83

THE BIOCHEMICAL ANALYSIS OF SOUTHERN AFRICAN RHINOCEROS POPULATIONS

COLLEEN O'RYAN

Thesis submitted in fulfillment of the requirements for Doctor of Philosophy degree in the Department of Chemical Pathology in the Faculty of Medicine at the University of Cape Town

March 1993

ACKNOWLEDGEMENTS

I wish to express my gratitude to:

My supervisor, Professor Eric Harley, for his proficient guidance and enthusiastic support.

I. Baumgarten who assisted with the maintenance and curation of the rhinoceros cell cultures.

I. Espie, A. Hall-Martin, M. Lindeque, P. Morkel, C. Raath, P. Rogers and T. Sandwith for the collection of rhinoceros samples, and the Natal Parks Board, the National Parks Board and the Namibian Ministry of National Conservation and Tourism for logistical assistance.

My colleagues for their assistance and providing a congenial working atmosphere.

My friends for proofreading my thesis.

My parents, for allowing me the freedom to pursue my academic goals.

ABSTRACT

The drastic decline in the numbers of the five extant species of rhinoceroses world-wide, mainly as a result of poaching, have placed these species in imminent danger of extinction. This emphasizes the need to understand the relationships among the different species of rhinoceros. The advances in molecular biology have allowed the application of DNA-based genetic techniques to address a number of aspects of rhinoceros biology which have both academic interest and practical value to conservation management.

There are four aspects to this study: Firstly, restriction endonuclease maps of mitochondrial DNA were constructed to estimate the time of divergence of *Diceros bicornis* (black rhinoceros) and *Ceratotherium simum* (white rhinoceros) from their common ancestor. Secondly, a population genetic study of the relationships among four subspecies of *D. bicornis*. Thirdly, the application of DNA fingerprinting to examine the intra- and inter-population relatedness in *D. bicornis* populations. Fourthly, a practical application of PCR to identify the origin of an unknown sample of DNA.

Restriction maps were constructed for 18 restriction endonucleases which were aligned on two Sac II sites and a Hpa I site which are invariant in almost all vertebrates. The proportion of shared sites (0.667) from which the sequence divergence (6.79%) was calculated translates to a time of divergence of 3.4 (+/-0.8) million years, assuming a calibration of sequence divergence against time of 2% per million years. In addition to the above, a 300 base pair fragment of the cytochrome-b region of mitochondrial DNA was amplified and sequenced. The sequence divergence between black rhinoceros and white rhinoceros was calculated from this sequence information.

There is much controversy about the validity of the various subspecific designations currently applied to the black rhinoceros. To address this question, the restriction endonuclease maps of four Southern Africa subspecies of black rhinoceros were studied from 33 individual representing four subspecies: *D. b. minor*, *D. b. bicornis*, *D. b. michaeli* and *D. b. chobiensis*. The DNA restriction maps constructed for each subspecies, using 16 restriction enzymes, showed only two differences between *D. b. minor* and *D. b. bicornis*, between *D. b. minor* and *D. b. michaeli*, and between *D. b. michaeli* and *D. b. bicornis*. The maps were monomorphic within subspecies and therefore, provide effective diagnostic markers. The map from the single

individual of *D. b. chobiensis* was identical to that of *D. b. minor*.

The application of DNA fingerprinting was explored as to whether it could be usefully employed to determine within and between population variation in four black rhinoceros subspecies. Even with a relatively simple statistical approach, the results obtained suggest the technique will have practical value for monitoring the loss of genetic diversity in small wild populations.

A practical application of this study was the novel use of the polymerase chain reaction to determine the species of origin in trace amounts of biological material. Mitochondrial DNA was amplified and sequenced from blood and tissue found on an axe, thought to have been used in the poaching of rhinoceros. The DNA sequences obtained from the axe were compared to mitochondrial DNA sequences from black rhinoceros and white rhinoceros, as well as to bovine DNA sequences. The unknown sample was unequivocally identified as being of bovine origin.

These results demonstrate the usefulness of DNA-based techniques in addressing a number of academic and practical questions of rhinoceros biology.

LIST OF ABBREVIATIONS

dATP (A)	Deoxyadenosine triphosphate
dCTP (C)	Deoxycytidine triphosphate
dGTP (G)	Deoxyguanosine triphosphate
dTTP (T)	Deoxythymidine triphosphate
^{32}P-dCTP	Deoxycytidine triphosphate, radioactively labeled with ^{32}Phosphorus (on the alpha phosphate)
^{32}S-d-ATP	Deoxyadenosine triphosphate, radioactively labeled with ^{35}Sulphur (on the alpha phosphate)
kb	kilobase pair
bp	base pair
DNA	Deoxyribonucleic acid
mtDNA	mitochondrial deoxyribonucleic acid
D-loop	Displacement loop
PCR	Polymerase chain reaction
UV	Ultra violet
MLP	multi-loci probe
SLP	single locus probe
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism

UNITS

k, m, u, n,	kilo-, milli-, micro-, nano- (prefixes)
m, l, g	metre, litre, gram
Ci	Curie
^{0}C	Degrees Celsius
D	Daltons
mol	Moles
M	Molar
N	Normal
g	Centrifugal force
V	Volts

LIST OF TABLES

PAGE

TABLE

2.1:A list of the restriction enzymes mapped and the codes used. 45

4.1:The primers used to amplify the cytochrome-b region of mtDNA. 82

5.I:Samples used for DNA extraction and amplification. 106

6.1:Five extant species of the family Rhinoceroditae. 117

6.2:The seven subspecies of black rhinoceros identified on the basis of morphological characteristics. 125

6.3:A summary of the mtDNA restriction enzyme profiles. 133

6.4:Summary of the origin of the black rhinoceros subspecies sampled. 134

8.1:Summary of the band sharing coefficient
between the different individuals of
the four subspecies. 165

8.2:Summary of the average band sharing coefficient
between the four different subspecies. 166

LIST OF FIGURES

PAGE

FIGURE

1.1:The schematic representation of mammalian mtDNA. 6

1.2:The time over which mitochondrial DNA and nuclear DNA are useful for comparisons 10

1.3:The rate of DNA substitution in (a) mtDNA and (b) single copy DNA. between species or genera. 13

1.4:The geographical separation of the horseshoe crab, *Limulus polyphemus*, mtDNA morphs along the eastern coast of the United States. 19

1.5:An example of the geographical mtDNA distribution in the rock hyrax, *Procavia capensis*. 20

1.6:The mtDNA distribution of North Atlantic eels. 21

2.1:The standard curve of lambda Hind III marker DNA plotted on semi-logarithmic graph paper. 35

2.2:The main menu screen of the RESOLVE program. 38

2.3:A typical display of the RESOLVE program's file status. 39

2.4:A display of restriction enzyme (Bam HI and Bgl II) data being entered into the RESOLVE program. 40

2.5:A display of the fragment analysis and comparison by the RESOLVE program. 41

2.6:The solution for the Bam HI and Bgl II digestion. 42

2.7:An example of unfixed sites within a fragment of DNA. 43

2.8:Single digest patterns of black rhinoceros and white rhinoceros. 46

2.9:Double digestion patterns for black rhinoceros. 47

2.10:The restriction enzyme maps of black rhinoceros and white rhinoceros aligned on the invariant Sac II site at position 676 in *Bos taurus*. 48

2.11:Restriction enzyme alignments for individual restriction enzymes. 49

2.12:Restriction enzyme digest of white rhinoceros. 51

3.1:A schematic representation of the polymerase chain reaction (PCR). 57

3.2:The triple method for sequencing region V of human DNA. 67

3.3:An outline of the procedure for generating ssDNA by PCR. 69

3.4:The accumulation of dsDNA and ssDNA at primer ratios of 100:1. 70

4.1:A sequencing gel after the amplified DNA was purified using: 1. PAGE 2. Millipore filter unit. 93

4.2:A sequencing gel after the amplified DNA was purified using QIAEX. 94

4.3:A sequencing gel after the amplified DNA of black rhinoceros was purified using QIAEX. 95

4.4:A sequencing gel of black rhinoceros and white rhinoceros DNA. 97

4.5:The aligned DNA sequences of: 1.The reference black rhinoceros 2. black rhinoceros 3. white rhinoceros 4. the reference *Bos taurus*. 98

4.6:Alignments of the amino acid sequences derived from: 1. The reference black rhinoceros 2. black rhinoceros 3.white rhinoceros 4. the reference *Bos taurus*. 106

5.1:Sequencing gel of axe DNA directly directly after purification on a Millipore filter unit. 110

5.2:Sequencing gel showing the DNA sequences from three different DNA clones obtained from the axe. 112

5.3:Sequence alignment of the eight axe DNA clones with the *Bos taurus* and the black rhinoceros reference sequences from Genbank. 113

6.1:The distribution of *Diceros bicornis*. 118

6.2:The distribution of *Ceratotherium simum*. 119

6.3:A graphical representation of the loss of genetic diversity in small populations. 122

6.4:The distribution of *Diceros bicornis* subspecies based on morphological characteristics. 123

6.5:Restriction enzyme maps of the four subspecies of black rhinoceros. 135

6.6:A gel showing the monomorphic restriction enzyme pattern of the black rhinoceros subspecies. 137

6.7:The diagnostic patterns that differentiate the four subspecies of black rhinoceros. 138

7.1:A model for minisatellite generation which is promoted by chi-like core sequences. 145

8.1:The map of M13 DNA. 154

8.2:The map of pGEM-3. 155

8.3:Fingerprinting pattern obtained after hybridization of $(CAC)_5$ probe to rhinoceros DNA. 162

8.4:Fingerprinting pattern obtained after hybridization of M13 probe to rhinoceros DNA. 163

SECTION A

COMPARISON OF THE MITOCHONDRIAL DNA IN *DICEROS BICORNIS MINOR* (BLACK RHINOCEROS) AND *CERATOTHERIUM SIMUM SIMUM* (WHITE RHINOCEROS)

INTRODUCTION

Although they are nearly extinct today, rhinoceroses were one of the most widespread and successful groups of mammals for over forty million years. The distinguishing feature of a rhinoceros are its horns. The horn is not only the source of the name "rhinoceros", but also the chief reason this animal has been hunted to the brink of extinction. The rapid decline of rhinoceros populations emphasizes the need to understand the relationships among the different species of rhinoceros, and how best to manage the remaining populations. Although fossil material of rhinoceroses is fairly well represented, the systematics of this group is not well studied.

With the advances in molecular biology in the past decade, a large number of new techniques are now available to the

evolutionary biochemist when constructing phylogenetic relationships among different groups of animals. The analysis of mitochondrial (mt) DNA is specifically relevant to molecular taxonomy. MtDNA analyses are best suited for elucidating relationships between groups of animals that diverged from a common ancestor 0.5 to 20 million years ago (Brown *et al* 1979).

MtDNA was used to address a number of features of rhinoceros evolutionary genetics. The time of divergence between black rhinoceros and white rhinoceros was estimated from mtDNA restriction endonuclease maps and DNA sequence data. The polymerase chain reaction was used to amplify the cytochrome-b region of mtDNA, after which the DNA was sequenced.

CHAPTER 1

THE USE OF MITOCHONDRIAL DNA IN EVOLUTIONARY GENETICS

1.1 INTRODUCTION

In the study of evolution, there are two tasks. One is to elucidate the phylogenetic histories of various organisms and the other is to understand mechanisms of evolution. Palaeontologists, embryologists and systematists have traditionally studied the first problem and population geneticists the second. Until the introduction of molecular techniques in the mid-1960s, the above questions have been addressed by examining fossil records, morphological and physiological characteristics. The debate over whether molecular or morphological characteristics are inherently better sources of information remains unresolved (Patterson, 1987). Comparative studies have shown that morphological changes and molecular divergence are independent and respond to different evolutionary pressures (Hillis & Moritz, 1990).

1.2 EARLIER MOLECULAR TECHNIQUES

The earliest molecular techniques employed the differences in the structure of proteins to unravel some of the questions facing evolutionary biologists. Hubby and Lewontin (1966) were among the first to utilize gel electrophoresis of proteins to examine genetic variation. Gel electrophoresis is a technique which examines protein structure by comparing the migration of proteins through an electrical field. This technique is relatively simple and cheap, however, only about 30% of the detectable loci are variable (Hartl & Clark, 1989).

Although protein electrophoresis has some limitations for example, low levels of variation in some organisms, different rates of mutation at different loci and inherent sampling limitations (Hillis & Moritz, 1990), it is still a very useful technique. Population structure (Eanes & Koehn, 1978), bottlenecks (Parkin & Cole, 1985), inbreeding (O'Brien *et al.*, 1987), dispersal (Brown, 1985), species boundaries (Wake *et al.*, 1983) and hybridization (Lamb & Avise, 1976) can all effectively be studied using protein electrophoresis.

Other molecular techniques include immunological methods, cytogenetics and DNA:DNA hybridization (Britten, 1986 and

Hillis & Moritz, 1990). Immunological techniques can be used to quantitatively estimate the extent of immunological cross-reactivity between proteins from different taxa (Thorpe, 1982). DNA:DNA hybridization is based on the assumption that the genetic relatedness of organisms is reflected in the similarity of their DNA base pair sequence. This can be measured by hybridizing strands of DNA from different species and the melting point of these duplexes can then be determined. The thermostability of the hybrid duplex DNA relative to that of the homoduplex DNA, serves as an index of divergence between two species (Sheldon, 1987).

1.3 MITOCHONDRIAL DNA

1.3.1 GENOMIC ORGANISATION

In recent years there has been a remarkable increase in the application of DNA-based techniques to the analysis of problems in population genetics and systematics. Animal mtDNA has become one of the most extensively studied components of eukaryotic DNA. Its buoyant density in the presence of EtBr and high copy number render mtDNA easy to purify. (Brown & Vinograd, 1974 and Lansman, *et al.*, 1981). MtDNA is distributed universally in the animal kingdom and is uniform in gene content, although not in gene order (Palevsky *et al.*, 1988, Clark-Walker, 1989 and Johansen *et*

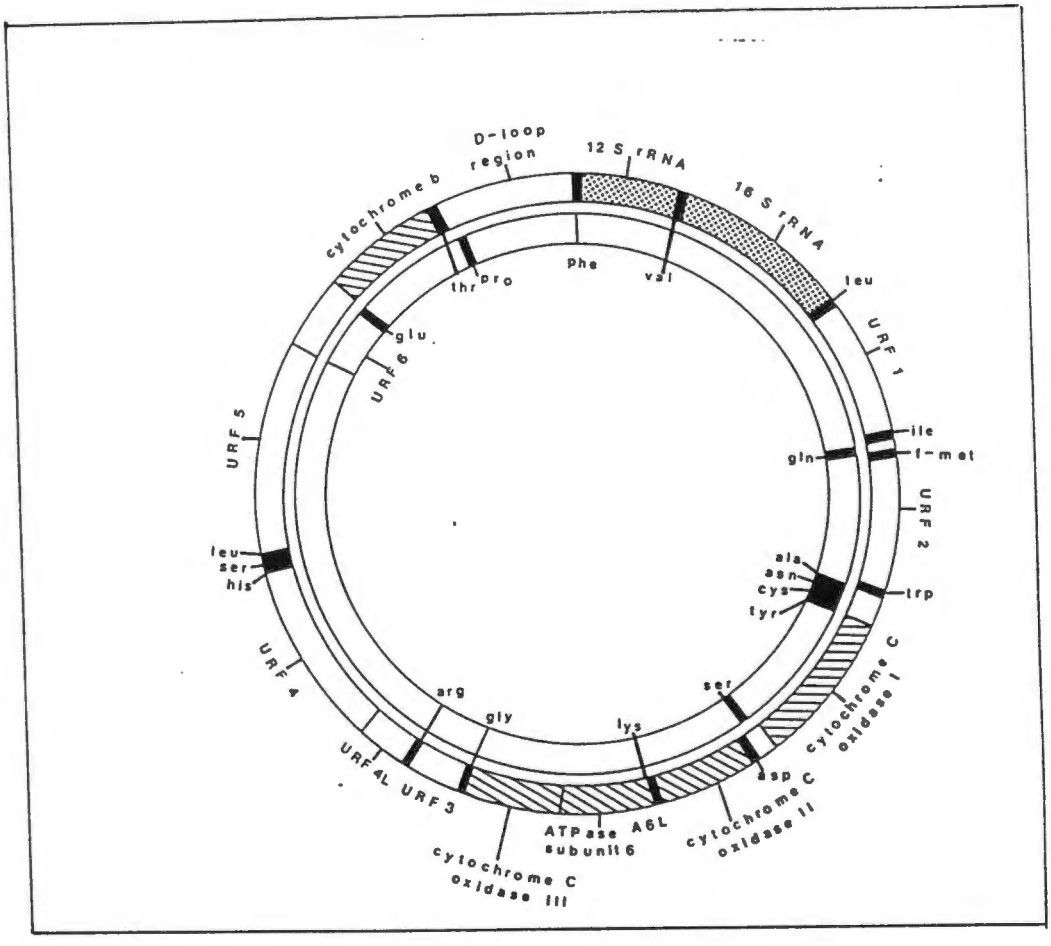

FIGURE 1.1

The schematic representation of mammalian mitochodrial DNA. The URFs code for components of NADH dehydrogenase (Reproduced from Clayton *et al.*, 1984)

al., 1990). Mammalian mtDNA is a covalently closed circular molecule that is 5um long, contains approximately 16 500 base pairs (bp) and has a molecular mass of 10^7 daltons (Giles *et al.*, 1980).

MtDNA of multicellular animals consists of two ribosomal RNA genes, 22 transfer RNA genes and 13 genes that code for proteins involved in electron transport or ATP synthesis. Each molecule has a control region or the displacement-loop (D-loop) region (figure 1.1). This contains sequences that function in the initiation of replication and transcription (Clayton, 1984). The gene order and general structure of mammalian mtDNA is highly conserved and has been found to be identical for those species sequenced: human (Anderson *et al.*, 1981), bovine (Anderson *et al.*, 1982), mouse (Bibb *et al.*, 1981), rat (Gadaleta *et al.*, 1981) and fin whale (Arnason *et al.*, 1991). The mtDNA molecules in plants and fungi are larger in size (typically 150 000 bp) and more complex in gene organisation (Hillis & Moritz, 1990).

1.3.2 MODE OF INHERITANCE

Animal mtDNA is haploid and maternally inherited (Gyllensten *et al.*, 1985). A typical mammalian somatic cell contains hundreds to thousands of mtDNA molecules. An ovum contains 10^5 molecules, while in the midpiece of sperm there

are only about 50 mtDNA molecules (Wilson *et al.*, 1985). Most studies of possible paternal inheritance of mtDNA have produced negative results (Giles *et al.*, 1980), but low level (one molecule in 1000) leakage of paternal sequences of mtDNA have been detected in experimental backcross hybrids of mice (Gyllensten *et al.*, 1991) and in hybrids of marine mussels (Zouros *et al.*, 1992).

The maternal inheritance of mtDNA has important consequences. Only a fraction of a population will pass on its mtDNA, so that the effective population size of mtDNA is smaller than that for nuclear genes (Harrison, 1989). Because mtDNA reflects matriarchal phylogeny, it is an excellent marker of patterns of colonization events, including founder effects (Glen-Hall & Muralidharan, 1989 and Smith, 1991). Recombination does not occur in mammalian mtDNA, with the exception of the displacement loop (Wilson *et al.*, 1985). This lack of recombination enhances the attractiveness of mtDNA as a genealogical tool.

1.3.3 HETEROPLASMY

There are many copies of mtDNA molecules in a cell and it is inevitable that new mutations will periodically generate different molecules so that the individual will have two or more genotypes of mtDNA co-existing. This condition results

in a heteroplasmic population of mtDNA molecules in that individual (Ashley *et al.*, 1989 and Fos *et al.*, 1990). It was originally thought that heteroplasmy might be extensive, but empirical experience has proved this not to be the case, with heteroplasmy being uncommon. The majority of individuals tested appear effectively haploid. Presumably, this is as a result of it becoming fixed for one or the other mtDNA morphs within one or very few generations. (Avise *et al.*, 1987).

1.3.4 EVOLUTION OF MTDNA EVOLUTION

Some regions in mtDNA are more evolutionarily labile than others probably due to different selective constraints. The mean rate of divergence averaged over the whole molecule is about 2% per million years in primates (Brown *et al.*, 1979), the genus *Ekes* (George & Ryder, 1986) and frogs (Carr *et al.*, 1987), as well as rhinoceroses, rodents, artiodactyls, gallinaceous birds, geese and salmonid fishes (Wilson *et al.*, 1985). This rate is 5 - 10 times faster than the single-copy fraction of the nuclear genome (figure 1.2).

However, this rate of mtDNA evolution is not necessarily true for other groups of animals. The substitution rate in *Drosophila* is twice as fast as its single copy nuclear DNA (Sharp & Li, 1989). There are also some suggestions that the

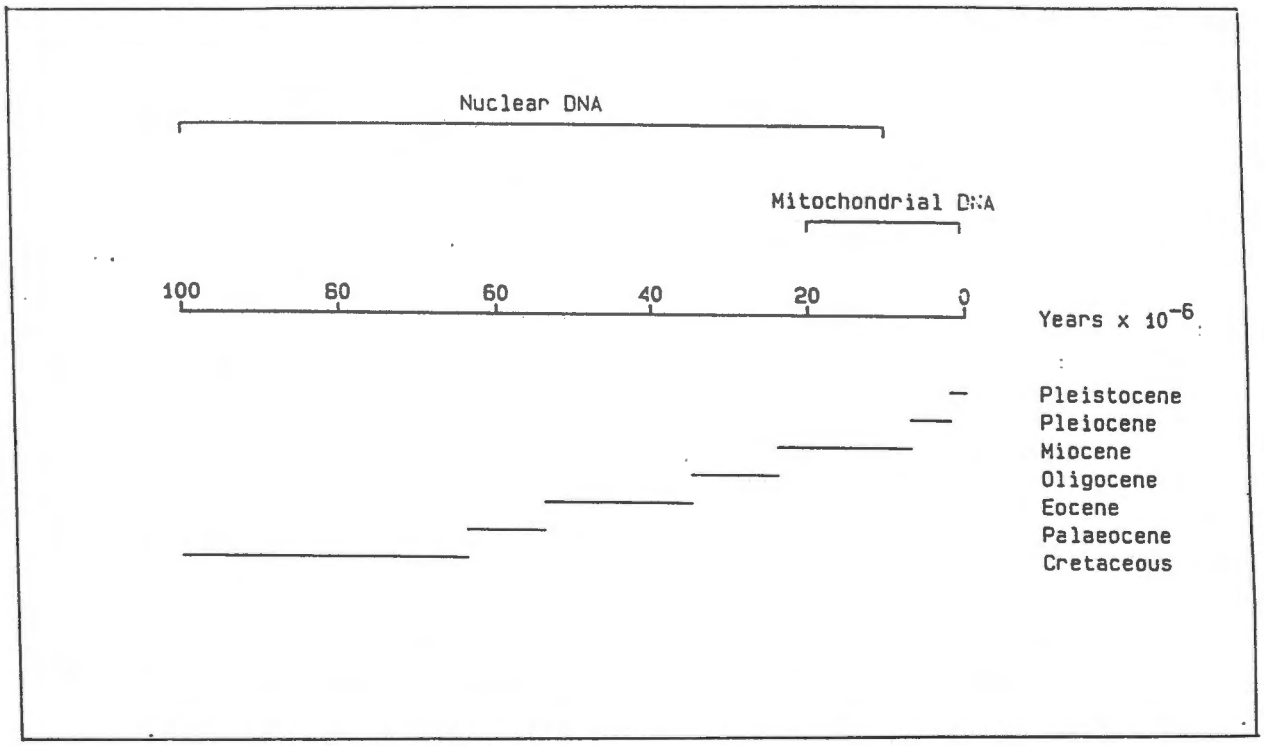

FIGURE 1.2

The time over which mitochondrial DNA and nuclear DNA are useful for comparisons between species or genera.

relative rates may vary in different mammals. Along the lineage leading to higher primates, mitochondrially encoded subunit II of cytochrome oxidase has been reported to have undergone at least a five-fold acceleration in the rate of change at the amino acid level (Avise *et al.*, 1987). However, the relevance of this to human mtDNA has however, been challenged (Easteal *et al., 1991*).

1.3.5 MOLECULAR CLOCK

The concept of the molecular clock was introduced by Zuckerkandl and Pauling (1962) when it was observed that the number of amino acid substitutions between a pair of species increases approximately linearly with time. However, there is no universal molecular clock and careful calibration of the clock for a specific lineage is essential (Cunningham *et al.*, 1992) before mtDNA can be used to estimate divergence for a particular group. Different rates of mtDNA evolution have been reported. Slower rates, for example, in sharks (Martin *et al.*, 1992) and faster rates, for example, in the false truffles (Bruns *et al.*, 1989). Therefore, macromolecules offer a quasi-clock being different for different groups of animals based on stochastic changes, and not on a perfect metronomic clock (Lewin, 1988).

The initial rate of mtDNA evolution (2% sequence divergence per million years in mammals) is a result of mutations primarily in the displacement loop region of the molecule, or in third codon positions ("wobble" position) which does not elicit an amino acid replacement.

Further mtDNA differences accumulate more slowly. For recently diverged taxa, base substitutions occur predominantly in the D-loop region or at sites within the coding region that do not result in a change in the amino acid encoded (Harrison, 1989). Transitions have been estimated to occur about ten times more frequently than transversions in mtDNA (Brown, *et al.*, 1979). As new mutations arise they are either lost or become fixed through random genetic drift (Fos *et al.*, 1990). The rate of nucleotide substitution appears to slow down when these sites are saturated, presumably because there are strong functional constraints on the tRNA and rRNA sequences and on the proteins encoded by the mtDNA.

The overall effect is that after about 8 - 10 million years, a plot of mtDNA sequence divergence against time (figure 1.3) becomes curvilinear eventually approaching a plateau (Brown *et al.*, 1979).

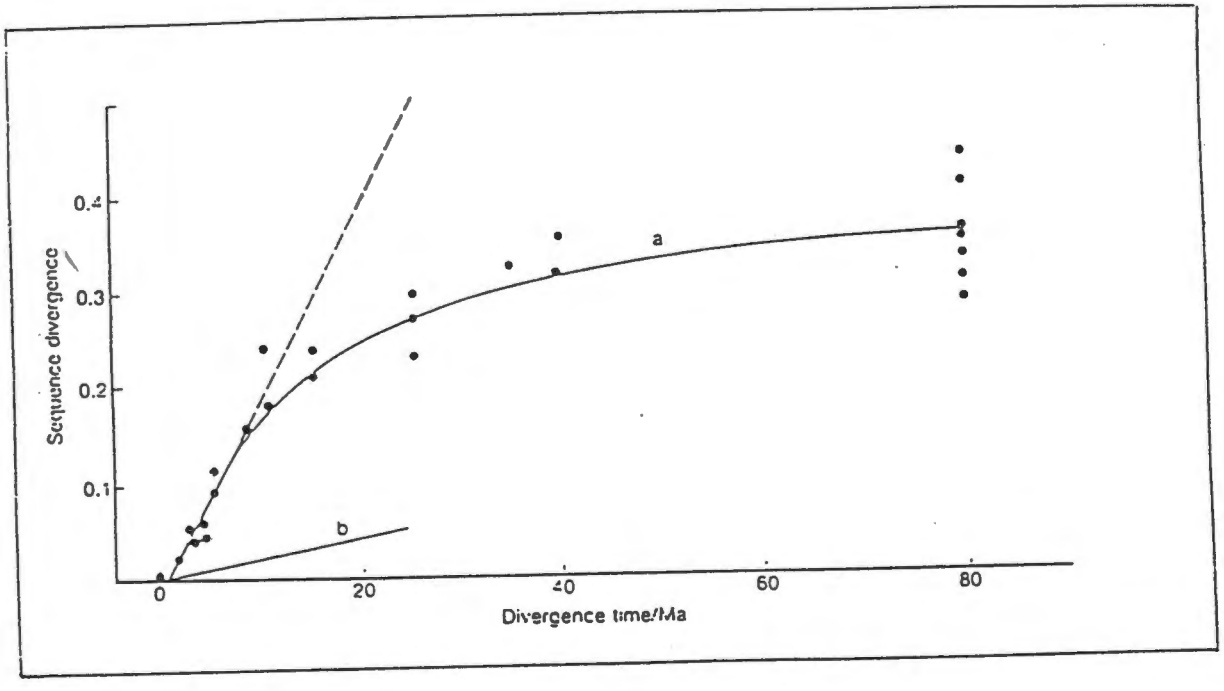

FIGURE 1.3

The rate of nucleotide substitution in (a) mtDNA and (b) single copy DNA. (Reproduced from Brown *et al.*, 1979)

1.4 MTDNA AS A GENETIC MARKER

Many of the criteria required by the ideal molecular system for phylogenetic analysis (Avise, *et al.*, 1987) are fulfilled by the molecular properties of mtDNA. These criteria of mt DNA are that it: (i) is distinctive, yet ubiquitously distributed so that homologous comparisons can be made among a wide variety of organisms; (ii) is easy to isolate; (iii) has a simple genetic structure; (iv) has a straightforward mode of genetic transmission; (v) provides qualitative character states so that phylogenetic relationships can be inferred by parsimony analysis; (vi) evolves at a rapid pace such that new character states arise within the lifespan of a species. However, it must be kept in mind that a mtDNA phylogeny represents a molecular genealogy and a maternally inherited and haploid one at that and not a species phylogeny, although it may frequently be a good estimate of the latter.

The methods for analyzing mtDNA include the use of restriction fragment length polymorphisms (RFLP), restriction enzyme site mapping and nucleotide sequencing (of cloned fragments or polymerase chain reaction amplified fragments, discussed in chapters 2, 4 and 5).

The use of RFLPs is the quickest and simplest way of analyzing mtDNA, but is suitable only for closely related

taxa. This approach is especially useful at the subspecies level (De Salle *et al.*, 1986, Honeycutt *et al.*, 1987, Cronin *et al.*, 1988 and Chang *et al.*, 1989) and for inter- and intra- population studies (Meyer *et al.*, 1990, Wayne *et al.*, 1990, de Villiers *et al.*, 1992 and Prinsloo & Robinson, 1992). Once the proportion of shared fragments becomes too low this method is inaccurate (Hillis and Moritz, 1990). In this technique mtDNA is cut into a number of fragment, the sizes of which are compared after gel electrophoresis. The proportion of shared fragments can be used to estimate sequence divergence between two groups of animals (Nei & Li, 1979). However, fragments are not suitable for use as character data in cladistic methods of phylogeny construction, since they violate the principle of independence of characters required by cladistic methods (Hillis & Moritz, 1990).

In restriction enzyme site mapping the specific cutting sites of different restriction enzymes are mapped relative to one another. The methodology is discussed in chapter 2. Mapping is more time-consuming than a RFLP study, but gives good resolution at the species and genus level (George & Ryder, 1986, Solignac *et al.*, 1986, Carr *et al.*, 1987, Cronin, 199 and Hillis *et al.*, 1992). Restriction sites can be used as valid characters (unlike restriction fragments) in cladistically based phylogeny construction methods.

Sequence information from mtDNA was first obtained after the cloning of a fragment of the mitochondrial DNA molecule (Higuchi *et al.*, 1984). The development of the polymerase chain reaction or PCR (Mullis & Faloona, 1987) has made the comparison of mtDNA sequences much easier (Higuchi *et al.*, 1987, Bruns & Palmer, 1989, Kocher *et al.*, 1989, and Miyamoto *et al.*, 1990). Phylogeneticists now have an excellent means of studying extinct animals whose remains are represented in museum collections. PCR facilitates the amplification and analysis of trace quantities of badly damaged DNA (Thomas *et al.*, 1990). Novel phylogenetic placements of 30 million year old DNA from an extinct taxon of termite, fossilized in amber (De Salle *et al.*, 1992) and 7 million year old DNA from magnolia (Golenberg *et al.*, 1990) have been made. The first case of a new avian species of shrike was identified solely on the basis of its mtDNA sequence (after amplification) was recently documented, without the standard type specimen being deposited in a museum (Hughes, 1992).

1.5 POPULATION GENETICS

1.5.1 PHYLOGEOGRAPHY

MtDNA exhibits considerable variation among individuals, both within and between populations. This has led to the use of the term "intraspecific phylogeography" to describe the use of mtDNA to bridge the gap between population genetics and systematics (Avise, 1986 and Avise *et al.*, 1987). The distribution of mtDNA genotypes has provided data useful for defining geographical structuring of populations.

There often appears to be major discontinuity in the distribution of mtDNA morphs of some groups of animals which presumably, reflect current environmental and/or historical influences (Avise *et al.*, 1979, Nicols, 1989, and Baker *et al.*, 1990).

Illustrated are examples of the horseshoe crab (figure 1.4) that showed extensive geographic mtDNA variation along the eastern coast of the United States (Saunders *et al.*, 1986) and the rock hyrax in southern Africa (figure 1.5) who also showed strong geographical partitioning (Prinsloo & Robinson, 1992). In contrast, the coyotes *Canis latrans* (Lehman & Wayne, 1991) and the American eel (figure 1.6)

Anguilla (Avise *et al.*, 1986) both showed no geographical differentiation across their respective ranges.

The American eel was sampled from seven different localities along the coast of North America as well as from two localities from the British Isles (figure 1.5). The European mtDNA genotype was clearly distinguishable from the American mtDNA genotype, but no variation was found along the 4 000km stretch of American coast. It is thought that there is a single spawning in the western mid-Atlantic Ocean and that there the subsequent widespread dispersal of larvae is by ocean currents. This study emphasizes the importance of life history in shaping genetic structure.

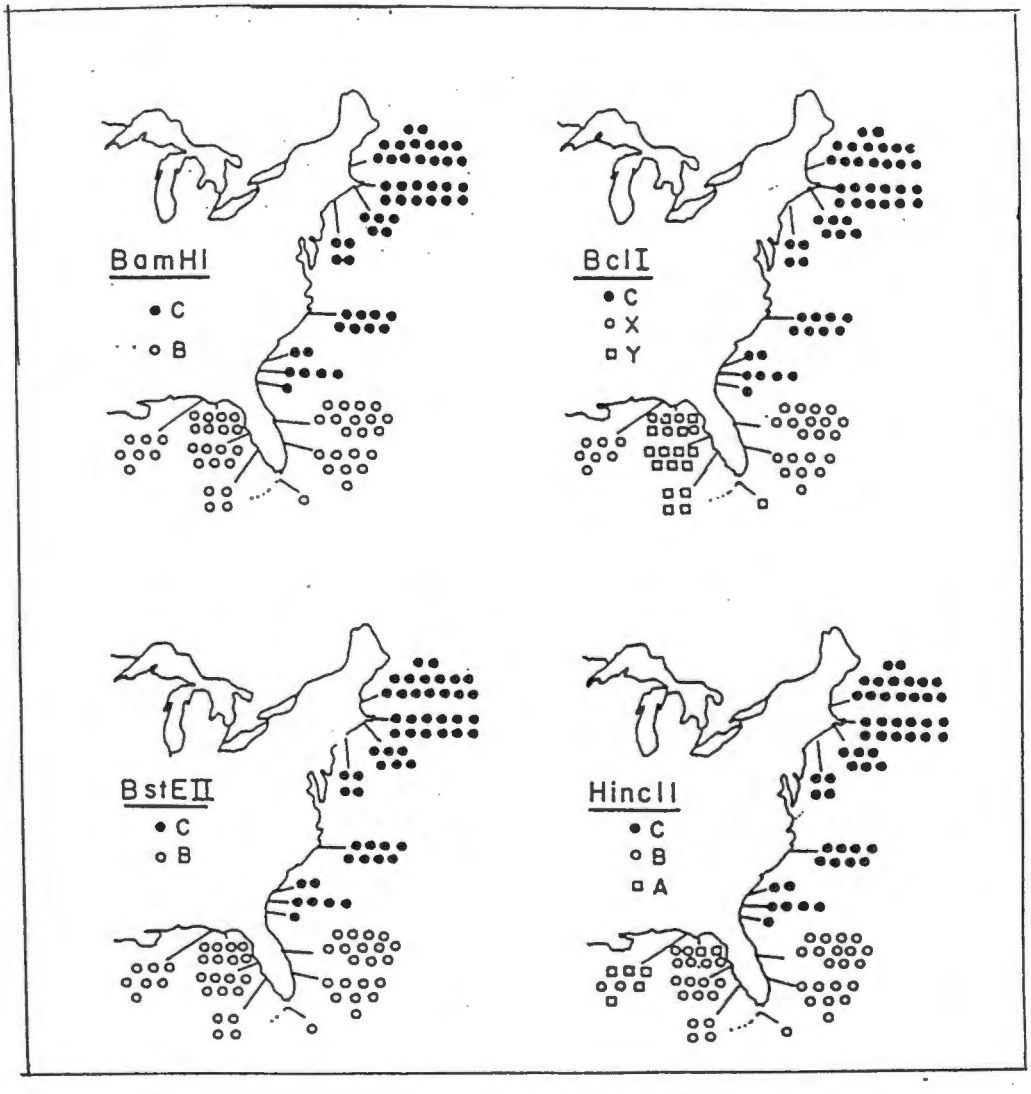

FIGURE 1.4

The geographical separation of the horseshoe crab, *Limulus polyphemus*, mtDNA morphs along the eastern coast of the United States. The filled circles, open circles and the open squares represent the different mtDNA morphs for four restriction enzymes. (Reproduced from Saunders *et al.*, 1986)

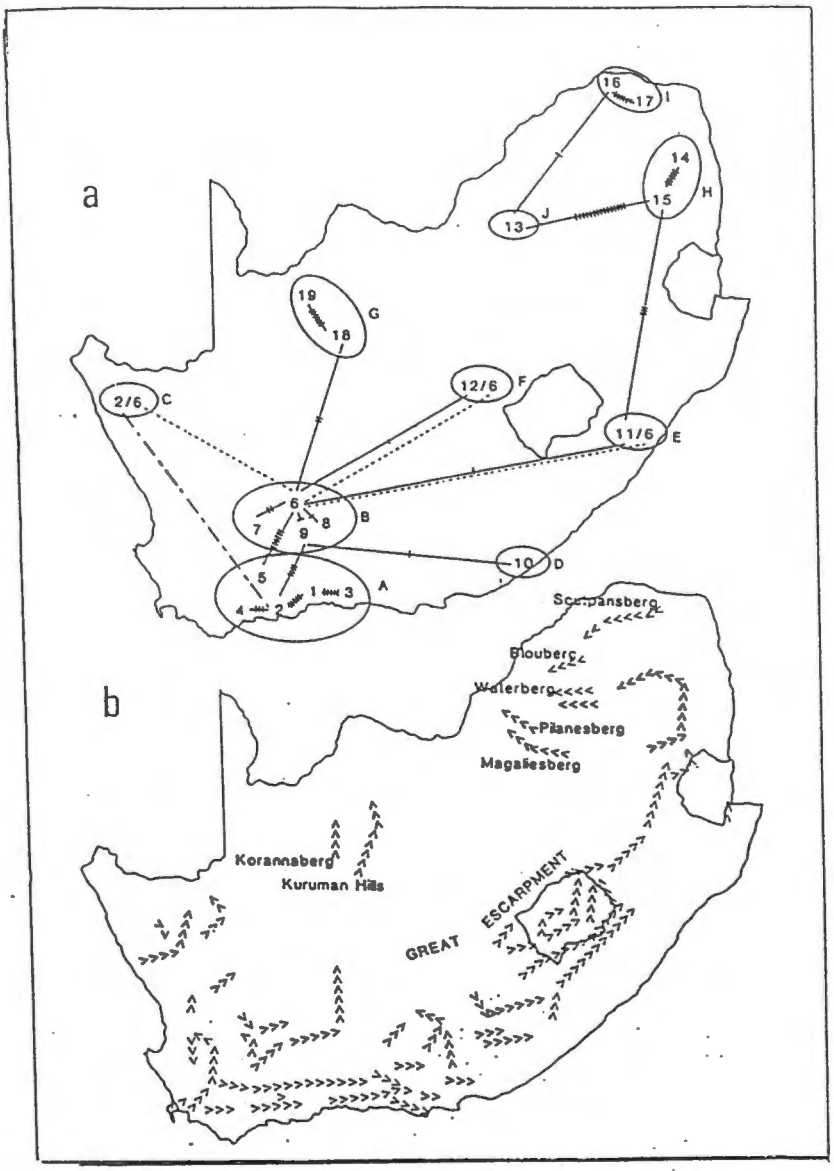

FIGURE 1.5

An example of the geographical mtDNA distribution in the rock hyrax, *Procavia capensis*. (a) The geographic overlay of minimum mutational steps between the rock hyrax clones. Crossbars represent single mutational steps, large circles encompass sample localities and their respective clones and broken lines designate identical clones in different populations. (b) The mountain ranges of South Africa. (Reproduced from Prinsloo & Robinson, 1992)

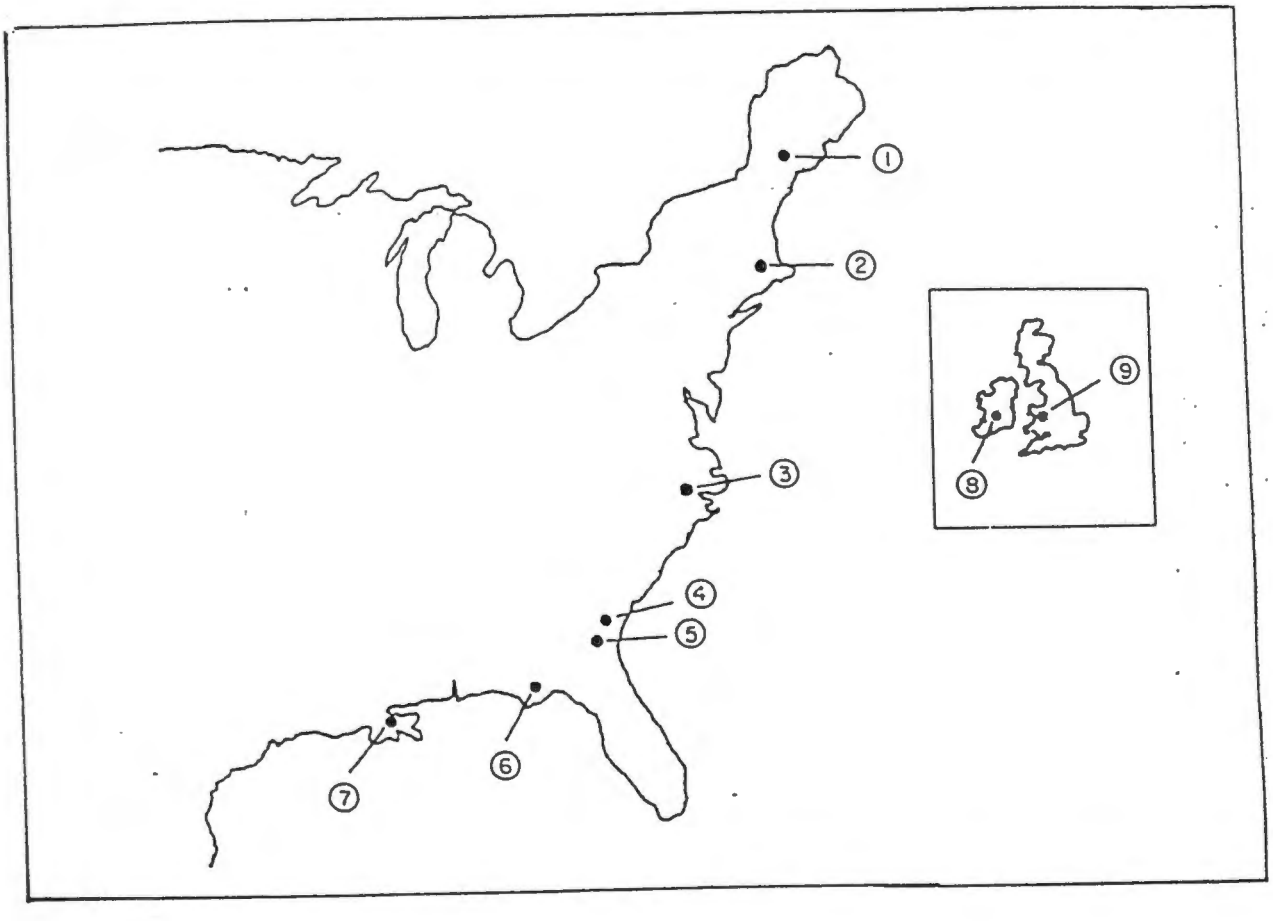

FIGURE 1.6

The mtDNA distribution of North Atlantic eels. The eels showed no genetic variation in its mtDNA among samples from a 4000 km stretch of North Atlantic coastline. (Reproduced from Avise *et al.*, 1986)

1.5.2 TIME OF DIVERGENCE AND POPULATION SIZE

Our own species shows few signs of geographical structuring. This suggests that existing human populations migrated from a common centre in the relatively recent past (Cann *et al.*, 1987 and Vigilant *et al.*, 1991). If we compare the mtDNA from existing humans and accept a divergence rate of 20 x 10^{-9} per site per year, we can conclude that the single common ancestor in the female lineage existed some 200 000-400 000 years or 10 000- 20 000 generations ago (Cann *et al.*, 1987). However, this an area engaging a great deal of controversy and criticism as how to best interpret mtDNA restriction maps or sequence data and consensus has yet to be reached (Vigilant *et al.*, 1989 and Gee, 1992).

MtDNA may also help estimate population size of past populations. From the mean number of years per generation and the percentage sequence divergence observed in mtDNA, the effective population size can be calculated (Wilson *et al.*, 1985).

1.5.3 FOUNDER EVENTS AND BOTTLENECKS

The maternal inheritance of mtDNA and the absence of recombination makes it useful for analysis of colonization events and population bottlenecks. The monophyletic origin of the Lake Victoria cichlid fishes (Meyer *et al.*, 1990) were shown by mtDNA as well as the African origin of honey bees in the Americas (Smith *et al.*, 1989 and Hall *et al.*, 1989).

Bottlenecks have been studied using mtDNA. For example, it has been suggested that *Homo sapiens* could have passed through a severe bottleneck 180 000 year ago (Brown, 1980). The capture of female *Xenopus laevis* from the vicinity of Cape Town in the 1930s and 1940s for pregnancy testing resulted in the suggestion that there had been a severe bottleneck in the population size of *Xenopus laevis* (Carr *et al.*, 1983).

1.5.4 HYBRIDIZATION AND INTROGRESSION

Different species are normally reproductively isolated and there is no gene flow between them. However, it has been reported that mtDNA can cross species boundaries (Wilson *et al.*, 1985, Cronin, 1991 and Wayne & Jenks, 1991). Often, in

these areas of hybridization, differential introgression occurs. This is when there are individuals or populations bearing the nuclear genes of one species and the mtDNA of a different species.

The first reports of introgression came from a study done on mice (Ferris *et al.*, 1982). A recent study on North American grey wolf populations has also demonstrated the introgression of coyote mtDNA into the grey wolf (Lehman *et al.*, 1991). These results suggest that in disturbed areas, previously distinct species may interbreed if one is rare and the other abundant. In large, highly mobile carnivores such as coyotes and grey wolves introgression can be rapid and occur over broad areas.

It is not clear how often introgression of mtDNA occurs, but closely related species of mobile terrestrial vertebrates have the potential for extensive genetic exchange when ecological conditions change suddenly.

1.6 CONSERVATION IMPLICATIONS

In conservation biology, most discussions have focussed on how best to preserve biotic diversity and variability within threatened species (Avise, 1989). Molecular genetic information is of considerable use to both define evolutionary distinct units (Bowen, *et al.*, 1991) and to

measure the existing variability in threatened species and to monitor this over time. These parameters were examined in the endangered Isle Royale grey wolves (Wayne *et al.*, 1991) so that better informed decisions can be made about their management.

Even complex behavioural characteristics like the nesting patterns of the green sea turtle, which has an effect on the survival of the animals and has conservation implications as their nesting habitats are becoming increasingly threatened, can be studied using molecular methods (Bowen *et al.*, 1989 and Meylan *et at*, 1990).

1.7 THE USE OF MTDNA IN THIS STUDY

1. Restriction enzyme sites maps of *Diceros bicornis* (black rhinoceros) and *Ceratotherium simum* (white rhinoceros) were constructed. The overall sequence divergence in mtDNA could be measured, thus given certain assumptions, enabling the time of divergence from their common ancestor to be estimated.

2. Part of the cytochrome-b region of mtDNA from *D. bicornis* and *C. simum* was sequenced, after amplification by PCR. This constitutes a data set of rhinoceros sequence information which can, when other members of the family are included, be

used for systematic studies. It can also be used as a separate estimate of the time of divergence between the black rhinoceros and white rhinoceros and provides reference data for subsequent applications such as the studies described below.

3. Given the vulnerable situation of black rhinoceros populations, mtDNA variation among four subspecies of black rhinoceroses were examined in order to clarify their subspecific status.

4. In a practical application of the sequence information, the species of origin of trace amounts of tissue on an axe head was determined, after amplification by the polymerase chain reaction of a fragment of mtDNA and sequencing of the cloned product.

CHAPTER 2

THE ESTIMATION OF THE TIME OF DIVERGENCE BETWEEN BLACK RHINOCEROS AND WHITE RHINOCEROS USING RESTRICTION ENZYMES MAPS

2.1 PURIFICATION OF MTDNA

Purified mtDNA was isolated from frozen tissue by differential centrifugation (Brown *et al.*, 1980, Lansman *et al.*, 1981, and Cummings *et al.*, 1987).

2.1.1 ISOLATION OF CRUDE MITOCHONDRIA

Rhinoceros bulk tissue samples were obtained from opportunistic deaths. The *C. simum* and *D. b. minor* hearts were obtained from Natal, RSA and the *D. b. michaeli* heart was from the Eastern Cape, RSA. Heart tissue was stored at -20^0C. Approximately 30g of frozen tissue was thawed on ice

and minced in a meat-grinder at 4^0C. All subsequent steps were carried out at 0-4^0C in order to minimize hydrolytic enzyme activity. The minced tissue was suspended in 4.5ml per gram of tissue of cold extraction buffer (100mM Tris-HCl pH 7.4, 150mM NaCl, 20mM EDTA, 10% (w/v) sucrose) and then homogenized in a blender for 30 seconds.

The homogenate was centrifuged in a swing-out bucket rotor at 1 000 g for 10 minutes at 4^0C and the supernatant was carefully decanted. This step was usually repeated. This was done to remove the nuclei and the cellular debris. If the sample contained significant quantities of lipid, the supernatant was filtered through a layer of cheese cloth. The supernatant was centrifuged at 10 000 g for 30 minutes at 4^0C to pellet the mitochondria. The pellet was then washed, resuspended in 30ml extraction buffer and recentrifuged at 20 000 g for 15 minutes at 4^0C to produce the crude mtDNA pellet.

2.1.2 MtDNA PURIFICATION

The crude mtDNA pellet was suspended in 3.5ml STE (100mM NaCl, 10mM Tris-HCl pH 8.0, 1mM EDTA) buffer. Subsequent steps were done at room temperature. SDS was added to the mtDNA pellet to a final concentration of 1% in order to lyse the mitochondrial membranes. After 15 minutes at room

temperature, CsCl was added to give a final concentration of 1M, and allowed to stand for a further 15 minutes at room temperature to facilitate the precipitation of the contaminating proteins. The proteins were pelleted by centrifugation at 10 000 g for 15 minutes at 20^0C in a Sigma 2MK bench-top centrifuge. CsCl was added to the supernatant to a final concentration of 1g per ml (less the amount of CsCl added earlier to precipitate the proteins) and ethidium bromide (EtBr) to a concentration of 80 ul/ml (stock solution 10mg/ml). The density of the samples were adjusted to 1.56 g/ml and loaded into Beckman quick-seal tubes with a syringe and centrifuged for 16 - 18 hours in a VTi65 vertical rotor at 50 000 rpm at 20^0C.

Covalently closed, circular mtDNA has a greater buoyant density in the presence of EtBr than linear nuclear DNA, and thus, can be separated into two distinct bands by CsCl density gradient centrifugation. The lower mtDNA band was visualized by a long wavelength UV lamp and collected. Salt-saturated isoamyl alcohol was used to remove the EtBr from the samples. An equal volume of salt-saturated isoamyl alcohol was added to the sample, mixed and then the two phases were allowed to separate. The upper organic phase (the EtBr-containing phase) was discarded. This extraction process was repeated until all visible traces of EtBr was removed. The mtDNA was ethanol precipitated at -20^0C for at least 60 minutes and was pelleted by centrifugation at 10

000 g for 15 minutes at 4^0C in a Sigma 2MK bench-top centrifuge. The mtDNA pellet was then resuspended in 200 - 400 ul TE buffer (10mM Tris-HCl pH 8.0, 1m EDTA), aliquoted and stored at -20^0C.

2.2 RESTRICTION ENZYME DIGESTS

Restriction enzyme digests were performed as recommended by the suppliers. It was important that the correct salt concentration was used to ensure complete digestion patterns for each restriction enzyme. A "universal" RE buffer, called KGB (in 2x KGB: 200mM potassium glutamate, 25mM Tris-acetate pH 7.6, 10mM magnesium acetate, 50ug/ml bovine serum albumin, 0.5M 2-mercaptoethanol) was used (McClleland, *et al.*, 1988). This universal buffer allowed one to perform double digestions in one microfuge tube, without DNA precipitation or dialysis steps between the two RE additions, in order to alter salt concentrations.

Enzyme dilutions were performed to ensure good restriction digestion patterns, using the minimum quantity of restriction enzyme. A restriction enzyme digest contained the following components, added sequentially, while working on ice: (i) approximately 5 ng of mtDNA (in a volume of 1-4 ul of TE buffer) (ii) 2 units of restriction enzyme diluted

with KGB and (iii) KGB buffer to a final volume of 15 ul. The reaction was incubated at 37^0C for 2 hours.

2.3 ENDLABELING MT DNA FRAGMENTS

The endlabeling technique (Brown 1980) was performed to visualize the mtDNA digestion fragments. After restriction enzyme digestion, 0.5-1 units of the Klenow fragment was added to each digest and incubated at room temperature for 15 minutes.

This incubation, without the deoxynucleotide substrates allows the labeling of restriction enzyme fragments whose recognition sequences do not contain dCTP or which have blunt ends or 3'-overhangs. This allowed the use of a single radioactively labeled dNTP (in this case dCTP) for all end-labeling reactions, irrespective of the recognition sequence of the restriction enzyme, because of the 3'-5'exonuclease activity of Klenow (large fragment of DNA polymerase).

The preincubation was followed by the addition of dATP, dGTP and dTTP to concentration of 2mM and the labeled nucleotide, [alpha-^{32}P]dCTP to a concentration of 1 uCi (specific activity of 370 MB/ml) to each digest. The end-labeled reaction was stopped by adding 4ul of EDTA-containing loading buffer (0.25% bromophenol blue, 40% (w/v) sucrose).

2.3.1 ENDLABELING LAMBDA HIND III MARKER DNA

To calculate the molecular sizes of the mtDNA fragments, the relative mobility of the fragments in an agarose gel were compared to the mobility of the known fragments of a marker. Lambda phage DNA digested with Hind III was used (molecular sizes in Appendix).

To endlabel lambda, 10 ug of Hind III-digested DNA was diluted in 40 ul KGB to which 6 units of Klenow fragment was added. This was incubated at room temperature for 15 minutes followed by the addition of the unlabeled dATP, dGTP and dTTP to 2 mM and 10 uCi of [alpha- ^{32}P]dCTP. The reaction was incubated for an additional 15 minutes at room temperature. This reaction served as the labeled marker stock solution, of which 1 ul (20 ng) was diluted with sterile distilled water to 15 ul; loading buffer was added and the mixture loaded into the slots of the agarose gel.

2.4 GEL ELECTROPHORESIS AND AUTORADIOGRAPHY

Agarose gels (1% - 1.7%) were used to separate DNA digestion fragments. The correct quantity of agarose was calculated for the specific percentage gel and dissolved in 150 ml 1x TAE buffer (50x buffer: 40 mM Tris-HCL pH 7.7, 40mM acetic

acid, 1mM EDTA). The agarose was dissolved by heating to above 65^0C and poured into 15 x 20 cm casting trays. The labeled DNA fragments were electrophoresed at 35V (2 Vcm^{-1}) for 16 to 20 hours using 1x TAE buffer as the running buffer.

After electrophoresis, the gels were dried and autoradiographed. The gels were vacuum dried at 55^0C on a slab gel drier for 90 minutes. The dried gel was loaded into a light-proof X-ray cassette with intensifying screens (Dupont), exposed to X-ray film (Hyperfilm-MP from Amersham) and developed after 6 - 24 hours exposure at -70^0C.

2.5 RESTRICTION SITE MAPPING

2.5.1 INTRODUCTION

To map a region of DNA, the DNA is digested with a number of restriction enzymes. Each restriction enzyme recognizes a specific short sequence (6 base pair recognition enzymes were used throughout this study) which cleaves the DNA at this specific recognition site. The resultant fragments are separated on agarose gels, visualized using radioactive dCTP, and the sizes calculated by reference to the molecular weight marker.

The mammalian mtDNA is a circular genome of approximately 16 400 base pairs, so when it is cleaved with a restriction enzyme that recognizes two sites, then two fragments result from this restriction enzyme digestion.

When mapping mtDNA, the sample is digested with a restriction enzyme that recognizes, for example, two restriction enzyme sites. Let us assume that both the resultant two fragments constitute a pattern "A". Then a second restriction enzyme that recognizes only a single site was chosen, and this single fragment pattern constitutes pattern "B". A third digestion is done with both restriction enzymes and the resultant pattern, A + B, is used to help map the one restriction enzyme site relative to the other. This is done for a whole array of restriction enzymes, all of which are mapped relative to one another. In this way one is able to map a range of enzymes on a stretch of DNA.

The mapping of the black rhinoceros and the white rhinoceros was facilitated by a computer program called RESOLVE (Harley, 1992a) which stores temporary solutions to mapped sites and computes the many different solutions found.

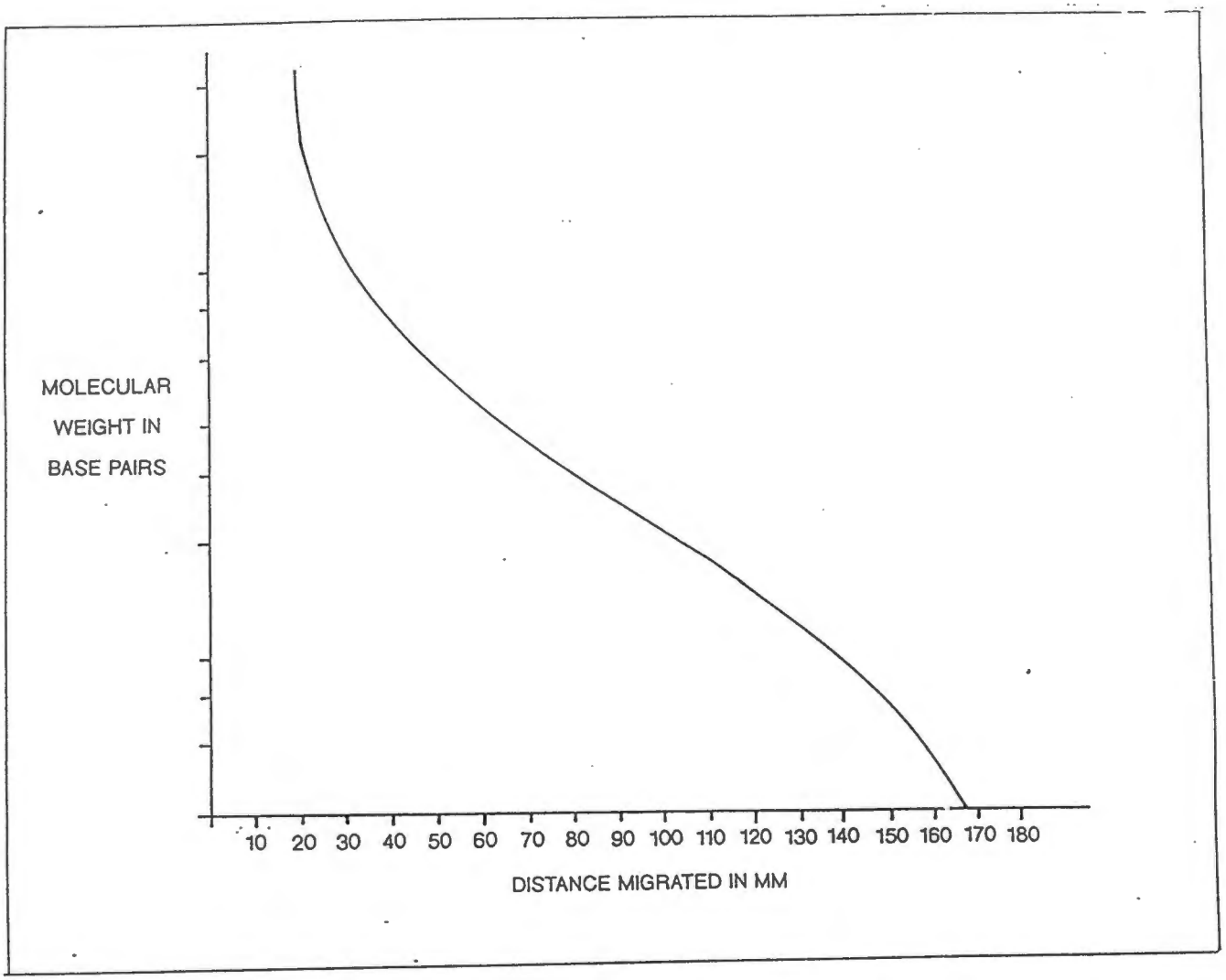

FIGURE 2.1

The standard curve of lambda Hind III marker DNA plotted on semi-logarithmic graph paper. The molecular sizes of the DNA are plotted on the y-axis(logarithmic scale) and the distance migrated from the origin is plotted on the x-axis.

2.5.2 DETERMINATION OF MOLECULAR SIZE

The relative mobility of the Hind III fragments of lambda DNA are plotted against the molecular weight of each fragment on semi-logarithmic paper. A standard curve is drawn (figure 2.1) and using the relative mobility of each mtDNA fragment, the molecular weight is read from the standard curve. The relationship of molecular weight to mobility is not linear, especially in the high molecular weight range. This non-linearity of the standard curves make the accurate determination of the size of large fragments (larger than 7000 bp) very difficult.

2.5.3 RESTRICTION ENZYME DATA MANAGEMENT

Restriction fragment sizes determined from single- (using one restriction enzyme), and double-digests (using two restriction enzymes) were entered into the computer program, RESOLVE (Harley, 1992a). This program allows for the storage and manipulation of the data. When the fragment sizes from a double digest were entered, the program searches for all possible combinations to fit (map) the two restriction enzyme fragment patterns of the double digest relative to each other. When the search for the solution (or solutions, if there is more than one that is found within the set error limits) to this double digest is complete, it can be stored in a temporary file (figure 2.2 to 2.6). Unfixed sites are

those where two or more "A" fragments lie totally within a "B" fragment, such that there are a number of different ways to position these sites within the "B" fragment (figure 2.7).

2.5.4 THREE-WAY CONSENSUS ANALYSIS

The three-way analysis was used to resolve unfixed sites, partial solutions to double digest data and to fit the solutions of more than one double-digest together. To perform a three-way analysis, the appropriate combination of double-digests are required. This consists of all three sets of double digest fragment patterns and the single digest fragment patterns for three restriction enzymes. For example, DNA was digested with Bam HI, Bgl II and Eco RV singly, and in the different double-digest combinations of Bam HI + Bgl II, Bam HI + Eco RV and Bgl II + Eco RV. The three-way analysis performs a compatibility analysis to find a map for the three restriction enzymes consistent with the single fragment patterns and each one of the three double digest patterns. This final result is stored in the final map files. RESOLVE has simplified the data management of restriction enzyme mapping as well as producing unbiased, rigorous restriction enzyme maps.

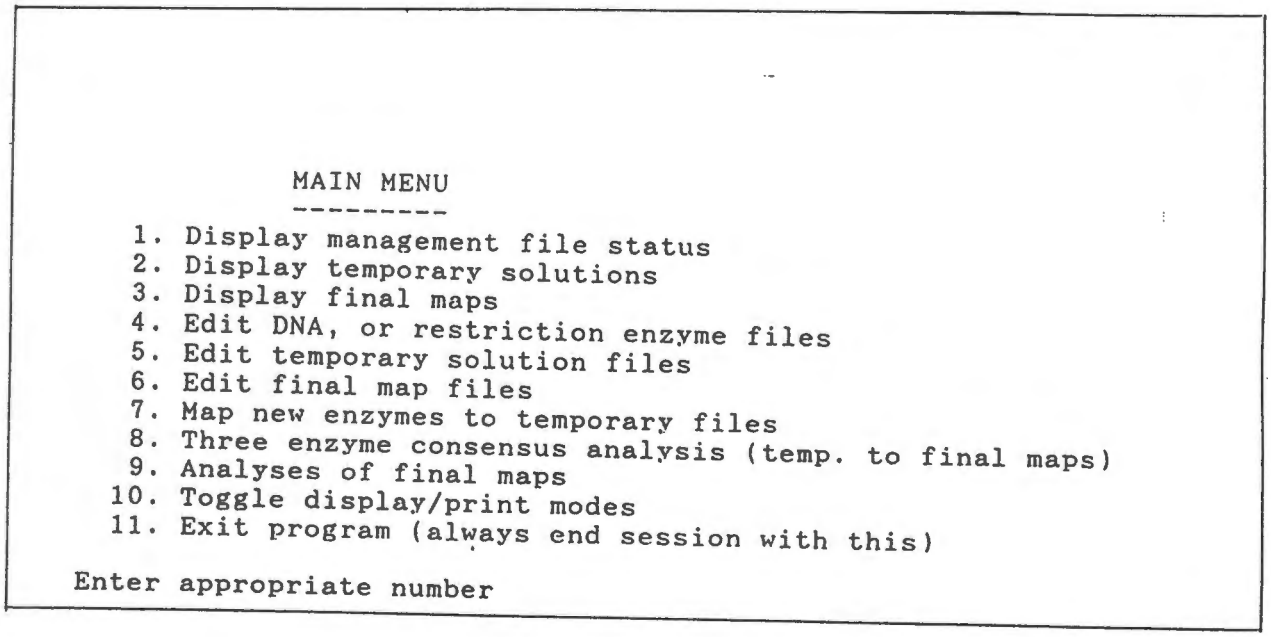

FIGURE 2.2

The main menu screen of the RESOLVE program.

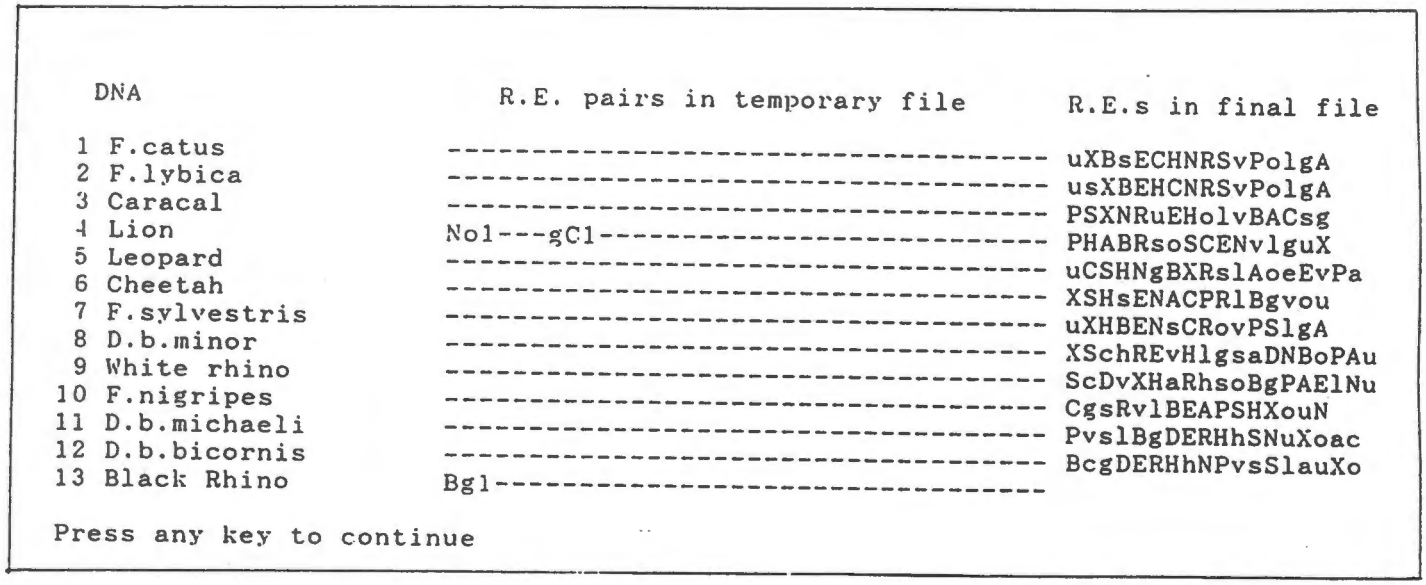

FIGURE 2.3

A typical display of the RESOLVE program's file status. The names of the restriction enzymes are abbreviated.

```
**************  Black Rhino        BamHI      & Bgl II    **************
% error =  2

Now enter the fragment sizes for BamHI      (in any order)
? 8400
? 8000

Now enter the fragment sizes for Bgl II
? 16400
```

FIGURE 2.4

A display of restriction enzyme (Bam HI and Bgl II) data being entered into the RESOLVE program.

```
Any corrections ? (Y/N)
? n
Normalise the total (and hence all fragment sizes) ? (Y/N) n
Show details of working (informative, but takes longer) ? (Y/N) y
A no  1 = 8400  pairs with fragments          2  &  3          =  8413
A no  2 = 8000  pairs with fragment           1                =  7997
    Total fits =  2

Partial Solution  1  :
A fragment  1 consists of AB fragments  2 & 3
A fragment  2 consists of AB fragments  1
One partial solution found for A

B no  1 = 16400                pairs with fragments          1 & 2 & 3
                                                                =  16410
    Total fits =  1

Partial Solution  1  :
B fragment  1 consists of AB fragments  1 & 2 & 3
One partial solution found for B
```

FIGURE 2.5

A display of the fragment analysis and comparison by the RESOLVE program.

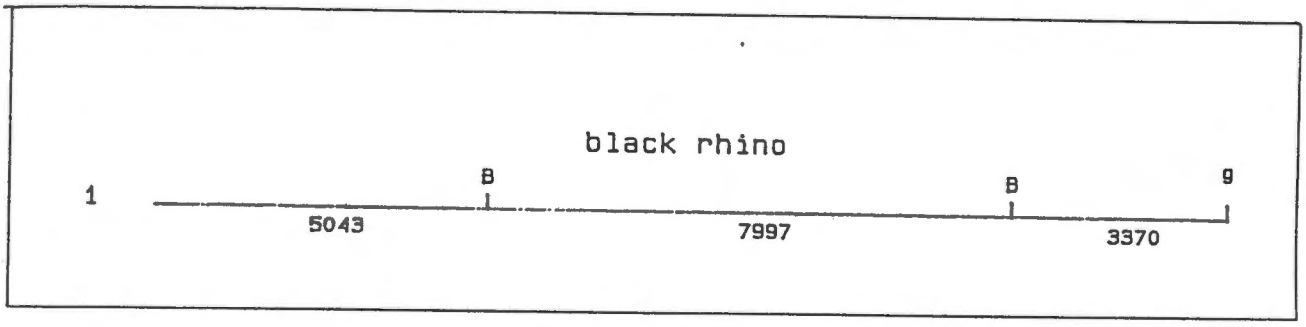

FIGURE 2.6

The solution for the Bam HI and Bgl II digestion.

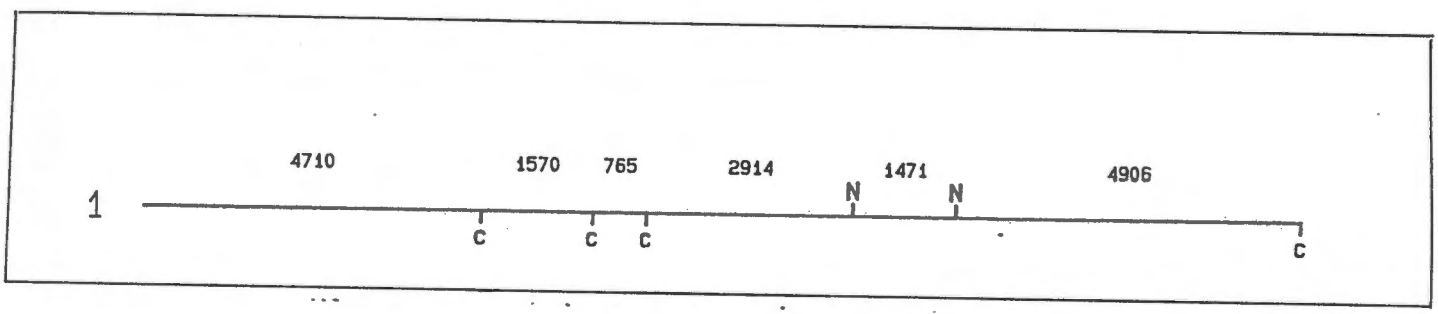

FIGURE 2.7

An example of unfixed sites within a fragment of DNA.

2.6 RESULTS AND DISCUSSION

Both the black rhinoceros (*Diceros bicornis minor*) and the white rhinoceros (*Cerathotherium simum simum*) mtDNAs were digested with 18 restriction enzymes (listed in Table 2.1). Both single and double digestions were performed. Figures 2.8 and 2.9 illustrate autoradiographs of single- and double-digestion patterns respectively. The sizes of each fragment were calculated from the molecular standard, lambda Hind III.

The restriction enzyme maps (figure 2.10) that were constructed by the double-digestion technique were aligned and oriented using the two Sac II sites and the one Hpa I site at positions 676, 2364 and 5480 respectively in the published bovine mtDNA sequence (Anderson, 1982). These sites are invariant throughout most of the vertebrates (Carr, 1987). When restriction enzyme sites from the two species were aligned to within 1% of the total map length, they were interpreted as being shared sites. The total map length was estimated to be 16417 (+/- 298) and 16411 (+/- 225) for black rhinoceros and white rhinoceros respectively (figure 2.11).

TABLE 2.1

A list of the restriction enzymes mapped and the code used.

Enzyme	Code used
Bam HI	B
Bcl I	c
Bgl II	g
Hind III	H
Hpa I	h
Eco RI	E
Eco RV	R
Dra I	D
Nco I	N
Pst I	P
Pvu II	v
Sac I	s
Sac II	S
Sal I	l
Sca I	c
Stu I	u
Xba I	X
Xho I	o

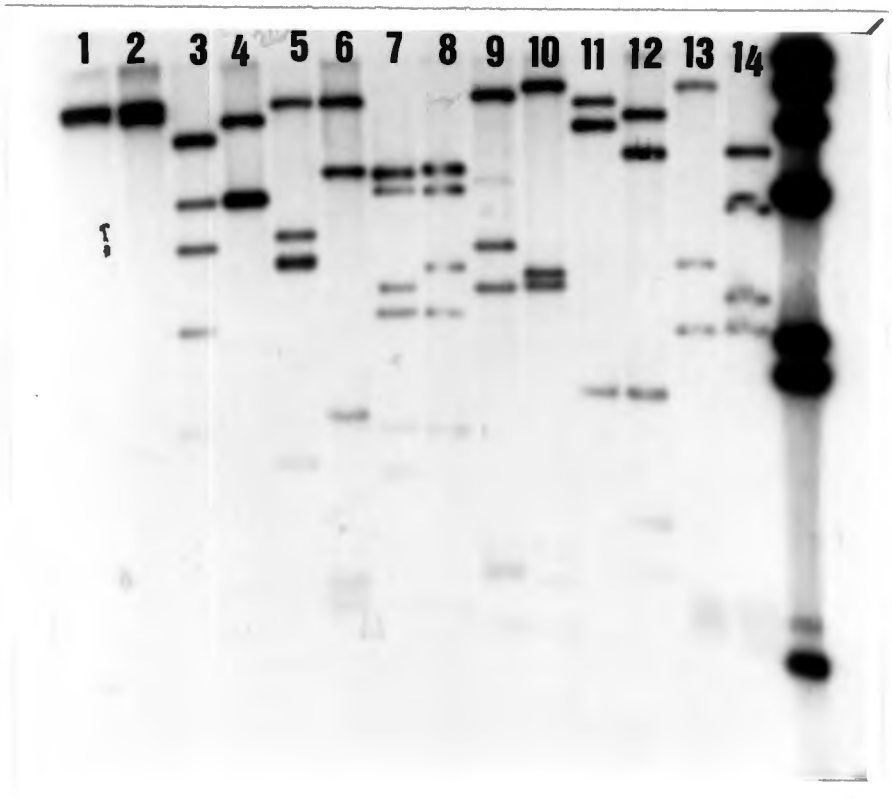

FIGURE 2.8

Agarose gel electrophoresis showing single digest patterns of black rhinoceros (odd numbers) and white rhinoceros(even numbers). Lanes 1 & 2: Bam HI, 3 & 4: Hind III, 5 & 6: Xha I, 7 & 8: Stu I, 9 & 10: Pvu II, 11 & 12: Dra I and 13 & 14: Bcl I.

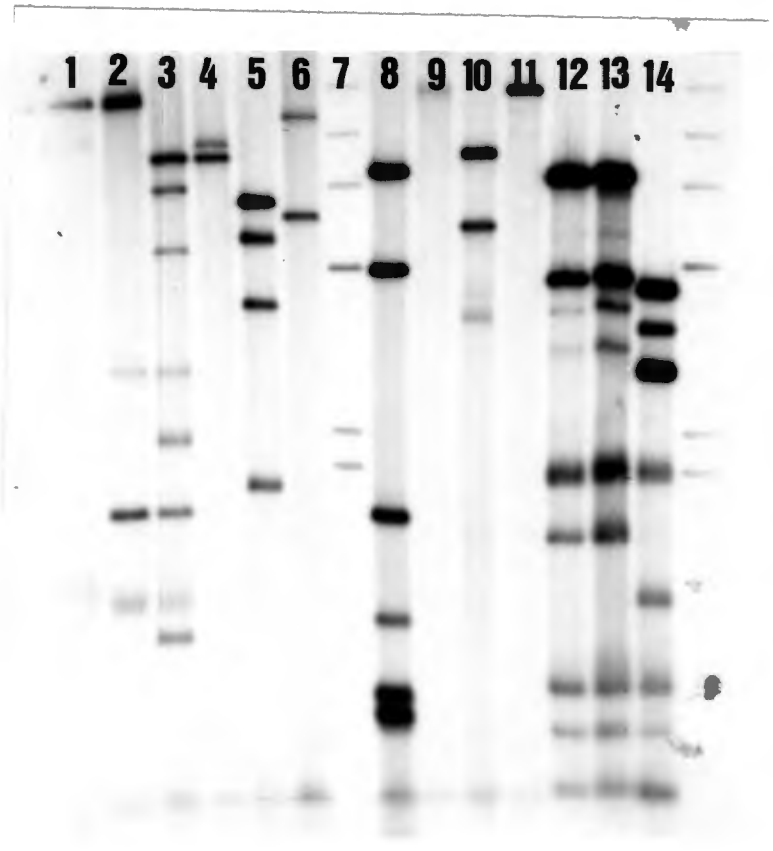

FIGURE 2.9

Some representative double digestion patterns for black rhinoceros. The DNA has been cut with: Lane 1. Sac II + BgI II, 2. Sac II. 3. Sac II + Bam HI, 4. Bam HI, 5. Bam HI + Eco RI, 6. Eco RI, 8. Bam HI + Xho I, 9. Xho I, 10. Bam HI + BgI II, 11. Sca I + BgI II, 12. Sca I and 13. Sca I + Bam HI. Lane 7 and 14. Hind III lambda DNA.

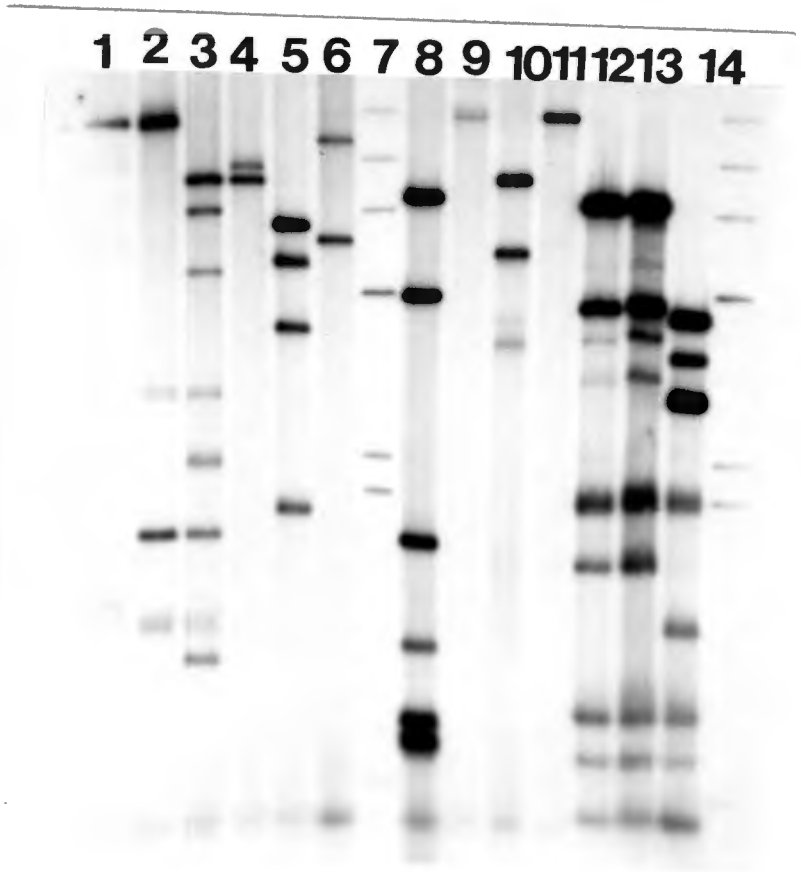

FIGURE 2.9

Some representative double digestion patterns for black rhinoceros. The DNA has been cut with: Lane 1. Sac II + BgI II, 2. Sac II. 3. Sac II + Bam HI, 4. Bam HI, 5. Bam HI + Eco RI, 6. Eco RI, 8. Bam HI + Xho I, 9. Xho I, 10. Bam HI + BgI II, 11. Sca I + BgI II, 12. Sca I and 13. Sca I + Bam HI. Lane 7 and 14. Hind III lambda DNA.

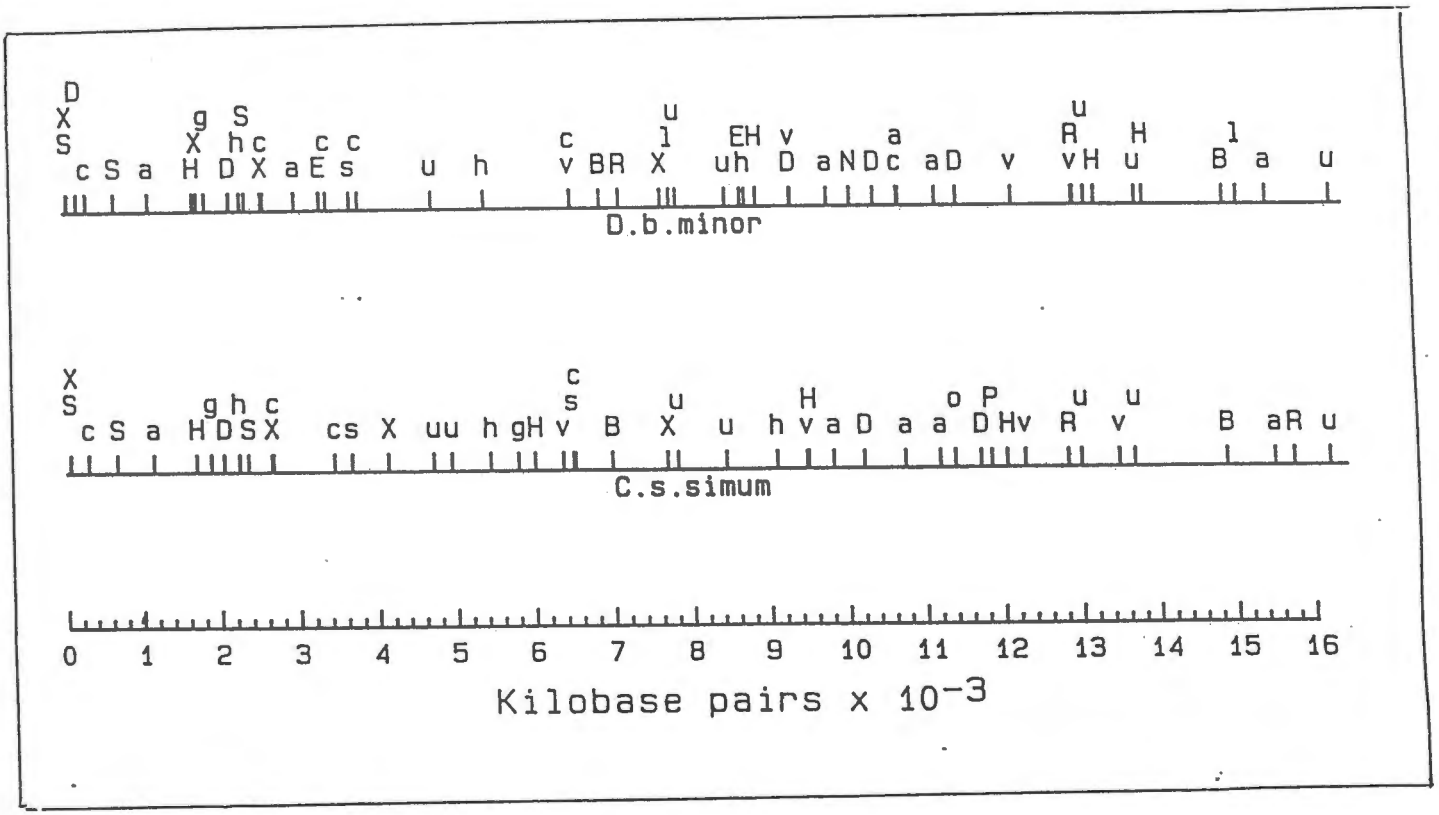

FIGURE 2.10

The restriction enzyme maps of black rhinoceros and white rhinoceros aligned on the invariant Sac II site at position 676 in the *Bos taurus* sequence.

FIGURE 2.10

The restriction enzyme maps of black rhinoceros and white rhinoceros aligned on the invariant Sac II site at position 676 in the *Bos taurus* sequence.

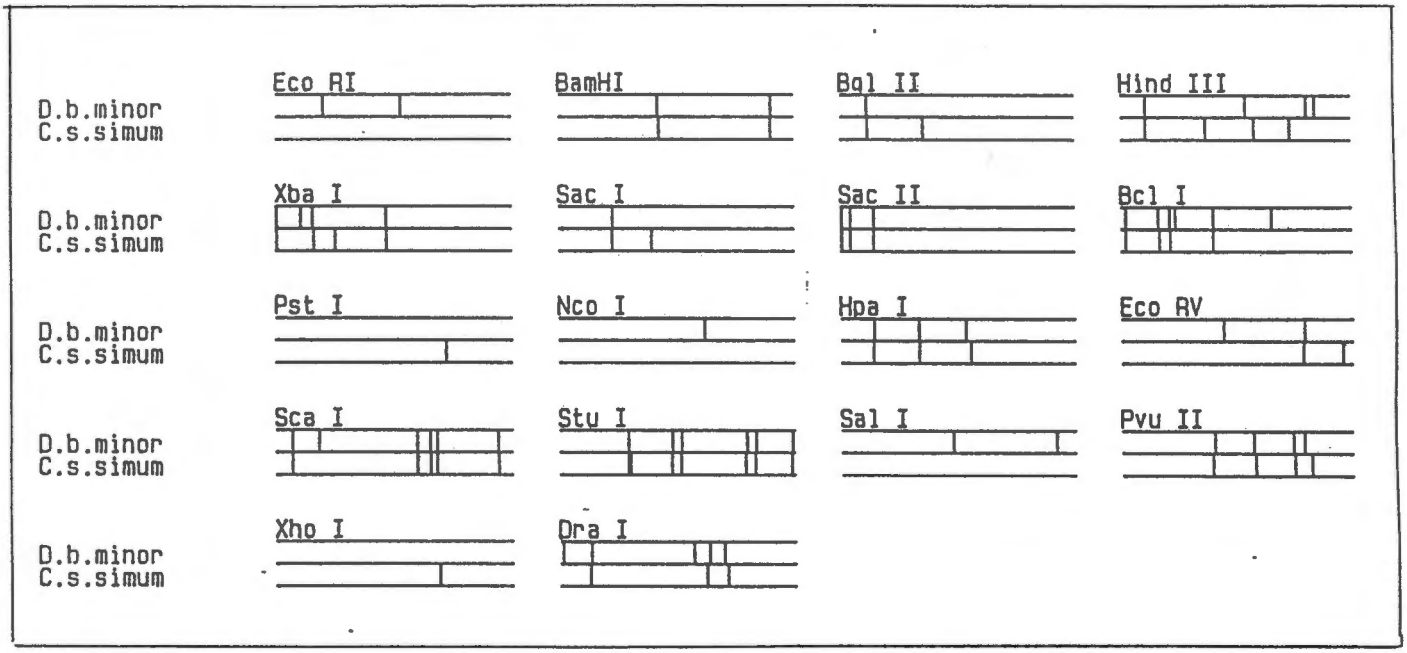

FIGURE 2.11

Individual restriction enzyme site alignments for *D. b. minor* and *C. s. simim.*

No polymorphic sites were found for any enzyme among the 24 black rhinoceroses (*D. b. minor*) samples, whilst only one Sca I polymorphic site was found amongst the 4 white rhinoceroses (*C. s. simum*) sampled. The polymorphic Sca I site is shown in figure 2.12.

The proportion of shared sites (s) between black rhinoceros and white rhinoceros was estimated by:

$$s = 2N_{xy}/(N_x+N_y),$$

N_y is the number of sites in the black rhinoceros,
N_x is the number of sites in the white rhinoceros and
N_{xy} is the number of sites shared (Nei and Li, 1979).

With $N_x = 52$ and $N_y = 42$ and $N_{xy} = 31$, the proportion of shared sites was estimated to be 0.667. The sequence divergence was calculated from this value using formula 9 of Nei and Li (1979), where sequence divergence, d, is given by:

$$d = (-3/2) \times \log [(4 \times s^{1/(2 \times R)})-1)/3], \text{ where}$$

R = mean length of the restriction enzyme recognition sequence (6 in this instance). The sequence divergence between the black rhinoceros and the white rhinoceros was 6.79% (with a standard deviation of 1.63%).

The initial rate of sequence divergence between two mammalian mtDNA lineages was estimated to be approximately 2% per million years (Brown et al 1979).

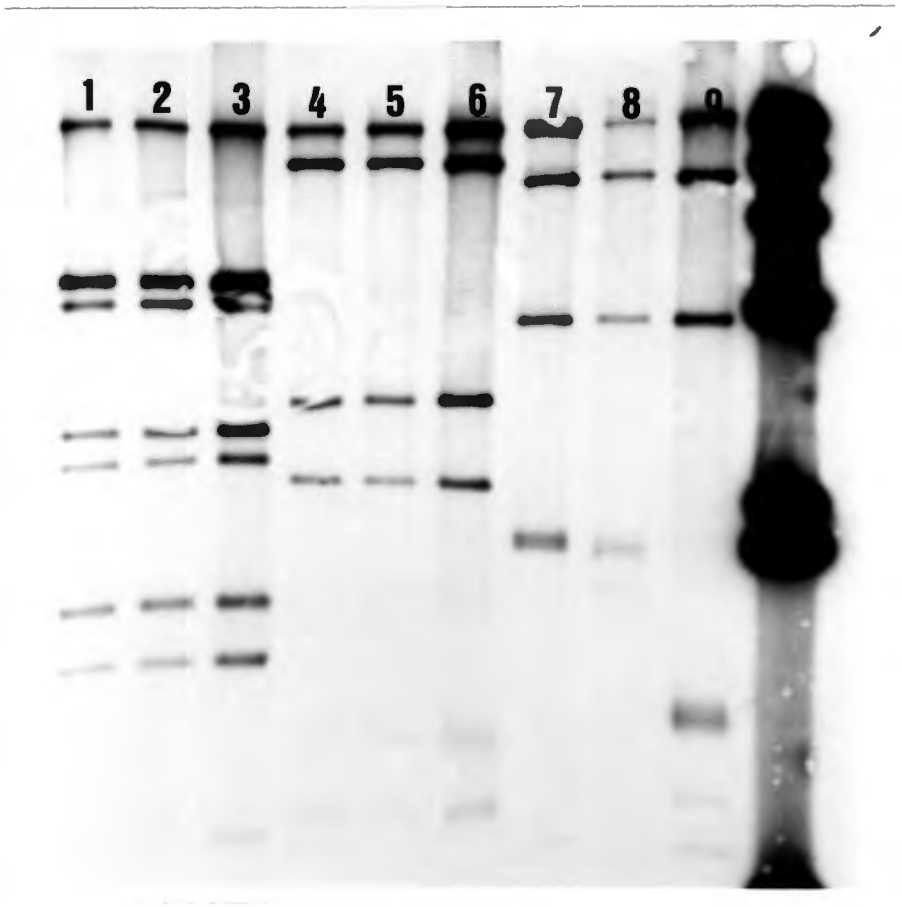

FIGURE 2.12

Restriction enzyme digest of white rhinoceros. DNA from three different individuals were digested with Stu I (lanes 1 - 3), Bcl I (4 - 6) and Sca I (7 - 9). The marker, lambda Hind III is in lane 10. The Sca I polymorphism is seen in lane 9, where the 2060 fragment is cut into two equal sized fragments.

If this holds true for the Rhinocerotidae (Wilson *et al.*, 1985), then it gives a time of divergence for these two lineages of 3.4 (+/- 0.8) million years. It must be made clear, however, that absolute time estimates for divergence events calculated in this manner are only approximate. Factors contributing to the uncertainty are not only the stochastic nature of the mutational process, but also the applicability of the calibration of the rate of sequence divergence against time for the group under study, and the amount of within species divergence, especially in the ancestral group at the time of the population separation leading to speciation.

Nevertheless, the value of 3.4 million year as a time of divergence between the black rhinoceros and the white rhinoceros is in accordance with the value of 3.5 million years suggested by George (1987) who used a restriction enzyme fragment comparison study to estimate the divergence time. The use of restriction enzyme site data is preferred over that of restriction fragments because two species may share restriction fragments, without necessarily sharing the same enzyme sites and *vice versa*. This could result in an over-estimation of the divergence time between two species. Hence, the use of restriction fragments are not recommended for the estimation of sequence divergence, especially between distantly related species. In addition, fragments are not suitable for use as character data in cladistic

methods of phylogeny construction, since they violate the principle of independence of characters required by cladistic methods (Hillis & Moritz, 1990).

The molecular differences between the black rhinoceroses and the white rhinoceros have also been examined using blood proteins and allozyme variation. When studying their blood proteins, no variation was found within black rhinoceros and very low levels of variation were found within white rhinoceros (Osterhoff & Keep, 1970). In a later study, 25 to 31 allozyme loci were studied in African and Indian rhinoceroses, where a lower than expected divergence between the black rhinoceros and the white rhinoceros was found (Merenlender *et al.*, 1989). This divergence value was used to calibrate the rhinoceros allozyme clock which yielded an amino acid substitution rate slower than other mammals, but comparable with some fish and reptiles.

The date obtained from mtDNA is concordant with fossil evidence. Although the fossil record of the Rhinocerotidae is fragmentary, there is a description of *Ceratotherium praecox* from deposits of about 4 million years ago (Hooijer, 1972). Because there is much similarity between *C. praecox* and both *D. bicornis* and *C. simum*, it is used to support the proposal that *Ceratotherium* split off from the *Diceros* lineage sometime during the Pliocene.

George and Ryder (1986) used restriction enzyme site comparisons of mtDNA in another family in the Perissodactyla to estimate that the common ancestor of the Equidae was present about 3.9 million years ago. This similarity to the figure of 3.4 million years in the African Rhinocerotidae may be coincidental, but contributes to the accumulation of a data set which may define major radiation episodes of African mammals in the Pliocene and the Pleistocene.

CHAPTER 3

INTRODUCTION TO THE POLYMERASE CHAIN REACTION

3.1 INTRODUCTION

The polymerase chain reaction (PCR) has revolutionized molecular biology and many seemingly-unrelated fields. The theoretical basis of PCR was probably first expounded in a paper by Kleppe *et al* (1971), but since the ingenious concept struck Kary Mullis in 1983 (Mullis 1990), PCR has rapidly spread to become a technique with widespread applications throughout the biological sciences.

This technique has enabled scientists to generate up to 100 billion copies of DNA from as little as one molecule in a few hours. PCR is the logarithmic amplification of a specific segment of DNA resulting in billions of copies of the segment being synthesized. The specificity is conferred by the annealing of a pair of oligonucleotide primers, which delineate the region of DNA to be amplified, to a heat

denatured DNA template, followed by the extension of the primers by a DNA-dependent DNA polymerase. The principle of PCR is illustrated in figure 3.1. The product synthesized in this first cycle serves as template for subsequent cycles, resulting in the exponential increase in number of a specific fragment. The theoretical amplification can therefore be expressed as 2^n, where n = the number of amplification cycles performed.

The first PCR experiments published (Saiki *et al.*, 1985, Saiki *et al.*, 1986 and Mullis & Faloona, 1987) were performed with DNA polymerase I or the Klenow fragment from *E. coli*. However, this enzyme has some problems: it lacks stability at the high temperatures (>90^0C) required to separate the DNA strands hence, necessitating the addition of fresh enzyme at the start of each extension cycle. Because the temperature optimum of Klenow is 37^0C, extension at this low temperature allows for significant non-specific annealing of the primers resulting in the synthesis of non-specific fragments (Saiki *et al.*, 1985).

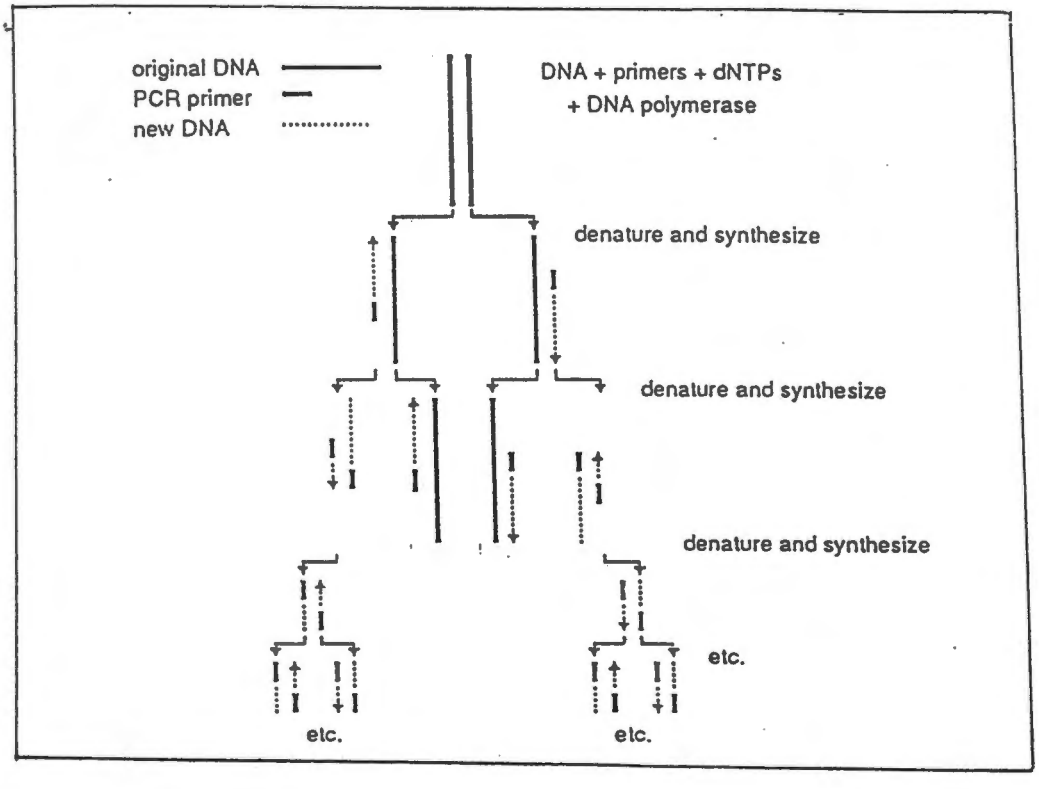

FIGURE 3.1

A schematic representation of the polymerase chain reaction (PCR).

3.2 TAQ POLYMERASE

The above disadvantages of using the Klenow fragment of DNA polymerase were resolved by the introduction of a thermostable DNA polymerase isolated from *Thermus aquaticus* (Taq polymerase). Saiki *et al* (1988) described the replacement of the *E. coli* DNA polymerase, Klenow, with Taq polymerase which can survive extended incubation at 95^0C. *Thermus aquaticus* strain YT1, a thermophilic, eubacterial microorganism capable of growth at 70 to 75^0C, was isolated from hot springs in Yellowstone National Park and was first described twenty years ago (Gelfand, 1989). This modification not only simplified the amplification procedure, making it amenable to automation, it also substantially improved the overall performance of the reaction.

Although Taq polymerase has a very limited ability to synthesize DNA above 90^0C (its optimum temperature for synthesis is 70^0C), the enzyme is relatively stable and is not rapidly denatured at temperatures above 90^0C. In the PCR reaction Taq polymerase retains about 50% of its activity after 130 minutes, 40 minutes, and 5 - 6 minutes at 92.5^0, 95^0, 97.5^0C, respectively. The use of Taq polymerase increases the sensitivity, specificity, yield and length of the target that can be amplified. The misincorporation rate

per nucleotide per cycle for Taq polymerase was estimated at 2×10^{-4}, which is somewhat greater than the 8×10^{-5} observed for Klenow. This increased misincorporation rate occurs because Taq polymerase lacks a 3′ to 5′ exonuclease proofreading function (Saiki *et al.*, 1988).

DNA polymerases from a number of other thermophilic organisms have also been purified and have also used in PCR reactions, but publications which make use of these polymerases have been rare (Skerra, 1992).

Vent DNA polymerase was isolated from *Thermococcus litoralis*, a themophilic marine archaebacterium which lives in submarine thermal vents and can grow at temperatures up to 98^0C (Neuner *et al.*, 1990). This polymerase has a 3′-5′ proofreading exonuclease activity and is more thermostable than Taq polymerase with a half-life of 60 minutes at 100^0C, whereas Taq is completely deactivated after 5 minutes at 100^0C.

Pfu DNA polymerase was isolated from the thermophilic marine archaebacterium, *Pyrococcus furiosus* (Marsh *et al.*, 1992). This multifunctional thermostable enzyme has the 3′-5′ exonuclease activity. The proofreading activity associated with the Pfu DNA polymerase results in a 12-fold increase in the fidelity of DNA synthesis over Taq DNA polymerase (Marsh *et al.*, 1992).

3.3 FACTORS THAT AFFECT PCR

3.3.1. PRIMER SELECTION

The length and the nucleotide sequence of the primers are probably the most important determinant for a successful amplification. Although a number of computer programs aiding the selection of primers are now available, the actual design of the primers remains largely empirical.

The following set of "rules" are useful to observe when designing primers (Rybicki, 1992):

1. Primers should be 17 to 28 bases in length;
2. The base composition should be 50 to 60% (G+C);
3. The primers should approach a random base distribution, with a GC content similar to that of the fragment being amplified;
4. Primers should have at the 3'-end a G or C, or CG, or GC because this prevents "breathing" of the end and increases the efficiency of priming;
5. A T_m (melting temperature) between 55 to 80^0C is preferred;
6. The 3'-ends of primers must not be complementary because this would result in primer-dimers being synthesized preferentially to any other product;

7. Primer self complementarity (ability to form secondary structures such as hairpins) and sequences with significant secondary structure should be avoided;
8. Stretches of three or more Cs or Gs at the 3′-ends of primers may produce mispriming at G or C-rich sequences (because of the stability of annealing), and should be avoided.

For phylogenetic studies, the amplification of DNA from a number of related organisms sometimes requires the use of general or "degenerate" primers. These primers are able to amplify the same gene from a number of different, but related organisms. "Degenerate" primers are a set of primers which have a number of options at several positions in the sequence, in order to allow the annealing to, and the amplification of, a variety of related sequences (Compton, 1990). Because mismatched nucleotides at the 3′-ultimate and 3′-penultimate positions have previously been shown to be detrimental to amplification, deoxyinosine (a deoxynucleotide analogue with inosine in the place of the base) has been successfully substituted in degenerate primers (Batzer *et al.*, 1991). The inclusion of deoxyinosine in the primer sequence, resulted in an amplification product from a range of non-human primates that diverged from one another as long ago as 25 million years ago.

3.3.2 REACTION BUFFER

The constituents of the PCR buffer have a marked effect on the success of the reaction. The concentration of the Mg^{2+} ions being the most important, influencing the specificity and the yield of the amplification product.

The standard reaction is done in a 50 or 100ul volume which, in addition to the DNA sample, contains 50mM KCl, 10mM Tris.HCl (pH8.4), 1.5mM $MgCl_2$ (this value is normally optimal with 200uM of each dNTP), 100ug/ml gelatin, 0.25uM of each primer, 200uM of each deoxynucleotide and 2.5 units Taq polymerase (Saiki, 1989). The concentration of the $MgCl_2$ is crucial for the success of the PCR. The $MgCl_2$ concentration has to be determined empirically for every primer-template set. Generally, an excess of $MgCl_2$ will result in the accumulation of non-specific amplification products, while insufficient $MgCl_2$ will reduce the yield.

Non-ionic detergents such as Tween-20, Triton X-100 and Nonidet NP-40 may also be incorporated into the reaction buffer to increase the specificity of the PCR (Innes and Gelfand, 1990). Formamide has also been used to increase the specificity of PCR when the DNA fragments have a high GC content (Sarker *et al.*, 1990) and glycerol has been used to

improve the amplification of G+C rich templates (Smith *et al.*, 1990). The PCR reaction can also be performed in buffers designed primarily for reverse transcriptase and *vice versa* (Fuqua *et al.*, 1990). This results in one-tube protocols where cDNA synthesis is followed by subsequent PCR with a minimum of separate steps.

3.3.3 CYCLING PARAMETERS

DNA amplification is performed by incubating the reaction samples sequentially at three different temperatures. The DNA was denatured briefly at 90 - 95^{0}C, the primers are allowed to anneal to the denatured template by cooling to 40-60^{0}C, followed by heating the annealed primers to 70-75^{0}C to extend the annealed primers with the Taq polymerase. The PCR normally comprises 30 cycles with more cycles added if trace quantities of old or degraded material is being amplified, in forensic cases or from ancient material. However increasing the number of cycles does not always necessarily improve the yield of the target DNA, but often increases the non-specific higher molecular weight fragments (Bell & DeMarini, 1991).

Each step in the cycle requires a minimum amount of time to be effective, whilst too much time at each step is both

wasteful of, and deleterious to, the DNA polymerase. If the amount of time in each step is reduced, the enzyme will retain a larger proportion of its activity because Taq polymerase activity decreases rapidly at temperatures higher than 94^0C (Gelfand 1989).

The upper limit of the annealing temperature depends on the "melting temperature" (T_m) of the primers. The T_m is related both to the length and the GC content of the primers and can be calculated [$T_m = 4(G+C) + 2(A+T)^0$C] (Maniatis *et al.*, 1982). At annealing temperatures lower than the T_m, non-specific binding and hence, non-specific amplification may occur. For primers with a high GC content, a higher annealing temperature would be necessary. A more rigorous treatment for the determination of annealing temperature is given by Rychlik *et al* (1990).

The extension time for the reaction is dependent on the length of the target being amplified. At 70^0C, the extension rate has been calculated to be greater than 60 nucleotides per second (Innes *et al.*, 1988). Theoretically this means that 3600 base pairs could be amplified in one minute. When amplifying larger fragments, increasing the extension time up to 15 minutes can improve the yield, but any longer extension times do not appear to have any significant effect on the yield (Jeffreys *et al.*, 1988).

3.4 DIRECT SEQUENCING OF PCR PRODUCTS

The primary utilization of PCR has been to provide DNA for subsequent sequencing. PCR has greatly simplified DNA sequencing, by eliminating many of the tedious and often technically challenging cloning steps.

Until recently, the amplified products were cloned before sequencing (Wong *et al.*, 1986 and Wong *et al.*, 1987a). However, direct sequencing of the amplified products has two major advantages over conventional cloning of PCR fragments into plasmid or viral vectors (Gyllensten and Erlich, 1988). Direct sequencing can be more readily standardized (and thus made more amenable to automation) since it is an *in vitro* system that does not depend on living organisms. It is faster and more reliable since normally only a single sequence needs to be determined for each sample (Ruano & Kidd, 1991).

Taq polymerase has no proofreading activity and a high misincorporation rate, therefore, when sequencing cloned products, more than one clone has to be sequenced to find the "consensus" sequence. This can be minimized by the use of the higher fidelity polymerases (Vent and Pfu) which has a proofreading exonuclease activity.

Direct sequencing produces a consensus directly, since any specific misincorporations will represent a very small subset of the amplified product molecules and will therefore not be seen. However, the direct sequencing of amplified products is not as straight-forward as it theoretically appears. The major difficulty of sequencing short double stranded fragments of DNA, as is typically produced by PCR, is the tendency of the complementary strands of the template to reanneal. A large number of methods are available to circumvent this problem in order to directly sequence the amplified products.

The "triple primer" method was utilized by Wrischnik *et al.*, (1987) to sequence double stranded template when investigating mutations in human mitochondrial DNA. In this approach, primer A and primer B are used to amplify the target DNA, whilst a third primer, primer C, is used to initiate sequencing of the target DNA. Primer C (in the cited example) is located 29 base pairs downstream of primer B (figure 3.2).

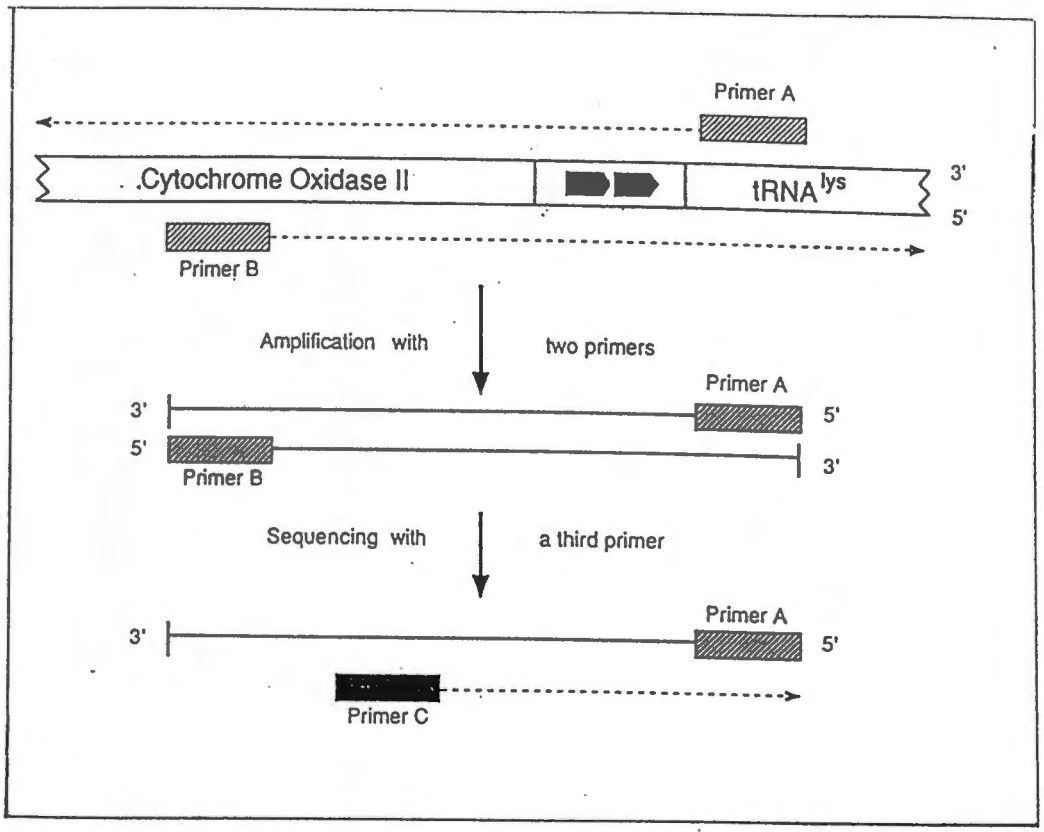

FIGURE 3.2

The triple method for sequencing region V of human DNA. Two 20 bp primers (A + B) match invariant sequences flanking region V. After amplification (with primers A + B), primer C is used to sequence the DNA directly. (Reproduced from Wrischnik *et al.*, 1987)

The generation of single stranded DNA fragments as the final product of PCR (asymmetric PCR) is a different approach to direct sequencing of PCR products. The single stranded DNA is generated by the use of unequal molar amounts of the two amplification primers (Gyllensten & Erlich, 1988 and Allard *et al.*, 1991).

The principle of single stranded amplification is shown in figure 3.3. The two primers are present in a ratio of 100:1. During the first few cycles double stranded DNA will be exclusively produced. However, when the primer added in limiting amounts has been used up, an excess of single stranded DNA will be produced by linear amplification in each subsequent cycle. Theoretically, the amount of double stranded DNA should increase exponentially, whereas the production of the single stranded DNA should follow a linear increase. Therefore, it is important that the production of double stranded DNA is allowed to reach an adequate level before single stranded DNA production is initiated (figure 3.4).

A number of different modifications to standard methods are available for the direct sequencing of amplified products. These include direct sequencing of double stranded product purified from low-melting point agarose (Kretz *et al.*, 1988); the production of single stranded template by exonuclease digestion after the PCR reaction (Higuchi &

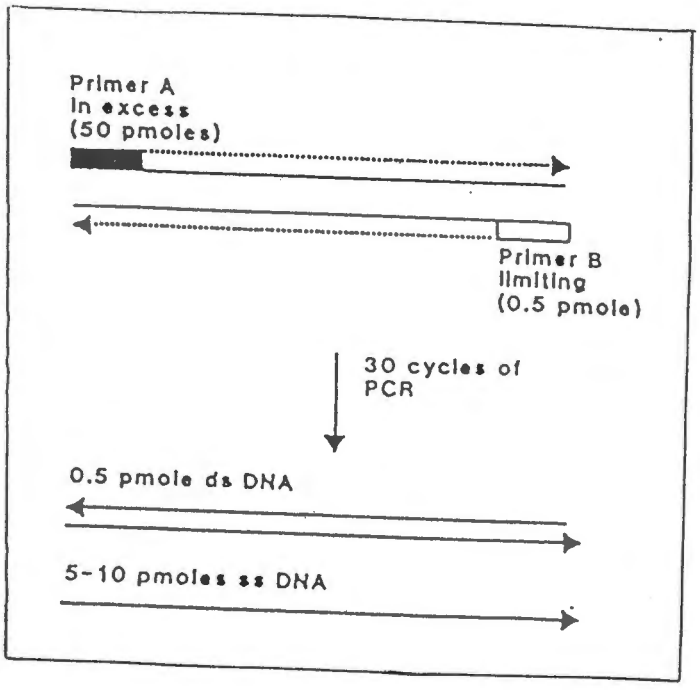

FIGURE 3.3

An outline of the procedure for generating ssDNA by PCR. The concentrations of the primers used in the amplification were initially set at a 100:1 ratio. After 0.5pmol of dsDNA had been generated, ssDNA started to accumulate. (Reproduced from Gyllensten & Erlich, 1989)

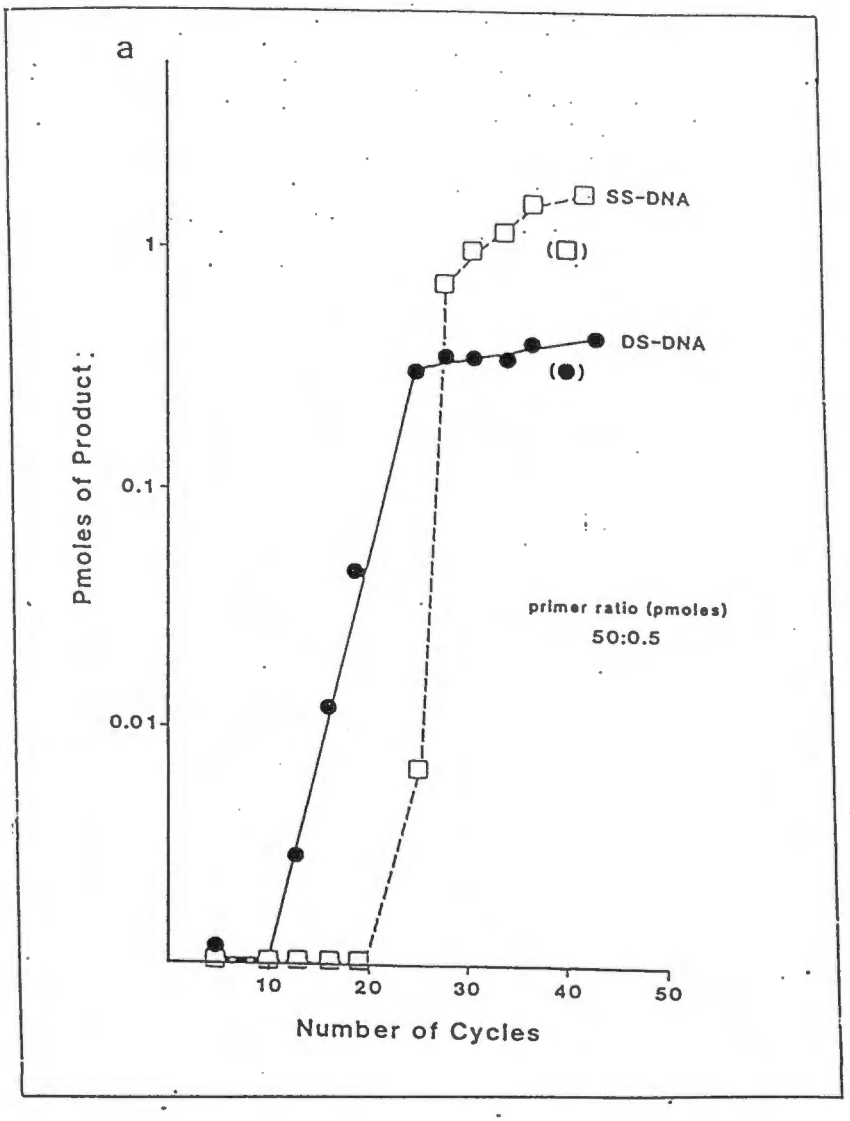

FIGURE 3.4

The accumulation of dsDNA and ssDNA at primer ratios of 100:1. The single stranded product is only seen from cycle 20, after which there is a linear increase in product. (Reproduced from Gyllensten & Erlich, 1989)

Ochman, 1988); the use of DMSO during the sequencing reaction (Winship, 1989); direct sequencing of the double stranded product after the removal of the complementary strand with single stranded DNA of a M13 clone (Gal & Hohn, 1990) and the use of detergent in the sequencing reactions (Bachmann *et al.*, 1990).

The optimum conditions for the direct sequencing of double stranded PCR products have been investigated (Casanova *et al.*, 1990). The following set of parameters were found to be crucial for the successful sequencing of double stranded products with Sequenase (modified T7 polymerase): (i) snap cooling to $-70^{0}C$ for the annealing of the primer to the template DNA, (ii) an incubation of 45 seconds for the labeling reaction and (iii) a primer:template ratio of at least 20:1.

All of the above methodologies employ the standard dideoxy method (Sanger *et al.*, 1977) of sequencing using the Sequenase sequencing kit (United States Biochemical). A detailed protocol was published by Innes *et al.*, (1988) describing the amplification and the direct sequencing of the amplification products using Taq polymerase. The advantage of using Taq polymerase for sequencing is that the enzyme operates at a higher temperature than the Klenow fragment and the sequencing ladders are less susceptible to background bands or idiosyncrasies due to secondary

structure in the template. The sequencing ladder should be readable over a longer distance and have uniform intensity. The ability of Taq polymerase to operate at higher temperatures and low salt conditions allows the heat destabilization of secondary structures such as hairpin loops in the DNA. Although Taq polymerase has a higher misincorporation rate than the Klenow fragment, the direct sequencing of the PCR products by any method is a "consensus" sequence. Those bases that occur at any given position in the majority of molecules will be the most visible on an autoradiograph and any bases at low frequency will be undetectable.

3.5 APPLICATION OF PCR TO EVOLUTIONARY GENETICS

Molecular approaches in evolutionary genetics have two principle objectives: i) the reconstruction of phylogenies and ii) the analysis of populations. Zuckerkandl and Pauling (1962) first proposed that the comparisons of the sequences of proteins and genes could be used as a molecular clock to date the divergence time of extant species from their common ancestor. Since then, a variety of biochemical methods have been used to assist in the reconstruction of phylogenies, to calculate the time of divergence between extant species and for the analysis of genetic variation in populations within a species.

Protein electrophoresis, immunological comparisons, protein sequencing, DNA:DNA hybridization, mitochondrial DNA fragment length comparisons and restriction site mapping have all made major contributions to systematics (Hillis & Moritz, 1990). However, most of these techniques are somewhat limited because they estimate sequence divergence values by indirect means, rather than by the direct comparison of DNA sequences.

Population geneticists require techniques which are able to discriminate between different individuals or different populations when screening a large number of individuals. Until recently, it has been impractical to generate sequence data for more than just a few organisms because of the effort required to clone and sequence DNA from many individuals. The polymerase chain reaction and direct sequencing now renders it feasible to screen large numbers of individuals for genetic variability in their DNA sequences, for evolutionary studies.

Although PCR requires a set of specific primers to amplify the target sequence, the lack of complete sequence information does not impede the application of PCR to evolutionary studies. The use of "universal" primers circumvents this apparent lack of prior sequence knowledge

of the new species under scrutiny (Kocher *et al.*, 1989). By choosing conserved sequence elements that bracket the region to be sequenced, primers can be designed that will amplify the target gene fragment from all the members of the taxonomic group of interest.

Such "universal" primers were designed to amplify the mitochondrial cytochrome-b gene of three orders of hoofed mammals (Kocher *et al.*, 1989). This study obtained 17 complete gene sequences directly from the amplified products. The phylogeny obtained, was consistent with those obtained from previous methods. This sequence information from the cytochrome-b gene can also be used to understand how the rate of mutation in the different genes of mitochondria varies and contributes to the understanding of the evolution of organelle DNA.

Mutations have been studied using PCR to examine polymorphisms in the mtDNA of human populations (Vigilant *et al.*, 1988) and have also been utilized in the ongoing debate, as to the origins of populations of modern man (sometimes popularized under the misleading title of "mitochondrial Eve") (Cann *et al.*, 1987). After amplification, sequence data from 189 people of diverse geographical origins, including 121 native Africans were obtained and the phylogenetic relationships were

reconstructed (Vigilant *et al.*, 1991). Analysis of this sequence data produced, was used to support the idea of an African root for the tree relating human mtDNAs.

Another contentious and seemingly intractable problem in mammalian systematics has been the phylogenetic resolution of the human, chimpanzee and gorilla trichotomy. The mtDNA sequences of the cytochrome oxidase subunit II gene were amplified, sequenced and compared for the above three organisms (Ruvolo *et al.*, 1991). Data from this study provide clear-cut evidence for the separation of the African ape trichotomy into two evolutionary lineages, one leading to gorillas and the other to humans and chimpanzees.

The same approach has been used to amplify nuclear genes as well as rRNA genes from a diverse range of organisms. Additional examples of recent studies are: sequences of different amplified genes were used to determine the phylogeny of antlered deer (Miyamoto *et al.*, 1990), Xantusiid lizards (Hedges *et al.*, 1991) and African mole-rats (Allard & Honeycutt, 1991).

The sensitivity of PCR allows for the amplification and sequencing of extremely small quantities of material. Amplification from single hairs (Huguchi *et al.*, 1988 and Vigilant *et al.*, 1989) and from single cells (Jeffreys *et*

al., 1988) have been reported. The mtDNA amplified from single hairs (Vigilant *et al.*, 1989) was used to analyse sequence variation among 15 members of the !Kung (Bushman) population found in Southern Africa.

The enormous sensitivity of PCR allows access to a wealth of material, previously inaccessible to evolutionary biologists: namely, skin and bone (Hagelberg *et al.*, 1991) from museum specimens and even from extinct organisms. The advent of PCR has made it possible to isolate DNA sequences from a few copies of intact DNA present in specimens where the majority of the molecules are damaged to an extent, that prevents analysis by traditional DNA techniques. However, the sensitivity of PCR makes it liable to amplify even very minor contaminants of contemporary DNA present in the ancient samples. This is especially problematic for human studies. Modifications in the old template DNA, as well as other components in the extracts of tissues may inhibit the DNA polymerase and may require modifications to the standard procedures to overcome the problem.

DNA has been extracted from a 7 000 year old human brain excavated from a sinkhole in Florida (Paabo *et al.*, 1988) and the sequences obtained showed that this ancient individual belonged to a mitochondrial lineage which is rare in the Old World and not previously known to have been involved in the prehistoric colonization of the New World.

DNA has been successfully extracted and amplified from samples of soft tissues that varies in age from 4 to 13 000 years (Paabo, 1988). "Ancient" DNA has been extracted from magnolia leaves of about 7 x 10^6 years old (Golenberg *et al.*, 1990) and from a termite preserved in amber of greater than 20 x 10^6 years old (De Salle *et al.*, 1992). Although the DNA extracted from ancient material is often badly degraded and present only as low molecular weight fragments, it is possible to amplify DNA fragments larger than any of the molecules present in the starting sample because Taq polymerase is able to "jump" between templates during amplification (Paabo *et al.*, 1990b).

3.6 OTHER APPLICATIONS OF PCR

The other research and medical applications of PCR are numerous and the list of the applications for PCR increases daily (Erlich, 1989). Although sequence determination is the most direct and comprehensive method for analyzing the genetic variation in the amplified DNA, a variety of other indirect ways for detecting mutations have also been

developed. A powerful indirect method for detecting mutations or polymorphisms is the use of denaturing gradient gel electrophoresis (DGGE) (Lessa, 1992). PCR is also a powerful tool for the detection and quantitation of gene expression (Kawasaki *et al.*, 1989). Even the PCR constraint of having some sequence knowledge of the gene of interest, is overcome in a number of ways. "Anchored PCR" (Loh *et al.*, 1989) and "RACE" (Frohman *et al.*, 1988) are some of the approaches used to amplify DNA when little or no prior sequence information is available. "Inverse PCR" or "inside-out" PCR has been developed for the analysis of sequences that flank a known region (Ochman *et al.*, 1989). The use of "universal" and degenerate primers has been discussed earlier to amplify DNA from related organisms.

3.7 USE OF PCR IN THIS STUDY

In this study PCR was used to amplify segments of mtDNA from the cytochrome-b region. The amplified DNA was sequenced and this information was used to supplement the restriction mapping estimates of the time of divergence from a common ancestor between the black rhinoceros and the white rhinoceros. The sequences will also be used to determine the genetic variability among four of the Southern Africa subspecies of black rhinoceros. An additional novel

application of PCR in this study was a forensic application, in which DNA from an axe-head purportedly used by poachers in the dehorning of a rhinoceros in Natal. This is a useful demonstration of how the results of initially academic research can be used to address practical problems of conservation management.

CHAPTER 4

THE USE OF DNA SEQUENCES TO ESTIMATE THE TIME OF DIVERGENCE BETWEEN BLACK RHINOCEROS AND WHITE RHINOCEROS

4.1: INTRODUCTION

A 300 bp fragment of the cytochrome-b region of mtDNA was amplified by the polymerase chain reaction (PCR). The amplified DNA was sequenced. The sequence information was used to estimate the time of divergence between the black rhinoceros and the white rhinoceros, and more importantly to contribute to an eventual comprehensive data set which can be used in a wider systematic context when comparable regions from related taxa are sequenced.

4.2: PRIMER DESIGN

The primers used for the amplification of the cytochrome-b fragment were chosen after the examination of published primer sequences (Kocher *et al.*, 1989) and computer-generated sequence homology searches. The published sequences of the cytochrome-b fragment of bovine, human and mouse mtDNA were compared for short stretches of homology. This sequence information, used in conjunction with the published "universal" primers (Kocher *et al.*, 1989), aided the design of the primers used in this study. When choosing primers it is important to avoid complementary stretches of bases at the 3'-OH end of the primers, long runs of the same base, and primer self-complementarity that would promote secondary structures or too low a melting temperature. Table 4.1 shows the sequences of the primers used to amplify and sequence the cytochrome-b region of mtDNA.

4.3: STANDARD PCR REACTION

4.3.1: DOUBLE STRANDED PCR

The reactions were done in 100ul final volumes. The following reagents were needed for each PCR reaction:

i) The primers were each used at a concentration of 100pmol.

TABLE 4.1

The primers used to amplify the cytochrome-b region of mtDNA. Primers A and B amplified a 300 bp fragment and primer C was used as an internal primer to sequence part of the amplified fragment.

Gene of mtDNA	Primer name	Sequence
cytochrome-b	primer A	GACCTCCTAGGAGACCCAGA
	primer B	AATGGGTGTTCTACTGGTTG
	primer C	GTTTATTCTTTATAGTGAG

ii) The four deoxynucleotides were each used at a concentration of 150uM. iii) A 10x buffer was used at a 1x final concentration in the reaction. The buffer components were 50mM KCl, 10mM Tris pH 8.0, 0.01% gelatin and 0.1% Triton X-100. iv) The concentration of $MgCl_2$ was empirically determined for each template-primer combination. The majority of the black rhinoceros samples required 4 to 5mM $MgCl_2$ for amplification. v) Taq polymerase (Promega) was used at a concentration of 1 to 2 units per sample. vi) To determine the optimum DNA concentration for each PCR, a template titration was done for each new sample amplified. vii) Sterile, distilled water was used to make the reaction volume to 100ul. viii) When all the components of the reaction had been assembled, the PCR reaction was covered with 100ul of mineral oil.

4.3.2 **SINGLE STRANDED PCR**

The single stranded PCR reaction was performed in a similar manner to that of the double stranded PCR, except that only one primer was used during the amplification. A double stranded PCR reaction was done to establish the conditions for the amplification and the product was visualized by gel electrophoresis.

An aliquot, normally 1ul, of the amplified product was used as template for a second, single stranded PCR. In single stranded PCR, 100pmol of only the one primer was used.

An alternative way to produce single stranded product was by asymmetric PCR. An excess of one amplified strand (relative to its complement) was generated by the addition of one primer in excess over the limiting primer. The PCR reaction was set up as for the double stranded PCR, but a 100:1 ratio of the two primers were used to generate the single stranded product.

4.4: PCR CYCLING CONDITIONS

The PCR cycling parameters are dependent on a number of factors: the level of homology between the template and the primer, the nucleotide composition of the primer, and the distance between the primers. The annealing temperature of the reaction dictates the specificity of the product being produced, which is determined by the nucleotide composition of the primers. An indication to the optimum annealing temperature was the Tm calculated from the primer base pair composition.

The cycling parameters for the amplification of the cytochrome-b fragment of mtDNA were: denaturation at $94^{0}C$

for 1 minute, annealing at 52^0C for 1 minute and extension at 72^0C for 1 minute, for 30 cycles in a thermal cycler (Techne or Hybaid).

After the PCR reaction, 1/10 of the total reaction was run on a 2% agarose and EtBr-stained gel to ascertain the success of the reaction.

4.5 PURIFICATION OF PCR PRODUCTS

4.5.1 INTRODUCTION

The amplified mtDNA fragments were sequenced directly after amplification, without any intermediate cloning steps. The same set of primers, as used for amplification were used for the sequencing of the mtDNA PCR products. Purification of the amplified product was required to separate the product from the template DNA, primers and other reaction components.

The purified DNA was used either directly for sequencing or for cloning into pUC-18 vector before sequencing (section 5.3).

4.5.2 QIAEX METHOD

This gel purification method has been employed with great success on double stranded DNA. QIAEX (Qiagen) is a kit which contains all the components required to extract the DNA from an agarose gel slice. QIAEX is a suspension of silica beads to which DNA can bind, or be eluted from, depending on the salt and temperature conditions. The QIAEX method of DNA purification is described below:

i) The amplification products of several successful PCR reactions were pooled and the volume was concentrated to 20ul by ethanol precipitation. The concentrated sample was electrophoresed on a 2% agarose gel, EtBr-stained and visualized using a hand-held long-wavelength UV lamp. The amplified DNA was excised from the gel in a wedge of agarose. Exposure to the UV was limited, to prevent UV damage to the DNA.

ii) For every 100mg of gel, 300ul of solubilization QX1 buffer was added. The QIAEX suspension was vortexed until a homogeneous suspension was obtained of which 10ul was added to the gel slice and the QX1 buffer mixture. The mix was incubated at 50^{0}C for 10 minutes to solubilize the agarose. The sample was agitated, by flicking it every 2 minutes, to keep the QIAEX in suspension.

iii) The sample, which now contained the DNA bound to the QIAEX, was centrifuged for 30 seconds in a micro-centrifuge. The supernatant was discarded.

iv) The pellet was resuspended in 500ul of the QX2 buffer and centrifuged for 30 seconds. The supernatant was discarded and this step was repeated.

v) The pellet was resuspended in 500ul QX3 buffer, centrifuged for 30 seconds and the supernatant was discarded. This wash step was repeated. The pellet was then centrifuged for an additional 30 seconds and any remaining QX3 buffer was removed . The pellet was air-dried for 10 minutes.

vi) To elute the DNA from the QIAEX, the dried pellet was resuspended in 20ul of TE buffer and incubated at room temperature for 5 minutes. The tube was periodically flicked to keep the QIAEX in suspension.

vii) The sample was centrifuged for 30 seconds and the supernatant was transferred to a clean microfuge tube.

viii) The elution steps (vi) and (vii) were repeated and the two elutes were combined. The purified DNA was then used in the sequencing reactions. As much as 70% of the DNA was recovered in this manner.

4.5.3 PAGE PURIFICATION

This gel purification method was utilized when purifying single stranded from double stranded amplification products. A 4% to 6% PAGE (PolyAcrylamide Gel Electrophoresis) gel was run to separate the single stranded from the double stranded DNA. Because single stranded and double stranded DNAs have different tertiary structures, they migrate to different positions on a gel. PAGE gels were used to separate the two forms of DNA because better resolution was obtained with PAGE gels than with agarose gels. The amplification product of several PCR reactions were pooled, concentrated by ethanol precipitation and run on a 4% to 6% PAGE gel. The gel was EtBr-stained and the two different bands (one being the single stranded and the other the double stranded PCR product) of DNA was excised from the gel. The DNA was eluted from the gel slice by a standard method described in Maniatis *et al* (1982). Only 40% of the DNA was recovered in this manner.

4.5.4 MILLIPORE FILTER UNITS

This method operates on a size exclusion principle. Small filter units (Millipore), which fit into microfuge tubes, exclude DNA larger than a particular molecular weight cut-off from passing through the filter. After centrifugation,

the molecules larger than the molecular cut-off are retained above the filter and the smaller ones are passed through the filter. For the mtDNA amplification products a Millipore filter unit of 30 000 molecular weight cut-off was used.

The amplification products of several PCR reactions were used. The samples were loaded onto the filter unit. The maximum volume was 400ul per filter unit. The filter unit was centrifuged in a micro-centrifuge at 6 500 rpm until all of the sample had been filtered. The filtrate was discarded and the filter unit, containing the larger DNA product molecules, was washed twice with 400ul TE buffer. The DNA was removed from the filter with 20 ul of TE buffer.

4.6 SEQUENCING

The purified mtDNA fragments were sequenced by the dideoxy method (Sanger, 1977). The Sequenase version 2.0 kit (United States Biochemical Corporation, Cleveland, Ohio) was used and the following modifications were used when sequencing the amplification product.

4.6.1 Sequencing reactions

i) Annealing reaction: A minimum primer:template ratio of 2:1 was used (Casanova, 1990). Reaction buffer was supplied in the kit and for double stranded template 2ul of the detergent mix (Nonidet NP-40 and Tween-20) was added. The final volume of 10ul for the reaction was made up with distilled water.

The double stranded template was denatured at 94^0C for 3 minutes and snap cooled in a dry ice/ethanol bath. The annealed template was kept frozen until required for the labeling reaction. The single stranded template was denatured by heating to 65^0C and no detergent mix was added to the reaction.

ii) Labeling reaction: The labeling reaction was done as recommended by the manufacturers of the Sequenase kit, with only the concentration of labeling mix of the unlabeled deoxynucleotides being changed. The template was labeled using ^{35}S-dATP. The dilution factor of the diluted labeling mix was lower (1:15) when sequencing close to the primer and higher (1:4) when sequencing further from the primer. The labeling reaction was performed at 20^0C for 2 minutes or at 12^0C for 4 minutes (Casanova 1990).

iii) Extension-termination reaction: This reaction was performed as recommended in the Sequenase kit. For double stranded template, 0.5ul of the detergent mix was added to the dideoxynucleotides.

4.6.2 Sequencing gel

The labeled fragments of the template were separated on a 6% denaturing polyacrylamide gel. The denaturant, urea, was included in the gel at a concentration of 8M. The sequencing reactions were denatured at 94^0C for 2 minutes before loading onto the gel, and were electrophoresed until the xylene cyanol marker (in the loading buffer) had reached the end of the gels. Bands higher in the sequencing ladder were resolved with longer electrophoresis runs.

The sequencing gels were fixed (5% methanol and 10% acetic acid) for 30 minutes in the fixing solution and dried at 80^0C on a vacuum-pump gel drier before autoradiography.

4.5 RESULTS AND DISCUSSION

The amplified fragment of cytochrome-b was sequenced directly after amplification as well as after cloning (methods in chapter 5). The 300 bp fragment was sequenced from both directions, using the amplification primers (primer A and B) as the sequencing primers.

A number of different methods were used to purify the amplified product before direct sequencing. Figures 4.1 and 4.2 show the sequencing gel after different purification methods. Purification for direct sequencing required the successful removal of the amplification primers. When single-stranded DNA was amplified, a variety of purification methods could be used because there was only one primer in the PCR reaction and the other primer could be used for sequencing. However, when sequencing double-stranded DNA, the gel-based purification methods (PAGE or QIAEX) gave better results (figure 4.3). An additional problem when sequencing directly after amplification, is the quality of the primers. If the same set of primers is used for both amplification and sequencing, the quality of the primers may be adequate for amplification, but not for sequencing.

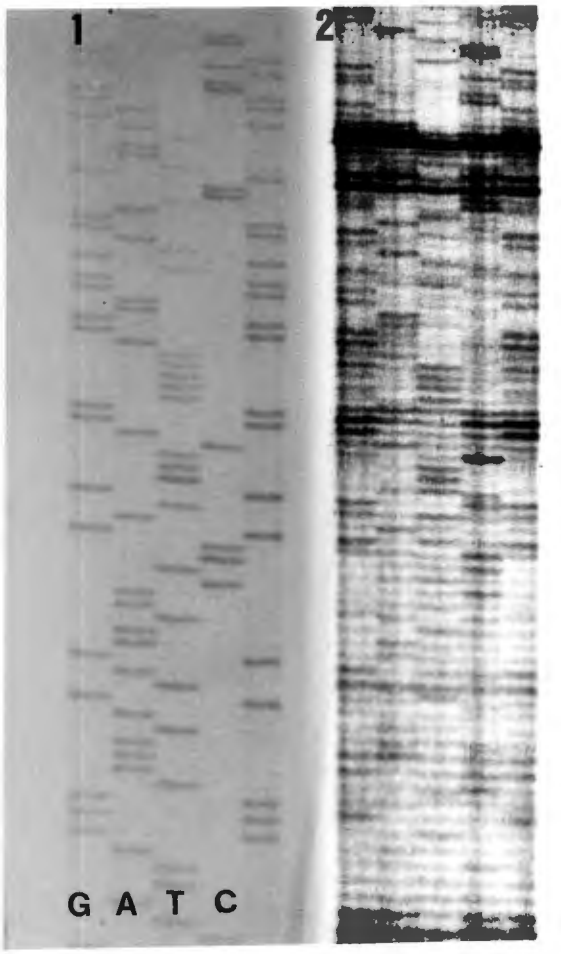

FIGURE 4.1

A sequencing gel after the amplified DNA was purified using:

1. PAGE 2. Millipore filter unit

FIGURE 4.2

A sequencing gel after the amplified DNA was purified using QIAEX.

FIGURE 4.3

A sequencing gel after the amplified DNA of black rhinoceros was purified using QIAEX.

This may be a result from a problem in the synthesis and subsequent purification of the primers, if sequences that are shorter than the complete primers are not effectively removed or the slow degradation of the primers by DNAses. When such primers are used for sequencing, bands in more than one lane of the sequencing gel results.

The mtDNA sequence of black rhinoceros and white rhinoceros, sequenced in this study (figure 4.4), were entered into a computer program, DAPSA, which stores, displays and manipulates the data (Harley, 1992b). The published sequences from black rhinoceros (Irwin *et al.*, 1991) and cattle (*Bos taurus*) from Genbank were also loaded into DAPSA. The sequences of the above four species were aligned and the percentage of sequence similarity was calculated (figure 4.5). In an earlier study which looked at dynamics of the mtDNA evolution in a diverse group of animals, the sequences of 111 of the 239 base positions of cytochrome-b were found to be invariant (Kocher *et al.*, 1989). Cytochrome-b is a relatively conserved region of the mitochondrial genome.

FIGURE 4.4

A sequencing gel of black rhinoceros and white rhinoceros DNA.

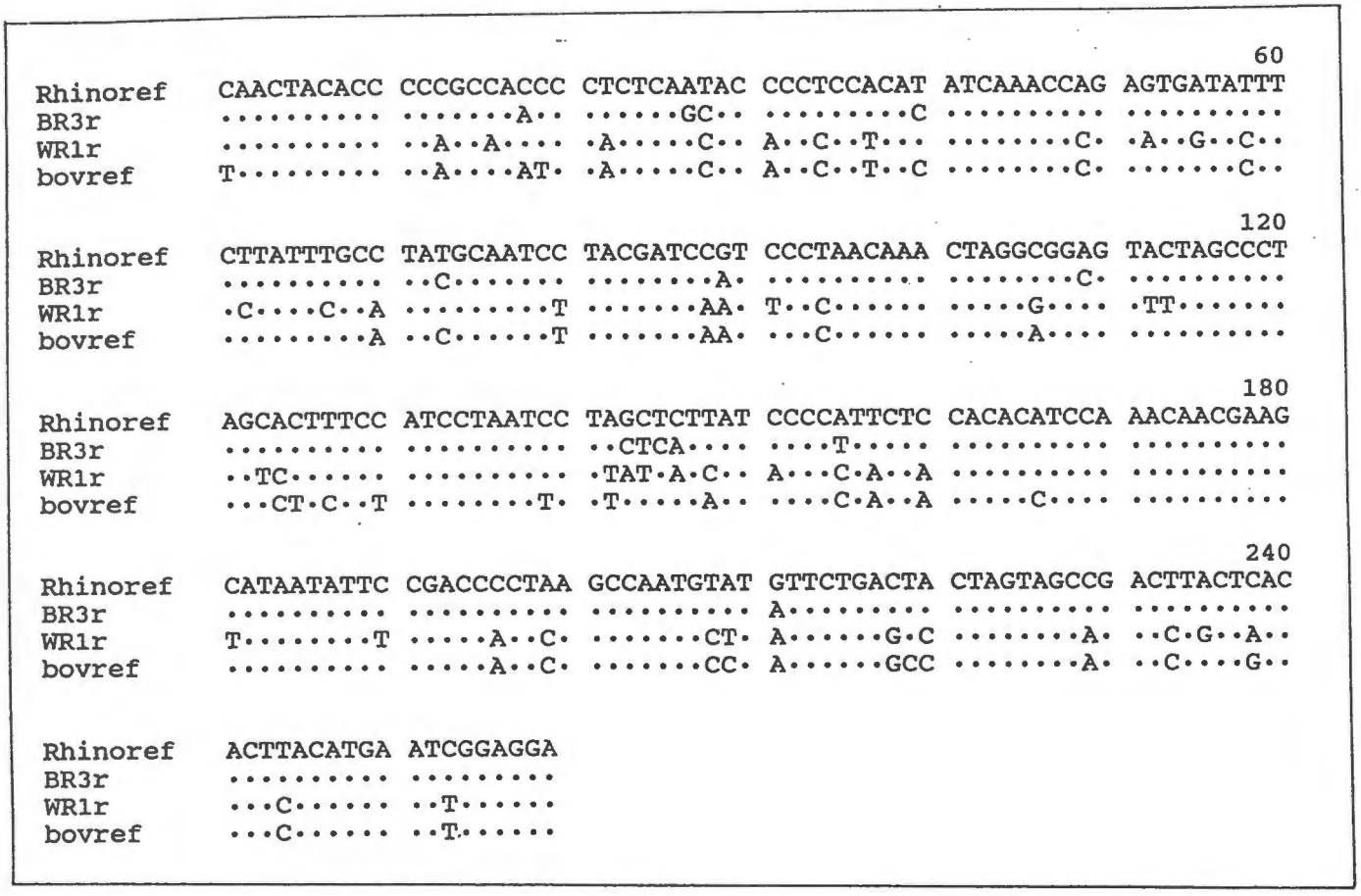

FIGURE 4.5

The aligned DNA sequences of: 1. The reference black rhinoceros, 2. black rhinoceros, 3. white rhinoceros and 4. the reference *Bos taurus* (the reference sequences were obtained from Genbank). Where a base in sequences 2 to 4 is the same as the reference sequence on the top line, it is depicted as a dot, in order to emphasize base differences in the sequence.

The percentage of sequence similarity between the published black rhinoceros sequence (Irwin *et al.*, 1991), and the black rhinoceros sequenced in this study (figure 4.5) was 94.9%, which was surprisingly low, a more expected value would have been 99-100%. The changes between the published black rhinoceros sequence and the sequence found in this study could be broken down into 3 first codon, 3 second codon and 8 third codon changes which would result in 6 amino acid replacements (figure 4.7). These amino acid replacements were always of the same "type", i.e, that a basic amino acid was always replaced with another basic amino acid, an acidic with an acidic, a hydrophobic with a hydrophobic and a hydrophilic with a hydrophilic. The black rhinoceros sequenced in this study was of the subspecies *D. b. minor*, whilst the subspecies designation of the published black rhinoceros sequence was unreported.

The imperfect match between the two black rhinoceroseses sequences may have a number of explanations. Firstly, they may not be of the same subspecies. This explanation would, however, only account for a small amount of sequence divergence between the two black rhinoceroses. Secondly, misreading the sequencing gel, especially in the higher molecular size region of the gel where the bands are compressed, could theoretically be a source of error for the sequence differences.

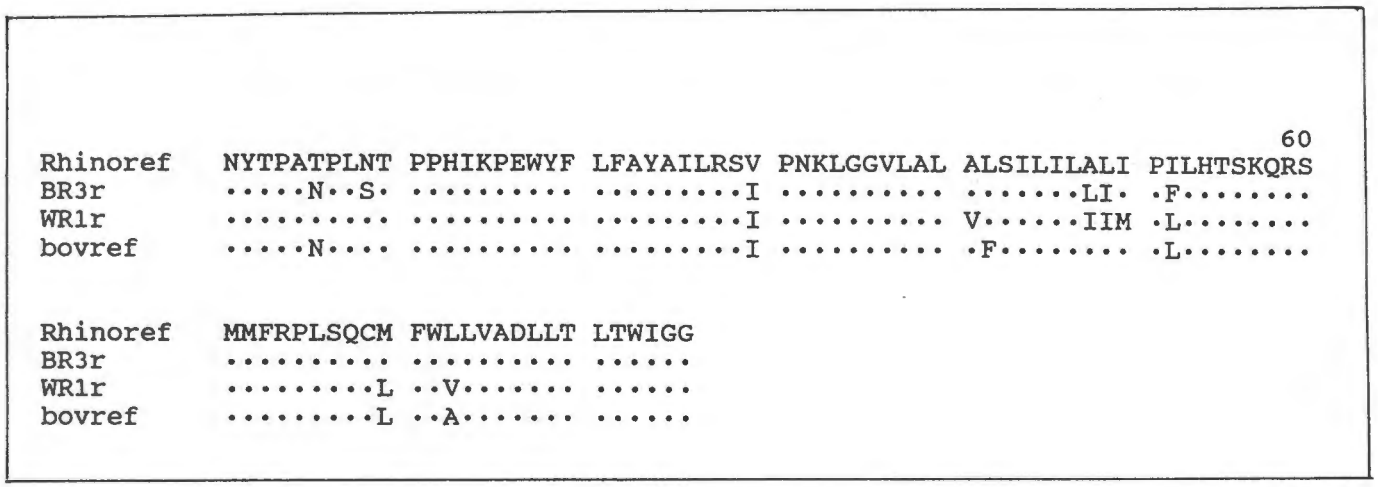

FIGURE 4.6

Alignments of the amino acid sequences derived from the DNA sequences: 1. The reference black rhinoceros, 2. black rhinoceros, 3. white rhinoceros and 4. the reference *Bos taurus* . Where an amino acid in sequences 2 to 4 is the same as the reference sequence on the top line, it is depicted as a dot, in order to emphasize the differences in the sequence.

However, this is unlikely in our sequence since the regions of sequence where the differences between the two black rhinoceroses occur, could be unambiguously read (figure 4.4) and the sequences were reproducible. The sequence differences were also confirmed with the other primer and reading the complementary strand.

The percentage of bases aligned between black rhinoceros and white rhinoceros was 81.0% (figure 4.5), which corresponds to a sequence dissimilarity of 19%. This corresponds to 8 first codon, 3 second codon and 38 third codon changes which result in 8 amino acid replacements of the same "type" (figure 4.6).

The percentage of sequence similarity (81.0%) between black rhinoceros and white rhinoceros can be transformed into distance data. Evolutionary distance or sequence divergence can be calculated from sequence similarity (Hillis & Moritz, 1999).

Sequence similarity, s = number of shares bases/length of aligned sequence, $d = -b \ln [1 - (1-s)]$

and, $b = 1 - [(f_A+f_G)^2+(f_C+f_T)^2]$

where d = evolutionary distance and b is the dissimilarity of infinitely diverged (completely randomized) sequences.

For example, if nucleotide sequences are being compared and all four nucleotides are equally likely, then b = 0.75 (Jukes & Cantor, 1969). In the present comparison, b was calculated to be 0.72. The sequence divergence between black rhinoceros and white rhinoceros was calculated to be 21.9%. The sequence divergence value obtained from restriction enzyme mapping (chapter 2) was 6.79%. However, due to time constraints, since this was the last component of the project to be performed, only two black rhinoceroses and one white rhinoceros were sequenced. To confirm this discrepancy between the above two methods, more samples will have to be sequenced, as well as a longer fragment of DNA sequenced.

Although the sequence divergence value between black rhinoceros and white rhinoceros, obtained from restriction enzyme mapping and sequence data is different, there is no precedent with which our results can be compared. The reason for this is that, until recently, sequencing (before the discovery of PCR) has been a relatively difficult task.

Hopefully, results like those presented on this study, where both restriction enzyme mapping and sequence information are presented, will contribute to a database which can be used to investigate the merits of the two different methods and their relative suitability in addressing the appropriate evolutionary question.

CHAPTER 5

NOVEL APPLICATION OF PCR FOR THE IDENTIFICATION OF THE SPECIES OF ORIGIN IN TRACE QUANTITIES OF BIOLOGICAL MATERIAL

5.1 INTRODUCTION

The long term survival of the rhinoceros is being threatened by habitat destruction and by continued poaching. The reasons for the increase in the poaching of rhinoceroses include the financial remunerations for the poachers and inadequate policing of the animals by game reserve rangers (as a result of financial and manpower constraints). In addition to the above reasons, the relatively lenient penalties and sentences for those found guilty, as well as the difficulty in convicting the alleged poachers contributes to the decrease in numbers of rhinoceros populations.

A number of measures have been instituted to protect rhinoceroses from poaching. Recently, an experimental dehorning program has been attempted by Zimbabwe to protect their decreasing rhinoceros populations (Leader-Williams, 1989).

It would be valuable to have an effective and sensitive genetic method of distinguishing one species from another, when trace amounts of blood or tissue is found on the clothing or weapons of suspected poachers. This could result in the more effective prosecution of poachers. It can be extremely difficult to isolate DNA of sufficient quality and quantity from the remains of poached animals for genetic analysis. The DNA extracted from dead animals is often so badly degraded that it cannot be used in the standard methods of identification (mtDNA restriction analyses and DNA fingerprinting).

The use of PCR was explored as a means to identify the species of origin from miniscule amounts of blood or tissue found on clothing, shoes or the implements used by the alleged poachers.

Three white rhinoceroses were killed and dehorned in the Umfolozi game reserve in October 1992. Suspects were found in possession of two axes, one bloodied, and blood stains on their shoes and shoelaces (although these may have been washed before being confiscated by the police). These samples were sent to our laboratory for analysis.

DNA was extracted from the samples, amplified using PCR, sequenced and the sequences were then compared to reference DNA sequences.

5.2 DNA AMPLIFICATION

5.2.1 DNA EXTRACTION

DNA was extracted from a number of samples (Table 5.1) under sterile conditions in a laminar flow hood. This was done in a separate laboratory from that used for rhinoceros DNA analysis. The DNA was extracted using a standard extraction protocol, but which was modified for the extraction of trace amounts of DNA (Paabo *et al.*, 1988). After the proteinase-K digestion and phenol treatment, the ethanol precipitation step was omitted. The aqueous, DNA-containing phase was concentrated through a Millipore filter unit (which retained molecules larger than 30 000 daltons). A blank extraction was also done to check for DNA contamination in the buffers. An aliquot from this DNA concentrate was used for amplification.

5.2.2 PCR

A 300 bp fragment of mtDNA of the cytochrome-b gene was amplified. The primers used for amplification were the general mammalian primers (Table 4.1) which have also been used to amplify rhinoceros DNA (Kocher *et al.*, 1989).

TABLE 5.1

The samples used for DNA extraction and amplification.

1.	axe, a
2.	axe, b
3.	shoes
4.	shoelaces
5.	stones, a
6.	stones, b
7.	debris from bag

The conditions of amplification have been discussed in chapter 4. No product was seen after the first 30 amplification cycles, but reamplification for a further 20 cycles yielded positive results. The axe head sample was the only one which resulted in an amplification product .

5.2.3 DIRECT SEQUENCING OF PCR PRODUCTS

The amplification products of six reactions were pooled and purified using the QIAEX method. The purified DNA was sequenced using the Sequenase-sequencing kit.

An internal primer (primer C), which annealed to both rhinoceros and bovine DNA, was designed to assist in sequencing part of the 300 bp fragment of cytochrome-b region of mtDNA (Table 4.1)

5.3 CLONING OF THE PCR FRAGMENT

The amplification product was cloned because an ambiguous sequence was produced by the direct sequencing method. A possible reason for this was that a contaminating fragment of DNA may have been amplified along with the cytochrome-b fragment. The 300 bp PCR fragment was purified and cloned into pUC 18.

The "Sma I-cloning" method was used (Lui & Schwarts, 1992). This method allows the cloning of blunt-ended fragments that lack Sma I restriction enzyme sites. The PCR fragment is ligated to the Sma I site of the multiple cloning site of the vector and vector self-ligation is prevented by the addition of Sma I into the ligation mix.

5.3.1 PREPARATION OF THE PCR PRODUCT AND VECTOR

The PCR fragment was run on a 2% agarose gel, the band excised and purified using the QIAEX gel purification method. The purified fragment was "polished" by blunt-ending it with the Klenow fragment of DNA polymerase. The vector, pUC 18, was also digested with Sma I.

5.3.2 LIGATION AND TRANSFORMATION

The ligation reaction was prepared in a volume of 10 to 20 ul, containing Sma I-digested vector, the blunt-ended insert in a molar ratio of 1:5 and 2 to 3 units of Sma I. The ligation reaction was incubated overnight at room temperature.

Competent *E. coli* (DH 5-alpha) cells were transformed with the ligation mix, and screened for positive

transformants. The blue/white X-gal method of selection was used (Maniatis *et al.*, 1989).

5.3.3 SEQUENCING OF THE CLONED PRODUCTS

The positive clones were used to inoculate 50 ml of LB media (containing 30ug/ml of ampicillin) and grown overnight. The cells were harvested and prepared for sequencing using the method of Faqua *et al* (1990). The purified amplification products were sequenced with the Sequenase sequencing kit, using the same primers as those used for the amplification.

5.4 RESULTS AND DISCUSSION

The suspects claimed that the blood on their clothing and the axes were from a calf that they had recently killed. The cytochrome-b region of mtDNA was sequenced since these primers were available in our laboratory. This region of mtDNA is relatively conserved across species with 88.0% sequence homology between bovine and white rhinoceros DNA. Although there was a large degree of homology between the two species, a 12% difference, corresponding to 25 individual base differences in the

FIGURE 5.1

Sequencing gel of the axe DNA directly after purification on a Millipore filter unit. The DNA was sequenced with the internal primer (primer C)

300 bp DNA sequence, was sufficient to differentiate between these two species with complete certainty.

Using the amplification primers as sequencing primers, the direct sequencing of the 300 bp amplified product yielded ambiguous results (figure 5.1). There often appeared to be a band in more than one lane on the same gel. Although clear motifs of the cytochrome-b region could be seen, there appeared to be two different populations of molecules of DNA present in the sequencing reactions, one of which was not cytochrome-b DNA molecule.

This could be the result of amplification after the non-specific annealing of the cytochrome-b primers to a fragment of DNA other than mtDNA. An internal sequencing primer was designed by comparing the cytochrome-b sequences of bovine and white rhinoceros (Table 4.1). It was thought that sequencing the amplification product with this internal primer would result in an unambiguous sequence. However, the internal primer also yielded unclear sequencing results, and it was not possible to resolve the source of the unknown DNA.

It was therefore decided to clone the amplification product before sequencing. This was done by the blunt-ended ligation of the amplified fragment into pUC-18

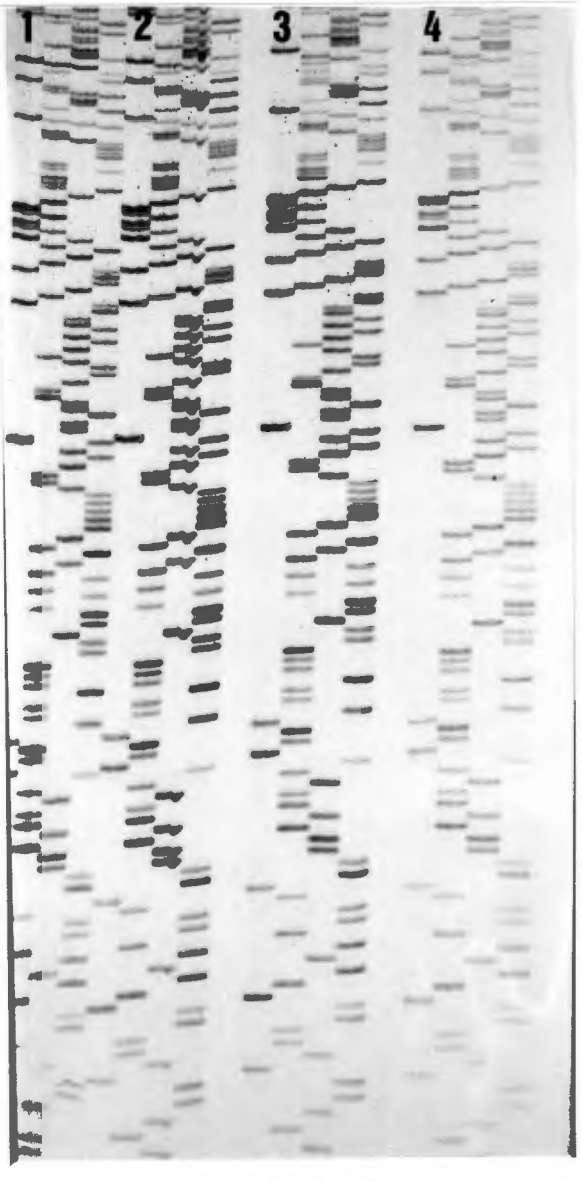

FIGURE 5.1

Sequencing gel of the axe DNA directly after purification on a Millipore filter unit. The DNA was sequenced with the internal primer (primer C)

```
                                                                                                  60
bovine      TCACATCAAA CCCGAGTGAT ACTTCTTATT TGCATACGCA ATCTTACGAT CAATCCCCAA
axe8        .......... .......... .......... .......... .......... ..........
axe12       .......... .......... .......... .......... .......C.. ..........
axe13       .......... .......... .......... .......... .......... ..........
axe19       .......... .......... .......... .......... .......... ..........
axe24       .......... .......... .......... .......... .......... ..........
axe25       .......... .......... .......... .......C.. .......... ..........
axe26       .......... .......... .......... .......... .......... ..........
axe27       .......... .......... .......... .......... .......... ..........
Wrhino      ...T...... .....A..G. .....C.... C.....T... .......... ....T.....

                                                                                                 120
bovine      CAAACTAGGA GGAGTACTAG CCCTAGCCTT CTCTATCCTA ATTCTTGCTC TAATCCCCCT
axe8        .......... .......... .......... .C........ .......... ..........
axe12       .......... .......... .......... .......... .......... ..........
axe13       .......... .......... .......... .......... .......... ..........
axe19       .......... .......... .......... .......... .......... ..........
axe24       .......... .......... .......... .......... .......... ..........
axe25       .......... .......... .......... .......... .......... ..........
axe26       .......... .......... .......... .......... .......... ..........
axe27       .......... .......... .......... .......... .......... ..........
Wrhino      .........G .....TT... ......T.C. T..C...... ..C...AT.A .C..A.....

                                                                                                 180
bovine      ACTACACACC TCCAAACAAC GAAGCATAAT ATTCCGACCA CTCAGCCAAT GCCTATTCTG
axe8        .......... .......... .......... .......... .......... ..........
axe12       .......... .......... .......... .......... .......... ..........
axe13       .......... .......... .......... .......... .......... ..........
axe19       .......... .......... .......... .......... .......... ..........
axe24       .......... .......... .......... .......... .......... ..........
axe25       .......... .......... .......... .......... .......... ..........
axe26       .......... .......... .......... .......... .......... ..........
axe27       .......... .......... .......... .......... .......... ..........
Wrhino      .........A .......... ....T..... ...T...... .......... ..T.......


bovine      AGCCCTAGTA GCA
axe8        .......... ...
axe12       .......... ...
axe13       .......... ...
axe19       .......... ...
axe24       .......... ...
axe25       .......... ...
axe26       .......... ...
axe27       .......... ...
Wrhino      ..T....... ...
```

FIGURE 5.3

Sequence alignment of the eight axe DNA clones with the bovine and the white rhinoceros reference sequences from Genbank.

vector. Eight positive clones were identified and sequenced using primer A of the amplification primers (figure 5.2). The sequences of the cloned products were compared to the reference cytochrome-b sequence from Genbank and to the white rhinoceros sequence done in our laboratory (chapter 4). The amplified DNA had 100% sequence homology with the bovine DNA, after corrections for Taq polymerase errors in some clones were accounted for (figure 5.3). This positively identified the DNA as being of bovine origin.

Therefore, it has been demonstrated that PCR and sequencing has a practical application, in the identification of the species of origin in trace quantities of blood or tissues. To my knowledge, this is the first forensic use of DNA sequences in a veterinary/legal context. With the rapid decline of rhinoceros populations, largely as a consequence of poaching, this forensic application of PCR is particularly useful. This method is the ideal means to identify the species of origin from trace amounts of biological material and thus, assist in the prosecution or defence of suspected poachers. This may serve as a deterrent to poachers, which may contribute to the stabilization of rhinoceros populations.

SECTION B

VARIATION IN MITOCHONDRIAL DNA MAPS OF FOUR SUBSPECIES OF *DICEROS BICORNIS* (BLACK RHINOCEROS)

CHAPTER 6

THE USE OF MT DNA TO DETERMINE THE GENETIC VARIATION BETWEEN FOUR SUBSPECIES OF BLACK RHINOCEROS

6.1 INTRODUCTION

The mammalian family Rhinocerotidae is one of the three families that comprise the order Perissodactyla. There are five extant species of rhinoceros (Table 6.1), all of which are listed as endangered in the 1986 IUCN Red list of threatened animals. *Diceros bicornis* (figure 6.1) and *Ceratotherium simum* (figure 6.2) are restricted to the African continent where they were once widely distributed in the sub-Saharan region, but with the increase in settlement of Africa from the seventeenth century both species have suffered a severe decline in numbers and in the extent of their range (Smitherns 1983).

TABLE 6.1

The five extant species of the family Rhinoceroditae

SUBSPECIES	SUBSPECIES	ESTIMATED POPULATION SIZE
Sumatran	3	594 -757
Dicerorhinus sumtrensis		
Great Indian	1	1709
Rhinoceros unicornis		
Javan	1	60 -70
Rhinoceros sondaicus		
Black rhinoceros	7	3392
Diceros bicornis		
White rhinoceros	2	4773
Ceratotherium simum		

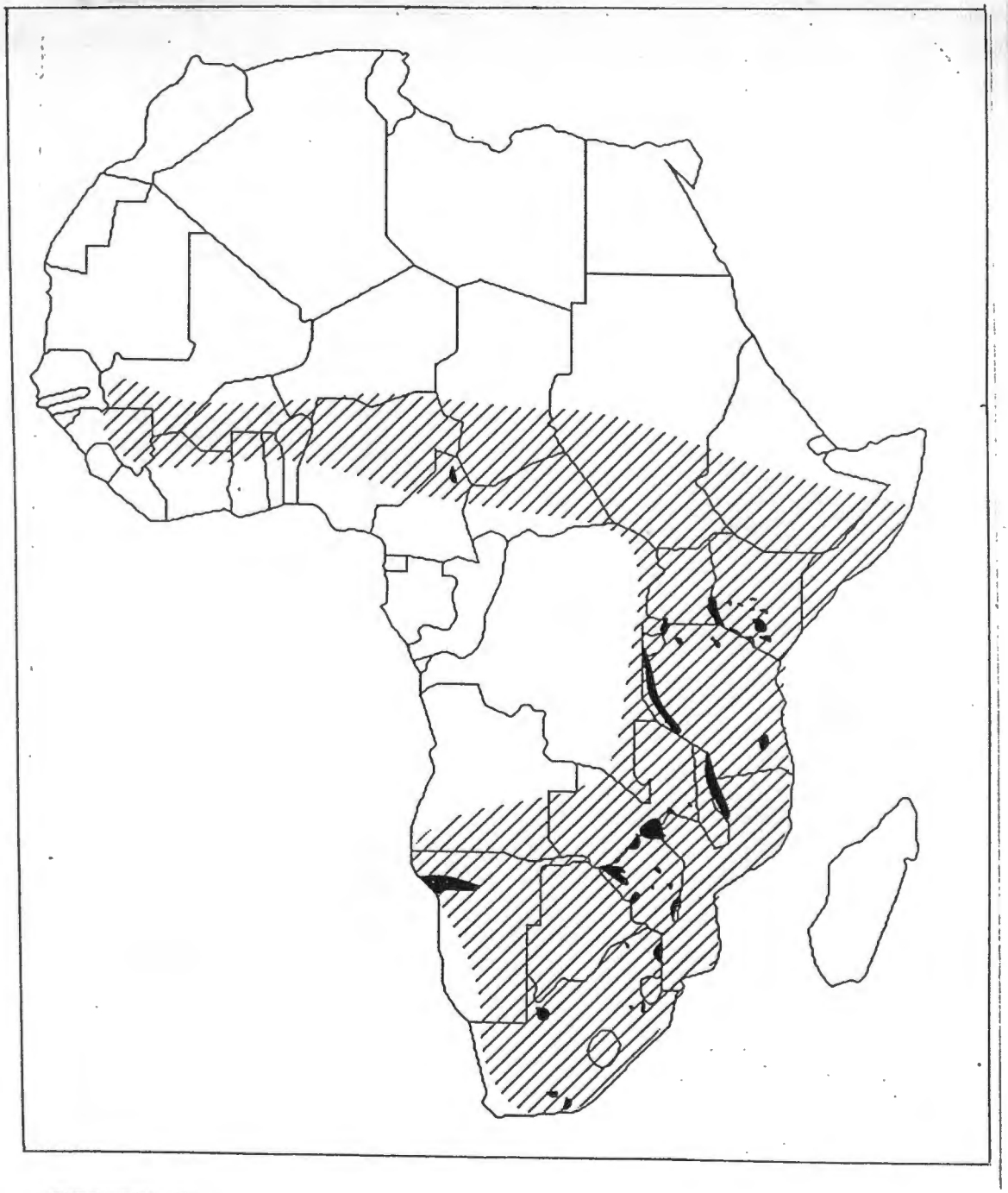

FIGURE 6.1

The distribution of *Diceros bicornis*. The diagonals represent its range in the early 1900s and the black area its current range (Ryder, 1992).

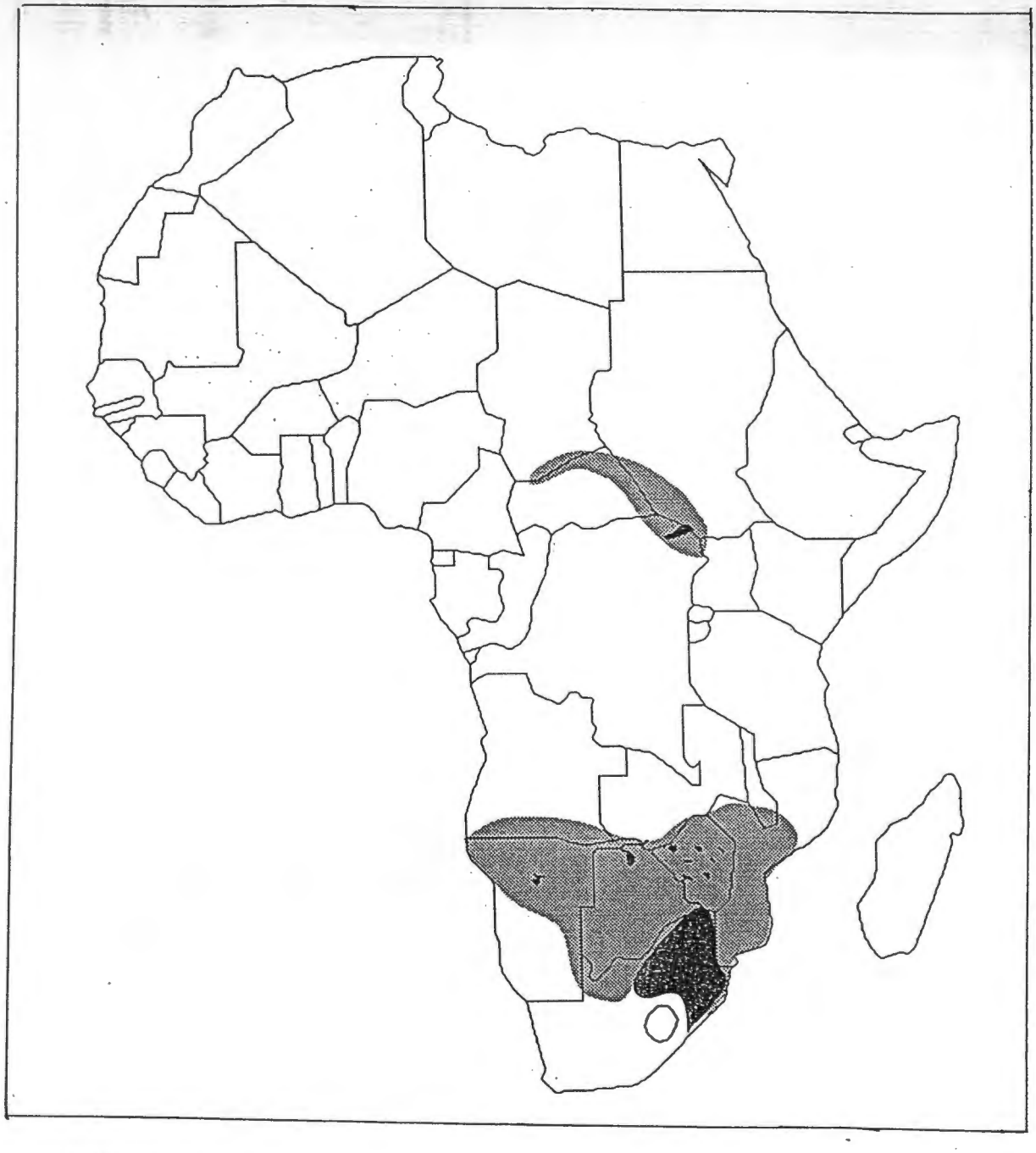

FIGURE 6.2

The distribution of *Ceratotherium simum*. The grey shaded area represents its range in the early 1900s and the black area its current range (Ryder, 1992).

Black rhinoceros numbers have declined more rapidly over the past 20 years than those for any other large African mammal. Numbers have fallen from 65 000 in 1970 to less than 4 000 in 1987 which represents a decrease of approximately 90% in the past decade (du Toit *et al.*, 1987).

This decrease is largely as a result of poaching for rhinoceros horn. In the early eighties, 50% of rhinoceros horn on the market was destined for North Yemen, where it is used for ornamental dagger handles, while the remaining 50% went to Eastern Asia for the production of traditional medicines.

Solutions to the decline in black rhinoceros numbers are currently being sought. They include the control of trading of rhinoceros horn, protection of black rhinoceros in game reserves by armed patrols, stringent anti-poaching laws and penalties and rhinoceros dehorning experiments. The dehorning experiment conducted in Damaraland, Namibia, is an example of one of the more extreme measures used in an attempt to thwart poachers (Leader-Williams 1989).

An additional problem that faces black rhinoceros is that of increasingly fragmentation into and small isolated populations. The animals' survival is vulnerable to genetic drift and demographic perturbations, even if protected from poachers (Gilpin & Soule, 1986 and du Toit, *et al.*, 1987).

Small populations lose genetic diversity rapidly (figure 6.3). Genetic diversity is important at the population level to allow adaptation to a continually changing environment and is important at the individual level to maintain the "vigour" of the animal.

However, this is not just a theoretical problem. In the Addo Elephant National Park in the Eastern Cape Province, there is a small population of 17 *D. b. michaeli* that originated from four Kenyan individuals. Both conservationists and national park managers are concerned about the "genetic fitness" of such a population.

Although many such small isolated populations exist throughout Southern Africa, current policy is to maintain the integrity of the different populations or "ecotypes" as separate populations. These populations may represent different subspecies of black rhinoceros, each of which is assumed to possess a set of genetic traits that constitute specific ecological adaptations.

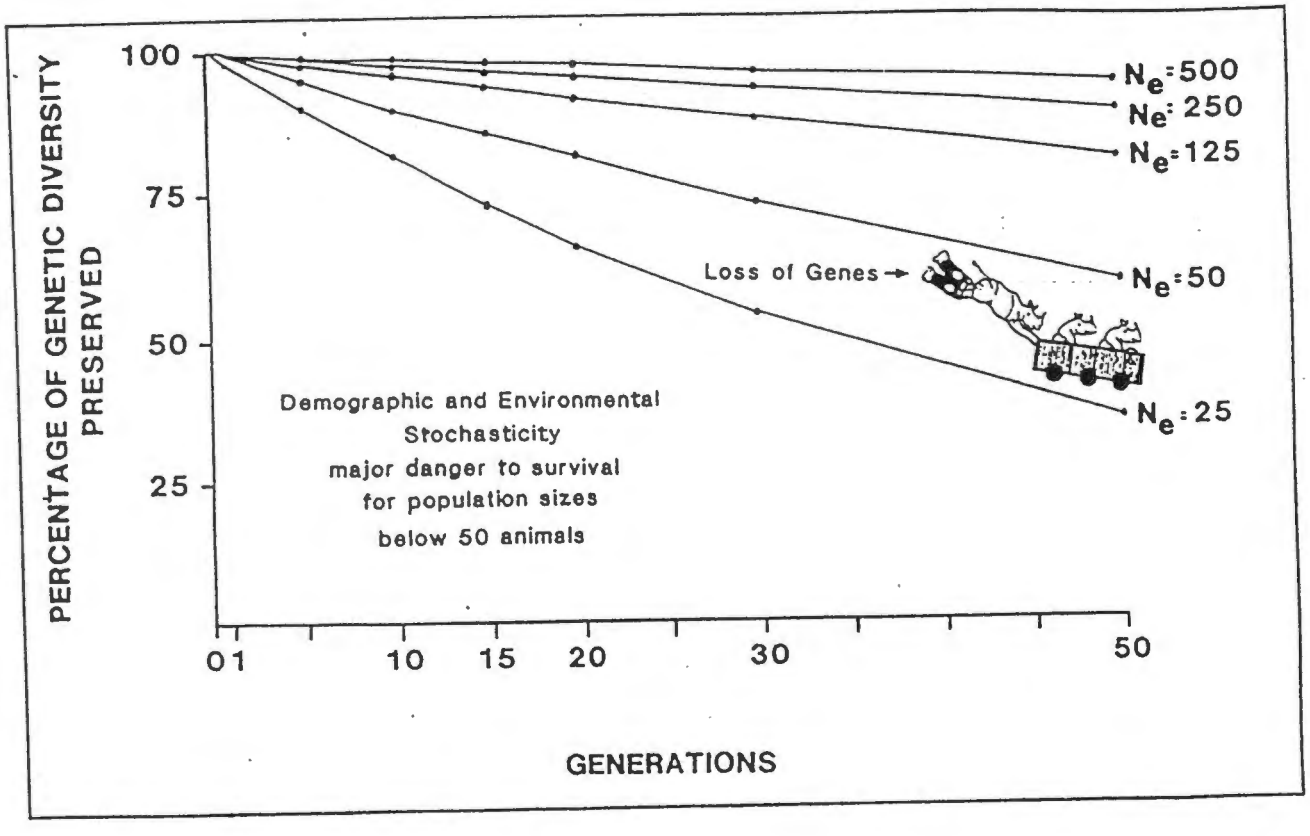

FIGURE 6.3

Graphical representation of the lost of genetic diversity in small populations. N_e is the effective population size. (Reproduced from du Toit, 1987)

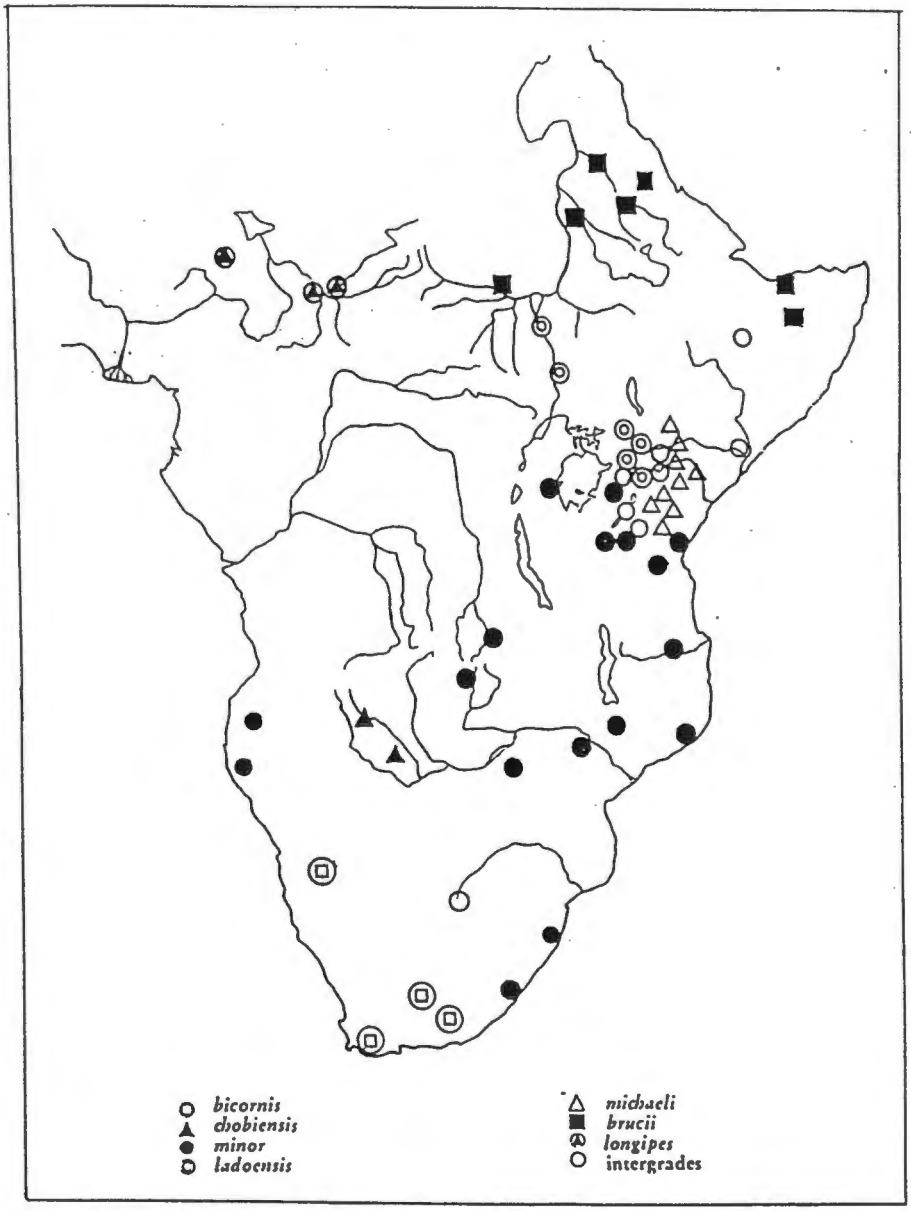

FIGURE 6.4

The distribution of *Diceros bicornis* subspecies based on morphological characteristics. (Reproduced from Groves, 1967)

To stabilize black rhinoceros numbers and to maintain the existing genetic variability, it is necessary to establish the "evolutionary units" within the species. Traditionally, subspecies status has been conferred on a group of animals using morphological characteristics. In the case of black rhinoceros, this consisted of minor cranial and pelage differences (Groves, 1967). These characteristics were not subjected to detailed statistical analysis. Consequently, subspecies distinctions are somewhat inconclusive, especially among neigbouring subspecies with continuous distributions (figure 6.4) .

Seven subspecies were identified (Table 6.2) on the basis of skull size (Groves 1967). However, in recent years it has been agreed (African Rhino Workshop in Cincinnati, 1986) that four basic ecological groups or conservation units should be recognized for practical management purposes and be kept separate from one another. These units are :

(i) the southwestern populations in Namibia

(ii) the southern-central populations extending from Natal through Zimbabwe and Zambia into southern Tanzania

(iii) the eastern populations in Kenya and northern Tanzania

(iv) the north-western populations extending from the horn of Africa to the Central African Republic and Cameroon.

TABLE 6.2

The seven subspecies of black rhinoceros identified on the basis of morphological characteristics (Groves, 1967).

SUBSPECIES	DISTRIBUTION
D.b.minor	Kenya to South Africa
D.b.michaeli	Kenya and Tanzania
D.b.bicornis	South Africa and Namibia
D.b.chobiensis	Angola
D.b.longipes	Central African Republic
D.b.ladoensis	Northern Kenya and Sudan
D.b.brucii	Ethiopia and Somalia

In order to clarify the black rhinoceros subspecies status, mtDNA of a number of different populations have been compared. Because this study required a large sample set of each different population, our original source of mtDNA from which the restriction enzyme maps were produced, heart tissue could not utilized. In conjunction with National Parks Board managers it was decided that whenever black rhinoceros were immobilize for translocation or veterinary purposes, small skin biopsies (usually a nick taken from the ear) would be taken from the animals and sent to our laboratory. Cell lines were established from the ear nicks, total DNA was extracted and probed with purified radiolabeled mtDNA.

6.2 ESTABLISHMENT OF PRIMARY CELL CULTURES

When the ear-nicks reached the laboratory, they were scrubbed clean of debris using an anti-bacterial (Hibitane) cleanser followed by a saline rinse. All subsequent manipulation were performed in a sterile laminar flow hood.

A small sliver of skin was cut from the ear-nick with sterile blades into a small petri-dish. The skin was then cut into very small pieces and a drop of enriched tissue culture medium [DMEM tissue culture medium containing 10% fetal calf serum (PHIS), penicillin,

streptomycin and neomycin (PSN)] was added and then aliquoted into a number of petri-dishes. A sterile coverslip was layered onto each drop of the mixture and 2.5 ml of enriched tissue culture medium was added. These primary cultures were incubated in a Hotpac incubator with a humidifier at 10% CO_2 at 37^0C. Twice a week the growth of the cells was checked and the media was changed.

When the cells were frozen for long term storage, two large flasks of confluent cells were trypsinized. Then 10 ml of media containing 20% PHIS was added to the trypsinized cells and this was centrifuged at 1 000 rpm at 20^0C for 10 minutes in a Sigma 3E1 centrifuge. The cellular pellets were washed, recentrifuged and resuspended in media containing 10% PHIS + 10% glycerol, which was dispensed in 1 ml tubes (Nunc) and frozen in a biological freezer. The frozen cells were stored in liquid nitrogen.

The frozen cells were checked to see whether they had been left undamaged by the freezing process. An aliquot of frozen cells was quickly thawed in a 40^0C waterbath and diluted into a flask of equilibrated media + 10% PHIS. The medium had been allowed to equilibrate with the CO_2 in the incubator for 30 minutes before the cells were added. The cells were incubated for two days after which they were checked for growth and for the presence of micoorganisms such as mycoplasmas.

All the long term rhino cell lines were transformed by using a SV40 recombinant virus (Van Doren & Gluzman, 1984). The transformed cell lines grew better and yielded more DNA than the untransformed cell lines. For transformation of a cell line, 1ul of the virus was diluted into 15ml media containing 10% PHIS and PSN. The cells were incubated in this virus-containing medium for 3 to 5 days, after which the cells were maintained in the standard manner.

6.3 THE MYCOPLASMA TEST

All cell lines were tested for mycoplasmas. The cells were grown on a coverslip for 5 days and were then stained for mycoplasma using Hoechst 33258. All the rhinoceros cell lines used in this study were mycoplasma-negative.

6.4 EXTRACTION OF DNA FROM CELL CULTURES

Confluent cells were washed twice with sterile saline. Then 5ml of the digestion buffer (0.9ml TE buffer, 0.1ml, 10% SDS solution and 100ug of proteinase-K per ml) was added. The cells were incubated at 37^0C for 2 hours. After the incubation, sodium perchlorate was added to 0.1M. Then an equal volume equal volume of phenol:chloroform (25:24) was added to the lysate. The

solution was mixed well and then centrifuged for 10 minutes in a microfuge (12 000 g). The aqueous phase was transferred into a clean microfuge tube and the phenol:chloroform extraction was repeated. This was followed by two chloroform:octanol extractions. Then two volumes of absolute (100) ethanol was added to the aqueous phase. The DNA precipitated immediately and was spooled out of solution with a glass rod. The DNA was resuspended in 200-500ul of TE and stored at -20^{0}C.

SOUTHERN BLOTTING OF DNA FROM TISSUE CULTURE WITH PURIFIED mtDNA

In order to visualize the mtDNA fragments in a total DNA preparation, an aliquot of the latter was electrophoresed on agarose gels, Southern blotted (Southern 1975) onto nylon membranes and hybridized to a labeled mtDNA probe.

6.5 SOUTHERN BLOTTING

Total DNA (1 - 5ug) was digested with 10 units of an appropriate restriction enzyme and separated on agarose gels (1-1.7%) at 35V for 16-20 hours. After electrophoresis, the DNA was denatured, neutralized (buffers in appendix) and transferred onto a nylon membrane (Hybond-N^{+}) by capillary blotting. The

alkaline method of transfer was employed. Before the DNA in the gel was transferred, the gel was soaked for 10 minutes in a 0.4M NaOH solution. The NaOH-soaked gel was then transferred onto the nylon, Hybond-N^+ membrane. After 12 to 16 hours of transfer, the membrane was gently washed in 4X SSC and was either stored until needed or used immediately for hybridization.

6.6 LABELING THE mtDNA PROBE

The mtDNA was purified from heart tissue of *Diceros bicornis minor* and labeled using the random-priming method (Feinberg and Vogelstein 1984).

The reaction was performed as recommended by the suppliers of the random-primed DNA labeling kit (Boehringer Mannheim). Template DNA (25 to 50 ng) was linearized by digestion with Bcl I and then heat-denatured. The following reagents were added sequentially to a microfuge tube on ice: linearized, denatured mtDNA; the dATP, dGTP, dTTP mixture and 50uCi [alpha ^{32}-P]dCTP and 2 units of Klenow enzyme. The volume of the reaction was made up to 20ul with sterile, distilled water. The reaction was incubated for 90 minutes at 37^0C, after which it was terminated by the addition of 2ul of a 0.2M EDTA (pH8) solution.

The labeled mtDNA was separated from the unincorporated labeled dCTP using a Sephadex G-50 column and the fractions containing the radioactive mtDNA were pooled. Probes labeled in this manner, routinely had a specific activity from 1 x 10^7 to 1 x 10^9 dpm /ug of DNA.

6.7 HYBRIDIZATION AND WASHING CONDITIONS OF THE HYBRIDIZATION MEMBRANE

The transfer membranes were prehybridized for at least 2 hours at 65^0C in a hybridization oven (Hybaid), in hybridization buffer containing 6x SSC, 0.1% SDS, 0.06% sodium pyrophosphate and 0.25% BLOTTO (low fat dried milk powder). After prehybridization, the heated-denatured mtDNA probe was added to the membrane and hybridized in 15ml of hybridization buffer at 65^0C for 16 to 18 hours.

The membranes were washed the next day in solutions of decreasing ionic strength at 65^0C. The membrane was washed in the following solutions: (i) three times in 6x SSC and 0.1% SDS, (ii) once in 3x SSC and 0.1% SDS and (iii) once in 0.1X SSC and 0.1% SDS.

After the last post-hybridization wash, the membrane was blotted dry between two sheets of filter paper and then sealed in a plastic bag and exposed to X-ray film (Hyperfilm from Amersham, UK) for autoradiography.

6.8 RESULTS AND DISCUSSION

The mtDNA restriction enzyme profiles are summarized in Table 6.3. Between one and seven restriction fragments were obtained with the 16 enzymes used. The most frequent fragment patterns observed were assigned the letter A, and subsequent patterns used B. For 13 of the 16 enzymes used, no mtDNA variability among the four subspecies was observed. Only three enzymes gave patterns which differed among the four subspecies. They were Bcl I, Dra I and Stu I.

Table 6.4 lists the source of the black rhinoceroses sampled. Of the 26 *D. b. minor* and five *D. b. bicornis* screened, all were monomorphic for their respective mtDNA genotypes with the 16 enzymes tested. Although only one *D. b. chobiensis* and one *D. b. michaeli* were tested, Ashley screened 11 *D. b. michaeli* individuals in a mtDNA fragment size analysis using a similar panel of restriction enzymes (Ashley, Melnick & Western, 1990). The documented fragment patterns of these 11 *D. b.michaeli* individuals tested were also found to be monomorphic and are consistent with our mapped restriction sites for *D. b. michaeli* (figure 6.5 for maps).

TABLE 6.3

A summary of the mt DNA restriction enzyme profiles. The most frequent fragment patterns observed were assigned the letter "A" and subsequent patterns, the letter "B".

ENZYME	NUMBER OF SITES	D. B. CHOBIENSIS N = 1	D. B. MINOR N = 26	D. B. MICHAELI N = 1	D. B. BICORNIS N = 5
Bcl I	6	A	A	B	B
Dra I	5	A	A	A	B
Stu I	6	A	A	B	A
Bam HI	2	A	A	A	A
Bgl II	1	A	A	A	A
Eco RI	2	A	A	A	A
Eco RV	2	A	A	A	A
Hind III	4	A	A	A	A
Hpa I	2	A	A	A	A
Nco I	1	A	A	A	A
Pvu II	4	A	A	A	A
Sac I	1	A	A	A	A
Sac II	3	A	A	A	A
Sal I	2	A	A	A	A
Sca I	7	A	A	A	A
Xba I	4	A	A	A	A

TABLE 6.4

Summary of the origin of the black rhinoceros subspecies sampled. "Number": denotes the number individuals sampled from each subspecies.

This individual was found four kilometers from the unfenced Angolan/Namibian border, east of the Kavango river barrier.

* The Addo population of *D. b. michaeli* was derived from a founder population of two cows and two bulls translocated from the Kiboko region in south-east Kenya in 1961/2.

SUBSPECIES	NUMBER	LOCATION
D.b.chobiensis	1	#Caprivi,Namibia
D.b.michaeli	1	*Addo, E. Cape,RSA
D.b.bicornis	5	Etosha, Namibia
D.b.minor	6	Mkuzi, Natal, RSA
	17	Umfolozi,Natal, RSA
	3	Eastern Cape, RSA

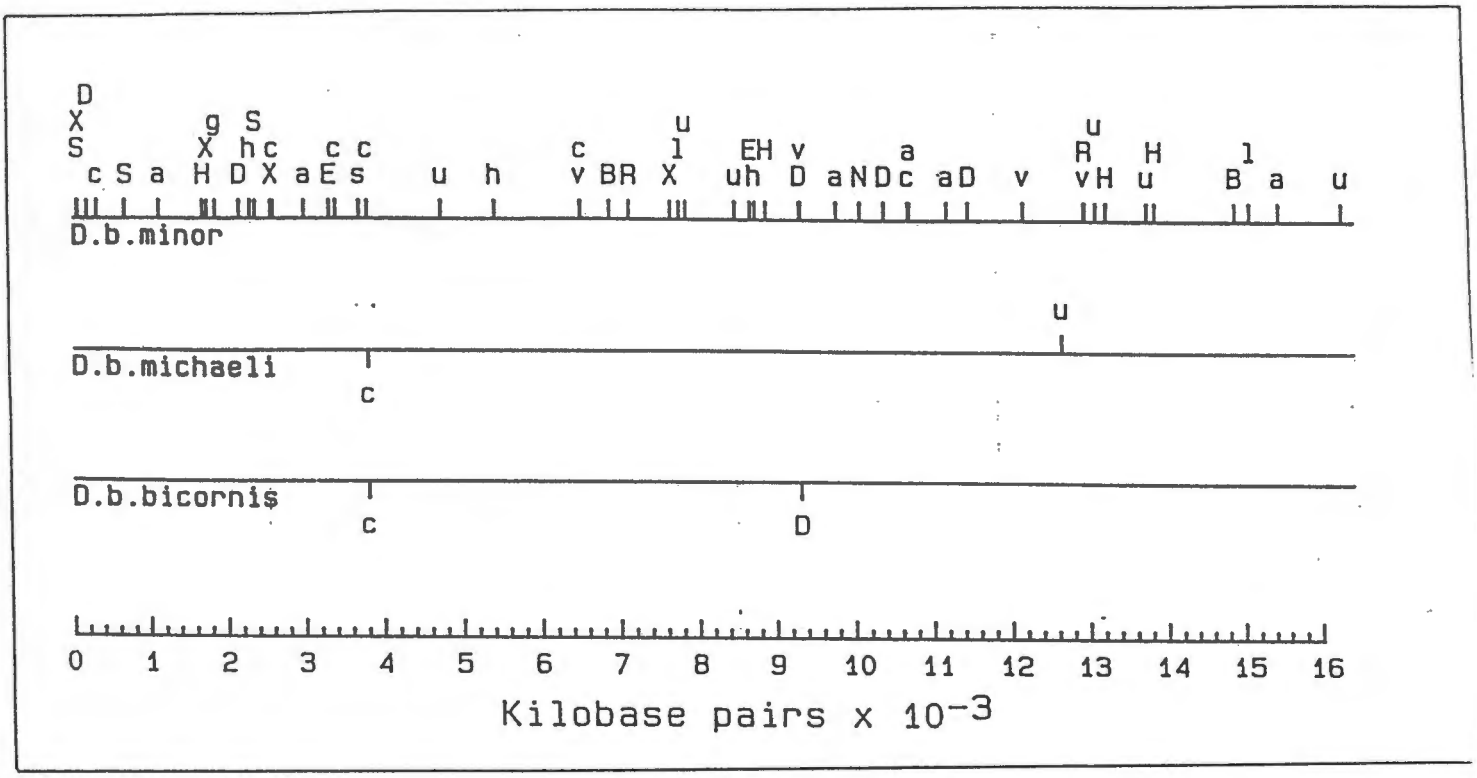

FIGURE 6.5

Restriction enzyme maps of the four subspecies of black rhinoceros. The maps have been aligned on the invariant Sac II position at 676 in the *Bos taurus* sequences. The maps for *D. b. minor* and *D. b. chobiensis* were identical. In *D. b. michaeli* and *D. b. bicornis* only the site differences were shown. Site gains were shown above the line and site losses below the line.

These monomorphic patterns therefore, identify diagnostic mtDNA haplotypes for *D. b. michaeli*, *D. b. minor* and *D. b. bicornis* (figure 6.6). *D. b. chobiensis* and *D. b. minor* were found to have identical restriction fragment patterns. These diagnostic fragment patterns are shown in figure 6.7 and the site differences giving rise to the variant fragment patterns in the three subspecies are shown in figure 6.5.

Two site changes differentiate *D. b. minor* from *D. b. michaeli*, two differentiate *D. b. minor* from *D. b. bicornis* and two differentiate *D. b. michaeli* from *D. b. bicornis*. In each case this corresponds to an estimated sequence divergence of 0.4% between any two of the subspecies (Nei & Li, 1979). Although the variation among the subspecies *minor*, *bicornis* and *michaeli* was low, the differences found appear to be monomorphic for each of these subspecies. Thus conclusions can be made with confidence for *D. b. minor* where 26 individuals have been studied from a number of localities and with moderate confidence for *D. b. bicornis* where 5 individuals were studied. It can also be made with moderate confidence for *D. b. michaeli*, although only one individual was studied here, our mtDNA map is consistent with the fragment size analysis of Ashley *et al* (1991) who used a similar panel of restriction enzymes and found no variation among the 11 individuals

FIGURE 6.6

Agarose gel electrophoresis of DNA from a number of individual of two subspecies of *D. bicornis* digested with Stu I. MtDNA fragments were visualized after Southern blotting, using a labeled rhinoceros probe. Lanes 1-4: *D. b. bicornis*, lane 5: marker lambda Hind III DNA and lanes 6-15: *D. b. minor*.

FIGURE 6.7 : Agarose gel electrophoresis demonstrating the diagnostic patterns that differentiate the four subspecies of black rhinoceros. Lanes 1, 5 and 9, *D. b. chobiensis*; lanes 2, 6 and 10, *D. b. minor*; lanes 3, 7 and 11, *D. b. michaeli* and lanes 4, 8, and 12, *D. b. bicornis*. DNA was digested with lanes 1 - 4, Stu I; lanes 5 - 8, Dra I, lanes 9 - 12, Bcl I. Lane 14 is lambda Hind III DNA.

of *D. b. michaeli* from Zimbabwe. It is likely that these different mtDNA morphs are fixed for each of these subspecies and therefore, would constitute diagnostic markers for each subspecies.

It is of interest that the *chobiensis* subspecies gives an identical pattern to *D. b. minor*. Although the distributional range of *D. b. chobiensis* is closer to that of *D. b. bicornis* , their bioclimatic habitat is more similar to that of *D. b. minor*. Both *D. b. minor* and *D. b. chobiensis* occur in savanna as opposed to the more arid, semi-desert habitat of *D. b. bicornis*. The similarity between the mtDNA of *D. b. chobiensis* and *D. b. minor* would appear to support a relatively recent link, probably through the lush mesic Zambezi/Chobe corridor.

There are only a very small number of individuals of the *D. b. chobiensis* subspecies remaining. If translocations are required on the basis of habitat loss or to protect them from poachers, it would seem logical to allow them to interbreed with *D. b. minor* populations, if the population of the translocated *D. b. chobiensis* is too low for their establishment as a separate breeding population.

The second result of practical value to emerge from this study is the small amount of mtDNA diversity among

the subspecies. The amount of diversity found is no more than that typically found among members of any large panmictic mammalian population (Avise *et al.*, 1986, Cann *et al., 1987*, Essop *et al.*, 1991 & Wayne *et al* 1991). On this basis alone it would be unlikely that interbreeding between these subspecies would result in any decrease in fitness or fecundity in the offspring (outbreeding depression).

On the other hand, it is necessary to confirm that there are no significant chromosomal differences among the subspecies, since this could result in decreased fecundity of offspring of such crosses (Laikre & Ryman, 1990; O'Brien *et al.*, 1985 and Packer *et al.*, 1990). This is important because a preliminary result (Ryder, personal communication) has indicated that there may be some chromosomal differences between *D. b. minor* and *D. b. michaeli*. Major chromosomal differences can be a cause of sterility or decreased fecundity of offspring of such crosses. If no chromosomal differences are confirmed, and the subspecies are to be managed as separate breeding entities, the justification will need to be on the basis of preserving some desirable feature of morphology or adaptive specialization in the subspecies.

These justifications will need to be rigorously defined, since keeping small populations of the subspecies separate, at least on a regional basis,

require greater management planning and could contribute to increasing the loss of genetic diversity, than if all the *D. bicornis* populations were allowed to interbreed.

SECTION C

FEASIBILITY OF THE USE OF DNA FINGERPRINTING AS A MEANS TO DETERMINE GENETIC DIVERSITY WITHIN AND BETWEEN FOUR SUBSPECIES OF BLACK RHINOCEROS

CHAPTER 7

INTRODUCTION TO DNA FINGERPRINTING

7.1 GENERAL INTRODUCTION

The detection of genetic polymorphisms is an important tool for identification and for the elucidation of relationships among individuals. Protein polymorphisms have been extensively used, but this (indirect detection) is limited by the amount of variability in the system. The direct analysis of DNA gives a direct measure of mutational events, and has revealed polymorphic loci in the genome (Cawood, 1989). The first of these hypervariable loci were identified in 1980 (Wyman & White, 1980) and subsequently, many other polymorphic regions have been found scattered throughout the human genome (Nakamura *et al.*, 1987).

7.2 JEFFREYS PROBE

In the early 1980s, a new type of DNA element was isolated from the human myoglobin gene that consisted of arrays of tandemly repeated short sequence units, normally G+C-rich (Jeffreys *et al.*, 1985a). A number of these arrays have been discovered and have become known as minisatellites, hypervarible regions (HVRs) or VNTRs (variable number tandem repeats).

Minisatellites display a high degree of allelic variation in the number of repeats in the array, which is detectable as length polymorphisms in a Southern blot. The degree of heterozygosity at HVRs is high and thus, HVRs have the potential for providing highly informative genetic markers.

Minisatellites exist as "families", the members of which are related by similarity in the nucleotide sequence of the repeat unit and in the motif of their repeated units, which allows cloned DNA segments containing these minisatellites to detect multiple loci simultaneously. The resultant pattern of hybridization fragments constitutes an individual-specific "DNA fingerprint" (Jeffreys, *et al.*, 1985b) and the fragments of the fingerprints are inherited in a Mendelian fashion (Jeffreys *et al.*, 1986).

7.3 FUNCTION OF MININSATELLITE DNA

Although minisatellite DNA has been found scattered throughout the genome, its function remains unknown. However, there are several lines of evidence which suggest that the core sequence of HVRs might serve as a recombination signal in human DNA (Jeffreys, 1987). The core sequence in Jeffreys probe is similar in length and G-content to Chi, the cross-over hotspot initiator sequence of *Escherichia coli*. Figure 7.1 illustrates the model for minisatellite generation promoted by chi-like core sequences. Because some HVRs do not contain the Chi-like sequence, alternative explanations, which include replication slippage, have been put forward for the role of minisatellite DNA (Jarman *et al* 1989, Dover, 1989 & Krowczynska *et al.*, 1990).

7.4 APPLICATIONS OF JEFFREYS PROBE

Most DNA fingerprinting applications have been based on individual- or clonal- specificity (Fey *et al.*, 1988). Fingerprinting has been used to determine paternity in humans (Helminen *et al.*, 1988 and Jeffreys *et al.*, 1988a), even when the putative father is unavailable for testing (Odelberg *et al.*, 1988). The individual-specificity of fingerprinting in humans lends itself to forensic applications (Gill *et al.*, 1985), although the

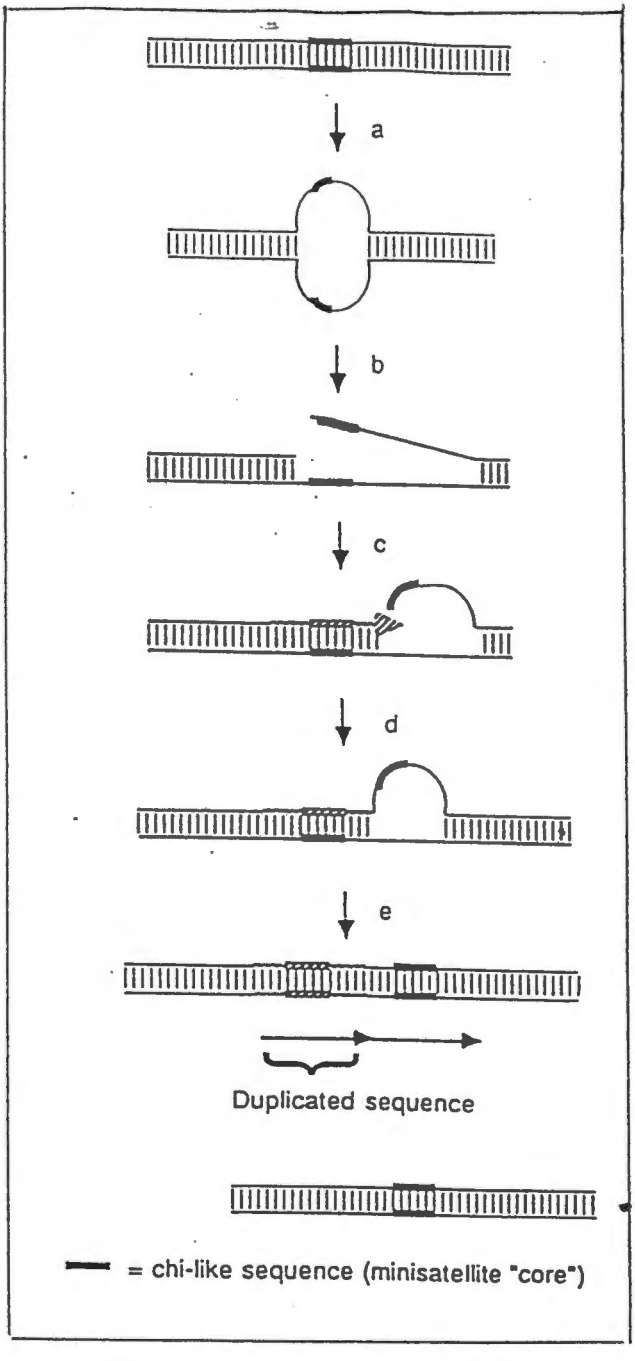

FIGURE 7.1

A model for minisatellite generation which is promoted by Chi-like core sequences. (a) Specific unwinding of the duplex by an enzyme that recognizes the core; (b) nicking to produce a single stranded projection; (c) DNA repair synthesis; (d) ligation to form a tandem duplication containing the core; (e) replication. (Reproduced from Jarman & Wells, 1989)

legal complexities have yet to be resolved (Neufeld *et al.*, 1990, Thompson *et al.*, 1990 and Roberts, 1992). Because DNA is stable for some time after death, fingerprints have been obtained from postmortem material (Bar *et al.*, 1988). The enzymatic amplification of human DNA by the polymerase chain reaction enables even DNA from a single cell to be analyzed by DNA fingerprinting (Jeffreys *et al.*, 1988b).

Although, DNA fingerprinting was initially developed to determine human parentage, the technique has been adapted for use in studies of many other animals (Jeffreys *et al.*, 1987, Jeffreys & Morton, 1987 and Parkin, 1989), particularly birds (Wetton *et al.*, 1987, Burke & Bruford, 1987, Burke *et al.*, 1989 and Meng *et al.*, 1990).

7.5 OTHER MULTILOCUS FINGERPRINTING PROBES

Several new developments have evolved in the DNA fingerprinting realm over the last few years, including the isolation of a number of other probes which are able to detect hypervariable regions in the genome. Alec Jeffreys and his co-workers have now developed a method to examine sequence variation of the minisatellite structure (Jeffreys *et al.*, 1990 and Jeffreys *et al.*, 1991).

Other multilocus probes isolated include: a 15 base pair repeat within the protein III gene of M13 bacteriophage (Vassart *et al.*, 1987), which is able to hybridize to a range of organisms (Ryskov *et al.*, 1988); a human clone pV47-2, which was isolated from a Carny 40 human chromosome-16 library after screening with M13 (Longmire *et al.*, 1990); a commercially available 22mer oligoprobe containing a SP1 binding site (Menzel *et al.*, 1990); a tandem repeat of 28 base pairs downstream of the human c-Ha-ras-1 oncogene (Washio *et al.*, 1989) and a number of synthetic oligonucleotide repeat sequences (Ali *et al.*, 1986, Schafer *et al.*, 1988a, Schafer *et al.*, 1988b and Vergnaud, 1989).

7.6 MODIFICATIONS OF FINGERPRINTING

Further refinements of DNA fingerprinting has been the development of single locus probes (SLP), micro-satellite fingerprinting and random amplified polymorphic DNA (RAPDs, pronounced "rapids").

7.6.1 SINGLE LOCUS PROBES

Several families of minisatellite DNA exit. Most contain a "core-like" sequence, but have different

flanking DNA between the repeats. Under the appropriate (low stringency) conditions, MULTILOCUS probes (MLP) (Jeffreys probe, M13, and the synthetic oligonucleotides) simultaneously detect alleles from many loci by hybridizing to the minisatellite "core" region. This results in a complex pattern of a large number of bands, which makes the interpretation and statistical analyses of the fingerprint extremely difficult.

A SLP is typically isolated from an individual clone containing the minisatellite core sequence as well as a unique flanking region of DNA (Wong *et al.*, 1987). This unique flanking region of DNA renders the SLP species specific for that locus. The SLP is hybridized at high stringency and detects alleles from one locus, resulting typically, in two fragments on the autoradiograph in a heterozygote. Species-specific SLPs have been isolated for humans (Wong *et al.*, 1987b) and for peacocks (Hanotte *et al.*, 1991).

SLPs offer several advantages over MLPs (Pemberton *et al.*, 1991). They are significantly more sensitive, technically easier to use and the interpretation of SLP profiles are simpler. The advantage of being able to obtain allele frequency estimates allows useful population genetic questions to be addressed. The data obtained is also amenable to database compilation and

allows for comparisons between experiments to be made more easily.

The major drawback of SLPs is their limited availability. Although hybridization occurs to closely related species, at present the SLPs isolated are predominantly species-specific (Kempenaers *et al.*, 1992). Consequently, a different SLP has to be isolated from a complex genomic library for every new species studied, which is a long and laborious exercise (Saito & Stark, 1986 and Armour *et al.*, 1990).

7.6.2 SIMPLE DI- AND TRI- NUCLEOTIDE REPEATS

Simple, tandemly, repeated di-, and tri- nucleotides sequences, called microsatellites, have been demonstrated to be polymorphic in length in a number of eukaryotic genomes. The frequency with which they occur (once every 50 000 to 60 000 bp), the high degree of polymorphism displayed, and their random distribution across the entire genome, make them potentially very useful as DNA markers. This method uses PCR to detect length polymorphisms (Rassman *et al.*, 1992). Primers which flank the polymorphic locus are used to amplify the repeat fragments from genomic DNA and the amplified product is resolved on a suitable polyacrylamide gel.

This technique has several advantages: One only needs small amounts of DNA which may be partially degraded. The resolving power of the gels is such that a single nucleotide difference can be detected. Because one analyses individual loci, it is possible to identify both alleles of the locus in a heterozygote. The major obstacle of the microsatellite approach is that it is necessary to find suitable simple sequence loci in order to design the PCR primers, but methods have been developed to isolate these loci (Rassman *et al.*, 1992). These polymorphic simple sequence loci have been shown to be conserved in cetacean species (Schlotterer *et al.*, 1991), as well as in numerous other closely related species (Moore *et al.*, 1991). Another method to isolate these loci is based on the amplification of micosatellites at the 3′ end of an ALU sequence (Charlieu *et al.*, 1992).

7.6.3 RAPD

RAPD is a PCR-based technique for assaying individual-specific variation (Williams *et al.*, 1991). It has been demonstrated that amplification with short (10 bp) primers are able to detect polymorphisms in the absence of specific nucleotide sequence information. These polymorphisms are generally inherited in a Mendelian fashion and can be used to construct genetic maps in a variety of species. The disadvantages associated with

the method are repeatability, expense and dominance (the inability to distinguish between the homozygote and the heterzygote at a locus) (Hendrik, 1992).

7.7 THE USE OF FINGERPRINTING IN THIS STUDY

In this study, DNA fingerprinting with M13 and a synthetic oligonucleotide (both MLPs) were attempted to investigate whether the method could be useful in quantitating the degree of relatedness among different populations of black rhinoceros. This information has potential to predict the risk of inbreeding in small populations of rhinoceros, so that the appropriate management measures can be made before this happens.

CHAPTER 8

THE USE OF FINGERPRINTING TO EXAMINE THE RELATEDNESS BETWEEN FOUR SUBSPECIES OF BLACK RHINICEROS

8.1 INTRODUCTION

The feasibility of DNA fingerprinting was explored to characterize the genetic variability within and between four Southern African subspecies of black rhinoceroses. The DNA used in this study were isolated from cell culture. Table 6.4 lists the source of the animals used.

In order to ascertain the genetic variability within and between four subspecies of black rhinoceros, M13 and a synthetic oligonucleotide were used as DNA fingerprinting probes. The synthetic oligonucleotide was synthesized and the tandem repeat of M13 that hybridizes to the hypervariable regions of the target genome, was cloned into pGem-3. The reason for this was to label the M13 tandem repeat to a higher specific activity.

8.2 SUBCLONING OF M13 INTO pGEM-3

8.2.1 DIGESTION AND PURIFICATION OF M13 AND pGEM DNA

The tandem repeat sequence of the protein III gene of M13 (figure 8.1) was cut from the bacteriophage and cloned into the pGEM-3 plasmid (figure 8.2). M13 DNA was digested with both Cla I and Sau 3A which resulted in appropriate sticky ends for cloning into the Bam HI and Acc I sites of pGEM-3.

Using the "shotgun" approach 5ug of M13 was digested with Cla I and Sau 3A, the digestion products were extracted with phenol:chloroform, phenol:octanol and then ethanol precipitated. This mix of purified M13 DNA fragments was cloned into pGEM-3.

A second cloning strategy was also employed. Only the specific protein III gene consensus sequence fragment was purified and cloned. M13 DNA (10ug) was digested with Cla I and Sau 3A and the digestion products were separated on a 5% polyacrylamide gel. The fragments were EtBr-stained and visualized under UV light. The 800 bp M13 fragment was excised from the gel and eluted from the gel as described in Maniatis (1982).

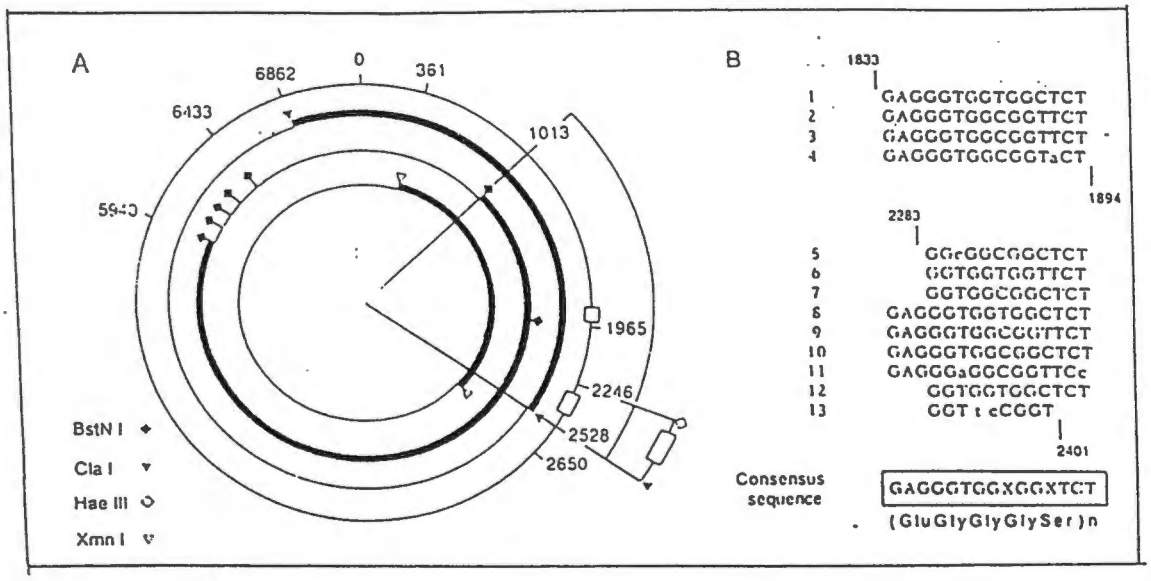

FIGURE 8.1

The map of M13 DNA. (a) M13 restriction enzyme map. The thick lines show restriction fragments that were preferentially recognized by the affinity-purified M13 sequences. The open boxes give the localization of the tandem repeats. (b) Alignment of the tandem repeats present in the protein III gene of M13. (Reproduced from Vassart *et al*, 1987)

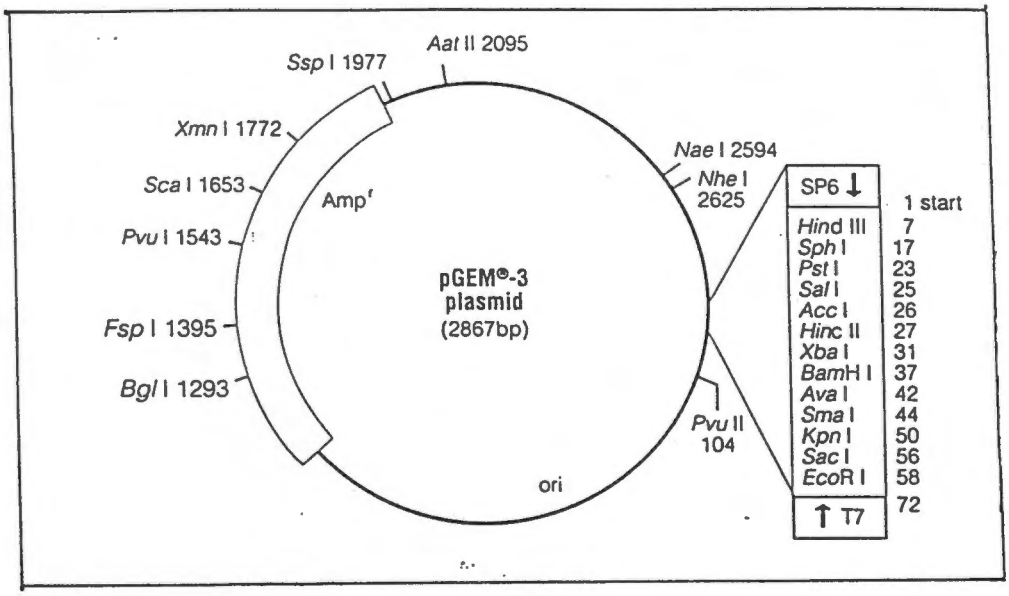

FIGURE 8.2

Restriction map of pGEM-3, showing the multiple cloning sites in the box.

The pGEM-3 DNA was extracted by the alkaline-lysis method (Maniatis) from an overnight culture, inoculated with a single colony of the vector. The purified DNA was digested with Bam HI and Acc I. The digestion products were extracted with phenol:chloroform, phenol:octanol and precipitated with ethanol.

8.2.2 LIGATION AND TRANSFORMATION

The appropriate molar ratio of the insert to vector was calculated using the j:i ratio to optimize the ligation reaction and to maximize the number of recombinants (Maniatis *et al*, 1982). The appropriate amounts the plasmid and the insert DNA were ethanol precipitated in the same microfuge tube. The pellet was dissolved in 18ul of distilled water. To ligate the insert to the plasmid, 2ul of ligation buffer and 2U of ligase were added. The ligation reaction was incubated overnight at 12^0C.

Compentent *E. coli* DK-1 cells were transformed with 1/10 of the ligation mix. The transformed cells were plated onto agar plates containing ampicillin and incubated overnight at 37^0C. The colonies were screened for positive clones containing the correct-sized insert fragment.

8.2.3 SCREENING THE RECOMBINANTS

Mini-prep overnight cultures (5ml) were inoculated with a single recombinant and the plasmid DNA was extracted by the alkaline lysis method. The plasmid with an insert were screened for the correct fragment of M13. The positive clones were then digested with a number of diagnostic enzymes to confirm for a correct sized insert. Positive clones were grown and the DNA extracted, purified and used as a fingerprinting probe.

8.3 LABELING OF M13 PROBE

The insert was released from the purified plasmid by a Hind III and Eco RI digestion. The fragments were electrophoresized on a agarose gel and the insert fragment was purified using the QIAEX method. The insert DNA was labeled to specific activity of 1 x 10^9 dpm\ug, using the random-priming method and fractionated on a Sephadex G-50 column previously described. The labeled M13 DNA was alkaline-denatured by the addition of NaOH to a final concentration of 0.2M. The denatured DNA was then hybridised to the fingerprint membrane.

8.4 LABELING OF OLIGONUCLEOTIDE PROBE

The oligonucleotide probe, (CAC_5), was synthesized by Genetic Designs, Inc. The 15-mer nucleotide was chemically synthesized with a free 5'-OH end. The oligonucleotide was labeled using the kinase labeling reaction.

The template DNA was diluted so that the concentration was 20 pmol/ul. The kinase reaction contained the following components: 1ul template DNA, 1ul of 10x polynucleotide kinase reaction buffer (0.5M Tris-HCl pH 7.6, 0.1M $MgCl_2$, 50mM DTT, 1mM spermidine, 1mM EDTA), 1ul ^{32}P-dATP (20 pmol of gamma-labeled ^{32}P), 2 units of T_4-polynucleotide kinase and sterile, distilled water to a final volume of 10ul. The labeling reaction was incubated at 37^0C for 60 minutes.

The unincorporated label was removed from the labeled probe by separation through a Sephadex G-10 column. Probes of a high specific activity (1 x 10^8 to 1 x 10^9dpm/ug) were obtained in this manner. The probe was heat-denatured before hybridization.

8.5 HYBRIDIZATION OF FINGERPRINTING PROBES

8.5.1 SOUTHERN BLOTTING OF DNA

Total DNA was extracted from black rhinoceros cell cultures, using the method described in chapter 7. The purified DNA was digested to completion with Hae III, Hinf I or Alu I. The reaction was set up as follows: 10ug of DNA, 1/10 of the final reaction volume of the 10X restriction enzyme buffer and 20 units of the enzyme were added. The DNA was incubated overnight at $37^{0}C$.

The DNA was electrophoresed on a 30 cm long, 0.9% agarose gel. The gel was run in TBE buffer for approximately 60 hours at 40 V. This meant that the 2 000 bp fragment of the marker, lambda Hind III DNA, had reached the end of the gel.

The gel was blotted onto nylon blotting membranes (Biorad), as previously described. The DNA was fixed onto the nylon membranes by baking the membranes in an oven at $80^{0}C$ for two hours.

8.5.2 HYBRIDIZATION AND WASHING OF THE MEMRANES

The fingerprinting membrane was prehybridized for 2 to 6 hours in prehybridization buffer (7% SDS, 1mM EDTA, 0.263M Na_2HPO_4, 1% BSA), at 65^0C for the M13 probe and 45^0C for the oligonucleotide probe, in a Hybaid hybridization oven. The denatured probe was then added to the prehybridization buffer and the probe was hybridised to the DNA overnight at the prehybridization temperature.

The probe was washed in the post hybridization buffers. The membrane hybridized to the M13 probe was washed in 2x SSC for 30 minutes at room temperature and then 15 minutes at 65^0C. The membrane hybridized to the oligonucleotide probe was washed in 6x SSC for 30 minutes at room temperature and then 15 minutes at 45^0C. After the last wash, the membrane was sealed in plastic and autoradiographed.

8.6 RESULTS AND DISCUSSION

Hybridization of the oligonucleotide probe, $(CAC)_5$, to black rhinoceros DNA resulted in fingerprinting gels that had a dark background. The hybridization patterns were not reproducible and the bands were not well resolved on the gel. Generally, the hydridization patterns obtained with the $(CAC)_5$ probe were difficult to score and to analyse (figure 8.3). A range of different buffers and hybridization conditions were attempted to resolve the fingerprinting bands on the gel. However, a clear fingerprinting gel could not be obtained using the $(CAC)_5$ probe. It has been reported that some probe-species combinations would not yield usable fingerprinting data. The success of a fingerprinting experiment depends upon the species-probe-restriction enzyme combination, as well as the extent of inter-individual variation within a species (Georges *et al*, 1988).

Fingerprinting with the cloned fragment of M13 was more informative. Once again, the optimum experimental conditions had to be empirically determined. This would appear to be the case with every new species-probe combination.

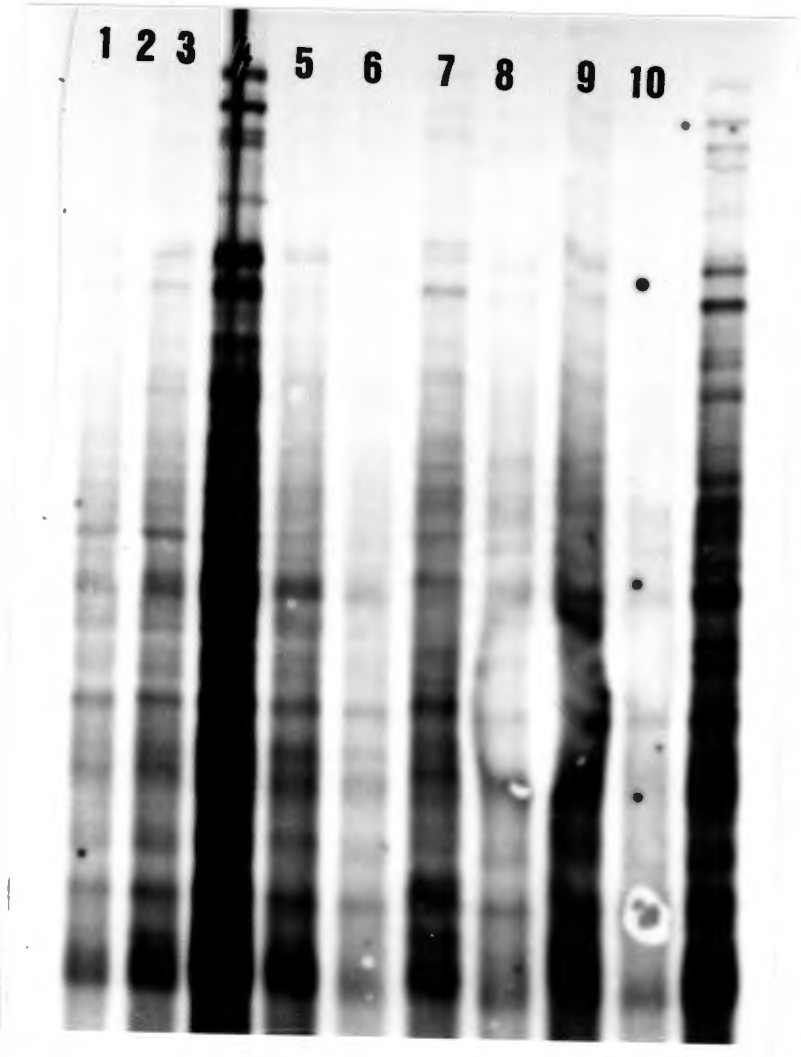

FIGURE 8.3

A Fingerprinting pattern obtained after hybridization of the multilocus $(CAC)_5$ probe to rhinoceros DNA. Lanes 1-3: *D. b. bicornis* and lanes 4-10: *D. b. minor*.

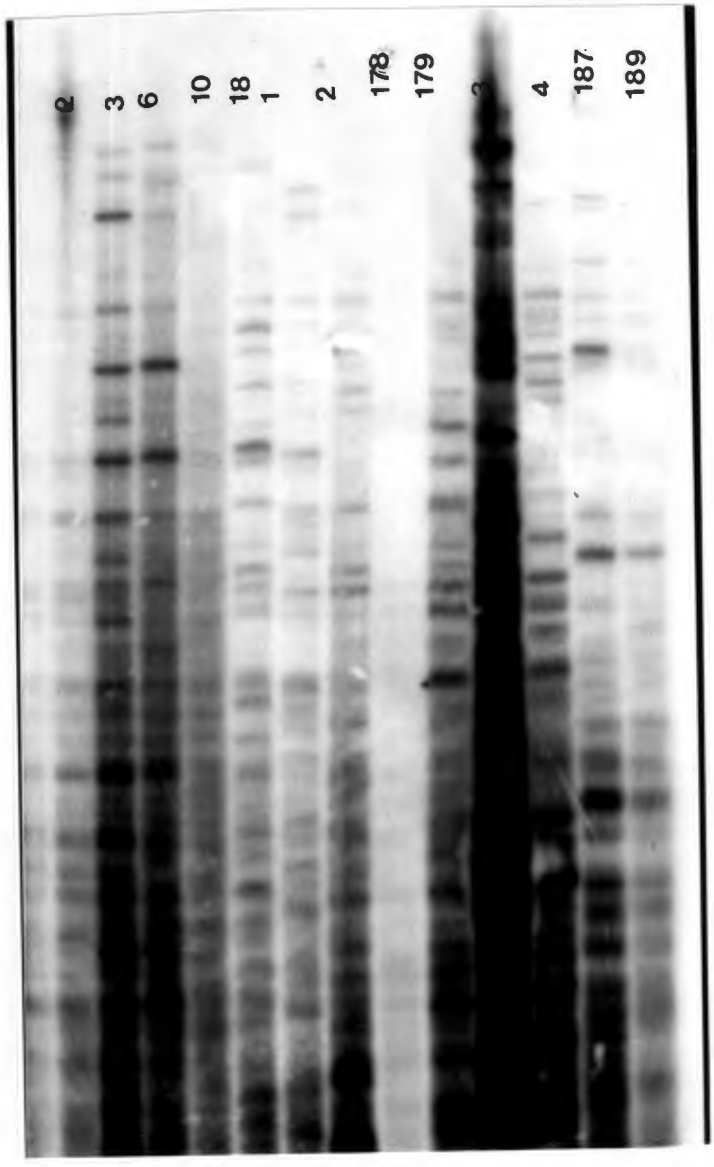

FIGURE 8.4

Fingerprinting pattern obtained after hybridization of multilocus M13 probe to rhinoceros DNA. From left to right. Lane: 2-5: *D. b. bicornis*, 6: *D. b. chobiensis*, 7: *D. b. michaeli*, 8-14: *D. b. minor*. Lanes 1,9 and 11 were not used.

Figure 8.4 shows the fingerprinting gel of the four different subspecies of black rhinoceros probed with M13. A number of bands could be discerned and an average of 30.8 bands (range 28 to 35) were scored. Bands larger that 10 000 bp and those smaller than 2 000 bp were not considered in the analysis. Bands in each lane were independently scored and a band sharing coefficient or F (the proportion of bands that are shared between two individuals) between each pair of individuals was calculated (Lansman *et al*, 1981).

A band sharing coefficient, $F = 2N_{AB}/(N_A + N_B)$, was calculated, where N_{AB} is the number of bands shared by two individuals having N_A and N_B bands each. The band sharing coefficient for black rhinoceroses ranged from 0.16 (between an individual from *D. b. michaeli* and one from *D. b. chobiensis*) to 0.81 (between two individuals from the same subspecies, *D. b. bicornis*) (Table 8.1).

The band sharing coefficient within the *D. b. bicornis* population averaged 0.76 and within the *D. b. minor* population it averaged 0.54. These values are consistent with the relative population sizes and distribution of these two subspecies. *D. b. minor* has a widespread distribution in Natal, while *D. b. bicornis* has a relatively limited distribution in the Etosha Game Reserve in Namibia. It is important to note that the black rhinoceroses are currently

TABLE 8.1

Summary of the band sharing coefficient between the different individuals of the four subspecies. Samples ER1, ER2, ER6 AND A10 are *D. b. bicornis*, A18 is *D. b. chobiensis*, AR1 is *D. b. michaeli* and BR2, A179, BR3, BR4, A187 and A189 are *D. b. minor*

	ER2 n=28	ER3 N=34	ER6 n=29	A18 n=32	AR1 n=29	BR2 n=29	A179 n=35	BR4 n=35	A187 n=33	A189 n=28
ER2										
ER3	0.81									
ER6	0.77	0.70								
A18	0.53	0.45	0.36							
AR1	0.63	0.51	0.41	0.16						
BR2	0.42	0.44	0.41	0.17	0.37					
A179	0.67	0.52	0.34	0.45	0.38	0.63				
BR4	0.48	0.38	0.45	0.45	0.30	0.30	0.37			
A187	0.49	0.33	0.29	0.46	0.36	0.41	0.41	0.38		
A189	0.54	0.24	0.28	0.40	0.38	0.45	0.41	0.72	0.72	

TABLE 8.2

Summary of the average band sharing coefficient between the four different subspecies.

SUMMARY OF BETWEEN - POPULATION VARIATION

	D.b. bicornis	D.b. chobiensis	D.b. michaeli	D.b. minor
D.b. bicornis	X	0.45	0.52	0.45
D.b. chobiensis	X	X	0.16	0.39
D.b. michaeli	X	X	X	0.36
D.b. minor	X	X	X	X

VALUES INDICATE F (BAND SHARING COEFFICIENT)

going through their bottleneck and the expected loss of genetic diversity would only be seen in later generations.

The band sharing coefficients between the different subspecies are summarized in Table 8.2. *D. b. chobiensis* and *D. b. michaeli* appeared to be genetically the most divergent pair of the four subspecies. They shared 16% (F = 0.16) of their fingerprinting bands. This value would be consistent with their geographical separation, but this estimate has limited value since it is based on only the one measurement. The values for the *D. b. chobiensis* comparisons fall within the expected range of values given by *D. b. minor* and the other two subspecies. This would be consistent with the observations that the restriction maps of *D. b. minor* and *D. b. chobiensis* are identical.

The sample from the *D. b. michaeli* individual was collected from the Addo Elephant Park in the eastern Cape, whose population was founded by four animals that were translocated from Kenya in the 1960's. The sample from the *D. b. chobiensis* individual was collected from the Kiboko region near the Namibian/Angolan border, where the last remaining wild population of this subspecies is thought to occur, if they are not now extinct.

The pair of subspecies that share the highest proportion of bands (F = 0.52) is *D. b. bicornis* and *D. b. michaeli*.

Similarly, a relatively high proportion of bands are shared between *D. b. bicornis* and *D. b. chobiensis* (F = 0.45) and *D. b. bicornis* and *D. b. minor* (F = 0.45).

The analyses and the statistical difficulties associated with DNA fingerprinting are well documented (Cohen, 1990, Evett *et al*, 1989, Lynch, 1988 and Morris *et al*, 1989). It has been shown that unbiased estimates of relatedness cannot be obtained at the individual level, without some knowledge of the allelic distribution in both the individuals of interest and the base population unless the proportion of the shared alleles between the unrelated individuals is essentially zero (Lynch, 1988). A number of technical complications also need to be taken into account when analysing DNA fingerprinting data. These include: the co-migration of different fragments to the same position on the gel, certain marker loci may be linked, band shifts due to DNA on the gel not running uniformly, the exclusion of the appropriate internal molecular size markers, as well as a host of other reasons. The proportion of shared fragments may merely reflect the proximity of any lanes on the gel and hence, the increased probability that more bands will match in two adjacent lanes than when two lanes are further apart (Lynch, 1988).

DNA fingerprinting has been used mostly in paternity or forensic analysis (Cawood, 1989 and Weatherhead &

Montgomerie 1991). Also, most of DNA fingerprinting research has concentrated on human studies. The major exception has been the determination of parentage in polyandrous birds who have complex social systems where extra-pair mating and fertilization occurs (Burke *et al*, 1989, Kempenaers *et al*, 1992 and Oring *et al*, 1992). DNA fingerprinting has been utilized for paternity estimation in other mammals (Dixson *et al*, 1988 and Packer *et al*, 1991).

Despite some of the limitations associated with DNA fingerprinting as a means to determine relatedness within and between different populations, it has been used in studies on whales (Amos & Dover, 1990), foxes (Gilbert *et al*, 1990), swans (Meng *et al*, 1990), naked mole rats (Reeve *et al*, 1990) and introgression in breeding programs of farm animals (Hillel *et al*, 1990). However, the consensus is that DNA fingerprinting with multiloci probes (Jeffreys probe, M13 and the synthetic oligonucleotides) would only have a limited application in the sphere of animal population genetics, and that single locus probes (SLP) could prove to be more appropriate (Pemberton *et al, 1991*). SLPs are significantly more sensitive (as they are species specific), technically easier to use, the interpretation of the SLP profiles are easy and the ability to obtain allele frequency estimates allows important population genetic questions to be addressed. The data obtained in this manner are also amenable to database formation, which would allow

comparisons between experiments to be made (Amour *et al*, 1990).

DNA fingerprinting is still a young, rapidly evolving technique, with many of its applications and the analysis of the data requiring further development. It does, however, hold much promise as a powerful tool in population genetics and population geneticists and behavioural biologists should be patient with the learning phase of DNA fingerprinting, since in the long run it may still fulfill many of its earlier promises.

The main interest of this limited study on the subspecies of black rhinoceros was the exploration of the value of DNA fingerprinting as a way of estimating: (a) the relative amount of genetic diversity within a rhinoceros subspecies or population, since most animal studies so far have been on birds and few studies have been done on large, wild mammal populations and (b) the amount of diversity between the subspecies.

It is possible to conclude from the data on *D. b. bicornis* (three individuals, F = 0.76) and *D. b. minor* (five individuals, F = 0.54) only that the results are consistent with former having a lesser genetic diversity than the latter. The amount of band sharing between subspecies was generally lower than that within subspecies, which is

consistent with their subspecific status. A larger number of individuals will have to be analysed before more definitive statements can be made, and the above results should be considered only as feasibility study.

With this database now available, it would be useful to sample, for example, the Addo population and determine whether there is a significant reduction in the band sharing coeffecients in this population. The degree of band sharing will help management decisions and whether to introduce additional individuals to expand the gene pool.

APPENDIX

RESTRICTION FRAGMENT SIZES FOR LAMBDA DNA DIGESTED WITH HIND III (in kb)

23 130
9 416
6 557
4 361
2 322
2 027
564
125

BUFFERS

1. EXTRACTION BUFFER

100mM Tris.HCL (pH 8)
150mM NaCl
20mM EDTA (pH 8)
10% w/v sucrose

2. TE BUFFER (pH 8)

10mM Tris.HCl (pH 8)
1mM EDTA (pH 8)

3. STE BUFFER (pH 8)

10mM Tris.Cl (pH 8)
1mM EDTA (pH 8)
100mM NaCl

4. TAE BUFFER (50X) for 1l

242g Tris base
57.1ml Glacial acetic acid
100ml 0.5M EDTA (pH 8)
5ml 10% Sodium pyrophosphate

5. KGB (POTASSIUM GLUTAMATE) BUFFER (2X)

200mM Potassium glutamate
50mM Tris acetate (pH7.6)
20mM Magnesium acetate
2mg/ml Bovine serum albumin
1mM 2-mercaptoethanol

6. DIGESTION BUFFER

100Mm NaCl
10mM Tris.HCl (pH 8)
25mM EDTA
0.5% SDS
0.1mg/ml proteinase K

7. 20X SSC for 1l

175.3 g NaCl
88.2g Sodium citrate
pH to 7 with 10M NaOH

8. DENATURATION BUFFER

1.5M NaCl
0.5M NaOH

9. NEUTRALIZATION BUFFER

20x SSC
0.5M Tris.HCl (pH 7.5)

10. PREHYBRIDIZATION BUFFERS

A) mtDNA PROBE

6X SSC
0.1% SDS
0.06% Sodium pyrophosphate
0.25% Blotto (low fat dried milk powder)

B) M13 PROBE AND $(CAC)_5$ PROBE

7% SDS
1mM EDTA
0.263M Na_2HPO_4
1% Bovine serum albumin

11. GEL LOADING BUFFER

0.25% w/v Bromophenol blue
40% w/v Sucrose
20mM EDTA

12. TBE BUFFER (5X) for 1l

54g Tris base
27.5g Boric acid
20ml 0.5M EDTA (pH 8)