTTTTAAGAA AGCTATAACA AAGTAGCCAG TATCCACTGA GACCCCCACA CCTCAAAGTA
CTTAAATGTA ATTAGATCCG ACAACGTCGA TCCATGAACC AAGCTACTCC AGGGATAACA
GCGCAATCCC CTTCAAGAGC CCCTATC

C.R.
GCGTAATCAT TTGTTCTTTA AATAAGGACC AGTATGAAGG GCTAAACGAG GTCTTAACTG
TCTCCTGTAG GTAATCTATG AAATTAGTAT TCCCGTGCAA AAACGAGAAT GTGAACATAA
GACGAGAAGA CCCTGTGGAA CTTTAAAATC ACGACCACCT TACAACCTTA CAAGCCCCAC
TGGGTCCACC CACACATAAA CCCCTGGTCG ACATTTTTCG GTTGGGGCGA CCTTGGAGAA
AAAAAAATCC TCCAAACCCA CAGACCACAA CTCTTCACTA AGACCAACTC CTCAAAGTAC
CAACAGTAAC CAGACCCAAT ATAATTGAGC AATGGACCAA GCTACCCCAG GGATAACAGC
GCAATCTCCT CCAAGAGCCC ATATC

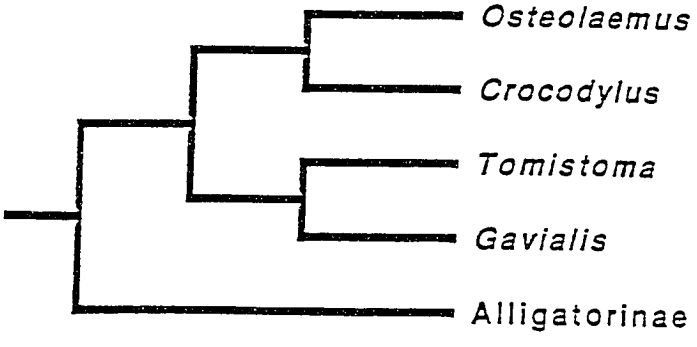
G.G.
GCGTAATCAC TTGTTCTTTA AATAAGGACC AGTATGAAAG GCTAAACGAG GTCTCAACTG
TCTCTTGCAG ATAATCAATG AAATTGATCT TCCTGTGCAA AAGCAGGAAT ATAAACATAA
GACGAGAAGA CCCTGTGGAA CTTAAAAACC AAGGATCAAT GCACTCCCAC CCAAACCTAC
TAAGGCTCAC CACTCCTGCA ACCAACTGAT CCTCGTTTTT TGGTTGGGGC GACCTTGGAG
AAAAAAGAAT CCTCCAAAAA CAAGACCACA ACTCTTAACT AAGAGCCACT CCTCAAAGTG
CCAACAGCGA CCAGACCCAA TATAATTGAT TAATGGACCA AGCTACCCCA GGGATAACAG
CGCAATCTCC TTCAAGAGCC CCTATC

T.S.
GCGTAATCAC TTGTTCTTTA AATGGGGACT AGTATGAAAG GCTAAACGAG AGTCCATCTG
TCTCTTGCAG GCAGCCAATG AAATTGATCT CCCTGTGCAA AAGCAGGGAT TTACACATCA
GACGAGAAGA CCCTGTGAAA CTTTAAACCC CCTAGGCCAC AACAAATGTA ATCCTTTTCC
CCAACCCGGG GCCAACTACC ATTAAAACTA TTGACCTAGT GTTTTTGGTT GGGGCGACCC
CAAAATAAAA AAAACTTTCC TGGAAAACAG TAACATAACT ACTACTAACT AAGACCTACA
CCCCAAAGTG CTTAAATGTA ATCAGATCCG GCATGCACCG ATCTATGGAC CAAGCTACTC
CAGGGATAAC AGCGCAATCC CCCTCAAGAG CCCCTATC

FIGURE 5. Alternative hypotheses on the relationship of logirostrine crocodilians based on morphological characters (A), and biochemical data (B) (from Gatesy and Amato, 1992).

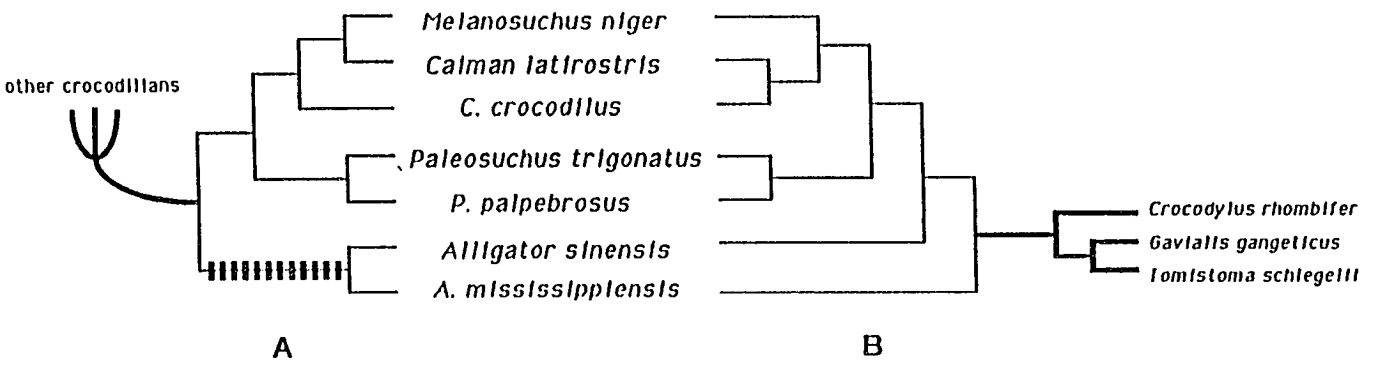


A



B

FIGURE 6. Maximum parsimony tree for higher level crocodilian relationships (from Gatesy et al., 1993).



CHAPTER 3

THE PHYLOGENETIC RELATIONSHIPS OF EXTANT RHINOCEROSES

SUMMARY

A higher level analysis of extant species of rhinoceros is presented. Maximum parsimony cladograms were derived by analysis of 12S and 16S mitochondrial genes both separately and combined. Five equally parsimonious alignments derived by use of different gap costs were explored and used to weight sites based on ambiguity of homology (elision) (Wheeler et al. in press). Additionally, a subset of Groves (1989) morphological characters were added to the molecular characters for a total evidence analysis.

This study supports a sister taxa relationship between the two Asian genera. It is also demonstrated that enough molecular variation exists in the investigated regions of mitochondrial DNA to support the utility of this area for a species/units-of-conservation study employing the previously (Introduction, this volume) described paradigm.

INTRODUCTION

There are five extant species of rhinoceros. The black rhinoceros (*Diceros bicornis*) and the white rhinoceros (*Ceratotherium simum*) are found in Africa; while the Indian (*Rhinoceros unicornis*), Javan (*Rhinoceros sondaicus*), and Sumatran (*Dicerorhinus sumatrensis*) rhinoceroses have an Asian distribution. These five species are all that remain of an extremely diverse group of perissodactyls that arose about 50 million years ago (Prothero et al. 1989). Approximately 4 million years ago, Rhinocerotoids disappeared from

North America and began to decline in the Old World. However, a number of unique species did persist until the end of the last glacial period, including the specialized woolly rhinoceros (*Coelodonta*) and the enormous *Elasmotherium*, which had a single giant horn on its forehead.

The relationships of the five extant species have recently been investigated by a number of authors. Guerin (1982), Groves (1983), and Prothero et al. (1986) have analyzed morphological characters and Ryder, Benviniste and George (unpublished) and Amato et al. (1993) have used molecular techniques. Two controversies have arisen from this research. Guerin's (1982) analysis supports a sister taxa relationship between Dicerorhinus and the African species while the other studies support a sister taxa relationship between the two Asian genera (Figure 7). The other controversy concerns the dating of lineage splitting suggested by the DNA-DNA hybridization study (Ryder et al. unpublished). Fossil material strongly supports the splitting of modern lineages approximately 10-15 million years ago (Prothero 1989; Groves 1983) rather than the remarkably recent dates suggested by the DNA-DNA hybridization study. However, questions about calibration of molecular clocks will not be addressed here.

The higher level study in this chapter adds additional samples and characters to the preliminary study of Amato et al. (1993) in addressing the relationships of the living taxa. Such higher level studies employing the same techniques and underlying assumptions as used in studies concerned with diagnosing units of conservation can provide important information about which regions of the genome are likely to prove informative for species identification (Amato et al. 1993; Amato and Gatesy in press). Additionally, a weighting scheme of DNA characters based on ambiguity of homology is explored, as well as a total evidence analysis including both molecular and morphological

characters.

MATERIALS AND METHODS

Forty-four rhinoceros representing the four living genera (Table 6) and two tapirs (Tapirus indicus) were sequenced for approximately 958 bases of 12S and 16S ribosomal mitochondrial genes. Individuals were sampled in a variety of manners as dictated by specific circumstances in the field and international collections. Samples included frozen blood, frozen tissue, blood preserved in RT buffer (100 mM Tris, 100 mM EDTA, and 2%SDS) and stored at room temperature, and shed hair and skin kept dry at room temperature. All samples were obtained without harm to the study animals. Total genomic DNA was isolated for all of the blood samples by previously described standard phenol/chloroform isolation procedures (Caccone et al. 1987). A method employing a chelating resin (Chelex 100, BioRad) optimized for forensics samples (Walsh et al. 1991) was used to isolate DNA from the shed hair and skin samples.

Fragments of the 12S and 16S ribosomal mitochondrial genes were PCR amplified with modified universal vertebrate primers (Kocher et al. 1989; Palumbi et al. 1990). PCR reactions were carried out in 100 ul reaction volumes with reagents from Perkin-Elmer Cetus Gene Amp Kit. Reactions were performed in a Perkin-Elmer Cetus DNA Thermal Cycler with approximately 250 ngs of template DNA and a magnesium concentration of 1.5 mM. Cycling conditions were 94^{o}C for 1 min., 55^{o}C for 1.5 min., and 72^{o}C for 2 min. for forty cycles. Most often, unbalanced primers were used to accomplish asymmetric PCR (Gyllensten and Erlich 1988). Single stranded PCR products were cleaned and concentrated with centricon-30 columns (Amicon) and directly

sequenced by the dideoxy method with reagents and protocol from USB's Sequenase 2.0 sequencing kit (Gatesy and Amato 1992). Some sequences were obtained using an automated sequencer (Applied Biosystems Model 373A) following the manufacturer's protocols. Both strands were sequenced to insure accuracy. For some samples, where only one strand could be read close to the primer, that stretch of sequence was coded as missing.

Individuals with identical sequences were grouped together as a single operational taxonomic unit (OTU). Different OTU sequences were aligned with the parsimony based multiple alignment program MALIGN (Wheeler and Gladstein 1992). Five alignments were generated by varying gap costs. Maximum parsimony cladograms were constructed using exact searches with the branch and bound setting of PAUP 3.1 for the Macintosh (Swofford 1991). The tapir (*Tapirus indicus*) was used as an outgroup. Choice of the tapir as an outgroup is based on extensive morphological character analyses of extant and fossil species (Prothero and Schoch 1989; Groves 1983). Bootstrap resamplings (Felsentein 1985) were performed as a frame of reference for inferring robustness.

In addition to the sequence data, morphological characters (Groves 1983) were selected based on Groves' determination of observed polarity in reference to the outgroup (Table 7). Only those characters that were phylogenetically informative and could be polarized were used in these analyses. These characters were coded as present or absent and included in a total evidence parsimony analysis.

A method of weighting nucleotide positions based on levels of ambiguity (elision) (Wheeler et al. 1994) was used. Essentially, this involves "stacking" the five alignments on top of each other and assigning weights to each position from one to five depending on whether a position is identical, or shares the

same base as any of the other four alignments. This weighting scheme involves the least number of assumptions about the relative "value" of positions and is objective and repeatable.

RESULTS

Approximately 401 bases of 12S and 553 bases of 16S ribosomal mitochondrial gene fragments were sequenced for the 42 rhinos and two tapirs (Tables 6 and 8). These 44 sequences represent 10 unique mitochondrial haplotypes designated as black rhino (BR), black rhino* (BR*), northern white rhino (NWR), southern white rhino (SWR), Indian rhino (IR), Sumatran rhino-West Malaysia (SRWM), Sumatran rhino-Sumatra (SRS), Sumatran rhino-Sumatra* (SRS*), Sumatran rhino-Borneo (SRB), and Malayan tapir (MT). These ten types were used as OTU's in generating cladograms. A detailed description of inter and intra population sequence divergence is found in Chapter 4.

Five different alignments for 12S and 16S sequences were obtained by doubling gap costs from 2 to 16. Each alignment generated a single most parsimonious tree. A gap cost of 2 resulted in a 12S tree that placed white rhinos and Sumatran rhinos as sister taxa (Figure 8). This tree was the only tree generated that appeared nonsensical based on the strong morphological support uniting the African lineages. The remaining four alignments support the African lineages as sister taxa and place the Indian rhino as sister taxon to all other rhinos (Figure 8). In contrast to the 12S trees, all five 16S alignments support a sister taxa relationship between the African lineages but also a sister taxa relationship between the Asian lineages (Figure 9). Bootstrap values indicate that the placement of the Indian rhino is the most weakly supported (Figures 8 and 9).

Based on bootstrap values, 12S and 16S alignments with a gap cost of 16 were combined for a total molecular analysis. Two equally parsimonious trees were generated (Figure 10). Not surprisingly, the trees differed on the placement of the Indian rhino. However, a total evidence analysis with equal weighting of all sequence data and the 21 morphological characters (Groves 1983) for which polarity could be established, resulted in a single most parsimonious tree (Figure 11). This cladogram places the Indian rhino as the sister taxon to the Sumatran rhinos. Bootstrap values calculated for 1000 replications yielded values of 100% for every node with the exception of 92% for the previously unresolved Indian rhino/Sumatran rhino node. This tree has a Consistency index excluding uninformative characters of .88 and a Retention index of .93.

12S and 16S trees constructed with sites weighted based on ambiguity of homology (elision) had identical topologies to the other 12S and 16S trees (Figure 12). However, bootstrap values, CIs, and RIs were higher with this weighting scheme. When both sequences were combined for a total molecular data analysis with sites weighted by elision, only a single most parsimonious tree was derived (Figure 12). This tree has the same topology as the 16S trees and the total evidence molecular/morphological tree (Indian rhinos as sister taxon to Sumatran rhinos).

DISCUSSION

The regions selected for study of conservation units (12S and 16S ribosomal mitochondrial sequences) clearly contain phylogenetically informative characters. Interestingly, the ambiguous placement of the Indian rhino mirrors the conflict between two morphological analyses (Guerin 1982; Groves 1983). Since the split of the Indian and Sumatran rhino lineages, based

on fossil evidence, is thought to be much older than the African lineages (Prothero et al. 1986; 1989), it appears that this "conserved" mitochondrial region is most informative for the more recent events. Additional ongoing studies (Amato et al. unpublished) on more quickly evolving regions (cytochrome b and D-Loop) are not likely to address this ancient split.

Combining the 16S and 12S data without weighting did not further resolve the Indian rhino placement. However, when the elision weighting scheme is applied to the combined molecular data, the result is a single well supported tree that agrees with the total molecular/morphological tree, the morphological trees of Groves (1983) and Prothero (1986), and the DNA-DNA hybridization tree (Ryder et al. unpublished). This result supports the importance of identifying homology of sites, and the use of weighting based on ambiguity of homology assessment.

The importance of a total evidence analysis is supported by the strong resolution in the total unweighted molecular/morphological tree. In this case the combined molecular data (954 bases) resulted in two equally parsimonious trees. However, with the addition of only 21 morphological characters, we have much greater resolution. This is contrary to the criticism of total evidence analysis which suggests that smaller data sets will be overwhelmed by the larger data sets. Since our only alternatives in handling different data sets are to consense or combine, and consensing has been shown to be inferior (Donoghue 1993), I chose to rely on total evidence results.

This analysis offers the first total evidence support for a sister taxa relationship between the two Asian lineages. Agreement with results obtained from two different morphological studies and a molecular study using different data adds additional confidence to this assertion. I believe that concordance between results of different studies offers more support than concordance of

trees derived from alternative analyses of the same data using algorithms and assumptions that are logically inferior.

TABLE 6. Rhinoceros samples included in this study.

IDENTIFICATION NUMBER	COUNTRY/LOCATION
Dicerorhinus sumatrensis	
6	Sumatra
22	Sumatra
24	Sumatra
27	Sumatra
28	Sumatra
33	Sumatra
17	Borneo
26	Borneo
31	Borneo
38	Borneo
1	West Malaysia
7	West Malaysia
13	West Malaysia
15	West Malaysia
19	West Malaysia
20	West Malaysia
23	West Malaysia
Diceros bicornis	
k1	Kenya
k2	Kenya
k3	Kenya
k4	Kenya
k5	Kenya
k6	Kenya
k7	Kenya
k8	Kenya
k9	Kenya
z10	Kenya
z11	Zimbabwe
z12	Zimbabwe
z13	Zimbabwe
z14	Zimbabwe
z15	Zimbabwe
z16	Zimbabwe
z18	Zimbabwe
z19	Zimbabwe
z20	Zimbabwe
Ceratotherium simum	
s1	San Diego Zoo

s2	San Diego Zoo
n3	San Diego Zoo
Rhinoceros unicornis	
b1	Bronx Zoo
b2	Bronx Zoo
Tapirus indicus	
b3	Bronx Zoo
b4	Bronx Zoo

TABLE 7. Rhinoceros morphological characters obtained from Groves (1983) used for this study. Characters were chosen from his list of 42 that were polarized. His criteria for polarity and character states for all characters were used. Only those characters in his 1983 publication were chosen where he discusses polarity. There may be other characters in his list of 42 that are valid for this study, but no mention of polarity is made in his description of these characters so they were omitted from the present analysis.

1. Orientation of the occipital crest. V=vertical; F=forward; B=backwards; not ordered and vertical is listed as ancestral.

2.-8. No mention of polarity for these characters.

9. Anterior migration of the orbit; Anterior=A; Posterior=P; Anterior is described as ancestral.

10. Supraorbital bony shelf. +=present; -=absent; absent is listed as ancestral.

11. Postorbital process; +=prominent; -=no trace of bony division; - is listed as ancestral.

12. No information on ancestral state.

13. Vomer characteristics; +=convex ridge; -=no ridge; ridge is derived.

14. Pterygoid plates; +=extended at free ends; -=not extended; not extended is ancestral.

15. Pterygoid plates shortened posteriorly=+; plate ends at M2/3=-; ending at M2/3 is primitive.

16. No information on primitive state.

17. Mastoid region; +=inflated; -=not inflated; primitive state is no inflation.

18. No information on polarity.

19. Mandibular incisor-premolar diastema; +=thin ridge on either side; -=no ridges; thin ridge is primitive.

20. Inferior margin of the mandibular corpus; +=convex; -=straight; straight is primitive.

21. Ascending ramus; vertical=0; slopes forward=1; slopes backward=2; vertical is primitive.

22. Sympheseal region of the mandible; lingual contour is round and U-shaped=-; inner margins come together to form a V in front=+; U-shaped is primitive.

23. No polarity data.

24. Premaxillae; maintains a horizontal course down=+; upraised=-; missing =?; down is primitive.

25. Upper incisors; lost=-; present=+; primitive is present.

26. Procumbent mandibular tusks; retained and well developed=+; lost=-; retained is ancestral.

27. No polarity information.

28. Crochet and small loph from anterior margin of the metaloph; emerges=+; absent=-; absence is most likely primitive.

29. No polarity assessment.

30. Entrance to the median valleys on the cheekteeth; wide=+; prtocone and hypocones are expanded=-; expanded protocones and hypocones are primitive.

31. No polarity assessment.

32. Buccal pillar; missing=-; present=+; missing is primitive.

33. Anterocrochet on molars; present=+; absent=-; presence is primitive.

34. M3 shape; trapezoidal=+; subtriangular=-; subtriangular is primitive.

35. Lower molar valleys; U-shaped=+; V-shaped=-; primitive state is U-shaped.

36. Lower molar valleys; nearly of equal depth=+; not equally deep=-; not equal is primitive state.

37. Rudimentary anterior premolar; shed before adulthood=+; retained=-; retained is primitive.

38. No polarity information.

39. Disposition of the proximal end of the fibula; fibular head is short and blunt=+, fibular head elongated and emerges proximal of the tibia's lateral condylar surface=-; elongated is primitive.

40. Proportions of long bones; proximal limb segments are long=-; shortened=+; long is primitive.

41. No polarity information.

42. No polarity information.

TABLE 8. 12s and 16S mitochondrial DNA sequences for 42 rhinoceros and 2 tapirs (BR= *Diceros bicornis*, SWR= *Ceratotherium simum simum*, NWR= *C. s. cottoni*, IR= *Rinoceros unicornis*, SRS= *Dicerorhinus sumatrensis* (Sumatra), SRWM= *D. s.* (West Malaysia), SRB= *D.s.* (Borneo), and MT= *Tapirus indicus.*

BR 1

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 2

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 3

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 4

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 5

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 6

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 7

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 8

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 9

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 10

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 11

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 12

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 13

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 14

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 15

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 16

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 17

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 18

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 19

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGA GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

BR 20*

```
  1 GCCTAGCCTT AAACTTAAAT AATTTTCCCA ACAAAATTAT TCGCTAGAGT ACTACAAGCA
 61 ACAGCTTAAA ACTCAAAAGA CTTGGCGGTG CTTTATATCC CCCTAGAGGT GCCTGTTCCA
121 TAATCGATAA ACCCCGATAA ACCCTACCAG CCCTTGCTAA TTCAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAACAAG GAACTAAAGT AAGCACAAGT ATAAGACATA AAAACGTTAG
241 GTCAAGGTGT AGCTTATGGG ATGGAGAGAA ATGGGCTACA TTTTCTACTC TAAGAACAAC
301 AATTACCCAA ACGAAAGTTT CCATGAAACC AAAAACTAAA GGAGGATTTA GCAGTAAATT
361 AAGAATAGAG AGCTTAATTG AACCAGGCCA TAAAGCACGC
```

IR 1

```
  1 GCTTAGCCCC AAACTCAAAT AATTCTTCCC AACAAAATTA TTCGCCAGAG TACTACTAGC
 61 AACAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA GCCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCCA ACCCTAAAAA GGAACCAAAG TAAGCACAAG TATAAGACAT AAAAACGTTA
241 GGTCAAGGTG TAGCTTATGG GATGGAGAGA AATGGGCTAC ATTTTCTACT TCAAGAACAA
301 CAACTACCCA AACGAAGGCT TTTATGAAAT TAAAAGCTAA AGGAGGATTT AGCAGTAAAT
361 TAAGAATAGA GAGCTTAATT GAACCAGGCC ATAAAGCACG C
```

IR 2

```
  1 GCTTAGCCCC AAACTCAAAT AATTCTTCCC AACAAAATTA TTCGCCAGAG TACTACTAGC
 61 AACAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA GCCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCCA ACCCTAAAAA GGAACCAAAG TAAGCACAAG TATAAGACAT AAAAACGTTA
241 GGTCAAGGTG TAGCTTATGG GATGGAGAGA AATGGGCTAC ATTTTCTACT TCAAGAACAA
301 CAACTACCCA AACGAAGGCT TTTATGAAAT TAAAAGCTAA AGGAGGATTT AGCAGTAAAT
361 TAAGAATAGA GAGCTTAATT GAACCAGGCC ATAAAGCACG C
```

MT 1

```
  1 GCCTAGCCAT AAACCAAAAT AATCTCCATA ACAAAATTAT TCGCCAGAGT ACTACTAGCA
 61 ACAGCCTAAA ACTCAAAGGA CTTGGCGGTG CTTTATATCC CTCTAGAGGA GCCTGTTCCG
121 TAATCGATAA ACCCCGATAA ACCTTACCAC CCCTTGCCAA TACAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAAAAAA GGAAACAAAG TAAGCATAAG CATAGGACAT AAAAACGTTA
241 GGTCAAGGTG TAGCTTATGA GGTGGAGAGA AATGGGCTAC ATTTTCTAAC CAAGAACAAC
301 ACATCCCATG ACACGAAAGT TTTTATGAAA CTAAAAACTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAATAG AGAGCTTAAT TGAACTAGGC CATGAAGCAC GC
```

MT 2

```
  1 GCCTAGCCAT AAACCAAAAT AATCTCCATA ACAAAATTAT TCGCCAGAGT ACTACTAGCA
 61 ACAGCCTAAA ACTCAAAGGA CTTGGCGGTG CTTTATATCC CTCTAGAGGA GCCTGTTCCG
121 TAATCGATAA ACCCCGATAA ACCTTACCAC CCCTTGCCAA TACAGCCTAT ATACCGCCAT
181 CTTCAGCAAA CCCTAAAAAA GGAAACAAAG TAAGCATAAG CATAGGACAT AAAAACGTTA
241 GGTCAAGGTG TAGCTTATGA GGTGGAGAGA AATGGGCTAC ATTTTCTAAC CAAGAACAAC
301 ACATCCCATG ACACGAAAGT TTTTATGAAA CTAAAAACTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAATAG AGAGCTTAAT TGAACTAGGC CATGAAGCAC GC
```

SRS 1

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRS 2

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRS 3

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRS 4

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRS 5

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRS 6*

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 1

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 2

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 3

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 4

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 5

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 6

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRWM 7

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AAGCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCCA
181 TCTTCAGCAA ACCCTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRB 1

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCGA
181 TCTTCAGCAA ACCGTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRB 2

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCGA
181 TCTTCAGCAA ACCGTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRB 3

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCGA
181 TCTTCAGCAA ACCGTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

SRB 4

```
  1 GCTTAGCCCT AAACCTAAAT GATTTCCCCC AACAAAATCA TTCGCCAGAG TACTACTAGC
 61 AATAGCCTAA AACTCAAAGG ACTTGGCGGT GCTTTATATC CCCCTAGAGG AGCCTGTTCC
121 ATAACCGATA AACCCCGATA AACCTTACCA ACCCTTGCTA ATTCAGCCTA TATACCGCGA
181 TCTTCAGCAA ACCGTAAAAA AGGAACTAAA GTAAGCACAA GTATAAGACA TAAAAACGTT
241 AGGTCAAGGT GTAGCTTATG GGATGGAGAG AAATGGGCTA CATTTTCTAC TACAAGAACA
301 ACAATTATCC AAACGAAAGC CCCCATGAAA CTAAGGGCTA AAGGAGGATT TAGCAGTAAA
361 TTAAGAACAG AGAGCTTAAT TGAACAAGGC CATAAAGCAC GC
```

NWR 1

```
  1 TAGCAACAGC CTAAAACTCA AAGGACTTGG ACAGTGCTTT ATATCCCCCT AGAGGAGCCT
 61 GTTCCATAAC CGATAAACCC CGATAAACCC CACCAACCCT TGCTAATTCA GCCTATATAC
121 CGCCATCTTC AGCAAACCCT AAAAAGGAAC TAAAGTAAGC ACAAGTATAA ACATAAAAAC
181 GTTAGGTCAA GGTATAGCTT ATGGGATGGA GAGAAATGGG CTACATTTTC TATTTTAAGA
241 A
```

SWR 1

```
  1 TAGCAACAGC CTAAAACTCA AAGGACTTGG ACAGTGCTTT ATATCCCCCT AGAGGAGCCT
 61 GTTCCATAAC CGATAAACCC CGATAAACCC CACCAACCCT TGCTAATTCA GCCTATATAC
121 CGCCATCTTC AGCAAACCCT AAAAAGGAAC TAAAGTAAGC ACAAGTATAA AACATAAAAA
181 CGTTAGGTCA AGGTGTAGCT TATGGGATGG AGAGAAATGG GCTACATTTT CTATTTTAAG
241 AA
```

SWR 2

```
  1 TAGCAACAGC CTAAAACTCA AAGGACTTGG ACAGTGCTTT ATATCCCCCT AGAGGAGCCT
 61 GTTCCATAAC CGATAAACCC CGATAAACCC CACCAACCCT TGCTAATTCA GCCTATATAC
121 CGCCATCTTC AGCAAACCCT AAAAAGGAAC TAAAGTAAGC ACAAGTATAA AACATAAAAA
181 CGTTAGGTCA AGGTGTAGCT TATGGGATGG AGAGAAATGG GCTACATTTT CTATTTTAAG
241 AA
```

16S SEQUENCES START HERE

BR 1

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 2

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 3

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 4

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 5

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 6

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 7

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 8

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 9

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 10

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 11

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 12

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 13

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 14

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 15

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

BR 16

60 CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT

BR 17

60 CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT

BR 18

60 CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT

BR 19

60 CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT

BR 20*

```
60  CACCTCCAGC ATATCTAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTNCA AAGGTAGCGT AATCACTTGT TCCCTAAATA GGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATGACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAGCTAA
300 TTCACAAAAA TAAAATCTCT AACCTACATC CAAGGGATAA CAAAACTTTG ACTGAATCAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AACCAACCTC CGAATGATTA AATTCCAGNC
420 TTACCAGTCN AAAATATCAC ATCACTTATT GATCCAAACT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA CCCCAATGGT GNAACCGCTA TTAATGGNTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

IR 1

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACTTGTA
180 TGANTGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATCAACTAA
300 TTCACAAAAT AAAATCTTCA ACCTACATAT AAGGGATAAC AAAATTTTAA CTGAATTAAC
360 GATTTCGGTT GGGGTGACCT CGGAGAACAA AACAACCTCC GAGTGATTAA ATTCCAGACT
420 TACTAGTCAA AAATATTACA TCACTTATTG ATCCAAATTA TTGATCAACG GAACAAGTTA
480 CCCTAGGGAT AACAGCGCAA TCCTATTCTA GAGTCCATAT CGACAATAGG GTTTACGACC
540 TCGATGTTGG ATCAGGACAT CCCAATGGTG CAACCGCTAT TAATGGTTCG TTTGTTCAAC
552 GATTAAAGTC CT
```

IR 2

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACTTGTA
180 TGANTGGCCA CACGAGGGTT TTACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACA CAATAAGACG AGAAGACCCT ATGGAGCTTT AATCAACTAA
300 TTCACAAAAT AAAATCTTCA ACCTACATAT AAGGGATAAC AAAATTTTAA CTGAATTAAC
360 GATTTCGGTT GGGGTGACCT CGGAGAACAA AACAACCTCC GAGTGATTAA ATTCCAGACT
420 TACTAGTCAA AAATATTACA TCACTTATTG ATCCAAATTA TTGATCAACG GAACAAGTTA
480 CCCTAGGGAT AACAGCGCAA TCCTATTCTA GAGTCCATAT CGACAATAGG GTTTACGACC
540 TCGATGTTGG ATCAGGACAT CCCAATGGTG CAACCGCTAT TAATGGTTCG TTTGTTCAAC
552 GATTAAAGTC CT
```

MT 1

```
60  CACCTCTAGC ATTACCAATA TTAGAGGCAC TGCCTGCCCA GTGACACCTG TTTAAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACTTGTA
180 TGAATGGCCA CACGAGGGTT TAACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATAACA AAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAATCAA
300 TTCACTAAAA ATAAACTTTC AACCTACCAG GTATAACAGA ACTTTAACTG AATTGACAAT
344 TTCGGTTGGG GTGACCTCGG NGAATAAACA ACCCCCGAGT GATT
```

MT 2

```
60  CACCTCTAGC ATTACCAATA TTAGAGGCAC TGCCTGCCCA GTGACACCTG TTTAAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACTTGTA
180 TGAATGGCCA CACGAGGGTT TAACTGTCTC TTACTTTCAA TCAGTGAAAT TGACCTTCCC
240 GTGAAGAGGC GGGAATAACA AAATAAGACG AGAAGACCCT ATGGAGCTTT AATTAATCAA
300 TTCACTAAAA ATAAACTTTC AACCTACCAG GTATAACAGA ACTTTAACTG AATTGACAAT
344 TTCGGTTGGG GTGACCTCGG NGAATAAACA ACCCCCGAGT GATT
```

SRB 1

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGCGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRB 2

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGCGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRB 3

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGCGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRB 4

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGCGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRS 1

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRS 2

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRS 3

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRS 4

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRS 5

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRS 6

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 1

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 2

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 3

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 4

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 5

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 6

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

SRWM 7

```
60  CACCTCTAGC ATACCCAGTA TTAGAGGCAC TGCCTGCCCA GTGACATCTG TTTCAACGGC
120 CGCGGTATCC TAACCGTGCA AAGGTAGCAT AATCACTTGT TCTCTAAATA AGGACCTGTA
180 TGAATGGCCA CACGAGGGTT TTACTGTCTC TTACCTTCAA TCAGTGAAAT TGACCTCCCC
240 GTGAAGAGGC GGGGATAACG CAACAAGACG AGAAGACCCT ATGGAGCTTC AATTAACTAA
300 TTCACAAAAA CAAAACCTTC AACCTATATC TAAGGAATAA CAAAATTTCG ATTGAATTAG
360 CAATTTCGGT TGGGGTGACC TCGGAGAACA AAACAACCTC CGAGTGATTA AATTCTAGAC
420 TAACCAGTCA AAAATAATAC ATCACTTATT GATCCAAATT ATTGATCAAC GGAACAAGTT
480 ACCCTAGGGA TAACAGCGCA ATCCTATTCT AGAGTCCATA TCGACAATAG GGTTTACGAC
540 CTCGATGTTG GATCAGGACA TCCTAATGGT GTAACCGCTA TTAATGGTTC GTTTGTTCAA
553 CGATTAAAGT CCT
```

NWR 1

```
60  GGTAGCGTAA TCACTTGTTC CCTAAATAGG GACCTGTATG AATGGCCACA CGAGGGTTTT
120 ACTGTCTCTT ACTTTCAATC AGTGAAATTG ACCTTCCCGT GAAGAGGCGG GAATAGCACA
180 ATAAGACGAG AAGACCCTAT GGAGCTTTAA TTAACTAATT CACAAAAAAT AAAATCTCTA
240 ACCCACATCC AGGGGATAAC AAAACTTTGA CTGAATTAGT AATTTCGGTT GGGGTGACCT
300 CGGAGAACAA ACCAACCTCC GAGTGATTAA ATTCCAGACT TACCAGTCAA AAATATTACA
357 TCACTTATTG ATCCAAACCA TTGATCAACG GAATAAGTTA CCCTAGGATA ACACGCG
```

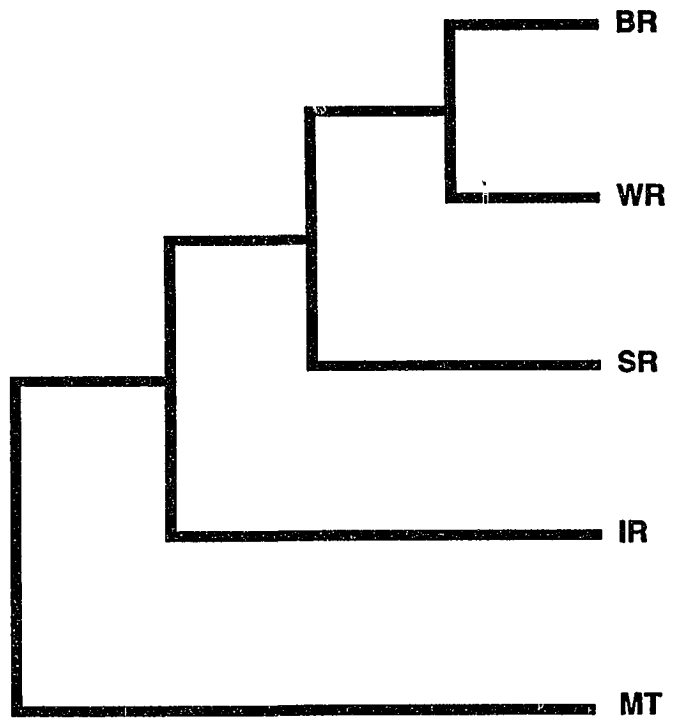
SWR 1

```
60  GGTAGCGTAA TCACTTGTTC CCTAAATAGG GACCTGTATG AATGGCCACA CGAGGGTTTT
120 ACTGTCTCTT ACTTTCAATC AGTGAAATTG ACCTTCCCGT GAAGAGGCGG GAATAGCACA
180 ATAAGACGAG AAGACCCTAT GGAGCTTTAA TTAACTAATT CACAAAAAAT AAAATCTCTA
240 ACCTACATCC AGGGGATAAC AAAACTTTGA CTGAATTAGT AATTTCGGTT GGGGTGACCT
300 CGGAGAACAA GCCAACCTCC GAGTGATTAA ATTCCAGACT TACCAGTCAA AAATATTACA
357 TCACTTATTG ATCCAAACCA TTGATCAACG GAATAAGTTA CCCTAGGATA ACACGCG
```

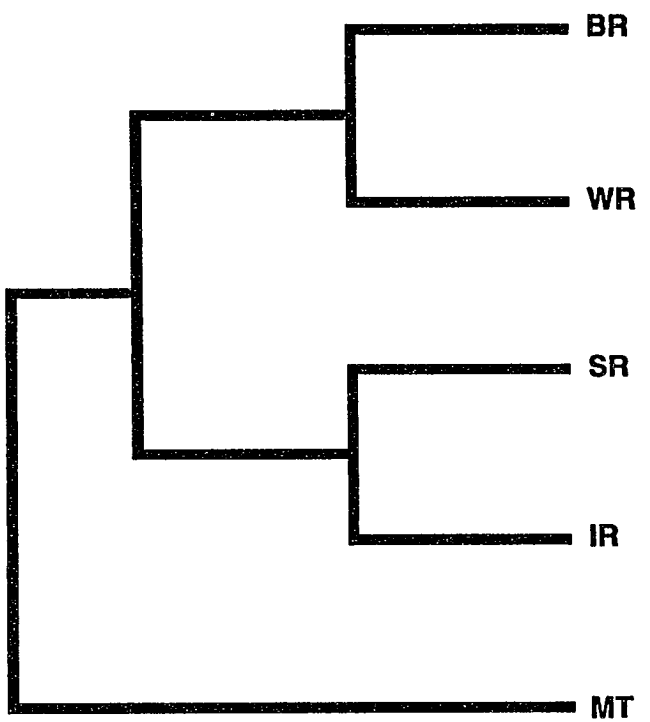
SWR 2

```
60  GGTAGCGTAA TCACTTGTTC CCTAAATAGG GACCTGTATG AATGGCCACA CGAGGGTTTT
120 ACTGTCTCTT ACTTTCAATC AGTGAAATTG ACCTTCCCGT GAAGAGGCGG GAATAGCACA
180 ATAAGACGAG AAGACCCTAT GGAGCTTTAA TTAACTAATT CACAAAAAAT AAAATCTCTA
240 ACCTACATCC AGGGGATAAC AAAACTTTGA CTGAATTAGT AATTTCGGTT GGGGTGACCT
300 CGGAGAACAA GCCAACCTCC GAGTGATTAA ATTCCAGACT TACCAGTCAA AAATATTACA
357 TCACTTATTG ATCCAAACCA TTGATCAACG GAATAAGTTA CCCTAGGATA ACACGCG
```

FIGURE 7. Alternative hypotheses of higher level relationships of rhinoceros based on Guerin (1982) (A) and Groves (1983) (B). (BR=*Diceros bicornis*; WR=*Ceratotherium simum*; SR=*Dicerorhinus sumatrensis*; IR=*Rhinoceros unicornis*; MT=*Tapirus indicus*).



A



B

FIGURE 8. Maximum parsimony trees based on 12S sequences for five alignments with gap costs of 2, 4, 8, 16, and 32. A gap cost of 2 resulted in topology (A) and had bootstrap values of .49 for the node connecting white rhinos and Sumatran rhinos and .47 for the node connecting Indian rhinos to the SR/WR clade. All other nodes on this tree had bootstrap values greater than .96. The other four gap costs result in topology (B). A gap cost of 16 resulted in bootstrap values of .76 for the node connecting the African rhinos and .36 for the node connecting African rhinos and Sumatran rhinos. (BR = Black rhino; BR* = Black rhino #10; NWR = Northern white rhino; SWR = Southern white rhino; SRS = Sumatran rhino-Sumatra; SRS* = Sumatran rhino-Sumatra #6; SRWM = Sumatran rhino-West Malaysia; SRB = Sumatran rhino-Borneo; IR = Indian rhino; MT = Malayan tapir).

B

A

FIGURE 9. Maximum parsimony tree based on 16S sequences for five alignments with gap costs of 2, 4, 8, 16, and 32. All alignments produce the same topology. A gap cost of 16 resulted in a bootstrap value of .59 for the node connecting Indian rhinos and Sumatran rhinos. All other nodes have bootstrap values of 100. (BR = Black rhino; BR* = Black rhino #10; NWR = Northern white rhino; SWR = Southern white rhino; SRS = Sumatran rhino-Sumatra; SRS* = Sumatran rhino-Sumatra #6; SRWM = Sumatran rhino-West Malaysia; SRB = Sumatran rhino-Borneo; IR = Indian rhino; MT = Malayan tapir).

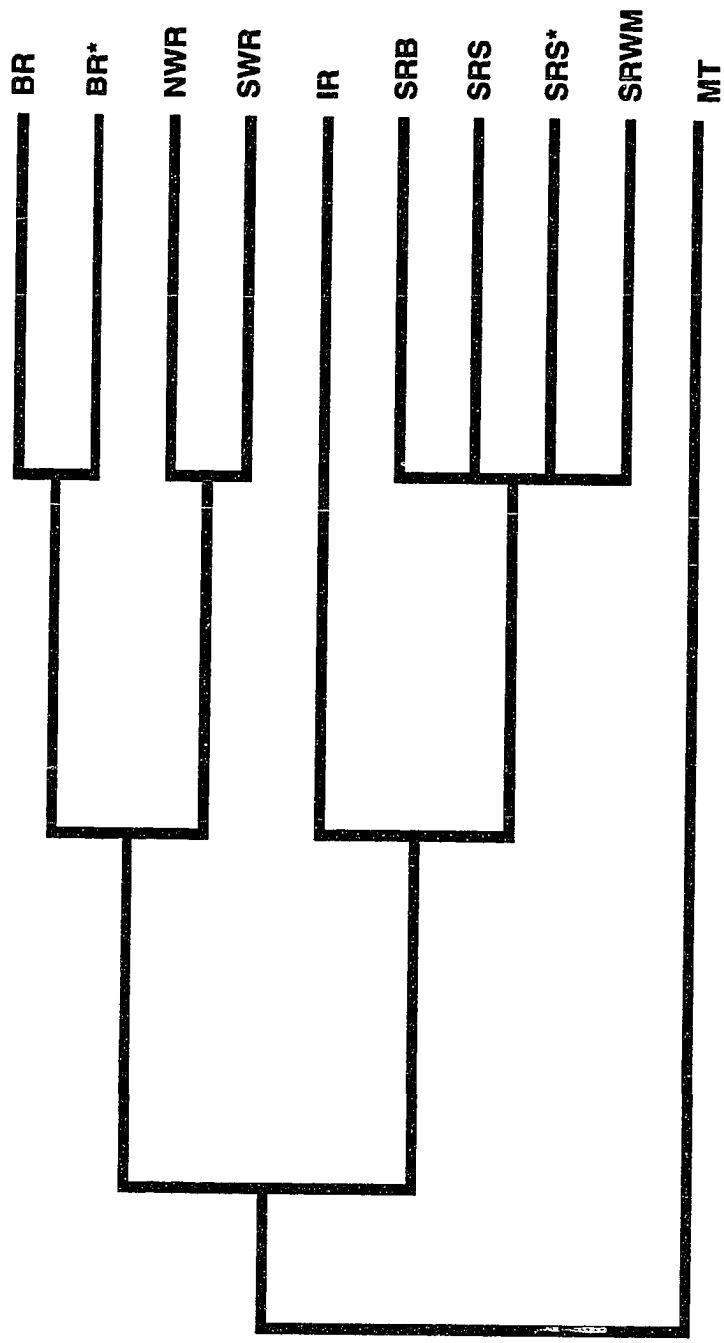


FIGURE 10. Two equally most parsimonious trees are recovered in an analysis of unweighted total molecular data. In tree (A) the node connecting Indian rhinos to all other rhinos has a bootstrap value less than .50. All other nodes have bootstrap values of 100. In tree (B) the node connecting Indian rhinos and Sumatran rhinos has a bootstrap value of .58. All other nodes have bootstrap values of 100. (BR = Black rhino; BR* = Black rhino #10; NWR = Northern white rhino; SWR = Southern white rhino; SRS = Sumatran rhino-Sumatra; SRS* = Sumatran rhino-Sumatra #6; SRWM = Sumatran rhino-West Malaysia; SRB = Sumatran rhino-Borneo; IR = Indian rhino; MT = Malayan tapir).

FIGURE 11. A single most parsimonious tree derived from total evidence (12S, 16S, and morphological characters). The node connecting Indian rhino and Sumatran rhinos has a bootstrap value of .92. All other nodes have bootstrap values of 100. The consistency index (CI) is .92 and the retention index (RI) is .93. (BR = Black rhino; BR* = Black rhino #10; NWR = Northern white rhino; SWR = Southern white rhino; SRS = Sumatran rhino-Sumatra; SRS* = Sumatran rhino-Sumatra #6; SRWM = Sumatran rhino-West Malaysia; SRB = Sumatran rhino-Borneo; IR = Indian rhino; MT = Malayan tapir).

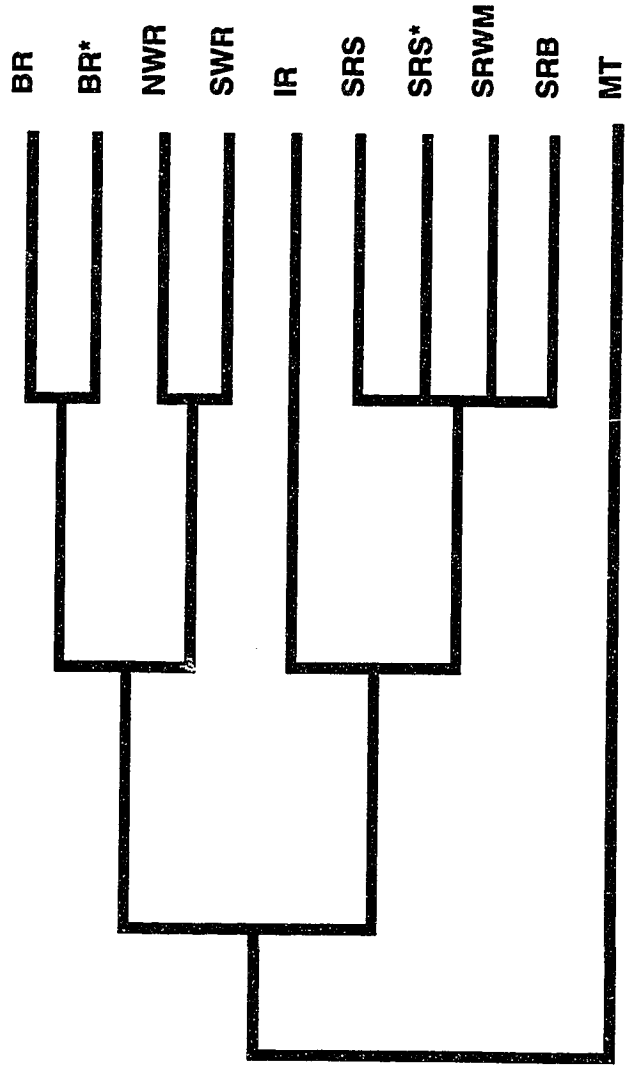
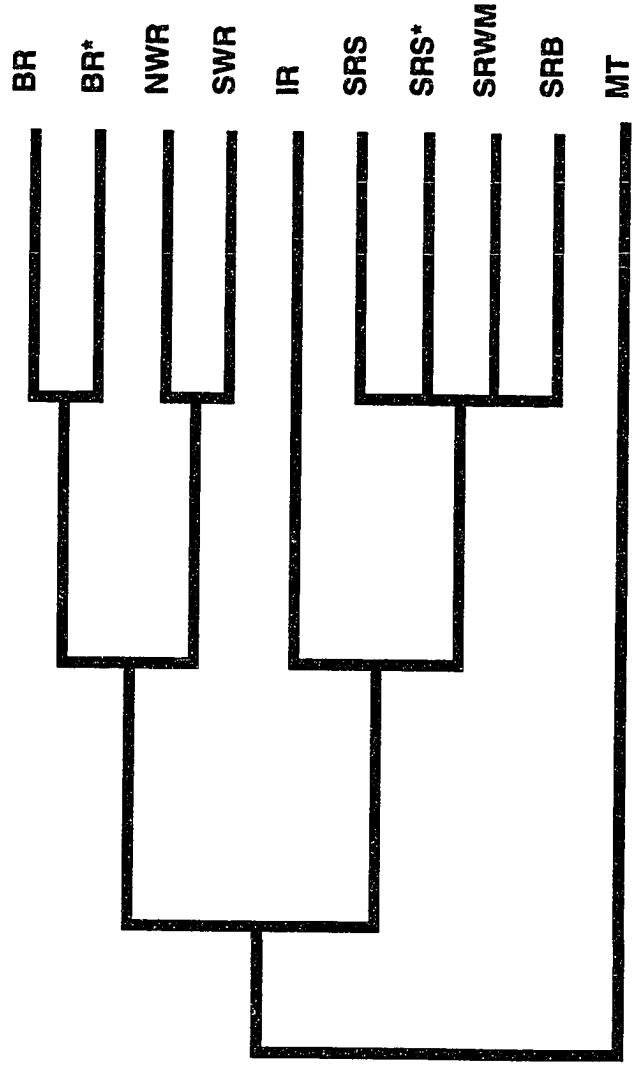


FIGURE 12. A single most parsimonious tree derived from total molecular evidence weighted by elision. The node connecting Indian rhino and Sumatran rhinos has a bootstrap value of .52. All other nodes have bootstrap values of 100. The consistency index (CI) is .92 and the retention index (RI) is .92. (BR = Black rhino; BR* = Black rhino #10; NWR = Northern white rhino; SWR = Southern white rhino; SRS = Sumatran rhino-Sumatra; SRS* = Sumatran rhino-Sumatra #6; SRWM = Sumatran rhino-West Malaysia; SRB = Sumatran rhino-Borneo; IR = Indian rhino; MT = Malayan tapir).



CHAPTER 4

AN ASSESSMENT OF CONSERVATION UNITS FOR THE SUMATRAN RHINOCEROS (*Dicerorhinus sumatrensis*)

SUMMARY

An assessment of conservation units for the Sumatran rhinoceros (*Dicerorhinus sumatrensis*) was conducted using a population aggregation analysis (PAA) of mitochondrial DNA site substitutions. Populations were defined as the three geographically separated regions of West Malaysia, Sumatra, and Borneo.

Individual DNA positions were not diagnostic for any population. A single haplotype provided a character as support for diagnosing the West Malaysian and Bornean population. The haplotypes on West Malaysia and Sumatra were more similar to each other than either were to the one on Borneo. These data, and a review of the morphological characters, do not support the designation of more than one conservation unit for Sumatran rhinos.

INTRODUCTION

The Sumatran rhinoceros (*Dicerorhinus sumatrensis*) is a highly endangered species currently confined to a few remnant upland forest areas in peninsular Malaysia, Sumatra, and Borneo. Like the other extant rhinos, the Sumatran rhino originally had an extensive distribution (Figure 13). Until the beginning of this century it ranged from Assam and Bengal, through Burma, Thailand, Cambodia, Laos, Viet Nam, China, Malaysia, Sumatra, and Borneo (Groves 1982).

While historically the rhinos used habitats that included lowland forests

and natural clearings, their presence in upland forest and mountainous regions explains why the species has persisted in more areas and in larger numbers than the sympatric Javan rhino (*Rhinoceros sondaicus*) which is confined to lowland forests (Santiapillai and Mackinnon 1993; Penny 1988). These mountainous areas are the last to be deforested and the most difficult in which to hunt the surprisingly nimble animal (Santiapillai and Mackinnon 1993; Khan et al. 1993). Sumatran rhino tracks have been found up to 2,000 meters in elevation.

Currently, the only viable populations persist in Sumatra (Gunung Leuser National Park, Torgamba Forest, Kerinci-Seblat National Park, Barisan Selatan National Park, and Gunung Patah), in West Malaysia (Taman Negara and Endau Rompin), and in Borneo. However, the total number of animals is probably under 500 (including 24 animals in captive breeding programs). In addition to deforestation, the rhinos are threatened by commercial hunting for their horn in both protected and non-protected areas. In 1990 at least ten rhinos were poached in Kerinci-Seblat National Park in Sumatra (Santiapillai and Mackinnon 1993). An organized conservation program is essential for the survival of this species.

The governments of Indonesia and Malaysia, as well as international conservation organizations (The Wildlife Conservation Society, The Worldwide Fund for Nature, IUCN Captive Breeding Specialist Group, and The Sumatran Rhino Trust) have mounted a major effort to conserve this species. Management plans include research, greater protection of wild populations, and a captive breeding program. Since management strategies may include translocating animals or gametes, the question of conservation units is of great importance.

Groves (1967a) divides the species into three subspecies [*D.s.*

sumatrensis (Sumatra and Malaysia), *D.s. harrissoni* (Borneo), and *D.s. lasiotis* (Burma and India)] based on measurements of eight morphological characters. There is no evidence that the populations referred to as D.s. lasiotis still exist. For conservation purposes, I have investigated the three geographically separated populations even though Groves (1967a) groups Sumatra and Malaysia together.

MATERIALS AND METHODS

Seventeen Sumatran rhinos representing the three populations (Table 10) were sequenced for 953 bases of 12S and 16S mitochondrial sequences. Individuals were sampled in a variety of manners as dictated by specific circumstances in the field and international collections. Samples included frozen blood, frozen tissue, blood preserved in RT buffer (100 mM Tris, 100 mM EDTA, and 2%SDS) and stored at room temperature, and shed hair and skin kept dry and at room temperature. All samples were obtained without harm to the study animals. Total genomic DNA was isolated for all of the blood samples by previously described standard phenol/chloroform isolation procedures (Caccone et al. 1987). A method employing a chelating resin (Chelex 100, BioRad) optimized for forensics samples (Walsh et al. 1991) was used to isolate DNA from the shed hair and skin samples.

Fragments of the 12S and 16S ribosomal mitochondrial genes were PCR amplified with modified universal vertebrate primers (Kocher et al. 1989; Palumbi et al. 1990). PCR reactions were carried out in 100 ul reaction volumes with reagents from Perkin-Elmer Cetus Gene Amp Kit. Reactions were performed in a Perkin-Elmer Cetus DNA Thermal Cycler with approximately 250 ngs of template DNA and a magnesium concentration of 1.5 mM. Cycling

conditions were 94°C for 1 min., 55°C for 1.5 min., and 72°C for 2 min. for forty cycles. Most often, unbalanced primers were used to accomplish asymmetric PCR (Gyllensten and Erlich 1988). Single stranded PCR products were cleaned and concentrated with centricon-30 columns (Amicon) and directly sequenced by the dideoxy method with reagents and protocol from USB's Sequenase 2.0 sequencing kit (Gatesy and Amato 1992). Some sequences were obtained using an automated sequencer (Applied Biosystems Model 373A) following the manufacturer's protocols. Both strands were sequenced to assure accuracy.

Sequences were assigned to local populations defined by geographical location (i.e. West Malaysia, Sumatra, and Borneo) (Table 9). Base substitutions were assessed as either characters or traits as defined by Davis and Nixon (1992). This method, population aggregation analysis (PAA), involves successive searches for fixed differences between aggregations of local populations. Characters are attributes that are not polymorphic and are unique within populations. Traits are attributes that may be polymorphic or are not unique to a population. An assessment of conservation units for Sumatran rhinoceros was considered in light of the population aggregation analysis.

RESULTS

Four haplotypes were identified from the seventeen Sumatran rhinos sampled. Only one haplotype was found in the samples from Borneo and one from West Malaysia, and two haplotypes from the animals on Sumatra. Four sites were variable (Table 10). These sites were position #133, 179, and 194 in the 12S sequence and position #313 in the 16S fragment. In total, the Bornean haplotype differed by two positions from Sumatran and three positions from

West Malaysian. West Malaysia and Sumatra vary by one position for one of the Sumatran haplotypes and by two positions for the other Sumatran haplotype.

None of the positions, when considered individually, fit the definition of character as defined by Davis and Nixon (1992) (Table 10). Rather, they would be considered traits. If the suite of substitutions is considered an attribute, then one character supports the separation of the three defined populations (with a polymorphic Sumatra).

These few variable sites show a greater similarity between West Malaysia and Sumatra then either of those populations compared to Borneo. Position #179 and 194 supports Groves' (1967a) subspecies designation placing the Malayan and Sumatran populations together as *D.s. sumatrensis* with the Borneo population as *D.s. harriosoni*.

Discussion

The results of the population aggregation analysis (PAA) of Sumatran rhinos for determining conservation units were equivocal. Single sites were homoplastic and thus not characters by a PAA definition. The use of an entire haplotype as a single character is complicated by the fact that the population on Sumatra is represented by two haplotypes. If we consider these two haplotypes as character states, then we have a single character support for three phylogenetic species at the minimum level of distinction.

It is interesting, but not surprising that the populations on West Malaysia and Sumatra appear slightly more similar than either does to Borneo. The isolation of Borneo by the submersion of the Sunda Shelf probably occurred a little earlier than the isolation of Sumatra from the mainland (Whitten et al. 1987). In general, there is a trend of increasing morphological differences in

birds and mammals as one proceeds from mainland Southeast Asia out along the Indonesian archipelago until the abrupt change that occurs in Sulawesi (Whitten et al. 1987). A number of authors have described this as originally reflecting a cline through the areas that are part of the Sunda shelf that were last connected about 12,000 years ago.

The question of determining conservation units is complicated in this particular case (Amato et al. 1993; Amato and Ryder 1993; Amato and Wharton in press). The populations are currently isolated on the mainland (West Malaysia) and on two islands (Borneo and Sumatra). This temporal and spatial separation is sufficient reason to refer to these populations as separate for taxonomic purposes. However, with the goal of preserving the evolutionary novelty that is represented in the Dicerorhine lineage, can we consider the three populations as part of the same conservation unit? Applying the PAA assessment of phylogenetic species does not argue against diagnosing them as a single conservation unit unless we consider the Sumatran haplotypes as character states. If we consider the haplotypes as a single character supporting three phylogenetic species, it clearly is the weakest support possible from this data set. Expanding the research to more variable regions is problematic due to the available number of samples. Since the three existing populations are greatly reduced in number, the chances of identifying highly variable characters that unite them simply because the intermediates are missing is likely. Also, traits that unite small fragmented populations can reflect inbreeding or the localized presence of a rare mutation in related individuals.

Groves' (1967a and 1993) subspecies designations are based on only eight morphological characters (all measurements as opposed to presence or absence) using a smaller sample size than this study. His West Malaysian and Sumatran measurements overlap extensively. Only Borneo is less similar. The

results reported in this chapter are not in serious conflict with the results from Groves' (1967a) morphological data.

The only other large mammal that has a similar distribution, and has been assessed on status as subspecies/conservation unit, is the orangutan (*Pongo pygmaeus*). Orangutans are found on both Sumatra and Borneo (and prehistoric remains have been found on the mainland), and may be assumed to have been isolated for the same length of time. Two studies (Caccone and Powell 1989; Janczewski et al. 1991) support the division of the two orang populations into minimally distinct species. This apparent conflict with the Sumatran rhino results may reflect such factors as generation time, the orang's obligate arboreal life style and differences in dispersal abilities among others. It is worth noting that the two orang populations interbreed readily and successfully in captivity with no signs of reduced fitness after several generations.

It is also worth noting that rhinoceros are chromosomally very conservative (Houck et al. 1994). Indian, Sumatran and white rhinos all have a karyotype of 2n=82 even though they last shared a common ancestor more than 15 million years ago. This chromosomal conservation reduces concerns about cytogenetic incompatibility.

There is no strong evidence supporting more than one conservation unit for Sumatran rhinos. Chromosomal conservation and degree of sequence divergence makes outbreeding depression (Templeton 1986) an unlikely outcome if individuals, or their gametes, are translocated as part of a conservation management plan. While this research, like all scientific research is falsifiable by the addition of further data, it is unwise to be paralyzed into inaction while waiting for more studies. The question of when enough studies have been conducted to "prove" that there is only one conservation unit

becomes a question of trying to prove the null hypothesis. This is an epistomological problem rather than a scientific problem and should not prevent us from developing a conservation management plan to preserve this unique taxon.

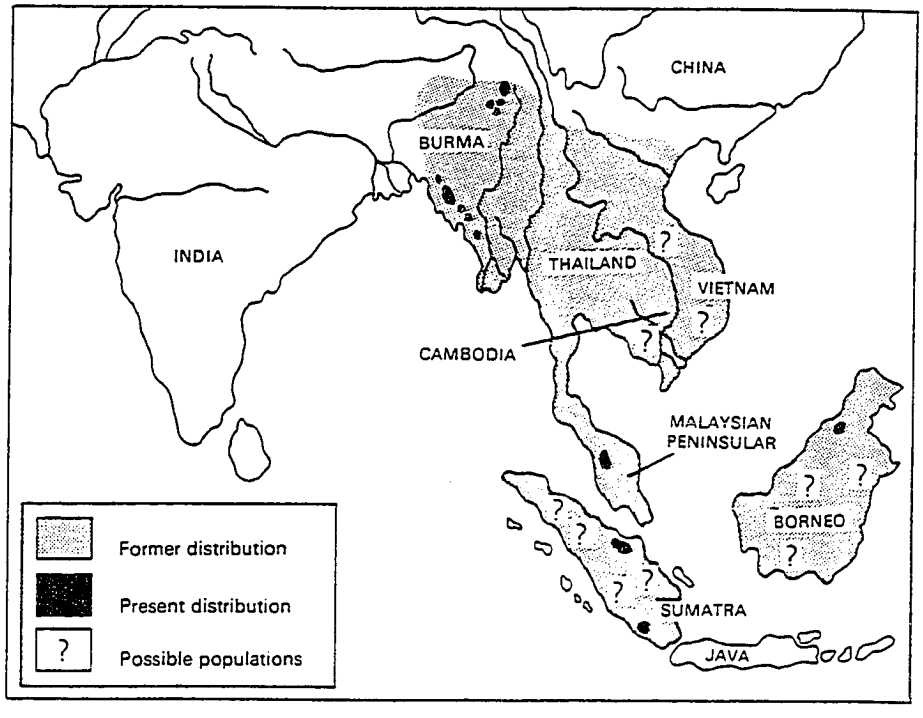
TABLE 9. Sumatran rhinoceros samples included in this study.

INTERNATIONAL STUDBOOK NUMBER	LOCATION
Dicerorhinus sumatrensis	
6	Sumatra
22	Sumatra
24	Sumatra
27	Sumatra
28	Sumatra
33	Sumatra
17	Borneo
26	Borneo
31	Borneo
38	Borneo
1	West Malaysia
7	West Malaysia
13	West Malaysia
15	West Malaysia
19	West Malaysia
20	West Malaysia
23	West Malaysia

TABLE 10. Sumatran rhino variable nucleotide sites. (SS=Sumatran rhinos # 22, 24, 27, 28, 33; SS*=Sumatran rhino #6; SW=Sumatran rhinos #1, 7, 13, 15, 19, 20, 23; SB=Sumatran rhinos #17, 26, 31, 38).

		SS	SS*	SW	SB
Site#	133(12S)	C	C	G	C
	179(12S)	C	C	C	G
	194(12S)	C	C	C	G
	313(16S)	C	G	G	C

FIGURE 13. Distribution map for Sumatran rhinos from Penny (1988)



Map 6: Past and present distribution of the Sumatran rhinoceros

CONCLUSIONS

Now that the approach of using phylogenetic species as a guide for identifying units of conservation has been proposed (Vogler and DeSalle in press; Barrowclough and Flesness in press; Amato 1991; Wharton and Amato, in press), a need to clearly identify any weaknesses inherent in this approach is evident. Davis and Nixon (1992) suggest a methodology for assessing which attributes are phylogenetic characters and which are traits. This methodology (population aggregation analysis) involves successive searches for fixed differences between aggregations of local populations. Characters are attributes that are not polymorphic and are unique within populations. Traits are attributes that may be polymorphic or are not unique to a population. Davis and Nixon (1992) point out that cladistic methodology allows for the discovering of hierarchy in terminal units by scoring attributes, but that only characters, not traits, can be used for determining a hierarchy that has phylogenetic information. Characters are what identify phylogenetic species. Furthermore, they observe that small sample sizes can affect our ability to discriminate between a character and a trait. This is especially problematic for many endangered species that exist in small, fragmented, remnant populations (Amato and Wharton in press).

Cracraft (1991; and pers. comm.) has argued that the only way to objectively identify units of conservation is to use a phylogenetic species/lower level systematics approach. In the most rigorous application of this approach, assessing the significance of diagnostic characters would be subjective. While I feel that this approach should provide the framework for identifying units of conservation, I am concerned by some of the limitations. Of special concern are discriminating patterns of attributes in a species that now exists only in highly

fragmented, small populations in comparison to its historic distribution. In these cases, all of the individuals in a population may be closely related and share attributes that reflect those familial relationships. These attributes would not be informative about the population's evolutionary distinctiveness or history. For example, the black rhinoceros population was estimated to be over 60,000 in 1970 (already greatly reduced from that in 1900), and now is less than 2,500. To survey scattered populations that may number as few as ten animals for diagnostic characters may yield patterns that do not reflect evolutionary events (Amato et al. 1993). Character data may have to be discussed in the context of additional data from museum specimens representing areas where the animals have been locally extirpated. Additional data about the taxon's original distribution and densities may be used with mathematical models to predict the likelihood of loss for variously sampled traits (Amato and Powell in prep.).

While this may sound too subjective to some authors, I would argue that decisions on units of conservation are ultimately subjective. That does not prevent us from employing a rigorous framework from which to improve what admittedly will be a subjective decision. Davis and Nixon (1992) attempt to overcome some problems in identifying phylogenetic species with their assessment of attributes as traits or characters. However, Davis and Nixon (1992) acknowledge the problem of sample size in this analysis; and more importantly, acknowledge the subjective nature of operationally identifying what constitutes a population. If a rigorous, objective phylogenetic species definition rests on a subjective assessment of a population's boundary, can the units of conservation ever be objective?

While the approach outlined rests on a sound foundation of logic and rigorous science, it is clear that it suffers some limitations when applied to very closely related taxa, small data sets, and operationally defined

populations--precisely the situations we most frequently confront in conservation. However, it is still a less problematic approach than using overall similarity or genetic distances which have been shown to be largely uninformative for identifying taxonomic rank or units of conservation (Vogler and DeSalle in press; Avise and Aquadro 1982).

The study of the *Caiman crocodilus* complex in this thesis, demonstrates a case that is amenable to the outlined framework. By having a large number of samples from a group still present in much of its original range, many of the acknowledged problems are avoided. A wide distribution of diagnostic molecular characters that are congruent with biogeographical data support the division of the *C. crocodilus* complex into three phylogenetic species-conservation units.

The results from the Sumatran rhino study are more equivocal. Individual sites are not diagnostic, but suites of sites (haplotypes) are diagnostic if the two Sumatran haplotypes are considered character states. Additionally, this study is complicated by the fact that the remaining animals represent a tiny fraction of the species original distribution. Identifying evolutionary patterns is difficult due to missing individuals, rather than missing data from the individuals still present. However, the available results do not support the designation of more than one conservation unit if our goal is to preserve the evolutionary novelty present in the Dicerorhinus lineage.

Higher level studies are increasingly important in understanding conservation units problems. The higher level studies in this dissertation demonstrate the usefulness of 12S and 16S mitochondrial sequences for phylogenetic studies. Identification of higher level relationships can also allow us to root specific taxa to elucidate lower level phylogenetic patterns.

The difficulty in identifying conservation units has, in a large part, been

due to the notion that the conservative approach is to maximally "split" and only reluctantly "lump". The result is paralysis in moving from research and data to implementing necessary management decisions. As stated in chapter 4, the problem with indefinitely postponing management decisions while awaiting additional studies to "prove" there are no differences between purported subspecies, becomes an epistomolgical problem rather than a scientific one. This problem has been pervasive in conservation. Templeton's paper on "Coadaptation and Outbreeding Depression" (1986) is the most often cited argument for this continued "conservative" approach. However, in this paper, Templeton (1986) largely discounts the likelihood of outbreeding depression resulting in the loss of a managed species. In fact he uses the opportunity to argue for a much more "progressive" approach that I strongly concur with. Templeton (1986) suggests that our goal should be "preserving evolutionary lineages", not a current endangered "constellation of present-day traits that defines a rigid category we call species".

REFERENCES

Amato, G.D. and J. Gatesy. (in press). PCR assays of variable nucleotide sites for identification of conservation units. In Molecular Approaches to Ecology and Evolution. Schierwater, B., B. Streit, G. Vagner, and R. DeSalle eds.

Amato, G.D. and D. Wharton. (in press). A systematic approach to identifying units of conservation: Examples of progress and problems. Proceedings of 1993 American Association of Zoos and Aquariums Conference, Omaha, Nebraska.

Amato, George D. 1991. Species hybridization and protection of endangered animals. Science 253 (5017):250.

Amato, George D., Mary V. Ashley, and John Gatesy. 1993. Molecular evolution in living species of rhinoceros:Implications for conservation. Proceedings of the International Conference on Rhinoceros Conservation and Biology. O.A. Ryder editor. Zoological Society of San Diego.

Avise, J.C. and R.M. Ball. 1992. Principles of genealogical concordance in species concepts and biological taxonomy. in Oxford Surveys in Evolutionary Biology. Antonovics, J. and D. Futuma eds. Oxford Univ. Press. London.

Avise, J.C. and C.F. Aquadro. 1982. A comparative summary of genetic distances in the vertebrates. Evolutionary Biology 15:151-185.

Ayala, F.J. 1976. Molecular Evolution. Sinauer Associates, Sunderland Mass.

Barrowclough, G.F. and N.R. Flesness. 1993. Species, subspecies, and races:The problem of units of management in conservation. In M. Allen and H. Harris (eds.) Wild Mammals in Captivity. University of Chicago Press. In press.

Brazaitis, P. 1973. The identification of living crocodilians. Zoologica. 58:59-101.

Brazaitis, P., G. Rebelo, and C. Yamashita. 1990. A summary report of the CITES central South American caiman study: Phase I: Brazil. In Crocodiles, Proceedings of the 9th Working Meeting of the Crocodile Specialist Group, IUCN. Gland< Switzerland. 1:100-115.

Brazaitis, P., G. Robelo, and C. Yamashita. 1992. Report of the WWF/Traffic USA Survey of Brazilian Amazonia Crocodilians: Survey period July 1988 to January 1992.

Caccone, A., G.D. Amato and J.R. Powell. 1987. Intraspecific DNA divergence in Drosophila: a case study in parthenogenetic D. mercatorum. Mol. Biol. Evol. 4:343-350.

Caccone, A., and J.R. Powell. 1989. DNA divergence among hominoids. Evolution 43:925-942.

Carvalho, A.L. 1955. *Os Jacares do Brasil.* Arquivos do Museu Nacional (Rio de Janeiro). 42 (1):125-139.

Cracraft, J.A. 1983. Species concept and speciation analysis. Current Ornithology 1:159-187.

Cracraft, J.A. and R. Prum. 1988. Patterns and processes of diversification: Speciation and historical congruence in some Neotropical birds. Evolution 42: 603-620.

Cracraft, J.A. 1989. Speciation and ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Pp. 28-59 in: D. Otte and J. A. Endler (eds.) Speciation and its Consequences. Sinauer Associates, Sunderland, MA.

Cracraft, J. 1991. Systematics, species concepts, and conservation biology. Abstract. Society for Conservation Biology, 5th Annual meeting

Daudin, F. 1802. *Histoire naturelle, generale et particuliere des reptiles.* Paris. Vol. 2, pp 399.

Davis, J.I., and K. C. Nixon. 1992. Populations, Genetic variation, and the delimitation of phylogenetic species. Systematic Biology vol.41 4:421-435.

Densmore, L.D. 1983. Biochemical and immunological systematics of the order Crocodilia. Evolutionary biology, Vol. 16. Hecht, M.K., B. Wallace, and G.H. Prance eds. Plenum Press, New York, New York pp. 397-465.

Densmore, L.D. and H.C. Dessauer. 1984. Low levels of protein divergence detected between Gavialis and Tomistoma: evidence for crocodilaian monophyly? Comp. Biochem. Physiol. 77B:715-720.

Densmore, L.D., and R.D. Owen. 1989. Molecular systematics of the order Crocodilia. Amer. Zool. 29:831-841.

DeSalle, R., A.R. Templeton, I. Mori, S. Pletscher, and J.S. Johnson. 1987. Temporal and spatial heterogeneity of mtDNA polymorphisms in natural populations of *Drosophila mercatorum*. Genetics, 116:215-223.

Dixon, J. 1979. Origin and distribution of reptiles in lowland tropical rainforests of South America. In, The South American Herpetofauna: Its Origin, Evolution and Dispersal. W.E. Duellman ed. Museum of Natural History. University of Kansas Press. Lawrence Ka. Mon. 7. pp 217-240.

Dobzhansky, Th. 1951. Genetics and the Origin of Species. Columbia University Press, New York.

Ellerman, J.R. and T.C.S. Morrison-Scott. 1951. Checklist of Palaearctic and Indian mammals 1758 to 1946 (British Museum Nat. Hist., London).

Farris, J.S. 1989. The retention index and the rescaled consistency index. Cladistics 5:417-419.

Felsenstein, J. 1990. PHYLIP: Phylogeny Inference Package, version 3.3, University of Washington, Seattle, WA.

Frair, W., and J. Behler. 1983. Book Review: *Liste Der Rezenten Amphibian Und Reptilian. Testudines, Crocodylia, Rhynchocephalia* by H. Wermuth and R. Mertens. Herp. Review. 14(1):23--25.

Fuchs, K. 1974. *Die Krokodilhaute; Ein wichtiger Merkmaltrager bei der Identifizierung von Krokodil-Arten.* Eduard Roether Verlag. Darmstadt, Germany. 183pp.

Gatesy, J., and G.D. Amato. 1992. Sequence similarity of 12S ribosomal segment of mitochondrial DNAs of gharial and false gharial. Copeia. (1):241-243.

Gatesy, J., D. Yelon, R. DeSalle, and E.S. Verba. 1992. Phylogeny of the Bovidae (Artiodactyla, Mammalia), based on mitochondrial ribosomal DNA sequences. Mol. Biol. Evol. 9(3):433-466.

Gatesy, J., R. DeSalle, and W. Wheeler. 1993. Alignment-ambiguous nucleotide sites and the exclusion of systematic data. Molec. Phylo. Evol. vol.2 2:152-157.

Garza, J.C., and D.S. Woodruff. 1992. A phylogenetic study of the gibbons (Hylobates) using DNA obtained noninvasively from hair. Molecular Phylogenetics and Evolution vol. 1 3:202-210.

Gloger, C.L. 1833. *Das Abandern der Vogel durch Einfluss des Klimas* (Breslau).

Groves, C.P. 1993. Testing rhinoceros subspecies by multivariate analysis. Proceedings of the International Conference on Rhinoceros Conservation and Biology. O.A. Ryder editor. Zoological Society of San Diego, pp. 92-97.

Groves, C.P. 1983. Phylogeny of the living species of rhinoceros. *Zeitschrift fur Zoologische Systematik und Evolutionforshung* 21 (4):293-313.

Groves, C.P. 1967a. The rhinoceroses of Southeast Asia. *Saugetierkundl. Mitt.* 15:221-237.

Guerin, C. 1982. *Les Rhinocerotidae (Mammalia, Perissodactyla) du Miocene terminal au Pleistocene superieur en Europe occidentale compares aux especes actuelles: tendances evolutives et relations phylogenetiques. Geobios* 15 (4): 599-605.

Gyllensten, U.B., and H.A. Erlich. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc. Natl. Acad. Sci. U.S. 85:7652-7656.

Hennig, W. 1966. Phylogentic Systematics. Univ. of Illinois Press, Urbana.

Houck, M.L., O.A. Ryder, J. Vahala, R.A. Kock, and J.E. Oosterhuis. 1994. Diploid chromosome number and chromosomal variation in the white rhinoceros (*Ceratotherium simum*). Journal of Heredity 85:30-34.

Hudson, R.R. and N.L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147-164.

Huxley, J.S. 1940. ed., The New Systematics. Clarendon Press, Oxford.

Irwin, D.M., T.D. Kocher, and A.C. Wilson. 1991. Evolution of the cytochrome b gene of mammals. J. Mol. Evol. 32:128-144.

Janczewski, D., W.B. Karesh, H. Frazier-Taylor, D. Sajuthi, M. Andau, and F. Gombek. 1991. Genetic characterization of free ranging orang utans of Borneo and Sumatra. Proceedings of the Great Apes Conference, Conservation of the great apes in the new world order of the environment, section 6.3.5.

Khan, M.K.M., B.H.M. Nor, E. Yusof, and M.A. Rahman. 1993. *In-situ* conservation of the Sumatran rhinoceros (*Dicerorhinus sumatrensis*) A Malaysian experience. Proceedings of the International Conference on Rhinoceros Conservation and Biology. O.A. Ryder editor. Zoological Society of San Diego, pp. 238-247.

King, F.W., and R.L. Burk. 1989. Crocodilian, tuatara, and turtle species of the world: A taxonomic and geopraphic refference. Association of Systematics Collections. Washington D.C. pp.1-15.

King, F.W., and V. Roca. 1987. The caimans of Bolivia. A preliminary report on a CITES and Centro de Dessarrollo Forestal sponsored survey of species distribution and status. 38 pp, App.i-ii, I-II. Report to CITES Secretariat, Lausanne, Switzerland.

Kleinschmidt, O. 1900. *Arten oder Formenkreise?* J. Ornithology. 48:134-139.

Kocher, T.C., W.K. Thomas, A. Meyer, S.V. Edwards, S. Paabo, F.X. Villablanca and A.C. Wilson. 1989. Dynamics of mitochondrial evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA 86:377-382.

Lessa, E.P. and G. Applebaum. 1993. Screening techniques for detecting allelic variation in DNA sequences. Mol. Ecol. 2:119-129.

Linnaeus, C. 1758. *Systema Naturae. Regnum Animale.* (10th ed. tomus I; L. Salvii, Holminae).

Mayr, E. 1963. Animal Species and Evolution. Oxford University Press, London.

Mayr, E. 1951. Speciation in birds. Proceedings of the Xth Intern. Ornithol. Congr., Uppsala, 1950, 91-131.

Mayr, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York.

Medem, F. 1960. Notes on the Paraguay caiman, *Caiman yacare* Daudin. Mitt. aus dem Zool. Mus. in Berlin. 36(1):129-142.

Medem, F. 1983. *Los Crocodylia de Sur America. Universidad Nacional de Colombia and Colciencias.* Bogota. Clombia. Vol. 2, 270 pp.

Medem, F., and H. Marx. 1955. An artificial key to the New-World species of crocodilians. Copeia. 1:1-2.

Mook, C.C. 1934. The evolution and classification of the Crocodilia. J. Geol. 42:295-304.

Myers, R.M., T. Maniatis, and L.S. Lerman. 1986. Detection and localization of single base changes by denaturing gradient gel electrophoresis. Meth. in Enzym. 155:501-527.

Myers, R.M., V.C. Sheffield, and D.R. Cox. 1989a. Mutation detection, GC-clamps, and denaturing gradient gel electrophoresis. In: PCR technology: Principles and applications for DNA amplification. (ed. H.A. Erlich). pp. 71-88. Stockton Press, New York.

Myers, R.M., V.C. Sheffield, and D.R. Cox. 1989b. Polymerase chain reaction and denaturing gradient gel electrophoresis. In: Polymerase chain reaction (H.A. Erlich, R. Gibbs, and H.H. Kazazian eds.) pp. 177-181. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Nelson, G. J., and N.I. Platnick. 1981. Systematics and biogeography: cladistics and vicariance. Columbia University Press, New York.

Newton, C.R., A. Graham, I.E. Heptinstall, S.J. Powell, C. Summers and N. Kalsheker. 1989. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res. 17:2503-2515.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nichols, W.C., J.J. Liepnicks, V.A. McKusick and M.D. Benson. 1989. Direct sequencing of the gene for Maryland/German familial amyloidotic polyneuropathy type II and genotyping by allele-specific amplification. Genomics 5:535-540.

Nixon, K.C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics 6:212-223.

Norell, M.A. 1988. Cladistic approaches to evolution and paleobiology as applied to the phylogeny of alligatorids. Ph.D. thesis, Yale University, New Haven, CT.

Norell, M.A. 1989. The higher level relationships of the extant Crocodylia. J. Herp. 23:325-335.

O'Brien, S.J. and E. Mayr. 1991. Bureaucratic mischief: recognizing endangered species and subspecies. Science 251:1187-1188.

O'Brien, S.J. 1990. National Geographic Research 6:485.

Okayama, H., D.T. Curiel, M.L. Brantly, M.D. Holmes and R.G. Crystal. 1989. Rapid nonradioactive detection of mutations in the human genome by allele-specific amplification. J. Lab. Clin. Med. 114:105-113.

Orita, M., H. Iwahana, H. Kanazawa, K. Hayashi, and T. Sekiya. 1989. Detection of polymorphisms of human DNA by gel electrophoresis and single-strand conformation polymorphisms. Proc. Nat. Acad Sci. USA. 86:2766-2770.

Paterson, H.E.H. 1985. The recognition concept of species. pp 21-29 in Species and Speciation, Vrba,E.S. ed. Transvaal Museum Monograph No. 4, Pretoria.

Penny, M. 1988. Rhinos: endangered species. Facts On File publications, New York, New York.

Powell, J.R. 1991. Monophyly/paraphyly/polyphyly and gene/species trees: an example from *Drosophila*. Mol. Biol. Evol. 8(6):892-896.

Prothero, D.R. 1993. Fifty million years of rhinoceros evolution. Proceedings of the International Conference on Rhinoceros Conservation and Biology. O.A. Ryder editor. Zoological Society of San Diego, pp. 82-89.

Prothero, D.R., C. Guerin, and E. Manning. 1989. The history of the Rhinocerotoidea, in Prothero, D.R. and R.M. Schoch, eds., The Evolution of Perissodactyls. New York, Oxford University Press, pp. 322-340.

Prothero, D.R., E. Manning, and C.B. Hanson. 1986. The phylogeny of the Rhinocerotoidea (Mammalia: Perissodactyla). Zoological Journal of the Linnean Society 87: 341-366.

de Queiroz, K. and M.J. Donoghue. 1990. Phylogenetic systematics and the phylogenetic species concept. Cladistics 4:317-338.

Remington, C.L., M.M. Cary, A.B. Klots, B.P. Beirne, E. Munroe, and L.P. Gray. 1951. Geographic subspeciation in the Lepidoptera: a symposium. Lepidopterists' News 5:17-35.

Rensch, B. 1929. *Das Prinzip geographischer Rassenkreise und das Problem der Artbildung* (Borntraeger, Berlin).

Rojas, M. 1992. The species problem and conservation: what are we protecting? Conser, Biol. 6:170-178.

Rothschild, W., E. Hartert, and K. Jordan. 1894. Note of the editors. *Novit Zool.* 1:1.

Ruano, G., and K.K. Kidd. 1991a. Coupled amplification and sequencing of genomic DNA. Proc. Nat. Acad. Sci. USA. 88:2815-2819.

Ruano, G., and K.K. Kidd. 1991b. Genotyping and haplotyping of polymorphisms directly from genomic DNA via coupled amplification and sequencing (CAS). Nuc. Acids Res. 19:6877-6882.

Santiapilla, C. and K. MacKinnon. 1993. Conservation and management of Sumatran rhino (*Dicerorhinus sumatrensis*) in Indosnesa. Proceedings of the International Conference on Rhinoceros Conservation and Biology. O.A. Ryder editor. Zoological Society of San Diego, pp. 257-263.

Selander, R.K. 1976. Genetic variation in natural populations. In Molecular Evolution, Ayala, F.J. ed. pp 21-45. Sinauer Associates, Sunderland, MA.

Simpson, B. 1979. Quaternary biogeography of the high montane regions of South America. In, The South American Herpetofauna: Its Origin, Evolution and Dispersal. W.E. Duellman ed. Museum of Natural History. University of Kansas Press. Lawrence Ka. Mon. 7. pp 107-188.

Sommer, S.S., A.R. Groszbach, and C.D.K. Bottema. 1992. PCR amplification of specific alleles (PASA) is a general method for rapidly detecting known single-base changes. BioTechniques 12:82-87.

Swofford, D.L. 1990. PAUP: Phylogenetic Analysis using Parsimony, version 3.0.Computer program distributed by Illinois Natural History Survey. Champaign IL.

Takahata, N. and S.R. Palumbi. 1986. Extranuclear differentiation and gene flow in the finite island model. Genetics 109: 441-457.

Tarsitano, S.F., E. Frey, and J. R. Riess. 1989. The evolution of the Crocodilia: a conflict between morphological and biochemical data. Amer. Zool. 29:843-856.

Templeton, A.R. 1989. The meaning of species and speciation: a genetic perspective. pp 3-27 In Speciation and Its Consequences, Otte, D. and J.A. Endler eds. Sinauer Associates, Sunderland, MA.

Templeton, A.R. 1986. Coadaptation and outbreeding depression. Conservation Biology: The Science of scarcity and diversity. M.E. Soule ed. Sinauer Ass. Inc., Sunderland, MA. pp. 105-116.

Vaurie, C. 1955. Pseudo-subspecies. Acta XI Congr. Intern. Ornitol., Basel. 1954, 369-380.

Vogler, A.P. and R. DeSalle. (in press) Diagnosing units of conservation management. Conser. Biol.

Voipio, P. 1950. Evolution at the population level with special reference to the game animals and practical game management. Papers Game Research, Helsinki 5:1-176.

Vrana, P. and W. Wheeler. 1992. Individual organisms as terminal entities:laying the species problem to rest. Cladistics 8:67-72.

Walsh, P.S., D.A. Metzger and R. Higuchi. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 10(4):506-513.

Wartell, R.M., S.H. Hosseini, and C.P. Moran. 1990. Detecting base pair substitutions in DNA fragments by temperature gradient gel electrophoresis. Nuc. Acids Res. 18:2699-2705.

Wermuth, H. and R. Mertens. 1977. Testudines, Crocodylia, and Rhyncocephalia. Das Tierreich 100:1-174.

Wheeler, W., R. DeSalle, and J. Gatesy. (in press). Elision; a method for accomodating multiple sequence alignments. Molec. Phylog. Evol.

Wheeler, W. and D. Gladstein. 1991. MALIGN: Program and documentation-version 1.5, American Museum of Natural History, New York, New York.

Whitten, A.J, S.J. Damanik, J. Anwar, and N. Hisyam. 1987. The Ecology of Sumatra. Yogyakarta; Gadjah Mada University Press.

Wu, W.Y., L. Ugozzoli, B.K. Pal, and R.B. Wallace. 1989. Allele-specific enzymatic amplification of β-globin genomic DNA for diagnosis of sickle cell anemia. Proc. Nat. Acad. Sci. USA 86:2757-2760.