

https://dspace.org/

Mitochondrial DNA polymorphisms in Southern African populations.

Soodyall, Himladevi

1993
http://hdl.handle.net/10539/22319

Downloaded from a DSpace repository. https://dspace.org/

MITOCHONDRIAL DNA POLYMORPHISMS IN SOUTHERN AFRICAN POPULATIONS

HIMLADEVI SOODYALL

A thesis submitted to the Faculty of Medicine, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy

Johannesburg, 1993

ABSTRACT

The subject of this thesis is mitochondrial DNA (mtDNA) variation in southern African populations. The purpose of this study was twofold. Firstly, mtDNA variations were used to investigate the genetic affinities of Negroid, Khoisan, Caucasoid and "Coloured" populations in an attempt to refine theories on southern African population affinities and prehistory.

MtDNA variations were detected using two different methods. The first method makes use of restriction fragment length polymorphisms (RFLPs) detected with the restriction enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II in 795 unrelated individuals from twenty ethnic groups within the Khoisan, Negroid, Caucasoid and "Coloured" populations from South Africa and Namibia. The combinations of the various restriction enzyme patterns (morphs) for the enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II (in this order), were used to derive the mtDNA type for each individual studied. This resulted in the discovery of 62 distinct mtDNA types: 30 of which had been previously reported, 28 out of 32 resulted from new combinations of enzyme morphs and 4/32 were due to the discovery of new enzyme morphs (*Msp*I-17 in the Ashkenazi Jewish population and *Ava*II-31, *Ava*II-32 and *Ava*II-33 in the South African "Coloured" population).

The second method involves sequencing approximately 750

base pairs of mtDNA contained within the two hypervariable segments within the non-coding control region of the mtDNA molecule in 144 individuals, most of whom where investigated for mtDNA RFLP variations. Pairwise comparisons of mtDNA sequences revealed 119 variant sites which gave rise to 129 unique mtDNA types.

Secondly, this study makes use of the mtDNA diversity found in the indigenous southern African populations to shed more light on the debate concerning the origins of modern humans. In particular, the mtDNA data obtained in this study were used to evaluate the African origin (single-replacement) hypothesis versus the multiregional hypothesis for the origin of modern humans. There are two pieces of evidence gleaned from the findings of this thesis that lend additional support to the hypothesis that the mtDNA in modern humans evolved in Africa. Firstly, both mtDNA RFLP data and control region sequence data reveal that southern African populations are genetically more diverse than Caucasoid and Asian populations. In fact, using the sequence divergence calculated for Khoisan populations (0.256%) from RFLP data, it was estimated that this group may have diverged from the common ancestor approximately 128 000 to 64 000 years ago compared with 104 000 to 52 000 year ago for Negroids (0.208%) and 48 000 to 24 000 years ago for Caucasoids (0.096%). Although the divergence times estimated from sequence data is somewhat higher (167 796 years for Khoisan, 148 588 for Negroid and 123 729 for

Caucasoids) than that estimated from RFLP data, these estimates lend additional support for the suggestion that Khoisan populations are more diverse than Negroid populations and may contain in their genes the molecular events characterizing the evolutionary prehistory of modern humans.

The second piece of evidence in support of the African origin model was revealed from phylogenetic analyses. By making use of the UPGMA and neighbour-joining methods (RFLP data) and neighbour-joining and parsimony methods (sequence data), the distribution of mtDNA types found in southern African populations were found to be more concentrated in the deepest branches of trees compared with mtDNA types from Caucasoid and Asian populations commonly found in the shallower lineages of the tree.

The findings of this thesis, therefore, are consistent with the African origin (single-replacement) hypothesis for the origin of modern humans which claims that the human mtDNA ancestor lived in Africa, probably southern Africa, fairly recently (less than 150 000 years ago). This would suggest that the tranformation from archaic to modern humans occurred in southern Africa and that modern humans then spread from Africa to replace the resident populations of Europe and Asia.

DECLARATION

I declare that this dissertion is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University. I declare that ths work has been approved by the Ethics committee of the University of the Witwatersrand, and the certificate number is 13/2/89.

H Soodyall

Himladevi Soodyall

11th day of January, 1993

DEDICATION

This work is dedicated to my parents and brothers for their encouragement and support throughout the years.

ACKNOWLEDGEMENTS

There are several people to whom I am indebted to for making this research project possible:

To Professor Trefor Jenkins, my supervisor and mentor, who has afforded me the opportunity and the use of facilities in his laboratory (Department of Human Genetics, South African Institute for Medical Research) to advance my research goals. I am most grateful to him for his encouragement, guidance and unstinting support in helping me mould a research career since I joined the Department in 1987.

To the volunteers who willingly consented to participate in this study by donating blood samples; to Ms Linda Dodds of Windhoek State Laboratories and Dr Seidel of Namibia Blood Transfusion Service for their assistance with the collection of blood samples from the Dama, Herero and Nama populations from Namibia; Dr Meyer of the South African Defence Force who was based at the Omega Camp for collecting blood samples from the Sekele San and the Kwengo populations; Highveld Blood Transfusion Service who assisted in the collection of blood samples from Negroid and Indian populations in the Johannesburg region; and to colleagues in the Department of Human Genetics who participated in the collection and processing of blood samples collected during field trips.

To the Pennsylvania State University, USA, for awarding me a SHARE* Fellowship (August 1990 to December 1990) which afforded me the opportunity of working with scientists in the USA. I thank Dr Lamarr-Kopp and Tannaz Rahman (Office for International Programs) for their friendship and the superb manner in which they looked after my welfare throughout my stay at Penn State.

To Professor Kenneth A. Johnson (Department of Biochemistry, Penn State) who accepted the role of mentor as part of the arrangement of the SHARE fellowship, I should like to express my sincerest gratitude for giving me a "home" for those five months. Under the expert guidance of Dr Smita Patel in his laboratory, I acquired valuable insight and knowledge in the fields of protein chemistry and enzyme kinetics. I am most grateful to Smita and Ken for allowing me to bring to South Africa the gene 5 protein which I had isolated and purified in their Department and have used for DNA sequencing in this research project. I also thank all those in Ken Johnson's laboratory for their kindness and friendship, in particular Smita, for making my stay at Penn State a most memorable one.

I also benefitted greatly by spending another three months working with Professor Mark Stoneking, Department of Anthropology, The Pennsylvania State University, during which time I had launched the sequencing work of this project. I am most grateful to Prof. Stoneking for

allowing me the use of his laboratory facilities and for his guidance and support throughout this time, and for his subsequent collaborations. I should also like to express my gratitude to all those in the Department of Anthropology for their kindness and friendship. I am most grateful to Mark Stoneking and the South African Institute for Medical Research for the financial support during this period.

I should like to thank Dr Stewart Grant (Department of Genetics, the University of the Witwatersrand) for his assistance with computational analysis and discussions during the earlier phase of this project. In recent times I have benefitted greatly from assistance from Stephen Sherry (Department of Anthropology, Penn State University) and Professor Andrew G. Clark (Department of Biology, Penn State University) who have helped me in the analysis of the data.

To Professor Masatoshi Nei (Institute for Molecular Evolutionary Genetics, Penn State Univeristy) for the RESTSITE, DNADIST and Neighbour-joining computer software.

To Professor George Nurse and Professor Trefor Jenkins for reading drafts of this thesis and for their valued comments and suggestions.

To Etienne Swannepoel and Arden Erasmus (Illustration Unit) and Mike Lanesman and Guy Hall (Photographic Unit) who assisted with the illustrations and photography, respectively.

To the South African Institute for Medical Research and the Medical Research Council for their generous contributions which afforded me the opportunities of attending scientific meetings both locally and abroad.

Last, but not least, I thank all my colleagues, friends and family for their helpfulness, encouragement and support. In particular, I wish to express by deepest gratitude to my parents who have made personal sacrifices to afford me the opportunities to pursue my academic career. I am most grateful to my brother Anesh for being so patient, tolerant and understanding and for supporting me through some very difficult times and for being a pillar of support, my best friend and companion over the years.

*SHARE (**S**cholarship, **H**elp, **A**cademic exchanges, **R**eview of equity holdings and **E**ducational programs)

TABLE OF CONTENTS **PAGE**

ABSTRACT..... ii

DECLARATION.....v

DEDICATION.....vi

ACKNOWLEGEMENTS.....vii

TABLE OF CONTENTS.....xi

LIST OF TABLES.....xviii

LIST OF FIGURES.....xxii

LIST OF ABBREVIATIONS.....xxvi

PUBLICATIONS.....xxvii

CHAPTER ONE: INTRODUCTION.....1

1.1 Development of evolutionary theory.....1

1.2 Aims of study and outline of thesis.....5

1.3 Advantages of using mtDNA.....6

1.4 Organization of human mtDNA.....8

1.5 Methods for determining mtDNA variations.....13

1.5.1 Low-resolution mapping.....13

1.5.2 High-resolution mapping.....14

1.5.3 DNA amplification and sequencing.....15

CHAPTER TWO: SUBJECTS AND METHODS.....17

2.1 Subjects.....17

2.1.1 Classification of populations.....21

2.1.2 Bantu-speaking Negroids.....21

2.1.2.1 Southeastern Bantu-speakers.....21

2.1.2.2 Venda and Lemba.....24

2.1.2.3 Southwestern Bantu-speakers.....25
2.1.3 Khoisan speakers.....27
2.1.4 Khoisan-speaking Negroids.....28
2.1.5 Caucasoid groups.....28
2.1.5.1 European immigrants.....28
2.1.5.2 Asian Indians.....29
2.1.6 Hybrid communities.....30
2.2 Methods.....31
2.2.1 Reagents and recipes.....31
2.2.2 Collection of blood.....31
2.2.3 Genomic DNA extraction.....32
2.2.4 DNA digestion and electrophoresis.....33
2.2.5 Southern blotting.....35
2.2.6 MtDNA extraction and purification.....37
2.2.7 Radiolabelling of mtDNA.....39
2.2.8 Hybridization and autoradiography.....41
2.2.8.1 Biodyne membranes.....41
(a) Prehybridization.....41
(b) Hybridization.....42
(c) Post-hybridization washes.....42
2.2.8.2 Hybond-N membranes.....43
(a) Prehybridization.....43
(b) Hybridization.....43
(c) Post-hybridization washes.....43
2.2.8.3 Reuse of hybridization mixes.....44
2.2.8.4 Autoradiography.....44
2.2.9 RFLP analysis.....44

2.2.10 Amplification of double stranded mtDNA from genomic DNA.....45
2.2.11 Preparation of single stranded DNA for sequencing.....48
2.2.11.1 Asymmetric PCR.....48
2.2.11.2 PCR with one phosphorylated primer followed by λ exonuclease digestion.....50
(a) Phosphorylation of primers.....50
(b) PCR with one phosphorylated primer.....51
2.2.12 Sequencing of single-stranded DNA.....52
2.2.12.1 Reconstitution of T7 DNA polymerase.....52
2.2.12.2 Sequencing protocol.....53
2.2.12.3 Sequencing gel electrophoresis.....54
2.3 Statistical analysis.....56
2.3.1 MtDNA variation within populations.....56
2.3.1.1 Genetic diversity.....56
2.3.1.2 Nucleotide diversity.....56
2.3.2 MtDNA variation between populations.....57
2.3.2.1 G_{ST}.....57
2.3.2.2 Nucleotide divergence.....58
2.4 Neutrality tests.....60
2.5 Phylogenetic analysis.....61

CHAPTER THREE: RESTRICTION FRAGMENT LENGTH POLYMORPHISM...62
3.1 Background.....62
3.2 Restriction enzyme analysis.....63
3.2.1 Description of morphs.....63
3.2.1.1 *Hpa*I.....63

3.2.1.2 *Bam*HI.....71
3.2.1.3 *Hae*II.....76
3.2.1.4 *Msp*I.....82
3.2.1.5 *Ava*II.....91
3.2.1.6 *Hinc*II.....102
3.2.2 MtDNA types.....108
3.2.3 Restriction site polymorphism.....125
3.3 Statistical analysis.....131
3.3.1 Genetic diversity.....131
3.3.2 Extent of population subdivision, G_{ST}.....134
3.3.3 Nucleotide sequence divergence.....138
3.4 Phylogenetic analysis.....140
3.4.1 Phylogenetic network of mtDNA types.....140
3.4.2 UPGMA analysis of mtDNA types.....147
3.4.3 Cluster analysis.....152
3.4.4 Do mtDNA mutations provide a selective advantage?.....154
3.5 Conclusions.....158

CHAPTER FOUR: CONTROL REGION SEQUENCE VARIATION IN SOUTHERN AFRICAN POPULATIONS.....161
4.1 Background.....161
4.2 Sequence analysis.....163
4.2.1 Variant sites.....163
4.2.2 Characterization of mutations.....170
4.2.2.1 Substitutions.....170
4.2.2.2 Length mutations.....172
4.3 MtDNA sequence types.....174

4.4 Phylogenetic analysis.....177
4.4.1 Pooling of data.....177
4.4.2 Parsimony analysis.....178
4.4.3 Neighbour-joining analysis.....186
4.5 Cluster analysis.....190
4.6 Rate of mtDNA evolution and age of the common mtDNA ancestor.....193
4.7 Frequency distribution of pairwise sequence difference.....196
4.8 Conclusions.....205

CHAPTER FIVE: ETHNOGRAPHIC PREHISTORY OF SOUTHERN AFRICAN POPULATIONS: BRINGING TOGETHER THE ARCHAEOLOGAL, LINGUISTIC AND GENETIC EVIDENCE.....208
5.1 Linguistic evidence.....209
5.1.1 Bantu languages.....210
5.1.2 Khoisan languages.....214
5.2 Archaeological evidence.....221
5.3 Genetic data.....233
5.3.1 MtDNA variation in Negroids.....234
5.3.1.1 Southeastern vs southwestern Bantu-speakers.....234
5.3.1.2 Relationship between the Herero and Ambo....235
5.3.1.3 Matrilineal inheritance in the Herero.....236
5.3.1.4 Origins of the Dama.....238
5.3.2 Genetic relationships between the San and the Khoi.....240
5.3.3 MtDNA polymorphism in "Coloureds".....248

5.3.4 MtDNA variation in the South African Indian population.....253
5.3.5 MtDNA structure of European Caucasoids in South Africa.....265
5.4 Conclusions.....269

CHAPTER SIX: EVOLUTION OF HUMAN MTDNA: SOUTHERN AFRICAN EVIDENCE.....272
6.1 Hominid evolution: fossil evidence.....272
6.2 Evolution of anatomically modern humans.....276
6.2.1 Multiregional evolution.....276
6.2.2 Replacement evolution.....277
6.2.3 Testing the two models of evolution.....278
6.3 Evolution of modern human: genetic evidence.....283
6.4 Southern African evidence.....284

REFERENCES.....290

APPENDIX 1: The sample I.D., ethnicity, language group, sex, place of collection and laboratory reference number of individuals investigated for mtDNA restriction fragment length polymorphisms.....309

APPENDIX II: Recipes and solutions.....323

APPENDIX 111: The restriction enzyme patterns (morphs) and the corresponding mtDNA type (mitotype) obtained for 795 individuals from southern Africa.....326

APPENDNDIX IV: MtDNA sequence types in southern African populations.....341

LIST OF TABLES **PAGE**

2.1. The sample sizes and ethnicities of southern African populations investigated in this study.....20

2.2. The specifications of the restriction enzymes used for mtDNA RFLP analysis.....34

2.3. Oligonucleotide primers used in the PCR and sequencing reactions.....47

3.1. The numbers and the frequencies (percentages) of *Hpa*I morphs in twenty ethnic groups from southern Africa.....70

3.2. The numbers and the frequencies (percentages) of *Bam*HI morphs in twenty ethnic groups from southern Africa.....75

3.3. The numbers and the frequencies (percentages) of *Hae*II morphs in twenty ethnic groups from southern Africa.....81

3.4. The numbers and the frequencies (percentages) of *Msp*I morphs in twenty ethnic groups from southern Africa.....89

3.5. The numbers and the frequencies (percentages) of *Ava*II morphs in twenty ethnic groups from southern Africa.....100

3.6. The numbers and the frequencies (percentages) of *Hinc*II morphs in twenty ethnic groups from southern Africa.....107

3.7. The numbers and the frequencies of mtDNA types discovered in twenty southern African populations.....109

3.8. Comparisons of variant sites detected with the enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II for all mtDNA types discovered in this study with corresponding positions in the reference sequence.....126

3.9. Correlation between the number of mtDNA types, genetic diversity and percent sequence divergence within ethnic groups from southern Africa.....133

3.10. Index of interpopulation differences estimated by calculating G_{ST}.....137

3.11. The average nucleotide divergence calculated from the mean pairwise differences for all pairs of mtDNA types found within (diagonal) and between the twenty ethnic groups investigated from southern Africa.....139

3.12. The Ewens-Watterson and the Tajima tests of neutrality on the restriction site variation in twenty southern African populations.....157

4.1. The sample sizes and the origins of the various ethnic groups investigated for mtDNA control region sequence variation.....164

4.2. Frequencies of the variant sites in the Khoisan, Negroid and Caucasoid populations determined by counting the number of individuals (M) who differed from the reference sequence at the positions given divided by the number of individuals from whom sequence data was available (N).....166

4.3. Characterization of mutations in the two hypervariable regions contained within the control region of the mtDNA molecule.....171

4.4. MtDNA types represented by more than one individual.....176

4.6. The numbers and the distribution of mtDNA types found in Negroid, Khoisan, Caucasoid, Asian and Papua New Guinean populations as inferred from parsimony analysis.....181a

4.7. The numbers and the distribution of mtDNA types found in the Negroid, Khoisan, Caucasoid, Asian and Papua New Guinean populations as inferred from neighbour-joining analysis.....187b

4.8. Expansion times of southern African populations calculated from the parameters of the Rogers and Harpending model.....204

5.1. The numbers and the frequencies (in percent) of the various mtDNA types found in one Negroid and three Khoisan populations from southern Africa, together with an index of genetic diversity (*h*)..247

5.2. The numbers and the frequencies of mtDNA types found in the Richtersveld and the Johannesburg "Coloured" populations from South Africa.....250

5.3. MtDNA variations detected in four S.A. Indian linguistic/religious groups, including the pooled sample, with six restriction enzymes.....259

5.4. The numbers and the frequencies (in percentage) of mtDNA types found in four linguistic/religious South African Indian populations and Hindus from Nepal and India.....260

5.5. Correlation between number of mtDNA types and homogeneity (*F*) in Indian groups from South Africa, Nepal and India.....263

5.6. The frequencies (percentages) of mtDNA types found in European Caucasoids compared with South African Caucasoids of European origin.....268

6.1. The estimated age, sites at which discovered and the cranial capacities of *Australopithecus* and *Homo* fossils discovered in Africa.....274

6.2. Theoretical predications from models of *Homo sapiens* evolution.....279

6.3. Estimated divergence times for the Negroid, Khoisan and Caucasoid populations from the average amount of sequence divergence as estimated from mtDNA RFLP and sequence data.....287

LIST OF FIGURES **PAGE**

1.1. The organization of the human mitochondrial DNA genome.....11

1.2. Structure of the control region showing the relative locations of the functional segments.....12

2.1. Map of southern Africa showing the locations of the different ethnic groups investigated for mtDNA variation in this study.....19

3.1. (a). Fragment patterns of *Hpa*I morphs discovered in southern African populations.....68

(b). Phylogeny showing the relationship of the eight *Hpa*I morphs discovered to date.....69

3.2. (a). Fragment patterns of *Bam*HI morphs discovered in southern African populations.....73

(b). Phylogeny showing the relationship of the five *Bam*HI morphs discovered to date.....74

3.3. (a). Fragment patterns of *Hae*II morphs discovered in southern African populations.....79

(b). Phylogeny showing the relationship of the twelve *Hae*II morphs discovered to date.....80

3.4. (a). Fragment patterns of *Msp*I morphs discovered in southern African populations.....86

(b). Phylogeny showing the relationship of the seventeen *Msp*I morphs discovered to date.....87

3.5. (a). Fragment patterns of *Ava*II morphs discovered in southern African populations.....96

(b). Phylogeny showing the relationship of all *Ava*II morphs discovered to date.....98

3.6. (a). Fragment patterns of four *Hinc*II morphs discovered in southern African populations.....104

(b). Phylogeny showing the relationship of all *Hinc*II morphs discovered to date.....105

3.7. Partial phylogeny of mtDNA types showing the relationship of some of the mtDNA types discovered to date.....144

3.8. (a). UPGMA tree showing the relationship of the 62 types discovered in this study (bold) when compared with the 149 mtDNA types discovered to date using the six enzyme system.....148

(b). UPGMA dendrograms of all mtDNA types described in Fig. 3.8.a showing the relationship of mtDNA types found in the pooled African, Asian, Caucasian and Amerindian populations.....150

3.9. Neighbour-joining cluster dendrogram showing the genetic relationship of twenty southern African populations investigated for mtDNA variations....153

4.1. Frequency of variant sites contained within the two hypervariable regions of the mtDNA molecule in Khoisan, Negroid and Caucasoid populations.....169

4.2. Length mutations caused by an insertion of a C residue succeeding the T at position 862.....173

4.3. Phylogenetic tree showing the relationship of the 250 unique mtDNA sequence types as inferred by parsimony analysis.....181a

4.4. Phylogenetic tree showing the relationship of the 250 unique mtDNA sequence types found in African, Asian, Caucasoid and Papua New Guinean populations as inferred by the neighbour-joining method.....187

4.5. (a). Neighbour-joining cluster dendrogram showing the genetic affinities of the seven southern African populations investigated for sequence variation within the control region of the mtDNA molecule.....192b

(b). Neighbour-joining cluster dendrogram showing the relationship of the African, Asian, Caucasoid and Papua New Guinean populations based on the sequence variation contained within the two hypervariable regions of the mtDNA molecule.....192c

4.6. Histograms showing the frequencies of the number of mutational differences obtained by pairwise comparisons of mtDNA sequence types in: (a) total sample of 144 individuals, (b) the southeastern Bantu-speaking Negroid group referred to as "Bantu" in this study, (c) Herero, (d) Dama, (e) Nama and (f) San populations.....201

5.1. Distribution of Bantu languages in Africa.....212

5.2. Maps of southern Africa showing the location/s (present or past) of the various Bantu-speaking Negroid ethnic groups: (a) southeastern Bantu-speakers, (b) southwestern Bantu-speakers and (c) central Bantu-speakers.....215

5.3. The locations (present or past) of Khoisan populations in southern Africa.....219

5.4. Distribution of Wilton (W) and Smithfield B (B) archaeological sites which have been associated with regions occupied by prehistoric Bush-speaking and Hottentot-speaking peoples, respectively.....224

5.5. Arrangement of the (a) Central Cattle Pattern, (b) Forest Pattern and (c) Ila House Plan of Early Iron Age settlement sites which have been associated with southeastern, southwestern and a mixture of southeastern and southwestern Bantu-speaking Negroid groups, respectively, in southern Africa.....229

5.6. Map of southern Africa showing the possible routes of migration of eastern Bantu-speaking Negroids into southern Africa based on the correlation of ceramic styles and language.....231

5.7. Map showing the regions in India from where the South African Hindi, Tamil, Gujerati and Moslem speaking Indian groups originated.....254

LIST OF ABBREVIATIONS

ACD	acid citrate dextrose
APS	ammonium persulphate
bis	N, N^{1}-methylene-bis-acrylamide
bp	base pairs
cpm	counts per minute
°C	degrees Celsius
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetic acid
g	grams
kb	kilobase
µg	microgram
µl	microlitre
µM	micromolar
mg	milligram
ml	millilitre
mM	millimolar
min	minute
mtDNA	mitochondrial DNA
M	molar
PCR	polymerase chain reaction
RFLP	restriction fragment length polymorphism
SDS	sodium dodecyl sulphate
ssDNA	single-stranded DNA
TEMED	N,N,N^{1},N^{1}, tetramethylethylenediamine

PUBLICATIONS

Merriwether, D.A., Clark, A.G., Ballinger, S.W., Schurr, T.G., Soodyall, H., Jenkins, T., Sherry, S.T., Wallace, D.C. (1991). The structure of human mitochondrial DNA variation. *Journal of Molecular Evolution* 33: 543-545.

Sherry, S.T., Rogers, A.R., Harpending, H., Soodyall, H., Jenkins, T. and Stoneking, M. (1992). Pairwise differences of mtDNA reveal recent human population expansions (submitted to *Proceedings of the National Academy of Science, USA*.).

Soodyall, H. and Jenkins, T. (1992a). Mitochondrial DNA polymorphisms in Khoisan populations from southern Africa. *Annals of Human Genetics* (in press).

Soodyall, H. and Jenkins, T. (1992b). Mitochondrial DNA polymorphisms in Negroid populations from Namibia: New light on the origins of Dama, Herero and Ambo. *Annals of Human Biology* (in press).

Soodyall, H. and Jenkins, T. (1992c). Mitochondrial DNA polymorphisms in the South African Indian population. (submitted to *Gene Geography*).

Soodyall, H. Clark, A.G. and Jenkins, T. Ethnographic history of Southern African populations: Mitochondrial DNA evidence (in preparation).

Soodyall, H., Stoneking, M. and Jenkins, T. Evolution of mitochondrial DNA in Africa (in preperation).

CHAPTER 1:

INTRODUCTION

1.1 Development of evolutionary theory

By the middle of the eighteenth century, after centuries of speculation on the origins of life, nearly all scientists and philosophers agreed that however living things were created, they were created in their present form and were immutable and quite separate from one another. This doctrine, known as the "fixity of species" was typified in the work of Carolus Linnaeus (1707-1778) who developed the first comprehensive classification system for living things. The result was a neat, rigid hierarchy with all of nature was fixed in an orderly and static arrangement (Jolly and Plog 1987).

This doctrine was questioned by Georges Louis de Buffon (1707-1788), and later Jean Baptiste de Lamarck (1744-1829), both of whom argued that organisms were constantly changing physiologically in order to adapt better with their environments. These acquired changes were then passed on to their offspring, ultimately, introducing changes within a species. Unfortunately, this type of scientific innovation was not accepted during the time of the French Revolution since they were regarded as antireligious and radical.

By the end of the eighteenth century, it was geology which provided the major stimulus for the development of evolutionary theory. Georges Cuvier (1769-1832) proposed the "catastrophism" theory to explain the successive geological layers of the earth and their "strange" fossil content. This theory was extended by James Hutton (1726-1797) who proposed that the earth is extremely ancient and that its rocks are continually laid down, uplifted and eroded. On the other hand Charles Lyell (1797-1875) who argued that there had been no catastrophes, suggested that these changes of the earth accumulated gradually due to natural changes (Lyell 1873).

Inspired by the work of Lyell, Charles Darwin (1809-1882) developed the theory of natural selection. Darwin realized that more organisms are born than the environment can support and only a minority succeed in reproducing, thus causing constant competition within each species. Some adaptive traits give individuals an advantage in their struggle, allowing them to produce the most offspring (Darwin 1898).

One major weakness of Darwin's theory was that there was no systematic explanation of how "favoured" characteristics were inherited. The contemporary belief, was that each individual inherited a blend of its parents' characteristics, but in that case newly appearing advantageous variations would be lost before natural

selection could act on them. Gregor Mendel (1822-1884) provided the genetic foundation that Darwin's theory lacked. Mendel, who carried out experiments in an effort to ascertain how traits were transmitted from one generation to the next, discovered that individual units of breeding information (genes) were passed on as discrete units. In one individual, a gene's effect might be blended or even suppressed by the effects of other genes, but the gene itself remained intact, and could be passed on to the next generation, to be expressed and thus exposed to natural selection.

In the first decade of this century, shortly after the rediscovery of Mendels's work, another major discovery appeared to provide a solution to the problem: genes sometimes underwent random changes, or mutations, which could produce new traits which were then transmitted to the offspring.

Since that time, it has been recognized that both mutation and natural selection are essential to evolutionary process. In the 1920s and 1930s mathematical evolutionists like Ronald Fisher, J.B.S. Haldane and Sewall Wright, began to build a new body of evolutionary theory based on a synthesis of Darwin's and Mendel's discoveries. The introduction of mathematical genetics revealed that even small differences among individuals could produce sufficient evolutionary changes to account for the observed

diversity of the natural world (Jolly and Plog 1987). This body of evolutionary theory allows us to trace the biological changes that gradually transformed our early primate history of 70 million years ago - from tree dwelling creatures to modern *Homo sapiens*.

In view of the limited number of fossils discovered to date, the recent innovation has been to use genetic polymorphisms found in living populations, rather than relying solely on the paleontological evidence, to reconstruct the evolution of modern humans. Using the variation contained in the mitochondrial DNA (mtDNA) genome, it has been suggested that all mtDNA types found in modern humans can be traced back to a single ancestor who lived in Africa some 200 000 years ago (Cann *et al.* 1987). The most critical and controversial feature of this hypothesis is the relatively recent age (approximately 100 000 years) for the human mtDNA ancestor. This would suggest that the transformation of humans from archaic to modern took place fairly recently in a single region of the world and then replaced the resident *Homo erectus* hominids in Europe and Asia (Stringer and Andrews 1988; Stoneking and Cann 1989). If this estimate is wrong, and the actual age of the common ancestor instead approaches 1 million years BP, then it would support the view that the transformation of modern humans occurred at about the same time in different parts of the world (Wolpoff 1989).

This thesis examines the mtDNA variations of various southern African populations to shed more light on this debate.

1.2 Aims of study and outline of thesis

The aim of this investigation has been to investigate mtDNA variation in southern African populations to reveal the genetic structure of southern African populations, and to refine theories on population affinities previously propounded on the basis of linguistic and archaeological evidence as well as by the use of classical genetic markers. This will allow us to provide answers to the following questions:

1. What are the relationships between the different Bantu-speaking groups? Can we infer migration patterns of Bantu-speakers to support the theory of migration and the spread of the Bantu languages from the West African coast?
2. What are the relationships between the San and Khoi groups?
3. What are the relationships between the Khoisan and Negroid populations?
4. What is the genetic composition of the southern African "Coloured" population?

In addition, the patterns of mtDNA variations between southern African populations and populations from other

geographic regions (Asia, Europe, Papua New Guinea as well as from other parts of Africa) would be compared in order to provide a better understanding of the origin and differentiation of human mtDNA.

In the remaining sections of this chapter, a brief description of mtDNA structure, function and anthropological value are described. Chapter 2 describes the methodology used to obtain the data and also gives a brief description of the populations sampled. Two different methods were used to study the mtDNA variation in southern African populations: restriction fragment length polmorphisms (RFLP's) using six restriction enzymes and mtDNA control region sequencing. The findings of this study using these methods are presented in Chapter 3 and Chapter 4, respectively. The contribution of mtDNA polymorphisms, in conjunction with linguistic, archaeological and historical data, in revealing the prehistory of southern Africa are disscussed in Chapter 5. The findings of this thesis with respect to the southern African evidence concerning the origin and differentiation of human mtDNA are elaborated on in Chapter 6.

1.3 Advantages of using mtDNA

MtDNA has several unique properties which have been exploited in evolutionary studies:

1. It has a high copy number per cell with an average

cell containing approximately 10 000 mitochondria, each with about 5 to 10 mtDNA molecules. The small size and location of mtDNA within a cytoplasmic organelle, makes it more accessible to isolation and purification.

2. The complete mtDNA nucleotide sequence of one individual has been determined by Anderson *et al.* (1981) and is referred to as the reference sequence. All new mutations may be mapped to a specific position on the reference sequence either by restriction enzyme mapping or DNA sequencing. In this way the exact nature of the mutation and often its consequence (if in a coding region) can be established, and this makes it possible to examine the evolutionary history of different regions of the mtDNA genome (Cann *et al.* 1984).

3. Perhaps the most useful property of mtDNA is its strict maternal inheritance and in the absence of recombination (Giles *et al.* 1980; Zuckerman *et al.* 1984) this means that phylogenetic trees that show the relationship of all mtDNA types, may be interpreted as genealogies which reflect the maternal history of the species.

4. MtDNA accumulates mutations at a rate approximately 5 to 10 times faster than nuclear DNA. It has been estimated that the average rate of mtDNA sequence divergence is about 2 to 4% per million years compared with 0.5% per million years for nuclear DNA (Brown *et*

al. 1979; 1982; Wilson *et al*. 1985, 1987; Stoneking 1986). In this way population differentiation is accentuated facilitating the identification of subtle differences both within and between groups, making mtDNA extremely useful in evolutionary studies.

1.4 Organization of human mtDNA

Human mtDNA is a closed circular molecule which consists of 16 596 base pairs (bp). It is made up of two complementary single DNA strands that differ in their guanine and thymine content (G + T) and can be separated by alkaline caesium chloride gradients into the heavy (H) and light (L) strands (Brown 1980). The genetic information in human mtDNA is organized in a highly economical way composed primarily of coding sequences (approximately 90%) with no introns. Repetitive sequences are confined to the "control" or noncoding region and intergenic sequences are rarely more than 10 bp in length.

The mtDNA molecule encodes two ribosomal RNAs (12S and 16S rRNA) and twenty-two transfer RNAs (tRNA) which together function to decode the endogenous message, and thirteen proteins which function in electron transport and oxidative phosphorylation. The protein coding genes encode the cytochrome C oxidase subunits I, II and III, ATPase subunit 6, cytochrome b, seven genes encoding subunits of the dehydrogenase complex (ND1-4, 4L, 5 and 6), and one gene

(A6L) which codes for the subunit of the ATP synthetase complex (Fig. 1.1) (Anderson *et al.* 1981; Chomyn *et al.* 1985,1986).

The major non-coding region (control region) is bordered by the tRNA for proline on the one side and phenylalanine on the other (Fig. 1.2). The control region consists of 1112 bp and is believed to be involved in functions associated with transcription and replication (Clayton 1982; Attardi *et al.* 1985). Although mostly noncoding, eight segments of functional significance are contained within this region (Fig. 1.2). These include the origins of transcription of both the light (LSP) and heavy (HSP) strands, the termination associated sequence (TAS), binding sites of mitochondrial transcription factors (mtTF-L and mtTF-H) and three conserved sequence blocks (CSB-1, CSB-2, and CSB-3), apparently associated with the initiation of DNA synthesis (Walberg *et al.* 1981; Chang and Clayton 1984; Hixon and Clayton 1985). Although their locations may vary, these segments display a high degree of conservation in different species (Walberg *et al.* 1981; Hixson and Clayton 1985; Brown *et al.* 1986; King and Low 1987).

Replication of the complementary H- and L-strands occurs in an asymmetric, continuous, and unidirectional fashion, with the sites of initiation of both the H- and L- strands located at different regions in the molecule (see Fig. 1.1). Heavy strand replication begins at a site (O_H) in

the control region with the formation of the displacement loop (D-loop) by the synthesis of a short piece of DNA (7S DNA). This product displaces the parent or homologous H-strand, and new H-strand synthesis proceeds away from the rRNA genes. The origin of L-strand synthesis (O_L) is located approximately 5 700 bp away from the origin of the H-strand synthesis (see Fig. 1.1). Light strand synthesis does not occur until H-strand synthesis reaches the L-strand origin, and then proceeds in the opposite direction.

Although the mitochondrial genome replicates autonomously and encodes several protein subunits used either in electron transport or ATP synthesis, most proteins required for mitochondrial metabolism, replication, transcription, and translation are encoded by DNA in the nucleus. The majority of mitochondrial proteins are synthesised on cytoplasmic ribosomes and need to be transported into the mitochondrion to initiate their function. This could be to either of its two membranes or inner and outer spaces. Thus, mechanisms related to uptake of these proteins by the outer membrane and subsequent internalization, all nuclear-directed functions, are extremely crucial to the proper functioning of mitochondria. Failure leads to varying degrees of mitochondrial-associated diseases (Wallace 1989; 1992).

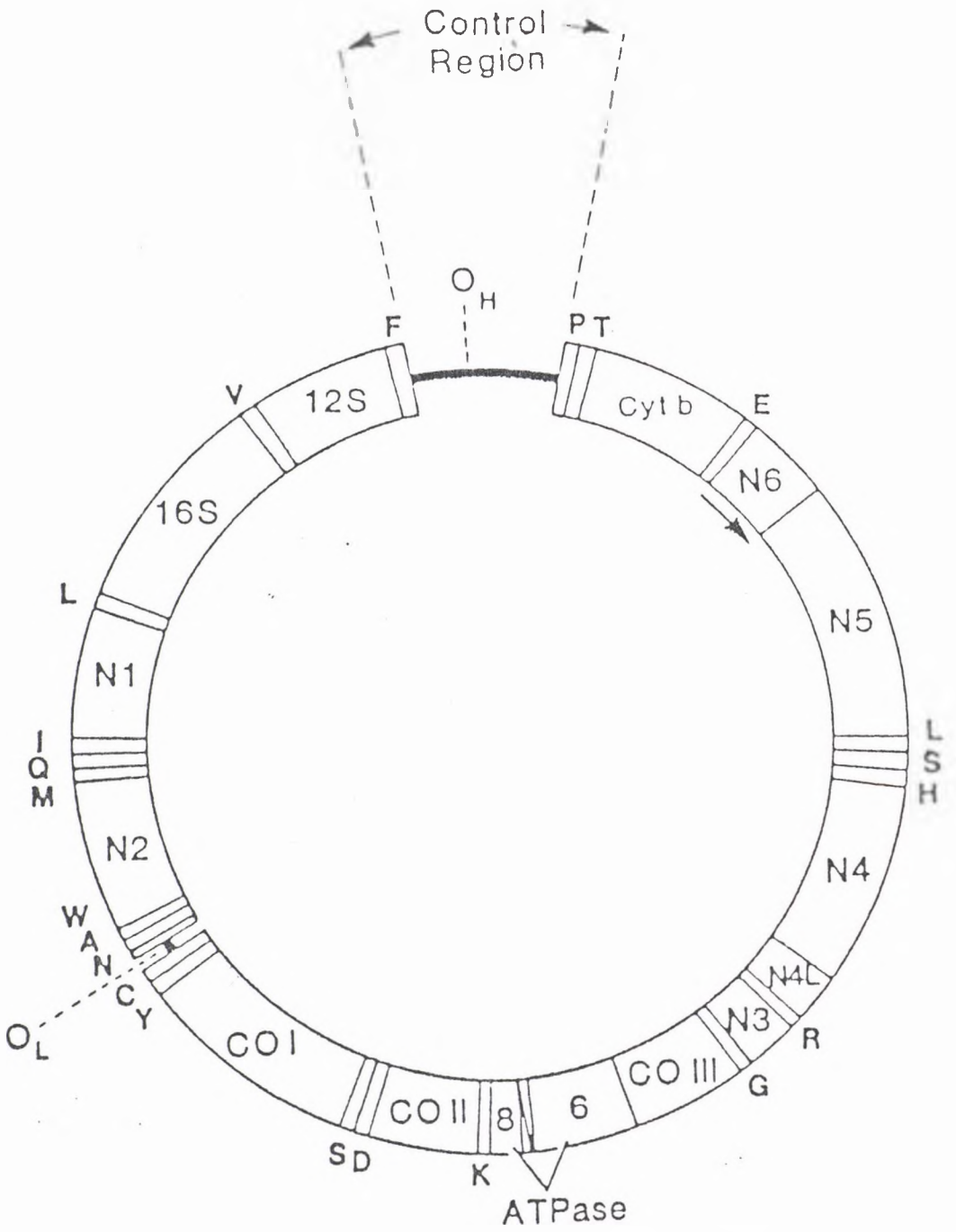
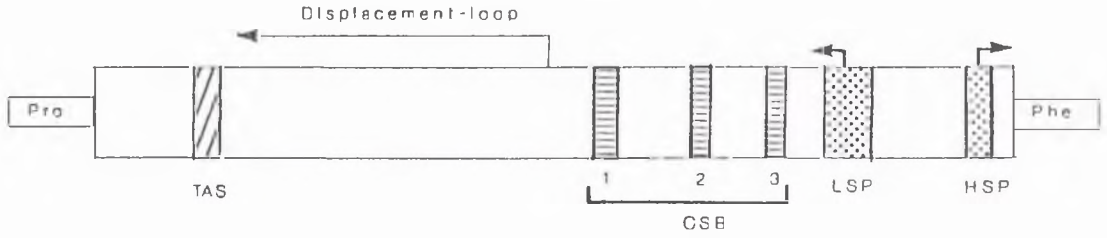


Fig. 1.1. The organization of the human mtDNA genome. Abbreviations for the genes are as follows: two ribosomal RNA genes (12S and 16S); seven genes for NADH dehydrogenase subunits (N1-N6, N4L); three genes for cytochrome oxidase subunits (COI-COIII); two genes for ATPase subunits (6 and 8); the gene for cytochrome b (Cyt b); and single letter codes for the 22 transfer RNA genes. The origins of replication of the heavy (O_H) and light (O_L) strands are designated within the solid bars representing noncoding regions. The large noncoding region known as the control region is labelled. (Taken from Vigilant 1990).



Pro tRNA-

TTCTTTCATG GGGAAGCAGA TTTGGGTACC ACCCAAGTAT TGACTCACCC
ATCAACAACC GCTATGTATT TCGTACATTA CTGCCAGCCA CCATGAATAT
TGTACGGTAC CATAAATACT TGACCACCTG TAGTACATAA AAACCCAATC
CACATCAAAA CCCCCTCCCC ATGCTTACAA GCAAGTACAG CAATCAACCC
TCAACTATCA CACATCAACT GCAACTCCAA AGCCACCCCT CACCCACTAG
GATACCAACA AACCTACCCA CCCTTAACAG TACATAGTAC ATAAAGCCAT
TTACCGTACA TAGCACATTA CAGTCAAATC CCTTCTCGTC CCCATGGATG
ACCCCCCTCA GATAGGGGTC CCTTGACCAC CATCCTCCGT GAAATCAATA
TCCCGCACAA GAGTGCTACT CTCCTCGCTC CGGGCCCATA ACACCTTGGG
GTAGCTAAAG TGAACTGTAT CCGACATCTG GTTCCTACTT CAGGGTCATA
AAGCCTAAAT AGCCCACACG TTCCCCTTAA ATAAGACATC ACGATGGATC
ACAGGTCTAT CACCCTATTA ACCACTCACG GGAGCTCTCC ATGCATTTGG
TATTTTCGTC TGGGGGGTAT GCACGCGATA GCATTGCGAG ACGCTGGAGC
CGGAGCACCC TATGTCGCAG TATCTGTCTT TGATTCCTGC CTCATCCTAT
TATTTATCGC ACCTACGTTC AATATTACAG GCGAACATAC TTACTAAAGT
GTGTTAATTA ATTAATGCTT GTAGGACATA ATAATAACAA TTGAATGTCT
GCACAGCCAC TTTCCACACA GACATCATAA CAAAAAATTT CCACCAAACC
CCCCCTCCCC CGCTTCTGGC CACAGCACTT AAACACATCT CTGCCAAACC
CCAAAAACAA AGAACCCTAA CACCAGCCTA ACCAGATTTC AAATTTTATC
TTTTGGCGGT ATGCACTTTT AACAGTCACC CCCCAACTAA CACATTATTT
TCCCCTCCCA CTCCCATACT ACTAATCTCA TCAATACAAC CCCCGCCCAT
CCTACCCAGC ACACACACAC CGCTGCTAAC CCCATACCCC GAACCAACCA
AACCCCAAAG ACACCCCCCA CA - Phe tRNA

Fig. 1.2. Structure of the control region showing the relative locations of the functional segments. Underlined regions indicate the termination-associated sequence (TAS), the conserved sequence blocks (CBS 1-3) and the light and heavy strand transcription promotors (LSP and HSP). Small arrows indicate the directions of transcrition.

1.5 Methods for determining mtDNA variation

Several methods have been used to assess the extent of mtDNA variation in human populations. Usually, the source of the starting material determines which approach is adopted. The merits and limitations of each approach are considered below:

1.5.1 Low-resolution mapping

In this approach, the starting material is whole blood. Total genomic DNA which consists of nuclear DNA and mtDNA is extracted and digested with six-base enzymes that cleave the mtDNA infrequently (possibly at 4 to 5 sites). The digested fragments are then separated by agarose gel electrophoresis and then transferred to nylon membranes by the method of Southern (1975). These membranes are then hybridized with a radiolabelled mitochondrial DNA "probe" and the mitochondrial restriction patterns are then analyzed following autoradiography. Fragments smaller than 250 bp cannot be visualized by this method, since these fragments migrate faster during electrophoresis and migrate off the gel under the electrophoretic conditions used. Since only enzymes which cleave mtDNA infrequently are used, a fewer number of sites are investigated resulting in a small proportion of the mtDNA genome being investigated.

This method first introduced by Douglas Wallace's laboratory has been the method of choice for most researchers in this field, since blood samples are usually the source of material at hand (Johnson *et al.* 1983; Bonne-Tamir *et al.* 1986; Brega *et al.* 1986a,b; De-Benedictis *et al.* 1989a,b; Santachiara-Benerecetti *et al.* 1988, Sartoris *et al.* 1988; Scozzari *et al.* 1988; Torroni *et al.* 1990; Vikki *et al.* 1988).

1.5.2 High-resolution mapping

This method is so called because of its higher resolving power over the previous method. It makes use of pure mtDNA as its source for restriction fragment length polymorphism (RFLP) analysis. This means that mtDNA has to be purified from tissue rich in mitochondria, which has to be obtained by non-invasive procedures without harming the subject. The best source of mitochondria which satisfies these conditions has been placental tissue. Unfortunately, placental tissue is not readily available.

Tissue purified mtDNA can be digested with four-base enzymes which digest the mtDNA at many sites resulting in about 25 fragments, thereby making it possible to examine several nucleotide positions thereby increasing the resolving power of the technique. In addition, digested fragments are end-labelled and separated on polyacrylamide gels, permitting the visualization of smaller fragments.

This method enjoys the advantage over Southern blot analysis in that smaller quantities of DNA, 10 ng rather than 5 μg, can be used for analysis with restriction enzymes. Yet another advantage over Southern blot analysis is that the sensitivity of the high resolution mapping technique allows the discrimination of sequences that differ by as little as 0.05 percent compared to the 0.5 percent detected by the Southern blot method (Wilson *et al.* 1985).

This method first introduced by researchers in Alan Wilson's laboratory at Berkeley has since been widely used (Brown 1980; Cann and Wilson 1983; Cann *et al.* 1984, 1987; Stoneking 1986 a,b; Stoneking *et al.* 1990; Horai and Matsunuga 1986; Harihara *et al.* 1986, 1988; Horai *et al.* 1986).

1.5.3 DNA amplification and sequencing

The technological advances of the past decade have led to the introduction of more sophisticated molecular methods of analysis; the polymerase chain reaction (PCR) is probably the most useful and this permits the amplification of small amounts of DNA from different material sources such as root hairs, ancient bones, mummified tissue, forensic material such as blood stains, buccal cells from mouth washes, etc. (Saiki *et al.* 1988). This obviates the need for large quantities of blood and other tissue samples which were

previously required to obtain sufficient amounts of DNA for nuclear DNA and mtDNA analyses.

Mitochondrial DNA can be specifically amplified from a variety of sources (as mentioned above) and then sequenced to obtain an unambiguous nucleotide sequence of the region under investigation. In addition, the exact nature of the mutation can be determined; this is crucial in the development of methods for phylogenetic analyses. Sequence analysis of the two hypervariable regions contained within the non-coding control region has been investigated extensively (Wrishcnik *et al*. 1987; Vigilant *et al*. 1988, 1989, 1991; Horai and Hayasaka 1990; Vigilant 1990; Di Rienzo and Wilson 1991).

Another approach which makes use of PCR has been introduced by Wallace's group. In this method mtDNA is amplified using fourteen pairs of primers which independently amplify overlapping regions of the mtDNA molecule such that all 16 569 bp of the mtDNA molecule can be analysed. These PCR products are digested with various restriction enzymes and the digestion products examined by agarose gel electrophoresis (Schurr *et al*. 1990; Torroni *et al*. 1992). In this way many more nucleotide positions throughout the molecule may be investigated for variations by restriction mapping.

CHAPTER TWO:

SUBJECTS AND METHODS

2.1 Subjects

The sample consists of 795 individuals belonging to 20 ethnic groups distinguished on the basis of language, culture and their geographic location/origins in southern Africa (Fig. 2.1). Southern Africa is taken as consisting of the whole of the Republics of South Africa, Botswana and Namibia, and the Kingdoms of Lesotho and Swaziland and that part of Mozambique which lies south of the Limpopo River (Nurse *et al*. 1985).

2.1.1 Classification of populations

Race, physical type, language, and culture are at present convenient means of catergorizing people. However, prior to the immigration of Caucasoids into South Africa (S.A.), southern Africa was occupied exclusively by Khoisan and Negroid peoples who spoke Khoisan and Bantu-languages, respectively.

The term Khoisan was coined by Schultze (1928) to embody the "Hottentot" peoples who referred to themselves as Khoi and the "Bushmen" whom the Khoi referred to as San. Southern African Negroids speak Bantu languages belonging to the Benue-Niger family of languages within the Niger-

Congo group as defined by Greenberg (1963). According to Greenberg, the 800 and more languages spoken in Africa can be broadly categorized into four different groups: Afroasiatic, Sudanic, Niger-Congo (including Bantu) and Click (Khoisan). Westphal (1963) argues that the term Khoisan has *"no linguistic usefulness whatsoever"* and that the words "Khoisan" and "click" have been misused when usually implying non-Bantu languages. He prefers the use of the term "Bush" languages to refer to the dialects spoken by the "Bushmen" people and the use of the term "Hottentot" to refer to the Nama and Dama languages (see section 5.1.2).

For the purpose of the present study, the author has used the classification suggested by Greenberg for the major groups and a combination of the classification used by Westphal (1963) and Doke (1954) to sub-classify the Bantu-speaking groups (Table 2.1). In addition, the suggested nomenclature of San for "Bushmen", Nama for "Hottentot" or Khoikhoi, Bantu implying language and Negroid for population group as suggested by Jenkins and Tobias (1977) has been adopted. A comprehensive table with information on sample identity, sex, language group (corresponding to that of the subjects mother) and the location and date at which samples were obtained, is given in Appendix 1.

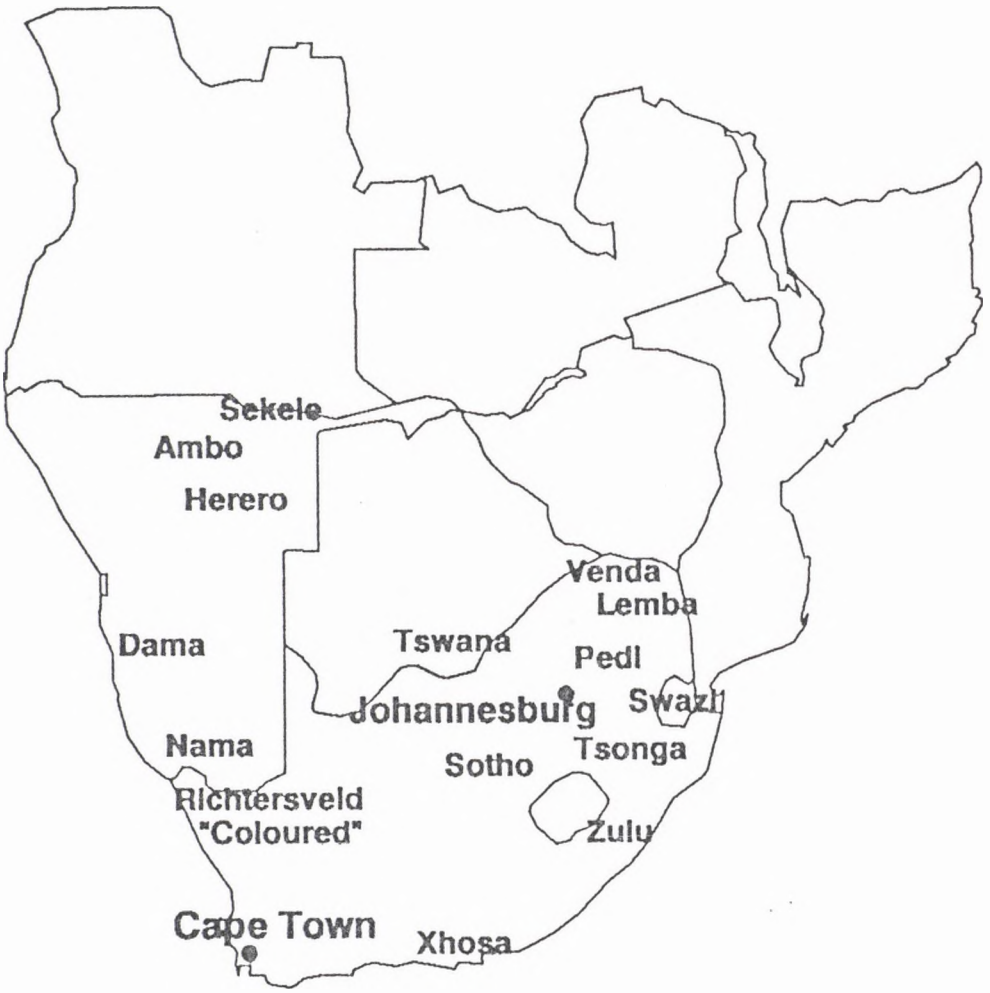


Fig. 2.1. Map of southern Africa showing the locations of the different ethnic groups investigated for mtDNA variation in this study. The English and Afrikaans-speaking "Whites", Ashkenazi Jewish, Johannesburg "Coloured" and Indian populations were obtained from Johannesburg (these populations are not shown in the map).

Table 2.1. The sample sizes and ethnicities of southern African populations investigated in this study.

POPULATION		SAMPLE SIZE	
BANTU-SPEAKING NEGROIDS*			349
NGUNI		89	
Zulu	30		
Swazi	41		
Xhosa	18		
TSONGA		34	
SOTHO/TSWANA		67	
Sotho	29		
Pedi	23		
Tswana	15		
VENDA		30	
LEMBA		53	
HERERO		54	
AMBO		22	
KHOISAN-SPEAKING NEGROIDS**			43
DAMA		43	
KHOISAN			105
SAN (Sekele)		59	
KHOI (Nama)		46	
CAUCASOID			192
ENGLISH-SPEAKING "WHITES"		36	
AFRIKAAN-SPEAKING "WHITES"		51	
ASHKENAZI JEWISH		55	
INDIAN		50	
HYBRID PEOPLES			106
RICHTERSVELD "COLOURED"		35	
JOHANNESBURG "COLOURED"		71	
TOTAL			**795**

*Populations were classified into ethnic groups using their linguistic affiliations as described by Westphal (1963) and Doke (1954).

**This category was suggested by Nurse et al. (1976)

2.1.2 Bantu-speaking Negroids

It has been suggested that Bantu languages originated in the Cameroon region in West Africa approximately 3 000 years ago and then spread to other parts of Africa (Greenberg 1963). In southern Africa two major Bantu linguistic divisions are found: those which are derived from southeastern and those from southwestern groups, respectively. Linguistic (Ehret 1973, 1982a,b; Vansina 1984) and archaeological (Phillipson 1977, 1985; Huffman 1982, 1989) data supporting these divisions are expanded in Chapter 5.

It is hoped that by studying southern African populations at a genetic level, some of the theories on the migration and spread of Bantu-speakers (Negroids) may be refined. The correlation between linguistic, archaeological, historical and genetic data in studies on southern African population prehistory will be addressed in detail in Chapter 5, but for the present, an introduction and discussion of some of the linguistic and cultural affiliations of the groups used in this study, will be presented in the next few sections.

2.1.2.1 Southeastern Bantu-speakers

There are three principal groups of Negroids who comprise this division of Bantu-speakers: the Nguni, the

Sotho/Tswana and the Tsonga. Although the Nguni people share some common cultural traits, their classification as Nguni is based on a broad linguistic uniformity. There are many local differences of dialect, but the dialects are closely related. At present, the Nguni peoples occupy an area ranging from present-day Ciskei to Swaziland. For purposes of convenience the Nguni group has been divided into three ethnic categories: Zulu, Swazi and Xhosa (Doke 1954; Westphal 1963).

The economic base of the Nguni centres around herding, cultivation, hunting, manufacture of ornaments, pottery and making of clothing. The fundamental unit of Nguni society is the lineage which comprises all the descendants of a common male ancestor. The lineages claiming descent from a common male ancestor constitute a clan. The exogamy rule applies with respect to clans. The chiefdom represents the main political unit, with the chief being a pivotal figure who has military, judicial and religious powers (Maylam 1986).

"Sotho/Tswana" is a broad generic term used to designate a large group of people who display linguistic and some cultural similarities. They occupy the interior plateau of South Africa, and are also conveniently subdivided into three main chiefdoms (Doke 1954): the western Sotho or Tswana group made up collectively of the Tlharo, Tlhaping, Fokeng, Kgatla, Koena, Kgalagadi and Tlharo groups; the

northern Sotho or Pedi group which comprise the Pedi, Tawana and the Ngwato groups; and the southern Sotho (Basotho) who occupy present-day Lesotho and adjacent areas (Nurse *et al.* 1985).

Three major differences between the Nguni and Sotho/Tswana group may be identified: firstly, although both are Bantu-speakers, their respective languages are distinctly separate components of the greater Bantu-language group. Secondly, their settlement patterns have differed enormously. Whereas the Nguni settlement has tended to be widely dispersed in scattered homesteads and villages (a pattern which has been called a "checkerboard realm"), Sotho have lived in more concentrated settlements ("tribal estate"), far removed from their agricultural, pastoral and hunting grounds (Sansom 1974). The third major difference centres around kinship; with the Sotho practising a system of preferred cousin marriages whereas the Nguni clans have been exogamous.

The Tsonga comprise the third main division of the southeastern Bantu-speakers following Doke's (1954) classification. Bryant (1929) is of the opinion that there has been a considerable amount of admixture between the Tsonga and the Nguni who cannot be distinguished with any degree of certainty. Tsonga have presently settled on the escarpment of the Drakensberg and penetrated to the plateau intermixing there with the Sotho inhabitants (Nurse *et al.*

1985).

2.1.2.2 Venda and Lemba

The Venda and Lemba occupy an area in the far north of the Transvaal Province of South Africa, with the Venda presently inhabiting the fertile Zoutpansberg mountains (Maylam 1986). Until recently, it was commonly believed that the "true" Venda were Shona immigrants who established themselves over the original non-Venda inhabitants, the Mbedzi and Ngona. Recent archaeological and linguistic research have failed to confirm this theory Beach (1980).

The Lemba who have lived among the Venda and Shona have displayed certain "peculiar" characteristics that have led scholars to set them apart from the Nguni and Sotho/Tswana peoples. Their economic base is unusual and depends more on manufacturing and trading than on cultivation and herding. They have not constituted any particular kind of socio-political entity, preferring to live dispersed among the Shona, Venda and Pedi. Although they generally speak the language of those among whom they reside, they tend otherwise to keep apart from the local inhabitants and practise strict endogamy. These "peculiarities" have raised several confusing theories as to the origins of the Lemba.

It has long been maintained that the Lemba were descendants of Muslim traders (Stayt 1931; Van Warmelo 1974). Lemba follow the Muslim traditions of male circumcision and not eating pork, and texts of Lemba prayers have revealed a strong Muslim influence. Parsons (1982) has doubted the Muslim origins, pointing out that male circumcision and the pork taboo are "*also found among neighbouring peoples*". The most likely explanation has been put forward by Beach (1980) who suggests that the Shona-speaking Muslim traders of the interior plateau were in decline in the seventeenth century and that "*the Muslim community was gradually becoming more and more absorbed into the Shona world, to become the Lemba groups scattered across the Plateau - groups that retained little more than fragments of the Islamic faith and culture*".

2.1.2.3 Southwestern Bantu-speakers

The other major group of southern Bantu-speakers consists of Negroids who speak southwestern Bantu languages. These include the Herero, Ambo and Kavango of Namibia. Herero constitute approximately 6.8% of the total population (estimated to be 1 million) of Namibia (Malan 1980). The nomadic movements of the Herero led them to many parts of the country where they clashed with other groups, primarily due to conflicts over grazing for their cattle. This resulted in their dispersion over a wide area stretching from southwestern Angola through Kaokoland and the central

parts of Namibia to Ngamiland in northern Botswana.

The Herero trace descent bilaterally, that is, through their paternal (*otuzo*) and through their maternal (*omaanda*) lines. According to oral tradition, all Herero descended from the daughters of one mother and it has been suggested that a maximum of eight *omaanda* (pl. of *eanda*) may be found among the Herero of Botswana (Pennington 1990).

The Ambo chiefdoms belong to the southwestern Bantu group, but are culturally closely related to the matrilineal agriculturalists of Central Africa. The present Ambo tribal divisions in Namibia are: Kwanyama (36.6%), Ndonga (28.7%), Kwambi (11.8%), Ngandjera (7.7%), Mbalantu (7.4%), Kwaluudhi (5.1%) and Eunda and Nkolonkadhi (2.7%) (Nurse *et al*. 1987). According to legend, the Ambo and Herero are descendants of two brothers Nangombe and Kathu, respectively. Kathu and his followers separated from his brother in search of suitable pastures for their large herds of cattle to form what is believed to be the Herero group, whereas Nangombe and his followers (Ambo) settled in the fertile plains of the Ovambo territory to tend their crops (Hahn 1928).

The Ambo and the Kavango people who live south of the Kavango River together represent the most southerly south-western extension of Central African Negroes practising matrilineality (see Nurse and Jenkins 1977 for details).

2.1.3 Khoisan speakers

It is generally accepted that southern Africa was occupied exclusively by Khoisan peoples prior to the appearance of Bantu-speaking Negroids in southern Africa (Ehret 1982a,b). Although the San and Khoikhoi are not always distinguishable in physical appearance, certain cultural and linguistic differences are apparent. The San are hunter-gatherers and keep no animals for their own use whereas the Khoikhoi are (or were until recently) pastoralists who herd large flocks of sheep and cattle (Vedder 1928; Ehret 1982a,b). Prehistoric and recent contact between Khoisan and Negroid populations have resulted in several cultural and linguistic exchanges between them (Westphal 1963; de Almeida 1965; Ehret 1982a,b), making the study of the origins of the Khoisan difficult.

The sample of 95 Khoisan included here for mtDNA studies include a Khoi sample (46) of Nama-speakers from Namibia who will hitherto be referred to as Nama. The San group investigated consists of a sample of 49 Sekele ("Yellow Bushmen") and are compared to a !Kung group (34) from Botswana speaking a related language (Johnson *et al.* 1983). A small group of Kwengo ("Black Bushmen") who resided close to the Omega Military base which was once situated near the Angola-Namibia border are included for comparison and to establish their ethnographic status.

2.1.4 Khoisan-speaking Negroids

The Dama constitute the second largest group in Namibia (8.5%) and are morphologically Negroid but speak the same language as the Nama (Nurse *et al.* 1976). The name Dama is derived from the Hottentot word *Daman* used by the "Hottentots" to refer to these *dark-skinned* people, who have been known by a variety of names, the two most common being Damara and Bergdama (Vedder 1928; Nurse *et al.* 1976). There is no consensus in the literature concerning the origins of the Dama and it was hoped that mtDNA studies on this group of Negroids would shed some light on their origins.

2.1.5 Caucasoid groups

2.1.5.1 European immigrants

In the nearly 350 years since the settling of southern Africa by the Dutch, South Africa has been host to several immigrant Caucasoid groups from western Europe and the United Kingdom. Consequently, the genetic pool of present-day "White" South Africans, consists of a multitude of ethnic groups whose origins include Dutch, English, German, Greek, Portuguese, Jewish, Italian, French and to a lesser extent, individuals from other European countries. Descendants of the original settlers as well as of later immigrants, have been incorporated into either the English-

speaking community or the Afrikaans-speaking community of South Africa. Since both English and Afrikaans are official languages in South Africa, Afrikaners are distinguished from the other "White" immigrants studied to show descent from the original Dutch settlers and later the French Huguenots, whereas English-speakers were those individuals whose ancestry was traced to any other European country and the United Kingdom. In most cases this information was obtained from the subject donating blood samples.

Ashkenazi Jewish immigration to South Africa occurred mainly between 1880 and 1910, from Lithuania, Latvia, Poland and Russia, and later in the 1930s from Germany (Saron 1965). A sample of unrelated Ashkenazi Jews living in the Johannesburg area was included in this study.

2.1.5.2 Asian Indians

Indentured Indians, primarily from Madras, Calcutta and Bombay, were brought to South Africa between 1860 and 1910 to tend the sugar-cane plantations in Natal. In addition, merchants from the districts of Kathiawar, Surat, Porbander and from the northern Indian provinces, arrived in Natal as normal fare-paying British subjects, and are usually referred to as "passenger" Indians (Bhana and Brain 1990). About 90% of the indentured Indians were Hindus whereas the majority of "passenger" Indians were Moslems.

2.1.6 Hybrid communities

The South African "Coloured" group according to Marais (1939), "*is the history of the contact of aboriginal Africans (and a few Orientals) with Europeans. Without this contact there would have been no Coloured folk*". From his interpretation of government records, missionaries' letters, census figures and interviews with travellers, Marais concluded that the "Coloured" people of today are derived from four "elements": Caucasoid, Negroid, Khoisan and slaves. The "slave strain" were imported from the west African coast in 1658 and in subsequent years came from Madagascar, Mozambique and the East (India, Ceylon and the Malay Archipelago). Several different "Hybrid communities" may be identified.

Those in whom the female gene contribution has been derived from a diverse ancestry (Khoisan, Negroid, and to a lesser extent from Indonesian), are referred to as the non-Malay Cape "Coloured", and they constitute the majority of the Johannesburg community. The Cape Malay are a very distinct group - endogamous Sunni Muslims descended from Bengali or Malay-speaking slaves or political exiles from Malaysia. Other groups of hybrid communities are descended primarily from Khoi females and include the Rehoboth Basters in Namibia, other Baster groups from Botswana and the northern Cape, and the Richtersveld "Coloured" group also from the northern Cape. Only two of these groups: a Johannesburg

"Coloured" and a Richtersveld "Coloured" group have been investigated in this study.

2.2 Methods

2.2.1 Reagents and recipes

All chemicals, solutions, media and recipes used in this section are described in Appendix 2.

2.2.2 Collection of blood

Blood samples were obtained from several locations as a result of a series of organized departmental field trips over the years. In addition, the author once again wishes to acknowledge the collaboration of Highveld Blood Transfusion Service for their assistance in the collection of blood from rural areas in South Africa, Windhoek State Laboratories for their assistance in collection of blood samples from the Nama, Dama and Ambo populations from Namibia, and Dr F. Meyer of the South African Defence Force for his participation in collecting blood from the San at the Omega Base at the Angola/Namibia Border.

Between 30-50 ml of blood was drawn from individuals who consented to participate in the study. In the case of minors, parental consent was first obtained. Whole blood was collected into tubes containing ACD solution and

transported at 4°C in cooler boxes containing ice packs so that it reached the laboratory in good condition within two days of collection.

In the laboratory, whole blood was separated into its components by centrifugation in a bench top Beckman Model TJ6 centrifuge at 2 700 rpm for 10 min. Plasma and buffy coat were collected into 5 ml sterile plastic tubes, whereas red blood cells were preserved in a solution containing glycerol (see Appendix 2). Buffy coats were stored at -20°C until required. Samples were coded using alphabetical abbreviations to denote population and a binomial numeric system to denote individual number (see Appendix 1).

2.2.3 Genomic DNA extraction

Buffy coats were stored at -20°C for at least 24 hr before DNA extraction. Total genomic DNA was extracted using the method described by Sykes (1983). Frozen buffy coats were allowed to thaw at room temperature and then transferred to 50 ml Nunc tubes. Approximately 2.5 volumes of cold Triton-X/NaCl solution (see Appendix 2) was added to each tube and mixed thoroughly by vigorous shaking to aid lysis of red blood cells trapped in the buffy coat during separation. The white blood cells were then packed by centrifugation at 2 700 rpm in a bench top centrifuge for 15 min at room temperature. Following centrifugation, the

supernatant was poured off and discarded. The pellet was washed at least twice with 30 ml Triton-X/NaCl (see Appendix 2) solution followed by centrifugation as before, until the resultant pellet was free of red blood cells. This pellet was homogenized with a sterilized glass rod in the presence of lysing buffer (Appendix 2), which was added dropwise to a final volume of 10 ml. The homogenized cells were then completely lysed by the addition of 2 ml of a 10% SDS solution followed by incubation at 37°C for 10 min. DNA thus released was retrieved from solution following phenol-chloroform extractions and ethanol precipitation as described by Maniatis *et al.* (1982) and then dissolved in 0.5-1.0 ml of TE (pH 8.0).

2.2.4 DNA digestion and electrophoresis

DNA extracted in the above consists of nuclear DNA and mtDNA. Digestion of total genomic DNA to completion ensures complete digestion of mtDNA. Details of enzymes and their respective gel concentrations are summarized in Table 2.2. Approximately 5 μg of total DNA was digested with the enzymes listed in Table 2.2 following the manufacturers' recommended conditions in a total volume of 50 μl. The reaction was carried out by incubation of the reaction mix at 37°C in a regulated temperature controlled incubator for a minimum of 10 hr.

Table 2.2. The specifications of the restriction enzymes used for mtDNA RFLP analysis. The recognition sequence of the enzymes as well as the number of sites on the reference sequence (N) are given for each restriction enzyme used.

Enzyme	Recognition Sequence	N	Supplier	Agarose gel concentration (%)
HpaI	GTT/ACC	3	Promega	0.7 or 0.8
BamHI	G/GATCC	1	Anglian	0.7
HaeII	(AG)/GCGC/(CT)	7	Promega	0.8
Msp	C/CGG	23	Promega	1.2
AvaII	G/G(AT)CC	8	Anglian	1
HincII	GT(CT)/AGAC	8	Amersham	1.2

Agarose gels were prepared by dissolving the appropriate amount of Seakem HGT agarose (see Table 2.2 for the different gel concentrations used) in 1X TBE buffer (Appendix 2) by heating until boiling, with intermittent shaking. For every 100 ml of cooled dissolved agarose, 3 μl of ethidium bromide (10 mg/ml) was added. This mixture was then poured into a gel mould (20 X 18 cm) prepared as described by Maniatis *et al*. (1980). A sample (5 μl) of the digestion mix was checked on an agarose gel to ensure that the genomic DNA was completely digested. When completely digested, these DNA samples (45 μl) were mixed with 5 μl of loading dye (Appendix 1) and loaded into the wells of the gel. Lambda DNA markers were loaded on either side of the reaction samples to assist in the determination of fragment sizes of digested products. Electrophoresis was performed at 40 mA for 14-16 hr or until the bromophenol blue dye reached the end of the gel.

2.2.5 Southern blotting

Following electrophoresis, the gels were placed on a UV transilluminator and photographed using Polaroid type 667 film. Gels were transferred to tupperware dishes and rinsed with distilled water to remove traces of electrophoresis buffer. Since electrophoresed products were transferred to two different types of nylon membranes (Biodyne and Hybond-N), transfer conditions were varied to meet the requirements suggested by the manufacturers.

When Biodyne nylon membranes were used, DNA was first denatured in a 1 M NaCl/0.5 M NaOH solution for 30 min with gentle shaking at room temperature. This solution was poured off and discarded and the gel rinsed twice with distilled water and then neutralized in a 3 M sodium acetate (pH 5.5) solution for 30 min. After neutralization, gels were rinsed twice with distilled water before soaking in 10X SSC solution (Appendix 2) for 10 min.

When Hybond-N membranes were used, denaturing was as before. Gels were neutralized in a Tris-NaCl buffer (1.5 M NaCl, 0.5 M Tris-HCl pH 7.2, 1 mM EDTA) for 30 min. The washes with distilled water between steps were also done and the gels were finally soaked in 20X SSC for at least 10 min before blotting.

The transfer of DNA from the gel to the membrane was accomplished by the established Southern blot method (Southern 1975). Transfer by capillary action (see Maniatis *et al*. 1982) was allowed to proceed for a minimum of 6 hr, although this was usually done overnight for convenience.

Subsequent to transfer, the absorbent paper towels and filter papers covering the membrane were removed and the wells of the gel were marked on the membrane using a waterproof permanent marker pen. Additional details pertaining to enzyme used, sample identity, gel

concentration, type of membrane and date were recorded at the bottom of the membrane. These membranes were then rinsed in 2X SSC solution to remove gel debris or any SSC solution that had crystallized, air dried for about 15 min and then baked for 1 hr at 80°C under vacuum to fix the DNA to the membrane. Baked blots were cooled at room temperature and then sealed in plastic bags between filter sheets until required for hybridization.

2.2.6 MtDNA extraction and purification

Mitochondria were isolated from placental tissue obtained with informed consent from a Caucasoid donor. The freshly obtained placental sample was stored on ice in a cooler box and transported to the laboratory within an hour after obtaining it. The fibrous tissue was first removed and approximately 50 g pieces of the tissue were washed in 0.9% saline solution, sealed in plastic bags and stored at -70°C until required.

Mitochondria were purified and mtDNA extracted and purified following a modified procedure of Cann (1982). Frozen placenta was partially thawed on ice and chopped into fine pieces on a glass plate placed in a tray filled with ice. The chopped pieces were transferred to a petri dish and minced into a pulp with cold extraction buffer (Appendix 2) which was added dropwise to the tissue. When thoroughly minced the pulverized mixture was transferred to a glass

homogenizer (Ultraturox) and the cells were homogenized with extraction buffer. It was crucial to perform all of these steps on ice to support the viability of the mitochondria.

When completely homogenized (even slurry), the cells were precipitated by centrifugation at 3 800 rpm for 10 min at 4°C in a Sorval centrifuge. The supernatant was poured off and the centrifugation repeated to retrieve unpelleted cells and other cellular debris. The cell free supernatant was transferred to a 50 ml Beckman centrifuge tube and centrifuged at 10 000 rpm for 20 min to pellet the mitochondria. The supernatants were pooled and the centrifugation repeated to recover residual mitochondria. The resultant pellet was resuspended in MS buffer (Appendix 2) and centrifuged as before.

The pellet thus obtained was resuspended in 3 ml of TE (pH 7.4). Mitochondria were lysed by the addition of 300 μl of 10% SDS and incubated at room temperature for 15 min. The equivalent of 1 M CsCl was added to the lysed cells, mixed, and incubated on ice for 15 min. The lysed mitochondrial debris was collected by centrifugation at 10 000 rpm for 15 min in ultra-centrifuge tubes spun in a Beckman Ty65 rotor. The mtDNA which was present in the supernatant was purified by CsCl-ethidium bromide centrifugation as follows: 0.72g of CsCl and 80 μl of 10 mg/ml ethidium bromide were added per ml of DNA solution. This mix was adjusted to the

correct density such that 1 ml of solution weighed between 1.55-1.57 g. When this density was obtained either by the addition of more CsCl or TE, the balanced tubes were centrifuged at 50 000 rpm for 16 hr at 20°C in a Beckman ultracentrifuge using the Ty65 rotor.

Following ultracentrifugation, the mtDNA band which was visualized against a UV light source was removed from the gradient with a 1 ml syringe and a 22 gauge needle. The ethidium bromide was removed from the DNA by isoamyl alcohol extractions until there were no remaining trace of it. The DNA was dialysed in TE in a petri dish using Millipore minidialysis filters with a pore size of 0.0025 μm for 1 hr to remove CsCl.

2.2.7 Radiolabelling of mtDNA

Purified mtDNA was radiolabelled with [^{32}P]-dCTP by random primed hexanucleotide oligolabelling (Feinberg and Vogelstein 1983) as follows: 10 μl of mtDNA was aliquoted into an Eppendorf tube and heated to boiling in a preheated waterbath for 5-7 min. This mix was immediately cooled by incubating on ice for 1 min to prevent the reannealing of single stranded templates. All solutions for labelling were provided in a multiprime labelling kit (Amersham). A 50 μl reaction mix resulted from the addition of 10 μl labelling buffer (5X), 5 μl hexanucleotide primer mix, 5μl [^{32}P]-dCTP, 1 μl spermidine (0.1 M), 17 μl distilled water

and 2 μl Klenow polymerase enzyme to the 10 μl of mtDNA in the Eppendorf tube. The reaction was allowed to proceed at room temperature for at least 1 hr.

The extent of incorporation was monitored at 15 min intervals by spotting 1 μl of reaction mix onto Millipore filter discs followed by washing with TCA solution and checking the counts with a Geiger counter. Filter discs were first washed with a 10% TCA solution (three times) followed by a 5% TCA solution (three times) to separate the incorporated material (remaining on filter) from the unincorporated, using a Millipore filtration apparatus connected to a vacuum pump. When the reaction was complete, *i.e.*, when the counts of the spotted sample before washing were comparable to the counts after washing, the reaction tube was placed on ice to stop the reaction.

To this mix 150 μl of a 100 μg/ml salmon sperm DNA was added to increase the volume of the labelled mix for phenol/chloroform purifications. The aqueous phase containing the labelled DNA was collected into a sterile Eppendorf tube. An additional 200 μl of salmon sperm DNA was added to the organic phase (phenol) and the centrifugation repeated to recover any labelled DNA lost at the interface. This aqueous phase was pooled with the first phase, mixed with a few crystals of acridine orange and loaded onto a packed Sephadex G_{50} column (equilibrated with 1X TE solution). The eluate was fraction collected

into Eppendorf tubes. Tubes containing the radiolabelled peak (determined by checking each fraction with the Geiger counter) were pooled and the amount of radioactive incorporation measured by scintillation counting of a 10 μl aliquot of this pooled eluate. The remaining fraction was stored at -20°C until required for use as a "probe" in hybridization reactions.

2.2.8 Hybridization and autoradiography

Hybridization conditions differed for Biodyne and Hybond-N membranes as follow:

2.2.8.1 Biodyne membranes

(a) Prehybridization

Blots were soaked in 2X SSC solution for 10 min to hydrate them. For bag hybridizations, an appropriate length of plastic roll was cut and 3 blots were aligned one on top of another and the bag was sealed at three ends. Excess fluid was squeezed out from the open end by rolling a pipette over the top of the bag. Thirty ml of 1X prehybridization mix (Appendix 2) was added to the bag from the open end and distributed evenly over the blots. Bubbles were removed and the bag was then sealed at the open end. The blots were incubated in prehybridization solution in a 42°C incubator for a minimum of 6 hr with shaking.

For hybridization in tupperware dishes, a maximum of 20 blots were used at a time and prehybridized in 150 ml of prehybridization mix.

(b) Hybridization

The bag was cut at one corner and the prehybridization mixture removed with a pipette. A 2X hybridization mix (Appendix 2) was mixed with an equal volume of deionized formamide and 20 ml of this mix was added to the bag. Appropriate volumes of the mtDNA probe and labelled λDNA (10^8 cpm/blot) were mixed together in an Eppendorf tube and heat denatured at 70°C for 5 min, cooled rapidly on ice and then added to the hybridization mix in the bag. This mix was distributed evenly over all the blots and the corner sealed again. These blots were incubated at 42°C for 18-24 hr with shaking. Tupperware-dish hybridizations were the same except that 100 ml of hybridization mix was used.

(c) Post-hybridization washes

The hybridization mixes were recovered and stored in 50 ml Nunc tubes for reuse. With bag hybridizations the blots were removed from the plastic bag and transferred to tupperware dishes. The washes were as follows: three rinses with 2X SSC solution for 5 min each at room temperature, two washes with 2X SSC solution for 15 min each at room temperature, one wash with 0.1X SSC + 0.1% SDS

at 42°C for 30 min, two washes at 65°C with 0.1X SSC + 0.1% SDS for 30 min each, and finally three rinses with 0.1X SSC solution.

2.2.8.2 Hybond-N membranes

(a) Prehybridization

The blots were treated in the same way for bag and tub hybridizations as for the Biodyne membranes except that the mixes were different (see Appendix 2). However, only 1 hr incubation at 42°C was needed.

(b) Hybridization

Denatured labelled mtDNA probe and λDNA used at the same concentrations given for Biodyne membranes were added directly to the prehybridization mix. Incubation conditions were the same as those used for Biodyne membranes.

(c) Post-hybridization washes

Hybridization mix was collected and stored at -20°C for reuse. The blots were first rinsed three times with 2X SSPE + 0.1%SDS solution at room temperature followed by two washes with the same solution for 15 min each at room temperature. The next wash involved incubation at 65°C in

a 1X SSPE + 0.1% SDS solution for 30 min. The final wash was with 0.1% SSPE (three rinses) at room temperature.

2.2.8.3 Reuse of hybridization mixes

Hybridization mixes (containing labelled DNA) which were stored at -20°C were heated at 70°C for 10 min and then cooled on ice before adding to prehybridized blots. Successful hybridization was obtained with mixes that were used for a maximum of three times.

2.2.8.4 Autoradiography

Following washing, individual blots were sealed in plastic bags and placed in cassettes (Dupont) with intensifying screens and X-Ray film (either 3M X-Ray or Kodak X-Omat RP) and exposed at -70°C. The duration of exposure varied from 18-24 hr followed by longer exposures for 3-5 days to ensure that all weak bands would appear.

2.2.9 RFLP analysis

The size of each mtDNA restriction enzyme fragment could be estimated by comparing its mobility relative to the λDNA markers used. Fragments smaller than 500 bp could not be resolved on these gels, an obvious weakness of this approach. Restriction fragment length patterns (morphs) were compared to published data, and when a pattern

corresponding to that previously published was observed, the morph number corresponding to that pattern was recorded for the individual in whom it was found. When an unpublished pattern was observed, a new number, maintaining the chronology in number assignment, was given for that pattern. Unpublished mutations were mapped using the published sequence as a reference (Anderson *et al.* 1981). The restriction pattern observed for each individual with the 6 enzymes used were recorded and the combined morphs with the enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II (in this order) were used to derive the mtDNA type (mitotype) for that particular individual (Blanc *et al.* 1983; Johnson *et al.* 1983).

2.2.10 Amplification of double stranded mtDNA from genomic DNA

The 1.1 kb control region of the mtDNA molecule was amplified by the polymerase chain reaction (PCR) from 144 individuals (Appendix 4) using the primers L15966 and H408 (Table 2.3). All buffers and recipes are given in Appendix 2.

A typical 50 μl reaction was set up by first preparing a reaction mix that consisted of 5 μl of 10X PCR buffer, 1 μl of a 10 mM dNTP mix, 1 μl each of primers L15996 and H408 from 10 μM stocks, 0.2 μl Promega Taq polymerase (1 unit) and 36.8 μl of double distilled water. A cocktail of this

reaction mix was made for the number of PCR reactions plus control reactions required. Forty-five μl of this reaction cocktail were added into sterile PCR reaction tubes. Five μl of total DNA (500 ng) were added to the test reaction tubes whilst 5 μl of double distilled water were added to the negative control reaction tube. These reaction mixes were covered with 1-2 droplets of light mineral oil to prevent evaporation. The reaction mixes were centrifuged briefly in an Eppendorf microfuge (30 sec) to mix the contents and the Eppendorf tubes and then transferred to a programmable thermocycler (Perkin Elmer Cetus).

Each amplification cycle consisted of denaturation at 94°C for 1 min, annealing at 56°C for 1 min and extension at 74°C for 1 min programmed for a total of 30 such cycles. After completion, a 5 μl aliquot of the PCR products were mixed with loading dye and loaded on a 1% agarose gel containing 0.1 μg/ml ethidium bromide to check for the successful amplification of the 1.1 kb fragment. One lane on the gel was loaded with 1 μg of the 1.0 kb DNA ladder (BRL) that was used as a marker to estimate the size of the PCR product. Following electrophoresis in TBE buffer, products were visualized with the aid of a UV transilluminator.

Fifteen μl of PCR products were loaded in a 1% low melting gel (SeaPlaque) containing 0.1 μg ethidium bromide. A DNA plug free of nuclear DNA, dNTP's and primers was aspirated from the gel with a sterile yellow tip cut at the end.

Care was taken not to expose the DNA samples to long exposure to UV, which causes nicks in the DNA. The aspirated plug was then dissolved in 500 μl of double distilled water and stored at 4°C until required.

Table 2.3. Oligonucleotide primers used in the PCR and sequencing reactions.

Primer	strand	position of 3' end	sequence (5' - 3')
L15996*	light	15996	CTCCACCATTAGCACCCAAAGC
H408*	heavy	408	CTGTTAAAAGTGCATACCGCCA
H16401	heavy	16401	TGATTTCACGGAGGATGGTG
L29	light	29	GGTCTATCACCCTCTTAACCAC

*primer set used in the PCR

2.2.11 Preparation of single stranded DNA for sequencing

Several procedures were used to obtain high quality single-stranded DNA templates for sequencing. This was because of the difficulty in obtaining single-stranded products of the light strand solely by asymmetric PCR. Thus, single-stranded DNA products were obtained from a combination of the following procedures:

2.2.11.1 Asymmetric PCR

PCR using unequal molar amounts of PCR primers successfully yielded the 1.1 kb single-stranded product for the H-strand. To set up the reactions, gel-purified DNA dissolved in distilled water was heated at 65°C for 5 min, vortexed to mix thoroughly, and a 10 μl aliquot dispensed into a sterile PCR reaction tube. To this was added 90 μl of reaction cocktail made up of: 2 μl 10 mM dNTP mix, 10 μl 10X buffer, 2 μl excess primer (10 μM), 2 μl limiting primer (0.2 μM), 0.4 μl Taq polymerase (2 units) and 73.6 μl of double distilled water. The reaction mixes were covered with 1-2 drops of mineral oil, centrifuged for 1 min and transferred to the thermocycler. PCR conditions were as for the double-stranded product given above, except that the number of cycles was increased to 35.

The PCR product was purified and concentrated by Ultrafree-

MC filter unit centrifugation (Millipore, 30 000 cut off) as follows: 350 μl of sterile TE were first added to the cartridge followed by approximately 90 μl of PCR product (oil removed by chloroform extraction). The units were then centrifuged at 5 000 rpm for 8 min in an Eppendorf microfuge. The filtrate which had collected in the Eppendorf tube was discarded and 450 μl of TE were added to the cartridge to wash the DNA further. The centrifugation step was repeated as before. DNA trapped on the filter was resuspended in 50 μl TE and stored at 4°C until required for sequencing.

For poor yields of single-stranded PCR products (as observed on the trial gel), DNA on the filter was washed with two volumes of 50 μl TE and then precipitated using salt (1/10 volume of 7.5 M sodium acetate, pH 5.0) and cold ethanol either at -20°C for 2 hr or -70°C for 30 min. Following centrifugation at 11 000 rpm in an Eppendorf microfuge for 5 min, the supernatant was removed by aspiration and the pellet washed in 70% ethanol for 3 min using the same centrifugation conditions. This pellet was then dissolved in 20 μl of TE, such that there was sufficient DNA for at least two sequencing reactions.

2.2.11.2 PCR with one phosphorylated primer followed by λ exonuclease digestion

Since the light strand was sometimes difficult to obtain by the asymmetric method, single-stranded products containing light strands were obtained using λ exonuclease. This reaction is based on the increased affinity of λ exonuclease for DNA strands phosphorylated at the 5′-terminus. In this instance primer H408 was phosphorylated and used together with primer L15996 to amplify the control region.

(a) Phosphorylation of primers

Approximately 20 μg of primer H408 (27.5 μl of a 100 μM stock) were placed in an Eppendorf tube and made up to 40 μl by the addition of the following reagents to the tube: 4 μl 10X kinase buffer, 0.5 μl ATP (100 mM), 2 μl T_4 polynucleotide kinase (BRL) and 6 μl of sterile double distilled water. This reaction was incubated at 37°C for 1 hr and then stopped by inactivating the enzyme phenol followed by one chloroform extraction. A Biospin 30 column (Biorad) was packed by centrifugation at 2 500 rpm in a Sorval centrifuge. The packed column was then washed twice with a TE solution containing in addition 50 mM NaCl. The sample was applied to the column and the eluate collected into an Eppendorf tube by centrifugation at 2 500 rpm. This phosphorylated primer was stored at -20°C until

required for PCR.

(b) PCR with one phosphorylated primer

A 100 µl PCR mix was prepared using 10 µM each of the phophorylated H408 primer and the unphosphorylated L15996 primer in the presence of dNTP's, enzyme, buffer and DNA (500 ng) in the same molar ratios as used for the double stranded PCR reaction described above. The reaction conditions were also the same. PCR products were purified by Ultrafree-MC 30,000 filter unit centrifugation in the same manner as described before. Finally, the DNA was recovered from the filter by mixing with 70 µl of TE.

Sixty µl of this purified product were digested with λ exonuclease in a digestion mix containing 7 µl of 10X exo buffer and 8 units of λ exonuclease enzyme in a total volume of 70 µl. This mix was incubated at 37°C for 30 min. The reaction was stopped by column purification using Centricon-100 units. The larger pore size of these filters allows the enzyme and other reaction constituents to pass through but retains the DNA. The sample was added to the cartridge of the Centricon unit and the volume was made up to 3 ml by the addition of TE buffer. Centrifugation was at 2 500 rpm in an IEC B-20A centrifuge with a fixed angle rotor (Type IEC 870).

A 5 μl aliquot of the retinate (approximately 30 μl) was electrophoresed together with a 5 μl aliquot of the double stranded product on a 1% agarose gel to ascertain whether the digestion was successful.

2.2.12 Sequencing of single-stranded DNA

DNA was sequenced using the Sanger dideoxy chain termination procedure (Sanger *et al*. 1977) and the Johnsonase T7 DNA polymerase. The *E.coli* clone A179/pGA1-14/pGp1 which carries the exonuclease-deficient (exo) mutant was a gift from Prof K.A. Johnson and Dr S. Patel (Pennsylvania State University). The construction and purification of the enzyme is described by Patel *et al*. (1991).

2.2.12.1 Reconstitution of T7 DNA polymerase

Gene 5 protein is only active when reconstituted with thioredoxin in the molar ratio of 1:20, respectively. Thioredoxin has to be first reduced with DTT to a final concentration of 5 mM in the presence of T7 buffer by incubating for 2-3 min on ice. Thioredoxin was a gift from Dr S. Patel, Pennsylvania State University. Gene 5 protein was then added to the reduced thioredoxin to form the active T7 DNA polymerase.

2.2.12.2 Sequencing protocol

The first step in the protocol was to anneal the sequencing primers to the single-stranded templates. To verify the sequence contained within the 1.2 kb control region, both the light and the heavy strands were sequenced from each individual studied. For the light strand templates primers H408 and H16401 were used while for heavy strand templates primers L29 and L15996 were used (see Table 2.3). A 10 μl reaction mix was made by adding a 7 μl aliquot of single-stranded purified PCR product, 2 μl 5X annealing buffer and 1 μl primer (1 μM) in a 0.65 ml microfuge tube. This mixture was spun briefly in a microfuge and transferred to the thermocycler programmed to heat the sample at 65°C for 5 min, and then cooled to 37°C over a period of 30 min. This ensured the efficient annealing of the primer to the single-stranded template. After this cycle, the tubes were centrifuged in the microfuge for 30 sec to collect the products of evaporation.

A cocktail of "labelling mix" consisting of 2 μl of a 1/20 dilution of a "G" nucleotide mix (USB nucleotide kit), 1 μl of 0.1 M DTT, 1 μl [^{35}S]-dATP and 2 μl reconstituted T7 DNA polymerase per reaction, was made for the number of sequencing reactions required. After mixing, 6 μl of this mix were added and mixed thoroughly with the annealed primer/template product. This reaction was incubated at room temperature for 5 min. While this reaction was being

plate. After the 5 min incubation, 3.5 μl of the this primer/template/enzyme/-[^{35}S]-dATP/"G" mix was added to each termination mix containing both deoxy- and dideoxy-nucleotides (USB) and incubated for 5 min at room temperature. The reactions were stopped by the addition of 6 μl of stop solution. The microtiter plates were stored at -20°C.

2.2.12.3 Sequencing gel electrophoresis

Seventy-five ml of a working acrlyamide solution consisting of 6% polyacrylamide and 7 M urea were used to prepare a sequencing gel of dimensions 40 x 80 cm (OWL Scientific gel apparatus). The plates were cleaned thoroughly with ethanol and water and the notched plate treated with Sigmacote (Sigma) to siliconize it. These plates were assembled with 0.4 mm spacers between them and securely taped together with casting tape (OWL Scientific). A 10% solution of APS (200 μl) and 45 μl of TEMED were added and mixed into the aliquoted acrylamide solution and dispensed gently into the gel mould with the aid of a 50 ml syringe to prevent the formation of bubbles. A 60-well comb was inserted with the flattened end into the gel and the gel left for about an hour at room temperature to polymerize.

After polymerization the comb was removed, cleaned and inserted such that the teeth of the comb just made contact with the gel surface to form the wells. The tape from the

bottom edge of the gel mould was removed and the plates assembled to the aluminium backed sequencing gel system (OWL Scientific). TBE buffer was added to the appropriate buffer chambers and the wells were cleaned by flushing with TBE buffer. Stop solution (with dye) was loaded at random across the width of the gel to inspect the quality of the wells. These gels were prewarmed by electrophoresis at 40 mA for 45 min.

Samples were heat denatured by placing the microtiter plate on a preheated inverted heating block (90°C) for 5 min. Three μl of denatured sample were loaded into the wells in the order G, A, T, C. Electrophoresis in TBE buffer was at 45 mA (corresponding to a gel temperature of approximately 55°C) for 2.5-3 hr. Following electrophoresis, gels were fixed in 10% glacial acetic acid/10% methanol mix for 20 min and then dried on Whatman 3M filter paper in a gel dryer at 80°C for 1 hr. Dried gels were exposed to X-ray film (Kodak XAR) in Sigma cassettes for 12-72 hr at room temperature.

2.3 Statistical Analyses

2.3.1 MtDNA variation within populations

2.3.1.1 Genetic diversity

The amount of variation (genetic diversity) within ethnic groups was calculated from the frequencies of mtDNA types found within each ethnic group using the method suggested by Nei (1987). According to this method, gene diversity (h) at a locus is defined as follows:

$$h = 1 - \Sigma x_i^2$$

where x_i is the frequency of the allele i for the total number of alleles m at the locus.

To overcome sampling biases, an unbiased estimate of gene diversity may be computed from the equation:

$$h = (N/N-1)(1 - \Sigma x_i^2)$$

where N is the total number of individuals in the sample (Nei 1987).

2.3.1.2 Nucleotide diversity

For restriction site and DNA sequence data, a more appropriate measure of polymorphism within a population is the average number of nucleotide differences per site between any two randomly chosen sequences. This measure is called the nucleotide diversity and is denoted by π (Nei

and Li 1979) and is defined as:

$$\pi = \Sigma\ x_i x_j \pi_{ij}$$

where x_i and x_j are the frequencies of the ith and jth type of DNA sequences, respectively, and π_{ij} is the proportion of different nucleotides between the ith and jth type of DNA sequence.

2.3.2 MtDNA variation between populations

2.3.2.1 G_{ST}

G_{ST} gives an indication of the fraction of genetic variation within an entire population which is due to interdeme genetic differences. It has been defined by Nei (1973) as the coefficient of population differentiation. For mtDNA data, the entire molecule is treated as a single locus, with the number of mtDNA types equivalent to the number of alleles and is calculated as follows:

$$G_{ST} = (H_T - H_S)/H_T$$

where H_S (average diversity among subpopulations) is calculated as $H_S = (\sum h_k)/k$ for k subpopulations and $H_T = 1 - \sum x_i^2$, where x_i is the average frequency across subpopulations.

G_{ST} is a convenient measure of population differentiation because it reaches equilibrium quickly (Crow and Aoki 1984) and is relatively insensitive to the number of population subdivisions (Takahata and Palumbi 1985).

Takahata and Palumbi (1985) developed a model which describes the behaviour of a single locus on extranuclear genomes, of which mtDNA is an example. Using the parameters defined in this model, they have modified Nei's method for estimating G_{ST} which could be applied directly using restriction site data. In this method each restriction site is treated as a separate locus with two alleles (presence or absence of a site).

Substituting the model parameters [equations 17 and 18, Takahata and Palumbi (1985)] in the G_{ST} equation described in the Nei method, the equilibrium value for G_{ST} using this method is given by:

$$G_{ST} = [1 + 2N_e (L/L-1)^2 \{m_e + (L/L-1)v\}]^{-1}$$

where N_e is the effective population number; m_e the effective migration rate; L is the number of demes, and v is the mutation rate. The derivation of these parameters are discussed in Takahata and Palumbi (1985).

2.3.2.2 Nucleotide divergence

Nucleotide divergence between any two population groups was calculated by the method of Nei and Li (1979). The algorithms used in this method have been incorporated into a computer program called RESTSITE which was written by Nei and Miller (1990) and kindly supplied by Dr Nei (The Pennsylvania State University). In this method, nucleotide divergences were estimated from the number of shared sites

(assuming that they were derived from a common ancestral DNA sequence) and different restriction sites (new sites following their divergence from the ancestral sequence) and are expressed as the mean number of substitutions per nucleotide site (δ) as follows:

$$\delta = -(3/2)\ln[(4S^{1/2r} - 1)/3]$$

where $S = 2n_{XY}/(n_X + n_Y)$, [$n_X = n_{X1} + n_{X2}$, where n_{X1} and n_{X2} are the number of ancestral restriction sites and the number of new restriction site in population X, respectively; similarly n_Y is the corresponding values in population Y; and n_{XY} is the number of identical sites shared by the two populations]; r = the number of nucleotides in the recognition sequence of the restriction enzymes used.

The above method makes two assumptions: Firstly, the expected G+C content stays constant and, secondly, that nucleotide substitutions occurs randomly at a constant rate (λ) with time (t). Under these assumptions, the probability that an original restriction site remains unchanged over a period of time t is given by $P = e^{-r\lambda t}$ (Nei and Li 1979). Therefore, when $P = e^{-r\lambda t}$ is used, the mean number of nucleotide substitutions per site is given by:

$$\delta = 2\lambda t$$

Thus, if the mutation rate (λ) is known, t may be estimated from restriction site data.

2.4 Neutrality tests

The Ewens-Waterson test was used to test the fit of observed data to the infinite sites model (Ewens 1972; Watterson 1977). Sampling theory of the infinite sites model shows that the expected allele frequency distribution is fully specified by the sample size (n) and the number of observed alleles (k). One-tailed tests were done by scoring the fraction of the simulations whose test statistic was greater than the observed number. In addition to the traditional test of homozygosity or gene identity (F), the significance of departures from the expected number of singleton alleles (S) and from the frequency of the most common allele (C) was tested.

The neutrality test of Tajima (1989) was also performed for each population. The Tajima test is based on the concordance between estimates of $\theta = 4Nu$ obtained from the heterozygosity calculations and from the number of segregating sites.

Both the Ewens-Watterson and the Tajima tests were applied to the data making use of a computer program written and kindly supplied by Dr A.G. Clark (The Pennsylvania State University).

2.5 Phylogenetic analysis

There are many different methods available for the construction of phylogenetic trees from molecular data. Excellent reviews on the principles and usefulness of the various approaches are provided by Nei (1987) and Swofford and Olsen (1990). Both the UPGMA (unweighted pair-group method with arithmetic mean) and the neighbour-joining algorithms (Saitou and Nei 1987) were used to construct phylogenetic trees from the distance matrix generated by the Nei and Miller (1990) method for the pairwise comparisons of mtDNA types. Data from southern African populations were compared to other populations studied elsewhere by the same method (Merriwether *et al.* 1991).

A phylogenetic tree relating all mtDNA types obtained by control region sequencing using parsimony analysis, was made possible by using the computer program PAUP (Phylogenetic Analysis Using Parsimony, version 2.4), developed by Dr D. Swofford and kindly made available by Dr M. Stoneking (The Pennsylvania State University). Phylogenetically informative sites were identified using the MacClade computer program developed by Madison and Madison (1987). Data from Vigilant *et al.* (1991) were pooled with the sequence data obtained from the present study and analysed by the parsimony and neighbour-joining methods using the chimpanzee sequence (Foran *et al.* 1988) as an outgroup to root the tree.

CHAPTER THREE:

RESTRICTION FRAGMENT LENGTH POLYMORPHISMS

3.1 Background

Several studies have been made employing mtDNA RFLP analysis to study population variation and human mtDNA evolution. These studies have made use of low resolution mapping (Johnson *et al*. 1983; Bonne-Tamir *et al*. 1986; Brega *et al*. 1986a,b; Santachiara-Benerecetti *et al*. 1988, Sartoris *et al*. 1988; Scozzari *et al*. 1988; Vikki *et al*. 1988; De-Benedictis *et al*. 1989a,b; Torroni *et al*. 1990) and high resolution mapping (Brown 1980; Cann and Wilson 1983; Cann *et al*. 1984, 1987; Harihara *et al*. 1986, 1988; Horai *et al*. 1986; Horai and Matsunuga 1986; Stoneking 1986; Stoneking *et al*. 1990) as discussed in Chapter 1.8.

Several population groups consisting of 3 065 individuals (including the 795 from this study) have been investigated for mtDNA variations using the low resolution mapping approach. This method was used in the present study to examine the mtDNA variation in 795 individuals drawn from 20 ethnic groups from southern Africa.

3.2 Restriction enzyme analysis

Polymorphisms were detected with the enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II. Each polymorphic restriction pattern is called a morph and assigned a number in ascending chronological order as it was discovered (Johnson *et al*. 1983). New morphs are depicted with a superscript showing the abbreviation of the population group in which it is found, eg., *Msp*I-7^{Hindu}, is assigned to the new morph found in Hindus (Semino *et al*. 1991).

3.2.1 Description of morphs

3.2.1.1 *Hpa*I

*Hpa*I recognizes the sequence GTTAAC. There are three sites found at positions 5693, 10016 and 12408 in the reference sequence (Anderson *et al*. 1981). To date eight morphs have been described with this enzyme (Denaro *et al*. 1981; Santachiara-Benerecetti *et al*. 1988; Scozzari *et al*. 1988) of which five (Fig. 3.1a) are found at varying frequencies in the populations sampled in this investigation (Table 3.1). A diagram showing the relationship of all morphs discovered in this study is shown in Fig 3.1b.

Morph *Hpa*I-1 is associated with site *a*, localized within the asparagine tRNA anticodon sequence (np 5693), and site *b* which is found in the region which forms the stem of the

anticodon loop of glycine tRNA (np 10016). Both these sites are found in all human *Hpa*I morphs as well as in chimpanzees, gorillas and orangutans. Morph *Hpa*I-1 is found at very low frequencies in all human populations. It is found at a frequency of 12.5% in Orientals (Johnson *et al.* 1983) and at lower frequencies in South African Negroids (4.0%) (Johnson *et al.* 1983) and Senegalese Negroids (1.6%) (Scozzari *et al.* 1988). In the present study it was found in the Zulu, Swazi, Xhosa and Johannesburg "Coloured" groups, albeit at low frequencies (Table 3.1).

The observation that *Hpa*I-1 was found in chimpanzees, gorillas and orangutans (Ferris *et al.* 1981), led Johnson *et al.* (1983) to hypothesize that *Hpa*I-1 may be the ancestral morph from which all other *Hpa*I morphs found in humans have evolved. They also suggested that since this morph was only found in Asians and not in the San and Pygmy groups studied it meant that modern humans diverged recently from an Asian ancestor. More recent data based on the analysis of several restriction enzymes and mtDNA sequencing, however, suggest an African origin for the evolution of human mtDNAs (Cann *et al.* 1987; Vigilant *et al.* 1991).

*Hpa*I-2 differs from *Hpa*I-1 by a new site gained at position 12408 which is localized in an open reading frame, to which no function has been assigned (Johnson *et al.* 1983). It is

found at high frequencies in Caucasoids investigated in this study (Table 3.1) as well as in Caucasoids investigated from other geographic locations (Denaro *et al.* 1981; Johnson *et al.* 1983; Bònne-Tamir *et al.* 1986; Sartoris *et al.* 1988; Vikki *et al.* 1988; De Benedictis *et al.* 1989; Semino *et al.* 1989). *Hpa*I-2 is found at a higher frequency in the Herero (88.9%) and Dama (65.1%) groups when compared with other Negroids investigated in this study (see Table 3.1). It is also found at a lower frequency in Senegalese Negroids (32.3%) (Scozzari *et al.* 1988). On the other hand, *Hpa*I-2 is found at lower frequencies in the Nama (19.6%) than the San (8.5% in the Sekele and 2.9% in the !Kung) groups.

*Hpa*I-3 differs from morph *Hpa*I-2 by an additional site *d* located at position 3592. This site results from a C to T transition at position 3592 in the region which codes for a protein of unknown function. However, the first three nucleotides of the *Hpa*I recognition sequence code for the amino acid valine (GTC). The mutation at the third nucleotide position results in the codon GTT which also codes for valine, hence a silent mutation.

*Hpa*I-3 is the commonest morph in African populations, and is found at high frequencies in the San (97.1% in the !Kung and 93.1% in the Sekele). It is found at a comparatively lower frequency in the Nama (76.1%) and Negroid groups [65.1% in Senegalese (Scozzari *et al.* 1988), 70.8% in South

African Bantu-speaking Negroids previously analysed by Johnson *et al.* (1983) and at varying frequencies in Negroids investigated in the present study (see Table 3.1)].

In contrast, *Hpa*I-3 is rarely found in Caucasoids. In fact, Semino *et al.* (1989) have suggested that this morph could be used as a marker to measure the extent of gene flow from Africans into European Caucasoids. Using the frequencies of morph *Hpa*I-3 found in Negroids and the Caucasoid population sampled in Sicily, Semino *et al.* (1989) estimate that approximately 10.8% of genes in the Sicilian Caucasoid population have maternal Negroid ancestry. Moreover, it was suggested that the African markers were introduced into the Sicilian gene pool from different sources: The Negroid component could have been transmitted directly through the introduction of groups of Negro slaves brought into the island by Phoenicians and Romans and/or indirectly through Arabic migrations, since Arabs are known to have typical Negroid nuclear genes at high frequencies (Mourant *et al.* 1978).

In the present study *Hpa*I-3 was found among Caucasoids only in the Indian (4.1%) and the Afrikaans-speaking (2.0%) groups, suggesting that there has been some gene flow from the indigenous South African populations into these two ethnic groups. Blood group markers (Botha and Pritchard 1972; Moores 1980; Nurse *et al.* 1985) and immunoglobulin

variants (Jenkins *et al.* 1970) found more commonly in indigenous populations from South Africa have been found in immigrant Caucasoids also providing evidence for gene flow from indigenous southern Africans into the Caucasoid population in South Africa.

*Hpa*I-5 differs from *Hpa*I-2 by the addition of site *f* found at position 1113 in the region of the 12S RNA. This morph is rare and was found only in one individual in the present investigation: a member of the Ambo group.

*Hpa*I-7 was first observed in a Wolof individual from Senegal and is characterized by the fusion of the 2.4 and 7.8 kb fragments observed in morph 3 to produce a fragment of 10.2 kb (Fig. 3.1a). This site loss results from a T to C mutation at position 12408 (Scozzari *et al.* 1988). Morph 7 was found in two Nama individuals.

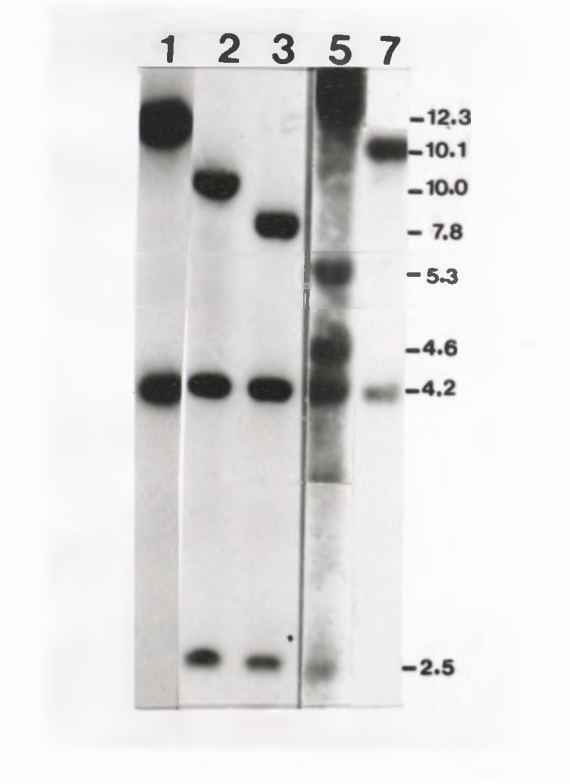


Fig. 3.1 (a) Fragment patterns of *Hpa*I morphs discovered in southern African populations (see text for details). Fragment sizes are given in kilobases.

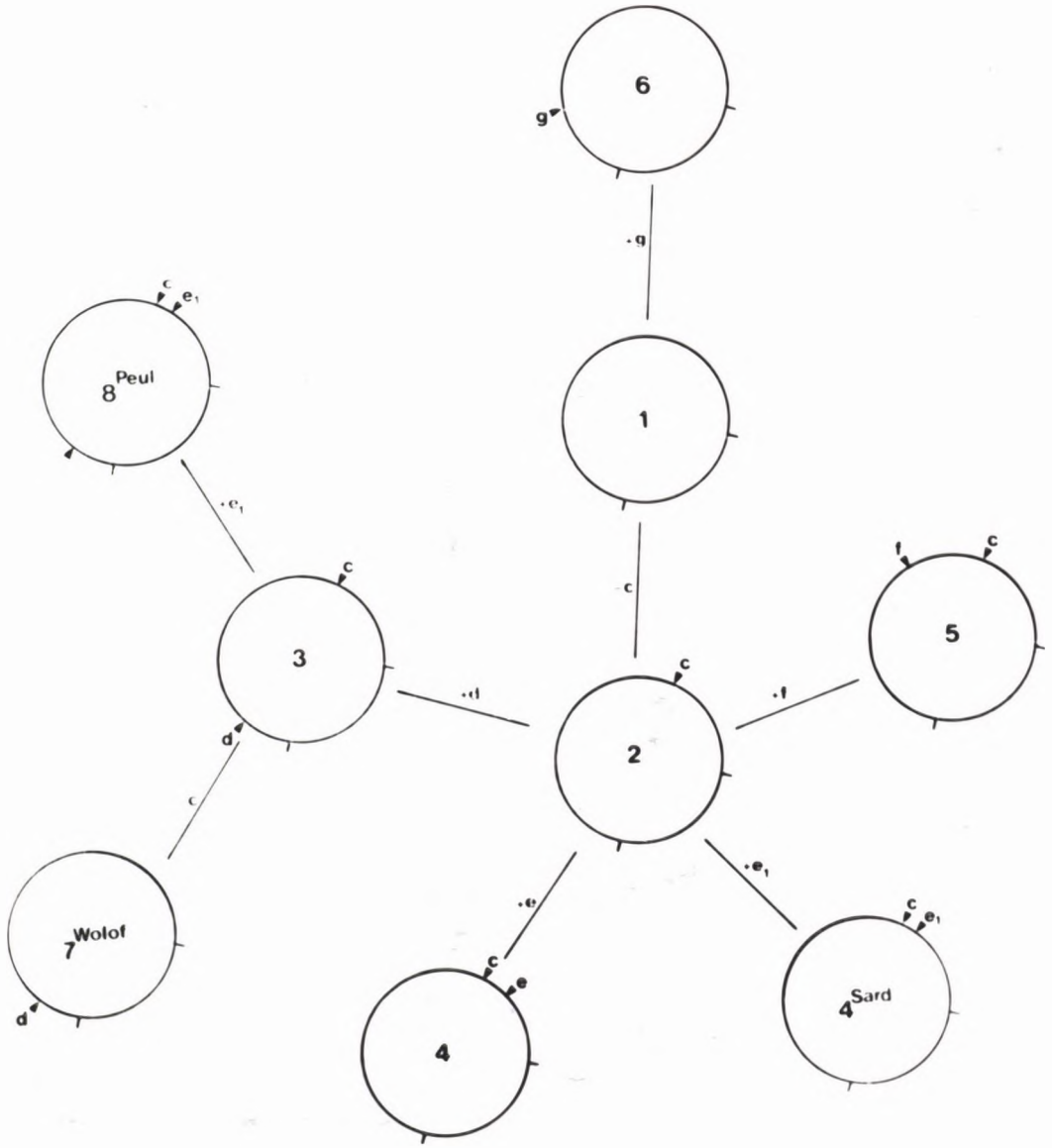


Fig. 3.1 (b). Phylogeny showing the relationship of the eight *Hpa*I morphs discovered to date. Arrowheads with letters on circle indicate variable sites whilst lines on circle indicate the invariant sites on the molecule. Lines connecting circles show how each morph can be derived with the letters showing the sites either gained (+) or lost (-). Data obtained from Denaro *et al.* (1981), Santachiara-Benerecetti *et al.* (1988) and Scozzari *et al.* (1988).

Table 3.1. The numbers and the frequencies (percentages) of *Hpa*I morphs in twenty ethnic groups from southern Africa. Standard errors are given in brackets.

		Hpal-1			Hpal-2			Hpal-3			Hpal-5			Hpal-7		
	N	No	Freq	(S.E)	No	Freq	(S.E)	No	Freq	(S.E)	No	Freq	(S.E)	No	Freq	(S.E)
Zulu	30	1	3.3	3.3	6	2.0	2.6	23	76.7	7.7	-	-	-	-	-	-
Swazi	41	1	2.4	2.4	17	41.5	7.7	23	56.1	7.8	-	-	-	-	-	-
Xhosa	18	1	5.6	5.4	4	22.2	9.8	13	72.2	10.6	-	-	-	-	-	-
Tsonga	34	-	-	-	11	32.4	8.0	23	67.6	8.0	-	-	-	-	-	-
Sotho	29	-	-	-	7	24.1	7.9	22	75.9	7.9	-	-	-	-	-	-
Pedi	23	-	-	-	2	8.7	5.9	21	91.3	5.9	-	-	-	-	-	-
Tswana	15	-	-	-	5	33.3	12.2	10	66.7	12.2	-	-	-	-	-	-
Venda	30	-	-	-	4	13.3	6.2	26	86.7	6.2	-	-	-	-	-	-
Lemba	53	-	-	-	9	17.0	5.2	44	83.0	5.2	-	-	-	-	-	-
Herero	54	-	-	-	48	88.9	4.3	6	11.1	4.3	-	-	-	-	-	-
Dama	43	-	-	-	28	65.1	7.3	15	34.9	7.3	-	-	-	-	-	-
Ambo	22	-	-	-	5	22.7	8.9	16	72.7	9.5	1	4.5	4.4	-	-	-
Sekele	49	-	-	-	3	6.1	3.4	26	93.9	7.1	-	-	-	-	-	-
Nama	46	-	-	-	9	19.6	5.9	35	76.1	6.3	-	-	-	2	4.3	3.0
R_"Col"	35	-	-	-	4	11.4	5.4	31	88.6	5.4	-	-	-	-	-	-
JHB_"Col"	71	2	2.8	2.0	17	23.9	5.1	52	73.2	5.3	-	-	-	-	-	-
English	36	-	-	-	36	100.0	0	-	-	-	-	-	-	-	-	-
Afrikaners	51	-	-	-	50	98.0	2.0	1	2.0	2.0	-	-	-	-	-	-
A_Jews	55	-	-	-	55	100.0	0	-	-	-	-	-	-	-	-	-
Indians	50	-	-	-	47	94.0	3.4	3	6.0	3.4	-	-	-	-	-	-

N = sample size

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

3.2.1.2 *Bam*HI

*Bam*HI recognises the sequence GGATCC and there is only one site (located at position 14258) in the reference sequence (Anderson *et al*. 1981). Thus far only five morphs have been discovered with this enzyme (Johnson *et al*. 1983; Semino *et al*. 1991), but only three are found at varying frequencies in the different populations sampled (Table 3.2).

*Bam*HI-1 corresponds to the published mtDNA sequence and this produces a single hybridization band equivalent to the linear mtDNA molecule (Fig 3.2a). *Bam*HI-2 is characterized by the generation of two fragments of sizes 15.7 and 2.1 kb which results from a site most likey gained at position 16350. This mutation occurs in the D-loop region of the molecule. *Bam*HI-3 is characterized by the presence of the site at position 16350 and a mutation creating a new *Bam*HI site at position 13368 within ND5 (Santachiara-Benerecetti *et al*. 1988). A phylogeny showing the relationship of all three morphs is shown in Fig 3.2b.

*Bam*HI is monomorphic in Negroids and Khoisan people and all individuals in these groups have morph *Bam*HI-1. *Bam*HI-1 is also monomorphic in other African populations (Johnson *et al*. 1983; Scozzari *et al*. 1988). Some variation is observed in Caucasoid populations of European descent

(Table 3.2). Whereas *Bam*HI-2 was found at low frequencies in all three European Caucasoids groups (English- and Afrikaans-speaking "Whites" and Ashkenazi Jews), *Bam*HI-3 was only found in the Ashkenazi Jewish group (1.8%).

There appears to be a concomitant change in sequence in the region at position 13368 that affects the *Ava*II site *d*, which when abolished by a G to A transition at position 13368 results in morph *Ava*II-9. The G to A transition at this position creates the new *Bam*HI site (see below) and this produces morph *Bam*HI-3. This mutation does not alter the amino acid codon for glycine thus not affecting the amino acid sequence coded by the ND5 gene (Santachiara-Benerecetti *et al*. 1988).

```
      Gly Ser                    Gly Ser
....GGG TCC....mutates to ... GGA TCC...

    AvaII site d              BamHI site a
```

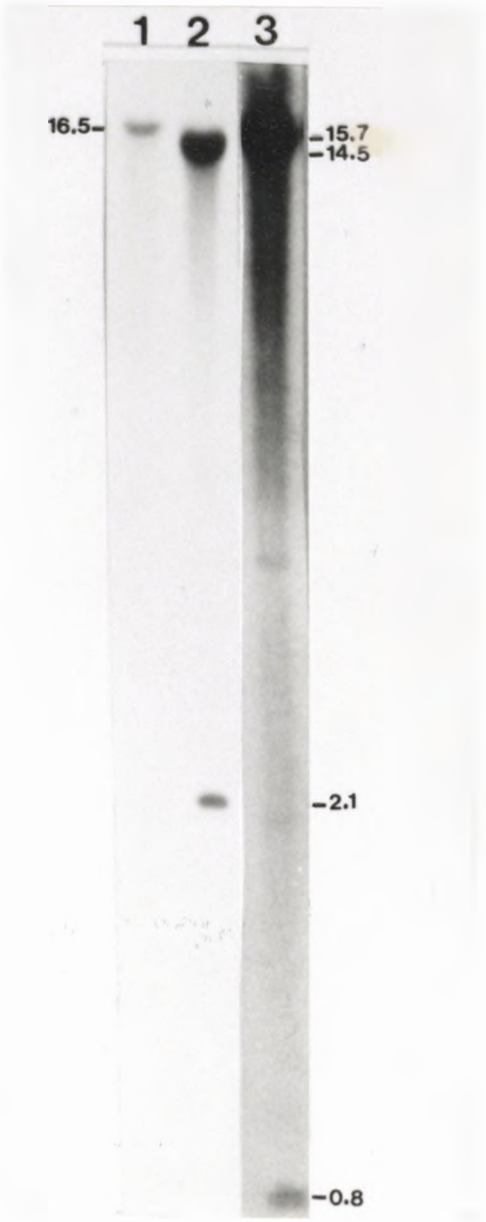


Fig. 3.2 (a). Fragment patterns of *Bam*HI morphs discovered in southern African populations (see text for details). Fragment sizes are given in kilobases.

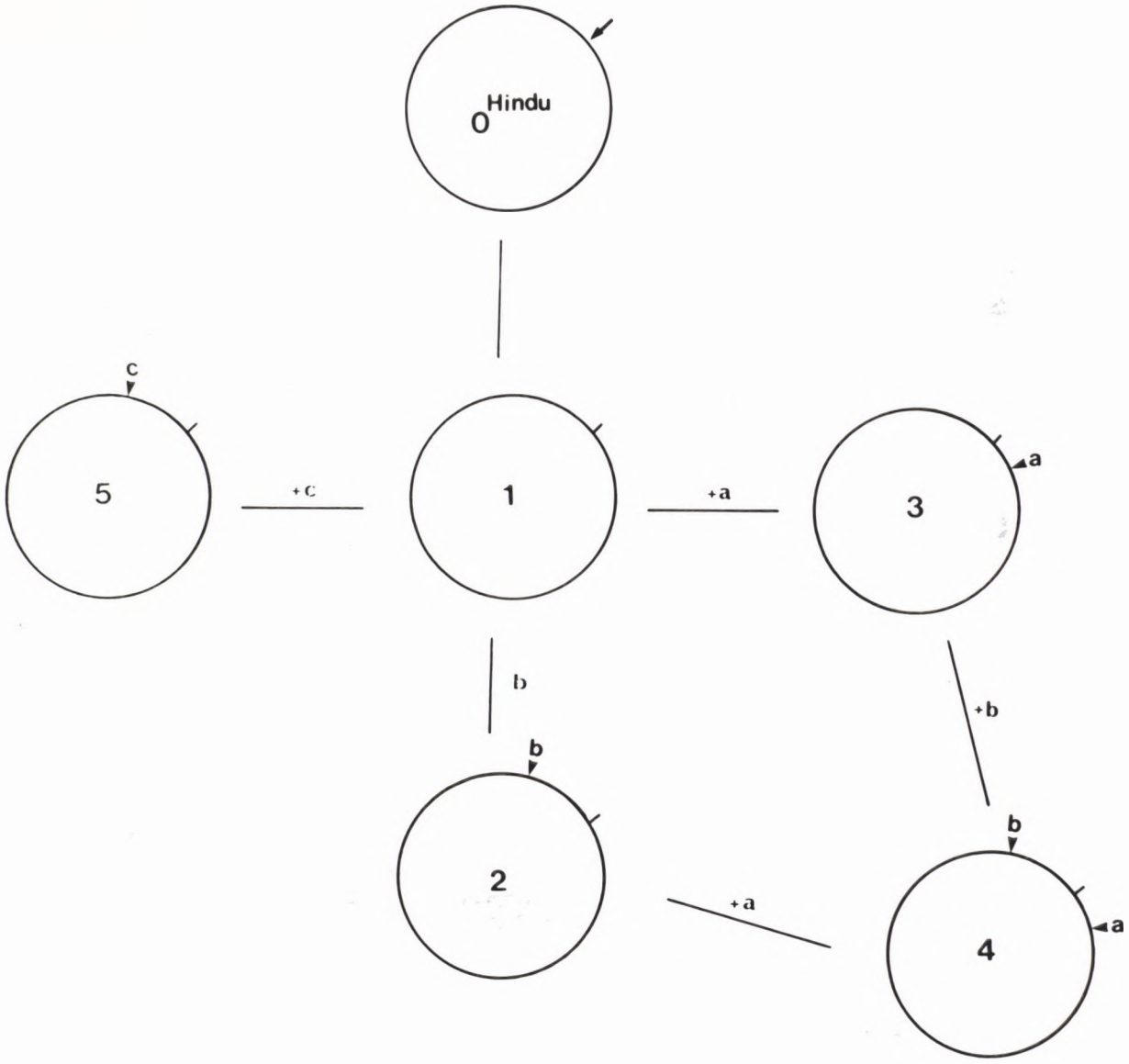


Fig. 3.2 (b). Phylogeny showing the relationship of the five *Bam*HI morphs discovered to date. Arrowheads with letters on circle indicate variable sites, lines on circle indicate the invariant sites on the molecule and arrow indicates an invariant site which has been deleted. Lines connecting circles show how each morph can be derived with the letters showing the sites either gained (+) or lost (-). Data obtained from Johnson *et al.* (1983), Santachiara-Benerecetti *et al.* (1988) and Scozzari *et al.* (1991).

Table 3.2. The numbers and the frequencies of *Bam*HI morphs in twenty southern African populations. Standard errors are given in brackets.

		BamHI-1			BamHI-2			BamHI-3		
	N	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
Zulu	30	20	100.0	-	-	-	-	-	-	-
Swazi	41	41	100.0	-	-	-	-	-	-	-
Xhosa	18	18	100.0	-	-	-	-	-	-	-
Tsonga	34	34	100.0	-	-	-	-	-	-	-
Sotho	29	29	100.0	-	-	-	-	-	-	-
Pedi	23	23	100.0	-	-	-	-	-	-	-
Tswana	15	15	100.0	-	-	-	-	-	-	-
Venda	30	30	100.0	-	-	-	-	-	-	-
Lemba	53	53	100.0	-	-	-	-	-	-	-
Herero	54	54	100.0	-	-	-	-	-	-	-
Dama	43	43	100.0	-	-	-	-	-	-	-
Ambo	22	22	100.0	-	-	-	-	-	-	-
Sekele	49	59	100.0	-	-	-	-	-	-	-
Nama	46	46	100.0	-	-	-	-	-	-	-
R_"Col"	35	35	100.0	-	-	-	-	-	-	-
JHB_"Col"	71	71	100.0	-	-	-	-	-	-	-
English	36	33	91.7	4.6	3	8.3	4.6	-	-	-
Afrikaners	51	47	92.2	3.8	4	7.8	3.8	-	-	-
A_Jews	55	53	96.4	2.5	1	1.8	1.8	1	1.8	1.8
Indians	50	50	100.0	-	-	-	-	-	-	-

N = sample size

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

3.2.1.3 *Hae*II

*Hae*II recognizes the sequence PGCGCQ (where P is a purine and Q a pyrimidine). Six of the twelve published morphs (Johnson *et al.* 1983; Horai *et al.* 1984; Bonne-Tamir *et al.* 1986; Santachiara-Benerecetti *et al.* 1988; Scozzari *et al.* 1988) were found in individuals sampled in this study (Fig. 3.3a). A phylogeny showing the relationship of all morphs discovered in this study is given in Fig. 3.3b.

*Hae*II-1 corresponds to the published sequence and is the commonest morph in all groups (Table 3.3). It is characterized by the presence of fragments 4.7, 4.5, 4.1, 1.4, 1.2, 0.4 and 0.1 kb (Fig. 3.3a).

*Hae*II-2 may be derived from *Hae*II-1 and is associated with the loss of the 4.5 and 4.1 kb fragments and the appearance of the 8.6 kb fragment. This is due to the alteration of a recognition site located at position 9052 of the reference sequence which is found in the region coding for ATPase 6. *Hae*II-2 is common in the Ashkenazi Jewish population (30.9%) and also very common in Israeli Jews (37.5%) (Bonne-Tamir *et al.* 1986). It is found at a lower frequency in other Caucasoids (English-speaking "Whites" and Afrikaans-speaking "Whites" at frequencies of 13.9% and 2.0%, respectively), in some Negroids (Xhosa, 11.1%; Tswana, 6.7%) and one "Coloured" individual investigated in the present study.

*Hae*II-3 is generated by replacement of the 1.4 and 4.5 kb fragments and the appearance of the 5.9 kb fragment (Fig. 3.3a) caused by the loss of a site at position 4529 found in ND2. This morph is rarely found and was only observed in three Caucasoid individuals (Table 3.3).

*Hae*II-6 was previously observed in one Caucasoid individual (Johnson *et al*. 1983). The mutation giving rise to this morph is thought to involve a new site gained at position 9250. This results in the loss of the 4.1 kb fragment and the generation of two new fragments approximately 3.9 kb and 0.2 kb in size. The 0.2 kb fragment cannot be resolved by agarose gel electrophoresis. This morph was found in two Nama individuals who also have the *Hpa*I-7 morph. Since the nucleotide positions at which both the *Hpa*I and *Hae*II sites mutations occur are far apart, there is no indication that a mutation caused at the one site may lead to a concomitant change at the other nucleotide position. These Nama individuals also have the same *Msp*I mutation (described later).

*Hae*II-8 results when the 4.1 kb fragment found in *Hae*II-1 is lost due to the gain of a new *Hae*II site at position 9689 and the appearance of the 3.4 and 0.7 kb fragments (Fig. 3.3a). It has been suggested that this mutation results from an A to G transition mutation at this position (Bonne-Tamir *et al*. 1986). Morph *Hae*II-8 was previously found in one Israeli Arab individual and is found in the

Swazi (2.4%), Venda (6.7%) and Lemba (3.8%) ethnic groups.

Morph *Hae*II-9 can be derived from morph 5 as the result of a site loss at position 14858; the 1.2 kb fragment disappears and an intensified 1.4 kb fragment is evident (Fig. 3.3a). This morph was first observed in one Japanese individual (Horai *et al*. 1984) and is found at a frequency of 11.9% in the Dama group and at lower frequencies in the Zulu (3.3%), Swazi (2.4%) and Sotho (3.4%). This morph was not found in Senegalese Negroids (Scozzari *et al*. 1988) nor was it found in the South African Negroid group previously analysed by Johnson *et al*. (1983).

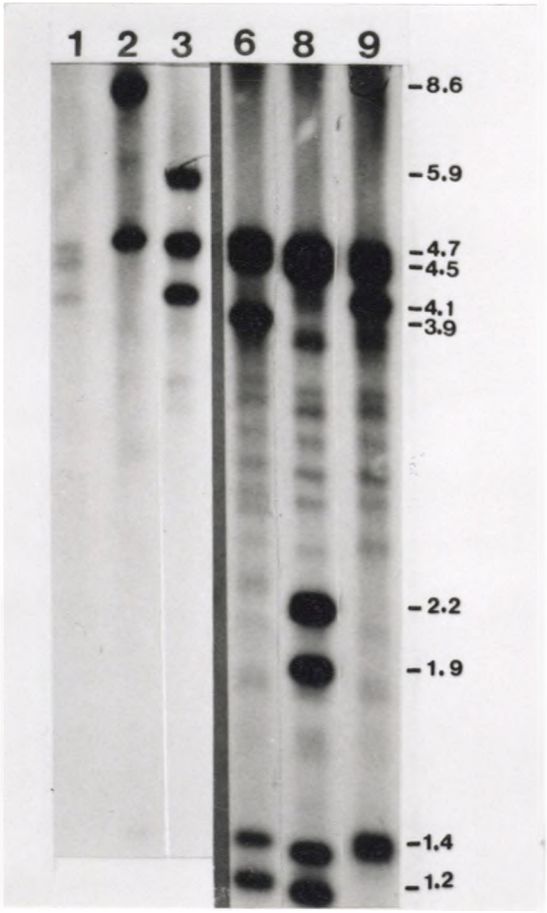


Fig. 3.3 (a). Fragment patterns of *Hae*HI morphs discovered in southern African populations (see text for details). Fragment sizes are given in kilobases.

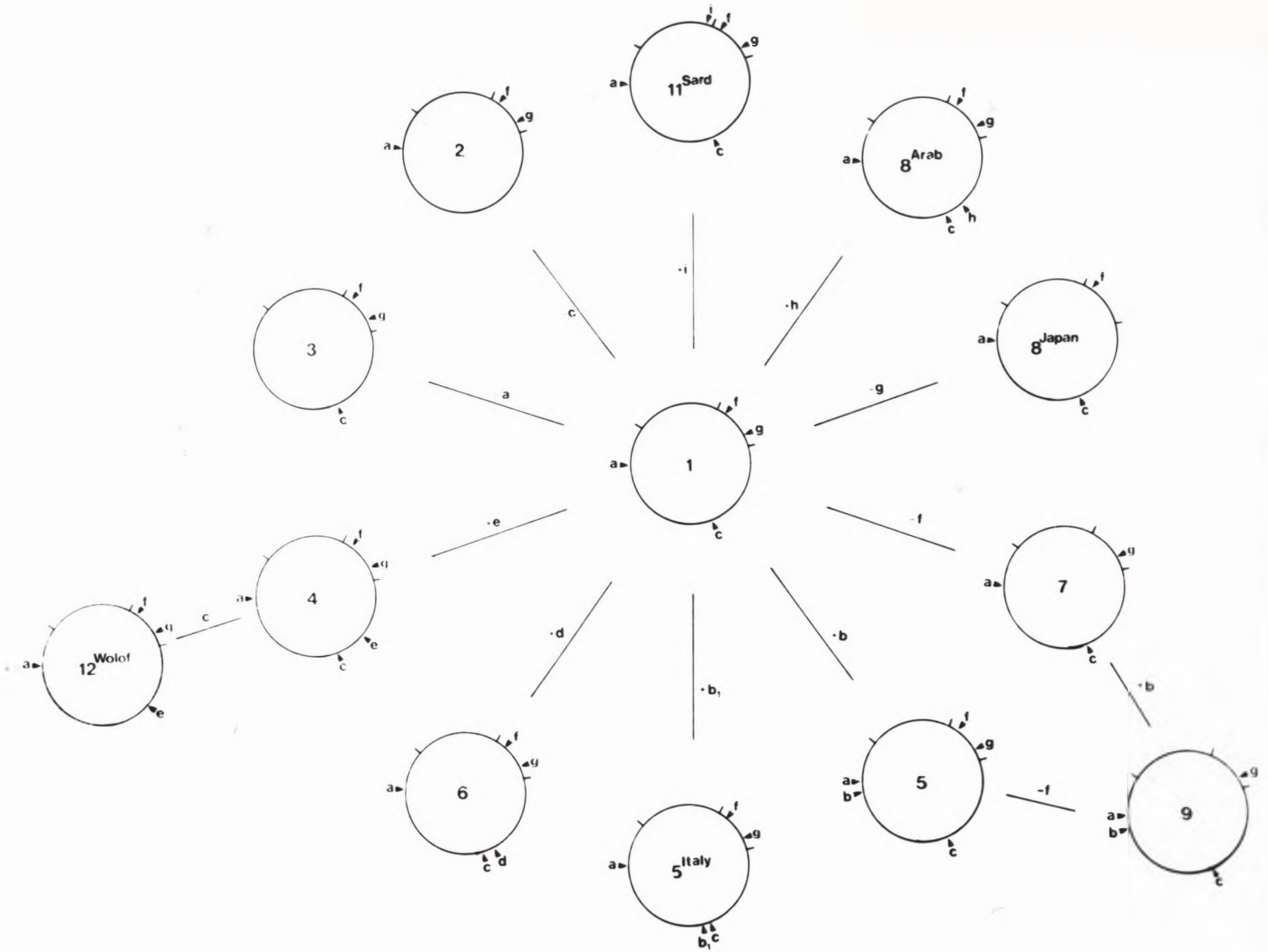


Fig. 3.3 (b). Phylogeny showing the relationship of the twelve *Hae*HI morphs discovered to date. Arrowheads with letters on circle indicate variable sites whilst lines on circle indicate the invariant sites on the molecule. Lines connecting circles show how each morph can be derived with the letters showing the sites either gained (+) or lost (-).

Table 3.3. The numbers and the frequencies (percentage) of *Hae*II morphs in twenty southern African populations. Standard errors are given in brackets.

		HaeII-1			HaeII-2			HaeII-3			HaeII-6			HaeII-8			HaeII-9		
	N	No	Freq	(S.E.)	No	Freq	(S.E.)	N	Freq	(S.E.)	No	Freq	(S.E.)	N	Freq	(S.E.)	No	Freq	(S.E.)
Zulu	30	29	96.7	3.3	-	-	-	-	-	-	-	-	-	-	-	-	1	3.3	3.3
Swazi	41	38	92.7	4.1	-	-	-	-	-	-	1	2.4	2.4	1	2.4	2.4	1	2.4	2.4
Xhosa	18	16	88.9	7.4	2	11.1	7.4	-	-	-	-	-	-	-	-	-	-	-	-
Tsonga	34	34	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sotho	29	28	96.6	3.4	-	-	-	-	-	-	-	-	-	-	-	-	1	3.4	3.4
Pedi	23	23	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tswana	15	14	93.3	6.5	1	6.7	6.5	-	-	-	-	-	-	-	-	-	-	-	-
Venda	30	28	93.3	4.6	-	-	-	-	-	-	-	-	-	2	6.7	4.5	-	-	-
Lemba	53	51	96.2	2.6	-	-	-	-	-	-	-	-	-	2	3.8	2.6	-	-	-
Herero	54	54	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Dama	43	38	88.4	4.9	-	-	-	-	-	-	-	-	-	-	-	-	5	12	4.9
Ambo	22	22	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sekele	49	59	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Nama	46	44	95.7	3.0	-	-	-	-	-	-	2	4.3	3.0	-	-	-	-	-	-
R_"Col"	35	35	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
JHB_"Col"	71	70	98.6	1.4	1	1.4	1.4	-	-	-	-	-	-	-	-	-	-	-	-
English	36	29	80.6	6.6	5	13.9	5.8	2	5.6	3.8	-	-	-	-	-	-	-	-	-
Afrikaners	51	50	98.0	2.0	1	2.0	2.0	-	-	-	-	-	-	-	-	-	-	-	-
A_Jews	55	37	67.3	6.3	17	30.9	6.2	1	1.8	1.8	-	-	-	-	-	-	-	-	-
Indian	50	50	100.0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

N = sample size
R_"Col" = Richtersveld "Coloured"
JHB_"Col" = Johannesburg "Coloured"
A_Jew = Ashkenazi Jewish

3.2.1.4 *Msp*I

This enzyme recognizes the sequence CCGG and there are twenty-three sites found on the reference sequence (Anderson *et al.* 1981). Nine of the sixteen previously described morphs (Johnson *et al.* 1983; Bònne-Tamir *et al.* 1986; Brega *et al.* 1986a,b; Scozzari *et al.* 1988; Semino *et al.* 1989; Semino *et al.* 1991) were observed and one new morph was discovered with this enzyme (Fig. 3.4a). The frequencies of the different morphs in the twenty ethnic groups studied in this investigation are shown in Table 3.4. A phylogeny showing the relationship of all morphs found in this study is shown in Fig. 3.4b.

*Msp*I-1 corresponds to the restriction enzyme pattern found in the reference sequence. Due to the limitations of agarose gel electrophoresis as employed in this study, fragments smaller than 450 bp could not be resolved (Fig. 3.4a). *Msp*I-1 is commonest in Caucasoid populations and found in low frequency in the Khoisan (Table 3.4). However, the frequencies of morph *Msp*I-1 also differed within Khoisan populations. Sekele have a higher frequency of *Msp*I-1 (45.8%) than the !Kung group (17.6%) and its frequency in the Nama is 34.8%.

Morph *Msp*I-2 is characterized by the fusion of the 1.1 and the 0.9 kb fragments to form the 2.04 kb fragment. It results from the loss of the *Msp*I site at position 8112

located within the gene coding for CO II (Anderson *et al.* 1981). *Msp*I-2 is the commonest morph in the Nama population and is found at a frequency of 60.9%. It is found at a lower frequency in the Sekele (16.9%) than the !Kung (52.9%). The higher frequency of this morph in the Richtersveld "Coloured" (82.9%) when compared to the JHB "Coloured" group (54.9%) suggests that there has been more contribution of female Nama genes into the former than the latter. The frequency of *Msp*-2 varies in the other populations studied (Table 3.4).

Morph 3 differs from morph 2 at a single site and is due to a C to G substitution at position 11457 contained within the ND4 gene. *Msp*I-3 is unique to the San and is found at a frequency of 35.6% in the Sekele and 26.5% in the !Kung. It is absent in the Nama and not present in Negroids or Caucasoids. It is observed in one other individual classified as "Coloured" in South Africa, in whom this morph may have been acquired due to gene flow from the San.

*Msp*I-4 may be derived from *Msp*I-1 due to a loss of a site at position 15925 which occurs within the region coding for threonine tRNA. It was only observed in European Caucasoids, albeit at low frequencies in this study (Table 3.4).

*Msp*I-5 may be derived from morph 2, in which a new site is

created within the 1.24 kb fragment, giving rise to fragments 0.95 and 0.29 kb in size (Fig. 3.4a). The new site occurs at position 13070 contained within the URF 5 gene. Only one individual from the Ambo group was found with this morph.

*Msp*I-7 is characterized by the gain of a site at position 15510 causing the 2.2 kb fragment observed in morph 1 to be lost and the generation of two fragments of 1.8 and 0.40 kb (Bònne-Tamir *et al.* 1986). This morph is rare, and was only found in one Zulu individual.

*Msp*I-12, recently found in two Wolof individuals from Senegal (Scozzari *et al.* 1988), is found in one San individual. This morph is characterized by the dissappearance of the 1.1 kb fragment observed in morph 1, and the appearance of two bands of about 0.65 and 0.5 kb. This new site was localized around positon 8646 within the ATPase 6 gene, and results from an A to G transition at this position (Scozzari *et al.* 1988).

*Msp*I-13, previously found in a Tukolor Negroid individual from Senegal (Scozzari *et al.* 1988), was found in some Negroid and Khoisan populations (Table 3.4). The gain of a site at position 11170 results in the loss of the 2.4 kb fragment observed in morph 1, and the generation of a 1.9 kb fragment and an enhanced 0.5 kb fragment (Fig. 3.4a). The two Nama individuals with this morph also had the *Hpa*I-

7 and *Hae*II-6 morphs. There is no connection between the *Msp*I-13 mutation and the *Hpa*I-7 and *Hae*II-6 mutations previously described.

One new morph (*Msp*I-17$^{S.A\ Jewish}$) was discovered with this enzyme. This results from the loss of a site at position 16453 causing the 0.53 kb and 0.11 kb fragments to fuse yielding the 0.64 kb fragment as seen in Fig. 3.4a. This new morph was found in two Ashkenazi Jewish individuals which suggests that it may have originated recently in the South African Ashkenazi Jewish population.

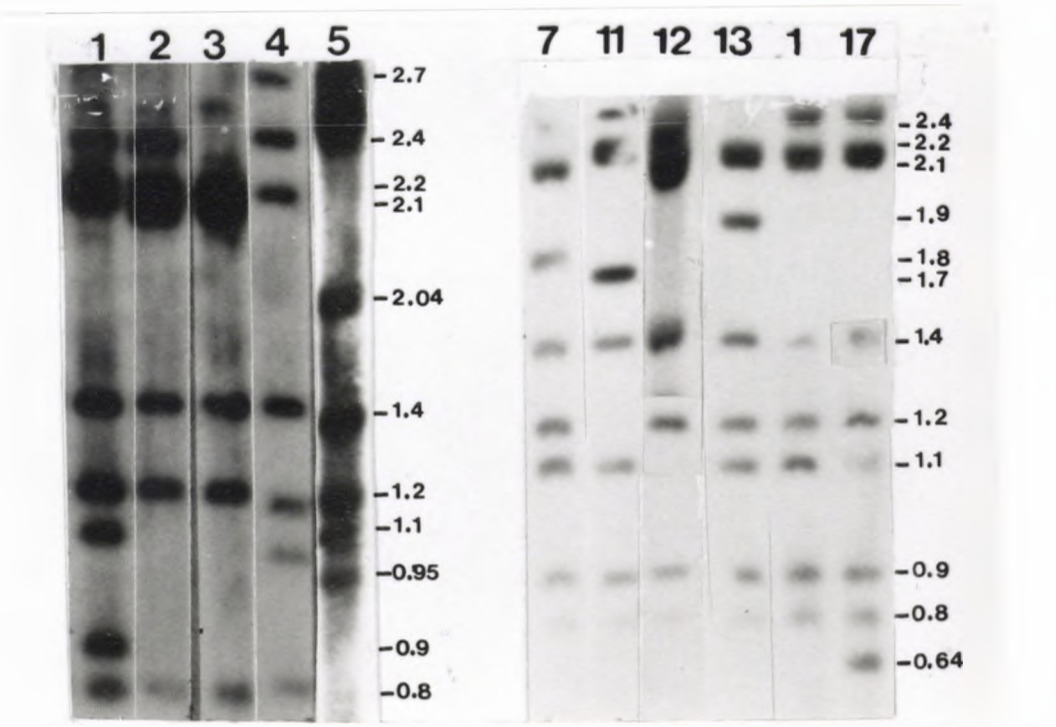


Fig. 3.4 (a). Fragment patterns of *Msp*I morphs discovered in southern African populations (see text for details). Morph *Msp*I-17 is new and was discovered in an individual from the South African Ashkenazi Jewish population. Fragment sizes are given in kilobases.

Fig. 3.4 (b). Phylogeny showing the relationship of the seventeen *Msp*I morphs discovered to date. Arrowheads with letters on circle indicate variable sites whilst lines on circle indicate the invariant sites on the molecule. Lines connecting circles show how each morph can be derived with the letters showing the sites either gained (+) or lost (-). The reference sequence (indicated by a star) has both sites a and b. It cannot be determined from RFLP analysis whether morph 1 has either site a or site b, therefore this is shown as a/b in the figure. Data obtained from Johnson *et al*. (1983), Bonne-Tamir *et al*. (1986), Brega *et al*. (1986a,b), Scozzari *et al*. (1988) and Semino *et al*. (1989, 1991).

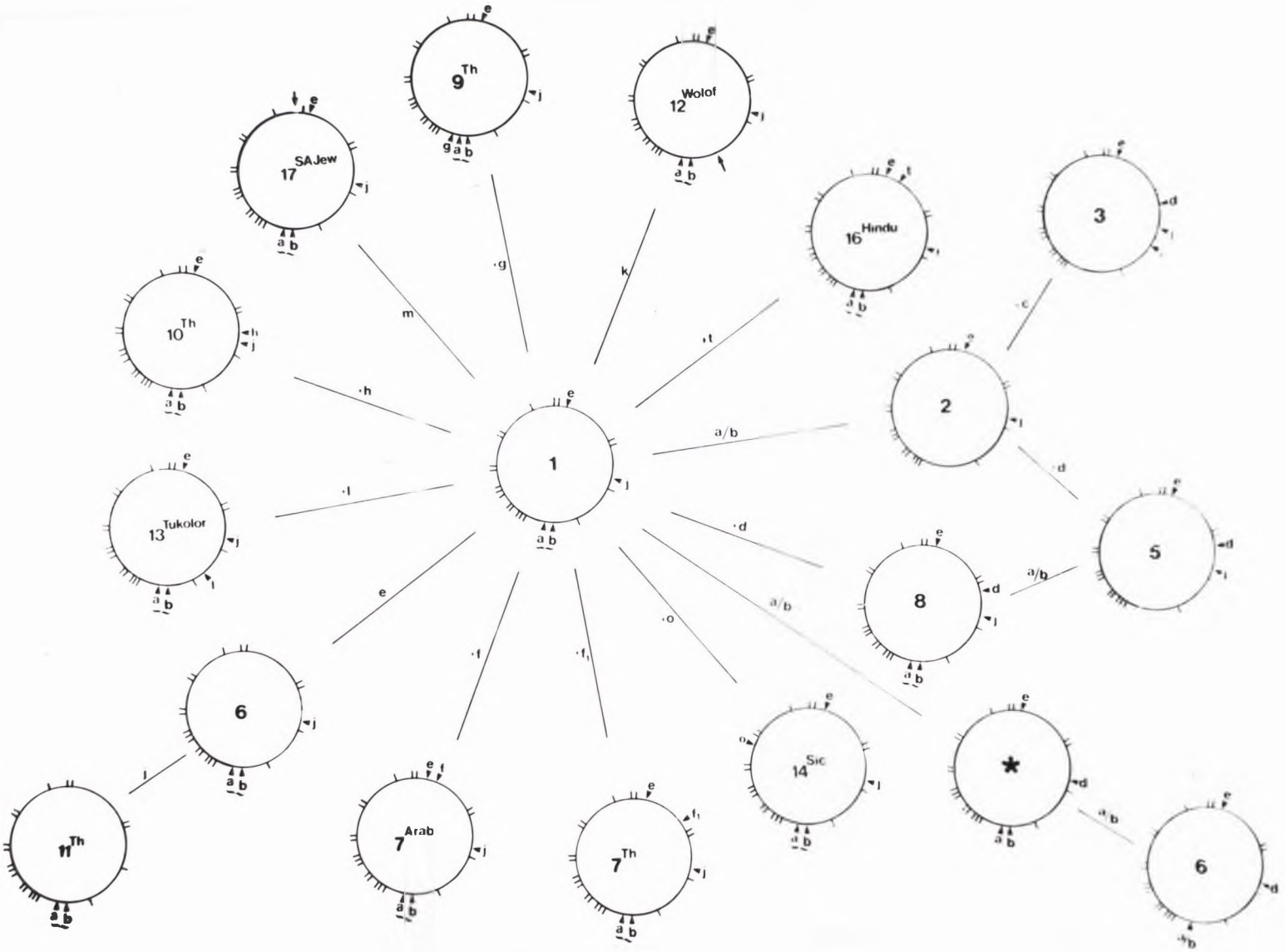

9 Th
12 Wolof
17 SAJew
16 Hindu
3
10 Th
2
1
13 Tukolor
8
5
6
11 Th
7 Arab
7 Th
14 Sic
*
6

Table 3.4. The numbers and the frequencies (percentage) of MspI morphs in twenty southern African populations. Standard errors given in brackets.

		MspI-1			MspI-2			MspI-3			MspI-4			MspI-5		
	N	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
Zulu	30	26	86.7	6.2	3	10.0	5.5	-	-	-	-	-	-	-	-	-
Swazi	41	38	92.7	4.1	3	7.3	4.1	-	-	-	-	-	-	-	-	-
Xhosa	18	16	88.9	7.4	2	11.1	7.4	-	-	-	-	-	-	-	-	-
Tsonga	34	31	91.2	4.9	1	2.9	2.9	-	-	-	-	-	-	-	-	-
Sotho	29	22	75.9	7.9	6	20.7	7.5	-	-	-	-	-	-	-	-	-
Pedi	23	18	78.3	8.6	5	21.7	8.9	-	-	-	-	-	-	-	-	-
Tswana	15	13	86.7	8.8	2	13.3	8.8	-	-	-	-	-	-	-	-	-
Venda	30	23	76.7	7.7	5	16.7	6.8	-	-	-	-	-	-	-	-	-
Lemba	53	44	83.0	5.2	8	15.1	4.9	-	-	-	-	-	-	-	-	-
Herero	54	49	90.7	4.0	5	9.3	4.0	-	-	-	-	-	-	-	-	-
Dama	43	31	72.1	6.9	8	18.6	5.9	-	-	-	-	-	-	-	-	-
Ambo	22	20	90.9	6.1	1	4.5	4.4	-	-	-	-	-	-	1	4.5	4.4
Sekele	49	19	38.8	7.0	8	16.3	5.3	21	35.6	6.8	-	-	-	-	-	-
Nama	46	16	34.8	7.0	28	60.9	7.2	-	-	-	-	-	-	-	-	-
R_"Col"	35	6	17.1	6.4	29	82.9	6.4	-	-	-	-	-	-	-	-	-
JHB_"Col"	71	31	43.7	5.9	39	54.9	5.9	1	1.4	1.4	-	-	-	-	-	-
English	36	32	88.9	5.2	-	-	-	-	-	-	4	11.1	5.2	-	-	-
Afrikaners	51	45	88.2	4.5	-	-	-	-	-	-	4	7.8	3.8	-	-	-
A_Jews	55	52	94.5	3.1	-	-	-	-	-	-	1	1.8	1.8	-	-	-
Indians	50	49	98.0	2.0	1	2.0	2.0	-	-	-	-	-	-	-	-	-

N = sample size

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

Table 3.4. continued

	N	Mspl-7			Mspl-11			Mspl-12			Mspl-13			Mspl-17		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
Zulu	30	1	3.3	3.3	-	-	-	-	-	-	-	-	-	-	-	-
Swazi	41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Xhosa	18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tsonga	34	-	-	-	1	2.9	2.9	-	-	-	1	2.9	2.9	-	-	-
Sotho	29	-	-	-	-	-	-	-	-	-	1	3.4	3.4	-	-	-
Pedi	23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tswana	15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Venda	30	-	-	-	-	-	-	-	-	-	2	6.7	4.6	-	-	-
Lemba	53	-	-	-	-	-	-	-	-	-	1	1.9	1.9	-	-	-
Herero	54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Dama	43	-	-	-	-	-	-	-	-	-	4	9.3	4.4	-	-	-
Ambo	22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sekele	49	-	-	-	-	-	-	1	1.7	1.9	-	-	-	-	-	-
Nama	46	-	-	-	-	-	-	-	-	-	2	4.3	3.0	-		-
R_"Col"	35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
JHB_"Col"	71	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
English	36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Afrikaners	51	-	-	-	2	3.9	2.8	-	-	-	-	-	-	-	-	-
A_Jews	55	-	-	-	-	-	-	-	-	-	-	-	-	2	3.6	2.5
Indians	50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

N = sample size

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

3.2.1.5 *Ava*II

This enzyme recognizes the sequence GG(A/T)CC and there are eight *Ava*II sites on the published reference sequence (Anderson *et al*. 1981). This enzyme detects the greatest amount of variation between mtDNA molecules in humans and, to date, thirty different morphs have been reported (Johnson *et al*. 1983; Horai *et al*. 1984; Wallace *et al*. 1985; Brega *et al*. 1986a,b; Santachiara-Benerecetti *et al*. 1988; Scozzari *et al*. 1988; Torroni *et al*. 1990; Semino *et al*. 1991). Of these only eight were found in this study (Table 3.5). In addition, three new morphs were discovered (Fig. 3.5a).

*Ava*II-1 which corresponds to the reference sequence is characterized by the presence of eight fragments, of which only the 9.8, 3.0, 1.0, 0.8 and 0.7 kb fragments can be resolved by agarose gel electrophoresis (Fig. 3.5a). It is commoner in Caucasoids than in Negroid or Khoisan populations (Table 3.5). *Ava*II-1 is found at an appreciably higher proportion in the Sekele (12/49) than in the !Kung (4/34), but these differences are not significant [X^2_1 = 2.09, p = 0.148; Yates correction X^2_1 =1.35, p = 0.245].

*Ava*II-2 may be derived from morph 1 which results from a site gain at position 8229 localized within the CO II coding region. It results in the loss of the 9.8 kb

fragment observed in morph 1 and the generation of two bands of approximately 5.4 and 4.4 kb in size (Fig. 3.5a). It is the commonest morph in both the Sekele (64.4%) and the Nama (60.9%) groups, with a comparable frequency in the !Kung (58.8%). On the contrary, Negroid populations have a low frequency of this morph [1.1% in Senegalese (Scozzari *et al.* 1988); 12.5% in Negroids from South Africa (Johnson *et al.* 1983) and varying frequencies in Negroid populations investigated in this study (see Table 3.5)]. *Ava*II-2 occurs only rarely in Caucasoids (Table 3.5).

*Ava*II-3 may be derived from *Ava*II-1 by the loss of a site at position 16390 within the D-loop region of the molecule. It is identified by the fusion of the 3.0 and 0.8 kb fragments to form the 3.8 kb fragment (see Fig. 3.5a). Whereas *Ava*II-3 is found at varying frequencies in southeastern Bantu speaking Negroids, southwestern Bantu-speakers (Herero) and Dama (Khoisan-speaking Negroids) lack this morph (Table 3.5). It is found at a frequency of 13.6% in the sample of Ambo who are southwestern Bantu-speakers, suggesting that they appear to be more closely related to southeastern than to southwestern Bantu-speakers. In addition, although the frequency of *Ava*II-1 is lower in the Dama (41.9%) and Herero (33.3%), its frequency in those branches of the Ambo included in this series (77.3%) is comparable to southeastern Bantu-speakers (Table 3.5), providing further support for the suggestion by Nurse (1983) that some sections of the Ambo are more

closely related to southeastern Bantu-speakers than to southwestern Bantu-speakers at the genetic level. The frequency of morph *Ava*II-3 is found to be lower in the Khoisan, except for the Sekele in whom it was found at a frequency of 30.5%. This is another piece of evidence to suggest that there has been more gene flow from Negroids into the Sekele than into the !Kung.

*Ava*II-5 appears to be a combination of morphs 2 and 3, having gained a site at position 8275 (as in morph 2) and having lost the *Ava*II site at position 16390 (as in morph 3). *Ava*II-5 is not found in Caucasoid populations but is found in Negroid and Khoisan populations at varying frequencies (Table 3.5).

In *Ava*II-6, a new site is found within the 3.0 kb fragment observed in morph 1, producing the new fragments of 2.5 and 0.5 kb in size. This new site has been localized near position 15890 and could occur within either the threonine tRNA, the proline tRNA, or cytochrome b genes. The exact position of this site has yet to be verified, possibly by sequencing (Johnson *et al*. 1983). *Ava*II-6 is found at a frequency of 14% in the Dama, and except for one Tsonga individual, it is not found in other Negroid groups. It is also found in one Nama individual and one Indian individual. It is possible that the Nama could have acquired this morph from the Dama due to gene flow from Dama females into the Nama group.

*Ava*II-8 is derived from morph *Ava*II-1 due to a site gain in the 9.8 kb fragment generating two new fragments of 7.8 and 2.0 kb in size. This site has been localized to position 4776 which occurs within the ND2 gene (Johnson *et al.* 1983). It is a rare morph and is found only in one Dama individual in this investigation.

*Ava*II-9 differs from morph 1 by the fusion of the two contiguous fragments of 3.0 and 0.7 kb in size, to form a new band of 3.7 kb. This morph could arise in two ways, either from morph 1 by a site loss at position 13367, or from morph 7, by a site gain at position 16390 separating the 0.8 kb fragment from the fused 3.0 and 0.7 kb fragments. Nucleotide position 13367 is located within the URF 5 gene, whereas 16390 is in the D-loop (Johnson *et al.* 1983). This morph was only found in one individual from the Jewish group.

*Ava*II-11 appears to be derived from the combination of morphs 2 and 6. As in morph 2, it has a new site at about position 8229 creating the 5.4 and 4.4 kb fragments and like morph 6, the 2.5 and 0.5 kb fragments are generated from the gain of a site localized at position 15870 (Fig. 3.5a). This morph is rare and is only found in one Lemba individual in this study.

Morph *Ava*II-15 results from a site gained within the 9.8 kb fragment generating the 8.4 and 1.5 kb fragments observed

in Fig. 3.5a. It is caused by a mutation at position 4280 giving rise to a new *Ava*II site. This morph was first found in an Israeli Jewish individual (Bonne-Tamir *et al.* 1986) and was found in one Herero individual in this study.

Three new morphs were discovered in the Johannesburg "Coloured" population sampled and were allocated numbers *Ava*II-31$^{S\,A\,Col}$, *Ava*II-32$^{S\,A\,Col}$ and *Ava*II-33$^{S\,A\,Col}$. In *Ava*II-31$^{S\,A\,Col}$, the 4.4 kb and 0.7 kb fragments found in morph 2 (Fig. 3.5a) fuse to give fragment 5.1 kb. This results from the loss of the site 12629 which occurs within the presumptive protein 5 gene. *Ava*II-32$^{S\,A\,Col}$ is also a derivative of morph 2 and results from a site gained at position 10774 located within the presumptive protein 4 gene. This results in the loss of the 4.4 kb fragment and the generation of two new fragments of sizes 2.5 and 1.9 kb. The third new morph, *Ava*II-33$^{S\,A\,Col}$, also a derivative of morph 2, results from a site gain at position 5165 causing the 5.4 kb fragment to be lost, thereby generating fragments 3.0 and 2.4 kb in size. This site occurs within the presumptive coding 2 gene.

Fig. 3.5 (a). Fragment patterns of *Ava*II morphs discovered in southern African populations. Morphs 31, 32 and 33 are new and were discovered in the South African "Coloured" population (see text for details).

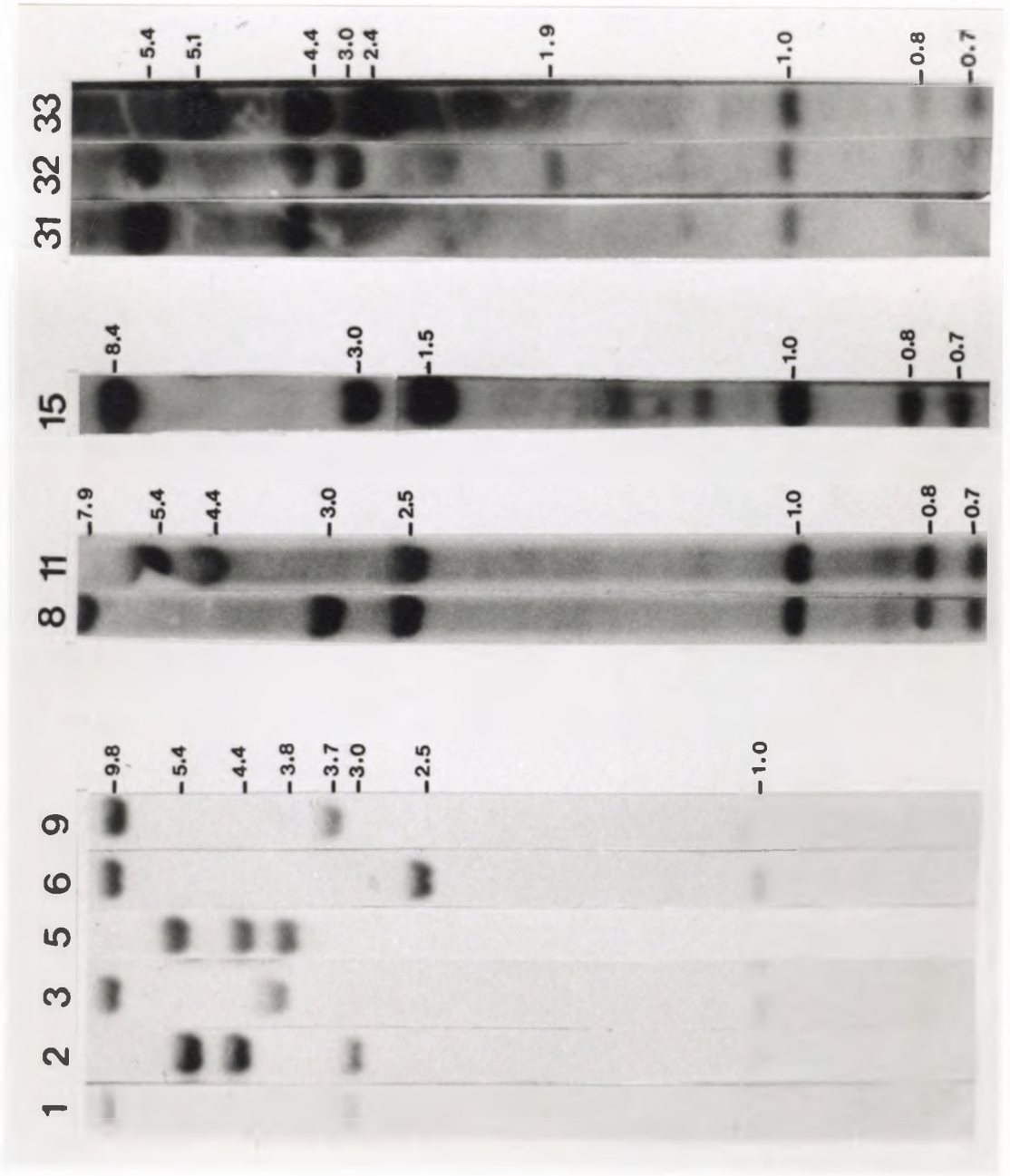

1 2 3 5 6 9
9.8
5.4
4.4
3.8
3.7
3.0
2.5
1.0
8 11
7.9
5.4
4.4
3.0
2.5
1.0
0.8
0.7
15
8.4
3.0
1.5
1.0
0.8
0.7
31 32 33
5.4
5.1
4.4
3.0
2.4
1.9
1.0
0.8
0.7

Fig. 3.5 (b). Phylogeny showing the relationship of the all *Ava*II morphs discovered to date. Arrowheads with letters on circle indicate variable sites whilst lines on circle indicate the invariant sites on the molecule. Lines connecting circles show how each morph can be derived with the letters showing the sites either gained (+) or lost (-). Data obtained from Johnson *et al*. (1983), Horai *et al*. (1984), Wallace *et al*. (1985), Brega *et al*. (1986a,b), Santachiara-Benerecetti *et al*. (1988), Scozzari *et al*. (1988), Torroni *et al*. (1990), Semino *et al*. (1991) and the present study.

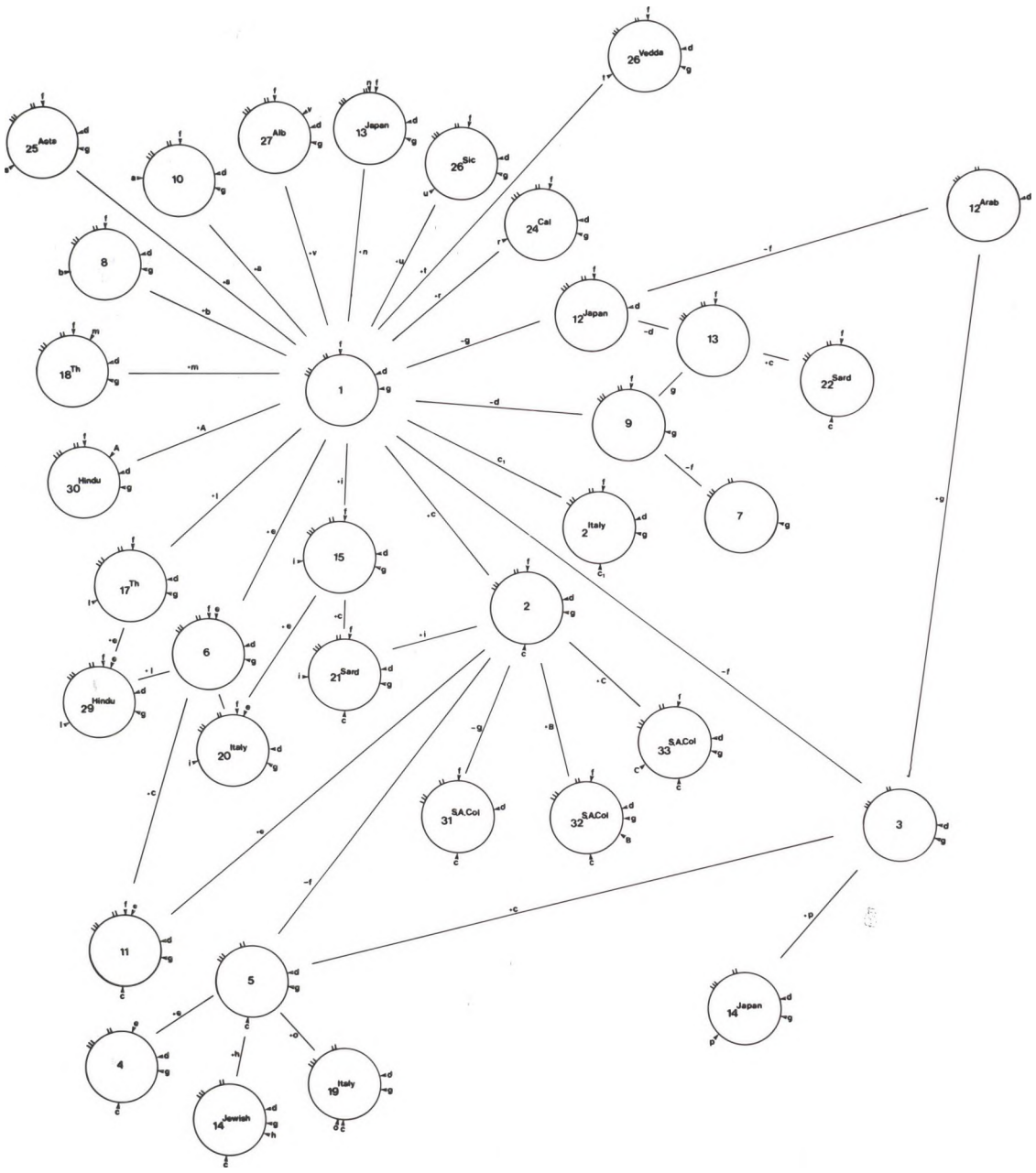

26 Vedda
25 Aete
27 Alb
13 Japan
28 Sic
10
24 Cal
12 Arab
8
12 Japan
13
18 Th
1
22 Sard
9
30 Hindu
2 Italy
7
17 Th
15
2
6
21 Sard
29 Hindu
20 Italy
33 S.A.Col
31 S.A.Col
32 S.A.Col
3
11
5
14 Japan
4
14 Jewish
19 Italy

Table 3.5. The numbers and the frequencies (percentage) of *Ava*II morphs in twenty southern African populations. Standard errors are given in brackets.

		AvaII-1			AvaII-2			AvaII-3			AvaII-5			AvaII-6			AvaII-8		
	N	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	S.E.)	No	Freq	(S.E.)
Zulu	30	16	53.3	9.1	4	13.3	6.2	10	33.3	8.6	-	-	-	-	-	-	-	-	-
Swazi	41	29	70.7	7.1	3	7.3	4.1	8	19.5	6.2	2	4.9	3.4	-	-	-	-	-	-
Xhosa	18	11	61.1	11.5	2	11.1	7.4	3	16.7	8.8	2	11.1	7.4	-	-	-	-	-	-
Tsonga	34	19	55.9	8.5	1	2.9	2.9	13	38.2	8.3	-	-	-	1	2.9	2.9	-	-	-
Sotho	29	14	48.3	9.3	9	31.0	8.6	5	17.2	7.0	1	3.4	3.4	-	-	-	-	-	-
Pedi	23	13	56.5	10.3	4	17.4	7.9	4	17.4	7.9	2	8.7	5.9	-	-	-	-	-	-
Tswana	15	8	53.3	12.9	1	6.7	6.5	4	26.7	11.4	2	13.3	8.8	-	-	-	-	-	-
Venda	30	11	36.7	8.9	6	20.0	7.3	13	43.3	9.0	-	-	-	-	-	-	-	-	-
Lemba	53	28	52.8	6.9	9	17.0	5.2	11	20.8	5.6	3	5.7	3.2	-	-	-	-	-	-
Herero	54	18	33.3	6.4	33	61.1	6.6	-	-	-	2	3.7	2.6	-	-	-	-	-	-
Dama	43	18	41.9	7.5	17	39.5	7.5	-	-	-	1	2.3	2.3	6	14.0	5.3	1	2.3	2.3
Ambo	22	17	77.3	8.9	2	9.1	6.1	3	-	-	-	-	-	-	-	-	-	-	-
Sekele	49	12	24.5	6.1	36	73.5	6.3	-	-	-	1	1.7	1.8	-	-	-	-	-	-
Nama	46	8	17.4	5.6	34	73.9	6.5	-	-	-	3	6.5	3.6	1	2.2	2.2	-	-	-
R_"Col"	35	5	14.3	5.9	25	71.4	7.6	-	-	-	5	14.3	5.9	-	-	-	-	-	-
JHB_"Col"	71	23	32.4	5.6	19	26.8	5.3	13	18.3	4.6	13	18.3	4.6	-	-	-	-	-	-
English	36	25	69.4	7.7	-	-	-	6	16.7	6.2	5	13.9	5.8	-	-	-	-	-	-
Afrikaners	51	43	84.3	5.1	3	5.9	3.3	5	9.8	4.2	-	-	-	-	-	-	-	-	-
A_Jews	55	54	98.2	1.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Indians	50	44	88.0	4.6	2	4.0	2.8	3	6.0	3.4	-	-	-	1	2.0	2.0	-	-	-

N = sample size

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

Table 3.5. (continued)

		Avall-9			Avall-11			Avall-15			Avall-31			Avall-32			Avall-33		
	N	No	Freq	S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
Zulu	30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Swazi	41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Xhosa	18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tsonga	34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sotho	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pedi	23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tswana	15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Venda	30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Lemba	53	-	-	-	1	1.9	1.9	-	-	-	-	-	-	-	-	-	-	-	-
Herero	54	-	-	-	-	-	-	1	1.9	1.9	-	-	-	-	-	-	-	-	-
Dama	43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ambo	22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sekele	49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Nama	46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
R_"Coi"	35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
JHB_"Col"	71	-	-	-	-	-	-	-	-	-	1	1.4	1.4	1	1.4	1.4	1	1.4	1.4
English	36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Afrikaners	51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A_Jews	55	1	1.8	1.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Indians	50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

N = sample size

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

3.2.1.6 *Hinc*II

This enzyme recognizes the sequence GTYRAC, where Y is either C or T and R is either G or A. This enzyme recognition sequence is closely associated with *Hpa*I which recognizes the sequence GTTAAC. Therefore all *Hpa*I sites are also *Hinc*II sites, but the converse does not apply. To date sixteen different *Hinc*II morphs have been identified (Blanc *et al.* 1983; Horai *et al.* 1984; Wallace *et al.* 1985; Brega *et al.* 1986a,b; Santachiara-Benerecetti *et al.* 1988; Scozzari *et al.* 1988; Semino *et al.* 1991) of which only four are found in this investigation (Fig. 3.6a). The frequency of the various morphs in the twenty groups is shown in Table 3.6, and a phylogeny relating all four morphs is shown in Fig. 3.6b.

*Hinc*II-2 is the commonest morph observed in this study. It is characterized by the presence of eleven fragments that correspond to the sites found in the reference sequence (see Fig. 3.6a). Except for the Swazi, Herero, Sekele and Richtersveld "Coloured" populations, all other groups are monomorphic for *Hinc*II-2.

*Hinc*II-1 differs from *Hinc*II-2 in the same region that the *Hpa*I site *c* is located. This site which occurs at position 12406 within the presumptive protein 5 gene is lost causing the 2.3 and 0.8 kb fragments seen in morph *Hinc*II-2 to fuse, thereby generating the 3.2 kb fragment.

*Hinc*II-1 is rare in southern African populations and is only found in one Swazi individual. This morph was previously encountered in Orientals at a frequency of 20% (Blanc *et al*. 1983), is not found in Caucasoids (Blanc *et al*. 1983, Semino *et al*. 1989) and is rare in Negroids (1.6% in Senegalese; Scozzari *et al*. 1988). The observation that this morph was also found in chimpanzees and orangutans, led to the suggestion that this morph was ancestral to other human mtDNAs (Blanc *et al*. 1983). This observation provided additional support for the hypothesis suggested by Denaro *et al*. (1981) that human mtDNA evolved in Asia.

*Hinc*II-5 differs from *Hinc*II-2 by the loss of fragments 1.9 kb and 0.1 kb caused by the loss of the site at position 7853 which is located within the cytochrome oxidase subunit 2 gene and the acquisition of the new fragment 2.0 kb in size. *Hinc*II-5 was found in the Herero and Richtersveld "Coloured" populations at frequencies of 1.9% and 11.4%, respectively. It was not found in other Negroids (Scozzari *et al*. 1988), but was found at low frequencies in Caucasoids (1.6%) and Orientals (1.8%) (Blanc *et al*. 1983).

*Hinc*II-9 differs from morph 2 by a single mutation due to the loss of the site at position 13259 which is located within the URF 5 gene. In this morph the two contiguous fragments 0.85 and 0.38 kb fuse to form the 1.2 kb fragment (Horai *et al*. 1984). This morph first found in Orientals (1.7%), is rarely found in individuals studied (Table 3.6).

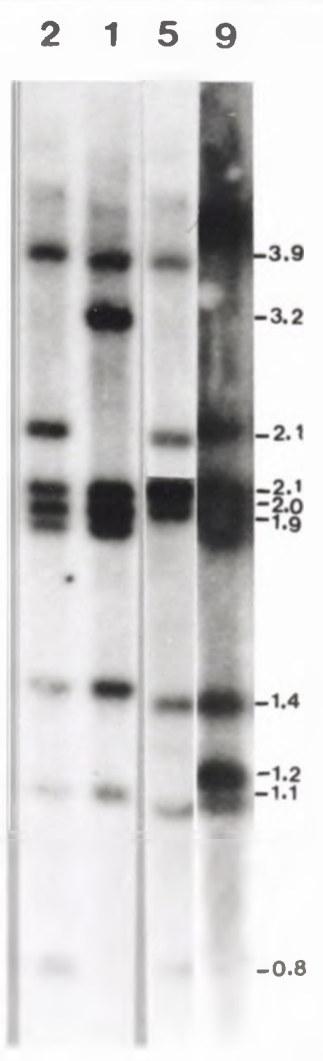


Fig. 3.6 (a). Fragment patterns of four *Hinc*II morphs discovered in southern African populations (see text for details).

Fig. 3.6 (b). Phylogeny showing the relationship of all *Hinc*II morphs discovered to date. Arrowheads with letters on circle indicate variable sites whilst lines on circle indicate the invariant sites on the molecule. Lines connecting circles show how each morph can be derived with the letters showing the sites either gained (+) or lost (-). Data obtained from Blanc *et al.* (1983), Horai *et al.* (1984), Wallace *et al.* (1985), Brega *et al.* (1986a,b), Santachiara-Benerecetti *et al.* (1988), Scozzari *et al.* (1988) and Semino *et al.* (1991).

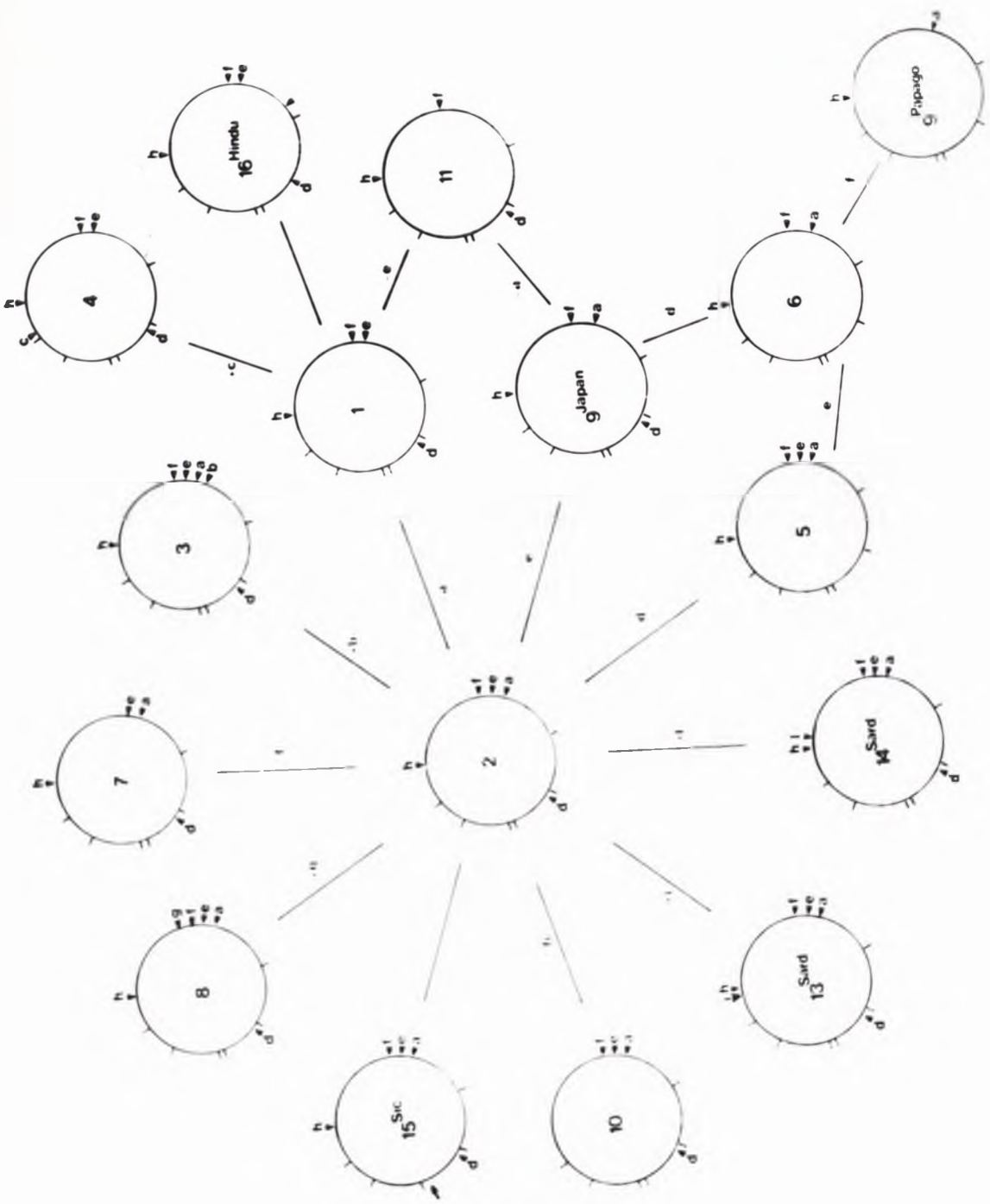

16 Hindu
11
9 Pupango
4
6
1
9 Japan
3
5
7
2
14 Sard
8
13 Sard
15 Src
10

Table 3.6. The numbers and the frequencies (percentage) of *Hinc* II morphs in twenty southern African populations. Standard errors given in brackets.

		HincII-1			HincII-2			HincII-5			HincII-9		
	N	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
Zulu	30	-	-	-	30	100.0	-	-	-	-	-	-	-
Swazi	41	1	2.4	2.4	39	95.1	3.4	-	-	-	1	2.4	2.4
Xhosa	18	-	-	-	18	100.0	-	-	-	-	-	-	-
Tsonga	34	-	-	-	34	100.0	-	-	-	-	-	-	-
Sotho	29	-	-	-	29	100.0	-	-	-	-	-	-	-
Pedi	23	-	-	-	23	100.0	-	-	-	-	-	-	-
Tswana	15	-	-	-	15	100.0	-	-	-	-	-	-	-
Venda	30	-	-	-	30	100.0	-	-	-	-	-	-	-
Lemba	53	-	-	-	53	100.0	-	-	-	-	-	-	-
Herero	54	-	-	-	53	98.1	1.9	1	1.9	1.9	-	-	-
Dama	43	-	-	-	43	100.0	-	-	-	-	-	-	-
Ambo	22	-	-	-	22	100.0	-	-	-	-	-	-	-
Sekele	49	-	-	-	48	98.0	2.0	-	-	-	1	2.0	2.0
Nama	46	-	-	-	46	100.0	0	-	-	-	-	-	-
R_"Col"	35	-	-	-	31	88.6	5.4	4	11.4	5.4	-	-	-
JHB_"Col"	71	-	-	-	71	100.0	-	-	-	-	-	-	-
English	36	-	-	-	36	100.0	-	-	-	-	-	-	-
Afrikaners	51	-	-	-	51	100.0	-	-	-	-	-	-	-
A_Jews	55	-	-	-	55	100.0	-	-	-	-	-	-	-
Indians	50	-	-	-	50	100.0	-	-	-	-	-	-	-

N = sample size
R_"Col" = Richtersveld "Coloured"
JHB_"Col" = Johannesburg "Coloured"
A_Jew = Ashkenazi Jewish

3.2.2 MtDNA types

The combination of the morphs denoted by various enzymes in the order *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II gives rise to a mtDNA type (mitotype) for each individual. Each unique mtDNA type is assigned a number in chronological order as it is discovered. To date, 119 mtDNA types have been reported (Merriwether *et al*. 1991). Of the 62 mtDNA types discovered in southern African populations, only 32 have been found before whilst other 30 are new. MtDNA type numbers 113 to 142 have been assigned to these new mtDNA types, contributing to a mtDNA pool of 149 mtDNA types discovered to date by this method of analysis.

Twenty-six out of the 30 new mtDNA types resulted from new combinations of previously existing morphs, whereas only 4 were due to the discovery of new morphs. The mtDNA morphs for the six enzymes used and the accompaning mtDNA type for each individual is given in Appendix 3. The frequency of mtDNA types in the 20 ethnic groups studied is given in Table 3.7.

Table 3.7. The numbers and the frequencies of mtDNA types discovered in twenty southern African populations. MtDNA type numbers were chronologically assigned to the unique combination of the enzyme morphs obtained for the restriction enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II in this order (Blanc *et al.* 1983; Johnson *et al.* 1983). Population abbreviations: R_"Col" = Richtersveld "Coloured", JHB_"Col" = Johannesburg "Coloured" and A_Jew = Ashkenazi Jewish.

Type	Mitotype	Zulu N=30			Swazi N=41			Xhosa N=18		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	5	16.7	6.8	12	29.3	7.1	3	16.7	8.8
1-9	2-1-1-1-1-9	-	-	-	1	2.4	2.4	-	-	-
2-2	3-1-1-1-3-2	9	30.0	8.4	7	17.7	6.0	3	16.7	8.8
3-2	3-1-1-2-2-2	3	10.0	5.5	1	2.4	2.4	1	5.6	5.4
3-5	3-1-1-2-2-5	-	-	-	-	-	-	-	-	-
4-2	3-1-1-3-2-2	-	-	-	-	-	-	-	-	-
4-9	3-1-1-3-2-9	-	-	-	-	-	-	-	-	-
5-2	3-1-1-2-5-2	-	-	-	1	2.4	2.4	1	5.6	5.4
5-5	3-1-1-2-5-5	-	-	-	-	-	-	-	-	-
6-2	2-1-2-1-1-2	-	-	-	-	-	-	1	5.6	5.4
7-2	3-1-1-1-1-2	9	30.0	8.4	10	24.4	6.7	6	33.3	11.1
8-2	1-1-1-1-1-2	1	3.3	3.3	-	-	-	1	5.6	5.4
8-1	1-1-1-1-1-1	-	-	-	1	2.4	2.4	-	-	-
9-2	1-1-2-1-1-2	-	-	-	-	-	-	-	-	-
10-2	3-1-1-1-2-2	-	-	-	2	2.9	2.6	-	-	-
11-2	2-2-3-1-5-2	-	-	-	-	-	-	-	-	-
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-
21-2	2-1-1-1-2-2	1	3.3	3.3	-	-	-	-	-	-
21-5	2-1-1-1-2-5	-	-	-	-	-	-	-	-	-
22-2	2-1-1-1-5-2	-	-	-	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-	-	-	-
26-2	2-1-6-1-1-2	-	-	-	1	2.4	2.4	-	-	-
29-2	2-1-1-4-1-2	-	-	-	-	-	-	-	-	-
30-2	3-1-1-1-11-2	-	-	-	-	-	-	-	-	-
31-2	3-1-1-1-5-1	-	-	-	1	2.4	2.4	1	5.6	5.4
33-2	3-1-1-2-3-2	-	-	-	-	-	-	-	-	-
44-2	2-1-1-4-3-2	-	-	-	-	-	-	-	-	-
45-2	2-1-8-1-1-2	-	-	-	1	2.4	2.4	-	-	-
47-2	2-1-1-1-3-2	-	-	-	1	2.4	2.4	-	-	-
51-2	2-1-1-11-1-2	-	-	-	-	-	-	-	-	-
56-2	2-1-1-1-6-2	-	-	-	-	-	-	-	-	-
70-2	2-1-1-13-1-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Zulu N=30			Swazi N=41			Xhosa N=18		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
113-2	3-1-1-12-1-2	-	-	-	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	-	-	-	-	-	-
115-2	3-1-1-2-31-2	-	-	-	-	-	-	-	-	-
116-2	2-1-1-2-2-2	-	-	-	-	-	-	-	-	-
117-2	3-1-1-1-15-2	-	-	-	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	-	-	-	-	-	-
119-2	3-1-1-1-8-2	-	-	-	-	-	-	-	-	-
120-2	3-1-1-3-32-2	-	-	-	-	-	-	-	-	-
121-2	3-1-1-2-33-2	-	-	-	-	-	-	-	-	-
122-2	2-1-8-1-3-2	-	-	-	-	-	-	-	-	-
123-2	3-1-1-13-3-2	-	-	-	-	-	-	-	-	-
124-2	2-1-1-2-1-2	-	-	-	1	2.4	2.4	-	-	-
125-2	3-1-1-11-1-2	-	-	-	-	-	-	-	-	-
126-2	2-1-1-17-1-2	-	-	-	-	-	-	-	-	-
127-2	2-1-2-17-1-2	-	-	-	-	-	-	-	-	-
128-2	2-2-3-1-1-2	-	-	-	-	-	-	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-	-	-	-
130-2	3-1-9-1-1-2	-	-	-	1	2.4	2.4	-	-	-
131-2	3-1-2-1-1-2	-	-	-	-	-	-	1	5.6	5.4
132-2	3-1-1-7-1-2	1	3.3	3.3	-	-	-	-	-	-
133-2	3-1-9-1-3-2	1	3.3	3.3	-	-	-	-	-	-
134-2	3-1-1-2-1-2	-	-	-	-	-	-	-	-	-
135-2	3-1-1-13-2-2	-	-	-	-	-	-	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-	-	-	-
137-2	3-1-1-2-11-2	-	-	-	-	-	-	-	-	-
138-2	3-1-9-1-8-2	-	-	-	-	-	-	-	-	-
139-2	7-1-6-13-1-2	-	-	-	-	-	-	-	-	-
140-2	3-1-1-1-6-2	-	-	-	-	-	-	-	-	-
141-2	2-2-1-4-3-2	-	-	-	-	-	-	-	-	-
142-2	3-2-1-11-2-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Tsonga N=34			Sotho N=29			Pedi N=23		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	9	26.5	7.6	5	17.2	7.0	2	8.7	5.9
1-9	2-1-1-1-1-9	-	-	-	-	-	-	-	-	-
2-2	3-1-1-1-3-2	13	38.2	8.3	5	17.2	7.0	4	17.4	7.9
3-2	3-1-1-2-2-2	1	2.9	2.9	4	13.8	6.4	4	17.4	7.9
3-5	3-1-1-2-2-5	-	-	-	-	-	-	-	-	-
4-2	3-1-1-3-2-2	-	-	-	-	-	-	-	-	-
4-9	3-1-1-3-2-9	-	-	-	-	-	-	-	-	-
5-2	3-1-1-2 5-2	-	-	-	1	3.4	3.4	1	4.3	4.2
5 5	3 1 1 2 5 5									
6-2	2-1-2-1-1-2	-	-	-	-	-	-	-	-	-
7-2	3-1-1-1-1-2	8	23.5	7.3	7	24.1	7.9	11	47.8	10.5
8-2	1-1-1-1-1-2	-	-	-	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-	-	-	-
9-2	1-1-2-1-1-2	-	-	-	-	-	-	-	-	-
10-2	3-1-1-1-2-2	-	-	-	3	10.3	5.6	-	-	-
11-2	2-2-3-1-5-2	-	-	-	-	-	-	-	-	-
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-
21-2	2-1-1-1-2-2	-	-	-	1	3.4	3.4	-	-	-
21-5	2-1-1-1-2-5	-	-	-	-	-	-	-	-	-
22-2	2-1-1-1-5-2	-	-	-	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-	-	-	-
26-2	2-1-6-1-1-2	-	-	-	-	-	-	-	-	-
29-2	2-1-1-4-1-2	-	-	-	-	-	-	-	-	-
30-2	3-1-1-1-11-2	-	-	-	-	-	-	-	-	-
31-2	3-1-1-1-5-1	-	-	-	-	-	-	1	4.3	4.2
33-2	3-1-1-2-3-2	-	-	-	-	-	-	-	-	-
44-2	2-1-1-4-3-2	-	-	-	-	-	-	-	-	-
45 2	2 1 8 1 1 2									
47 2	2 1 1 1-3 2				-	-				
51-2	2-1-1-11-1-2	-	-	-	-	-	-	-	-	-
56-2	2-1-1 1-6-2	1	2.9	2.9	-	-	-	-	-	-
70-2	2-1-1-13-1-2	1	2.9	2.9	-	-	-	-	-	-

Type	Mitotype	Tsonga N = 34			Sotho N = 29			Pedi N = 23		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
113-2	3-1-1-12-1-2	-	-	-	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	-	-	-	-	-	-
115-2	3-1-1-2-31-2	-	-	-	-	-	-	-	-	-
116-2	2-1-1-2-2-2	-	-	-	-	-	-	-	-	-
117-2	3-1-1-1-15-2	-	-	-	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	-	-	-	-	-	-
119-2	3-1-1-1-8-2	-	-	-	-	-	-	-	-	-
120-2	3-1-1-3-32-2	-	-	-	-	-	-	-	-	-
121-2	3-1-1-2-33-2	-	-	-	-	-	-	-	-	-
122-2	2-1-8-1-3-2	-	-	-	-	-	-	-	-	-
123-2	3-1-1-13-3-2	-	-	-	-	-	-	-	-	-
124-2	2-1-1-2-1-2	-	-	-	1	3.4	3.4	-	-	-
125-2	3-1-1-11-1-2	1	2.9	2.9	-	-	-	-	-	-
126-2	2-1-1-17-1-2	-	-	-	-	-	-	-	-	-
127-2	2-1-2-17-1-2	-	-	-	-	-	-	-	-	-
128-2	2-2-3-1-1-2	-	-	-	-	-	-	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-	-	-	-
130-2	3-1-9-1-1-2				1	3.4	3.4			
131-2	3-1-2-1-1-2	-	-	-	-	-	-	-	-	-
132-2	3-1-1-7-1-2	-	-	-	-	-	-	-	-	-
133-2	3-1-9-1-3-2	-	-	-	-	-	-	-	-	-
134-2	3-1-1-2-1-2	-	-	-	-	-	-	-	-	-
135-2	3-1-1-13-2-2	-	-	-	1	3.4	3.4	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-	-	-	-
137-2	3-1-1-2-11-2	-	-	-	-	-	-	-	-	-
138-2	3-1-9-1-8-2	-	-	-	-	-	-	-	-	-
139-2	7-1-6-13-1-2	-	-	-	-	-	-	-	-	-
140-2	3-1-1-1-6-2	-	-	-	-	-	-	-	-	-
141-2	2-2-1-4-3-2	-	-	-	-	-	-	-	-	-
142-2	3-2-1-11-2-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Tswana N=15			Venda N=30			Lemba N=53		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	3	20.0	10.3	2	6.7	4.6	7	13.2	4.7
1-9	2-1-1-1-1-9	-	-	-	-	-	-	-	-	-
2-2	3-1-1-1-3-2	3	20.0	10.3	10	33.3	8.6	11	20.8	5.6
3-2	3-1-1-2-2-2	-	-	-	3	10.0	5.5	5	9.4	4.0
3-5	3-1-1-2-2-5	-	-	-	-	-	-	-	-	-
4-2	3-1-1-3-2-2	-	-	-	-	-	-	-	-	-
4-9	3-1-1-3-2-9	-	-	-	-	-	-	-	-	-
5-2	3-1-1-2-5-2	1	6.7	6.5	-	-	-	1	1.9	1.9
5-5	3-1-1-2-5-5	-	-	-	-	-	-	-	-	-
6-2	2-1-2-1-1-2	-	-	-	-	-	-	-	-	-
7-2	3-1-1-1-1-2	3	20.0	10.3	5	16.6	6.8	18	34.0	6.5
8-2	1-1-1-1-1-2	-	-	-	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-	-	-	-
9-2	1-1-2-1-1-2	-	-	-	-	-	-	-	-	-
10-2	3-1-1-1-2-2	-	-	-	3	10.0	5.5	3	5.7	3.2
11-2	2-2-3-1-5-2	-	-	-	-	-	-	-	-	-
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-
21-2	2-1-1-1-2-2	1	6.7	6.5	-	-	-	1	1.9	1.9
21-5	2-1-1-1-2-5	-	-	-	-	-	-	-	-	-
22-2	2-1-1-1-5-2	-	-	-	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-	-	-	-
26-2	2-1-6-1-1-2	-	-	-	-	-	-	-	-	-
29-2	2-1-1-4-1 2	-	-	-	-	-	-	-	-	-
30-2	3-1-1-1-11-2	-	-	-	-	-	-	-	-	-
31-2	3-1-1-1-5-1	1	6.7	6.5	-	-	-	2	3.8	2.6
33-2	3-1-1-2-3-2	-	-	-	2	6.7	4.6	-	-	-
44-2	2-1-1-4-3-2	-	-	-	-	-	-	-	-	-
45-2	2-1-8-1-1-2	-	-	-	1	3.3	3.3	1	1.9	1.9
47-2	2-1-1-1-3-2	1	6.7	6.5	1	3.3	3.3	-	-	-
51-2	2-1-1-11-1-2	-	-	-	-	-	-	-	-	-
56-2	2-1-1-1-6-2	-	-	-	-	-	-	-	-	-
70-2	2-1-1-13-1-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Tswana N=15			Venda N=30			Lemba N=53		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
113-2	3-1-1-12-1-2	-	-	-	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	-	-	-	-	-	-
115-2	3-1-1-2-31-2	-	-	-	-	-	-	1	1.9	1.9
116-2	2-1-1-2-2-2	-	-	-	-	-	-	-	-	-
117-2	3-1-1-1-15-2	-	-	-	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	1	3.3	3.3	1	1.9	1.9
119-2	3-1-1-1-8-2	-	-	-	-	-	-	-	-	-
120-2	3-1-1-3-32-2	-	-	-	-	-	-	-	-	-
121-2	3-1-1-2-33-2	-	-	-	-	-	-	-	-	-
122-2	2-1-8-1-3-2	-	-	-	1	3.3	3.3	-	-	-
123-2	3-1-1-13-3-2	-	-	-	1	3.3	3.3	-	-	-
124-2	2-1-1-2-1-2	-	-	-	-	-	-	-	-	-
125-2	3-1-1-11-1-2	-	-	-	-	-	-	-	-	-
126-2	2-1-1-17-1-2	-	-	-	-	-	-	-	-	-
127-2	2-1-2-17-1-2	-	-	-	-	-	-	-	-	-
128-2	2-2-3-1-1-2	-	-	-	-	-	-	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-	-	-	-
130-2	3-1-9-1-1-2	-	-	-	-	-	-	-	-	-
131-2	3-1-2-1-1-2	1	6.7	6.5	-	-	-	-	-	-
132-2	3-1-1-7-1-2	-	-	-	-	-	-	-	-	-
133-2	3-1-9-1-3-2	-	-	-	-	-	-	-	-	-
134-2	3-1-1-2-1-2	1	6.7	6.5	-	-	-	-	-	-
135-2	3-1-1-13-2-2	-	-	-	-	-	-	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-	1	1.9	1.9
137-2	3-1-1-2-11-2	-	-	-	-	-	-	1	1.9	1.9
138-2	3-1-9-1-8-2	-	-	-	-	-	-	-	-	-
139-2	7-1-6-13-1-2	-	-	-	-	-	-	-	-	-
140-2	3-1-1-1-6-2	-	-	-	-	-	-	-	-	-
141-2	2-2-1-4-3-2	-	-	-	-	-	-	-	-	-
142-2	3-2-1-11-2-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Herero N=54			Dama N=43			Ambo N=22		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	18	33.3	6.4	9	20.9	6.2	3	13.6	7.3
1-9	2-1-1-1-1-9	-	-	-	-	-	-	-	-	-
2-2	3-1-1-1-3-2	-	-	-	-	-	-	3	13.6	7.3
3-2	3-1-1-2-2-2	3	5.6	3.1	1	2.3	2.3	1	4.5	4.4
3-5	3-1-1-2-2-5	-	-	-	-	-	-	-	-	-
4-2	3-1-1-3-2-2	-	-	-	-	-	-	-	-	-
4 9	3 1 1 3 2 9									
5 2	3 1 1 2 5 2	1	1.9	1.9	1	2.3	2.3			
5-5	3-1-1-2-5-5	-	-	-	-	-	-	-	-	-
6-2	2-1-2-1-1-2	-	-	-	-	-	-	-	-	-
7-2	3-1-1-1-1-2	-	-	-	1	2.3	2.3	12	54.5	10.6
8-2	1-1-1-1-1-2	-	-	-	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-	-	-	-
9-2	1-1-2-1-1-2	-	-	-	-	-	-	-	-	-
10-2	3-1-1-1-2-2	1	1.9	1.9	-	-	-	-	-	-
11-2	2-2-3-1-5-2	-	-	-	-	-	-	-	-	-
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-
21-2	2-1-1-1-2-2	27	50.0	6.8	14	32.6	7.1	1	4.5	4.4
21-5	2-1-1-1-2-5	1	1.9	1.9	-	-	-	-	-	-
22-2	2-1-1-1-5-2	1	1.9	1.9	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-	1	4.5	4.4
26-2	2-1-6-1-1-2	-	-	-	-	-	-	-	-	-
29-2	2-1-1-4-1-2	-	-	-	-	-	-	-	-	-
30-2	3-1-1-1-11-2	-	-	-	1	2.3	2.3	-	-	-
31-2	3-1-1-1-5-1	-	-	-	-	-	-	-	-	-
33 2	3 1 1-2-3-2									
44 2	2 1 1 4 3 2									
45-2	2-1-8-1-1-2	-	-	-	-	-	-	-	-	-
47-2	2-1-1-1-3-2	-	-	-	-	-	-	-	-	-
51-2	2-1-1-11-1-2	-	-	-	-	-	-	-	-	-
56-2	2-1-1-1-6-2	-	-	-	-	-	-	-	-	-
70-2	2-1-1-13-1-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Herero N=54			Dama N=43			Ambo N=22		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
113-2	3-1-1-12-1-2	-	-	-	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	-	-	-	-	-	-
115-2	3-1-1-2-31-2	-	-	-	-	-	-	-	-	-
116-2	2-1-1-2-2-2	1	1.9	1.9	2	4.7	3.2	-	-	-
117-2	3-1-1-1-15-2	1	1.9	1.9	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	4	9.3	4.4	-	-	-
119-2	3-1-1-1-8-2	-	-	-	1	2.3	2.3	-	-	-
120-2	3-1-1-3-32-2	-	-	-	-	-	-	-	-	-
121-2	3-1-1-2-33-2	-	-	-	-	-	-	-	-	-
122-2	2-1-8-1-3-2	-	-	-	-	-	-	-	-	-
123-2	3-1-1-13-3-2	-	-	-	-	-	-	-	-	-
124-2	2-1-1-2-1-2	-	-	-	3	7.0	3.9	-	-	-
125-2	3-1-1-11-1-2	-	-	-	-	-	-	-	-	-
126-2	2-1-1-17-1-2	-	-	-	-	-	-	-	-	-
127-2	2-1-2-17-1-2	-	-	-	-	-	-	-	-	-
128-2	2-2-3-1-1-2	-	-	-	-	-	-	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-	1	4.5	4.4
130-2	3-1-9-1-1-2	-	-	-	-	-	-	-	-	-
131-2	3-1-2-1-1-2	-	-	-	-	-	-	-	-	-
132-2	3-1-1-7-1-2	-	-	-	-	-	-	-	-	-
133-2	3-1-9-1-3-2	-	-	-	-	-	-	-	-	-
134-2	3-1-1-2-1-2	-	-	-	1	2.3	2.3	-	-	-
135-2	3-1-1-13-2-2	-	-	-	-	-	-	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-	-	-	-
137-2	3-1-1-2-11-2	-	-	-	-	-	-	-	-	-
138-2	3-1-9-1-8-2	-	-	-	5	11.6	4.9	-	-	-
139-2	7-1-6-13-1-2	-	-	-	-	-	-	-	-	-
140-2	3-1-1-1-6-2	-	-	-	-	-	-	-	-	-
141-2	2-2-1-4-3-2	-	-	-	-	-	-	-	-	-
142-2	3-2-1-11-2-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Sekele N=49			Nama N=46			R. "Col" N=35		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	1	2.0	2.0	3	6.5	3.6	2	5.7	3.9
1-9	2-1-1-1-1-9	-	-	-	-	-	-	-	-	-
2-2	3-1-1-1-3-2	-	-	-	-	-	-	-	-	-
3-2	3-1-1-2-2-2	7	14.3	4.9	25	54.3	7.3	20	57.1	8.4
3-5	3-1-1-2-2-5	-	-	-	-	-	-	2	5.7	3.9
4-2	3-1-1-3-2-2	20	40.8	7.0	-	-	-	-	-	-
4-9	3-1-1-3-2-9	1	2.0	2.0	-	-	-	-	-	-
5-2	3-1-1-2-5-2	1	2.0	2.0	3	6.5	3.6	4	11.4	5.4
5-5	3-1-1-2-5-5	-	-	-	-	-	-	1	2.9	2.9
6-2	2-1-2-1-1-2	-	-	-	-	-	-	-	-	-
7-2	3-1-1-1-1-2	10	20.4	5.6	3	6.5	3.6	2	5.7	3.9
8-2	1-1-1-1-1-2	-	-	-	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-	-	-	-
9-2	1-1-2-1-1-2	-	-	-	-	-	-	-	-	-
10-2	3-1-1-1-2-2	6	12.2	4.7	3	6.5	3.6	2	5.7	3.9
11-2	2-2-3-1-5-2	-	-	-	-	-	-	-	-	-
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-
21-2	2-1-1-1-2-2	2	4.1	2.8	6	13.0	4.9	-	-	-
21-5	2-1-1-1-2-5	-	-	-	-	-	-	-	-	-
22-2	2-1-1-1-5-2	-	-	-	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-	-	-	-
26-2	2-1-6-1-1-2	-	-	-	-	-	-	-	-	-
29-2	2-1-1-4-1-2	-	-	-	-	-	-	-	-	-
30-2	3-1-1-1-11-2	-	-	-	-	-	-	-	-	-
31-2	3-1-1-1-5-1	-	-	-	-	-	-	-	-	-
33-2	3-1-1-2-3-2	-	-	-	-	-	-	-	-	-
44-2	2-1-1-4-3-2	-	-	-	-	-	-	-	-	-
45-2	2-1-8-1-1-2	-	-	-	-	-	-	-	-	-
47-2	2-1-1-1-3-2	-	-	-	-	-	-	-	-	-
51-2	2-1-1-11-1-2	-	-	-	-	-	-	-	-	-
56-2	2-1-1-1-6-2	-	-	-	-	-	-	-	-	-
70-2	2-1-1-13-1-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	Sekele N=49			Nama N=46			R_"Col" N=35		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
113-2	3-1-1-12-1-2	1	2.0	2.0	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	-	-	-	-	-	-
115-2	3-1-1-2-31-2	-	-	-	-	-	-	-	-	-
116-2	2-1-1-2-2-2	-	-	-	-	-	-	1	2.9	2.9
117-2	3-1-1-1-15-2	-	-	-	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	-	-	-	-	-	-
119-2	3-1-1-1-8-2	-	-	-	-	-	-	-	-	-
120-2	3-1-1-3-32-2	-	-	-	-	-	-	-	-	-
121-2	3-1-1-2-33-2	-	-	-	-	-	-	-	-	-
122-2	2-1-8-1-3-2	-	-	-	-	-	-	-	-	-
123-2	3-1-1-13-3-2	-	-	-	-	-	-	-	-	-
124-2	2-1-1-2-1-2	-	-	-	-	-	-	1	2.9	2.9
125-2	3-1-1-11-1-2	-	-	-	-	-	-	-	-	-
126-2	2-1-1-17-1-2	-	-	-	-	-	-	-	-	-
127-2	2-1-2-17-1-2	-	-	-	-	-	-	-	-	-
128-2	2-2-3-1-1-2	-	-	-	-	-	-	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-	-	-	-
130-2	3-1-9-1-1-2	-	-	-	-	-	-	-	-	-
131-2	3-1-2-1-1-2	-	-	-	-	-	-	-	-	-
132-2	3-1-1-7-1-2	-	-	-	-	-	-	-	-	-
133-2	3-1-9-1-3-2	-	-	-	-	-	-	-	-	-
134-2	3-1-1-2-1-2	-	-	-	-	-	-	-	-	-
135-2	3-1-1-13-2-2	-	-	-	-	-	-	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-	-	-	-
137-2	3-1-1-2-11-2	-	-	-	-	-	-	-	-	-
138-2	3-1-9-1-8-2	-	-	-	-	-	-	-	-	-
139-2	7-1-6-13-1-2	-	-	-	2	4.3	3.0	-	-	-
140-2	3-1-1-1-6-2	-	-	-	1	2.2	2.1	-	-	-
141-2	2-2-1-4-3-2	-	-	-	-	-	-	-	-	-
142-2	3-2-1-11-2-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	JHB_"Col" N=71			English N=36			Afrikaners N=51		
		No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	13	18.3	4.6	23	63.9	8.0	41	80.4	5.6
1-9	2-1-1-1-1-9	-	-	-	-	-	-	-	-	-
2-2	3-1-1-1-3-2	8	11.3	3.8	-	-	-	-	-	-
3-2	3-1-1-2-2-2	17	23.9	5.1	-	-	-	-	-	-
3-5	3-1-1-2-2-5	-	-	-	-	-	-	-	-	-
4-2	3-1-1-3-2-2	-	-	-	-	-	-	-	-	-
4-9	3-1-1-3-2-9	-	-	-	-	-	-	-	-	-
5-2	3-1-1-2-5-2	13	18.3	4.6	-	-	-	-	-	-
5-5	3-1-1-2-5-5	-	-	-	-	-	-	-	-	-
6-2	2-1-2-1-1-2	-	-	-	3	8.3	4.5	1	2.0	2.0
7-2	3-1-1-1-1-2	8	11.3	3.8	-	-	-	-	-	-
8-2	1-1-1-1-1-2	1	1.4	1.4	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-	-	-	-
9-2	1-1-2-1-1-2	1	1.4	1.4	-	-	-	-	-	-
10-2	3-1-1-1-2-2	2	2.8	2.0	-	-	-	-	-	-
11-2	2-2-3-1-5-2	-	-	-	2	5.6	3.8	-	-	-
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-
21-2	2-1-1-1-2-2	-	-	-	-	-	-	2	3.9	2.7
21-5	2-1-1-1-2-5	-	-	-	-	-	-	-	-	-
22-2	2-1-1-1-5-2	-	-	-	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-	-	-	-
26-2	2-1-6-1-1-2	-	-	-	-	-	-	-	-	-
29-2	2-1-1-4-1-2	-	-	-	1	2.8	2.8	-	-	-
30-2	3-1-1-1-11-2	-	-	-	-	-	-	-	-	-
31-2	3-1-1-1-5-1	-	-	-	-	-	-	-	-	-
33-2	3-1-1-2-3-2	1	1.4	1.4	-	-	-	-	-	-
44-2	2-1-1-4-3-2	-	-	-	3	8.3	4.6	1	2.0	2.0
45-2	2-1-8-1-1-2	-	-	-	-	-	-	-	-	-
47-2	2-1-1-1-3-2	4	5.6	2.7	3	8.3	4.6	1	2.0	2.0
51-2	2-1-1-11-1-2	-	-	-	-	-	-	1	2.0	2.0
56-2	2-1-1-1-6-2	-	-	-	-	-	-	-	-	-
70-2	2-1-1-13-1-2	-	-	-	-	-	-	-	-	-

Type	Mitotype	JHB_"Col" N=71 No	Freq	(S.E.)	English N=36 No	Freq	(S.E.)	Afrikaners N=51 No	Freq	(S.E.)
113-2	3-1-1-12-1-2	-	-	-	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	1	2.8	2.7	-	-	-
115-2	3-1-1-2-31-2	1	1.4	1.4	-	-	-	-	-	-
116-2	2-1-1-2-2-2	-	-	-	-	-	-	-	-	-
117-2	3-1-1-1-15-2	-	-	-	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	-	-	-	-	-	-
119-2	3-1-1-1-8-2	-	-	-	-	-	-	-	-	-
120-2	3-1-1-3-32-2	1	1.4	1.4						
121-2	3-1-1-2-33-2	1	1.4	1.4	-	-	-			
122-2	2-1-8-1-3-2	-	-	-	-	-	-	-	-	-
123-2	3-1-1-13-3-2	-	-	-	-	-	-	-	-	-
124-2	2-1-1-2-1-2	-	-	-	-	-	-	-	-	-
125-2	3-1-1-11-1-2	-	-	-	-	-	-	-	-	-
126-2	2-1-1-17-1-2	-	-	-	-	-	-	-	-	-
127-2	2-1-2-17-1-2	-	-	-	-	-	-	-	-	-
128-2	2-2-3-1-1-2	-	-	-	-	-	-	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-	-	-	-
130-2	3-1-9-1-1-2	-	-	-	-	-	-	-	-	-
131-2	3-1-2-1-1-2	-	-	-	-	-	-	-	-	-
132-2	3-1-1-7-1-2	-	-	-	-	-	-	-	-	-
133-2	3-1-9-1-3-2	-	-	-	-	-	-	-	-	-
134-2	3-1-1-2-1-2	-	-	-	-	-	-	-	-	-
135-2	3-1-1-13-2-2	-	-	-	-	-	-	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-	-	-	-
137-2	3-1-1-2-11-2	-	-	-	-	-	-	-	-	-
138-2	3-1-9-1-8-2	-	-	-	-	-	-	-	-	-
139-2	7-1-6-13-1-2			-		-	-			
140-2	3-1-1-1-6-2			-		-	-			
141-2	2-2-1-4-3-2	-	-	-	-	-	-	3	5.9	3.3
142-2	3-2-1-11-2-2	-	-	-	-	-	-	1	2.0	2.0

Type	Mitotype	A Jew N=55			Indian N=50		
		No	Freq	(S.E.)	No	Freq	(S.E.)
1-2	2-1-1-1-1-2	34	61.8	6.6	44	88.8	4.5
1-9	2-1-1-1-1-9	-	-	-	-	-	-
2-2	3-1-1-1-3-2	-	-	-	1	2.0	2.0
3-2	3-1-1-2-2-2	-	-	-	1	-	-
3-5	3-1-1-2-2-5	-	-	-	-	2.0	2.0
4-2	3-1-1-3-2-2	-	-	-	-	-	-
4-9	3-1-1-3-2-9	-	-	-	-	-	-
5-2	3-1-1-2-5-2	-	-	-	-	-	-
5 5	3 1 1 2 5 5						
6-2	2-1-2-1-1-2	17	30.9	6.2	-	-	
7-2	3-1-1-1-1-2	-	-	-	-	-	-
8-2	1-1-1-1-1-2	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-
9-2	1-1-2-1-1-2	-	-	-	-	-	-
10-2	3-1-1-1-2-2	-	-	-	-	-	-
11-2	2-2-3-1-5-2	-	-	-	-	-	-
18-2	2-3-1-4-9-2	1	1.8	1.8	-	-	-
21-2	2-1-1-1-2-2	-	-	-	1	2.0	2.0
21-5	2-1-1-1-2-5	-	-	-	-	-	-
22-2	2-1-1-1-5-2	-	-	-	-	-	-
25-2	5-1-1-1-1-2	-	-	-	-	-	-
26-2	2-1-6-1-1-2	-	-	-	-	-	-
29-2	2-1-1-4-1-2	-	-	-	-	-	-
30-2	3-1-1-1-11-2	-	-	-	-	-	-
31-2	3-1-1-1-5-1	-	-	-	1	2.0	2.0
33-2	3-1-1-2-3-2	-	-	-	-	-	-
44-2	2-1-1-4-3-2	-	-	-	-	-	-
45 2	2 1 8 1 1 2	-	-	-	-		
47-2	2-1-1-1-3-2	-	-	-	1	2.0	2.0
51-2	2-1-1-11-1-2	-	-	-	-	-	-
56-2	2-1-1-1-6-2	-	-	-	1	2.0	2.0
70-2	2-1-1-13-1-2	-	-	-	-	-	-

Type	Mitotype	A Jew N=55 No	Freq	(S.E.)	Indian N=50 No	Freq	(S.E.)
113-2	3-1-1-12-1-2	-	-	-	-	-	-
114-2	2-2-2-1-5-2	-	-	-	-	-	-
115-2	3-1-1-2-31-2	-	-	-	-	-	-
116-2	2-1-1-2-2-2	-	-	-	-	-	-
117-2	3-1-1-1-15-2	-	-	-	-	-	-
118-2	3-1-1-13-1-2	-	-	-	-	-	-
119-2	3-1-1-1-8-2	-	-	-	-	-	-
120-2	3-1-1-3-32-2	-	-	-	-	-	-
121-2	3-1-1-2-33-2	-	-	-	-	-	-
122-2	2-1-8-1-3-2	-	-	-	-	-	-
123-2	3-1-1-13-3-2	-	-	-	-	-	-
124-2	2-1-1-2-1-2	-	-	-	-	-	-
125-2	3-1-1-11-1-2	-	-	-	-	-	-
126-2	2-1-1-17-1-2	1	1.8	1.8	-	-	-
127-2	2-1-2-17-1-2	1	1.8	1.8	-	-	-
128-2	2-2-3-1-1-2	1	1.8	1.8	-	-	-
129-2	2-1-1-5-1-2	-	-	-	-	-	-
130-2	3-1-9-1-1-2	-	-	-	-	-	-
131-2	3-1-2-1-1-2	-	-	-	-	-	-
132-2	3-1-1-7-1-2	-	-	-	-	-	-
133-2	3-1-9-1-3-2	-	-	-	-	-	-
134-2	3-1-1-2-1-2	-	-	-	-	-	-
135-2	3-1-1-13-2-2	-	-	-	-	-	-
136-2	3-1-8-1-1-2	-	-	-	-	-	-
137-2	3-1-1-2-11-2	-	-	-	-	-	-
138-2	3-1-9-1-8-2	-	-	-	-	-	-
139-2	7-1-6-13-1-2	-	-	-	-	-	-
140-2	3-1-1-1-6-2	-	-	-	-	-	-
141-2	2-2-1-4-3-2	-	-	-	-	-	-
142-2	3-2-1-11-2-2	-	-	-	-	-	-

3.2.3 Restriction site polymorphism

The six restriction enzymes (*Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II) together recognize 233 nucleotides on the mtDNA molecule. Of these only 81 are informative and give rise to the 149 mtDNA types described by Merriwether *et al.* (1991). Only 31 of these informative sites are found in individuals investigated in the present study (Table 3.8).

The 31 variant sites associated with the 62 different mtDNA types discovered in the present study were compared with the corresponding positions on the reference sequence. From this comparative analysis it was found that *Hpa*I site *d* (np 3592) which gives rise to morph *Hpa*I-3, was found in 54% of mtDNA types whilst the remaining 46% of types, like the reference sequence, did not have this site. Approximately 25% of mtDNA types lack the *Msp*I *a* site (np 8112) and 30% lack the *Msp*I *b* site (np 8150) which are found in the reference sequence. Whereas the majority of types lack the mutation at *Ava*II site *c* (np 8249), only 25% of types are found with this mutation. *Ava*II site *f* (np 16390) was found in 15% of mtDNA types but not in the other 85% of mtDNA types including the reference sequence. All other variant sites occur rarely (less than 10%).

Table 3.8. Comparisons of variant sites detected with the enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II for all mtDNA types discovered in this study with corresponding positions in the reference sequence. Sites are scored as (1) if they are the same as in the reference sequence and (0) if different in the reference sequence. The positions of these sites as well as the type of mutation [either a loss (-) or a gain (+)] are also indicated.

		HpaI			BamHI		HaeII					MspI							
	Site	c	d	f	a	b	a	b	c	d	g	a	b	c	d	e	f	k	l
		-	+	+		+	-	+	-	+	+	-	-	+	+	+	+	+	+
	Nucleotide	1			1	1								1	1	1	1		1
	Position	2	3	1	3	6	4	4	9	9	9	8	8	1	3	5	5	8	1
		4	5	0	3	3	5	8	0	2	6	1	1	4	0	9	5	6	1
		0	9	0	6	5	2	3	5	5	8	1	5	4	7	2	1	4	6
		6	2	4	8	0	9	0	2	0	7	2	0	0	0	5	0	9	4
Control sequence		1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
Type	Mitotype																		
1-2	2-1-1-1-1-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
1-9	2-1-1-1-1-9	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
2-2	3-1-1-1-3-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
3-2	3-1-1-2-2-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
3-5	3-1-1-2-2-5	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
4-2	3-1-1-3-2-2	1	1	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0
4-9	3-1-1-3-2-9	1	1	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0
5-2	3-1-1-2-5-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
5-5	3-1-1-2-5-5	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
6-2	2-1-2-1-1-2	1	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	0
7-2	3-1-1-1-1-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
8-2	1-1-1-1-1-2	0	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
8-1	1-1-1-1-1-1	0	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
9-2	1-1-2-1-1-2	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	0
10-2	3-1-1-1-2-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
11-2	2-2-3-1-5-2	1	0	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0
18-2	2-3-1-4-9-2	1	0	0	1	0	1	0	1	0	0	1	1	0	0	0	0	0	0
21-2	2-1-1-1-2-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
21-5	2-1-1-1-2-5	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
22-2	2-1-1-1-5-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
25-2	5-1-1-1-1-2	1	0	1	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
26-2	2-1-6-1-1-2	1	0	0	0	0	1	0	1	1	0	1	1	0	0	1	0	0	0
28-2	2-1-1-4-1-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0
30-2	3-1-1-1-11-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
31-2	3-1-1-1-5-1	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
33-2	3-1-1-2-3-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
44-2	2-1-1-4-3-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0
45-2	2-1-8-1-1-2	1	0	0	0	0	1	0	1	0	1	1	1	0	0	1	0	0	0
47-2	2-1-1-1-3-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
51-2	2-1-1-11-1-2	1	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0
56-2	2-1-1-1-6-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
70-2	2-1-1-13-1-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1

		HpaI			BamHI		HaeII					MspI							
	Site	c	d	f	a	b	a	b	c	d	g	a	b	c	d	e	f	k	l
		-	+	+	-	+	-	+	-	+	+	-	-	+	+	+	+	+	+
	Nucleotide	1			1	1								1	1	1	1		1
	Position	2	3	1	3	6	4	4	9	9	9	8	8	1	3	5	5	8	1
		4	5	0	3	3	5	8	0	2	6	1	1	4	0	9	5	6	1
		0	9	0	6	5	2	3	5	5	8	1	5	4	7	2	1	4	6
		6	2	4	8	0	9	0	2	0	7	2	0	0	0	5	0	9	4
Control sequence		1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
Type	Mitotype																		
113-2	3-1-1-12-1-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0
114-2	2-2-2-1-5-2	1	0	0	0	1	1	0	0	0	0	1	1	0	0	1	0	0	0
115-2	3-1-1-2-31-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
116-2	2-1-1-2-2-2	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
117-2	3-1-1-1-15-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
118-2	3-1-1-13-1-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1
119-2	3-1-1-1-8-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
120-2	3-1-1-3-32-2	1	1	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0
121-2	3-1-1-2-33-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
122-2	2-1-8-1-3-2	1	0	0	0	0	1	0	1	0	1	1	1	0	0	1	0	0	0
123-2	3-1-1-13-3-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1
124-2	2-1-1-2-1-2	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
125-2	3-1-1-11-1-2	1	1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0
126-2	2-1-1-17-1-2	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
127-2	2-1-2-17-1-2	1	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	0
128-2	2-2-3-1-1-2	1	0	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0
129-2	2-1-1-5-1-2	1	0	0	0	0	1	0	1	0	0	0	0	0	1	1	0	0	0
130-2	3-1-9-1-1-2	1	1	0	0	0	1	1	1	0	0	1	1	0	0	1	0	0	0
131-2	3-1-2-1-1-2	1	1	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	0
132-2	3-1-1-7-1-2	1	1	0	0	0	1	0	1	0	0	1	0	0	0	1	1	0	0
133-2	3-1-9-1-3-2	1	1	0	0	0	1	1	1	0	0	1	1	0	0	1	0	0	0
134-2	3-1-1-2-1-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
135-2	3-1-1-13-2-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1
136-2	3-1-8-1-1-2	1	1	0	0	0	1	0	1	0	1	1	1	0	0	1	0	0	0
137-2	3-1-1-2-11-2	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
138-2	3-1-9-1-8-2	1	1	0	0	0	1	1	1	0	0	1	1	0	0	1	0	0	0
139-2	7-1-6-13-1-2	0	1	0	0	0	1	0	1	1	0	1	1	0	0	1	0	0	1
140-2	3-1-1-1-6-2	1	1	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0
141-2	2-2-1-4-3-2	1	0	0	0	1	1	0	1	0	0	1	1	0	0	0	0	0	0
142-2	3-2-1-11-2-2	1	1	0	0	1	1	0	1	0	0	1	0	0	0	0	0	0	0

		AvaII									HincII			
	Site	b	c	d	e	f	i	ac	ab	ac	a	d	e	f
		+	+	-	+	-	+	+	+	+	-	-	-	+
	Nucleotide			1	1	1		1	1		1		1	1
	Position	4	8	3	5	6	4	2	0	5	2	7	3	3
		4	2	3	8	3	2	6	7	1	4	8	2	6
		7	4	6	9	9	8	2	7	6	0	5	5	3
		6	9	7	0	0	0	0	4	5	6	3	9	4
Control sequence		0	0	1	0	1	0	0	0	0	1	1	1	1
Type	Mitotype													
1-2	2-1-1-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
1-9	2-1-1-1-1-9	0	0	1	0	1	0	0	0	0	1	0	0	0
2-2	3-1-1-1-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
3-2	3-1-1-2-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1
3-5	3-1-1-2-2-5	0	1	1	0	1	0	0	0	0	1	0	1	1
4-2	3-1-1-3-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1
4-9	3-1-1-3-2-9	0	1	1	0	1	0	0	0	0	1	0	0	0
5-2	3-1-1 2-5-2	0	0	1	0	1	0	0	0	0	1	1	1	1
5-5	3-1-1-2-5-5	0	0	1	0	1	0	0	0	0	1	0	1	1
6-2	2-1-2-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
7-2	3-1-1-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
8-2	1-1-1-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
8-1	1-1-1-1-1 1	0	0	1	0	1	0	0	0	0	0	1	1	1
9-2	1-1-2-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
10-2	3-1-1-1-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1
11-2	2-2-3-1-5-2	0	0	1	0	1	0	0	0	0	1	1	1	1
18-2	2-3-1-4-9-2	0	0	1	1	1	0	0	0	0	1	1	1	1
21-2	2-1-1-1-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1
21-5	2-1-1-1-2-5	0	1	1	0	1	0	0	0	0	1	0	1	1
22-2	2-1-1-1-5-2	0	0	1	0	1	0	0	0	0	1	1	1	1
25-2	5-1-1-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
26-2	2-1-6-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
28-2	2-1-1-4-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
30-2	3-1-1-1-11-2	0	1	1	1	1	0	0	0	0	1	1	1	1
31-2	3-1-1-1-5-1	0	0	1	0	1	0	0	0	0	0	1	1	1
33-2	3-1-1-2-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
44-2	2-1-1-4-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
45-2	2-1-8-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
47 2	2 1-1 1-3 2	0	0	1	0	0	0	0	0	0	1	1	1	1
51 2	2-1-1 11-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
56-2	2-1-1-1-6-2	0	0	1	1	1	0	0	0	0	1	1	1	1
70-2	2-1-1-13-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1

		AvaII									HincII			
	Site	b	c	d	e	f	i	ac	ab	ac	a	d	e	f
		+	+	-	+	-	+	-	+	+	-	-	+	+
	Nucleotide			1	1	1		1	1		1		1	1
	Position	4	8	3	5	6	4	2	0	5	2	7	3	3
		4	2	3	8	3	2	6	7	1	4	8	2	6
		7	4	6	9	9	8	2	7	6	0	5	5	3
		6	9	7	0	0	0	0	4	5	6	3	9	4
Control sequence		0	0	1	0	1	0	0	0	0	1	1	1	1
Type	Mitotype													
113-2	3-1-1-12-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
114-2	2-2-2-1-5-2	0	0	1	0	1	0	0	0	0	1	1	1	1
115-2	3-1-1-2-31-2	0	0	1	0	1	0	1	0	0	1	1	1	1
116-2	2-1-1-2-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1
117-2	3-1-1-1-15-2	0	0	1	0	1	1	0	0	0	1	1	1	1
118-2	3-1-1-13-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
119-2	3-1-1-1-8-2	1	0	0	0	0	0	0	0	0	1	1	1	1
120-2	3-1-1-3-32-2	0	1	1	0	1	0	0	1	0	1	1	1	1
121-2	3-1-1-2-33-2	0	1	1	0	1	0	0	0	1	1	1	1	1
122-2	2-1-8-1-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
123-2	3-1-1-13-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
124-2	2-1-1-2-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
125-2	3-1-1-11-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
126-2	2-1-1-17-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
127-2	2-1-2-17-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
128-2	2-2-3-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
129-2	2-1-1-5-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
130-2	3-1-9-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
131-2	3-1-2-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
132-2	3-1-1-7-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
133-2	3-1-9-1-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
134-2	3-1-1-2-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
135-2	3-1-1-13-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1
136-2	3-1-8-1-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
137-2	3-1-1-2-11-2	0	1	1	1	1	0	0	0	0	1	1	1	1
138-2	3-1-9-1-8-2	1	0	0	0	0	0	0	0	0	1	1	1	1
139-2	7-1-6-13-1-2	0	0	1	0	1	0	0	0	0	1	1	1	1
140-2	3-1-1-1-6-2	0	0	1	1	1	0	0	0	0	1	1	1	1
141-2	2-2-1-4-3-2	0	0	1	0	0	0	0	0	0	1	1	1	1
142-2	3-2-1-11-2-2	0	1	1	0	1	0	0	0	0	1	1	1	1

3.3 Statistical Analysis

3.3.1 Genetic diversity

Genetic diversity (h) within population groups were computed using Nei's method (1987). Since h is dependant on the frequencies of different mtDNA types found within groups, this value increases as the number of types increases in a population (Table 3.9). Genetic diversity for haploid systems is related to homogeneity (F) which was first used by Johnson *et al.* (1983). While h is defined as $(1 - \Sigma x_i^2)$, F is calculated as Σx_i^2.

Caucasoid populations appear more homogeneous than non-Caucasoid populations (Table 3.9). In fact, the S.A. Indian population is more homogeneous (F = 0.77) when compared to Indians from Nepal (F = 0.35) and India (F = 0.51)(Semino *et al.* 1991), as well as other Caucasoids [0.63 in the Finnish (Vikki *et al.* 1988); 0.39 in Italians (Brega *et al.* 1986); 0.29 in Israeli Jews (Bonne-Tamir *et al.* 1986); and 0.36 in the Caucasoid sample investigated by Johnson *et al.* 1983]. The high homogeneity in S.A. Indians may be due to the relatively limited origins of the S.A. Indian populations, with sub-groups originating from a single village or town and consisting, in the main, of related individuals. Traditional religious and customary beliefs of strict caste endogamy may also be a factor influencing marriage pattern amongst Indian

religious/language groups. Although the S.A. European Caucasoid population is made up of several diverse European and U.K. populations, the higher homogeneity in these groups when compared to other European Caucasoids (see above) may suggest that the mtDNA lineages in S.A. European Caucasoids were derived from a few mtDNA lineages. Also, there has been very little gene flow from the indigenous populations into South African Caucasoids when compared to Caucasoids from other countries (Semino *et al*. 1989). This observation may be accounted for by the political barriers in South Africa, where marriages between individuals from different "race" groups were barred by law, until about two years ago.

That Negroid populations are more diverse than Khoisan and Caucasoid populations is indictated by the presence of several mtDNA types at frequencies which range between 20-30% (see Table 3.7). This may be due to more gene flow ocurring between the various ethnic groups within the Negroid groups than between Khoisan groups, with a concomitant introduction of variation within the Negroid group.

Table 3.9. Correlation between the number of mtDNA types, genetic diversity and percent sequence divergence within ethnic groups from southern Africa.

POPULATION	SAMPLE SIZE	NOS. OF TYPES	GENETIC DIVERSITY	%SEQUENCE DIVERGENCE
ZULU	30	8	0.805	0.194
SWAZI	41	14	0.838	0.194
XHOSA	18	9	0.863	0.231
TSONGA	34	7	0.747	0.155
SOTHO	29	10	0.877	0.251
PEDI	23	6	0.731	0.203
TSWANA	15	9	0.914	0.246
VENDA	30	11	0.855	0.214
LEMBA	53	13	0.824	0.198
HERERO	54	9	0.646	0.139
DAMA	43	12	0.837	0.322
AMBO	22	7	0.688	0.149
SAN	49	9	0.803	0.256
NAMA	46	8	0.683	0.257
RICHTERSVELD"COLOURED"	35	9	0.664	0.222
JOHANNESBURG "COLOURED	71	13	0.857	0.411
ENGLISH-SPEAKING	36	7	0.583	0.162
ASHKENAZI JEWISH	55	6	0.531	0.079
INDIAN	50	7	0.228	0.044
AFRIKAANS-SPEAKING	51	8	0.354	0.099
NEGROID	392	41		0.208
KHOISAN	95	11		0.256
CAUCASOID	192	19		0.096

3.3.2 Extent of population subdivision, G_{ST}

The variance due to differences between populations was estimated from the frequencies of mtDNA types using Nei's coefficient of population subdivision, G_{ST}. A G_{ST} value of 0.163 was calculated using the frequencies of the 62 mtDNA types discovered in the 20 ethnic groups investigated in this study. This value is comparable to the value of 0.180 estimated by the Takahata and Palumbi (1985) method, which is dependant on the probability of allelic identity using mtDNA restriction sites. This suggests that approximately 17% of the total variance observed in this study is due to differences found between populations whereas 83% of the variation occurs within populations (this of course has considerable bearing on how one is to define a "population").

To investigate the relative contribution of each group to the G_{ST} value, each population was removed in a stepwise manner and the G_{ST} recalculated (Table 3.10). Another estimate of G_{ST} was obtained by removing each of the four major population groups; Caucasoids, Khoisan, "Coloured" and Negroid in turn and recalulating G_{ST} (Table 3.10).

The stepwise removal of populations did not result in an appreciable change in the G_{ST} estimate (Table 3.10): eleven of these changes resulted in a slightly higher value affecting mainly the Negroid ethnic groups. When Khoisan

and Caucasoid populations were exluded from the calculation, slightly lower G_{ST} values were obtained. This suggests that the total variance in Negroids results very little from gene flow from Khoisan and/or Caucasoid populations.

Removal of the combined Caucasoid group (English- and Afrikaans-speaking "Whites", Indians and Jewish populations) resulted in a 42.9% decrease in G_{ST}. This means that very little variation occurs between African populations which may be attributable to geographic proximity of these populations. This also provides support for the biological unity of African populations and suggests that both Negroid and Khoisan populations have descended from a recent common female ancestor who lived in Africa. Removal of the pooled Khoisan and "Coloured" groups only changed the G_{ST} value by 5.5% and 2.5%, respectively, whereas removal of the pooled Negroid group increased the G_{ST} value by 51.5%. This would suggest that the mtDNA diversity found within the Negroid group lies intermediate between the Khoisan and Caucasoid, since when Negroids are removed from the G_{ST} calculation, apprecible differences between Caucasoid and the non-Caucasoid populations are obtained.

The G_{ST} value calculated for the 20 ethnic groups in this study is lower than the value of 0.35 calculated for 62 populations from different geographic regions, (including

the 20 invesitgated here) using the six enzyme system (Merriwether *et al*. 1991). It is also lower than the value of 0.31 obtained by Stoneking *et al*. (1990) who used high resolution restriction enzyme mapping to investigate variation in populations from seven different geographic regions. This may be attributable to the larger number of African compared with non-African populations investigated in this study when compared with the studies by Stoneking *et al*. (1990) and Merriwether *et al*. (1991). The extent of geographic structuring is more obvious in the studies by Stoneking *et al*. (1990) and Merriwether *et al*. (1991) since populations were drawn from several different continents, which results in more variance contributed by the different populations.

Table 3.10. Index of interpopulation differences estimated by calculating G_{ST}^{*}. The contribution of each population and the pooled Caucasoid, Khoisan, Negroid and "Coloured" populations to the G_{ST} value was calulated by removing each population in turn and then recalulating the value.

POPULATION	H_S	H_T	G_{ST}
Total Sample	0.856	0.716	0.163
Remove			
Zulu	0.856	0.712	0.169
Swazi	0.856	0.710	0.171
Xhosa	0.855	0.709	0.171
Tsonga	0.858	0.715	0.167
Sotho	0.855	0.708	0.172
Pedi	0.856	0.716	0.164
Tswana	0.853	0.706	0.173
Venda	0.851	0.709	0.167
Lemba	0.855	0.711	0.169
Herero	0.854	0.720	0.156
Dama	0.849	0.710	0.164
Ambo	0.857	0.718	0.162
San	0.849	0.712	0.162
Nama	0.851	0.718	0.156
Richtersveld "Coloured"	0.850	0.719	0.154
JHB "Coloured"	0.853	0.709	0.169
English	0.859	0.723	0.158
Ashkenazi Jewish	0.858	0.726	0.154
Indian	0.866	0.742	0.143
Afrikaners	0.863	0.735	0.148
Pooled Caucasoid	0.870	0.789	0.093
Pooled Khoisan	0.844	0.703	0.154
Pooled "Coloured"	0.846	0.711	0.159
Pooled Negroid	0.780	0.588	0.247

*Calulated using the method described by Nei (1987)

3.3.3 Nucleotide sequence divergence

Nucleotide divergence was estimated from the number of shared and different restriction sites for all pairs of types using the method of Nei and Miller (1990). By estimating the average nucleotide divergence between individuals drawn randomly from all possible pairs of populations, a distance matrix showing the amount of sequence divergence between groups was obtained (Table 3.11). The average amount of diversity within each group (π) was calculated using the method of Nei and Li (1979) and is shown on the diagonal of Table 3.11.

The pooled Khoisan group has a slightly higher (0.256) sequence divergence than the pooled Negroid group (0.208). The pooled Caucasoid group revealed the least amount of sequence variation (0.096) (refer to Table 3.9). Although there are more mtDNA types in the pooled Caucasoid group (19) compared to the pooled Khoisan group (11), mtDNA types found in the Khoisan may only be related to one another somewhat indirectly and involving several mutational events whereas mtDNA types found in Caucasoids may be related to one another by single mutational events. This may account for the lower sequence divergence found in Caucasoids when compared with Negroids and Khoisan populations.

Table 3.11. The average nucleotide divergence calculated from the mean pairwise differences for all pairs of mtDNA types found within (diagonal) and between the twenty ethnic groups investigated from southern Africa.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Zulu	**0.194**	0.209	0.226	0.183	0.244	0.219	0.229	0.216	0.204	0.343	0.339	0.183	0.419	0.468	0.607	0.427	0.338	0.337	0.262	0.301
Swazi		**0.194**	0.223	0.189	0.254	0.241	0.231	0.252	0.218	0.284	0.316	0.179	0.452	0.492	0.649	0.441	0.266	0.249	0.185	0.226
Xhosa			**0.231**	0.215	0.262	0.237	0.243	0.262	0.225	0.337	0.347	0.197	0.434	0.474	0.611	0.431	0.333	0.306	0.253	0.296
Tsonga				**0.155**	0.251	0.227	0.215	0.208	0.201	0.334	0.335	0.169	0.471	0.528	0.689	0.457	0.284	0.286	0.216	0.251
Sotho					**0.251**	0.238	0.267	0.266	0.234	0.319	0.331	0.223	0.362	0.383	0.495	0.389	0.388	0.381	0.304	0.359
Pedi						**0.203**	0.245	0.239	0.206	0.367	0.355	0.198	0.345	0.386	0.488	0.379	0.417	0.405	0.325	0.372
Tswana							**0.246**	0.257	0.233	0.336	0.339	0.208	0.442	0.476	0.609	0.426	0.327	0.317	0.256	0.294
Venda								**0.214**	0.226	0.394	0.372	0.221	0.424	0.469	0.595	0.434	0.398	0.417	0.334	0.371
Lemba									**0.198**	0.339	0.336	0.182	0.378	0.419	0.544	0.399	0.366	0.356	0.279	0.324
Herero										**0.139**	0.284	0.299	0.454	0.433	0.591	0.461	0.249	0.238	0.178	0.211
Dama											**0.322**	0.312	0.468	0.458	0.588	0.459	0.351	0.348	0.281	0.312
Ambo												**0.149**	0.399	0.448	0.594	0.412	0.296	0.274	0.206	0.251
Sekeie													**0.256**	0.298	0.331	0.405	0.656	0.645	0.553	0.601
Nama														**0.257**	0.271	0.386	0.675	0.669	0.576	0.621
R_"Col"															**0.222**	0.421	0.874	0.873	0.771	0.818
JHB_"Col"																**0.411**	0.588	0.592	0.508	0.549
English																	**0.163**	0.146	0.122	0.139
A_Jew																		**0.079**	0.088	0.118
Indian																			**0.044**	0.782
Afrikaners																				**0.099**

R_"Col" = Richtersveld "Coloured"

JHB_"Col" = Johannesburg "Coloured"

A_Jew = Ashkenazi Jewish

3.4 Phylogenetic analysis

3.4.1 Phylogenetic network of mtDNA types

A phylogentic network showing the relationship between all mtDNA types discovered in this study and those analysed by Excoffier (1990) has been established following the method used by Johnson *et al.* (1983). In this method, the best network is derived by relating the existing mtDNA types through single site changes, *i.e.*, two types differing by a single apparent mutation (generally leading to a single restriction site difference, with some exceptions) are linked together (Excoffier 1990; Johnson *et al.* 1983).

The mtDNA types found in this study were added to the phylogeny constructed by Excoffier (1990) and the revised phylogeny presented in Fig. 3.7 has the following modifications: populations are individually represented by symbols and new types discovered in this study have been added to the phylogeny with the assigned mtDNA numbers 112-142 (types 78-111 have been described by others and communicated to Merriwether *et al.* (1991) who used these data to construct a complete UPGMA tree described later in Fig. 3.8a). Types 78-111 are not shown in the phylogeny constructed in Fig. 3.7. Note that all types described in the phylogeny are given for the first five enzymes only, so that comparisons with

other published data, which only made use of the five enzymes, could be made.

In the new phylogeny, mtDNA type 32 (3-1-1-5-1) can be derived directly from type 134 (3-1-1-2-1), which was only discovered in the present study, instead of from a hypothesized intermediate derived from type 3 (3-1-1-2-1) as shown in Fig. 3.7. The phylogeny given in Fig. 3.7 also suggests that mtDNA type 11 (2-2-3-1-5) is derived from an intermediate type which in turn is derived from mtDNA type 23. With the discovery of mtDNA type 128 (2-2-3-1-1), mtDNA type 11 can now be derived directly from type 128.

There has been considerable overlap and confusion in the assignment of mtDNA type numbers, in particular, in the assignment of mtDNA types 64 and 65. These mtDNA type numbers were assigned by Vikki *et al.* (1988), (2-1-1-4-13) as mtDNA type 64 and (2-1-1-1-1) as mtDNA type 65. Scozzari *et al.* (1988) called (3-1-7-1-3) type 64 and (2-1-12-1-1) type 65. The situation was further complicated by Excoffier (1990) who used mtDNA type numbers 76 and 77 to describe the types 64 and 65 published by Vikki *et al.* (1988) as presented in Fig. 3.7. Type numbers 76 and 77 had been allocated previously by Semino *et al.* (1989) to describe what are usually accepted as types 76 (2-1-1-14-1) and 77 (2-1-1-1-26).

These areas of confusion in the assignment of mtDNA type numbers emphasize the need for an organized committee whose role would be in the assistance of data coordination and prior distribution before publication. To overcome the confusion, the author has used mtDNA types 64-71, as presented in the phylogeny in Fig. 3.7, to correspond to mtDNA types as published by Scozzari *et al.* (1988); mtDNA types 72-77 according to Semino *et al.* (1989) and mtDNA type 63 (2-2-12-1-5) according to Vikki *et al.* (1988). MtDNA type 64 as published by Vikki *et al.* (1988) is referred to here as 64^{F} (where F refers to the Finnish population). Type 65 described by Vikki *et al.* (1988) is similar to type 1 except that it shows variation with the enzyme *Hinc*II and should therefore be denoted as a subtype 1-15 (2-1-1-1-1-15).

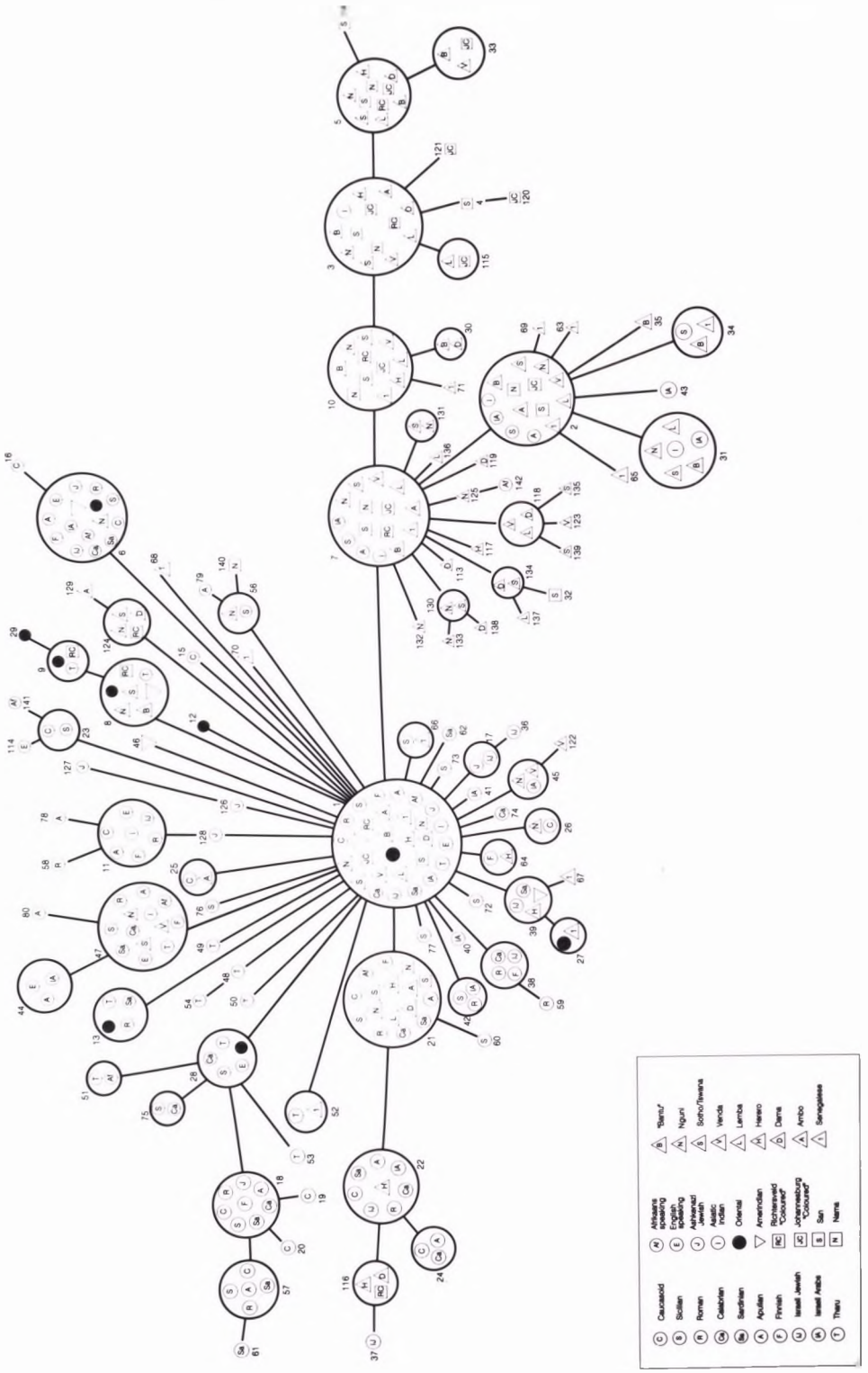
The distribution of mtDNA types in the phylogeny is consistent with the geographic origins and ethnicity of populations. Negroid and Khoisan populations from Africa are found in the longest branch of the network radiating from type 1, which includes those mtDNA types 7, 10, 3, 5 and 33 that are found in more than one ethnic group (see Fig. 3.7). The other common Negroid mtDNA type, type 2, is derived from type 7 from this branch.

In contrast there are several short lineages radiating from type 1 which consist mainly of non-African mtDNA types. The branch which includes mtDNA types 28, 18 and

57 (and mtDNA types which may be derived from them) is found predominantly in Caucasoid groups (see Fig. 3.7). Type 21 which was first discovered in Caucasoids (Johnson *et al*. 1983) is found at high frequency in southern African Herero and Dama ethnic groups (see Table 3.7). Types 22, 24 and 60 which are derived from type 21, consist mainly of mtDNA types found in Caucasoid populations, with the exception of type 22 which is found in one Herero individual. However, type 116 which is also derived from type 21 is found at low frequencies in the Herero, Dama and Richtersveld "Coloured" groups.

There are a few types (2,3,7,10 and 34) which have the *Hpa*I-3 morph which are more common in African populations but which are also found, albeit at lower frequencies, in some Caucasoid and Oriental populations. It was suggested by Semino *et al*. (1989) that the *Hpa*I-3 morph may be used to estimate the extent of the non-Caucasoid maternal ancestry in Caucasoid groups. They estimated that approximately 10.8% of mtDNA types found in Sicilians were derived from Negroids either directly or indirectly through Arabic migrations into Sicily. This is also observed with nuclear DNA marker frequencies.

Fig. 3.7. Partial phylogeny of mtDNA types showing the relationship of some of the mtDNA types discovered to date. MtDNA types 112-142 were discovered in the present study and types 1-77 were described previously in the phylogeny presented by Excoffier (1990). MtDNA types 78-111 are not shown in this phylogeny but are in Fig. 3.8a. Large circles represent the mtDNA type number and the symbols and the letter code describe the populations who were found to have the respective types. All types are related to one another by the least number of mutational steps according to Johnson *et al.* (1983).


Caucasoid
Sicilian
Roman
Calabrian
Sardinian
Apulian
Finnish
Israel Jewish
Israel Arabs
Thiaru
Afrikaans speaking
English speaking
Ashkenazi Jewish
Asiatic Indian
Oriental
Amerindian
Richtersveld "Coloured"
Johannesburg "Coloured"
San
Nama
'Bantu'
Nguni
Sotho/Tswana
Venda
Lemba
Herero
Dama
Ambo
Senegalese

Excoffier (1990) claimed that type 1 was the ancestral mtDNA type based on the following observations: firstly, it was the only mtDNA type found in all geographic and ethnic groups studied by this method to date; secondly, type 1 satisfied the criteria that the hypothesized ancestral type should possess all polymorphic sites in their most frequent states. Stoneking (1990) has opposed this claim on two grounds: firstly, the addition of more enzymes in the derivation of a mtDNA types resolved type 1 into several other subtypes (refer to Blanc *et al.* 1983). Secondly, the observation that type 1 was found in all population groups studied was another indication of the lack of resolution of this method of analysis.

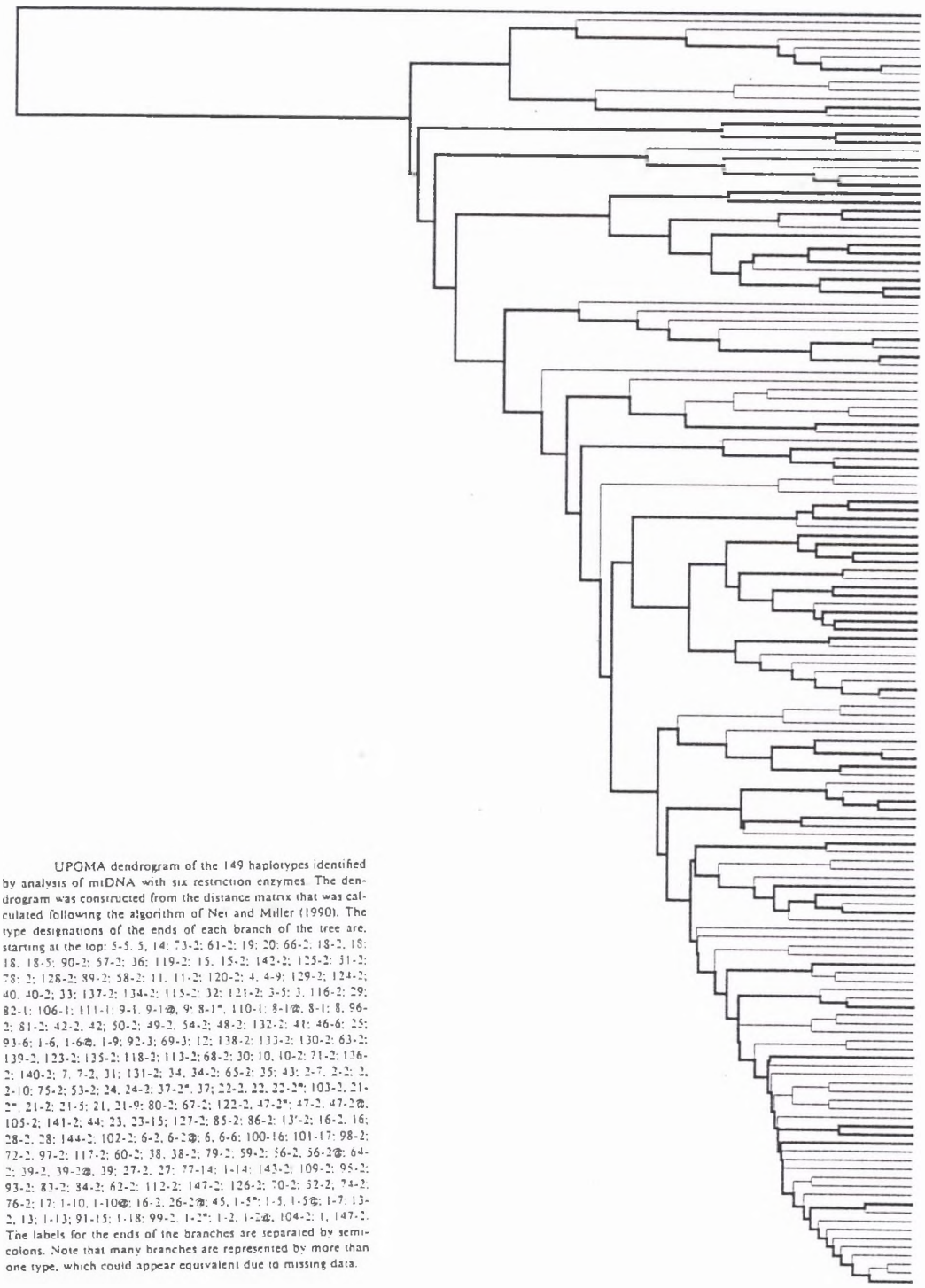
There are several factors (discussed in section 4.4.2) which have to be considered when assigning ancestral states using parsimony analysis. Such an investigation would require development of statistical methods of inference of ancestral states that rely on detailed models of migrations and evolutionary change (Maddison 1991). Since these criteria could not be satisfied by the parsimony method used in the construction of the phylogeny presented in Fig. 3.7, it was not possible to determine which mtDNA type is ancestral to all the others. Although mtDNA type 1 seems the most likely ancestor, this could not be determined with certainty.

3.4.2 UPGMA analysis of mtDNA types

A UPGMA phylogentic tree showing the relationship of all 149 mtDNA types defined by the 81 variant sites obtained when using the restriction enzymes *Hpa*I, *Bam*HI,*Hae*II, *Msp*I, *Ava*II and *Hinc*II is shown in Fig. 3.8a. The mtDNA types found in the present study are emphasized by the thickened lines in the figure. MtDNA types found in Africans, Europeans, Asians and Amerindians are shown in Fig. 3.8b.

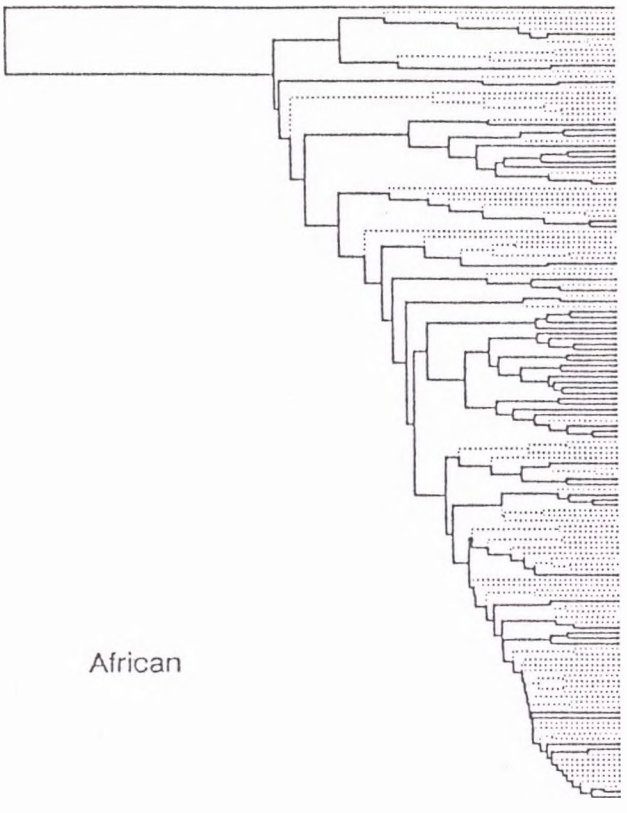
The tree given in Fig. 3.8a reveals two important features: Firstly, southern African mtDNA types are scattered throughout the tree suggesting that mtDNA types in these populations are associated with both "older" (more sequence divergence) and more recent mutations (less sequence divergence); Secondly, when compared to types found in other major groups, *viz.*, Caucasoids, Asians and Amerindians, African mtDNA types have the most diversity (see Fig. 3.8b). This observation is coupled with the scenario that the oldest parental population is expected to have the greatest diversity and that new daughter populations have less variability. These findings lend further support to the hypothesis that human mtDNA evolved in Africa (Cann *et al.* 1987).

Fig. 3.8 (a) UPGMA tree showing the relationship of the 62 types discovered in this study (bold) when compared with the 149 mtDNA types discovered to date using the six enzyme system (Merriwether *et al*. 1991).

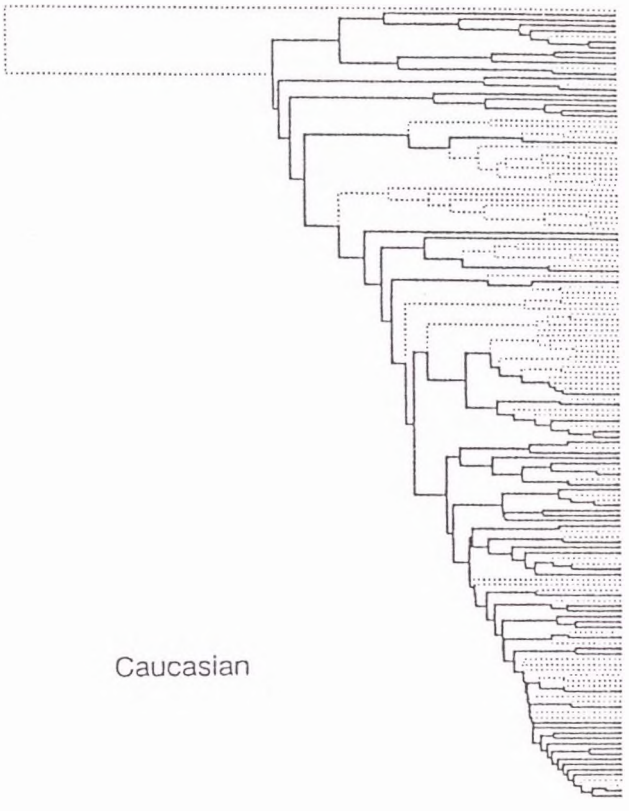


UPGMA dendrogram of the 149 haplotypes identified by analysis of mtDNA with six restriction enzymes. The dendrogram was constructed from the distance matrix that was calculated following the algorithm of Nei and Miller (1990). The type designations of the ends of each branch of the tree are, starting at the top: 5-5, 5, 14; 73-2; 61-2; 19; 20; 66-2; 18-2, 18; 18, 18-5; 90-2; 57-2; 36; 119-2; 15, 15-2; 142-2; 125-2; 31-2; 78; 2; 128-2; 89-2; 58-2; 11, 11-2; 120-2; 4, 4-9; 129-2; 124-2; 40, 40-2; 33; 137-2; 134-2; 115-2; 32; 121-2; 3-5; 3, 116-2; 29; 82-1; 106-1; 111-1; 9-1, 9-1@, 9; 8-1*, 110-1; 8-1@, 8-1; 8, 96-2; 81-2; 42-2, 42; 50-2; 49-2, 54-2; 48-2; 132-2; 41; 46-6; 25; 93-6; 1-6, 1-6@, 1-9; 92-3; 69-3; 12; 138-2; 133-2; 130-2; 63-2; 139-2, 123-2; 135-2; 118-2; 113-2; 68-2; 30; 10, 10-2; 71-2; 126-2; 140-2; 7, 7-2, 31; 131-2; 34, 34-2; 65-2; 35; 43; 2-7, 2-2; 2, 2-10; 75-2; 53-2; 24, 24-2; 37-2*, 37; 22-2, 22, 22-2*; 103-2, 21-2*, 21-2; 21-5; 21, 21-9; 80-2; 67-2; 122-2, 47-2*; 47-2, 47-2@, 105-2; 141-2; 44; 23, 23-15; 127-2; 85-2; 86-2; 13*-2; 16-2, 16; 28-2, 28; 144-2; 102-2; 6-2, 6-2@; 6, 6-6; 100-16; 101-17; 98-2; 72-2, 97-2; 117-2; 60-2; 38, 38-2; 79-2; 59-2; 56-2, 56-2@; 64-2; 39-2, 39-2@, 39; 27-2, 27; 77-14; 1-14; 143-2; 109-2; 95-2; 93-2; 83-2; 34-2; 62-2; 112-2; 147-2; 126-2; 70-2; 52-2; 74-2; 76-2; 17; 1-10, 1-10@; 16-2, 26-2@; 45, 1-5*; 1-5, 1-5@; 1-7; 13-2, 13; 1-13; 91-15; 1-18; 99-2, 1-2*; 1-2, 1-2@; 104-2; 1, 147-2. The labels for the ends of the branches are separated by semicolons. Note that many branches are represented by more than one type, which could appear equivalent due to missing data.

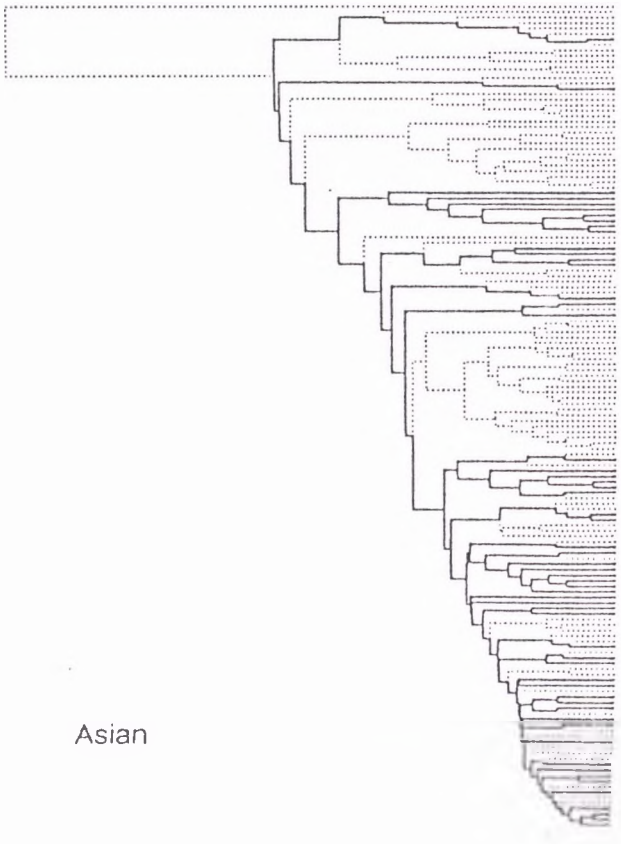
Fig. 3.8 (b). UPGMA dendrograms of all mtDNA types desribed in Fig. 3.8a showing the relationship of mtDNA types found in the pooled African, Asian, Caucasian and Amerindian populations (highlighted paths) (Merriwether *et al*. 1991).



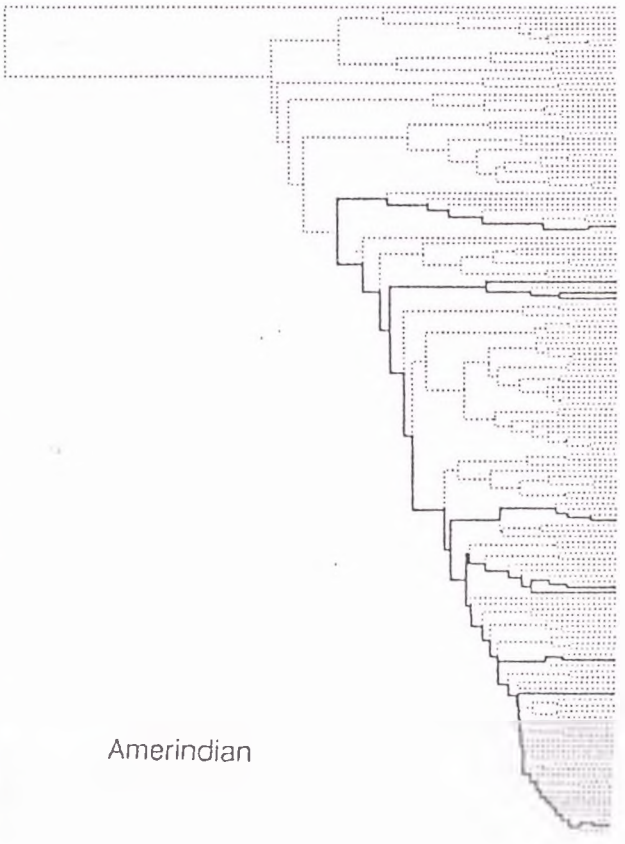
African



Caucasian



Asian

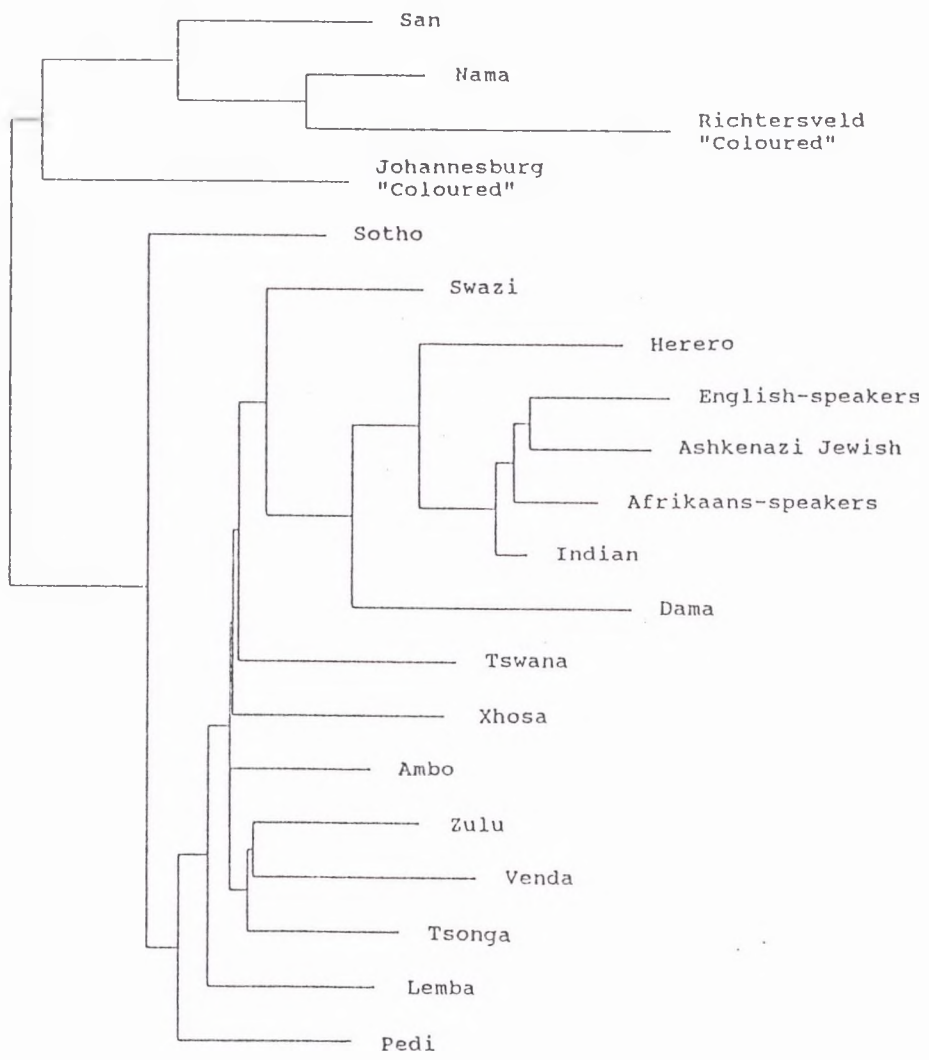


Amerindian

3.4.3 Cluster analysis

A neighbour-joining tree (Fig. 3.9) was constructed from the matrix of average distances to show the clustering affinites of the 20 ethnic groups investigated in the present study. The dendrogram reveals three different clusters. The first consists of the Khoisan and "Coloured" groups, the second of Negroid groups (excluding the Herero and Dama) and the third consisting of the Caucasoid groups. The Herero and Dama are more closely related to the Caucasoid cluster than to the Negroid cluster and this may be due to the high frequency of mtDNA types in these groups having morph *Hpa*I-2.

There are small differences between Negroid linguistic groups and this may be correlated with the relatively recent assignment of linguistic divisions in Bantu-speaking Negroids. A more detailed account of the significance of difference noted between ethnic groups is discussed in the Chapter 5.



Branch length 10mm = 0.43

Fig. 3.9. Neighbour-joining cluster dendrogram showing the genetic relationship of twenty southern African populations investigated for mtDNA variations.

3.4.4. Do mtDNA mutations provide a selective advantage?

Both the Ewens-Watterson (Ewens 1972; Watterson 1977) and the Tajima tests (Tajima 1989) reveal a theoretical departure from selective neutrality (Table 3.12). This is supported by an excess of homozygosity and an excess of segregating sites observed for most populations (Table 3.12). On the other hand, neutrality tests based on mtDNA variation analysed by high resolution mapping (Whittam *et al*. 1986) conformed to neutral expectations for most of the functional domains on the mtDNA molecule. However, some alleles occurred at much higher frequencies than is expected by the model, in particular, those regions that encode proteins and tRNAs.

Whittam *et al*. (1986) examined the assumptions of the Ewens-Watterson test and tried to explain why deviations from the neutral mutation model occurred with mtDNA restriction site data, and their findings are applicable to the data presented here. One of the assumptions underlying the statistical theory is that the distribution of allele frequency should achieve a steady state between mutational gain of new variants and the stochastic loss of alleles through genetic drift. Since human populations have undergone a recent expansion, such a steady state has not as yet been attained (Nei and Graur 1984). Theory predicts that during the period of population expansion, the expected number of neutral alleles increases more rapidly

than the expected genetic diversity (Nei and Li 1979). The excess of singletons observed in all populations (Table 3.12) studied here and those by (Whittam *et al*. 1986) supports this prediction, hence the deviation from neutrality.

Another assumption is that all alleles at a locus represent strictly neutral mutations whose dynamics have been affected solely by random genetic drift (Ohta and Kimura 1969). As an alternative to the effect of population expansion, the effect of purifying selection favouring the most common allele within each population may also contribute to the departure observed (Whittam *et al*. 1986). Excoffier (1990) claimed that selection is the main cause of departure from the neutral model. This was strongly critisized by Stoneking (1990) who suggested that the excess of homozygosity seen with the low resolution technique is due to the overestimation of the common type (type 1) which could be resolved further by the addition of more enzymes. On the other hand Excoffier (1990) demonstrated that the neutrality hypothesis is rejected even with high resolution mtDNA data. Thus, it seems unlikely that the departure from neutrality observed with the RFLP data in this study is merely an artifact of the low resolution technique as argued by Stoneking (1990).

Thus, there are several possible explanations for the rejection of neutrality. These include natural selection,

population growth or failure to attain equilibrium, population heterogeneity and inhomogeneity in the neutral substitution rate as discussed by Merriwether *et al*. 1991. Since the mitochondrial genome is largely coding, with the largest non-coding region limited to the control region, it is possible that some of the genetic variation could be the result of selection. This could, in part, account for the higher rate of mutation in the non-coding control region compared to other parts of the molecule. Although mtDNA mutations have been associated with a wide range of mitochondrial myopathies (Wallace *et al*. 1988), there is no evidence to date to suggest that this is due to selection.

Table 3.12. The Ewens-Watterson and the Tajima tests of neutrality on the restriction site variation observed in twenty southern African populations.

	Ewens-Watterson test								Tajima Test			
	N	h	F	*p*	S	*p*	C	*p*	s	k	D	*p*
Zulu	30	8	0.222		4		0.300		10	1.761	-0.958	
Swazi	41	14	0.183		10	+ +	0.293		28	3.324	-1.677	
Xhosa	18	9	0.185		6	+	0.333		12	2.111	-1.458	
Tsonga	34	7	0.275		4	+	0.382		11	1.405	-1.498	
Sotho	29	10	0.153		5		0.241		13	2.305	-1.011	
Pedi	23	6	0.301		2		0.478		9	1.881	-0.757	
Tswana	15	9	0.147		6		0.200		12	2.305	-1.463	
Venda	30	11	0.173		5		0.333		9	1.968	-0.416	
Lemba	53	13	0.192		7	+	0.340		14	1.795	-1.256	
Herero	54	9	0.366		6	+ +	0.500		27	1.926	-2.199	---
Dama	43	12	0.175		7		0.318		14	2.927	-0.281	
Ambo	22	7	0.343		4	+	0.545		8	1.364	-1.243	
Sekele	49	9	0.215		3		0.286		24	5.148	0.104	
Nama	46	8	0.318		2		0.532		12	2.296	-0.463	
R_"Col"	35	9	0.338	+	2		0.556	+	22	3.664	-1.053	
JHB_"Col"	71	13	0.155		6		0.239		17	3.630	-0.094	
English	36	7	0.434		2		0.639		12	1.499	-1.569	
Afrikaners	51	6	0.706	+ +	5	+ +	0.837	+ + +	10	0.842	-1.793	-
A_Jews	55	7	0.479		4	+	0.618		9	0.702	-1.771	-
Indians	50	7	0.777	+ +	6	+ +	0.880	+ + +	6	0.392	-1.801	

N = sample size
h = the number of distinct mtDNA types
F = the homogygosity
S = the number of singletons
C = the frequency of the most common allele
s = the number of segregating sites
k = the estimate of O derived from the infinite alleles sampling theory
D = Tajima's (1989) test statistic
p = probability of a random value exceeding the observed value; + ($p<0.05$), + + ($p<0.01$), + + + ($p<0.001$)
Similarly, negative signs indicate signficance of a one-tailed test with the observed values lower than expected

3.5 Conclusions

1. Using the restriction enzymes *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II, 62 mtDNA types were discovered in the 795 southern African individuals sampled. Only 32 of these mtDNA types were previously discovered, whilst the remaining 30 are new and arise mainly due to new combinations of existing morphs. There were only four new morphs discovered: *Msp*I-17$^{S.A\ Jew}$, *Ava*II-31$^{S.A\ Col}$, *Ava*II-32$^{S.A\ Col}$ and *Ava*II-33$^{S.A\ Col}$.

2. Khoisan populations were found to have more sequence divergence (0.256) than Negroid (0.208%) and Caucasoid (0.096%) groups, and in general, there was more variation between groups than within groups. The variance contributed by differences between populations were estimated using the G_{ST} index, and was found to be 0.163 by the Nei (1987) method and 0.180 by the Takahata and Palumbi (1985) method. When compared to the value of 0.350 obtained by Merriwether *et al.* (1991) on 62 populations (including the 20 studied here), the lower value for southern African populations may be due to the high number of indigenous Africans in the southern African sample studied here (497 out of 795) compared to (863 out of 3065) in the Merriwether *et al.* (1991) sample. The G_{ST} value is therefore a good indicator of the geographic structuring of populations.

3. Both an excess of homozygosity and an excess of segregating sites revealed a theoretical departure from neutral expectations. Thus, mtDNA data invalidate the assumption that the genetic variation is neutral and that the populations studied are in mutation drift equilibruim. There are several reasons for the departure from neutrality and these include selection, population growth or failure to attain equilibrium, population heterogeneity and inhomogeneity in the neutral substitution rate.

4. Phylogenetic analysis making use of the UPGMA algorithm reveal that mtDNA types found in southern African populations are scattered along several branches of the tree, which suggests than the groups sampled here have old and relatively new mutations which characterize the evolutionary history of human populations.

5. When analyzed in the form of a parsimony network, the mtDNA types are organized into distinct lineages which separate African (Negroid and Khoisan) from non-African (Caucasoid) mtDNA types. Though not the rule, the mtDNA types found in indigenous African populations are located predominantly in branches which require several mutational steps relative to type 1, whereas mtDNA types in Caucasoid populations are usually related to type 1 by a single mutational event. These findings provide additional support for the suggestion that

African populations are more diverse than non-African populations, which is one of the arguments for the hypothesis that modern humans originated in Africa.

6. Neighbour-joining tree analysis reveals three distinct clusters of southern African populations. The first, includes the two hybrid groups together with the Sekele and Nama groups. The second cluster, with the exception of the Ambo, groups the different Negroid populations classified as southeastern Bantu-speakers, who, as might be expected, are separated by very little genetic distance. The Herero and Dama groups cluster closely with the Caucasoid groups in cluster 3.

7. These findings indicate that, although the six-enzyme system used here has low resolution, thi s method of analysis is extremely powerful in elucidating the genetic structure of southern African populations.

CHAPTER FOUR:
CONTROL REGION SEQUENCE VARIATION IN SOUTHERN AFRICAN POPULATIONS

4.1 Background

In the previous chapter the usefulness of mtDNA RFLPs in elucidating population relationships among different ethnic groups in southern Africa was discussed. However, mention has been made of certain shortcomings of the Southern blot approach when compared with high resolution mapping (see Section 1.5.2). With the development of methodology for mtDNA control region sequencing (Greenberg *et al.* 1983), this method is presently the method of choice for studying the evolution of human mtDNA for the following reasons: Firstly, the control region is non-coding and appears to accumulate mutations at a rate approximately ten times faster than the rest of the molecule (Vigilant 1990). This suggests that the control region on its own would yield as much, if not more, information on population variation than would be obtained by restriction enzyme analyses of the entire mitochondrial genome.

Secondly, DNA sequence analysis allows one to characterize mutations unambiguously with respect to the position and the type of mutation. Such information is fundamental in elucidating the mechanisms involved in creating mutations in mtDNA.

Thirdly, DNA amplification by the polymerase chain reaction (PCR) alleviates the requirement for large amounts of starting material such as blood and/or placental tissue, since mtDNA may be amplified exponentially from nanogram quantities from a range of sources of DNA such as single hairs, blood spots and buccal cells from a mouth wash. The advantage of using material obtained from the latter sources are that the methods of collection are noninvasive and less traumatic to subjects and can be conducted by non-medically qualified people.

More recently, PCR has been used to amplify mtDNA from ancient humans (mummies, human tissue preserved in bogs and sketetal tissue) which have several important benefits in evolutionary studies. Firstly, the mtDNA structure of humans who lived several thousand years ago may be elucidated and compared with that of contemporary humans. The comparative analysis of contemporary versus ancient human mtDNAs in a known time interval may help refine estimates of the mutation rate of mtDNA. Secondly, the study of mtDNA variation in ancient tissue is extremely useful in the reconstruction of mankind's prehistory.

This chapter discusses the results obtained from mtDNA sequence analysis in 144 individuals, most of whom were previously analyzed by the RFLP method. Consequently, it was possible to make comparisons of the two methodologies.

4.2 Sequence analysis

4.2.1 Variant sites

A total of 144 mtDNA control region sequences were obtained from individuals sampled from the ethnic groups shown in Table 4.1. The primary objective of this section of the study was to investigate variation contained in the control region in indigenous southern African populations. At the same time, a few Caucasoids (7) and "Coloured" individuals (3) found to contain rare mtDNA types as deduced from RFLP studies were also sequenced and included here for comparison.

By pairwise comparisons of all mtDNA sequences, 119 out of approximately 750 bp sequenced, were variant. Only the informative sequence data are listed for all individuals sequenced in this study and is given in Appendix 4. A comparison in the frequencies of variant sites contained within the control region was made for Negroid, Khoisan and Caucasoid populations to establish whether there were any population-specific differences (Table 4.2 and Fig. 4.1). The comparisons based on the Caucasoid sample are insignificant due to the small number of individuals used in this comparison.

Table 4.1. The sample sizes and the origins of the various ethnic groups investigated for mtDNA control region sequence variation.

POPULATION	ETHNIC GROUP	SAMPLE SIZE		LOCATION	REFERENCE
NEGROID			**172**		
	"Bantu"	45		South Africa	present study
	Herero_1	12		Namibia	present study
	Herero_2	26		Botswana	Vigilant et al. 1991
	Dama	21		Namibia	present study
	Eastern Pygmies	20		Zaire	Vigilant et al. 1991
	Western Pygmies	17		Central African Republic	Vigilant et al. 1991
	Hadza	17		Tanzania	Vigilant et al. 1991
	Yoruban	14		Nigeria	Vigilant et al. 1991
KHOISAN			**82**		
	Sekele	38		Namibia	present study
	!Kung	25		Botswana	Vigilant et al. 1991
	Nama	18		Namibia	present study
	Naron	1		Botswana	Vigilant et al. 1991
CAUCASOID			**21**		
	South African	7		South Africa	present study
	European	14		Western European	Vigilant et al. 1991
ASIAN	-		**23**	China	Vigilant et al. 1991
PAPUA NEW GUINEAN	-		**20**	Papua New Guinea	Vigilant et al. 1991
"OTHER"			**12**		
	African American	8		USA	Vigilant et al. 1991
	Australian Aborigine	1		Australia	Vigilant et al. 1991
	South African "Coloured"	3		South Africa	present study

Of the 119 variant positions identified, 79 were found in the Khoisan, 95 in the Negroid and only 23 in the Caucasoid group. Only 16 of these sites are found in all three groups (see Fig. 4.1), usually at a frequency greater than 30% (see Table 4.2), suggesting that these sites were present in human ancestors before population differentiation. Though not the rule, the frequency of the common sites in all three groups are found at higher frequencies in the Khoisan and Negroid groups than in the Caucasoid group (see Fig. 4.1).

In addition to sites shared by all three groups, there were 45 variant sites found only in the Negroid and the Khoisan groups, whilst there were two sites which were found in Negroid and Caucasoid groups. On the other hand, only one site was common to the Khoisan and Caucasoid groups. There were more unique variant positions in the Negroid group (32) than in the Khoisan (16) and Caucasoid (4) groups.

Table 4.2. Frequencies of the variant sites in the Khoisan, Negroid and Caucasoid populations determined by counting the number of individuals (M) who differed from the reference sequence at the posititons given divided by the number of individuals from whom sequence data was available (N).

	Khoisan			Negroid			Caucasoid		
Site position	N	M	Freq	N	M	Freq	N	M	Freq
46	51	1	0.020	75	1	0.013	7	-	-
70	56	4	0.071	76	3	0.039	7	-	-
91	56	2	0.036	76	-	-	7	-	-
101	56	4	0.071	76	14	0.184	7	-	-
106	56	16	0.286	76	26	0.342	7	2	0.286
117	56	-	-	76	1	0.013	7	-	-
122	56	1	0.018	76	2	0.026	7	-	-
125	56	-	-	76	7	0.092	7	-	-
129	56	-	-	76	1	0.013	7	-	-
130	56	1	0.018	76	1	0.013	7	-	-
144	56	12	0.214	76	-	-	7	-	-
145	56	-	-	76	5	0.066	7	-	-
146	56	-	-	76	2	0.026	7	-	-
149	56	7	0.125	76	9	0.118	7	-	-
151	56	1	0.018	76	1	0.013	7	-	-
154	56	1	0.018	76	-	-	7	-	-
156	56	-	-	76	1	0.013	7	-	-
157	56	4	0.071	76	2	0.026	7	-	-
162	56	1	0.018	70	-	-	7	-	-
163	56	2	0.036	70	5	0.071	7	-	-
165	55	40	0.727	62	32	0.516	7	3	0.429
166	55	3	0.054	62	4	0.064	7	-	-
167	55	43	0.782	56	35	0.625	7	3	0.429
168	56	-	-	62	1	0.016	7	-	-
170	56	-	-	77	1	0.013	7	-	-
181	56	-	-	76	1	0.013	7	-	-
187	56	6	0.107	76	5	0.066	7	-	-
190	56	3	0.053	76	6	0.079	7	-	-
191	56	2	0.036	76	1	0.013	7	-	-
192	56	10	0.179	76	2	0.026	7	-	-
196	56	2	0.036	76	3	0.040	7	1	0.142
201	56	53	0.946	76	69	0.907	7	2	0.286
202	56	-	-	76	-	-	7	1	0.143
208	56	37	0.661	76	19	0.250	7	-	-
212	56	17	0.304	76	1	0.013	7	-	-
217	56	5	0.089	76	6	0.079	7	-	-
220	56	12	0.214	76	1	0.013	7	-	-

Site position	Khoisan			Negroid			Caucasoid		
	N	M	Freq	N	M	Freq	N	M	Freq
221	56	28	0.500	76	13	0.171	7	-	-
234	56	1	0.018	77	3	0.039	0	-	-
237	56	-	-	77	1	0.013	0	-	-
240	56	2	0.036	77	-	-	0	-	-
242	56	1	0.018	77	-	-	6	1	0.167
243	56	1	0.018	77	6	0.078	6	-	-
244	56	1	0.018	77	1	0.013	6	-	-
248	56	1	0.018	78	-	-	6	-	-
252	56	2	0.036	78	2	0.026	6	-	-
256	56	18	0.321	78	32	0.411	6	2	0.333
262	56	2	0.036	78	-	-	6	-	-
264	56	1	0.018	78	4	0.051	6	1	0.167
268	56	3	0.053	77	4	0.052	6	-	-
269	56	8	0.143	77	3	0.039	6	-	-
270	56	1	0.018	77	2	0.026	6	-	-
271	56	-	-	77	5	0.064	6	1	0.167
272	56	8	0.143	77	19	0.247	6	3	0.500
273	56	1	0.018	77	-	-	6	-	-
274	56	1	0.018	77	2	0.026	6	-	-
278	56	1	0.018	77	2	0.026	6	-	-
280	56	-	-	77	1	0.013	6	-	-
282	56	1	0.018	77	11	0.143	6	-	-
287	56	-	-	77	1	0.013	6	-	-
289	56	36	0.643	76	33	0.434	6	1	0.167
290	56	1	0.018	77	-	-	6	-	-
297	56	-	-	77	1	0.013	6	-	-
293	56	1	0.018	77	12	0.156	6	-	-
303	56	2	0.036	76	3	0.039	6	-	-
305	56	2	0.036	76	4	0.053	6	-	-
333	56	2	0.036	75	-	-	6	-	-
335	29	-	-	66	1	0.015	0	-	-
338	29	-	-	46	5	0.109	0	-	-
340	29	1	0.034	46	2	0.043	0	-	-
346	21	-	-	34	1	0.029	0	-	-
368	1	-	-	14	1	0.071	0	-	-
625	37	36	0.973	51	41	0.803	6	5	0.833
637	48	1	0.021	75	-	-	6	-	-
639	51	1	0.020	75	-	-	6	-	-
645	54	-	-	75	3	0.040	7	-	-
647	54	-	-	75	4	0.053	7	-	-
649	54	-	-	75	1	0.013	7	-	-
650	54	-	-	75	1	0.013	7	-	-

Site position	Khoisan			Negroid			Caucasoid		
	N	M	Freq	N	M	Freq	N	M	Freq
666	56	-	-	75	-	-	7	2	0.286
672	56	1	0.018	76	-	-	7	-	-
695	56	-	-	77	1	0.013	7	-	-
698	56	39	0.696	77	25	0.325	7	3	0.429
702	56	9	0.161	77	19	0.247	7	-	-
703	56	4	0.071	77	8	0.103	7	1	0.143
704	56	50	0.089	77	44	0.541	7	-	-
705	56	-	-	77	-	-	7	1	0.143
716	56	1	0.018	77	-	-	7	-	-
724	56	-	-	77	1	0.013	7	-	-
734	56	3	0.053	77	11	0.143	7	-	-
735	56	-	-	77	1	0.013	7	-	-
736	56	1	0.018	77	-	-	7	-	-
737	56	-	-	77	7	0.091	7	-	-
738	56	-	-	77	5	0.065	7	-	-
740	56	3	0.054	77	-	-	7	-	-
741	56	10	0.178	77	27	0.351	7	-	-
748	56	45	0.803	77	41	0.532	7	1	0.143
751	56	24	0.429	77	13	0.169	7	-	-
752	56	2	0.036	77	1	0.013	7	1	0.143
753	56	3	0.054	77	9	0.117	7	-	-
757	56	3	0.054	77	4	0.052	7	-	-
760	56	4	0.071	77	4	0.052	7	-	-
778	56	-	-	78	-	-	7	-	-
789	56	-	-	78	3	0.038	7	-	-
800	56	39	0.696	78	34	0.436	7	-	-
804	56	-	-	78	1	0.013	7	-	-
816	56	11	0.196	78	41	0.526	7	3	0.429
818	56	5	0.089	78	1	0.013	7	-	-
838	56	-	-	78	-	-	7	1	0.143
847	56	1	0.018	78	2	0.026	7	-	-
851	56	-	-	78	2	0.026	7	1	0.143
864	56	11	0.196	76	17	0.224	7	-	-
867	56	-	-	78	1	0.013	7	-	-
874	56	2	0.036	78	6	0.077	6	2	0.333
875	56	-	-	78	1	0.013	6	-	-
878	56	-	-	78	1	0.013	6	-	-
883	56	-	-	78	2	0.026	6	-	-
915	52	1	0.019	77	-	-	6	-	-

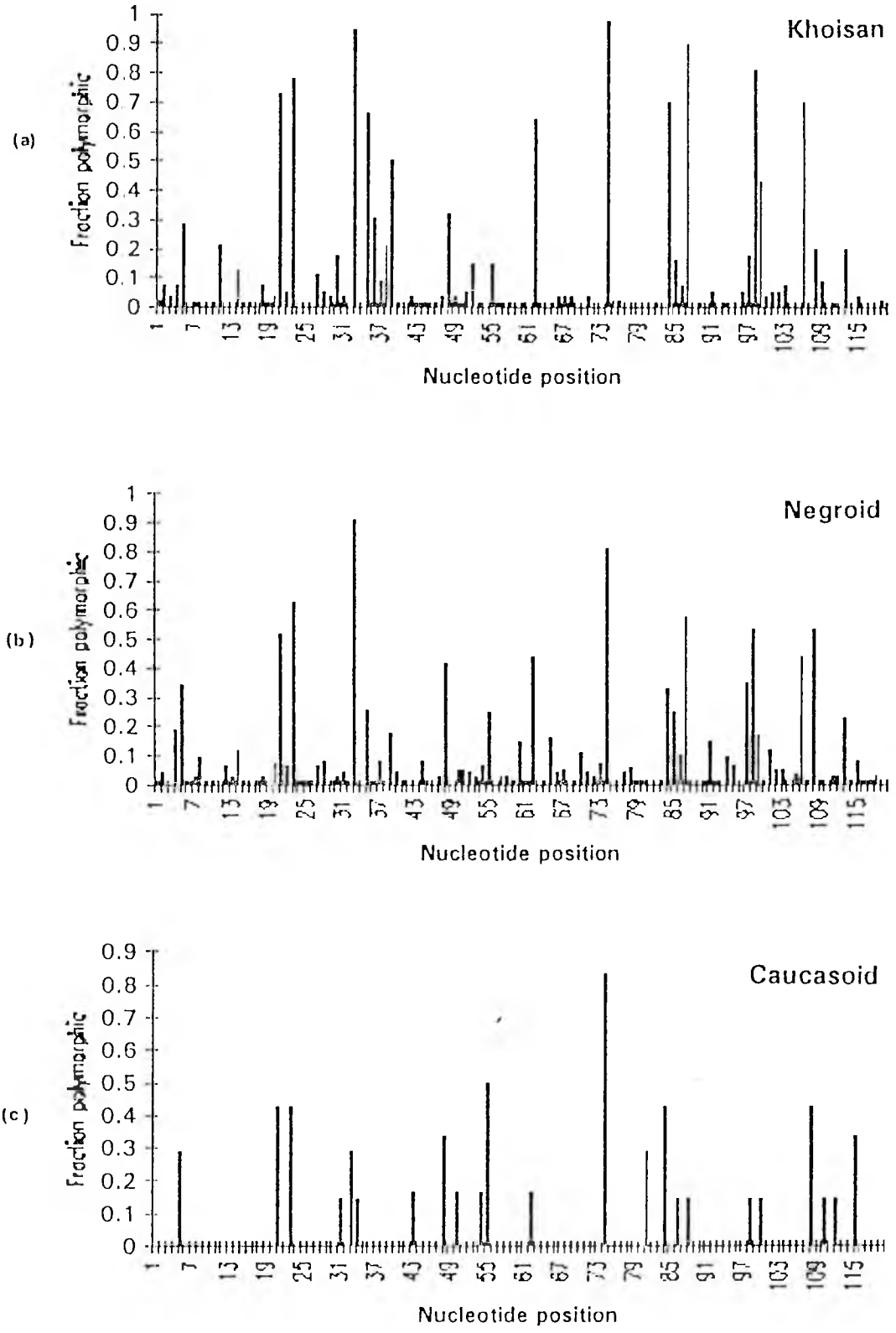


Fig. 4.1. Frequency of variant sites contained within the two hypervariable regions of the mtDNA molecule in Khoisan (a), Negroid (b) and Caucasoid (c) populations.

4.2.2 Characterization of mutations

4.2.2.1 Substitutions

The majority of mutations consists of single-base type substitutions which result from either transitions (purine to purine or pyrimidine to pyrimidine) or transversions (purine to pyrimidine). There were approximately twice as many substitution type mutations contained within HVS-I (73) than within HVS-II (44). Transition mutations (101) are commoner than transversions (9) with an approximate ratio of 11:1 (Table 4.3).

At some positions, more that one type of substitution resulted in a mutation (multiple mutations). Such mutations were found at positions 166, 234, 243, 264, 268, 269 and 748, and the types of substitutions at these positions are given in Table 4.3. Six of the seven multiple mutation sites occur in HVS-I whilst only one is found in HVS-II.

In fact, 79 of the 119 mutations found in the control region are located within the region which forms the displacement loop during H-strand synthesis. This region spans nucleotide positions 150-737. Four mutations were found in the termination-associated sequence (positions 135-149), one mutation in CSB-1 whilst a length mutation and two single substitution type mutations occur in CSB-2.

Table 4.3. Characterization of mutations in the two hypervariable regions contained within the control region of the mtDNA molecule.

Single base substitutions

Type of mutation	Hypervariable segment I	Hypervariable segment II	Total
Transitions			
C to T	32	11	43
T to C	16	7	23
A to G	9	11	20
G to A	7	8	15
Total	**64**	**37**	**101**
Transversions			
C to A	2	1	3
C to G	1	1	2
A to C		1	1
G to C		1	1
T to A		1	1
G to T		1	1
Total	**3**	**6**	**9**

Mutliple base substitutions

Position	Mutation	Number with mutation	Type of mutation
166	C to T	8	Transition
	C to G	1	Transversion
	C to A	1	Transversion
234	C to A	3	Transversion
	C to T	1	Transition
243	A to G	1	Transition
	A to C	5	Transversion
	A to T	1	Transversion
264	C to A	3	Transversion
	C to G	3	Transversion
268	C to T	6	Transition
	C to G	1	Transversion
269	C to T	3	Transition
	C to G	7	Transversion
748	T to C	89	Transition
	T to A	1	Transverion

No mutations were found in the other 5 regions to which replication initiation functions have been attributed.

Transitions involving substitutions from a C to T (43) and T to C (23) are commoner than A to G (20) and G to A (15) type substitutions. Although mitochondrial polymerases have been identified, no proofreading functions have been demonstrated for these enzymes, which may account for the high rate of mutation in mtDNA (Kornberg and Baker 1991).

4.2.2.2 Length mutations

In addition to substitution type mutations a length variation mutation was also observed. This occurs in the region which lies between positions 849 and 861 which encodes part of CSB-2. This region contains a stretch of seven C residues followed by a T and then another five C residues in the reference sequence (Anderson *et al.* 1981). All individuals sampled in the present investigation were found to have six C residues (instead of five) succeeding the T nucleotide at position 862. In 33.3% of sequences analyzed, the region preceding the T at position 862 consisted of eight C's instead of seven (Fig 4.2). This is due to the insertion of a C residue. This mutation occurs at a similar frequency in Vigilant's (1990) sample of 189 individuals. It has been suggested that the slipped mispairing mechanism reported by Streisinger *et al.* (1966) and Albertini *et al.* (1982) may be involved in the

generation of these length variations. Since the exact mechanism whereby insertion/deletion type mutations occur is not sufficiently understood, this mutation was omitted in the phylogenetic analyses.

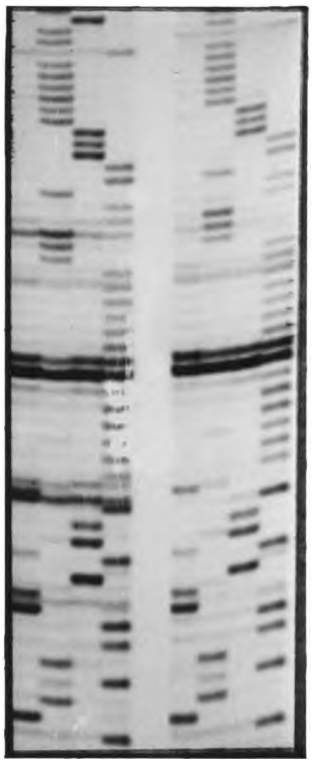


Fig. 4.2. Length mutation caused by an insertion of a C residue succeeding the T at position 862. This muation was found in 33.3% of the sequences analysed.

4.3 MtDNA sequence types

Pairwise comparisons of sequences revealed 129 unique mtDNA types (see Appendix 4). Of these 7 were shared either within or between linguistic groups (Table 4.4): three types were shared by the following individuals within the Sekele population ([Om30 and Om34]; [Om114, Om126, Om135, Om141 and Om148]; [Om123 and Om127]); two individuals from the "Bantu" group (Zu33 and Zu35) had a common type; one type is found in one Dama (Da18) and three Herero (He9, He27 and He28) individuals; another type is found in one Herero (He12), one "Bantu" (Zu37) and four Dama individuals (Da7, Da8, Da9 and Da26). The low extent of sharing of types between Negroid and Khoisan populations may be due either to the high resolving power of the sequencing approach or it may suggest that the divergence of Negroid and Khoisan groups from a common female ancestor may have occurred in prehistoric times, and that these populations have evolved separately in recent times producing distinct Khoisan-like and Negroid-like mtDNA types.

The relationship between the number of types found within each ethnic group, genetic diversity (h) and percent sequence divergence is shown in Table 4.5. Since there are only a small number of types shared within groups, the values of h obtained within groups are high, *i.e.*, approaching 1. Shared types within the Sekele, Dama and Herero have contributed to the marginally reduced genetic

diversity.

The amount of nucleotide divergence calculated for each ethnic group is shown in Table 4.5 and estimates for the pooled Khoisan and pooled Negroid groups were calculated to be 1.98% and 2.02%, respectively. This is in agreement with the value of 2.08% obtained by Vigilant *et al.* (1991) for their sample of 121 African individuals, and higher than that determined for Asians (1.75%) and European Caucasoids (1.08%).

Table 4.4. MtDNA types represented by more than one individual .

MTDNA TYPE	INDIVIDUALS INCLUDED	POPULATION/S
Zu 33	Zu 35	Zulu
Zu 37*	He 12, Da 07, Da 08, Da09, Da26	Zulu; Herero; Dama
Om 30	Om 34	Sekele
Om 114	Om 126, Om 135, Om 141, Om 148	Sekele
Om 123	Om 127	Sekele
Na 04	He 19	Nama, Herero
He 09	He 27, He 28, Da 18	Herero, Dama
He 06	Da 20	Herero, Dama
Xh 06	N 30 (Vigilant et al. 1991)	Xhosa, Yoroba

*Also found in 20 unrelated Herero frOm Botswana (Vigilant et al. 1991)

4.4 Phylogenetic analyses

4.4.1 Pooling of data

The 144 mtDNA control region sequence data obtained during the present study were pooled with the sequence data obtained by Vigilant *et al.* (1991) from 189 individuals (see Table 4.1). From pairwise comparisons of sequences, it was found that 22 out of a total of 333 mtDNA sequences were found in more than one individual, which resulted in the identification of 250 unique mtDNA sequences.

The following types were common in the pooled sample: one type was found in one "Bantu" (Xh6) and one Yoruba (N30) individual; another type which was found in 20 unrelated Herero individuals (Vigilant 1990) was also found in 1 Herero (He12), 1 "Bantu" (Zu37) and 4 Dama individuals (Da7, Da8, Da9, Da26) from the present study. Other shared types in this study were mentioned earlier and there were another 13 shared types described by Vigilant (1990).

Although there are shared types within the Sekele (present study) and the !Kung (Vigilant 1990), there were no types shared between these two groups. Also, the Nama population sampled did not have any types in common with either the Sekele or the !Kung. This is another piece of evidence to suggest that, other than linguistic and cultural differences, Nama and San do differ in their genetic

structure. This would reinforce the suggestion made from RFLP analysis that both the San and the Nama diverged from a recent Khoisan ancestor and have subsequently evolved separately.

The unique sequences were analysed in conjunction with the reference sequence and the chimpanzee sequence by both parsimony and neighbour-joining analysis:

4.4.2 Parsimony analysis

The principle of maximum parsimony or minimum evolution involves the identification of a tree that requires the smallest number of changes to explain the differences observed among the various mtDNA sequences under study referred to as operational taxonomic units (OTUs). Often more than one tree with the same minimun number of changes are found, so that no unique tree can be inferred (Swofford and Olsen 1990). This has been a major disadvantage in the analysis of mtDNA data since several hundred trees of equal parsimony are generated when analyzing mtDNA data using the PAUP program. Another problem is that all trees obtained from a single run are related and different trees sometimes of shorter lengths are obtained when the order of input of the sequence data is varied (Hedges *et al.* 1992; Maddison *et al.* 1992 and Templeton 1992).

The parsimony tree presented in Fig. 4.3 was generated from

a single run without varying the order of input of the 250 unique mtDNA sequences which was obtained by pooling the sequences obtained in this study with that published by Vigilant *et al.* (1991). Branch swopping procedures were performed to improve the length of the maximum parsimony trees saved. This resulted in the generation of 100 trees all of minimal length. Only 50 of the 100 trees saved were examined to assess the genealogical relationship of the mtDNA types found in the ethnic groups studied. One tree which accounted for 897 mutational events was selected at random and presented in Fig. 4.3.

It must be emphasized that for a data set this large (250 mtDNA types with 119 informative sites) the number of maximum parsimony trees is unknown, "but is almost certainly bigger than can be exhaustively determined in a reasonable amount of computer time" (Stoneking *et al.* 1992b). When using the Macintosh version of the PAUP program (version 3.0; Swofford 1990), it took 5 days to compute 100 trees of minimal length on a Macintosh SE/30 computer (kindly made available to me by Prof. Mark Stoneking, Dept. of Anthropology, The Pennsylvania State University). To compute 10 000 trees on the available computer resources would have taken approximately 500 days to analyze! Hedges *et al.* (1992) accomplished this feat on a Silicon Graphics 340/VGX workstation which took about 5 days of computational time, so to analyse the number of trees they cite is computationally possible whereas for the

purpose of this preliminary study, such extensive analyses were not performed.

Accepting that there are limitations associated with the parsimony analysis performed in this study, the genealogical relationships of the mtDNA sequences in the tree presented in Fig. 4.3 can be used to study the evolution of mtDNA in human populations. This was achieved by arbitrarily dividing the tree into eight different clusters based on the branching order of the tree. Although confidence limits on the branching pattern of the tree would be desirable, statistical tests like the bootstrapping method described by Felsenstein (1975) was not applied due to computational limitations. Thus, the inferences drawn from this analysis are not statistically supported.

The number of mtDNA types from Negroid, Khoisan, Asian, Caucasoid and Papua New Guinean populations were counted in each cluster and are given in Table 4.6.

Fig. 4.3. Phylogenetic tree showing the relationship of the 250 unique mtDNA sequence types as inferred by parsimony analysis. The data was obtained by pooling the sequence data obtained in the present study (144 individuals) with that of Vigilant *et al.* (1991) (189 individuals). The chimpanzee sequence was used to root the tree. The mtDNA types found in the different populations studied are represented symbolically using the key given in the figure. The tree is divided for convenience into eight clusters based on the branching order of the mtDNA types, and the number of mtDNA types found in the Negroid, Khoisan, Asian, Caucasoid and Papua New Guinean populations are given in Table 4.6.

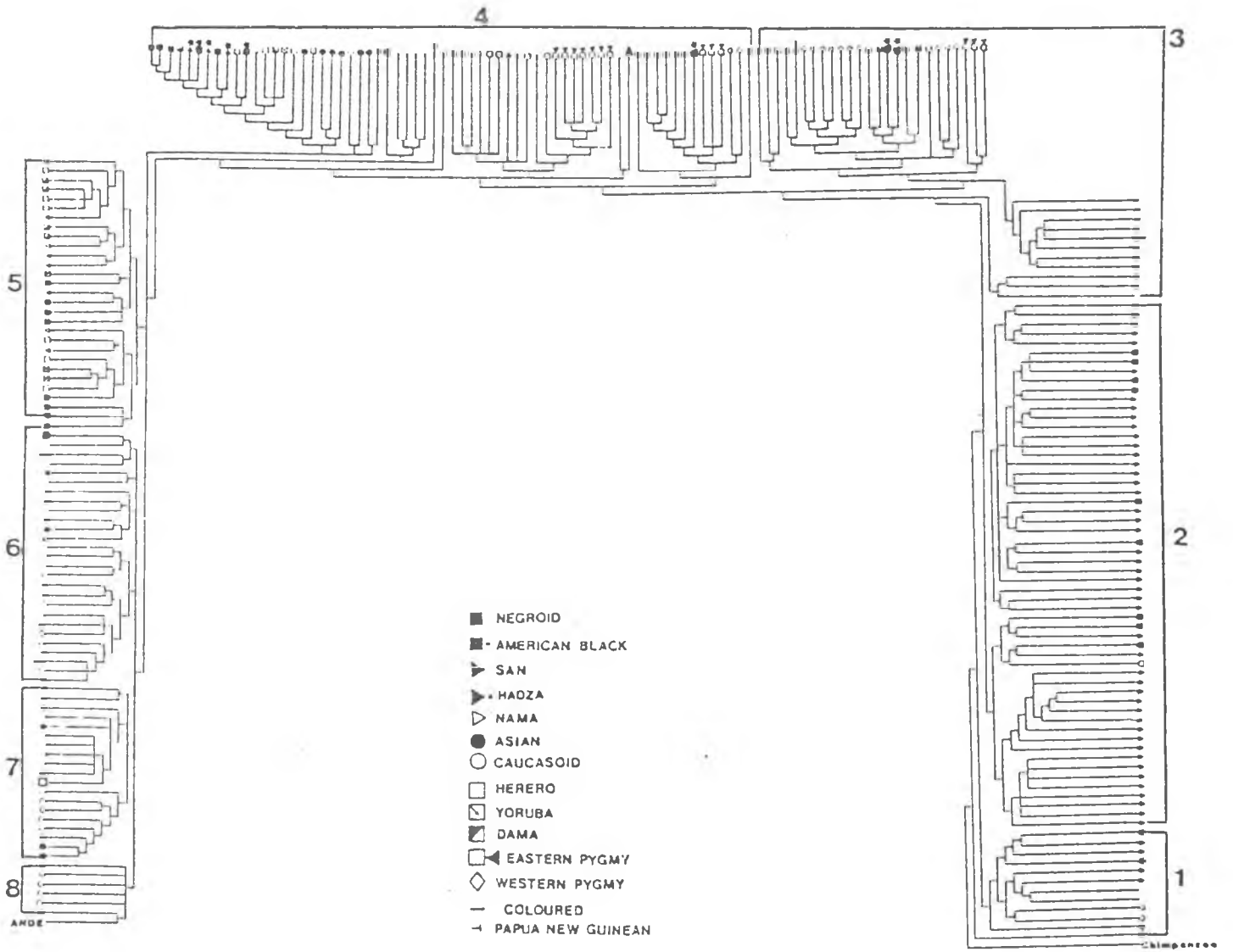
Table 4.6. The numbers and the distribution of mtDNA sequence types found in Negroid, Khoisan, Caucasoid, Asian and Papua New Guinean populations as inferred from parsimony analysis.

	Negroid (N=125)	Khoisan (N=63)	Caucasoid (N=21)	Asian (N=23)	PNG (N=14)	Other (N=4)	Total 250**
Cluster							
1	4	4	2	-	-	2*	12
2	12	44	-	-	-	-	56
3	32	1	-	2	-	1*	36
4	46	5	1	7	5	1***	65
5	21	8	-	-	-	-	29
6	5	1	9	8	4	-	27
7	4	-	5	5	5	-	19
8	1	-	4	1	-	-	6

*Refers to the "Coloured'

**Excludes the chimpanzee and reference sequence

"***Refers to the Australian Aborigine


4
3
5
6
2
7
1
8
ANDE
Chimpanzee
NEGROID
AMERICAN BLACK
SAN
HADZA
NAMA
ASIAN
CAUCASOID
HERERO
YORUBA
DAMA
EASTERN PYGMY
WESTERN PYGMY
COLOURED
PAPUA NEW GUINEAN

Cluster 1 consists predominantly of rare mtDNA types previously discovered in southern Africans from RFLP analysis. These types are found more commonly in the individuals of African descent (10 out of 12) except for two Caucasoids in this cluster. One of these Caucasoids (Afrik 40) was found to have mtDNA type 142-2 (3-2-1-11-2-2) from RFLP analysis. Since this type has the *Hpa*I-3 morph, it seems likely that this type was derived from an African mtDNA type (see Fig. 4.3) and introduced into the Afrikaans-speaking ethnic group by gene flow from either the Negroid or Khoisan groups. In fact, Botha and Pritchard 1972 have estimated that approximately 6-7% of blood group genes found in the South African Afrikaans-speaking community are derived from indigenous African populations. Mitochondrial DNA studies in this instance would confirm that indigenous African females have contributed to the gene pool of European Caucasoids in South Africa.

Cluster 2 consists of 44/56 (78.6%) Khoisan mtDNA types and only 12/56 (21.4%) Negroid mtDNA types (see Table 4.6). MtDNA types found in the Hadza from Tanzania do not group in this cluster. Previous data making use of the a single blood group system (Rhesus) suggested that the Hadza (who speak a click language) and the San were genetically closely related (Mourant 1983). MtDNA studies on the other hand do not support this suggestion (present study; Vigilant 1990). In fact, mtDNA types found in the Hadza

cluster more closely with Negroid mtDNA types (see Fig. 4.3) which would suggest that the Hadza are Negroid people who have borrowed words spoken by Khoisan peoples in ancient times and still hunt and gather for their subsistence. A more intensive genetic study from a larger sample size and including several other genetic markers, would help shed more light on the origins of the Hadza. In the succeeding discussion, the Hadza have been grouped with the Negroid populations.

Clusters 3 (32/36) and 4 (46/65) together contain 78 out of a total of 125 mtDNA types found in Negroid populations (Table 4.6). Several smaller ethnic specific clusters of mtDNA types are found scattered throughout the tree (Fig. 4.3). This is particularly true for clusters 2,3 and 4 (see Fig. 4.3). Though not the rule, there appears to be a high correlation between the patterns of branching and clustering of mtDNA types in the parsimony tree given in Fig. 4.3 and the geographic origins of the populations.

Clusters 1 and 2 appear to be "typical" Khoisan clusters since 48 out of a total of 63 mtDNA types found in the Khoisan (76.2%) are found in these clusters. On the other hand clusters 3,4, and 5 which together constitute 79.2% (99/125) of the total number of mtDNA types found in Negroids appear to be "typically" Negroid. Asian mtDNA types first appear in cluster 3 (2 out of a total of 23 Asian mtDNA types) and then in clusters 4,6,7 and 8 (see

Table 4.6). Caucasoid mtDNA types are found at a low frequency in cluster 4 (1.5%) but are the major types found in clusters 6 (33.3%), 7 (26.3%) and 8 (66.7%). MtDNA types in individuals from Papua New Guinea cluster closest to the Asian mtDNA types. This lends further support to the hypothesis that mtDNA types in Papua New Guineans have been derived from Asians (Stoneking *et al.* 1986a,b) and to the other evidence for their south-east Asian origins as discussed by Bellwood (1989).

The distribution of mtDNA lineages in the parsimony tree given in Fig. 4.3 would suggest placing the geographic origin of the mtDNA ancestor in Africa. It also suggests that Asians diverged before European Caucasoids, findings which are consistent with the data discussed by Horai *et al.* (1991). These observations support the suggestion put forward by Cann *et al.* (1987) and Vigilant *et al.* (1991) that human mtDNA evolved in Africa. One cannot however, exclude the possibility that modern humans evolved, for example, in Asia. If one considers the possibility that Asian populations underwent a bottleneck or a selective event, it could have the effect of reducing the diversity found in Asians, thereby biasing the placement of mtDNA types found in Asians in the phylogenetic trees. This possibility must always be borne in mind when evaluating the presence and absence of mtDNA types and their meaning for population affinities and population evolutionary histories.

Although the effectiveness of parsimony at estimating phylogenies has been extensively examined (Felsenstein 1988; Sober 1985), its strengths and weaknesses for estimating ancestral states have not been critically investigated (Maddison 1991). Maddison (1991) argues that parsimony inference of ancestral locations requires background assumptions about the nature of migrations; using the restriction site data published by Cann *et al.* (1987) he has shown that if one assumes unordered migration, five classes of trees can be generated from these data - all show that both Africa and Asia are equally parsimonious ancestral states, whereas if one assumes ordered migrations, only one class of trees exists in which the line of migration is from Africa to Europe, thence to Asia, New Guinea and Australia.

Another weakness of parsimony analysis as argued by Wilson *et al.* (1991) is the disregard of sample size when inferring geographic origins. Another problem associated with the branch swapping procedure in parsimony analysis is that, depending on the order of input of the data, trees of different length and ancestral regions are obtained. These weaknesses of parsimony analysis have introduced more debate on the placing of the human mtDNA ancestor (Hedges *et al.* 1992; Maddison *et al.* 1992; Templeton 1992).

In view of the weakness related to parsimony analysis eluded to in this section, the pooled mtDNA sequence data

were analysed using the neighbour-joining algorithm, and the inferences regarding population differentiation and human mtDNA evolution by the two methods are discussed in the next section.

4.4.3 Neighbour-joining analysis

Unlike the standard algorithm for minimum evolution trees, the neighbour-joining (NJ) method developed by Saitou and Nei (1987) keeps track of nodes on the tree rather than taxa or clusters of taxa. The raw data are provided as a matrix of distances between all pairs of taxa (terminal nodes) on the tree to be inferred. A modified distance matrix is constructed in which the separation between each pair of nodes is adjusted on the basis of their average divergence from all nodes (conceptually, the adjustment has the effect of normalizing the divergence of each taxon for its average clock rate). The tree is constructed by linking the least distant pair of nodes as defined by this modified matrix. When two nodes are linked, their common ancestral node is added to the tree and the terminal nodes with their respective branches are removed from the tree. This pruning process converts the newly added common ancestor to a terminal node on the tree of reduced size. At each stage of the process, two terminal nodes are replaced by one new node. The process is complete when two nodes remain, separated by a single branch (Swofford and Olsen 1990).

The branch lengths as well as the topology of a single tree can be obtained by this method. The phylogenetic tree obtained by means of the NJ method is shown in Fig. 4.4. A legitimate criticism of NJ analysis is that there is no objective criterion for determining how much better the NJ tree fits the data than some alternative topology, but bootstrap resampling of the data does provide a means of assessing the statistical significance of clades identified by NJ analysis (Stoneking *et al*. 1992b). Unfortunately, this test of significance was not performed due to computational limitations in the present study, but would be considered in future analyses.

Using the order of branching of mtDNA types in this tree, mtDNA types could be grouped together into three major clusters (Table 4.7). Of the 75 different mtDNA types found in cluster 1, 48 are found in the Khoisan population. This constitutes 76.2% (48/63) of the total number of mtDNA types in the Khoisan, whilst only 20/75 (26.6%) are found in the Negroid population. Only 2 out of a total of 21 unique mtDNA types found in Caucasoids are found in this cluster; both of these were South African Ashkenazi Jewish individuals (RJ5 and RJ17) who were also found to have rare mtDNA types 127-2(2-1-2-15-1-2) and 128-2(2-2-3-1-1-2), respectively. There were no Asian and Papua New Guinean mtDNA types in this cluster.

Fig. 4.4. Phylogenetic tree showing the relationship of 250 unique mtDNA sequence types found in African, Asian, Caucasoid and Papua New Guinean populations as inferred by the neighbour-joining method. The tree is arbitarily divided into three clusters and the number of Negroid, Khoisan, Caucasoid, Asian and Papua New Guinean mtDNA types in each of these clusters are given in Table 4.7.

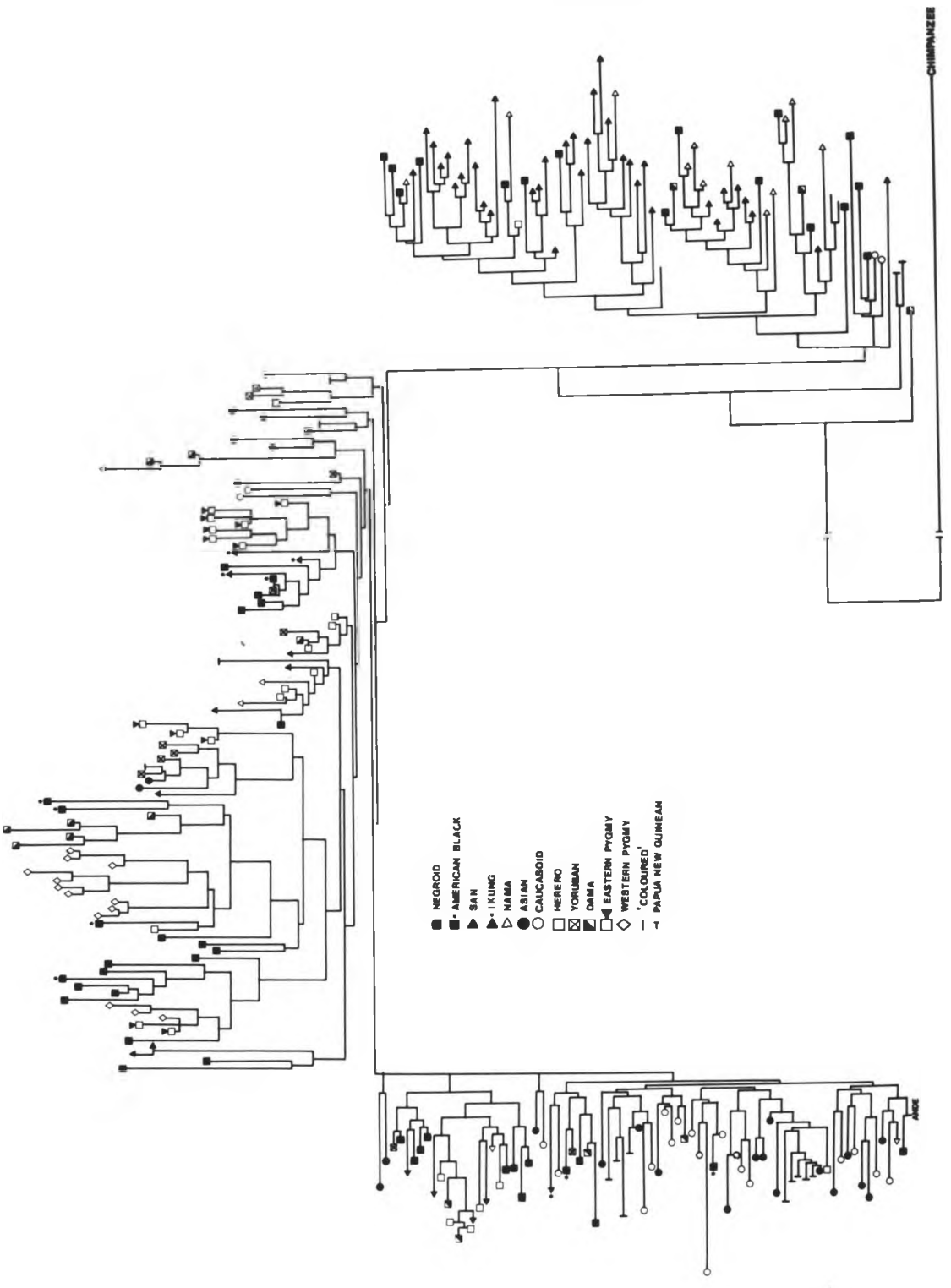
Table 4.7. The numbers and the distribution of mtDNA types found in the Negroid, Khoisan, Caucasoid, Asian and Papua New Guinean populations as inferred from neighbour-joining analysis.

	Negroid (N = 125)	Khoisan (N = 63)	Caucasoid (N = 21)	Asian (N = 23)	PNG (N = 14)	Other (N = 4)	Total 250**
Cluster							
1	20	48	2	-	2	3*	75
2	77	9	2	7	3	1***	99
3	28	6	17	16	9	-	76

*Refers to the "Coloured"

**Excludes the chimpanzee and reference sequence

***Refers to the Australian Aborigine


CHIMPANZEE
NEGROID
AMERICAN BLACK
SAN
!KUNG
NAMA
ASIAN
CAUCASOID
HERERO
YORUBAN
DAMA
EASTERN PYGMY
WESTERN PYGMY
'COLOURED'
PAPUA NEW GUINEAN
AKE

On the other hand, cluster 2 consists of 77/99 (77.8%) of the mtDNA types found in Negroid populations (61.6% of the total mtDNA types of Negroids) and only 9/99 (9.1%) are found in the Khoisan. Seven out of a total of 23 Asian mtDNA types are found in this cluster whereas only 2 out of a total of 21 Caucasoid mtDNA types are found in this cluster. A small number of mtDNA types (3/14) from Papua New Guineans are also found in this cluster.

A large proportion of the mtDNA types found in Caucasoids (19/21), Asians (16/23) and Papua New Guineans (9/14) are found in cluster 3. The close relationship between mtDNA types found in Papua New Guineans and Asians may be interpreted as due to the ancient Asian derivation of Papua New Guineans. A smaller proportion of the mtDNA types found in Negroid (28/125) and Khoisan (6/63) populations are found in cluster 3.

Although the inferences made on the basis of both the parsimony analysis and the NJ methods are not backed by statistical methods for assessing confidence, it would appear that the mtDNA types found in Khoisan populations are commonly found in the deepest branches of the tree whilst mtDNA types found in Caucasoids are restricted to the shallower parts of the tree. On the other hand the mtDNA types found in Negroid populations are placed intermediate between the Khoisan and the Asian/Caucasoid groups.

Since the NJ method has the advantage of producing the branch lengths between pairs of mtDNA sequences under the assumption that all lineages have evolved equal amounts, some inferences on population history can be made. There are several small clusters of mtDNA types linking mtDNA sequences found in the same ethnic group. This is particularly true for the Eastern and Western Pygmies, the Herero and the Khoisan populations in whom mtDNA types are found at terminal positions in the NJ tree (see Fig. 4.4). The limited distribution of mtDNA types found in these groups may be due to a recent population bottleneck or to the effects of genetic drift when several mtDNA types could have been lost.

In the parsimony tree (Fig.4.3) and the NJ tree (Fig. 4.4) many of the mtDNA lineages giving rise to mtDNA sequences in Papua New Guineans appear to be derived from Asian branches. However, in the distance analysis (Fig4.4) the branches associated with mtDNA types found in Papua New Guineans are somewhat lengthened whereas no differences in branch lengths can be observed in the parsimony tree (Fig. 4.3). Since the Papua New Guinean branches have the effect of lengthening the Asian lineages from which they are derived it would suggest that the mtDNA types found in Papua New Guineans have been derived recently from an expanding Asian population.

Both methods of analysis would suggest, therefore, that human mtDNA types are most diverse in Khoisan populations followed by Negroids, and that Asian mtDNA types are more diverse than Caucasoid mtDNA types. These observations lend further support for the hypothesis that human mtDNA originated in Africa.

4.5 Cluster analysis

The average sequence divergence between populations investigated in this study was used to construct a NJ tree (Fig. 4.5a). As with RFLP data, the Herero and Dama are closely associated with the Caucasoid group. Sequence data also support the suggestion made from RFLP analysis that southern African Negroids who presently speak different Bantu-languages may be distinguished into two groups; those who speak southeastern Bantu languages referred to as the "Bantu" group in this study and those who speak a southwestern Bantu language (Herero).

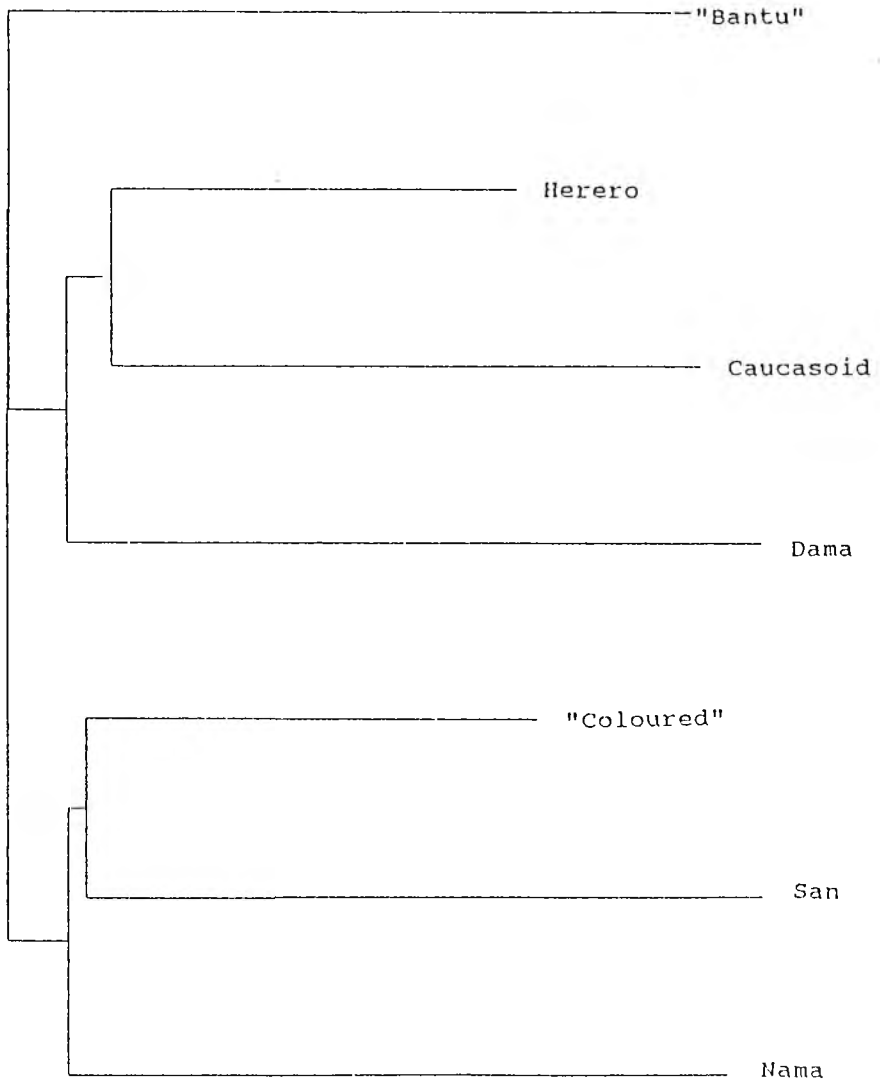
The three sequence types from the "Coloured" individuals sampled cluster with the Nama and the San in the same cluster (Fig. 4.5a). Although a small number of individuals were sampled from the different ethnic groups, mtDNA sequence data confirms the findings of RFLP studies that there are differences in the mtDNA genes of eastern and western Bantu-speaking Negroids and, that Nama and San populations are more closely related to one another than

either is to Negroid populations.

A cluster dendrogram showing the relationship of populations in the pooled sample (Vigilant *et al.* 1991 and the present study) is shown in Fig. 4.5b. In this analysis the reference sequence (Anderson *et al.* 1981) is included in the Caucasoid group and the Australian Aborigine, "Coloured" and Afroamerican individuals are excluded since their sample sizes are too small to consider them as populations (see Table 4.1). In this cluster dendrogram the Western and Eastern Pygmies are placed closer to the root of the tree and are separated from other Negroids. The southern African populations are found in two closely related clusters. The firsts consists of the Dama and the Herero and the second consists of the Nama and the San groups. The southeastern Bantu-speaking Negroids referred to here as "Bantu" cluster with the Khoisan populations rather than with the Herero and the Dama groups. The Yoruba from Nigeria and the Hadza from Tanzania are more closely related to the Asian and Caucasoid groups than to other African populations.

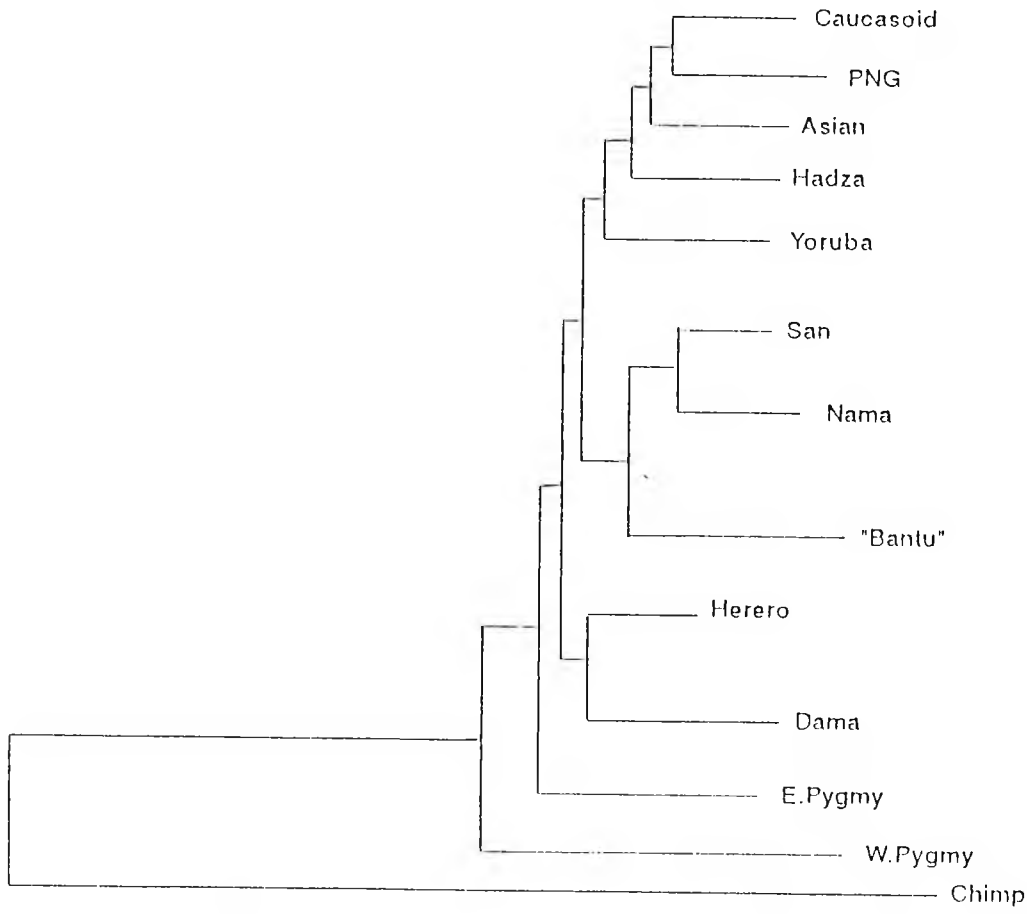
The inferences on population affinities based on cluster affinites should be considered with caution for two reasons. Firstly, mtDNA studies indicate only the maternal contribution of genes in a population. For example, the apparent closer relatedness of the Herero and Dama to Caucasoids than to other African populations noted in this

study has not been supported by Y chromosome RFLP studies (Spurdle 1992). Secondly, cluster dendrograms presented in Figs. 4.5 (a) and (b) are based on the average nucleotide divergence between all pairs of populations and is a population tree that should not be confused with the inferences drawn from the gene trees given in Figs. 4.3 and 4.4 (Nei 1987).



Branch length 10mm = 0.001

Fig. 4.5 (a). Neighbour-joining cluster dendrogram showing the genetic affinities of the seven southern African populations investigated for sequence variation within the control region of the mtDNA moleculeas.



Branch length 10mm = 0.003

Fig. 4.5 (b). Neighbour-joining cluster dendrogram showing the relationship of African, Asian, Caucasoid and Papua New Guinean populations based on the sequence variation contained within the two hypervariable regions of the mtDNA molecule. The data used in this analysis was obtained by pooling the sequences obtained in this study with that previously described by Vigilant *et al.* (1991). The chimpanzee sequence was used as an outgroup to root the tree.

4.6 Rate of mtDNA evolution and age of the common mtDNA ancestor

Much of the controversy concerning theories of human evolution can, in part, be attributed to varying estimates of the rate of mtDNA evolution. Using the estimate of 2 to 4% per million years, Cann *et al.* (1987) concluded that the present diversity found in human populations can be traced back to a common female ancestor who lived in Africa some 140 000 to 290 000 years ago. Vigilant *et al.* (1991) estimated the rate of sequence variation in the control region of the mtDNA molecule and, by using the average amount of nucleotide divergence between human and the chimpanzee control region sequences, calculated that nucleotides in the control region mutate at a rate of 15.1% per million years.

This estimate was based on the rate of transition to transversion mutations. Transitions are known to occur more frequently than transversions and are frequently underestimated due to back mutations. Vigilant *et al.* (1991) extrapolated the number of transition mutations that would have occurred in this region using the transition to transversion ratio that they obtained from parsimony analysis (15:1). By using the average number of transversions that occur between human and chimpanzee sequences (26.4), the adjusted value for the equivalent number of transitions in humans was calculated to be 396,

and the adjusted amount of sequence divergence in this region extrapolated to 69.2%. Using the 4 to 7 million year estimate for the divergence of humans and chimps (Hasegawa *et al.* 1989), the mutation rate in the control region was calculated to be approximately 11.5% to 17.3% per million years and the time of divergence from the common ancestor 166 000 to 249 000 years ago (Vigilant *et al.* 1991).

More recently, several authors have expressed concern about the inadequacies of parsimony analysis (Hedges *et al.* 1992; Maddison *et al.* 1992; Templeton 1992). To overcome these limitations, Stoneking *et al.* (1992) have devised a new method for dating the age of the common ancestor which is independent of human-chimpanzee comparisons and parsimony tree analysis. This method relies on the identification of monophyletic clusters or groups of mtDNA types specific to a defined geographic "region" of the world. This "region" of the world should have been colonized only once and at a definite time that is firmly established from archaeological or biogeographical evidence, with little or no back-migrations (Stoneking *et al.* 1992). While no human populations exactly meet these criteria, the colonization of Papua New Guinea comes reasonably close (Stoneking *et al.* 1986a,b).

It has been suggested that Papua New Guinea was first colonized approximately 40 000 years ago (Groube *et al.*

1986) and that the founders were of Asian origin (Stoneking 1986; Bellwood 1989). MtDNA control region sequence data obtained by sequencing the two hypervariable regions from 165 individuals of whom 50 where from Papua New Guinea, were used to calibrate the rate of human mtDNA control region sequence evolution and the age of the common ancestor (Stoneking *et al*. 1992). The average genetic distance in Papua New Guineans was found to be 0.71%. Assuming that the maximum time for the colonization of Papua New Guinea was 60 000 years ago, the rate of mutation in the control region was estimated to be 11.81% per million years with a corresponding estimate of the age of the common mtDNA ancestor equivalent to 133 000 years (Stoneking *et al*. 1992). Using another method which involves ascertaining the 95% confidence interval (based on 95% confidence intervals for the rate of control region sequence evolution and the amount of sequence evolution corresponding to the human mtDNA ancestor), these authors have estimated the coalescence time to be 63 000 to 356 000 years ago.

It is virtually impossible to use southern African populations to estimate the rate of mtDNA mutation since southern Africa does not satisfy the requirements for a "defined geographic region" since the exact time of "colonization" is not known and, there has been several migrations within Africa which has resulted in varying degrees of genetic admixtures between the various

populations.

4.7 Frequency distribution of pairwise sequence difference

Another means of analysis which is not dependant on tree construction, involves the use of pairwise sequence comparisons within groups. This is represented as a wave spectrum which shows the frequency of the number of mismatches within a group. The method, which was developed by Rogers and Harpending (1992), reveals episodes of population growth and decline, as well as population bottlenecks, which are inferred from the shape of the wave spectrum. Thus, both population demography as well as inferences on the age of the population can be extrapolated from the population signatures obtained using this method of analysis (for details see Rogers and Harpending 1992).

These waves travel from left to right at a rate of $2u$ per generation. Thus, t generations after an expansion begins, the crest of the wave will be near $2ut$. After the population has expanded by roughly two orders of magnitude, further increases in its size have little effect on the distribution until the first wave has had time to dissipate. Since waves dissipate only very slowly, they remain in view for thousands of generations. The mutation rate, u, is the probability that a mutation occurs at one or more of the nucleotides being sampled and can be

estimated from sequence data by μm_T, where μ is the mutation rate per nucleotide and m_T is the number of nucleotides being sequenced. For restriction site data u is estimated by $2\mu rk$, where rk is the number of nucleotide sites covered by the restriction sites in the sample (Rogers and Harpending 1992).

From the waves generated by the pairwise distribution of mtDNA sequence mismatches (see Fig. 4.6a-f), the three parameters of the Rogers and Harpending model (r, ν_0, and ν_1) can be computed and used to calculate the population expansion time for each group studied (Table 4.8) (present study; Sherry *et al.* 1992). The parameter (r) estimates the time (in units of $1/2u$ generations) since the population expanded; (ν_0) estimates the size (in units of $1/2u$ individuals) of the initial population and (ν_1) estimates the size of the population after expansion.

The wave profiles of the southeastern Bantu-speaking Negroids (Fig. 4.6b), Herero (Fig. 4.6c), Dama (Fig. 4.6d), Nama (Fig. 4.6e) and San (Fig. 4.6f) populations studied here reveal some very interesting growth patterns. The steep leading face of the San wave spectrum (Fig. 4.6f) and its high vertical intercept value (ν_1), suggests that the San group has expanded from a small founding population (Rogers and Harpending 1992). The crest of the San wave is close to 10 which when analyzed using the model parameters in Table 4.8, corresponds to a population expansion time of

39 600 years BP. The wave spectrum of the Nama group reveals that the crest of the wave (19) corresponds to approximately 66 900 years since its expansion.

The three Negroid groups studied here (Dama, Herero and "Bantu") all show different wave spectrums consistant with different evolutionary histories (Figs. 4.6b-d). The L-shaped curve for the Herero emphasizes a recent population bottleneck. The expansion time for this pooled Namibian and Botswana groups corresponds to 2 800 years BP. It has been shown from RFLP studies on the Namibian Herero sample (Chapter 3) that this group of Negroids has less diversity when compared to other southern African Negroids. It was suggested earlier that the Herero suffered large dissemination in numbers due to several wars, first with the Hottentots and later the Germans in the Battle of Waterburg (1904-1908). It has been estimated that there were 70 000 to 85 000 Herero in Namibia before the war and only 15 000 to 25 000 at the end of the war (Bridgeman 1981). The number of Herero who escaped to what is now called Botswana is unknown, but estimates of between 1 800 and 9 000 have been suggested (Pennington 1990).

The close genetic relationship between the Herero and Dama groups, as was suggested by RFLP studies, is confounded by the pairwise differences of mtDNA sequences in the Dama. In fact, the wave spectrum (Fig. 4.6d) indicates at least two, and possibly three, different waves for the Dama,

suggesting that there would have been at least two or three possible population expansions. Using this method of analysis, the expansion time for the Dama has been calculated to be 45 700 years BP.

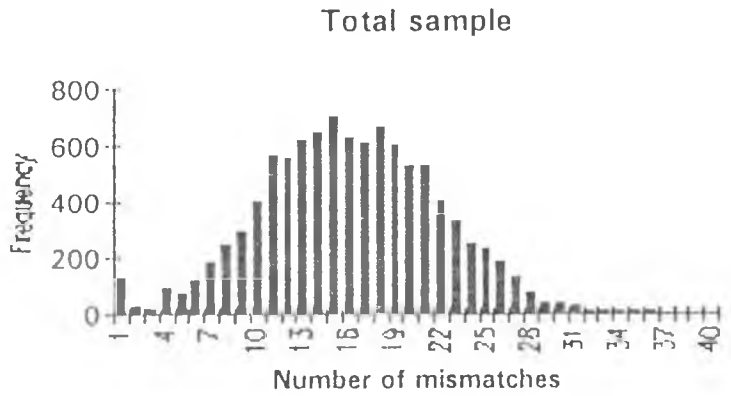
The southeastern Bantu-speaking group referred to here as "Bantu", has a unimodal wave spectrum with a peak at 17 (as shown in Fig. 4.6b) with a corresponding expansion time of 89 200 years BP. The different expansion times of the Khoisan groups (39 600 for San and 66 900 for Nama) and the Negroids is puzzling, since the expansion times inferred from this method suggests that Negroid populations diverged before Khoisan groups from the common ancestor. On the contrary, phylogenetic analyses based on both parsimony (Fig. 4.3) and neighbour-joining (Fig. 4.4) analyses suggests that Khoisan mtDNA sequences are commonest in the deepest branches of the tree which suggests that the Khoisan population diverged before the Negroid population from the common female ancestor.

Nevertheless, the results are consistent with major episodes of modern human expansion occurring approximately 100 000 years ago. On the other hand estimates based on RFLP (Cann *et al.* 1987) and sequencing data (Vigilant *et al.* 1991) suggest that modern humans evolved approximately 200 000 years ago. It is possible, therefore, that a human population may have persisted as a relatively small population for some time, even for 100 000 years, and then

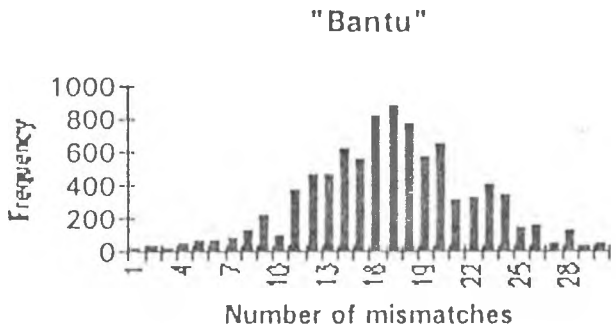
expanded rapidly.

Fig. 4.6. Histograms showing the frequencies of the number of mutational differences obtained by pairwise comparision of mtDNA sequence types in: (a) total sample of 149 individuals, (b) the southeastern Bantu-speaking Negroid group referred to as "Bantu" in this study, (c) Herero, (d) Dama, (e) Nama and (f) San populations.

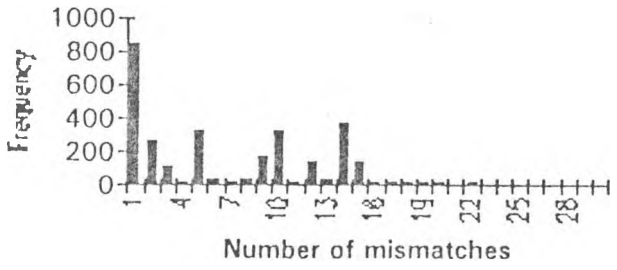
a.


Total sample
800
600
400
200
0
Frequency
1
4
7
10
13
16
19
22
25
28
31
34
37
40
Number of mismatches

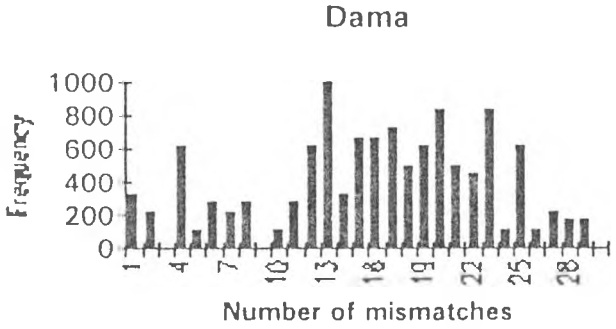
b.


"Bantu"
1000
800
600
400
200
0
Frequency
1
4
7
10
13
16
19
22
25
28
Number of mismatches

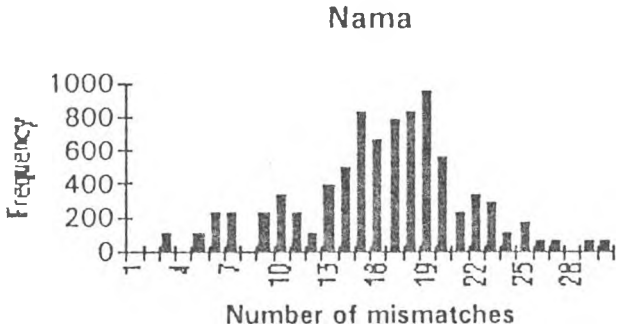
c.


Herero
1000
800
600
400
200
0
Frequency
1
4
7
10
13
16
19
22
25
28
Number of mismatches

d.


Dama
Frequency
1000
800
600
400
200
0
1
4
7
10
13
16
19
22
25
28
Number of mismatches

e.


Nama
Frequency
1000
800
600
400
200
0
1
4
7
10
13
16
19
22
25
28
Number of mismatches

f.

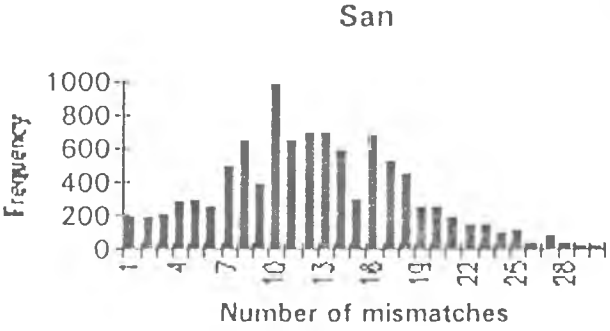

San
Frequency
1000
800
600
400
200
0
1
4
7
10
13
16
19
22
25
28
Number of mismatches

Table 4.8. Expansion times of southern African populations calculated from the parameters of the Rogers and Harpending model. Data from the pooled African, Asian, Caucasoid and Papua New Guinean (PNG) populations are given for comparisons[1].

POPULATION	N	r	θ	EL	r'	T(ka)
"Bantu"	41	9.05	1.75	441	15.46	89.2
Dama	20	6.72	5.24	638	7.93	45.7
Herero*	39	0.39	4.02	589	0.50	2.8
Herero_1	12	6.11	2.71	589	4.10	23.5
Herero_2	27	2.50	4.11	589	0.00	0.0
San*	64	5.15	2.94	464	6.87	39.6
Sekele_1	38	10.83	4.07	564	8.76	50.5
!Kung_2	26	4.69	3.13	564	1.88	10.8
Nama	18	7.52	3.06	488	11.60	66.9
African	249	6.99	2.39	398	13.22	76.0
Asian	80	3.69	1.44	241	11.54	66.6
Caucasoid	20	5.11	1.38	532	7.23	41.7
PNG	32	5.26	3.94	414	9.56	55.2

*Pooled data from the present study (1) and from Vigilant *et al.* (1991) (2).
[1]Data from Sherry *et al.* (1992).
N=sample size, r=the date of expansion measured in units of mutational time, θ=the pre-expansion female effective population size, EL=the number of nucleotides used in the analysis, r'=r/EL, T is the time of expansion and ka=1000 BP.

4.8. Conclusions

1. Approximately 750 bp contained within the two hypervariable segments within the control region of the mtDNA molecule were analyzed from each of 144 individuals from southern Africa. Pairwise comparisons of mtDNA sequences revealed 119 variant sites in the sample. Using only these informative sites, 129 unique mtDNA types were identified in the sample.

2. When pooling the sequence data from this study with that of Vigilant *et al.* (1991), 250 unique mtDNA types were discovered. One type which was common in one Herero, one "Bantu" and four Dama individuals from this study was also found in 20 Herero individuals from Botswana. This suggests that there has been gene flow between Negroid groups in southern Africa. Although there were some shared types within the Sekele and within the !Kung San groups, there were no types shared between the these two groups or between the San and the Nama. This suggests that the two San groups and the Nama have evolved separately with very little gene flow between them.

3. Both the parsimony and neighbour-joining trees which show the relationship of the 250 mtDNA sequence types, reveal that mtDNA types from African populations are found in the deepest branches of the tree. Asian

and Caucasoid mtDNA types are found mainly in branches with little sequence divergence.

(1) Using the branching order of the parsimony tree, approximately eight different clusters were identified. The majority (43/62) of Khoisan mtDNA types are contained within cluster 2 whilst clusters 2, 3, 4 and 5 contained, respectively, 13, 30, 45 and 21 types out of a total of 122 Negroid mtDNA types. Asian mtDNA types first appear in cluster 3, but are then found together with PNG mtDNA types in clusters 4, 6 and 7. Caucasoid mtDNA types are found predominantly in clusters 6, 7 and 8.

(2) The branching order of the neighbour-joining tree can be conveniently divided into three different clusters. The first contains 48 out of a total of 63 Khoisan mtDNA types whilst only 19 of the 122 Negroid mtDNA types are found in this cluster. On the other hand, 92/122 Negroid mtDNA types are found in cluster 2, which is made up of 124 mtDNA types, and it also contains some Khoisan (14/124), Asian (10/124), Caucasoid (3/124) and PNG (5/124) mtDNA types. Cluster 3 has a high proportion of Caucasoid (18 out of a total of 23 Caucasoid mtDNA types), Asian (14 out of a total of 24) and PNG (9 out of 16) mtDNA types.

4. The dendrograms showing the relationship of the ethnic groups analyzed by control region sequence variation, like those obtained from RFLP analysis, reveal that the San, Nama and "Coloured" (only 3 individuals) groups are closely related in one cluster. The "Bantu" cluster away from the Dama and Herero groups which are closely related to the Caucasoid group in another cluster. Although the sample sizes are small for some of the groups investigated here, it is evident that mtDNA sequence variation is a powerful tool for revealing genetic differences between ethnic groups found in southern Africa.

5. Using the pairwise distribution of sequence mismatches according to the Rogers and Harpending (1992) model for population growth and expansion, population expansion times for the populations sampled here range from 2 800 years BP for the Herero to 89 200 years BP for the Negroid group. The San wave profile, according to the predictions of the model, suggests that the San group expanded from a small founding female population approximately 39 600 years BP.

CHAPTER FIVE:

ETHNOGRAPHIC PREHISTORY OF SOUTHERN AFRICAN POPULATIONS: BRINGING TOGETHER THE ARCHAEOLOGICAL, LINGUISTIC AND GENETIC EVIDENCE

The indigenous peoples of southern Africa, in particular, South Africa, are derived from a multitude of ethnic groups each with a unique and colourful history. Unfortunately, as a consequence of the years of oppression introduced in this country by the apartheid policies, there is a growing tendency for individuals to conceal their cultural identities due to pressures imposed on them by their peers and community. As a result of this, many individuals are ignorant of their family and/or ethnic history. However, oral traditions have managed to ensure the continuity of some of the socio-cultural and religious practises of ethnic groups from generation to generation. Some of these are still, albeit to a lesser extent, practised by some ethnic groups, particularly in rural areas where inhabitants have been influenced to a lesser extent by "outside" forces.

In recent times, several studies have made use of the linguistic, archaeological and historical evidence to reconstruct the early history of the indigenous peoples of southern Africa. In most societies the norm is to "adopt" the culture and language of the dominating groups for social acceptance and status in a community, and the true

history of parental groups are often difficult to establish merely by linguistic and cultural studies.

On the other hand, the genetic elements are not prone to the effects of "borrowing" as is the case with linguistic and cultural systems. Thus, as intensive genetic study on the patterns of sequence variations found both within and between extant populations, would help in the elucidation of the complex processes associated with human diversity and the adaptations humans have made to changes in the environment. From such studies, a better understanding of the prehistory of southern African populations would emerge.

This chapter reviews some of the current theories concerning the prehistory and affinities of southern African populations which have been inferred from the lingistic, archaeological and historical data, and attempts to refine them by incorporating the genetic evidence obtained using mtDNA polymorphisms.

5.1 Linguistic evidence

Davidson (1991) very eloquently described the concept of language as "*a system of symbolic communication where an arbitrary relationship between the sign (usually a word) and the signified is defined by convention. Although language does not exist outside the actions of its users,*

it can be analysed through these actions as if they did".

Linguists make use of a variety of elements of languages, such as morphology, phonology, semantics, syntax and pragmatics, to study the similarities and differences of different languages. These tools have been used by a number of linguists (Bleek 1927; Guthrie 1962; Doke 1954; Westphal 1963; Ehret 1982a,b, Vansina 1984) to reveal the origins and differentiation of languages in southern Africa. Details pertaining to the linguistic affiliation of populations sampled in this investigation have been discussed in Chapter 2, but for continuity in this chapter, a few points require expansion here.

5.1.1 Bantu languages

Greenberg's (1963) claim that the 450 known Bantu languages spoken today may be traced to its nuclear area of origin in the Cameroon region of West Africa some 3000 years, is still not proven. In southern Africa, Bantu languages spoken by the majority of Negroid people, have been found to belong to either southeastern or southwestern Bantu divisions of the parental eastern and western Bantu languages, respectively, (Westphal 1963). There still remains some uncertainty in the taxonomy concerning the exact place of smaller groups, especially within the eastern Bantu category where the languages are much closer to each other than other languages (Nurse 1982). However,

distinctions between western Bantu language groups are easier to identify since they have diverged further from each other (Vansina 1984).

Having evolved in the Cameroon region, Bantu languages soon spread to East Africa towards the region of the Great Lakes to form what is referred to by linguists as eastern Bantu languages (also known as the Zambesi languages) (Vansina 1984). Since no trace of this "movement" can be found between Nigeria and the Great Lakes, some authors have used this as an argument to suggest that eastern Bantu must be derived from western Bantu (Ehret 1982a). Vansina (1984), on the other hand, argues that just as the languages of the autochthones in the western Bantu area have vanished, so too have all traces of the eastern Bantu in the west, their speakers having been absorbed by other linguistic groups. The distribution of Bantu linguistic groups throughout Africa is summarized in Fig. 5.1.

The eastern Bantu-speakers had farmed cereal crops ever since they left the Cameroon. In the region of the Great Lakes they had also mastered techniques for iron smelting and it was they who eventually carried cereal agriculture, cattle keeping and metallurgy throughout eastern and southern Africa in the first centuries of the Christian era (Huffman 1989).

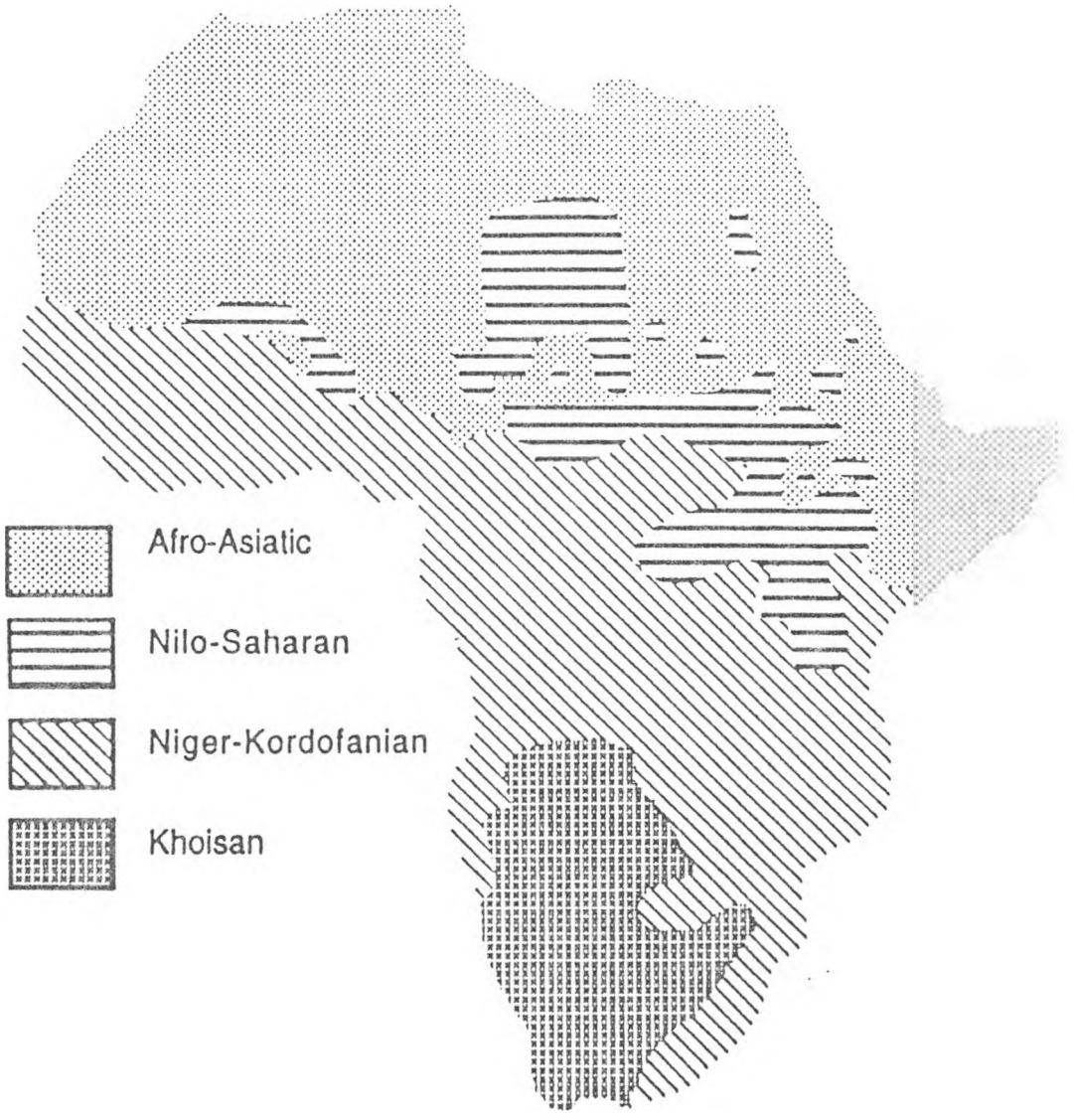


Fig. 5.1. Distribution of Bantu languages in Africa.

On the other hand, a different culture practised by western Bantu-speakers can be traced to sites in the Cameroon grassland which have been associated with farming communities who used stone tools and pottery but no metal (de Maret 1982). According to Vansina (1984), proto-western Bantu culture can be deduced from its vocabulary which suggests that these people farmed yams, oil palms, some gourds and minor domestic plants. They also fished, hunted and gathered wild produce, kept dogs and goats and were potters. It is important to realise that they did not farm cereal nor did they smelt iron, and that their technology was purely neolithic.

Farmers speaking western Bantu languages gradually occupied all of central Africa, and while the majority maintained their neolithic culture some developed an aquatic culture with varying degrees of specialization associated with fishing. It is not surprising that a division in the western Bantu group followed, giving rise to the northern Zaire languages (Vansina 1984). Meanwhile, other western Bantu-speakers proceeded southwards and then turned towards the Kwango River, ultimately giving rise to the southwestern Bantu-speakers found in most parts of Namibia, the Zambesi Valley, the Luangwa Valley, around northern Lake Malawi and Lake Tanganyika. Their expansion was probably arrested in the south-east when they met farmers speaking eastern Bantu languages (Vansina 1984).

Southern Bantu languages presently consists of southeastern, southwestern and central Bantu-speaking languages. Their distribution in southern Africa is given in Fig. 5.2a-c, respectively.

5.1.2 Khoisan languages

Traditionally Khoisan languages (Bush and Hottentot) were used to refer to languages spoken by the "Bushmen" and "Hottentot" people and which could be distinguished from the Bantu languages by the high frequency of click sounds. Westphal (1963) argues that clicks have no positive classifactory value, since clicks are also recognized at varying frequencies in southern Bantu languages such as Xhosa, Zulu, Ndebele and Sotho, introduced into these groups as a consequence of "borrowing" from either Bush or Hottentot languages.

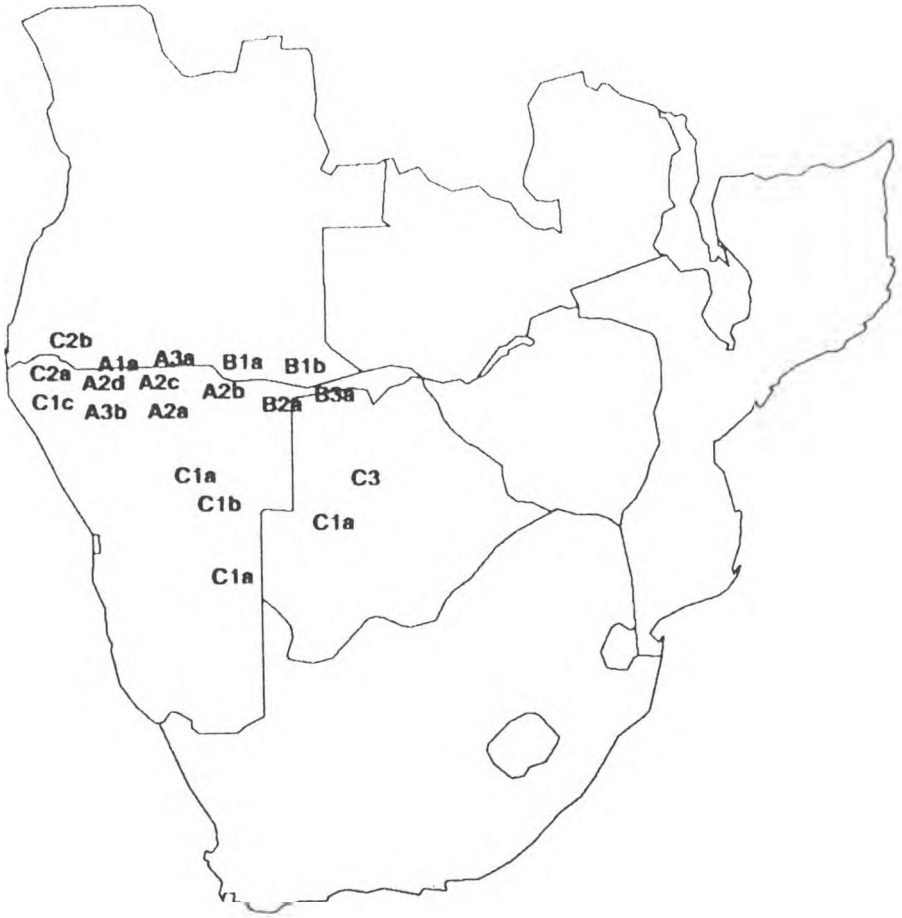
According to Westphal's classification of languages, languages spoken by the San and Nama people belong to the non-Bantu category rather than the click-speaking group. By analysing the languages spoken by the San, Westphal identified four different language types referred to by him as Bush "A", Bush "B", Bush "C" and Bush "D", compared with Bleek's classification of Central, Northern and Southern Bush languages (Bleek 1927). The distribution of the various Bush and Hottentot languages found in southern Africa is given in Fig. 5.3.

Fig. 5.2. Maps of southern Africa showing the location/s (present or past) of the various Bantu-speaking Negroid ethnic groups: (a) southeastern Bantu-speakers, (b) southwestern Bantu-speakers and (c) central Bantu-speakers. (These maps were compiled in collaboration with Professors Trefor Jenkins and George Nurse).



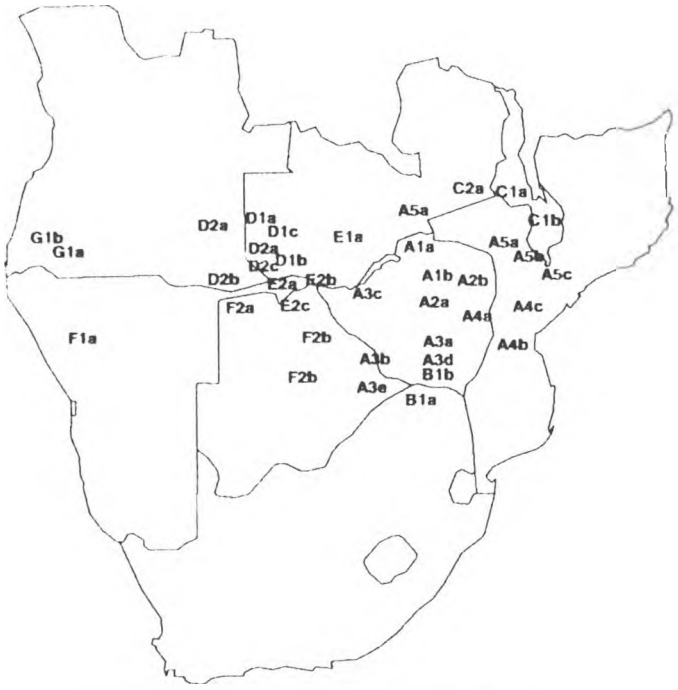
Distribution of Eastern Bantu-speaking Negroids

POPULATION AFFILIATION	ETHNIC GROUP	CODE
[A] SOTHO TSWANA		
South Sotho	South Sotho	A1a
	Taung	A1b
	Lozi	A1c
North Sotho	Pedi	A2a
	Kwena	A2b
	Phalaborwa (Pulana, Pai	A2c
	Lovedu	A2d
	Tlokwa	A2e
Tswana	Ngwato	A3a
	Ngwaketse	A3b
	Kgatla	A3c
	Tlharo	A3d
	Tlhaping	A3e
	Hurutshe	A3f
	Tawana	A3g
	Rolong	A3h
Kgalagadi		A4
[B] NGUNI		
Zulu	Lala	B1a
	Qwabe	B1b
	Ndebele	B1c
	Kora (Tegede)	B1d
Swati	Swati	B2a
	Bhaca	B2b
	Hlubi	B2c
	Ngoni	B2d
Xhosa	Pondo	B3a
	Mpondomise	B3b
	Gcaleka	B3c
	Thembu	B3d
	Bomvana	B3e
	Gqika	B3f
	Ndlambe	B3g
Mfengu	Fingo	B4
[C] TSONGA TSWA		
Tsonga	Tsonga	C1a
	Tswa	C1b
	Ronga	C1c
	Thonga	C1d
Chopi	Chopi	C2a
	Tonga	C2b
Shangana	Hlangana	C3a
	Lundau	C3b



Distribution of Western Bantu-speaking Negroids

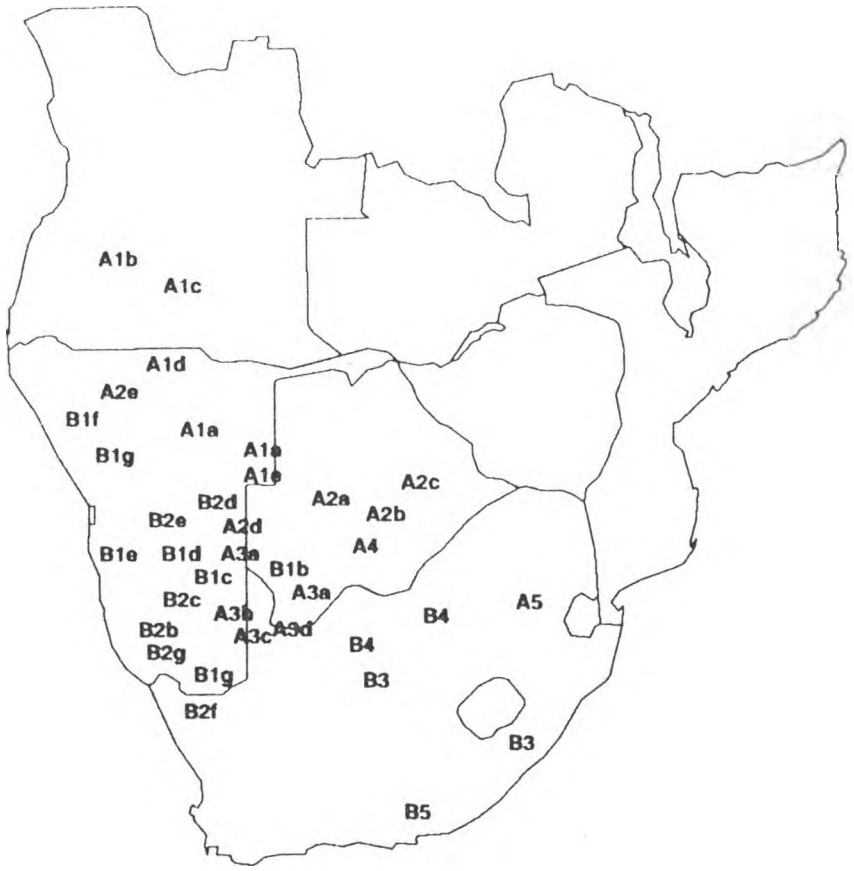
POPULATION		CODE
[A] AMBO		
Mbalantu	Mbalantu	A1a
Ndonga	Ndonga	A2a
	Kwambi	A2b
	Kwaluudhi	A2c
	Nkolonkudhi	A2d
Kwanyama	Kwanyama	A3a
	Ngandjera	A3b
[B] KAVANGO		
Kwangali	Kwangali	B1a
	Mbunja	B1b
Sambyu	Sambyu	B2a
Gciriku	Gciriku	B3a
[C] HERERO		
Herero	Herero	C1a
	Mbanderu	C1b
	Himba	C1c
Cimba	Cimba	C2a
	Kwisi	C2b
Yei		C3



Distribution of Central Bantu-speaking Negroids

POPULATION AFFILIATION	ETHNIC GROUP	CODE
(A) SHONA		
Tavara	Tavara	A1a
	Korekore	A1b
Zezuru	Zezuru	A2a
	Yowa	A2b
Karanga	Karanga	A3a
	Kalanga (Lilima)	A3b
	Rozwi	A3c
	Jena	A3d
	Birwa	A3e
Manyika	Manyika	A4a
	Ndau	A4b
	Rwe	A4c
Nyungwe	Nyungwe (Chikunda)	A5a
	Sena	A5b
	Podza	A5c
(B) VENDA		
Venda	Venda	B1a
	Lemba	B1b
(C) MARAVI		
Chewa	Cewa (Chipeta)	C1a
	Mang'anja	C1b
Nsenga	Nsenga	C2a
(D) LUYANA		
Luyana	Luyana	D1a
	Kololo	D1b
	Luyi	D1c
Mashi	Mashi (Kamashi,Mashiko)	D2a
	Mbukushu (Goba)	D2b
	Shanju	D2c
(E) TONGA		
Ila	Ila	E1a
Totela	Subiya	E2a
	Fwe (We)	E2b
	Ikuhane	E2c
(F) KHOISAN SPEAKING NEGROIDS		
Dama	Dama	F1a
Sarwa	Hukwe (Kwengo)	F2a
	Sarwa	F2b
(G) KWADI SPEAKING NEGROIDS		
Kwadi	Kwadi	G1a
	Kwepe	G1b

Fig. 5.3. The locations (present or past) of Khoisan populations in southern Africa. (This map was compiled in collaboration with Professors Trefor Jenkins and George Nurse).



Distribution of Khoisan speaking peoples in southern Africa

POPULATION AFFILIATION	ETHNIC GROUP	CODE
|A| SAN		
Northern	Zhu/oasi	A1a
	!Kung (Kwankhwala)	A1b
	!Kung (Sekele)	A1c
	!Kung (Glangg!ai)	A1d
	//au//e	A1e
Central	G/wi	A2a
	G//ana	A2b
	G//oro	A2c
	Nharo	A2d
	Hei//om	A2e
Southern	!Xo	A3a
	G!uoKx'ate	A3b
	N/huki	A3c
	/Xam	A3d
Eastern	– hua	A4
//Xegwi		A5
|B| Nama		
Naman	Gami = nang (Bondelswa	B1a
	!Karagekhoin (Fransman	B1b
	//Hababen (Veldschoendr	B1c
	Gei//Khauan (Root Natie)	B1d
	Mu//eing (= Ouinin)	B1e
	!Gumin (Strandlopers)	B1f
	//Khau/goan (Swartboois	B1g
Oorlams	!Aman (Groenmense)	B2a
	/Hei//Khauan	B2b
	/Hobesen (Witboois)	B2c
	Gei/Khauan (Amraals)	B2d
	//Aixa//ain (Afrikaners)	B2e
	Richtersveld	B2f
Griqua (!Karixun)		B3
Korona (!Ora)		B4
Xona		B5

Although click languages are confined to southern Africa and Tanzania (Hadza and Sandawe), skeletal remains reminiscent of San physical appearances found in some parts of north Africa led Tobias (1964) to believe that Khoisan people could have once lived outside of southern Africa. Unfortunately, the archaeological data from these regions are fragmentary (Morris 1984) and there is no support for the suggestions made by historians that the San originated in north Africa (Stow 1905; Theal 1910; Schapera 1930).

5.2 Archaeological evidence

Using the archaeological record of human settlement patterns in southern Africa, some of the earliest settlement patterns have been dated to about 20 000 years BP (Clark 1959). This was based on an analysis of surface assemblages of stone tools. The type of tool technology of the past 20 000 years which has been referred to as the Later Stone Age (LSA), have been found mainly on grassy banks, in front of caves or rock shelters or in the deflation hollow in recent wind blown sands (Parkington 1980). By studying the tool type in conjunction with radiocarbon dating of archaeological material from several sites throughout southern Africa, a better understanding of the chronology of activity differences of prehistoric peoples may be obtained (Parkington 1980).

There are several indications of prehistoric cultures which

may be gleaned from the archaeological evidence: faunal bone material, shells and stone tools found at the various sites examined hint at the dietary habits of the peoples occupying different settlements. It is argued that there have been different episodes of relative stable tool making industries during the past 10 000 years, which may be correlated with the culture of the people occupying the sites (Parkington 1980). For example, the painted shelters in the inland area of the eastern Cape (historically associated with "Bushmen"), contained Late Smithfield assemblages (characterized by the presence of end- and side-scrapers, bored stones and bone point) while artifacts from those in the coastal belt were predominantly Wilton (characterized by a subsistence pattern of hunters and gatherers based on plant foods, hunting of nocturnal solitary smaller antelope or as described by Deacon (1976), as an "adoption of a microlithic technology").

When comparing the distribution of Khoisan language groups with that of Later Stone Age industries, there is strong support for the association of Bush-speaking peoples with the Late Smithfield culture, on the one hand, and between Wilton culture and Hottentot-speaking people on the other hand. The only serious exception occurs in Bushmanland, where, according to both historical and archaeological distribution, Bush-speaking /Xam Bushmen produced Wilton tools. Late Smithfield culture occurs in most parts of the Cape and the Free State inland plateau region whilst the

Wilton zone covers the whole of Namibia, the western and north-western Cape, the western part of Bushmanland, the Cape Folded Mountain Belt, and the southern coastal regions as far east as East London (Fig. 5.4). The two zones may overlap at some places, as was observed in the eastern Cape mountain and coastal belt (Schoute-Vanneck and Walsh 1960; Rudner and Grattan-Bellen 1964).

Within the Khoisan category, there were those who were hunter-gatherers (San) and those who were pastoralists (Khoi). From the linguistic evidence, considerable emphasis has been placed on the vocabulary associated with livestock in the reconstruction of the prehistory of pastoralism in southern Africa. From an archaeological perspective, this is supported by the presence of bovine skeletal material at the various sites. The nature and extent of interactions between hunter-gatherers and patoralists is very much dependent on quantitative faunal analyses (Parkington 1980).

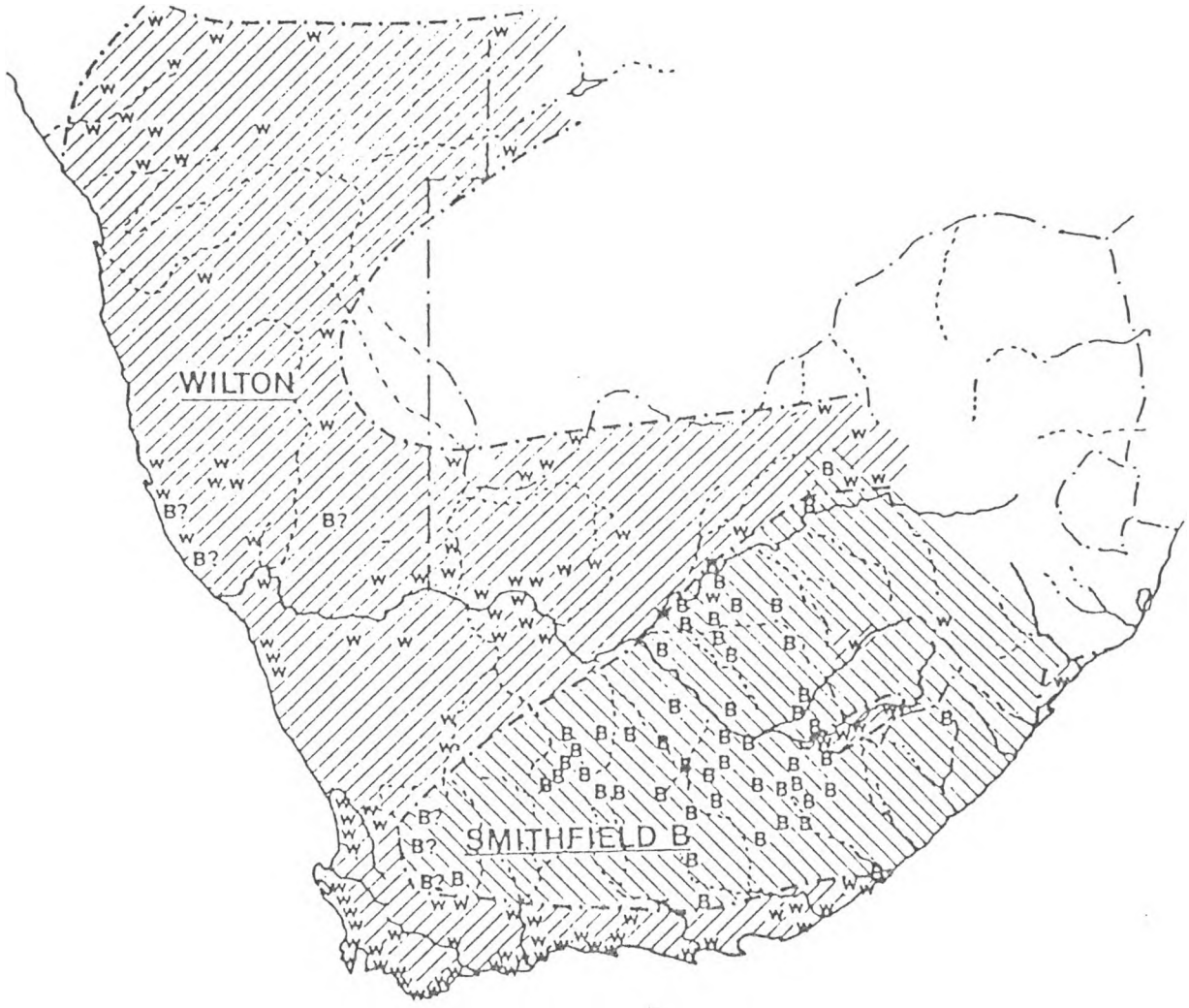


Fig. 5.4. Distribution of Wilton (W) and Smithfield B (B) archaeological sites which have been associated with regions occupied by prehistoric Bush-speaking and Hottentot-speaking peoples, respectively. Map obtained from Rudner (1979).

Bushman paintings have also provided an excellent marker for studying the regions occupied by "Bushmen" or San people. In the Cape, a major division in style can be distinguished between the coastal plain and the interior mountains. Although the physical distances involved are not great, there are marked differences between the two areas in climate, vegetation and archaeology. For example, in the mountains rock art sites are abundant and often contain large and varied sets of paintings. At the coast, however, suitable locations for paintings are far more limited and the art is often impoverished. Handprints constitute the major component of the art throughout the sandveld, although representational images do show an increase in frequency away from the coast.

Bushman artists were knowledgeable agents who depicted their social systems in their rock paintings. Although different interpretations of these symbolic representations are possible, some insight into the culture of prehistoric San populations may be obtained from these pictorial representations.

In addition to stone age culture, iron technology brought into southern Africa by the Bantu-speaking Negroids approximately 1500 years BP, may be used together with ceramic data to reconstruct the history of Negroid peoples in southern Africa. Most Iron Age archaeologists divide the last two thousand years in sub-equatorial Africa into

Early Iron Age (EIA) and Late Iron Age (LIA) periods (Huffman 1989). Although these are chronological terms, they also refer to archaeological units with certain kinds of economic systems and pottery styles.

Huffman (1980, 1982, 1989) has studied the settlement patterns and ceramic styles of pottery excavated from several locations in southern Africa in an attempt to determine cultural identity. In his opinion, settlement organization should reflect the society's attitudes with respect to politics, economy, rank, status and religion whilst the ceramic evidence may be used to study the stylistic continuity of an assemblage through successive phases. He argues that if local development took place, then the continuity of assemblage pattern should parallel the sequence of change that occurred in a related branch or branches - evidence for this could be reconstructed from the archaeological evidence.

Kuper (1980) has shown that the various Sotho-Tswana and Nguni speaking peoples share one common cultural system referred to as the southern Bantu Cultural system. This system incorporates interconnected attitudes about the political role of men, the spiritual role of ancestors and the importance of cattle. Cattle are the main avenue to wives and children and therefore to power, status and success. Most importantly, this southern Bantu Cultural system is expressed in the settlement pattern (Central

Cattle Pattern) with the following characteristics: a central cattle byre which contains grain storage pits and elite burials; a men's court next to or in the byre; an outer arc of houses with the principal house upslope of the byre and the remainder arranged to the right and left according to seniority; and within a house a right-male/left-female distinction which is usually orientated at right angles to a front-secular/back-sacred dichotomy (see Fig. 5.5a).

Metal currency was as important to western Bantu speakers as cattle were to eastern Bantu-speakers, and the status of metal workers was high (Vansina 1969). Smiths were often associated with clan origins, and chiefs were often referred to as master-smiths (Vansina 1955; de Maret 1985). Consequently, certain metal-working tools such as anvils and hammers, and objects such as ceremonial axes and iron gongs were symbols of leadership. Since eastern Bantu-speakers did not normally correlate metal-working with leadership, these symbolic items were a characteristic feature of western Bantu-speakers (Huffman 1989).

These special metal items together with extended burial in mass cemeteries, rectangular houses divided internally in a special way often incorporated separate kitchen and sleeping rooms (Baumann 1935). These rectangular houses were usually arranged in a rough rectangle or in parallel

rows on opposite sides of a "street" (Laman 1953) to make up the Forest pattern (Fig. 5.5b). The men's court was at the end of the street or open space neat the headman's house, as in the Central Cattle Pattern, but storage pits in the centre of the village did not feature - rather, V-shaped pits were sometimes dug behind the village for processing palm oil (Douglas 1963), and hollows for firing pots could be found near a potter's house (Eggert and Kanimba 1980). Shrines were placed at the doorway of a house and in front of a headman's house, rather than in back/private areas (Turner 1957). In contrast to the Central Cattle Pattern, the system of beliefs and values that underlies this spatial organization is not known.

A mixture of Central Cattle and Forest patterns are often found in southern Zambia and have been referred to as the Ila House Plan (after an Ila chief's village). The arrangement of the huts in a circle around a central courtyard resembles Zulu patterns seen in the nineteenth century (Parkington and Cronin 1979), whilst the interior of the Ila hut exemplifies the western Bantu spatial organization in which the spiritual area was at the threshold rather than at the back; a moulded fireplace was against the wall opposite the bed rather than in the centre; a line of storage jars down the middle of the hut created separate sleeping and cooking compartments (see Fig. 5.5c).

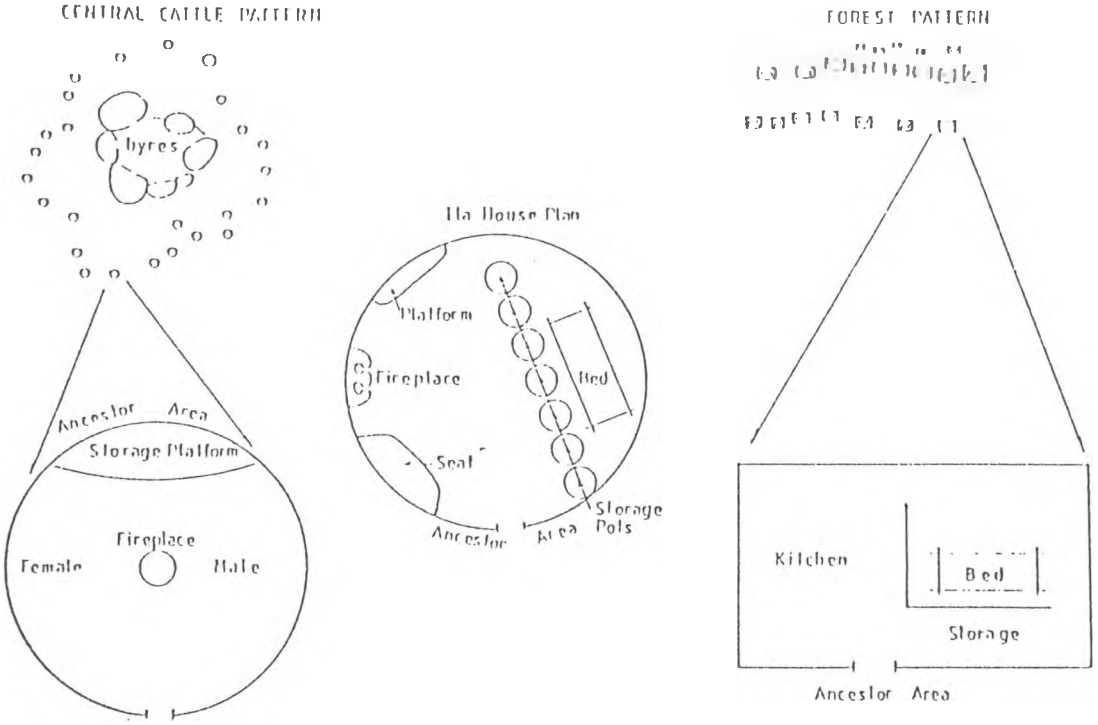


Fig. 5.5. Arrangement of the (a) Central Cattle Pattern, (b) Forest Pattern and (c) Ila House Plan, of Early Iron Age settlement sites which have been associated with southeastern, southwestern and a mixture of southeastern and southwestern Bantu-speaking Negroid groups, respectively, in southern Africa. (Obtained from Huffman 1989).

Using a combination of both the patterns of settlement and the ceramic styles (shape, design and motif combination), Huffman (1989) concludes that there is a relationship between Bantu-speaking groups and EIA, a feature Phillipson (1977) believed to be largely circumstantial. According to Huffman's theory, ceramic style is created and transmitted by groups of people through verbal communication and its distribution should reflect the distribution of a group of people speaking the same language. From his studies, Huffman concludes that eastern Bantu-speaking people have been associated with three different cultures which he refers to as the Nkope and Kwale branches of the Urewe Tradition (formerly referred to as eastern stream by Phillipson 1977) and the Kalundu Tradition (previously assigned by Phillipson to the western stream) (Fig. 5.6). Western Bantu-speaking people on the other hand are associated with the Luangwa styles which are thought to be derived from the Naviundu styles.

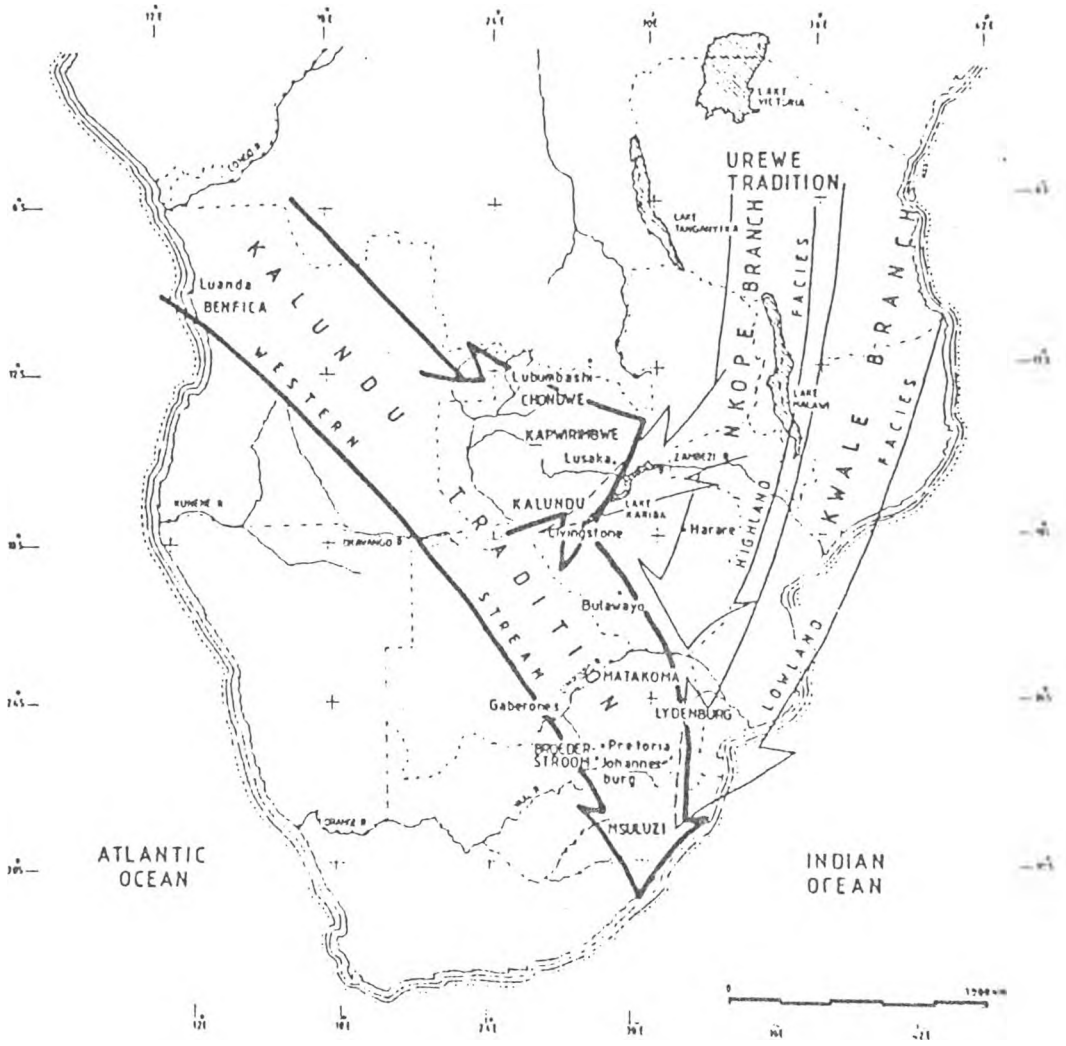


Fig. 5.6. Map of southern Africa showing the possible routes of migrations of eastern Bantu-speaking Negroids into southern Africa based on the correlation of ceramic styles and language. According to Huffman, the differences in the ceramic styles of pottery escavated from various archaeological sites in South Africa and Zimbabwe suggest that these patterns have been derived from two different traditions; Kalundu and Urewe. The distribution of these Traditions in sub-Saharan Africa are shown in the figure. (Obtained from Huffman 1989).

There are two major disjunctions of ceramic styles found in southern Africa after the first appearance of the Luangwa in Zambia (Fig. 5.6). The first or Blackburn group moved into Natal and the Transkei during the eleventh and twelfth centuries and is found in present-day Nguni peoples (Maggs 1984). The second or Moloko pottery occurs in the Phalaborwa region in the twelfth and thirteenth centuries and later on in the Transvaal plateau and are presently associated with Sotho-Tswana peoples (Evers 1982; Mason 1986). Both language families are part of eastern Bantu who practise patrilineal inheritance with the Central Cattle Pattern (Kuper 1982). They must therefore be derived from EIA cattle-keepers (Huffman 1989).

The above description is only a brief summary of the ways in which the archaeological evidence may be used to reconstruct southern African population prehistory, and the literature cited here represents only a small proportion of that which is actually published. Nonetheless, what is important, is to remember that there are several tools available with which archaeologists, in the absense of written records, can use to reconstruct the prehistory of the indigenous Khoisan and Negroid peoples.

As mentioned above, although both language and archaeology are powerful tools, they nevertheless are susceptible to change due to borrowing depending on the circumstances, with a result that often the true identity of individuals

or ethnic groups are masked. Using the linguistic and cultural affiliations of ethnic groups as a reference, the genetic similarities and differences of these groups may be used to provide a better understanding of population affinities and prehistory. This is discussed in the next section.

5.3 Genetic data

Prior to the advent of recombinant DNA technology, the genetic affinities of southern African populations were studied extensively using blood group markers and protein polymorphisms (Jenkins 1982; Nurse *et al.* 1985; Dunn 1990). With the availability of the latest molecular techniques in the 1980's, this study was expanded in our department to incorporate DNA polymorphisms using many different polymorphic markers. These included 30 different RFLP's of which 6 were linked to the cystic fibrosis gene (Morris *et al.* 1991), and variation at the α- and β-globin loci (Ramsay and Jenkins 1987; 1988).

The markers cited above are inherited in strict Mendelian fashion. We have also taken advantage of the maternally inherited mtDNA markers, the subject of this thesis, to investigate population variation contributed solely through female lineages. Likewise, the study of Y-chromosome RFLP's demonstrates the genetic contribution to diversity generated solely along male lineages and these results have

been reported by Spurdle and Jenkins (1992a,b,c). Variation in the number of tandem repeats (VNTR's) and $(AC)_n$ repeats have been exploited in population studies (Marques 1992).

The next section summarizes the usefulness of mtDNA polymorphisms, which, when used in conjunction with blood group markers, protein markers and other DNA markers, would lead to a better understanding of the genetic affinities of southern African peoples.

5.3.1. MtDNA variation in Negroids

The southern African Negroid sample consisted of 402 individuals (see Table 2.1) of whom 76 were southwestern Bantu-speakers (Herero and Ambo), 273 southeastern Bantu-speakers (Nguni, Tsonga, Sotho/Tswana, Venda and Lemba) and 53 Khoisan-speakers (Dama, Kwengo). The numbers and frequencies of mtDNA types found in various Negroid ethnic groups have been presented in Table 3.7.

5.3.1.1 Southeastern vs southwestern Bantu-speakers

The most striking differences between southeastern and southwestern Bantu-speakers was noted with mtDNA types 2-2 (3-1-1-1-3-2), 7-2 (3-1-1-1-1-2) and 21-2 (2-1-1-1-2-2). Types 7-2 and 2-2 are more common in southeastern Bantu-speakers than southwestern Bantu-speakers (Table 3.7).

Using these three mtDNA types the Ambo sample shows a closer genetic relationship with southeastern Bantu-speakers than with the Herero, whilst the Dama, who presently speak Nama are more closely related to the Herero than they are to the other Negroids or Khoisan populations. Type 2-2 which is found at a frequency of 32.5% in South African Negroids (Johnson *et al.* 1983) and 40% in Senegalese Negroids (Scozzari *et al.* 1988), and at varying frequencies in this study (Table 3.7), is found at a frequency of 13.5% in the Ambo but is not found at all in either the Herero or the Dama. In addition, type 21-2 which is found at a frequency of 50.0% in the Herero and 32.6% in the Dama, is found at a frequency of only 4.5% in the Ambo sample, and is not found at all in the other Negroids. These differences are statistically significant (Herero vs Ambo, $X^2_{[1]} = 13.88$, $p = 0.0002$; Dama vs Ambo, $X^2_{[1]} = 6.43$, $p = 0.0112$).

5.3.1.2 Relationship between the Herero and Ambo

The mtDNA findings have certain important historical implications regarding the traditional view that the Herero and Ambo are descended from two brothers. It is not clear whether Ngangombe and Kathu were uterine "brothers" or whether the term "brother", as translated by Hahn (1928), indicates only a vague relationship as is sometime the case in careless interpretations of kinship terms. This belief (based on oral tradition) was previously supported by sero-

genetic findings that the two groups had similar frequencies of a 6-phospho-gluconate dehydrogenase variant (*PGD*^{R}) and a minor haemoglobin variant (*Hbδ*^{B2}) (Nurse *et al.* 1987).

Variations detected on the Y chromosome may be used to test whether the Herero and Ambo are related patrilineally, since the Y chromosome is inherited only along patrilineal lineages. When using p49a/*Taq*I polymorphisms, a Y chromosome marker, Spurdle (1992) found minimal differences between the Herero and the Ambo. Due to the limited variability detected by another Y chromosome marker (p49a/*Pvu*II), it cannot be determined with certainty whether the male founders of the Ambo and Herero were genetically related (Spurdle 1992).

The distinction between Herero and Ambo mtDNA types, in particular the frequencies of mtDNA types 2,7, and 21 (Table 3.7), would suggest that at least some of the female founders of the Ambo population were of southeastern-Bantu origins whilst the female founders of the Herero have south western-Bantu origins.

5.3.1.3 Matrilineal inheritance in the Herero

It has been suggested that all Herero are descended from the daughters of one mother (Pennington 1990). Using family histories of a sample of Herero people living in

Botswana, Pennington (1990) has suggested that the subjects interviewed belonged to at least eight different *omaanda*. When using the mtDNA data obtained from a sample of Herero people from Namibia (present study), all mtDNA types found in this group of Herero could be traced to at least two female lineages. This suggestion is based on the observation that the majority of individuals within the Herero group sampled, have either mtDNA type 1-2 or type 21-2. (Incidentally, type 21 may be derived from type 1 due to a single mutation event involving the gain of an *Ava*II site at position 8229). Other types found in the Herero may be derived from either type 1 or type 21 by mutation or may have been acquired from other groups due to gene flow.

Since the Herero suffered large reductions in numbers due to numerous wars with the "Hottentots" and later the Germans (Vedder 1928), it is not inconceivable that this group could have suffered reduced variability even at the genetic level. It has been estimated that there were at least 70 000 to 85 000 Herero in Namibia (then South West Africa) before the Battle of Waterberg (1904-1908) and only 15 000 to 25 000 at its conclusion (Bridgeman 1981). The number of Herero who escaped to what is now Botswana is unknown, but the number of refugees has been estimated to be approximately 9 000 (Pennington 1990). In fact, control region sequence analysis using the Rogers and Harpending model (Rogers and Harpending 1992) also suggests that the

Herero population has experienced a severe population bottleneck (Fig. 4.8).

5.3.1.4 Origins of the Dama

The Dama who constitute approximately 8.5% of the total population of Namibia, which has been estimated to be approximately 1 million (Malan 1980), resemble Negroids in physical appearance but speak a Hottentot language, Nama. Since the Dama are biologically Negroid, Nurse and Jenkins (1977) have proposed that the Dama should be grouped in a distinct linguistic category of Khoisan-speaking Negroids. Dama have been known by a variety of names, the two most common being Damara and Bergdama (Vedder 1928; Nurse *et al.* 1976). There is no consensus in the literature concerning their origins. One school of thought claims that they were the aboriginal inhabitants of Namibia and, following the arrival of the "Hottentots", were forced into servitude to tend their cattle herds, lost their own language and acquired that of their masters (Vedder 1928; Knussmann and Knussmann 1969/1970). The observation that the Dama used words analogous to some in the Sudanese languages led Vedder (1928) to suggest that the "Hottentots", whom he thought were descended from north African "Hamites", captured African Negroid tribes (mainly males) to be used as slaves to tend their herds as they progressed southwards through the continent. Another hypothesis proposed by Jenny (1966) is that the Dama represent an amalgam of small

groups of a variety of Negro peoples who adopted the Hottentot language, though Nurse *et al.* (1977) have argued that there are no cultural anomalies to support this theory.

MtDNA studies reveal that the Dama and Herero have similar frequencies of mtDNA types (see Table 3.7) and the cluster analysis dendrograms shown in figures 3.9 and 4.5a,b also supports the suggestion that these two populations are genetically very closely related. In fact, it should be stressed that whereas most southeastern Bantu-speakers have a high frequency of morph *Hpa*I-3, Dama and Herero have a high frequency of morph *Hpa*I-2. Types 1 and 21, which have morph *Hpa*I-2, together constitute 85.2% and 53.5% of the mtDNA types found in Herero and Dama, respectively. This morph was previously found to be commoner in Caucasoid populations (Denaro *et al.* 1981; Johnson *et al.* 1983; Bonne-Tamir *et al.* 1986; Vikki *et al.* 1988; Sartoris *et al.* 1988; DeBenedictis *et al.* 1989a,b; Semino *et al.* 1989), whilst *Hpa*I-3 was found to be commonest in indigenous African populations (Johnson *et al.* 1983; Scozzari *et al.* 1988). Perhaps the high frequency of *Hpa*I-2 morphs in both these groups accounts for the closer cluster affinities of these groups when compared to other Negroid populations (Fig. 3.9).

As mentioned earlier, the Ambo who presently speak a western Bantu language, appear from the limited sample

studied here, to be derived from the same female founder/s as the present day eastern Bantu speakers, and their genetic classification is therefore different from their linguistic classification (Soodyall and Jenkins 1992b). While on a broad scale there is concordance between language groups and their genetic structure (Cavalli-Sforza *et al.* 1992), some discrepancies do exist.

5.3.2 Genetic relationships between the San and the Khoi

Several theories based on Khoisan history, culture, art, physical appearance and language, have been proposed to explain the origins of both the San and Khoikhoi but their origins remain an enigma. According to Stow (1905), San occupied the northern regions of Africa and were displaced by powerful warlike tribes, causing them to migrate into southern Africa. He claims that there were two main groups of San: sculptors and painters, who entered southern Africa along different routes. Sculptors entered southern Africa along the central portion of the country and settled in the northern Cape region between the Vaal and Riet Rivers, whereas the painter group advanced through Damaraland on the West Coast and, on reaching the great mountain ranges in the south, turned eastward and settled in the region of present-day Botswana. Theal (1910), who doubted Stow's theory, believed that the San had originated in Asia, and that these hunter-gatherers eventually entered Africa and spread across North Africa gradually proceeding to southern

Africa. He claims that the Pygmies of central Africa are descendants of the original San population. Schapera (1930) does not accept Theal's suggestion that the Pygmies and San are related, nor does he support the claim of an Asian origin of the San. He suggests that the San originally occupied areas as far north as Ethiopia and, when besieged by more powerful tribes, ventured southwards and entered southern Africa.

The observation that the click languages spoken by Khoikhoi people shared some characteristics in common with "Hamites", who were pastoral people living in East Africa, led Vedder (1928, 1938) to conclude that Khoikhoi were San people who came under the influence of the "Hamites" and migrated into southern Africa much later than the San.

Previous sero-genetic marker studies on the San and Nama have revealed differences in frequencies of certain ABO and Rhesus blood group alleles, as well as in the immunoglobulin allotypes (Zoutendyk *et al.* 1953,1965; Jenkins *et al.* 1971; Jenkins 1972). In fact, the immunoglobulin haplotypes $Gm^{1,21}$ and $Gm^{1,5,13,14,21}$ were only found in the San whereas haplotype $Gm^{1,2,21}$ was only found in Khoikhoi individuals (Steinberg *et al.* 1975). It has been suggested that the low frequencies of these unique haplotypes found in some southern African Negroid populations may be due to genetic admixture between these groups and the San and Nama groups (Jenkins 1972). Jenkins

(1972) has used the frequencies of these unique haplotypes as an index to estimate the extent of gene admixture from the Khoisan groups into the Negroid groups.

MtDNA studies have revealed that the San have a unique mtDNA type, type 4, found at a frequency of 42.8% in Sekele and 26.5% in the !Kung, but not found in the Nama (Table 5.1) or Negroid groups studied (Johnson *et al.* 1983; Scozzari *et al.* 1988). This observation leads to the suggestion that gene admixture between San and Negroids mainly involved male San and Negroid females. This suggestion is further supported by the observation that Negroid mtDNA types 2 and 7, found at zero and low levels (6.5%), respectively, in the Nama, have been found at varying frequencies in the San (Table 5.1), providing further evidence to support the suggestion that females from Negroid groups have been assimilated into the San group.

By way of contrast, there are no unique mtDNA types in the Nama, although type 3 which is found at a frequency of 54.3% in the Nama, is less common in the San (9.5% in the Sekele and 29.5% in the !Kung). These differences are statistically significant (Nama vs Sekele [$X^2_{[1]} = 15.3$, $p = 9.17 \times 10^{-5}$]; Nama vs !Kung [$X^2_{[1]} = 5.13$, $p = 0.024$]). Since type 3 is found at low frequencies in some Bantu-speaking Negroid groups in southern Africa (Johnson *et al.* 1983; this study) and not found in Senegalese Negroids (Scozzari

et al. 1988), it is likely that the Nama have contributed this type to extant Negroid groups in southern Africa. Studies using the *Gm* system confirmed significant gene flow from the Khoisan to the Bantu-speaking chiefdoms (Jenkins *et al.* 1970).

There are at present insufficient data to confirm whether prehistoric Negroid populations had type 3 in their genetic pool. If type 3 were found in the extant populations of East Africa, it would support the view that the common ancestors of both Khoisan and Negroids possessed type 3 and that the present differences in frequencies are due to genetic drift or selection. If type 3 is not found in these populations, it would confirm that this type originated in Khoisan populations and Negroids with this type would have acquired it as a result of gene flow from the Khoisan. Should this be so, then type 3 could be used to reconstruct the early history of Khoisan populations in Africa.

In some instances where serogenetic markers were found at higher frequencies in the Nama than in the San, it has been suggested that this may be due to the unilateral flow of genes as is expected in populations subjected to enslavement (Nurse *et al.* 1976). Sero-genetic studies revealed that several genetic markers (ABO blood group allele *B*, Rhesus allele *r*, immunoglobulin $Gm^{1,5,6}$ allele, 6-phosphogluconate dehydrogenase PGD^{c} allele, adenylate

kinase AK^2 allele, glucose 6-phosphate dehydrogenase Gd^A and haptoglobin Hp^1 allele) found at high frequencies in the Dama, are not found in the San but are found at appreciably frequencies in the Nama. The introduction of these genes into the Nama, it was suggested, resulted from the unidirectional flow of genes from the Dama into the Nama as a result of the enslavement of the Dama by the Nama (Nurse *et al.* 1976). Due to this enslavement, unions between Nama males (dominant group) and Dama females (subservient group) were "allowed", but not *vice versa*. It has been observed that these females and their offspring were incorporated into the Nama group, thereby increasing the frequency of alleles of the subservient group in the dominant group (Nurse *et al.* 1976).

Since mtDNA is maternally inherited, one would expect to find some Dama (subservient) mtDNA types being incorporated into the Nama (dominant) group. MtDNA type 21-2 (2-1-1-1-2-2), the commonest type in the Dama group (32.6%), has attained a frequency of 13.0% in the Nama, but is found at a low frequency in the Sekele (4.1%) and is absent in !Kung. This further supports the suggestion that there has been little gene flow from the Dama into the San (Nurse *et al.* 1976).

Not all types which show a disparity in frequency between San and Nama groups may be attributable to Dama enslavement. Type 3 which has a frequency of 54.3% in the

Nama is present at a low frequency in the Dama (2.3%). Gene admixture between San and Negroid populations has also contributed to the differences. Type 7 (3-1-1-1-2) which is very common in Negroids (Johnson *et al.* 1983; Scozzari *et al.* 1988; present study), is found at lower frequencies in Khoisan (7.1% in the Sekele; 5.5% in the !Kung; 6.5% in the Nama), suggesting that this type may have been acquired from Negroids. Since type 7 is not found in neighbouring Dama or Herero Negroid populations in Namibia, a likely source may be the Ambo in whom a frequency of 54.5% was found.

A striking difference between the Sekele !Kung from Namibia (present study) and the !Kung group from Botswana (Johnson *et al.* 1983), is in their frequencies of type 5 (3-1-1-2-5), which occurs at a frequency of 20.6% (7/34) in !Kung and only 4.7% (1/49) in the Sekele ($X^2_{[1]} = 5.94$, $p = 0.015$). It is found at a low frequency in the Nama (6.5%) and is uncommon in Negroids (Johnson *et al.* 1983, Scozzari *et al.*, 1988). Type 5 may be derived from type 3 by gaining an *Msp*I site at position 13070 of the reference sequence (Anderson *et al.* 1981). Since types 3 and 5 are found in both San and Nama populations, it is possible that both types may have been present in the mtDNA pool of extinct Khoisan founders. The higher frequency of type 5 observed in the !Kung may be due to genetic drift, a result of their relative isolation (Soodyall and Jenkins 1992a).

The higher frequency of type 3 in the !Kung (26.5%) when compared to the Sekele San (9.5%), suggests that there has been more gene flow from the Nama into the former than into the latter group. This observation was also made from sero-genetic studies in which San groups, excluding the Sekele, cluster closer to the Nama, with the difference in the Sekele being attributed to the contribution of Negroid genes to their genetic structure (Jenkins 1986).

Although the frequencies of mtDNA types 3 and 5 differ appreciably between the three Khoisan groups studied here, the difference in frequencies of the other types are not significant.

There appears to be some controversy in the literature regarding the ethnographic status of the Kwengo or "Black Bushmen". According to de Almeida (1965), they are a group of "Bushmen" who closely resemble Negroids in physical appearance, but who are culturally and linguistically more closely related to the Khoisan. Using sero-genetic markers, Nurse and Jenkins (1977) showed that the serogenetic profile of the Kwengo is eminently Negroid. Although the sample size of the Kwengo used in the present study is small, the Negroid-like mtDNA types 2 and 7 were found at frequencies of 20% and 60%, respectively, providing further support for the classification of this group as Khoisan-speaking Negroids (together with the Dama) rather than genetically Khoisan.

Table 5.1. The numbers and frequencies (in percent) of the various mtDNA types found in one Negroid and three Khoisan populations from southern Africa, together with an index of genetic diversity (*h*).

		SEKELE (N = 49)			NAMA (N = 46)			!KUNG (N = 34)			KWENGO (N = 10)		
TYPE	MITOTYPE	No	Freq	(S.E)	No	Freq	(S.E)	No	Freq	(S.E)	No	Freq	(S.E)
1*	2-1-1-1-1	1	2.0	(2.0)	3	6.5	(3.6)	1	2.9	(2.9)	2	20.0	(12.6)
2*	3-1-1-1-3	-	-	-	-	-	-	1	2.9	(2.9)	2	20.0	(12.6)
3*	3-1-1-2-2	7	14.3	(4.9)	25	54.3	(7.3)	9	26.5	(7.6)	2	20.0	(12.6)
4*	3-1-1-3-2	20	40.8	(7.0)	-	-	-	9	26.5	(7.6)	-	-	-
4-9	3-1-1-3-2-9	1	2.0	(2.0)	-	-	-	-	-	-	-	-	-
5*	3-1-1-2-5	1	2.0	(2.0)	3	6.5	(3.6)	7	20.6	(6.9)	-	-	-
7*	3-1-1-1-1	10	20.4	(5.6)	3	6.5	(3.6)	2	5.9	(4.0)	4	40.0	(15.5)
10*	3-1-1-1-2	6	12.2	(4.7)	3	6.5	(3.6)	2	5.9	(4.0)	-	-	-
14*	3-1-1-2-4	-	-	-	-	-	-	2	5.9	-	-	-	-
21-2	2-1-1-1-2-2	2	4.1	(2.8)	6	13.0	(4.9)	-	-	-	-	-	-
31*	3-1-1-5-1	-	-	-	-	-	-	1	2.9	(2.9)	-	-	-
113-2	3-1-1-12-1-2	1	2.0	(2.0)	-	-	-	-	-	-	-	-	-
139-2	7-1-6-13-1-2	-	-	-	2	4.3	(3.0)	-	-	-	-	-	-
140-2	3-1-1-1-6-2	-	-	-	1	2.2	(2.1)	-	-	-	-	-	-
	*h***		0.768			0.683			0.829			0.800	

.Data from Johnson *et al.* (1983).

*Types exclude *Hinc* II morphs since this enzyme was not used by Johnson *et al* (1983). In the present study, these types have morph *Hinc* II-2.

***h* = genetic diversity (Nei 1987).

5.3.3 MtDNA polymorphism in "Coloureds"

The term "Coloured" is supposedly used to indicate a people who has resulted from hybridization between Caucasoid males and non-Caucasoid females. However, several other patterns of genetic admixtures involving slaves (Negroid, Asian), indigenous peoples of Africa (Khoi, San, Negroid) as well as various peoples from Europe have also contributed to the genetic diversity of the "Coloured" population. There were significant local variations in the patterns of admixture. Since many of these unions did not result in wedlock and were temporary, it is difficult to document which processes of hybridization have actually given rise to the South African "Coloured" population.

Using blood group markers, Jenkins (1972) has investigated the genetic contributions of the parental stocks of the "Coloured" population living in the Johannesburg area, and estimates that their genetic pool consists of approximately 30% Caucasoid, 30% Negroid, 10% San and 30% Khoi genes. In another study, Botha and Pritchard (1972) showed that the genetic pool of "Coloureds" in the Cape corresponded with their ethnicity. They divided the Cape "Coloured" group into those who were Moslems (Cape Malays) and non-Moslems (non-Malay "Coloured") and, using the Rhesus blood group allele frequencies, they calculated that the genetic pool of the Cape Malays consists of 42% Asian, 33% Western European and 25% southern African genes, whilst the non-

Malay "Coloured" group consists of 46% southern African, 32% Western European and 22% Asian genes. These estimates do not indicate what proportions of genes have been contributed by males and/or females.

Since mtDNA is maternally inherited, it may be used to estimate the female contribution to the "Coloured" population. Table 5.2 gives the frequencies of the mtDNA types found in the two "Coloured" populations investigated in this study: from the Richtersveld area in the northern Cape and from the Johannesburg region in the Transvaal, referred to here as the Richtersveld "Coloured" and Johannesburg (JHB) "Coloured" groups, respectively. The frequencies of mtDNA types found in these "Coloured" groups are compared to those found in southern African Negroid and Khoisan populations as well as those found in Senegalese Negroids.

The pooled "Coloured" population consists of 17 mtDNA types of which only 5 (1-2, 3-2, 5-2, 7-2, 10-2) are common to both the Richtersveld and JHB "Coloured" populations. Although type 1-2 is found in all populations investigated to date, its frequency in the Richtersveld population (5.7%) is similar to that found in Khoisan populations (2-7%) whilst its frequency in the JHB "Coloured" population (18.3%) is similar to that found in southeast-Bantu speakers (17.3%) and the Dama (20.9%).

Table 5.2. The numbers and the frequencies of mtDNA types found in the Richtersveld and the Johannesburg "Coloured" populations from South Africa. The frequencies of these types in the pooled Negroid, Khoisan and Caucasoid populations from Africa are given for comparisons.

	S.E.BANTU		HERERO		DAMA		SENEGALESE		SEKELE		!KUNG		NAMA		RICHTERS		JHB		CAUCASOID	
	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq
	295		54		43		186		49		34		46		35		71		192	
1-2	51	17.3	18	33	9	21	41	22.0	1	2.0	1	2.9	3	6.5	2	5.7	13	18.3	142	74.0
2-2	68	23.1	-	-	-	-	71	38	-	-	1	2.9	-	-	-	-	8	11.3	1	0.5
3-2	23	7.8	3	5.6	1	2.3	-	-	7	14.3	9	26.5	25	54.3	20	57.1	17	23.9	1	0.5
3-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	5.7	-	-	-	-
5-2	6	2	1	1.9	1	2.3	-	-	1	2.0	7	20.6	3	6.5	4	11.4	13	18.3	-	-
5-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	2.9	-	-	-	-
7-2	89	30.2	-	-	1	2.3	45	24	10	20.4	2	5.9	3	6.5	2	5.7	8	11.3	-	-
8-2	2	0.7	-	-	-	-	3	1.6	-	-	-	-	-	-	-	-	1	1.4	-	-
9-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1.4	-	-
10-2	11	3.7	1	1.9	-	-	2	1.1	6	12.2	2	5.9	3	6.5	2	5.7	2	2.8	-	-
33-2	2	0.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1.4	-	-
47-2	3	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	5.6	5	2.6
115-2	1	0.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1.4	-	-
116-2	-	-	1	1.9	2	4.7	-	-	-	-	-	-	-	-	1	2.9	-	-	-	-
120-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1.4	-	-
121-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1.4	-	-
124-2	2	0.7	-	-	3	70.0	-	-	-	-	-	-	-	-	1	2.9	-	-	-	-

Although the frequency of type 1-2 is somewhat higher in the Herero than in the JHB "Coloured" population, the difference is not statistically significant. Thus, when using type 1-2 only, JHB "Coloureds" closely resemble South African Negroids whilst Richtersveld "Coloureds" have a frequency similar to that found in the Khoisan.

This observation is further supported by comparing the frequencies of types 2-2 and 7-2 which are found at a higher frequency in southeastern Bantu-speakers (Table 5.2) than in the Khoisan populations (see Table 5.1). Type 2-2 is not found in the Richtersveld "Coloured" population but is found at a frequency of 11.3% in the JHB "Coloured" population; type 7-2 is found at an appreciably lower frequency in the Richtersveld "Coloured" population (5.7%) when compared with the JHB "Coloured" population (11.3%).

Four mtDNA types: 3-2 (3-1-1-2-2-2), 4-2 (3-1-1-3-2-2), 5-2 (3-1-1-2-5-2) and 10-2 (3-1-1-1-2-2) appear to be commoner in Khoisan populations than Negroids (Table 5.2) and may be used to estimate the contribution of Khoisan genes to the "Coloured" populations investigated. Of these mtDNA types, type 4-2 is unique to the San (40.8% in the Sekele and 26.5% in the !Kung) and type 3-2 is found at a significantly higher frequency in Nama (25/46) than in the Sekele (7/49) [$X^2 = 15.3$, $p < 0.001$].

The frequencies of types 3,5 and 10 in Khoisan populations suggests that these types were found in the proto-Khoisan population and that type 4 evolved in the San following their hypothesized division into the Khoi and San groups. Type 3-2 is found at a significantly higher proportion in the Richtersveld "Coloured" population (20/35) when compared to the JHB "Coloured" population (17/71) [X^2 = 9.96, p=0.0016]. This observation supports the historical suggestion that the Richtersveld "Coloured" population is derived from gene admixture between Caucasoid males and Khoi females (Nurse *et al*. 1985).

Excluding the rare/singleton mtDNA types found in the Richtersveld "Coloured" (5.8%) and JHB "Coloured" (14.0%) populations, the mtDNA pool of Richtersveld "Coloureds" consists of 5.7% Negroid mtDNA types (types 2 and 7) and 82.8% Khoisan (types 3,5 and 10). The higher frequency of type 3 in the Nama (54.3%) when compared with the !Kung (26.5%) and the Sekele (14.3%) would suggest that this type was derived from gene flow from the Khoi into the Richtersveld "Coloured" population. Thus, if type 3 is used as an index to study the contribution of only Khoi genes, then approximately 62.8% of the Richtersveld "Coloured" population may have Khoi ancestry. In contrast, the JHB "Coloured" population is made up of 45.1% Khoisan genes of which 23.9% may be inherited from the Khoi (using type 3) and 21.3% from either the Khoi or the San.

Using types 2 and 7 as the index for studying the contribution of Negroid genes, the JHB "Coloured" population is made up of 22.6% Negroid genes. Thus, excluding types 1 and rare types, the female contribution of genes in the Richtersveld "Coloured" population is derived from approximately 94% Khoisan and only 6% Negroid mtDNAs whereas the JHB "Coloured" population has approximately 67% Khoisan and 33% Negroid mtDNAs.

5.3.4 MtDNA variation in the South African Indian population

The Indian population in South Africa has presently grown to approximately 1 million and constitutes approximately 3.0% of the total population (South African Statistics 1990). The vast majority of Indians (about 80%) reside in Natal, with only 15% and 4% having settled in the Transvaal and Cape Provinces, respectively (South African Statistics 1990). Until recently Indians were barred from living in the Orange Free State, unless prior consent was obtained from the Government, with the result that very few individuals of Indian origin reside there (53 out of a population of 821 361, South African Statistics 1990). A map showing the regions of origins of both indentured and "passenger" Indians from India, as well as their regions of settlement in South Africa is shown in Fig. 5.7.

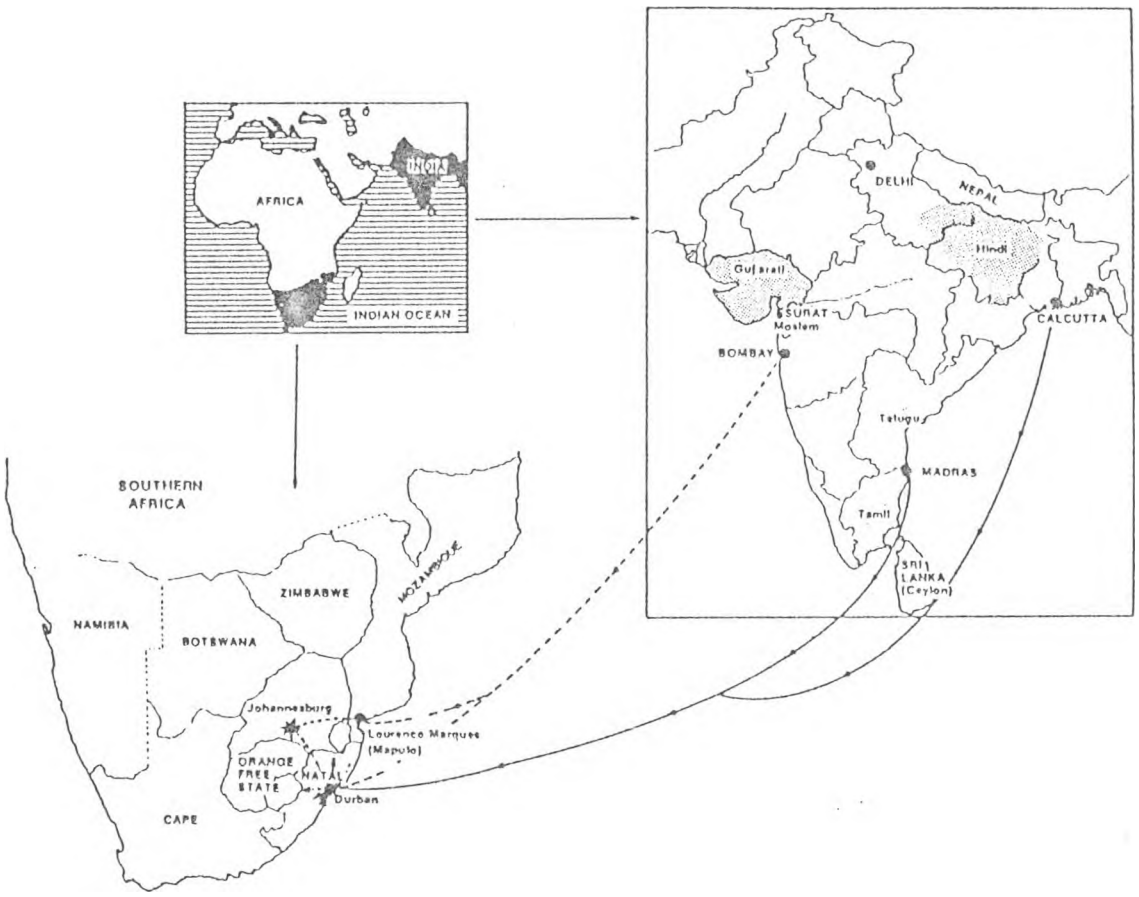


Fig. 5.7. Map showing the regions in India from where the South African Hindi, Tamil, Gujerati and Moslem speaking Indian groups originated. Indians first arrived in South Africa either as indentured labourers (——) or as traders (commonly referred to as "passenger" Indians (---)). See text for details.

South African Indians (S.A. Indians) can be conveniently subdivided by their religion and language into groups that correspond reasonably well with their known geographical centres of origin in India (Grierson 1927). The majority of Indians in South Africa belong to the Hindu religious group in which four linguistic groups may be identified: Tamil, Telugu, Hindi and Gujerati, whereas a smaller proportion of Indians, living predominantly in the Johannesburg area, are Moslems, who follow the Sunni tenets of Islam and are either Urdu- or Gujerati-speakers (mainly Gujerati). Tamil and Telugu are ancient vernacular languages of Dravidian origin in southern India, whereas Hindi and Gujerati belong to the Aryan or Indo-European family of languages (Bhana and Brain 1990). Hindi is based on Sanskrit which arose in north-western India, whereas Gujerati originated in western India.

Although the majority of "passenger" Indians were Moslem and came from the Indian State of Gujarat and spoke mainly Gujerati (Fig. 5.8), some came from more northerly States and spoke Urdu. The Urdu language is Aryan in origin and evolved in northwestern India following the Moslem invasions of the 15th and 16th centuries. The invaders came from Central Asia, Arabia, Mesopotamia and Iran. Urdu is a form of Hindi on which many Farsi and Arabic words and constructs have been grafted. It is spoken by most South African Moslems (Meer 1969).

The sample size of the Indian group discussed in Chapter 3 consisted of only 50 individuals. More recently, this RFLP study was extended to investigate the relationships between religious and linguistic groups living in South Africa. The extended sample consists of 147 unrelated Indians of whom 37 were Tamil-speaking Hindus (Tamil), 45 Hindi-speaking Hindus (Hindi), 24 Gujerati-speaking Hindus (Gujerati), and 41 Gujerati/Urdu-speaking Moslems (Moslem).

The frequencies of enzyme morphs for the four groups of South African (S.A.) Indians classified according to language and religion, as well as the pooled S.A. Indian sample are shown in Table 5.3. With the exception of six individuals who have morph *Hpa*I-3 and one individual with morph *Hpa*I-5, the S.A Indians are almost monomorphic for *Hpa*I-2 (95.2%). *Hpa*I-2 is also the commonest morph in all Caucasoid populations studied to date (Johnson *et al.* 1983; Santachiara Benerecetti *et al.* 1988; Sartoris *et al.* 1988; De Benedictis *et al.* 1989 a,b; Semino *et al.* 1989; Vikki *et al.* 1988).

The S.A. Indian population is monomorphic with the restriction enzymes *Bam*HI and *Hae*II (morphs *Bam*HI-1 and *Hae*II-1, respectively), whereas some variation was previously observed with these enzymes in Hindus from New Delhi and Nepal, albeit at low frequencies (Semino *et al.* 1991). *Msp*I-2 was observed in only one individual, whereas the rest had the common Caucasoid morph *Msp*I-1.

Six previously reported *Ava*II morphs were observed (Table 5.3). *Ava*II-1 was the commonest morph, found in 91.8% of the sample; morph *Ava*II-2 was found in all four groups, although at low frequencies (Table 5.3). Morphs *Ava*II-3 and *Ava*II-5 were found in only the Moslem group at frequencies of 7.3% and 2.4%, respectively, and only one Hindi individual was found with *Ava*II-6 (2.2%). *Ava*II-12, first found among Japanese (Horai and Matsunaga 1986) and later in one Hindu individual from New Delhi (Semino *et al.* 1991), was found in one Tamil individual in the present study.

Whereas the Hindi and Gujerati groups are monomorphic for morph *Hinc*II-2, some variation was observed in the Tamil and Moslem groups (Table 5.3). Morph *Hinc*II-7 was found in two individuals (one Tamil and one Moslem), whilst *Hinc*II-10 was found only in one Tamil individual. Although the majority of Indians (98.0% in the present study and 91.1% in Hindus from New Delhi and Nepal, Semino *et al.* 1991) were found to have morph *Hinc*II-2, three types (1, 18, and 21) showed variation with this enzyme, giving rise to four subtypes of type 1 and two subtypes each of types 18 and 21 (Table 5.4).

Table 5.4 shows the mtDNA types and their frequencies in the four S.A. Indian groups, including the pooled S.A. Indian population, and these are compared with the data on Indians from Nepal and New Delhi (Semino *et al.* 1991). Of

the twenty-six mtDNA types found in Indians, only five types (types 1-2, 1-7, 1-10, 47-2 and 72-2) are shared by both the pooled sample from South Africa and the pooled Hindu sample (Semino *et al*. 1991). The commonest mtDNA type in both Indian samples is type 1-2, with the frequency being higher in S.A. Indians (87.8%) than in Hindus from Nepal and India (64.6%).

Of the thirteen mtDNA types found in S.A. Indians, five types, 2-2 (3-1-1-1-3-2), 3-2 (3-1-1-2-2-2), 7-2 (3-1-1-1-1-2), 10-2 (3-1-1-1-2-2) and 31-2 (3-1-1-1-5-2), are defined by *Hpa*I-3, whereas all eighteen types found in the Hindu sample (Semino *et al*. 1991) have *Hpa*I-2. Previous studies have concluded that *Hpa*I-3 is unique to indigenous African populations (Johnson *et al*. 1983; Scozzari *et al*. 1988).

The frequency of this morph in non-African populations suggests the extent of gene admixture between indigenous African females and non-African males. In S.A. Indians, 4.2% of individuals had mtDNA types including *Hpa*I-3, and it is possible that individuals with these types may have acquired it as a result of miscegenation between Indian males and either Negroid or "Coloured" females, the latter containing mtDNA types from either indigenous Negroid or Khoisan populations in South Africa.

Table 5.3. MtDNA variations detected in four S.A. Indian linguistic/religious groups, including the pooled sample with six restriction enzymes.

		Gujerati N=24			Tamil N=37			Hindi N=45			Moslem N=41			S.A.Indian* N=147		
Enzyme	Morph	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.)	No	Freq	(S.E.
Hpal	-2	23	95.8	(4.1)	36	97.3	(2.7)	45	100		36	87.8	(5.1)	140	95.2	(1.8)
	-3	1	4.2	(4.1)	1	2.7	(2.7)	-	-		4	9.8	(4.6)	6	4.1	(1.6)
	-5	-	-		-	-		-	-		1	2.4	(2.4)	1	0.7	(0.7)
BamHI	-1	24	100		37	100		45	100		41	100		147	100	
HaeII	-1	24	100		37	100		45	100		41	100		147	100	
MspI	-1	24	100		37	100		45	100		40	97.6	(2.4)	146	99.7	(0.7)
	-2	-	-		-	-		-	-		1	2.4	(2.4)	1	0.7	(0.7)
AvaII	-1	23	95.8	(4.1)	35	91.9	(3.7)	43	95	(3.1)	34	82.9	(5.9)	135	91.8	(2.2)
	-2	1	4.2	(4.1)	1	2.7	(2.7)	1	2.2	(2.2)	3	7.3	(4.1)	6	4.1	(1.6)
	-3	-	-		-	-		-	-		3	7.3	(4.1)	3	2	(1.2)
	-5	-	-		-	-		-	-		1	2.4	(2.4)	1	0.7	(0.7)
	-6	-	-		-	-		1	2.2	(2.2)	-	-		1	0.7	(0.7)
	-12	-	-		1	2.7	(2.7)	-	-		-	-		1	0.7	(0.7)
HincII	-2	24	100		35	94.6	(3.7)	45	100		40	97.6	(2.4)	144	98	(1.2)
	-7	-	-		1	2.7	(2.7)	-	-		1	2.4	(2.4)	2	1.3	(0.9)
	-10	-	-		1	2.7	(2.7)	-	-		-	-		1	0.7	(0.7)

*Pooled South African Indian sample.

Table 5.4. The numbers and the frequencies (in percentage) of mtDNA types found in four linguistic/religious South African Indian populations and Hindus from Nepal and India.

		Gujerati N=24		Tamil N=37		Hindi N=45		Moslem N=41		S.A.Indian N=147		Nepal N=34		New Delhi N=45		Hindu N=79	
Type	Mitotype	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq	No	Freq
1-2	2-1-1-1-1-2	22	91.7	32	86.5	43	95.6	32	78.0	129	87.8	19	55.9	32	71.1	51	64.6
1-7	2-1-1-1-1-7	-	-	1	2.7	-	-	1	2.4	2	1.4	2	5.9	-	-	2	2.5
1-10	2-1-1-1-1-10	-	-	1	2.7	-	-	-	-	1	0.7	-	-	1	2.2	1	1.3
1-16	2-1-1-1-1-16	-	-	-	-	-	-	-	-	-	-	1	2.9	-	-	1	1.3
2-2	3-1-1-1-3-2	-	-	-	-	-	-	1	2.4	1	0.7	-	-	-	-	-	-
3-2	3-1-1-2-2-2	-	-	-	-	-	-	1	2.4	1	0.7	-	-	-	-	-	-
6-2	2-1-2-1-1-2	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
7-2	3-1-1-1-1-2	1	4.2	1	2.7	-	-	-	-	2	1.4	-	-	-	-	-	-
8-1	1-1-1-1-1-1	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
10-2	3-1-1-1-2-2	-	-	-	-	-	-	1	2.4	1	0.7	-	-	-	-	-	-
11-2	2-2-3-1-5-2	-	-	-	-	-	-	-	-	-	-	2	5.9	-	-	2	2.5
18-2	2-3-1-4-9-2	-	-	-	-	-	-	-	-	-	-	2	5.9	-	-	2	2.5
18-9	2-3-1-4-9-9	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
21-2	2-1-1-1-2-2	1	4.2	1	2.7	1	2.3	1	2.4	4	2.7	-	-	-	-	-	-
21-9	2-1-1-1-2-9	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
23-2	2-2-1-1-3-2	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
25-2	5-1-1-1-1-2	-	-	-	-	-	-	1	2.4	1	0.7	-	-	-	-	-	-
28-2	2-1-1-4-1-2	-	-	-	-	-	-	-	-	-	-	5	14.7	1	2.2	6	7.6
31-2	3-1-1-1-5-2	-	-	-	-	-	-	1	2.4	1	0.7	-	-	-	-	-	-
47-2	2-1-1-1-3-2	-	-	-	-	-	-	2	4.9	2	1.4	-	-	1	2.2	1	1.3
56-2	2-1-1-1-6-2	-	-	-	-	1	2.3	-	-	1	0.7	-	-	-	-	-	-
72-2	2-1-1-1-12-2	-	-	1	2.7	-	-	-	-	1	0.7	-	-	1	2.2	1	1.3
82-2	2-1-1-16-1-2	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
83-2	2-1-1-1-29-2	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3
84-2	2-0-1-1-30-2	-	-	-	-	-	-	-	-	-	-	3	8.8	1	2.2	4	5.1
85-2	2-1-1-7-18-2	-	-	-	-	-	-	-	-	-	-	-	-	1	2.2	1	1.3

Data from Scozzari *et al.* (1991).
MtDNA type numbers are given for the enzyme morphs *Hpa*I, *Bam*HI, *Hae*II, *Msp*I and *Ava*II (in this order). The addition of *Hinc*II morphs gives rise to a mtDNA subtype.

A correlation between mtDNA types and the extent of homogeneity as defined by Johnson *et al.* (1983) for the different groups is shown in Table 5.5. The high homogeneity may be related to the strict caste system which dictates that marriages occur only within the caste or clan, tribe or group of birth (Sanghvi 1966; Vyas *et al.* 1958). This rule was associated with the Hindu belief that the factors which govern "inheritance" were passed down to them only by their fathers. Dravidian-speaking Hindus (Tamil and Telugu) forbid marriages between the children of brothers, while permitting them between the children of sisters, and between uncles and nieces. Indo-Aryan-speaking Hindus (Hindi and Gujerati), on the other hand, forbade all cross- and parallel-cousin marriages. Moslem Indians permit both cross- and parallel-cousin marriages. The majority of Indians in South Africa still abide by these ancient customs, and tend to marry within religious and language groups.

Very little variation was observed between Indian religious and language groups (Tables 5.3 and 5.4). The majority of individuals in each group were found to have mtDNA type 1-2 (86.5% in the Tamils, 95.6% in the Hindi, 78.0% in the Moslem and 91.7% in the Gujerati), making it difficult to assess the correlation between mtDNA diversity and Indian marriage and cultural customs as alluded to earlier. Other types occur rarely and are found at low frequencies (Table 5.4).

Whereas the Hindi and Gujerati group consisted of three types each, six different mtDNA types were found in the Tamil group and nine in the Moslem group (Table 5.5). Five of the nine types found in the Moslem group, contained the African *Hpa*I-3 morph, whilst one contained *Hpa*I-5, previously observed in a Caucasoid individual (Johnson *et al*. 1983). Two individuals (one Tamil and one Gujerati) were found to have the *Hpa*I-3 morph, giving rise to type 7-2, which is common in Negroids (Johnson *et al*. 1983; Scozzari *et al*. 1988).

Indians with *Hpa*I-3 may have acquired it due to miscegenation between Indian males and "Coloured", Negroid or Khoisan females. It has been suggested that the genetic pool in the "Coloured" group has been largely derived from Negroids, San, Khoi and possibly slaves brought to the Cape from Indonesia and Malaysia (Jenkins 1972). The mtDNA pool in the Moslem group has presumably been influenced by gene flow from the Cape "Coloured" population, also known as the Cape Malay, who are also Moslem. Marriages between Cape Malays and Indians do occur, albeit infrequently, and this may account for the introduction of African mtDNA types into the Indian group. According to the Race Relations Survey (1990), during the previous year 52 Indians applied for "re-classification" of "race" to "Coloured" and 63 "Coloured" individuals applied for a change to "Indian", suggesting that some intermarriage between the two groups have occurred.

Table 5.5. Correlation between number of mtDNA types and homogeneity (*F*) in Indian groups from South Africa, Nepal and India.

Group	N	No of Types	*F*
Gujerati	24	3	0.84
Tamil	37	6	0.75
Hindi	45	3	0.91
Moslem	41	9	0.62
S.A.Indian	**147**	**13**	**0.77**
Nepal	34	7	0.35
New Delhi	45	14	0.51
Hindu	**79**	**18**	**0.58**

Data from Scozzari *et al*. (1991).
Test of homogeneity (Johnson *et al*. 1983).
N = Number of individuals.

Studies based on blood group markers (Moores 1980; Nurse *et al.* 1985) as well as on the immunoglobulin allotypes (Jenkins *et al.* 1970) support the hypothesis that gene admixture between S.A. Indians and Negroids has occurred. Using the $Gm^{1,5,13,14}$ haplotype found at a frequency of 50.9% in Natal Zulus, Jenkins *et al.* (1970) attributed the frequency of 0.06% of this haplotype found in Indians from Durban, to gene flow from the local Zulus. If it were assumed that this haplotype was not found in Indians when they arrived from India, the extent of Zulu admixture into Indians in Durban would be approximately 1% (Jenkins *et al.* 1970). Since 4.2% of mtDNA types found in S.A. Indians consists of the non-Caucasoid *Hpa*I-3 morph, the maternal contribution of indigenous African mtDNAs into S.A. Indians may be estimated to be approximately 4%. In contrast, no such gene admixture has been observed in Indians from India and Nepal (Semino *et al.* 1991), suggesting that the original mtDNA pool of Indians brought to South Africa might well have been monomorphic for morph *Hpa*I-2.

Although mtDNA does not provide a good marker system for distinguishing between Indian religious and language groups, blood group marker frequencies in both the Natal and Johannesburg Tamil, Telugu- and Hindi-speaking Indians may be used to distinguish Hindus from Moslems. For the most part, the frequencies of these markers resemble those of corresponding populations in north-eastern and south-eastern India, whereas the Moslem Indians formed a unit

corresponding to frequencies observed in western India (Moores 1980). In addition, Moslem Indians in the Johannesburg area also have a high frequency (29%) of an inverted Y-chromosome polymorphism (not found in other groups), and the origin of this polymorphism was traced to the city of Surat in the province of Gujarat (Bernstein *et al.* 1986), providing additional genetic support for the different origins of religious/language groups of Indians presently inhabiting South Africa (Soodyall and Jenkins 1992c).

5.3.5 MtDNA structure of European Caucasoids in South Africa

The "White" population of S.A. has grown to approximately 4.6 million since Jan van Riebeeck set up a Dutch Colony at the Cape in 1652 (South African Statistics 1990). The Afrikaner people of S.A. have since evolved from a gradual amalgamation of "Dutch" elements with an appreciable French element consisting of the Huguenots who settled in the Cape in the late 17th century. Dutch settlement in S.A. both antedated and followed the arrival of the French and by 1807, 48.6% of Afrikaners in S.A. were of Dutch origin, 1.4% Flemish, 25.8% German, 17.3% French, 3.0% Scandinavian and the remaining 4.0% were from the rest of Europe (Colenbrander 1902). Of the approximately 4.6 million "Whites" in S.A., apporoximately 2.5 million are Afrikaans-speakers, 1.6 million English-speakers and the rest either

speak Dutch (12 000), German (41 000), Greek (17 500), Italian (14 300), Portuguese (56 600), French (7 300) and other (93 500) languages (South African Statistics 1990).

Several mtDNA types found in S.A. "Whites" were also found in European Caucasoids (Table 5.6). MtDNA type 1 is the commonest type found at frequencies of 63.9% in English speakers, 80.4% in Afrikaners and 61.6% in Ashkenazi Jews. The lower frequency observed in the Ashkenazi Jewish population may be attributable to the higher frequency of mtDNA type 6 in Jewish populations (30.9% in the present study and 36.0% in Israeli Jews). Type 6 has been found at varying frequencies in other Caucasoids (Table 5.6). Types 44 and 47 are found in both the English and Afrikaans speaking "White" ethnic groups but are not found in the Ashkenazi Jewish group (Table 5.6).

Differences between Jewish and non-Jewish ethnic groups have been observed in a number of other genetic studies (Mourant *et al.* 1978; Lane *et al.* 1985). Jewish populations have had a particularly eventful history and the fact that they have remained a well-defined group with very little outbreeding has produced a gene pool which in some ways differs significantly from those of their surrounding populations. Inherited metabolic conditions such as Tay-Sachs disease, Gaucher's disease, and Niemann-Pick disease, for example, occur at high frequencies in the S.A. Ashkenazi Jewish population (Lane *et al.* 1985). Since

Ashkenazi Jews in different parts of the world have similar frequencies of these metabolic disorders, it has been suggested that the main features of the Ashkenazi gene pool can be traced from their mideastern origins to Eastern Europe before the large scale migrations to the USA, Canada, Israel and S.A. took place (Lane *et al*. 1985). This may account for the similarities in frequencies of mtDNA types 1 and 6 in S.A. Ashkenazi Jews and the Israeli Jews (Bònne-Tamir *et al*. 1986).

In general, the lower mtDNA diversity in Caucasoids worldwide, suggests that Caucasoids diverged more recently from the common female mtDNA ancestor.

Table 5.6. The frequencies (percentages) of mtDNA types found in European Caucasoids compared to South African Caucasoids of European origin.

		European Caucasoids								South African Caucasoids		
		Cauc	Sic	Rom	Sard	Cal	Apul	Fin	I_Jew	A_Jew	Eng	Afrik
		50	90	87	134	60	87	110	38	55	36	51
Type	Mitotype											
1	2-1-1-1-1	58.0	55.6	65.5	75.4	66.7	52.9	79.1	38.5	61.8	63.9	80.4
2	3-1-1-1-3	-	3.3	-	-	-	3.4	-	-	-	-	-
6	2-1-2-1-1	12.0	1.0	4.6	4.5	6.7	8.0	1.8	35.9	30.9	8.3	2.0
7	3-1-1-1-1	-	-	1.1	-	-	-	-	-	-	-	-
11	2-2-3-1-5	6.0	-	1.1	-	-	5.7	3.6	2.6	-	5.6	-
13	2-1-5-1-1	-	-	2.2	1.5	-	-	-	-	-	-	-
15	2-1-1-1-8	2.0	-	-	-	-	-	-	-	-	-	-
16	2-1-2-1-10	2.0	-	-	-	-	-	-	-	-	-	-
17	2-1-1-1-9	2.0	-	-	-	-	-	-	2.6	-	-	-
18	2-3-1-4-9	2.0	12.2	12.6	7.5	6.7	9.2	2.7	-	1.8	-	-
19	2-3-1-4-7	2.0	-	-	-	-	-	-	-	-	-	-
20	2-3-7-4-9	2.0	-	-	-	-	-	-	-	-	-	-
21	2-1-1-1-2	2.0	1.1	3.4	1.4	3.3	4.6	7.3	-	-	-	3.9
22	2-1-1-1-5	2.0	-	1.1	0.7	1.7	3.4	-	1-.2	-	-	-
23	2-2-1-1-3	2.0	1.1	-	-	-	-	-	-	-	-	-
24	2-1-1-4-2	2.0	-	-	-	1.7	1.2	-	-	-	-	-
25	5-1-1-1-1	2.0	-	-	-	-	-	-	-	-	-	-
26	2-1-6-1-1	2.0	-	-	-	-	-	-	-	-	-	-
28	2-1-1-4-1	-	-	-	0.7	1.7	-	-	-	-	2.8	-
34	3-1-2-1-3	-	1.1	-	-	-	-	-	-	-	-	-
36	2-1-1-1-13	-	-	-	-	-	-	-	2.6	-	-	-
37	2-1-1-1-14	-	-	-	-	-	-	-	2.6	-	-	-
38	2-1-1-1-15	-	-	2.2	-	-	1.2	1.8	2.6	-	-	-
39	2-1-4-1-1	-	-	-	1.5	-	-	-	2.6	-	-	-
42	2-1-1-8-1	-	1.1	1.1	-	-	-	-	-	-	-	-
44	2-1-1-4-3	-	-	-	-	-	-	-	-	-	8.3	2.0
47	2-1-1-1-3	-	5.6	1.1	0.7	5	1.2	1.8	-	-	8.3	2.0
51	2-1-1-11-1	-	-	-	-	-	-	-	-	-	-	2.0
56	2-1-1-1-6	-	1.1	-	-	-	-	-	-	-	-	-
57	2-3-1-4-13	-	2.2	1.1	3	1.7	3.4	-	-	-	-	-
58	2-2-3-1-19	-	-	1.1	-	-	-	-	-	-	-	-
59	2-1-1-1-20	-	-	1.1	-	-	-	-	-	-	-	-
60	2-1-1-1-21	-	-	-	0.7	-	-	-	-	-	-	-
61	2-3-1-4-22	-	-	-	0.7	-	-	-	-	-	-	-
62	2-1-11-1-1	-	-	-	1.5	-	-	-	-	-	-	-
63	3-1-7-1-3	-	-	-	-	-	-	0.9	-	-	-	-
64	2-1-12-1-1	-	-	-	-	-	-	0.9	-	-	-	
65	7-1-1-1-3	-	-	-	-	-	-	1.8	-	-		
72	2-1-5-1-1	-	1.1	-	-	-	-	-	-	-	-	-
73	2-1-1-1-2	-	1.1	-	-	-	-	-	-	-	-	-
74	2-1-1-4-25	-	-	-	-	1.7	-	-	-	-	-	-
75	2-1-1-3-1	-	1.1	-	-	3.3	-	-	-	-	-	-
76	2-2-12-1-5	-	1.1	-	-	-	-	-	-	-	-	-
77	2-1-1-4-13	-	1.1	-	-	-	-	-	-	-	-	-
114	2-2-2-1-5	-	-	-	-	-	-	-	-	-	2.8	-
126	2-1-1-17-1	-	-	-	-	-	-	-	-	1.8	-	-
127	2-1-2-17-1	-	-	-	-	-	-	-	-	1.8	-	-
128	2-2-3-1-1	-	-	-	-	-	-	-	-	1.8	-	-
141	2-2-1-4-3	-	-	-	-	-	-	-	-	-	-	5.9
142	3-2-1-11-2	-	-	-	-	-	-	-	-	-	-	2.0

Abbreviations: Cauc = Caucasoids, Sic = Sicilian, Rom = Roman, Sard = Sardinian, Cal = Calabrian, Apul = Apulian, Fin = Finnish, I_Jew = Israeli Jewish, A_Jew = Ashkenazi Jewish, Eng = English, Afrik = Afrikaans.

5.4 Conclusions

1. MtDNA RFLP analyses has uncovered certain differences from the theories on the relationship of the Herero, Dama and Ambo populations with other southern African populations sampled based on linguistic, historical and archaeological data. In particular:

   (1) When using mtDNA types 2, 7 and 21, the Herero and Dama have similar frequencies whereas the Ambo resemble southeast Bantu-speaking groups in its frequency. In particular, type 21 is found at frequencies of 32.6% in the Dama and 50.0% in the Herero whereas it is found at a significantly lower frequency in the Ambo (4.5%). In addition, the "Negroid-like" types 2 and 7 found at frequencies of 13.5% and 54.5% respectively in the Ambo, are rarely found in the Herero and Dama groups. These findings suggest that although the Ambo group sampled here are linguistically related to southwestern Bantu-speakers, their mtDNA genes have been derived predominantly from females within the southeast Bantu- speaking group and are therefore genetically very closely related to the southeastern Bantu group.

   (2) Again, using the frequencies of mtDNA types 2, 7 and 21 again, it would appear that the Dama and Herero groups have been derived from the same female founders

whereas all other South African Negroid groups, as well as the Ambo, have been derived from a different female lineage. These differences hint at a possible correlation with archaeological evidence regarding the migration of Bantu-speakers into southern Africa as postulated by Huffman (1989). If this is so, then it suggests that either the Negroid population was differentiated into mtDNA types 2, 7 and 21 in the Cameroon region prior to their exodus to other parts of Africa, including southern Africa, or that those females who migrated into eastern Africa and later southern Africa carried only mtDNA types 2 and 7, whereas those who migrated into central Africa had predominantly type 21 and may have lost types 2 and 7 due to genetic drift. The other interpretation would be that type 21 (which only differs from type 1 by one mutational event involving an *Ava*II site) arose recently in western Bantu-speakers in central Africa. The absence of type 21 in Senegalese Negroids (Scozzari *et al.* 1988) would tend to favour the latter explanation. However, in order to understand the differentiation of mtDNA in Negroid populations in Africa, populations along the hypothesized routes of migration from both eastern and central Africa need to be analyzed for mtDNA variation to check for the presence of these informative mtDNA types.

2. MtDNA types 3 and 4 are found at different frequencies in the San and Nama groups. Whereas type 4 is unique to San populations, mtDNA type 3 is found at a higher frequency in the Nama group (54.3%) than the Sekele group (14.3%). The presence of type 3 in Negroid groups would suggest the extent of gene flow from the Nama into these groups. The high frequency of this type in the Richtersveld "Coloured" group (62.8%) corroborates the historical relationships between Caucasoid males and Khoi females to give rise to this hybrid group. The JHB "Coloured" group, on the other hand, have a much lower frequency of type 3 (23.9%) reinforcing other evidence that the Khoi contributed much less to this group than to the Baster groups (Nurse *et al.* 1985; May and duToit 1986).

3. The South African Caucasoid population (European Caucasoids and Asian Indians), like other Caucasoid groups from other parts of the world, show less variation at the mtDNA level when compared to Negroid and Khoisan populations from Africa. This is another piece of evidence in support of the hypothesis that modern humans originated in Africa before spreading to other parts of the world.

CHAPTER SIX:

EVOLUTION OF HUMAN MTDNA: SOUTHERN AFRICAN EVIDENCE

6.1. Hominid evolution: fossil evidence

Darwin had recognized that "*In each great region of the world the living mammals are closely related to the extinct species of the same region. It is, therefore, probable that Africa was formerly inhabited by extinct apes closely allied to the gorilla and chimpanzee; and as these two species are man's nearest allies, it is somewhat more probable that our progenitors lived on the African continent than elsewhere*" (Darwin 1859).

The first tangible evidence that a two-legged primate existed in the distant past was obtained from the Transvaal region in South Africa. In 1925 Raymond Dart (University of the Witwatersrand, Johannesburg) announced that he had found the "missing link" between ape and human evolution with his discovery of the fossil skull from Taung, which he dubbed the Taung child, since the skull probably belonged to a five or six years old child. Dart named this specimen *Australopithecus africanus* (from *australis*, southern; and *pithekos*, ape) - African southern ape (Dart 1959).

Actually, it was also the first link in a long chain of fossil discoveries belonging to *Australopithecus* from other parts of South Africa (Sterkfontein, Kromdraai, Swartkrans

and Makapangsgat) and East Africa (Hadar, Laetoli, Kanapoi, Lothagam, Omo, Olduvai, Koobi Flora and Peninj) (see Table 6.1).

There is evidence from the archaeological record in the form of stone flakes, concentration of stone artifacts, broken bones representing food refuse and mixtures of stones and bones to suggest that a behaviourial component in hominids had developed about 2 myr ago (Isaac 1971). At approximately the same time a new hominid species *Homo habilis* appeared both in East Africa (Leakey *et al.* 1964) and South Africa (Jolly and Plog 1987) (see Table 6.1).

Table 6.1 The estimated age, sites at which discovered and the cranial capacities of *Australopithecus* and *Homo* fossils discovered in Africa.

	Fossils				
	A. afarensis	*A. Africanus*	*A. robustus*	*A. bosei*	*H. habilis*
Dates	4-3 mya	3-2.5 mya	2.2-1.5 mya	2.2-1 mya	2.2-1.6 mya
Sites	Hadar Omo Laetoli	Taung Sterkfontein Makapangsgat	Kromdraai Swartkranz	Olduvai Lake Turkana Omo	Olduvai Lake Turkana Omo Sterkfontein (?) Swartkrans (?)
Cranial capacity	380-500 cc	435-530 cc	520 cc (one specimen)	500-530 cc	500-800 cc

(?) query
cc cubic centimetres

More evolved hominids appeared in Africa before 1.5 myr ago and have been called *H. erectus*. There is evidence to suggest that *H. erectus* existed for at least 1 myr and some of their adaptations include the use of stone tools, hand axes and fires.

Fossils dated from between 300 000 to 200 000 years BP have been described as archaic *H. sapiens* have been found in Africa, Europe and Asia. By the beginning of the late Pleistocene, marked around 130 000 years by an abrupt change from full glacial to full interglacial conditions (Roberts 1984), it is clear that major changes had occurred. Northern latitudes of Eurasia were the home of Neanderthals, a large-brained *Homo* species with very characteristic cranial and postcranial morphology (Stringer 1984; Stringer and Andrews 1988; Trinkaus 1986). They were best known between 70 000 - 30 000 years ago in western Europe, although Neanderthal features are present in some European hominids older than 130 000 years, presumably ancestors of the former populations. There is no evidence from the fossil record to suggest that Neanderthals existed after about 34 000 years ago.

Fossils reminescent of anatomically modern humans were discovered in sub-Saharan Africa and dated to the last interglacial period between 130 000 to 75 000 years ago, with at least some fossils being older than 100 000 years in South Africa. In Ethiopia they may be as old as 125 000

years while in Israel they may be as old as 100 000 years. The first fossils bearing the closest resemblance to modern humans were discovered at a site in France and referred to as *Cro-magnon*. They were subsequently found at different places on the earth at different times, the earliest date ascribed to their emergence is about 40 000 years BP.

Although there have been several interpretations to explain hominid evolution (Tobias 1991), the present debate centres around the pattern of evolution that led to the development of geographic variants or "racial" difference in anatomically modern humans. In particular, did these differences develop in a single localized ancestral population, or were they the culmination of ancient evolutionary trends which led to the appearance of modern humans throughout the world?

6.2 Evolution of anatomically modern humans

Two conflicting models have been postulated to explain the evolution of anatomically modern humans from the archaics:

6.2.1 Multiregional evolution

The multiregional model predicts that racial features evolved over very long periods in the regions where they are found today, except in regions where comparatively recent migrations have occurred - such as in North America,

where Amerindians, Inuit, Europeans and Africans in turn brought their distinctive features with them during the past 20 000 years (Stringer 1990). According to this model, regional features are considered to have preceded the appearance of the *H. sapiens* morphology and to have been carried over from local *H. erectus* ancestors. Advocates of the multiregional model favour gene flow as well as natural selection as a means of maintaining similarities among human populations evolving on different continents.

The appearance of *H. sapiens* was thus primarily the result of a continuation of long-term trends in human evolution which took place approximately 1 million years ago (Wolpoff 1984; 1989), and which has occurred mainly through the re-sorting of the same genetic material under the effects of selection rather than by the evolution and radiation of novel genetic material and morphologies (Stringer and Andrews 1988).

6.2.2 Replacement evolution

In contrast, the replacement model (also referred to as the "single origin" and "African origin" models) proposes that there was a relatively recent common ancestral population for *H. sapiens* (approximately 100 000 years ago) which was racially undifferentiated that evolved in Africa and then spread out to other regions of the world, developing racial

features in the process (Cann *et al.* 1987; Stringer and Andrews 1988; Stringer 1990; Wilson and Cann 1992). This hypothesis implies that archaic lineages, such as the European Neantherthals and East Asian Solo Man, mixed little or not at all with early modern humans but were replaced by them (Stringer 1990). This suggests that archaic and anatomically modern humans, were at least for a time, contemporaries and competed for the same territory and resources before the archaic groups became extinct. It is of importance, therefore, to ascertain what was the impetus that caused archaic human to become extinct.

6.2.3 Testing the two models of evolution

Table 6.2 summarizes the paleontological and genetic evidence needed to support the two models for the evolution of modern humans. To test the validity of these models, the fossil evidence from Europe, Asia and Africa was assessed.

Table 6.2. Theoretical predictions from models of *Homo sapiens* evolution (from Stringer and Andrews 1988).

Aspect	Multiregional evolution	Replacement evolution
Geographic patterning of human evolution	Continuity of pattern from middle Pleistocene to present	Continuity of pattern only from late Pleistocene appearance of Homo sapiens to present
	Interpopulation differences are high, greatest between each peripheral area	Interpopulation differences relatively low, greatest between African and non-African populations
	Intrapopulation variation greatest at center of human range	Intrapopulation variations greatest in African populations
Regional continuity and the establishment of Homo sapiens	Transitional fossils widespread	Transitional fossils restricted to Africa, population replacement elsewhere
	Modern regional characters of high antiquity at peripheries	Modern regional characters of low antiquity at peripheries (except Africa)
	No consistent temporal pattern of appearance of Homo sapiens characters between areas	Phased establishment of H. sapiens suite of characters: Africa, S.W. Asia and other areas
Selective and behavioral factors invloved in the origin of Homo sapiens	Factors varied and widespread, perhaps related to technology; local behavioral continuity expected	Factors special and localized in Africa; behavioral discontinuities expected outside Africa

The are no fossils in Europe that unequivocally exhibit an appearance intermediate between the Neanderthals or the archaics and modern human populations (Smith 1984; Stringer *et al*. 1984). This suggests that anatomically modern humans did not evolve from the European Neanderthals as might be expected under the multiregional model. Fossils identified as Neanderthal have been found at Saint-Cesaire in France (Stringer *et al*. 1984) and at Hahnoferstand in Germany both dated to about 35 000 years ago, discoveries which contradict the multiregional hypothesis, for it is unlikely that the Neanderthals were ancestral to modern humans if they survived after the appearance of anatomically modern humans in the same place.

One piece of evidence used by the supporters of the multiregional hypothesis is the apparent continuity through time of specific regionally distributed physical traits, particularly in Asia. Examples of *H. erectus*, archaic *H. sapiens* and modern humans in the Far East share certain skeletal characteristics that are either rare of absent in populations from Europe and Africa (Wolpoff 1984). This evidence, it is claimed, shows that in the evolution of anatomically modern humans, local populations did not become extinct nor were they replaced by groups from elsewhere. Instead, local groups of archaic humans each evolved into modern humans more or less independently with local "racial" traits preserved over hundreds of thousands of years.

The recent discovery of two human crania in Middle Pleistocene deposits in China show a mixture of features associated with both *H. erectus* and with archaic *H. sapiens* (Tianyuan and Etler 1992), suggesting that the emergence of modern humans was not restricted to one region of the world alone, providing another piece of evidence to support the multiregional hypothesis.

The key difference between the multiregional hypothesis and the replacement hypothesis is that the multiregional hypothesis predicts continuity on all continents while the African replacement model predicts continuity only in Africa. Consequently, transitional fossils should, according to the multiregional model, be widely distributed while the replacement model predicts that transitional fossils should be restricted to Africa.

The patterns of diversity and variation found in the fossils are inconsistent with the multiregional model. The oldest examples of modern human fossils are found in southern Africa and the Middle East, supporting the African origin scenario (Stringer and Andrews 1988; Stringer *et al.* 1984). Further evidence is provided by the fossil evidence showing clear evidence of replacement in Western Europe (Stringer and Andrews 1988). The incompleteness of the fossil record in the Far East complicates claims for the local continuity in that area. In addition, there is disagreement among workers as to whether or not those

characters in some of the fossil specimens from Asia claimed to display continuity are merely primitive (Brauer 1989).

However, fossil specimens from Africa have morphological features intermediate between "archaic" and "modern" such as those found at Florisbad, Jebel Irhoud, Omo and Ngaloba as well as those which approximate the "modern anatomical" pattern (Stringer 1989). "Anatomically modern" fossils include those found at Klasies River Mouth in South Africa estimated to be approximately 70 000 years old, perhaps as much as 125 000 years old (Singer and Wymer 1982); the Border Cave remains are at least 50 000 years old and may be 130 000 years old (Beaumont *et al.* 1978; Rightmire 1979); and the Omo I fossil which has been more firmly dated to at least 110 000 years BP. These pieces of evidence corroborate the existence of transitional fossils in Africa.

Using the arguments presented above, there is more support for the replacement model over the multiregional model for explaining the evolutionary history of modern humans. In addition, the available paleontological evidence appears to support the view that modern humans probably originated i sub-Saharan Africa (Stringer and Andrew 1988).

6.3 Evolution of modern humans: genetic evidence

Despite the controversies over the paleontological evidence, there is increasing support from genetic data to support the hypothesis that modern humans evolved in Africa. When blood group markers were used, extant human populations were grouped into Euroafrican and Asian divisions, with the greatest diversity found in Asian populations (Cavalli-Sforza and Edwards 1964). This suggested that modern humans may have evolved in Asia. On the other hand, when protein and HLA markers were used, the major division was between African and Eurasian populations, with African populations revealing the highest amount of genetic diversity (Nei and Roychoudhury 1974; 1982). This observation suggested that Africa was the place of origin of modern human. More recently the application of several polymorphic nuclear DNA markers have reinforced this division of human populations into African and non-African populations, with African populations still showing greater genetic diversity than their non-African counterparts (Cavalli-Sforza *et al.* 1988; Bowcock *et al.* 1991).

More debate on this issue was introduced by the demonstration that all mtDNA lineages can be traced back to a common female ancestor who lived in Africa approximately 200 000 years ago (Cann *et al.* 1987; Vigilant *et al.* 1991;

Wolpoff 1989; Thorpe and Wolpoff 1992)).

To test the validity of this claim, the author has focused her attention on unravelling the prehistory of the peoples of southern Africa. The hypothesis is based on the scenario that since fossils reminiscent of anatomically modern humans have been discovered at Border Cave and Klasies River Mouth in South Africa (Rightmire 1984), it is possible that the indigenous populations of southern Africa may contain within their genomes evidence of molecular events characterizing the evolutionary history of modern humans.

6.4 Southern African evidence

One of the objectives of this study was to use the mtDNA variation found in southern African populations to test the two models, multiregional and single-replacement, proposed for the origins of modern humans. The criteria used in favour of each model were summarized in Table 6.2.

This study provides additional evidence to support the single-replacement theory which claims that the mtDNA found in modern humans evolved in Africa (Cann *et al.* 1987; Vigilant *et al.* 1991). There are two pieces of evidence to support this theory. Firstly, both the mtDNA RFLP data (Tables 3.9 and 3.11) and the control region sequence data (Table 4.5) reveal that African populations have a greater

genetic diversity than non-African populations. This observation is linked to the scenario that the greatest amount of genetic variation would be expected in the oldest populations, provided that the mtDNA mutates at a constant rate. This finding is consistent with that observed by Cann *et al.* (1987) and Vigilant *et al.* (1991).

Secondly, several inferences on human evolutionary history may be obtained from phylogenetic trees. Cann *et al.* (1987) and Vigilant *et al.* (1991) used parsimony trees to show the genetic relationship of human mtDNA types and found that the mtDNA types discovered in African populations were more commonly located in the deepest branches of the tree, whereas Asian and Caucasoid mtDNA types were mostly found in shallower branches of the tree. This was interpreted as evidence for the African origin of human mtDNA. However, there has been increasing criticism of the use of parsimony analysis when analysing mtDNA data (Maddison 1991, 1992; Hedges 1992; Templeton 1992). In particular, reanalysis of the sequence data obtained by Vigilant *et al.* (1991) revealed that the tree published by these authors was not the most parsimonious (Maddison 1992; Templeton 1992). Other problems associated with parsimony analysis, as discussed by Maddison (1991) and Wilson *et al.* (1991), have questioned the interpretation of mtDNA data analyzed by this method.

Having used both parsimony and neighbour-joining analysis of mtDNA sequence data in this study (Chapter 4), there appears to be good concordance between the findings of these two methods of analysis. Both the parsimony (Fig. 4.3) and the neighbour-joining (Fig. 4.4) trees generated from the same data set revealed that mtDNA types found in African populations are more concentrated in the deeper branches of the tree while Caucasoid and Asian mtDNA types are commoner in the shallower branches. In addition, when mtDNA types obtained from RFLP analysis were analyzed using the UPGMA method, the mtDNA types found in African populations were also found to be in the deepest branches of the tree (Fig. 3.8b). These findings are consistent with the single-replacement model concerning the origin of modern humans.

As discussed in section 4.6, the mtDNA data obtained in the present study cannot be used to calibrate the rate of mtDNA mutations since southern Africa, as well as other regions of Africa, do not satisfy the criteria of a "defined geographic region" whose time of "colonization" is known and where back migrations do not occur. However, estimates of the divergence times of southern African populations from the hypothesised ancestor can be made using the mutation rates calibrated in Papua New Guineans (Stoneking *et al.* 1986; 1992). For RFLP data the estimated rate of 2 to 4 percent per million years was used (Brown *et al.* 1979; Cann *et al.* 1987) and for control region sequence data, the

rate of 11.8±3.11% obtained by Stoneking *et al.* (1992a) was used. The time of divergence of the Negroid, Khoisan and Caucasoid populations were calculated from the average amount of sequence divergence estimated in Table 3.9 (RFLP data) and Table 4.5 (sequence data) and are summarized in Table 6.3.

Table 6.3. Estimated divergence times for the Negroid, Khoisan and Caucasoid populations from the average amount of sequence divergence as estimated from mtDNA RFLP and sequence data.

POPULATION	N		SEQUENCE DIVERGENCE (%)		DIVERGENCE TIMES (BP)
RFLP data[1]					
Negroid	392		0.208		104 000 - 50 000
Khoisan	95		0.256		128 000 - 64 000
Caucasoid	192		0.096		48 000 - 24 000
Sequence data[2]					
Negroid		78		1.75	148 588± 83 857
"Bantu"	45		2.13		180 508±102 067
Herero	12		1.06		89 831± 50 794
Dama	21		2.07		175 424± 99 191
Khoisan		56		1.98	167 796± 94 878
San (Sekele)	38		1.81		153 390± 86 732
Nama	18		2.15		182 203±103 024
Caucasoid	7		1.46		123 729± 69 961
Pairwise differences[3]					
Negroid		78			52 800
"Bantu"	45				89 200
Herero	12				23 500
Dama	21				45 700
Khoisan		56			58 700
San (Sekele)	38				50 500
Nama	18				66 900

[1]Based on a mutation rate of 2 to 4 percent per million years
[2]Based on a mutation rate of 11.8±3.11percent per million years
[3]Based on pairwise differences of mtDNA types according to the Rogers and Harpending model (see Table 4.8).

There is a general lack of concordance between the estimates of divergence times calculated from RFLP data and sequence data in the populations listed in Table 6.1. Although there are appreciable differences in divergence time estimates within Negroid and Khoisan groups, the results obtained by pooling the samples of the Negroid, the Khoisan and the Caucasoid populations suggest that the Khoisan population diverged from the hypothesized common ancestor before the Negroid populations.

There are several factors which may contribute to the discrepancies in the estimated divergence times: Firstly, the differences in the resolving power of the two techniques: there are a fewer number of variant sites (31 out of a total of 233 nucleotide bases) analyzed when using the RFLP method compared with 119 out of a total of 750 when analysing sequence data. Secondly, mutations are commoner in the non-coding control region than in the coding regions of the molecule; this would have the effect of over-estimating the divergence times for sequence data compared with the data obtained from the entire molecule using restriction enzyme analysis, producing the observed increase in divergence time as noted in Table 6.3.

Whereas estimates of divergence times derived from mtDNA sequence variation may be used to estimate the age of the common ancestor, the parameters of the Rogers and Harpending (1992) model estimates the time of expansion of

populations. The age of the common ancestor need not necessarily coincide with the expansion date. The human population might well have persisted as a relatively small population for some time and then expanded rapidly. Khoisan populations appear to have undergone a recent expansion approximately 58 700 years ago compared with 52 800 years ago for Negroid populations. The mean expansion time calculated from a worldsample of 357 individuals was estimated to be approximately 70 000 years ago (Sherry *et al*. 1992). This analysis would suggest that modern humans have expanded perhaps from a small founding population only fairly recently.

Although it would be desirable to obtain estimates of statistical significance for the data obtained in this study using methods (for example, bootstrapping) for estimating confidence limits on the phylogenetic trees presented, such tests could not be applied to the data due to computational limitations. However, such tests will be performed once the hardware and software computer resources become available. These tests would increase the robustness of the analysis and would also add significance to the data. Acknowledging that there is a lack of statistical support in the interpretation of the data presented in this study, the findings of this study tend to support the hypothesis that the mtDNA in modern humans evolved in Africa and, therefore, lends further support to the replacement model for the origins of modern humans.

REFERENCES

Albertini, A.M., Hofer, M., Calos, M.P. and Miller, J.H. (1982). On the formation of spontaneous deletions: The importance of short sequence homologies in the generation of large deletions. *Cell* 29: 319-328.

Anderson, S., Bankier, A.T., Barrel, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe. B,A., Sanger, F., Schreier, H., Smith, A.J.H., Staden, R. and Young, I.G. (1981). Sequence and organization of the mitochondrial genome. *Nature* 290: 457-465.

Attardi, G. (1985). Animal mitochondrial DNA: An extreme example of genetic economy. In *International Review of Cytology, Vol. 93: Genome Evolution in Prokaryotes and Eukaryotes*, D.C. Reanney and P. Chambon (eds), Academic Press, New York, pp 93-145.

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

Baumann, H. (1935). *Luanda: Bei Bauern und Jagern in Juner-Angola*. Wurfel Verlag, Berlin.

Beach, D.N. (1980). The Shona and Zimbabwe 900-1850. London, pp 260-263.

Beaumont, P., de Villiers, H. and Vogel, J.C. (1978). Modern man in sub-Sarahan Africa prior to 49 000 years BP. A review and evalution with particular reference to Border Cave. *South African Journal of Science* 74: 409-419.

Bellwood, P.S. (1989). The colonization of the Pacific: some current hypotheses. In *The colonization of the Pacific - a genetic trail*, A.V.S. Hill and S.W. Serjeantson (eds), Oxford University Press, Oxford, pp 1-59.

Bernstein, R., Wadee, A., Rosendorff, J., Wessels, A. and Jenkins, T. (1986). Inverted Y chromosome polymorphism in the Gujerati Muslim Indian population of South Africa. *Human Genetics* 74: 223-229.

Bhana, S. and Brain, J.B. (1990). *Setting down roots. Indian migrants in South Africa, 1860-1911*. Witwatersrand University Press, Johannesburg.

Blanc, H., Chen, K.-H., D'Amore, M.A. and Wallace, D.C. (1983). Amino acid change associated with the major polymorphic *Hinc*II site of Oriental and Caucasian mitochondrial DNAs. *American Journal of Human Genetics* 35: 167-176.

Bleek, D. (1927). The Distribution of Bushman languages in South Africa. In *Festschrift Meinhof*, pp55-56, cited in *The Khoisan Peoples of South Africa*, I. Schapera (ed), Routledge and Kegan Paul, London, pp 419-20.

Bonne-Tamir, B., Johnson, M.J., Natali, A., Wallace, D.C. and Cavalli-Sforza, L.L. (1986). Human mitochondrial DNA types in two Israeli populations - A comparative study at the DNA level. *American Journal of Human Genetics* 38: 341-351.

Botha, M.C. and Pritchard, J. (1972). Blood group gene frequencies: An indication of the genetic constitution of population samples in Cape Town. *South African Medical Journal* suppl: 1-26.

Bowcock, A.M., Kidd, J.R., Mountain, J.L., Herbert, J.M., Carotenuto, L, Kidd, K.K. and Cavalli-Sforza, L.L. (1991). Drift, admixture, and selection in human evolution: A study with DNA polymorphisms. *Proceedings of the National Academy of Science, USA*. 88: 839-843.

Brauer, G. (1989). The evolution of modern humans: a comparison of African and non-African evidence. In *The Human Revolution,* P. Mellars and C. Stringer (eds), Edinburgh University Press, Edinburgh, pp 123-154.

Brega, A., Gardella, R., Semino, O., Morpurgo, G., Astaldi-Ricotti, G.B., Wallace, D.C. and Santachiara-Benerecetti, A.S. (1986a). Genetic studies on the Tharu population of Nepal: restriction endonuclease polymorphisms of mitochondrial DNA. *American Journal of Human Genetics* 39: 502-512.

Brega, A., Scozzari, R., Maccioni, L., Iodice, C., Wallace, D.C., Bianco, I., Cao, A. and Santachiara-Benerecetti, A.S. (1986b). Mitochondrial DNA polymorphisms in Italy. I. Population data from Sardinia and Rome. *Annals of Human Genetics* 50: 327-338.

Bridgeman, J. (1981). *The Revolt of the Herero*. University of California Press, Berkeley.

Brown, W. M. (1980). Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proceedings of the National Academy of Science, USA*. 77: 3605-3609.

Brown, W.M., George, M. and Wilson, A.C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Science, USA*. 76: 1967-1971.

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

Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *Journal of Molecular Evolution* 18: 225-239.

Bryant, A.T. (1929). *Olden times in Zululand and Natal.* Longman, London.

Cann, R.L. (1982). *The evolution of human mitochondrial DNA.* PhD Thesis, University of California, Berkeley.

Cann, R.L. and Wilson, A.C. (1983). Length mutations in human mitochondrial DNA. *Genetics* 104: 699-711.

Cann, R.L., Brown, W.M. and Wilson, A.C. (1984). Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106: 479-499.

Cann, R.L., Stoneking, M. and Wilson, A.C. (1987). Mitochondrial DNA and human evolution. *Nature* 325: 31-36.

Cavalli-Sforza, L.L. and Edwards, A.W.F. (1964). Analysis of human evolution. Genetics today. *Proceedings of the 11th International Congress for Genetics* 2: 923-933.

Cavalli-Sforza, L.L., Piazza, A., Menozzi, P. and Mountain, J. (1988). Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Science, USA.* 85: 6002-6006.

Cavalli-Sforza, L.L., Minch, E. and Mountain, J.L. (1992). Coevolution of genes and languages revisited. *Proceedings of the National Academy of Sciences, USA.* 89: 5620-5624.

Chang, D.D. and Clayton, D.A. (1984). Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. *Cell* 36: 635-643.

Chomyn, A., Cleeter, W.A., Ragan, C.I., Riley, M., Doolittle, R.F. and Attardi, G. (1986). URF6, last unidentified reading frame of mtDNA codes for an NADH dehydrogenase subunit. *Science* 234: 614-617.

Chomyn, A., Marrioniti P, Cleeter, M., Ragan, F., Marsuno-Yagi, A., Hatefi, Y., Doolittle, R.F. and Attardi, G. (1985). Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory chain NADH dehydrogenase. *Nature* 314: 592-594.

Clark, J.D. (1959). *The prehistory of Southern Africa.* Penguin Books, Harmondsworth.

Clayton, D.A. (1982). Replication of animal mitochondrial DNA. *Cell* 28: 693-705.

Colenbrander, H.T. (1902). *De Afkomst der Boeren*. Het algemeen Nederlandsch Verbond, Kaapstad.

Crow, J.F. and Aoki, K. (1984). Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. *Proceedings of the National Academy of Science, USA*. 81: 6073-6077.

Dart, R. (1959). *Adventures with the Missing Link*. Viking Press.

Darwin, C.R. (1859). *On the Origin of Species by means of Natural Selection* (1st ed.). Murray, London.

Darwin, C.R. (1898). *On the Origin of Species by means of Natural Selection* (6th ed.). Appleton and Co., New York.

Davidson, I. (1991). The archaeology of language origins - a review. *Antiquity* 65: 39-48.

Deacon, H.J. (1976). *Where hunters gathered: a study of Holocene Stone Age people in the eastern Cape*. South African Archaeological Society Monograph 1, Claremont.

de Almeida, A. (1965). *Bushmen and other Non-Bantu peoples of Angola*. Witwatersrand University Press for the Institute for the Study of Man in Africa.

De Benedictis, G., Rose, G., Passarino, G. and Quagliariello, C. (1989a). Restriction fragment length polymorphism of human mitochondrial DNA in a sample population from Apulia (Southern Italy). *Annals of Human Genetics* 53: 311-318.

De Benedictis, G., Rose, G., Picardi, P., and Quagliariello, C. (1989b). Mitochondrial DNA polymorphism in Calabria (Southern Italy). *Gene Geography* 3: 33-40.

De Maret, P. (1982). The Iron Age in the West and South. In *The Archaeology of Central Africa*, F. von Noten (ed), Graz: Akademische Druk-u. Verlagsanstalt, pp 77-96.

De Maret, P. (1985). The Smiths Myth and the Origin of Leadership in Central Africa. In *African Iron Working: Ancient and Traditional*, R. Haaland and P. Shinnie (eds), Norwegian University Press, Oslo, pp 73-87.

Denaro, M., Blanc, H., Johnson, M.J., Chen, K.H., Wilmsen, E., Cavalli-Sforza, L.L. and Wallace, D.C. (1981). Ethnic variation in *Hpa*I endonuclease cleavage patterns of human mitochondrial DNA. *Proceedings of the National Academy of Science, USA*. 78: 5768-5772.

Di Renzo, A. and Wilson, A.C. (1991). Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proceedings of the National Academy of Science, USA*. 88: 1597-1601.

Doke, C.M. (1954). *The southern Bantu languages*. Oxford University Press for International African Institute, London.

Douglas, M. (1963). *The Lele of Kasai*. Oxford University Press for the International African Institute, London.

Dunn, S.D. (1990). *The genetic affinities of some indigenous southern African populations*. MSc thesis, University of the Witwatersrand, Johannesburg.

Eggert, M. and Kanimba, M. (1980). Aspects d'un metier traditionnel: L'industrie poterie a Ikenge. *Baessler-Archiv, Neue Folge*, 28: 387-340.

Ehret, C. (1973). Patterns of Bantu and Central Sudanic settlement in central and southern Africa. *Transafrican Journal of History* 3: 1-71.

Ehret, C. (1982a). The first spread of food production to southern Africa. In *The Archaeological and Linguistic Reconstruction of African History*, C. Ehret and M. Posnansky (eds), University of California Press, Berkeley, pp 158-181.

Ehret, C. (1982b). Linguistic inferences about early Bantu history. In *The Archaeological and Linguistic Reconstruction of African History*, C. Ehret and M. Posnansky (eds), University of California Press, Berkeley, pp 57-65.

Evers, T.M. (1982). Excavations at the Lydenburg Heads site, eastern Transvaal, South Africa. *South African Archaeological Bulletin* 37: 16-33.

Ewens, W.J. (1972). The sampling theory of selectively neutral alleles. *Theoretical Population Bioliogy* 3: 87-112.

Excoffier, L. (1990). Evolution of human mitochondrial DNA: Evidence for departure from a pure neutral model of populations at equilibrium. *Journal of Molecular Evolution* 30: 125-139.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction fragments to high specific activity. *Annals of Biochemistry* 1327: 266-267.

Felsenstein, J. (1988). Phylogenies from molecular sequences: Inference and reliability. *Annual Review of Genetics* 22: 521-565.

Ferris, S.D., Brown, W.M., Davidson, W.S. and Wilson, A.C. (1981). Extensive polymorphism in the mitochondrial DNA of apes. *Proceedings of the National Academy of Science, USA.* 78: 6319-6323.

Foran, D.R., Hixson, J.E. and Brown, W.M. (1988). Comparisons of ape and human sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. *Nucleic Acids Research* 16: 5841-5861.

Giles, R.E., Blanc, H., Cann, H.M. and Wallace, D.C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the Natational Academy Science, USA.* 77: 6715-6719.

Grierson, G.A. (1927). *1. The Dravidian family. 2. The Indo-European Family, the Aryan sub-family. Linguistic Survey of India 1*, Motilal Banarsidass, Delhi.

Greenberg, J.H. (1963). *The Languages of Africa*. Indiana University Press, Bloomington.

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

Groube, L.M., Chappell, J., Mike, J. and Price, D. (1986). A 40,000 year-old human occupation site at Huon Peninsula, Papua New Guinea. *Nature* 324: 453-455.

Guthrie, M. (1962). Some developments in the prehistory of the Bantu origins. *Journal of African History* 3: 273-282.

Hahn, C.H.L. (1928). The Ovambo. In: *The Native Tribes of South West Africa,* C.H.L. Hahn (ed), Cape Times Ltd, Cape Town.

Harihara,S., Hirai, M. and Omoto, K. (1986). Mitochondrial DNA polymorphisms in Japanese living in Hokkaido. *Japanese Journal of Human Genetics* 31: 73-83.

Harihara, S., Saitou, N., Hirai, M., Gojobori, T., Park, K.S., Misawa, S., Ellepola, S.B., Ishada, Y. and Omoto, K. (1988). Mitochondrial DNA polymorphisms among five Asian populations. *American Journal of Human Genetics* 43: 134-143.

Hasegawa, M., Kishino, H. and Yano, A. (1989). Estimation of branching dates among primates by molecular clocks of nuclear DNA which slowed down in Hominoidea. *Journal of Human Evolution* 18: 461-476.

Hedges, S.B., Kumar, S., Tamura, K. and Stoneking, M. (1992). Human origins and the analysis of mitochondrial DNA sequences. *Science* 255: 737-739.

Hixson, J.E. and Clayton, D.A. (1985). Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. *Proceedings of the National Academy of Science, USA*. 82: 2660-2664.

Horai, S. and Hayasaka, K. (1990). Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *American Journal of Human Genetics* 46: 828-842.

Horai, S. and Matsunaga, E. (1986). Mitochondrial DNA polymorphism in Japanese. *Human Genetics* 72: 105-117.

Horai, S., Gojobori, T. and Matsunaga, E. (1984). Mitochondrial DNA polymorphism in Japanese. I. Analysis with restriction enzymes of six base pair recognition. *Human Genetics* 68: 324-332.

Horai, S., Gojobori, T. and Matsunaga, E. (1986). Distinct clustering of mitochondrial DNA types among Japanese, Caucasians and Negroids. *Japanese Journal of Genetics* 61: 271-275.

Horai, S., Kondo, R., Murayama, K., Hayashi, S., Koike, H, and Nakai, N. (1991). Phylogenetic affiliation of ancient and contempory humans inferred from mitochondrial DNA. *Philosophical Transactions of the Royal Societ of London* 333: 409-417.

Huffman, T.N. (1980). Ceramics, classification and Iron Age entities. *African Studies* 39: 123-174.

Huffman, T.N. (1982). Archaeology and ethnohistory of the African Iron Age. *Annual Review of Archaeology* 11: 133-150.

Huffman, T.N. (1989). Ceramics, settlements and Late Iron Age migrations. *African Archaeological Reviews* 7: 155-182.

Isaac, G.L. (1971). The diet of early man: aspects of archaeological evidence from Lower and Middle Pleistocene sites in Africa. *World Archaeology* 2: 278-299.

Jenkins, T. (1972). *Genetic polymorphisms of Man in Southern Africa*. MD Thesis, London.

Jenkins, T. (1982). Human Evolution on Southern Africa. In *Human Genetics. Part A: The Unfolding Genome*, B. Bonne-Tamir (ed.), Alan R. Liss, New York, pp 227-253.

Jenkins, T. (1986). The prehistory of the San and Khoikhoi as recorded in their blood. In *Contemporary studies on Khoisan 2*, R. Vossen and K. Keuthmann (eds), Helmut Buske Verlag, Hamburg, pp 51-77.

Jenkins, T. and Tobias, P.V. (1977). Nomenclature of population groups in Southern Africa. *African studies* 36: 49-55.

Jenkins, T., Harpending, H.C., Gordon, H., Keraan, M.M., Johnston, S. (1971). Red cell enzyme polymorphisms in Khoisan peoples of southern Africa. *American Journal of Human Genetics* 23: 513-532.

Jenkins, T., Zoutendyk, A., and Steinberg, A.G. (1970). Gammaglobulin groups (Gm and Inv) of various southern African populations. *American Journal of Physical Anthropology* 32: 197-218.

Jenny, H. (1966). *Sudwestafrika, Land zwischen den Extremen*. Kohlhammer Verlag, Stuttgart.

Johnson, M.J., Wallace, D.C., Ferris, S.D., Rattazzi, M.C. and Cavalli-Sforza, L.L. (1983). Radiation of human mitochondria DNA types analyzed by restriction endonuclease cleavage patterns. *Journal of Molecular Evolution* 19: 255-271.

Jolly, C.J. and Plog, F. (1987). An Evolutionary Perspective on Human Species. In *Physical Anthropology and Archaeology (4th ed.)*, Alfred A. Knopf (ed), New York, pp 3-25.

King, T.C. and Low, R.L. (1987). Mapping of control elements in the displacement loop region of bovine mitochondrial DNA. *Journal of Biological Chemistry* 262: 6204-6213.

Knussman, R., and Knussman, R. (1969/1970). Die Dama-eine Altschicht in Sudwestafrika? *Journal of the South West Africa Scientific Society, Windhoek* 24: 9-32.

Kornberg, A. and Baker, T.A. (1991). *DNA replication*, (2nd ed), W.H. Freeman and Co., New York.

Kuper, A. (1980). Symbolic dimensions of the Southern Bantu homestead. *Africa* 50: 8-23.

Kuper, A. (1982). *Wives for Cattle: Bridewealth and Marriage in Southern Africa*. Routledge and Kegan Paul, London.

Laman, K. (1953). *The Kongo 1*. Studia Ethnographica Upsaliensia 4.

Lane, A.B., Dunn, D.S. And Jenkins, T. (1985). A genetic profile of the South African Ashkenazi Jewish population. *South African Medical Journal* 68: 935-939.

Leakey, L.S.B, Tobias, P.V. and Napier, J.R. (1964). A new species of the genus *Homo* from Olduvai Gorge. *Nature* 202: 7-9.

Lyell, C. (1873). *The Geological Evidences of the Antiquity of Man*. J. Murray, London.

Maddison, D.R. (1991). African origin of human mitochondrial DNA reexamined. *Systematic Zoology* 40: 355-363.

Maddison, W.P. and Maddison, D.R. (1987). MacClade 2.1, Museum of comparative Zoology, Harvard University, Cambridge.

Maddison, D.R., Ruvolo, M. and Swofford, D.L. (1992). Geographic origins of human mitochondrial DNA: Phylogenetic evidence from control region sequences. *Systematic Biology* 33: 543-555.

Maggs, T. (1984). The Iron Age South of the Zambesi. In *Southern African Prehistory and Paleoenvironments*, R.G. Klein (ed), A.A. Balkema, Rotterdam, pp 329-360.

Malan, J.S. (1980). *Peoples of South West Africa/Namibia*. Haum Publishers, Pretoria, Cape Town.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). *Molecular cloning - a laboratory manual*. Cold Spring Harbour laboratory, New York.

Marais, J.S. (1939). *The Cape Coloured People, 1652-1937*. Witwatersrand University Press, Johannesburg, 1968.

Marques, I. (1992). *Variable Number of Tandem Repeat Polymorphism of Man in southern Africa*. PhD thesis, University of the Witwatersrand, Johannesburg.

Mason, R.J. (1986). *Origins of Black People of Johannesburg and the Southern Western Central Transvaal AD 350-1880*, Occasional Paper 16, University of the Witwatersrand Archaeological Research Unit, Johannesburg.

May, R.M. and DuToit, E.D. (1989). Blood group gene frequencies of four population groups in the western Cape. *South African Medical Journal* 76: 647-650.

Maylam, P. (1986). *A History of the African People of SOuth Africa: from the Early Iron Age to the 1970s*. Creda Press, Cape Town.

Meer, F. (1969). *Portrait of Indian South Africans*. Premier Press, Durban.

Merriwether, D.A., Clark, A.G., Ballinger, S.W., Schurr, T.G., Soodyall, H., Jenkins, T., Sherry, S.T., Wallace, D.C. (1991). The structure of human mitochondrial DNA variation. *Journal Molecular Evolution* 33: 543-545.

Moores, P.P. (1980). *The blood groups of the Natal Indians*. PhD thesis, University of Natal, Natal.

Morris, A. (1984). *An osteological analysis of the protohistoric population of the northern Cape and western Orange Free State, South Africa*. PhD Thesis, University of the Witwatersrand, Johannesburg.

Morris, D.J., Heim, R.A., Verga, V., Denter, M., Dunn, D.S. and Jenkins, T. (1991). Study of 30 DNA markers in three Southern African populations. *Gene Geography* 5: 1-12.

Mourant, A.E. (1983). *Blood Relations*. Oxford University Press, Oxford.

Mourant, A.E., Kopec, A.C. and Domaniewska-Sobczak, K. (1978). *The distribution of the human blood groups and other polymorphisms*. Oxford University Press, London.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Science, USA*. 70: 3321-3323.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York, pp 177-253.

Nei, M. and Graur, D. (1984). Extent of protein polymorphism and the neutral mutation theory. *Evolutionary Biology* 17: 73-118.

Nei, M. and Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of Natational Academy of Science, USA*. 76: 5269-5273.

Nei, M. and Livshits, G. (1990). Evolutionary relationships of Europeans, Asians, and Africans at the molecular level. In *Population biology of genes and molecules*, N. Takahata and J. Crow (eds), Baifukan, Tokyo, pp 251-266.

Nei, M. and Miller, J.C. (1990). A simple method for estimating average number of nucleotide substitutions within and between populations from restriction site data. *Genetics* 125: 873-879.

Nei, M. and Roychoudhury, A.K. (1974). Genetic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids. *American Journal of Human Genetics* 26: 421-443.

Nei, M. and Roychoudhury, A.K. (1982). Genetic relationship and evolution of human races. *Evolutionary Biology* 14: 1-59.

Nurse, D. (1982). Bantu expansion into East Africa:linguistic evidence. In *The Archaeological and Linguistic Reconstruction of African History*, C. Ehret and M. Posnansky (eds), Berkeley, pp 199-238.

Nurse, G.T. (1983). Population movement around the northern Kalahari. *African Studies* 76: 153-163.

Nurse, G.T. and Jenkins, T. (1977). Serogenetic studies on the Kavango peoples of South West Africa. *Annals of Human Biology* 4: 465-478.

Nurse, G.T., Botha, M.C. and Jenkins, T. (1977). Sero-genetic studies on the San of South West Africa. *Human Heredity* 27: 81-98.

Nurse, G.T., Lane, A.B. and Jenkins, T. (1976). Sero-genetic studies on the Dama of South West Africa. *Annals of Human Biology* 3: 33-50.

Nurse, G.T., Weiner, J.S. and Jenkins, T. (1985). *The peoples of Southern Africa and their affinities*. Claredon press, Oxford.

Nurse, G.T., Dunn, D.S., Rootman, A.J., and Jenkins, T. (1987). Sero-genetic studies on the Ambo of Namibia. *Gene Geography* 1: 65-79.

Ohta, T. and Kimura, M. (1969). Linkage disequilibruim at steady state determined by random genetic drift and recurrent mutation. *Genetics* 63: 229-238.

Parkington, J. (1980). Time and place: Some observations on spatial and temporal patterning in the Later Stone Age sequence in southern Africa. *South African Archaeological Bulletin* 35: 73-83.

Parkington, J. and Cronin, M. (1979). The size and layout of Mgungundlovu 1829-1838. In *Iron Age Studies in Southern Africa*, N.J. van der Merwe and T.N. Huffman (eds), Goodwin Series 3, South African Archaeological Society, Cape Town, pp 133-148.

Parsons, N. (1982). *A new history of South Africa*. London.

Patel, S.S., Wong, I. and Johnson, K.A. (1991). Presteady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease deficient mutant. *Biochemistry* 30: 511-525.

Pennington, R. (1990). *Fertility of the Herero and Mbanderu pastoralists of Botswana*. PhD thesis, The Pennsylvania State University, Pennsylvania.

Phillipson, D.W. (1977). The spread of the Bantu language. *Scientific American* 236: 106-114.

Phillipson, D.W. (1985). *African Archaeology*. Cambridge University Press, Cambridge.

Ramsay, M. and Jenkins, T. (1987). Globin gene-associated restriction fragment length polymorphisms in southern African peoples. *American Journal of Human Genetics* 41: 1132-1144.

Ramsay, M. and Jenkins, T. (1988). Alpha-globin gene cluster haplotypes in the Kalahari San and southern African Bantu-speaking Blacks. *American Journal of Human Genetics* 43: 527-533.

Race Relations Survey (1990). South African Institute of Race Relations, Johannesburg.

Rightmire, G.P. (1979). Implications of the Border Cave skeletal remains for Later Pleistocene human evolution. *Current Anthropology* 20: 23-35.

Rightmire, G.P. (1984). A World survey of the Fossil Evidence. In *The Origins of Modern Humans*, F.H. Smith and F. Spencer (eds), Liss, New York, pp 295-325.

Roberts, N. (1984). Pleistocene environments in time and space. In *Hominid Evolution and Community Ecology*, R. Foley (eds.), Academic Press, New York, pp 25-53.

Rogers, A.R. and Harpending, H.C. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9: 552-569.

Rudner, J. (1979). The use of stone artefacts and pottery among the Khoisan peoples in historic and protohistoric times. *South African Archaeological Bulletin* 34: 3-17.

Rudner, J. and Grattan-Bellew, P. (1964). Archaeological sites along the southern coast of South West Africa. *South African Journal of Science* 60: 67-79.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R.G., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). Primer-directed enzymatic amplifications of DNA with a thermostable DNA polymerase. *Science* 239: 487-491.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: A new method for constructing phylogenetic trees. *Molecular and Biological Evolution* 4: 406-425.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, USA*. 74: 5463-5467.

Sansom, B. (1974). Traditional economic systems. In *The Bantu-speaking Peoples of Southern Africa*, W.D. Hammond-Tooke (ed), London.

Sanghvi, L.D. (1966). Inbreeding in India. *Eugenics Quarterly* 13: 291-301.

Santachiara-Benerecetti, A.S., Scozzari, R., Semino, O., Torroni, A., Brega, A. and Wallace, D.C. (1988). Mitochondrial DNA polymorphisms in Italy. II. Molecular analysis of new and rare morphs from Sardinia and Rome. *Annals of Human Genetics* 52: 39-56.

Saron, G. (1965). *The making of South African Jewry*, L. Feldberg (ed.), Alex White, Johannesburg, pp 9-18.

Sartoris, S., Varetto, O., Migone, N., Capello, N., Piazza, A., Ferrara, G.B. and Ceppellini, R. (1988). Mitochondrial DNA polymorphism in four Sardinian villages. *Annals of Human Genetics* 52: 327-340.

Schapera, I. (1930). *The Khoisan peoples of South Africa*. Routledge and Kegan Paul, London.

Schoute-Vanneck, C.A. and Walsch, R.C. (1960). The Tongaat variant of the Wilton culture. *South African Archaeological Bulletin* 15: 29-35.

Schultze, L. (1928). *Zur Kenntnis des Körpers der Hottentotten und Buschmänner*. (Jenaische Denkschriften, 17), Gustav Fischer, Jena, pp 147-228.

Schurr, T.G., Ballinger, S.W., Gan, Y.-Y, Hodge, J.A., Merriwether, D.A., Lawrence, L.M., Knowler, W.C., Weiss K.M. and Wallace, D.C. (1990). Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *American Journal of Human Genetics* 46: 613-623.

Scozzari, R., Torroni, A., Semino, O., Sirugo, G., Brega, A. and Santachiara-Benerecetti, A.S. (1988). Genetic studies on the Senegalese population. I. Mitochondrial DNA polymorphisms. *American Journal of Human Genetics* 43: 534-544.

Semino, O., Torroni, A., Scozzari, R., Brega, A., De Benedictis, G. and Santachiara-Benerecetti, A.S. (1989). Mitochondrial DNA polymorphisms in Italy. III. Population data from Sicily: a possible quantitation of maternal African ancestry. *American Journal of Human Genetics* 53: 193-202.

Semino, O., Torroni, A., Scozzari, R., Brega, A., Santachiara Benerecetti, A.S. (1991). Mitochondrial DNA polymorphisms among Hindus: a comparison with the Tharus of Nepal. *Annals of Human Genetics* 55: 123-136.

Sherry, S.T., Rogers, A.R., Harpending, H., Soodyall, H., Jenkins, T. and Stoneking, M. (1992). Pairwise differences of mtDNA reveal recent human population expansions (submitted to *Proceedings of the National Academy of Science, USA*.).

Smith, F.H. (1984). Fossil hominids from the Upper Pleistocene of Central Europe and the origin of modern Europeans. In *The Origins of Modern Humans: a World Survey of the Fossil Evidence*, F.H. Smith and F. Spencer (eds), Alan R. Liss, New York, pp 137-210.

Singer, R. and Wymer, J. (1982). *The Middle Stone Age at Klasies River Mouth in South Africa*. University of Chicago Press, Chicago.

Sober, E. (1985). A likelihood justification of parsimony. *Cladistics* 1: 209-233.

Soodyall, H. and Jenkins, T. (1992a). Mitochondrial DNA polymorphisms in Khoisan populations from southern Africa. *Annals of Human Genetics* 56: 315-324.

Soodyall, H. and Jenkins, T. (1992b). Mitochondrial DNA polymorphisms in Negroid populations from Namibia: New light on the origins of Dama, Herero and Ambo. *Annals of Human Biology* (in press).

Soodyall, H. and Jenkins, T. (1992c). Mitochondrial DNA polymorphisms in the South African Indian population. (submitted to *Gene Geography*).

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98: 503-517.

South African Statistics (1990). Central Statistical Services, The Government Printer, Pretoria.

Spurdle, A.B. (1992). *Y-specific restriction fragment length polymorphisms in southern African populations*. PhD thesis, University of the Witwatersrand, Johannesburg.

Spurdle, A.B and Jenkins, T. (1992a). Y chromosome probe p49a detects comples *Pvu*II haptotypes and many new *Taq*I haptotypes in southern African populations. *American Journal of Human Genetics* 50: 107-125.

Spurdle, A.B. and Jenkins, T (1992b). The Y chromosome as a tool for studying human evolution. *Current Genetics and Development* (in press).

Spurdle, A.B. and Jenkins, T (1992c). The inverted Y chromosome polymorphism in the Gujerati Muslim Indian population of South Africa has a single origin. *Human Heredity* (in press).

Stayt, H.A. (1931). Notes on the Balemba. *Journal of the Royal Anthropological Institute* LXI: 236.

Steinberg, A.G., Jenkins, T., Nurse, G.T., Harpending, H.C. (1975). Gammaglobulin groups of the Khoisan peoples of Southern Africa: Evidence for polymorphism for a $Gm^{1,5,13,14,21}$ haplotype among the San. *American Journal of Human Genetics* 27: 528-542.

Stoneking, M. (1986). *Human mitochondrial DNA evolution in Papua New Guinea*. PhD. thesis, University of California, Berkeley.

Stoneking, M. (1990). Departure of human mitochondrial DNA variation from neutral expectations: an alternative explanation. *Journal of Molecular Evolution* 31: 343-346.

Stoneking, M. and Cann, R.L. (1989). African origin of human mitochondrial DNA. In *The Human Revolution*, P. Mellars and C. Stringer (eds), Edinburgh University Press, Edinburgh, pp 17-30.

Stoneking, M., Bhatia, K. and Wilson, A.C. (1986a). Mitochondrial DNA variation in eastern highlanders of Papua New Guinea. In *Genetic Variation and its Maintenance*, D.F. Roberts and G.F. DeStefano (eds), Cambridge University Press, Cambridge, pp 87-100.

Stoneking, M., Bhatia, K. and Wilson, A.C. (1986b). Rate of sequence divergence estimated from restriction maps of mitochondrial DNAs from Papua New Guinea. In *Cold Spring Harbor Symposia on Quantitative Biology, Volume LI*, Cold Spring Harbor Laboratory, Cold Spring harbor, pp 433-439.

Stoneking, M., Jorde, L.B., Bhatia, K. and Wilson, A.C. (1990). Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics* 124: 717-733.

Stoneking, M., Sherry, T., Redd, A.J. and Vigilant, L. (1992a). New approaches to dating suggest a recent age for the human mtDNA ancestor. *Philosophical Transactions of the Royal Society B* (in press).

Stoneking, M., Sherry, S.T. and Vigilant, L. (1992b). Geographic origin of human mitochondrial DNA revisited. *Systematic Biology* 41: 384-391.

Stow, G.W. (1905). *The Native races of South Africa*. London: Swann Sonnenschein.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966). Frameshift mutations and the genetic code. In *Cold Spring harbor Symposia on Quantitative Biology, Volume XXXI*, Cold Spring Harbor Laboratory of Quantitative Biology, Cold Spring Harbor, pp 77-84.

Stringer, C.B. (1984). Human evolution and biological adaptation in the Pleistocene. In *Hominid Evolution and Community Ecology*, R. Foley (eds.), Academic Press, New York, pp 55-83.

Stringer, C.B. (1989). The origin of early modern humans: a comparison of the European and non-European evidence. In *The Human Revolution,* P. Mellars and C. Stringer (eds), Edinburgh University Press, Edinburgh, pp 232-244.

Stringer, C.B. (1990). The emergence of modern humans. *Scientific American* Dec: 68-74.

Stringer, C.B. and Andrews, P. (1988). Genetic and fossil evidence for the origin of modern humans. *Science* 239: 1263-1268.

Stringer, C.B., Hublin, J.-J. and Vandermeersch, B. (1984). The origin of anatomically modern humans in Western Europe. In *The Origins of Modern Humans: a World Survey of the Fossil Evidence*, F.H. Smith and F. Spencer (eds) , Alan R., Liss, New York, pp 51-135.

Swofford, D.L and Olsen, G.J. (1990). Phylogeny reconstruction. In *Molecular Systematics*, D.M. Hillis and C. Moritz (eds), Sinauer, Sunderland, Massachusetts, pp 411-501.

Sykes, B.C. (1983). DNA in heritable disease. *Lancet* II: 787-788.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Takahata, N. and Palumbi, S.R. (1985). Extranuclear differentiation and gene flow in the finite island model. *Genetics* 109: 441-457.

Templeton, A.R. (1992). Human origins and analysis of mitochondrial DNA sequences. *Science* 255: 737.

Theal, G. McC. (1910). *The yellow and dark-skinned people of Africa south of the Zambesi*. Swann Sonnenschein, London.

Thorpe, A.G. and Wolpoff, M.H. (1992). The multiregional evolution of humans. *Scientific American* 266:28-41.

Tianyuan, L. and Etler, D.A. (1992). New Middle Pleistocene hominid crania from Yunxian in China. *Nature* 357: 404-407.

Tobias, P.V. (1964). Bushman hunter-gatherers: a study in human ecology. In *Ecological Studies in Southern Africa*, D.H.S. Davies (ed), Junk, The Hague, pp 67-86.

Tobias, P.V. (1991). *Images of Humanity. The selected writings of Philip Tobias*. Ashanti Publishing, Creda Press, Rivonia.

Torroni, A., Semino, O., Rose, G., De Benedictis, G., Brancati, C. and Santachiara-Benerecetti, A.S. (1990). Mitochondrial DNA polymorphisms in the Albanian population of Calabria (Southern Italy). *International Journal of Anthropology* 5: 97-104.

Torroni, A., Schurr, T.G., Yang, C.C., Szathmary, E.J.E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowlder, W.C., Lawrence, D.N., Weiss, K.M. and Wallace, D.C. (1992). Amerindian mitochondrial DNA analysis reveals pre-Clovis colonization of the New World. *Genetics* 130: 153-162.

Trinkaus, E. (1986). The neanderthals and modern human origins. *Annual Review of Anthropology* 15: 193-218.

Turner, V. (1957). *Schism and Continuity in an African Society: A study of Ndembu Villiage Life*, Mancester University Press.

van Warmelo, N.J. (1974). The classification of cultural groups. In *Bantu-speaking Peoples*, W.D. Hammond-Tooke (ed), London, pp 78-83.

Vansina, J. (1955). Initiation rituals of the Bushongo. *Africa* 25: 138-153.

Vansina, J. (1969). The bells of Kings. *Journal of African History* 10: 187-197.

Vansina, J. (1984). Western Bantu expansion. *Journal of African History* 25: 129-145.

Vedder, H. (1928). The Bergdama. In: *The native tribes of South West Africa*, C.H.L. Hahn (ed), Cape Times Ltd, Cape Town.

Vedder, H. (1938). *South West Africa in Early Times.* Oxford: University Press.

Vigilant, L. (1990). *Control region sequences from African populations and the evolution of human mitochondrial DNA.* PhD thesis, University of California, Berkeley.

Vigilant, L., Stoneking, M. and Wilson, A.C. (1988). Conformational mutation in human mtDNA detected by direct sequencing of enzymatically amplified DNA. *Nucleic Acids Research* 16: 5945-5955.

Vigilant, L., Pennington, R., Harpending, H., Kocher, T.D. and Wilson, A.C. (1989). Mitochondrial DNA sequences in single hairs from a southern African population. *Proceedings of the Natational Academy of Science, USA.* 86: 9350-9354.

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. and Wilson, A.C. (1991). African populations and the evolution of human mitochondrial DNA. *Science* 253: 1503-1507.

Vikki, J., Savontaus, M.L. and Nikolaskelainen, E.K. (1988). Human mitochondrial DNA types in Finnland. *Human Genetics* 80: 317-321.

Vyas, G.N., Bhatia, H.M., Banker, D.D. and Purandare, N.M. (1958). Study of blood groups and other genetical characters in six Gujerati endogamous groups in Western India. *Annals of Human Genetics* 22: 185-199.

Walberg, M.W. and Clayton, D.A. (1981). Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Research* 9: 5411-5421.

Wallace, D.C. (1989). Report of the committee on human mitochondrial DNA. *Cytogenetics and Cell Genetics* 51: 612-621.

Wallace, D.C., Garrison, K. and Knowler, W.C. (1985). Dramatic founder effect in Amerindian mitochondrial DNAs. *American Journal of Physical Anthropology* 68: 149-155.

Wallace, D.C. (1992). Diseases of the mitochondrial DNA. *Annual Review of Biochemistry* 61: 1175-1212.

Watterson, G.A. (1977). The homozygosity test of neutrality. *Genetics* 88: 405-417.

Westphal, E.O.L. (1963). The linguistic prehistory of southern Africa: Bush, Kwadi, Hottentot, and Bantu linguistic relationships. *Africa*: 33: 237-265.

Wilson, A.C. and Cann, R.L. (1992). The recent African Genesis of humans. *Scientific American* 266: 22-27.

Wilson, A.C., Stoneking, M. and Cann, R.L. (1991). Ancestral geographic states and the peril of parsimony. *Systematic Zoology* 40: 363-365.

Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. and Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* 26: 375-400.

Wilson, A.C., Stoneking, M., Cann, R.L., Prager, E.M., Ferris, S.D., Wrishcnik, L.A. and Higuchi, R.G. (1987). Mitochondrial clans and the age of our common mother. In *Human Genetics*, F. Voge and K. Sperling (eds), Springer-Verlag, Berlin, pp 158-164.

Whittam, T.S., Clark, A.G., Stoneking, M., Cann R.L. and Wilson, A.C. (1986). Allelic variation in human mitochondrial genes based on patterns of restriction site polymorphism. *Proceedings of the National Academy of Science, USA*. 83: 9611-9615.

Wolpoff, M.H. (1984). Evolution of *Homo erectus*: the question of stasis. *Paleobiology* 10: 389-406.

Wolpoff, M.H. (1989). Multiregional evolution: The fossil alternative to Eden. In *The Human Revolution*, P Mellars and C. Stringer (eds), Edinburgh University Press, Edinburgh, pp. 62-108.

Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. and Wilson, A.C. (1987). Length mutations in human mitochondrial DNA: Direct sequencing of enzymatically amplified DNA. *Nucleic Acids Research* 15: 529-542.

Zoutendyk, A., Kopec, A.C. and Mourant, A.E. (1953). The blood groups of the Bushmen. *American Journal of Physical Anthropology* 11: 361-368.

Zoutendyk, A., Kopec, A.C. and Mourant, A.E. (1965). The blood groups of the Hottentots. *American Journal of Physical Anthropology* 13: 691-698.

Zuckerman, S.H., Solus, J.F., Gillespie, F.P. & Eisenstadt, J.M. (1984). Retention of both parental mitochondrial DNA species in mouse-Chinese hamster somatic cell hybrids. *Somatic Cell and Molecular Genetics* 10: 85-92.

APPENDIX I: The sample I.D., ethnicity, language group, sex, place of collection and laboratory reference number of individuals investigated for mtDNA restriction fragment length polymorphisms. Individuals were allocated identification numbers corresponding with their ethnicity (alphabetical code) and a binomial numbering system was used to differentiate samples as they were collected. Abbreviations used: JHB=Johannesburg, Sot/Tsw=Sotho/Tswana, Eksteenfon.=Eksteenfontein and Eng/Afrik=English and Afrikaans speaking.

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Zu38	Nguni	Zulu	M	JHB	HBTS 1096/89
Zu39	Nguni	Zulu	M	JHB	HBTS 1098/89
Zu41	Nguni	Zulu	M	JHB	HBTS 1108/89
Zu42	Nguni	Zulu	M	JHB	HBTS 1113/89
Zu43	Nguni	Zulu	M	JHB	HBTS 1148/89
Zu47	Nguni	Zulu	M	JHB	HBTS 1952/89
Zu48	Nguni	Zulu	M	JHB	HBTS 1154/89
Zu50	Nguni	Zulu	M	JHB	HBTS11162/89
Zu14	Nguni	Zulu	F	JHB	314/88
Zu16	Nguni	Zulu	M	JHB	343/88
Zu18	Nguni	Zulu	F	JHB	347/88
Zu19	Nguni	Zulu	M	JHB	353/88
Zu20	Nguni	Zulu	F	JHB	356/88
Zu22	Nguni	Zulu	F	JHB	1/89
Zu24	Nguni	Zulu	F	JHB	5/89
Zu25	Nguni	Zulu	F	JHB	10/89
Zu26	Nguni	Zulu	F	JHB	13/89
Zu27	Nguni	Zulu	F	JHB	22/89
Zu28	Nguni	Zulu	F	JHB	29/89
Zu49	Nguni	Zulu	M	JHB	HBTS11154/89
Zu50	Nguni	Zulu	M	JHB	HBTF11162/89
Zu51	Nguni	Zulu	M	JHB	HBTS11185/89
Zu52	Nguni	Zulu	M	JHB	HBTS41439/89
Zu56	Nguni	Zulu	M	JHB	HBTS41490/89
Zu58	Nguni	Zulu	M	JHB	HBTS41492/89
Zu67	Nguni	Zulu	M	JHB	76/89
Zu69	Nguni	Zulu	F	JHB	81/89
Zu72	Nguni	Zulu	F	JHB	94/89
Zu73	Nguni	Zulu	F	JHB	100/89
Zu77	Nguni	Zulu	M	JHB	143/89
Sw01	Nungi	Swazi	F	White River	Bh315/83
Sw02	Nguni	Swazi	F	White River	Bh337/83
Sw03	Nguni	Swazi	F	White River	Bh278/83
Sw04	Nguni	Swazi	F	White River	Bh324/83
Sw05	Nguni	Swazi	F	White River	Bh217/83
Sw06	Nguni	Swazi	F	White River	Bh325/83
Sw07	Nguni	Swazi	F	White River	Bh223/83
Sw11	Nguni	Swazi	F	White River	Bh216/83
Sw12	Nguni	Swazi	F	White River	Bh232/83
Sw14	Nguni	Swazi	F	White River	Bh286/83
Sw15	Nguni	Swazi	F	White River	Bh301/83
Sw16	Nguni	Swazi	F	White River	Bh304/83
Sw17	Nguni	Swazi	F	White River	Bh367/83
Sw18	Nguni	Swazi	M	White River	Bh256/83
Sw20	Nguni	Swazi	M	White River	Bh240/83
Sw21	Nguni	Swazi	F	White River	Bh316/83
Sw22	Nguni	Swazi	F	White River	Bh231/83
Sw23	Nguni	Swazi	F	White River	Bh313/83
Sw25	Nguni	Swazi	F	White River	Bh254/83
Sw26	Nguni	Swazi	M	White River	Bh347/83
Sw27	Nguni	Swazi	F	White River	Bh272/83
Sw28	Nguni	Swazi	F	White River	Bh362/83
Sw30	Nguni	Swazi	F	White River	Bh243/83
Sw31	Nguni	Swazi	F	White River	Bh342/83
Sw32	Nguni	Swazi	M	White River	Bh370/83
Sw33	Nguni	Swazi	M	White River	Bh263/83
Sw34	Nguni	Swazi	F	White River	Bh280/83
Sw36	Nguni	Swazi	F	White River	Bh277/83

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Sw37	Nguni	Swazi	F	White River	Bh282/83
Sw38	Nguni	Swazi	F	White River	Bh294/83
Sw39	Nguni	Swazi	F	White River	Bh312/83
Sw40	Nguni	Swazi	F	White River	Bh311/83
Sw41	Nguni	Swazi	F	White River	Bh323/83
Sw43	Nguni	Swazi	F	JHB	Bh321/83
Sw08	Nguni	Swazi	M	JHB	HBTS 1141/89
Sw10	Nguni	Swazi	M	JHB	HBTS 1146/89
Sw51	Nguni	Swazi	M	JHB	HBTS 1092/89
Sw52	Nguni	Swazi	M	JHB	HBTS 1095/89
Sw53	Nguni	Swazi	M	JHB	HBTS 1099/89
Sw54	Nguni	Swazi	M	JHB	HBTS 1116/89
Sw48	Nguni	Swazi	F	JHB	353/88
Xh01	Nguni	Xhosa	F	JHB	SAIMR/88
Xh02	Nguni	Xhosa	F	JHB	SAIMR/88
Xh03	Nguni	Xhosa	F	JHB	SAIMR/88
Xh04	Nguni	Xhosa	F	JHB	SAIMR/88
Xh05	Nguni	Xhosa	F	JHB	SAIMR/88
Xh07	Nguni	Xhosa	F	JHB	SAIMR/88
Xh10	Nguni	Xhosa	F	JHB	SAIMR/88
Xh11	Nguni	Xhosa	M	White River	Bh307/83
Xh20	Nguni	Xhosa	M	JHB	HBTS 1074/89
Xh22	Nguni	Xhosa	M	JHB	HBTS 1158/89
Xh14	Nguni	Xhosa	F	JHB	331/88
Xh15	Nguni	Xhosa	F	JHB	334/88
Xh16	Nguni	Xhosa	F	JHB	343/88
Xh17	Nguni	Xhosa	F	JHB	345/88
Xh18	Nguni	Xhosa	F	JHB	12/89
Xh24	Nguni	Xhosa	M	JHB	HBTS 41488/89
Xh27	Nguni	Xhosa	M	JHB	88/89
Xh28	Nguni	Xhosa	F	JHB	89/89
Tso01	Nguni	Tsonga	F	White River	Bh266/83
Tso02	Nguni	Tsonga	F	White River	Bh246/83
Tso03	Nguni	Tsonga	F	White River	Bh357/83
Tso05	Nguni	Tsonga	F	White River	Bh220/83
Tso07	Nguni	Tsonga	F	White River	Bh221/83
Tso09	Nguni	Tsonga	F	White River	Bh225/83
Tso10	Nguni	Tsonga	M	White River	Bh204/83
Tso11	Nguni	Tsonga	F	White River	Bh265/83
Tso12	Nguni	Tsonga	F	White River	Bh276/83
Tso15	Nguni	Tsonga	F	White River	Bh328/83
Tso16	Nguni	Tsonga	F	White River	Bh285/83
Tso17	Nguni	Tsonga	F	White River	Bh361/83
Tso18	Nguni	Tsonga	F	White River	Bh368/83
Tso19	Nguni	Tsonga	F	White River	Bh300/83
Tso20	Nguni	Tsonga	F	White River	Bh255/83
Tso21	Nguni	Tsonga	F	White River	Bh249/83
Tso22	Nguni	Tsonga	F	White River	Bh229/83
Tso23	Nguni	Tsonga	F	White River	Bh245/83
Tso24	Nguni	Tsonga	M	White River	Bh354/83
Tso25	Nguni	Tsonga	M	White River	Bh351/83
Tso20	Nguni	Tsonga	F	White River	Bh255/83
Tso27	Nguni	Tsonga	F	White River	Bh340/83
Tso28	Nguni	Tsonga	F	White River	Bh237/83
Tso29	Nguni	Tsonga	F	White River	Bh248/83
Tso30	Nguni	Tsonga	F	White River	Bh287/83
Tso31	Nguni	Tsonga	F	White River	Bh251/83
Tso32	Nguni	Tsonga	M	White River	Bh364/83
Tso06	Nguni	Tsonga	M	JHB	HBTS 1144/89
Tso08	Nguni	Tsonga	M	JHB	HBTS 1156/89
Tso37	Nguni	Tsonga	M	JHB	HBTS 1079/89
Tso13	Nguni	Tsonga	M	JHB	HBTS 11129/89

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Tso39	Nguni	Tsonga	M	JHB	HBTS 41486/89
Tso41	Nguni	Tsonga	M	JHB	HBTS 41507/89
Tso42	Nguni	Tsonga	F	JHB	64/89
Sot01	Sot/Tsw	Sotho	F	White River	Bh239/83
Sot02	Sot/Tsw	Sotho	F	White River	Bh305/83
Sot05	Sot/Tsw	Sotho	F	White River	Bh213/83
Sot06	Sot/Tsw	Sotho	F	White River	Bh333/83
Sot12	Sot/Tsw	Sotho	F	JHB	Bh314/83
Sot26	Sot/Tsw	Sotho	M	JHB	198/89
Sot27	Sot/Tsw	Sotho	F	JHB	SAIMR/88
Sot29	Sot/Tsw	Sotho	F	JHB	SAIMR/87
Sot30	Sot/Tsw	Sotho	F	JHB	SAIMR/87
Sot44	Sot/Tsw	Sotho	M	JHB	HBTS 1068/89
Sot45	Sot/Tsw	Sotho	M	JHB	HBTS 1106/89
Sot46	Sot/Tsw	Sotho	M	JHB	HBTS 1107/89
Sot48	Sot/Tsw	Sotho	M	JHB	HBTS 1112/89
Sot03	Sot/Tsw	Sotho	M	JHB	HBTS 1153/89
Sot53	Sot/Tsw	Sotho	M	JHB	SAIMR/90
Sot54	Sot/Tsw	Sotho	M	JHB	SAIMR/90
Sot07	Sot/Tsw	Sotho	M	JHB	HBTS 41466/89
Sot11	Sot/Tsw	Sotho	F	JHB	52/89
Sot17	Sot/Tsw	Sotho	F	JHB	93/89
Sot19	Sot/Tsw	Sotho	F	JHB	98/89
Sot32	Sot/Tsw	Sotho	M	JHB	306/88
Sot33	Sot/Tsw	Sotho	F	JHB	313/88
Sot35	Sot/Tsw	Sotho	M	JHB	325/88
Sot36	Sot/Tsw	Sotho	F	JHB	333/88
Sot37	Sot/Tsw	Sotho	F	JHB	335/88
Sot39	Sot/Tsw	Sotho	F	JHB	339/88
Sot40	Sot/Tsw	Sotho	F	JHB	350/88
Sot43	Sot/Tsw	Sotho	F	JHB	31/89
Sot13	Sot/Tsw	Sotho	M	JHB	52/89
Tsw01	Sot/Tsw	Tswana	F	JHB	SAIMR/88
Tsw02	Sot/Tsw	Tswana	F	JHB	SAIMR/88
Tsw05	Sot/Tsw	Tswana	F	JHB	SAIMR/88
Tsw06	Sot/Tsw	Tswana	M	JHB	SAIMR/88
Tsw07	Sot/Tsw	Tswana	F	JHB	SAIMR/88
Tsw22	Sot/Tsw	Tswana	M	JHB	HBTS 1086/89
Tsw23	Sot/Tsw	Tswana	M	JHB	HBTS 1100/89
Tsw24	Sot/Tsw	Tswana	M	JHB	HBTS 1105/89
Tsw25	Sot/Tsw	Tswana	M	JHB	HBTS 1109/89
Tsw31	Sot/Tsw	Tswana	F	JHB	51/89
Tsw27	Sot/Tsw	Tswana	F	JHB	76/89
Tsw32	Sot/Tsw	Tswana	M	JHB	90/89
Tsw33	Sot/Tsw	Tswana	F	JHB	91/89
Tsw34	Sot/Tsw	Tswana	F	JHB	92/89
Tsw35	Sot/Tsw	Tswana	M	JHB	98/89
Ped02	Sot/Tsw	Pedi	F	JHB	SAIMR/87
Ped03	Sot/Tsw	Pedi	F	JHB	SAIMR/87
Ped04	Sot/Tsw	Pedi	M	JHB	SAIMR/87
Ped05	Sot/Tsw	Pedi	F	JHB	SAIMR/87
Ped06	Sot/Tsw	Pedi	M	JHB	SAIMR/88
Ped07	Sot/Tsw	Pedi	F	JHB	SAIMR/88
Ped08	Sot/Tsw	Pedi	M	JHB	SAIMR/87
Ped09	Sot/Tsw	Pedi	M	JHB	SAIMR/88
Ped17	Sot/Tsw	Pedi	M	JHB	HBTS 1096/88
Ped19	Sot/Tsw	Pedi	M	JHB	HBTS 1156/88
Ped20	Sot/Tsw	Pedi	M	JHB	HBTS 1145/88
Ped01	Sot/Tsw	Pedi	M	JHB	HBTS 1154/88
Ped11	Sot/Tsw	Pedi	F	JHB	323/88
Ped15	Sot/Tsw	Pedi	F	JHB	348/88
Ped21	Sot/Tsw	Pedi	M	JHB	HBTS 11126/88

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Ped22	Sot/Tsw	Pedi	M	JHB	HBTS 41445/88
Ped23	Sot/Tsw	Pedi	M	JHB	HBTS 41454/88
Ped24	Sot/Tsw	Pedi	M	JHB	HBTS 41464/88
Ped25	Sot/Tsw	Pedi	M	JHB	HBTS 41465/88
Ped26	Sot/Tsw	Pedi	M	JHB	HBTS 11132/88
Ped27	Sot/Tsw	Pedi	M	JHB	HBTS 11133/88
Ped29	Sot/Tsw	Pedi	M	JHB	HBTS 41480/88
Ped30	Sot/Tsw	Pedi	M	JHB	HBTS 41482/88
Ve01	Venda	Venda	M	Matangari	1983/05/23
Ve02	Venda	Venda	M	Matangari	1983/05/23
Ve03	Venda	Venda	F	Matangari	1983/05/23
Ve04	Venda	Venda	M	Matangari	1983/05/23
Ve05	Venda	Venda	F	Matangari	1983/05/23
Ve06	Venda	Venda	F	Matangari	1983/05/23
Ve07	Venda	Venda	F	Matangari	1983/05/23
Ve08	Venda	Venda	F	Matangari	1983/05/23
Ve09	Venda	Venda	F	Matangari	1983/05/23
Ve10	Venda	Venda	M	Matangari	1983/05/23
Ve11	Venda	Venda	F	Matangari	1983/05/23
Ve12	Venda	Venda	F	Matangari	1983/05/23
Ve13	Venda	Venda	M	Matangari	1983/05/23
Ve14	Venda	Venda	M	Matangari	1983/05/23
Ve15	Venda	Venda	F	Matangari	1983/05/23
Ve16	Venda	Venda	F	Matangari	1983/05/23
Ve17	Venda	Venda	M	Matangari	1983/05/23
Ve18	Venda	Venda	F	Matangari	1983/05/23
Ve20	Venda	Venda	F	Matangari	1983/05/23
Ve21	Venda	Venda	M	Matangari	1983/05/23
Ve22	Venda	Venda	M	Matangari	1983/05/23
Ve24	Venda	Venda	F	Matangari	1983/05/23
Ve26	Venda	Venda	M	Matangari	1983/05/23
Ve27	Venda	Venda	F	Matangari	1983/05/23
Ve29	Venda	Venda	F	Matangari	1983/05/23
Ve31	Venda	Venda	F	Matangari	1983/05/23
Ve32	Venda	Venda	F	Matangari	1983/05/23
Ve33	Venda	Venda	M	Matangari	1983/05/23
Ve34	Venda	Venda	M	Matangari	1983/05/23
Ve35	Venda	Venda	M	Matangari	1983/05/23
Lem01	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem05	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem06	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem07	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem08	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem10	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem11	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem12	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem14	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem16	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem19	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem20	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem21	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem24	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem27	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem30	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem31	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem32	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem33	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem35	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem37	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem38	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem41	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem42	Lemba	Lemba	M	Pietersburg	1987/11/10

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Lem43	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem45	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem46	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem47	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem49	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem53	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem54	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem55	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem57	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem59	Lemba	Lemba	F	Pietersburg	1987/11/10
Lem60	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem61	Lemba	Lemba	M	Pietersburg	1987/11/10
Lem70	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem71	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem75	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem78	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem79	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem80	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem81	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem82	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem83	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem84	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem85	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem87	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem91	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem92	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem94	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem95	Lemba	Lemba	M	Pietersburg	1988/08/10
Lem98	Lemba	Lemba	M	Pietersburg	1988/08/10
He002	Herero	Herero	F	Namibia	1987/09/02
He003	Herero	Herero	F	Namibia	1987/09/02
He004	Herero	Herero	F	Namibia	1987/09/02
He006	Herero	Herero	F	Namibia	1987/09/02
He008	Herero	Herero	F	Namibia	1987/09/02
He010	Herero	Herero	F	Namibia	1987/09/02
He009	Herero	Herero	F	Namibia	1987/09/02
He012	Herero	Herero	F	Namibia	1987/09/02
He016	Herero	Herero	F	Namibia	1987/09/02
He017	Herero	Herero	F	Namibia	1987/09/02
He018	Herero	Herero	F	Namibia	1987/09/09
He019	Herero	Herero	F	Namibia	1987/09/09
He020	Herero	Herero	F	Namibia	1987/09/09
He021	Herero	Herero	F	Namibia	1987/09/09
He022	Herero	Herero	F	Namibia	1987/09/09
He023	Herero	Herero	F	Namibia	1987/09/09
He025	Herero	Herero	F	Namibia	1987/09/09
He027	Herero	Herero	F	Namibia	1987/09/09
He028	Herero	Herero	F	Namibia	1987/09/09
He029	Herero	Herero	F	Namibia	1987/09/09
He030	Herero	Herero	F	Namibia	1987/09/09
He031	Herero	Herero	F	Namibia	1987/09/09
He032	Herero	Herero	F	Namibia	1987/09/09
He033	Herero	Herero	F	Namibia	1987/09/09
He034	Herero	Herero	F	Namibia	1987/09/09
He068	Herero	Herero	F	Namibia	1987/10/22
He069	Herero	Herero	F	Namibia	1987/10/22
He070	Herero	Herero	F	Namibia	1987/10/22
He071	Herero	Herero	F	Namibia	1987/10/22
He072	Herero	Herero	F	Namibia	1987/10/22
He073	Herero	Herero	F	Namibia	1987/10/22
He076	Herero	Herero	F	Namibia	1987/10/22
He077	Herero	Herero	F	Namibia	1987/10/22

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
He078	Herero	Herero	F	Namibia	1987/10/22
He079	Herero	Herero	F	Namibia	1987/10/22
He080	Herero	Herero	F	Namibia	1987/10/22
He082	Herero	Herero	F	Namibia	1987/10/22
He083	Herero	Herero	F	Namibia	1987/10/22
He084	Herero	Herero	F	Namibia	1987/10/22
He085	Herero	Herero	F	Namibia	1987/10/22
He091	Herero	Herero	F	Namibia	1987/11/18
He093	Herero	Herero	F	Namibia	1987/11/18
He094	Herero	Herero	F	Namibia	1987/11/18
He095	Herero	Herero	F	Namibia	1987/11/18
He096	Herero	Herero	F	Namibia	1987/11/18
He097	Herero	Herero	F	Namibia	1987/11/25
He098	Herero	Herero	F	Namibia	1987/11/25
He099	Herero	Herero	F	Namibia	1987/11/25
He100	Herero	Herero	F	Namibia	1987/11/25
He101	Herero	Herero	F	Namibia	1987/11/25
He102	Herero	Herero	F	Namibia	1987/11/25
He103	Herero	Herero	F	Namibia	1987/11/25
He105	Herero	Herero	F	Namibia	1987/11/25
He106	Herero	Herero	F	Namibia	1987/11/25
Da01	Dama	Dama	F	Windhoek	1986/10/29
Da07	Dama	Dama	F	Windhoek	1986/10/29
Da10	Dama	Dama	M	Windhoek	1986/10/29
Da11	Dama	Dama	F	Windhoek	1986/10/29
Da12	Dama	Dama	F	Windhoek	1986/10/29
Da13	Dama	Dama	F	Windhoek	1986/10/29
Da14	Dama	Dama	F	Windhoek	1986/11/27
Da15	Dama	Dama	F	Windhoek	1986/11/27
Da16	Dama	Dama	F	Windhoek	1986/11/27
Da17	Dama	Dama	F	Windhoek	1986/11/27
Da18	Dama	Dama	M	Windhoek	1986/11/27
Da19	Dama	Dama	M	Windhoek	1986/11/27
Da20	Dama	Dama	M	Windhoek	1986/11/27
Da21	Dama	Dama	M	Windhoek	1986/11/27
Da22	Dama	Dama	M	Windhoek	1986/11/27
Da23	Dama	Dama	F	Windhoek	1986/11/05
Da24	Dama	Dama	F	Windhoek	1986/11/05
Da25	Dama	Dama	F	Windhoek	1986/11/05
Da26	Dama	Dama	F	Windhoek	1986/11/05
Da27	Dama	Dama	F	Windhoek	1986/11/05
Da28	Dama	Dama	F	Windhoek	1986/11/05
Da29	Dama	Dama	M	Windhoek	1986/11/05
Da30	Dama	Dama	F	Windhoek	1986/11/05
Da31	Dama	Dama	F	Windhoek	1986/11/05
Da32	Dama	Dama	F	Windhoek	1986/11/05
Da33	Dama	Dama	F	Windhoek	1986/11/05
Da34	Dama	Dama	F	Windhoek	1986/11/05
Da35	Dama	Dama	F	Windhoek	1986/11/05
Da36	Dama	Dama	F	Windhoek	1986/11/05
Da37	Dama	Dama	F	Windhoek	1986/11/05
Da51	Dama	Dama	M	Windhoek	1989/01/05
Da52	Dama	Dama	M	Windhoek	1989/01/05
Da53	Dama	Dama	M	Windhoek	1989/01/06
Da54	Dama	Dama	M	Windhoek	1989/01/06
Da58	Dama	Dama	M	Windhoek	1989/01/11
Da62	Dama	Dama	M	Windhoek	1989/01/13
Da65	Dama	Dama	M	Windhoek	1989/01/26
Da66	Dama	Dama	M	Windhoek	1989/01/27
Da67	Dama	Dama	M	Windhoek	1989/01/31
Da68	Dama	Dama	M	Windhoek	1989/01/31
Da71	Dama	Dama	M	Windhoek	1990/02/01

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Da72	Dama	Dama	M	Windhoek	1990/02/02
Da75	Dama	Dama	M	Windhoek	1990/02/02
Ov01	Ambo	Kwanyama	F	Windhoek	1988/04/21
Ov02	Ambo	Kwanyama	F	Windhoek	1988/04/21
Ov03	Ambo	Kwanyama	F	Windhoek	1988/04/21
Ov04	Ambo	Kwambe	M	Windhoek	1988/04/21
Ov05	Ambo	Kwanyama	M	Windhoek	1988/04/21
Ov06	Ambo	Indomga	M	Windhoek	1988/04/21
Ov07	Ambo	Kwanyama	F	Windhoek	1988/04/21
Ov08	Ambo	Kwanyama	F	Windhoek	1988/04/21
Ov09	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov10	Ambo	Ongandjera	F	Windhoek	1988/04/28
Ov11	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov13	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov14	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov15	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov16	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov17	Ambo	Kwanyama	F	Windhoek	1988/04/28
Ov18	Ambo	Kwanyama	M	Windhoek	1988/04/28
Ov19	Ambo	Ongandjera	M	Windhoek	1988/04/28
Ov20	Ambo	Kwanyama	F	Windhoek	1988/05/13
0v21	Ambo	Kwanyama	F	Windhoek	1988/05/13
Ov22	Ambo	Kwanyama	F	Windhoek	1988/05/13
Ov23	Ambo	Kwanyam	F	Windhoek	1988/05/13
Om010	Barakwena	!Kung	M	Omega Base	1985/10/30
Om019	Barakwena	!Kung	M	Omega Base	1985/10/30
Om024	Barakwena	!Kung	M	Omega Base	1985/11/27
Om031	Barakwena	!Kung	M	Omega Base	1985/10/30
Om033	Barakwena	!Kung	M	Omega Base	1985/10/30
Om034	Barakwena	!Kung	M	Omega Base	1985/10/30
Om035	Barakwena	!Kung	M	Omega Base	1985/10/30
Om038	Barakwena	!Kung	M	Omega Base	1985/10/30
Om039	Barakwena	!Kung	M	Omega Base	1985/10/30
Om042	Barakwena	!Kung	M	Omega Base	1985/10/30
Om054	Sekele	!Kung	M	Omega Base	1985/11/05
Om101	Sekele	!Kung	M	Omega Base	1985/11/27
Om103	Sekeke	!Kung	M	Omega Base	1985/11/27
Om104	Sekele	!Kung	M	Omega Base	1985/11/27
Om105	Sekele	!Kung	M	Omega Base	1985/11/27
Om106	Sekele	!Kung	M	Omega Base	1985/11/27
Om107	Sekele	!Kung	M	Omega Base	1985/11/27
Om108	Sekele	!Kung	M	Omega Base	1985/11/27
Om109	Sekele	!Kung	M	Omega Base	1985/11/27
Om110	Sekele	!Kung	M	Omega Base	1985/11/27
Om111	Sekele	!Kung	M	Omega Base	1985/11/27
Om112	Sekele	!Kung	M	Omega Base	1985/11/27
Om113	Sekele	!Kung	M	Omega Base	1985/11/27
Om114	Sekele	!Kung	M	Omega Base	1985/11/27
Om115	Sekele	!Kung	M	Omega Base	1985/11/27
Om116	Sekele	!Kung	M	Omega Base	1985/11/27
Om117	Sekele	!Kung	M	Omega Base	1985/11/27
Om118	Sekele	!Kung	M	Omega Base	1985/11/27
Om119	Sekele	!Kung	M	Omega Base	1985/11/27
Om120	Sekele	!Kung	M	Omega Base	1985/11/27
Om123	Sekele	!Kung	M	Omega Base	1985/11/27
Om124	Sekele	!Kung	M	Omega Base	1985/11/27
Om125	Sekele	!Kung	M	Omega Base	1985/11/27
Om126	Sekele	!Kung	M	Omega Base	1985/12/03
Om127	Sekele	!Kung	M	Omega Base	1985/12/03
Om128	Sekele	!Kung	M	Omega Base	1985/12/03
Om129	Sekele	!Kung	M	Omega Base	1985/12/03
Om131	Sekele	!Kung	M	Omega Base	1985/12/03

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Om132	Sekele	!Kung	M	Omega Base	1985/12/03
Om133	Sekele	!Kung	M	Omega Base	1985/12/03
Om134	Sekele	!Kung	M	Omega Base	1985/12/03
Om135	Sekele	!Kung	M	Omega Base	1985/12/03
Om136	Sekele	!Kung	M	Omega Base	1985/12/03
Om137	Sekele	!Kung	M	Omega Base	1985/12/03
Om138	Sekele	!Kung	M	Omega Base	1985/12/03
Om139	Sekele	!Kung	M	Omega Base	1985/12/03
Om140	Sekele	!Kung	M	Omega Base	1985/12/03
Om141	Sekele	!Kung	M	Omega Base	1985/12/03
Om142	Sekele	!Kung	M	Omega Base	1985/12/03
Om143	Sekele	!Kung	M	Omega Base	1985/12/03
Om144	Sekele	!Kung	M	Omega Base	1985/12/03
Om145	Sekele	!Kung	M	Omega Base	1985/12/03
Om147	Sekele	!Kung	M	Omega Base	1985/12/03
Om148	Sekele	!Kung	M	Omega Base	1985/12/03
Om149	Sekele	!Kung	M	Omega Base	1985/12/03
Om150	Sekele	!Kung	M	Omega Base	1985/12/03
Jm202	!Kung	!Kung	F	Tsumkwe	1981/06/03
Jm031	!Kung	!Kung	F	Tsumkwe	1981/06/03
Jm278	!Kung	!Kung	F	Tsumkwe	1981/06/03
Na01	Nama	Nama	F	Windhoek	1988/12/12
Na02	Nama	Nama	F	Windhoek	1988/12/12
Na03	Nama	Nama	F	Windhoek	1988/12/12
Na04	Nama	Nama	F	Windhoek	1988/12/12
Na05	Nama	Nama	F	Windhoek	1988/12/12
Na06	Nama	Nama	F	Windhoek	1988/12/12
Na09	Nama	Nama	M	Windhoek	1988/12/28
Na10	Nama	Nama	F	Windhoek	1988/12/28
Na11	Nama	Nama	M	Windhoek	1988/12/30
Na14	Nama	Nama	M	Windhoek	1989/01/10
Na15	Nama	Nama	M	Windhoek	1989/01/10
Na16	Nama	Nama	M	Windhoek	1989/01/10
Na20	Nama	Nama	M	Windhoek	1989/01/17
Na23	Nama	Nama	M	Windhoek	1989/01/19
Na24	Nama	Nama	M	Windhoek	1989/01/19
Na25	Nama	Nama	M	Windhoek	1989/01/19
Na26	Nama	Nama	M	Windhoek	1989/01/20
Na28	Nama	Nama	M	Windhoek	1989/01/24
Na07	Nama	Nama	F	Windhoek	1989/01/24
Na08	Nama	Nama	F	Windhoek	1988/12/19
Na12	Nama	Nama	F	Windhoek	1989/01/04
Na13	Nama	Nama	F	Windhoek	1989/01/04
Na17	Nama	Nama	F	Windhoek	1989/01/11
Na18	Nama	Nama	F	Windhoek	1989/01/11
Na19	Nama	Nama	F	Windhoek	1989/01/11
Na21	Nama	Nama	F	Windhoek	1989/01/18
Na22	Nama	Nama	F	Windhoek	1989/01/18
Na27	Nama	Nama	F	Windhoek	1989/01/24
Na29	Nama	Nama	M	Windhoek	1989/01/24
Na31	Nama	Nama	F	Windhoek	1989/01/26
Na32	Nama	Nama	M	Windhoek	1989/01/26
Na34	Nama	Nama	M	Windhoek	1989/02/01
Na35	Nama	Nama	F	Windhoek	1989/02/02
Na36	Nama	Nama	F	Windhoek	1989/02/02
Na37	Nama	Nama	F	Windhoek	1989/02/02
Na38	Nama	Nama	F	Windhoek	1989/02/14
Na39	Nama	Nama	F	Windhoek	1989/02/15
Na40	Nama	Nama	M	Windhoek	1989/02/24
Na41	Nama	Nama	F	Windhoek	1989/02/24
Na42	Nama	Nama	F	Windhoek	1989/03/02
Na43	Nama	Nama	F	Windhoek	1989/03/02

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Na44	Nama	Nama	F	Windhoek	1989/03/02
Na45	Nama	Nama	F	Windhoek	1989/03/02
Na46	Nama	Nama	F	Windhoek	1989/03/02
Na47	Nama	Nama	M	Windhoek	1989/03/03
Na48	Nama	Nama	M	Windhoek	1989/03/03
Rv01	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 506/89
Rv02	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 505/89
Rv03	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 504/89
Rv04	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 502/89
Rv05	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 501/89
Rv06	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 498/89
Rv07	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 497/89
Rv08	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 494/89
Rv09	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 493/89
Rv10	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 489/89
Rv11	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 483/89
Rv12	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 482/89
Rv13	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 476/89
Rv14	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 472/89
Rv15	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 471/89
Rv16	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 459/89
Rv17	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 467/89
Rv18	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 450/89
Rv19	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 448/89
Rv20	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 445/89
Rv21	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 410/89
Rv22	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 409/89
Rv23	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 404/89
Rv24	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 399/89
Rv25	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 398/89
Rv26	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 393/89
Rv27	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 387/89
Rv28	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 386/89
Rv29	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 380/89
Rv30	"Coloured"	Afrikaans	F	Eksteensfon.	JVW 350/89
Rv31	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 347/89
Rv32	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 372/89
Rv33	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 342/89
Rv35	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 495/89
Rv37	"Coloured"	Afrikaans	M	Eksteensfon.	JVW 490/89
Rc01	"Coloured"	Eng/Afrik	M	JHB	HBTS 15128/87
Rc02	"Coloured"	Eng/Afrik	M	JHB	HBTS 15123/87
Rc03	"Coloured"	Eng/Afrik	M	JHB	HBTS 15117/87
Rc04	"Coloured"	Eng/Afrik	M	JHB	HBTS 45330/87
Rc05	"Coloured"	Eng/Afrik	M	JHB	HBTS 45327/87
Rc06	"Coloured"	Eng/Afrik	F	JHB	HBTS 05629/87
Rc07	"Coloured"	Eng/Afrik	M	JHB	CM 1/87
Rc08	"Coloured"	Eng/Afrik	M	JHB	CM 2/87
Rc09	"Coloured"	Eng/Afrik	M	JHB	CM 3/87
Rc10	"Coloured"	Eng/Afrik	M	JHB	CM 4/87
Rc11	"Coloured"	Eng/Afrik	M	JHB	CM 5/87
Rc12	"Coloured"	Eng/Afrik	M	JHB	CM 6/87
Rc13	"Coloured"	Eng/Afrik	M	JHB	CM 7/87
Rc14	"Coloured"	Eng/Afrik	M	JHB	CM 8/87
Rc15	"Coloured"	Eng/Afrik	M	JHB	CM 9/87
Rc16	"Coloured"	Eng/Afrik	M	JHB	CM 10/87
Rc17	"Coloured"	Eng/Afrik	M	JHB	CM 11/87
Rc18	"Coloured"	Eng/Afrik	M	JHB	CM 12/87
Rc19	"Coloured"	Eng/Afrik	M	JHB	CM 13/87
Rc20	"Coloured"	Eng/Afrik	M	JHB	CM 14/87
Rc21	"Coloured"	Eng/Afrik	M	JHB	CM 15/87
Rc22	"Coloured"	Eng/Afrik	M	JHB	CM 16/87

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
Rc23	"Coloured"	Eng/Afrik	M	JHB	CM 17/87
Rc24	"Coloured"	Eng/Afrik	M	JHB	CM 18/87
Rc25	"Coloured"	Eng/Afrik	M	JHB	CM 19/87
Rc26	"Coloured"	Eng/Afrik	M	JHB	CM 20/87
Rc27	"Coloured"	Eng/Afrik	M	JHB	CM 21/87
Rc28	"Coloured"	Eng/Afrik	M	JHB	CM 22/87
Rc29	"Coloured"	Eng/Afrik	M	JHB	CM 23/87
Rc30	"Coloured"	Eng/Afrik	M	JHB	CM 24/87
Rc31	"Coloured"	Eng/Afrik	M	JHB	CM 25/87
Rc32	"Coloured"	Eng/Afrik	M	JHB	CM 26/87
Rc33	"Coloured"	Eng/Afrik	M	JHB	CM 27/87
Rc34	"Coloured"	Eng/Afrik	M	JHB	CM 28/87
Rc35	"Coloured"	Eng/Afrik	M	JHB	CM 29/87
Rc36	"Coloured"	Eng/Afrik	M	JHB	CM 30/87
Rc37	"Coloured"	Eng/Afrik	M	JHB	CM 31/87
Rc38	"Coloured"	Eng/Afrik	M	JHB	HBTS 35946/87
Rc39	"Coloured"	Eng/Afrik	M	JHB	HBTS 35951/87
Rc40	"Coloured"	Eng/Afrik	M	JHB	HBTS 35952/87
Rc41	"Coloured"	Eng/Afrik	M	JHB	HBTS 35955/87
Rc42	"Coloured"	Eng/Afrik	M	JHB	HBTS 35959/87
Rc43	"Coloured"	Eng/Afrik	M	JHB	HBTS 35956/87
Rc44	"Coloured"	Eng/Afrik	M	JHB	HBTS 35961/87
Rc45	"Coloured"	Eng/Afrik	M	JHB	HBTS 35962/87
Rc46	"Coloured"	Eng/Afrik	M	JHB	HBTS 35958/87
Rc47	"Coloured"	Eng/Afrik	M	JHB	HBTS 35965/87
Rc49	"Coloured"	Eng/Afrik	M	JHB	HBTS 25485/87
Rc50	"Coloured"	Eng/Afrik	F	JHB	HBTS 25476/87
Rc51	"Coloured"	Eng/Afrik	M	JHB	HBTS 25479/87
Rc52	"Coloured"	Eng/Afrik	F	JHB	HBTS 25475/87
Rc53	"Coloured"	Eng/Afrik	M	JHB	HBTS 25474/87
Rc54	"Coloured"	Eng/Afrik	M	JHB	HBTS 25477/87
Rc55	"Coloured"	Eng/Afrik	M	JHB	HBTS 56932/87
Rc57	"Coloured"	Eng/Afrik	M	JHB	HBTS 56948/87
Rc58	"Coloured"	Eng/Afrik	F	JHB	HBTS 56918/87
Rc59	"Coloured"	Eng/Afrik	F	JHB	HBTS 56917/87
Rc66	"Coloured"	Eng/Afrik	M	JHB	CM 33/87
Rc67	"Coloured"	Eng/Afrik	M	JHB	CM 34/87
Rc68	"Coloured"	Eng/Afrik	M	JHB	CM 35/87
Rc69	"Coloured"	Eng/Afrik	M	JHB	CM 36/87
Rc70	"Coloured"	Eng/Afrik	M	JHB	CM 37/87
Rc71	"Coloured"	Eng/Afrik	M	JHB	CM 38/87
Rc72	"Coloured"	Eng/Afrik	M	JHB	CM 39/87
Rc73	"Coloured"	Eng/Afrik	M	JHB	CM 40/87
Rc74	"Coloured"	Eng/Afrik	M	JHB	CM 41/87
Rc75	"Coloured"	Eng/Afrik	M	JHB	CM 42/87
Rc76	"Coloured"	Eng/Afrik	M	JHB	CM 43/87
Rc77	"Coloured"	Eng/Afrik	M	JHB	CM 44/87
Rc78	"Coloured"	Eng/Afrik	M	JHB	CM 45/87
Rc79	"Coloured"	Eng/Afrik	M	JHB	CM 46/87
RW01	"White"	English	M	JHB	1987/02/10
RW02	"White"	English	F	JHB	1987/02/10
RW03	"White"	English	F	JHB	1987/02/10
RW04	"White"	English	F	JHB	1987/02/10
RW05	"White"	English	F	JHB	1987/02/10
RW06	"White"	English	F	JHB	1987/02/10
RW07	"White"	English	M	JHB	1987/02/10
RW08	"White"	English	M	JHB	1987/02/10
RW09	"White"	English	F	JHB	1987/02/10
RW11	"White"	English	F	JHB	1987/02/10
RW13	"White"	English	F	JHB	1987/02/10
RW14	"White"	English	F	JHB	1987/02/10
RW15	"White"	English	F	JHB	1987/02/10

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
RW16	"White"	English	F	JHB	1987/02/10
RW17	"White"	English	F	JHB	1987/02/10
RW18	"White"	English	M	JHB	1986/05/06
RW21	"White"	English	M	JHB	1987/02/10
RW27	"White"	English	F	JHB	1987/02/10
RW30	"White"	English	F	JHB	1987/02/10
RW31	"White"	English	M	JHB	1987/02/10
RW33	"White"	English	F	JHB	1987/02/10
RW35	"White"	English	F	JHB	1987/02/10
RW37	"White"	English	F	JHB	1987/02/10
RW40	"White"	English	F	JHB	1987/02/10
RW42	"White"	English	F	JHB	1987/02/10
RW43	"White"	English	F	JHB	1987/02/10
RW44	"White"	English	F	JHB	1987/02/10
RW45	"White"	English	F	JHB	1986/05/06
RW46	"White"	English	F	JHB	1988/04/08
RW50	"White"	English	M	JHB	1987/02/10
RW51	"White"	English	M	JHB	1987/02/01
RW53	"White"	English	F	JHB	1987/02/01
RW54	"White"	English	F	JHB	1987/02/01
RW55	"White"	English	M	JHB	1987/02/01
RW56	"White"	English	F	JHB	1987/02/01
RW41	"White"	English	F	JHB	1987/02/01
RJ01	Ashkenazi	Jew	F	JHB	1988/08/26
RJ02	Ashkenazi	Jew	F	JHB	1988/08/26
RJ03	Ashkenazi	Jew	M	JHB	1988/08/26
RJ04	Ashkenazi	Jew	M	JHB	1988/08/26
RJ05	Ashkenazi	Jew	F	JHB	1988/08/26
RJ06	Ashkenazi	Jew	M	JHB	1988/08/26
RJ07	Ashkenazi	Jew	M	JHB	1988/08/26
RJ08	Ashkenazi	Jew	M	JHB	1988/08/26
RJ09	Ashkenazi	Jew	F	JHB	1988/08/26
RJ10	Ashkenazi	Jew	M	JHB	1988/08/26
RJ11	Ashkenazi	Jew	F	JHB	1988/08/26
RJ12	Ashkenazi	Jew	M	JHB	1988/08/26
RJ13	Ashkenazi	Jew	M	JHB	1988/08/26
RJ14	Ashkenazi	Jew	F	JHB	1988/08/26
RJ15	Ashkenazi	Jew	F	JHB	1988/08/26
RJ17	Ashkenazi	Jew	M	JHB	1988/08/26
RJ18	Ashkenazi	Jew	M	JHB	1988/08/26
RJ19	Ashkenazi	Jew	M	JHB	1988/08/26
RJ20	Ashkenazi	Jew	M	JHB	1988/08/26
RJ21	Ashkenazi	Jew	F	JHB	1988/08/26
RJ22	Ashkenazi	Jew	M	JHB	1988/08/26
RJ23	Ashkenazi	Jew	M	JHB	1988/08/26
RJ24	Ashkenazi	Jew	M	JHB	1988/08/26
RJ25	Ashkenazi	Jew	F	JHB	1988/08/26
RJ26	Ashkenazi	Jew	F	JHB	1988/08/26
RJ27	Ashkenazi	Jew	M	JHB	1988/08/26
RJ28	Ashkenazi	Jew	M	JHB	1988/08/26
RJ29	Ashkenazi	Jew	F	JHB	1988/08/26
RJ30	Ashkenazi	Jew	M	JHB	1988/08/26
RJ31	Ashkenazi	Jew	F	JHB	1989/05/15
RJ32	Ashkenazi	Jew	M	JHB	1989/05/16
RJ33	Ashkenazi	Jew	M	JHB	1989/05/18
RJ34	Ashkenazi	Jew	F	JHB	1989/05/18
RJ35	Ashkenazi	Jew	M	JHB	1989/05/19
RJ36	Ashkenazi	Jew	F	JHB	1989/05/19
RJ37	Ashkenazi	Jew	M	JHB	1989/05/25
RJ39	Ashkenazi	Jew	M	JHB	1989/05/26
RJ40	Ashkenazi	Jew	M	JHB	1989/05/30
RJ41	Ashkenazi	Jew	F	JHB	1989/05/30

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
RJ42	Ashkenazi Jew		F	JHB	1989/05/30
RJ43	Ashkenazi Jew		M	JHB	1989/05/30
RJ44	Ashkenazi Jew		M	JHB	1989/06/02
RJ45	Ashkenazi Jew		M	JHB	1989/06/15
RJ46	Ashkenazi Jew		M	JHB	1989/06/15
RJ47	Ashkenazi Jew		F	JHB	1989/06/15
RJ48	Ashkenazi Jew		M	JHB	1989/06/15
RJ49	Ashkenazi Jew		F	JHB	1989/06/21
RJ52	Ashkenazi Jew		F	JHB	1989/06/21
RJ53	Ashkenazi Jew		F	JHB	1987/02/04
RJ54	Ashkenazi Jew		M	JHB	1987/02/04
RJ55	Ashkenazi Jew		M	JHB	1987/02/04
RJ56	Ashkenazi Jew		F	JHB	1987/02/04
RJ57	Ashkenazi Jew		M	JHB	1987/02/04
RJ58	Ashkenazi Jew		M	JHB	1987/02/04
RJ16	Ashkenazi Jew		M	JHB	1988/08/26
RI01	Indian	Tamil	M	JHB	HBTS 23615/88
RI02	Indian	Hindi	M	JHB	HBTS 23590/88
RI04	Indian	Moslem	F	JHB	HBTS 34248/88
RI06	Indian	Hindi	M	JHB	HBTS 14279/88
RI07	Indian	Hindi	M	JHB	HBTS 14263/88
RI08	Indian	Hindi	M	JHB	HBTS 14261/88
RI09	Indian	Hindi	M	JHB	HBTS 4768/88
RI11	Indian	Moslem	M	JHB	HBTS 4896/88
RI12	Indian	Moslem	F	JHB	HBTS 4858/88
RI13	Indian	Hindi	M	JHB	HBTS 63763/88
RI14	Indian	Hindi	M	JHB	HBTS 34289/88
RI15	Indian	Tamil	F	JHB	HBTS 63777/88
RI16	Indian	Moslem	M	JHB	HBTS 14309/88
RI17	Indian	Hindi	M	JHB	HBTS 44713/88
RI18	Indian	Hindi	M	JHB	HBTS 63900/88
RI19	Indian	Moslem	M	JHB	HBTS 55136/88
RI20	Indian	Moslem	M	JHB	HBTS 63936/88
RI21	Indian	Moslem	F	JHB	HBTS 23943/88
RI22	Indian	Moslem	M	JHB	HBTS 14542/88
RI23	Indian	Moslem	F	JHB	HBTS 55319/88
RI24	Indian	Hindi	M	JHB	HBTS 64071/88
RI25	Indian	Tamil	M	JHB	HBTS 64076/88
RI26	Indian	Tamil	M	JHB	HBTS 14602/88
RI27	Indian	Tamil	M	JHB	HBTS 05230/88
RI28	Indian	Tamil	M	JHB	HBTS 05232/88
RI29	Indian	Moslem	M	JHB	HBTS 4719/88
RI30	Indian	Hindi	M	JHB	HBTS 14728/88
RI31	Indian	Hindi	M	JHB	HBTS 64278/88
RI32	Indian	Moslem	M	JHB	HBTS 45064/88
RI33	Indian	Tamil	M	JHB	HBTS 45066/88
RI34	Indian	Hindi	M	JHB	HBTS 34874/88
RI35	Indian	Tamil	M	JHB	HBTS 24361/88
RI36	Indian	Hindi	M	JHB	HBTS 55595/88
RI37	Indian	Moslem	M	JHB	HBTS 15085/88
RI38	Indian	Hindi	M	JHB	HBTS 24652/88
RI39	Indian	Hindi	M	JHB	HBTS 24662/88
RI40	Indian	Moslem	M	JHB	HBTS 35213/88
RI41	Indian	Hindi	F	JHB	HBTS 55694/88
RI43	Indian	Moslem	M	JHB	HBTS 24837/88
RI44	Indian	Hindi	M	JHB	HBTS 14212/88
RI45	Indian	Moslem	M	JHB	HBTS 45917/88
RI46	Indian	Hindi	M	JHB	HBTS 45916/88
RI47	Indian	Moslem	M	JHB	HBTS 45915/88
RI48	Indian	Tamil	M	JHB	HBTS 64989/88
RI49	Indian	Hindi	M	JHB	HBTS 46009/88
RI50	Indian	Tamil	M	JHB	HBTS 64983/88

Sample I.D.	Ethnic group	Language group	Sex	Place collected	Laboratory reference No.
RI51	Indian	Tamil	M	JHB	HBTS 64998/88
RI52	Indian	Tamil	M	JHB	HBTS 64982/88
RI53	Indian	Moslem	M	JHB	HBTS 64946/88
RI54	Indian	Hindi	F	JHB	SAIMR/88
RAC01	"White"	Afrikaans	M	JHB	36/89
RAC02	"White"	Afrikaans	F	JHB	11/89
RAC03	"White"	Afrikaans	F	JHB	84/89
RAC04	"White"	Afrikaans	M	JHB	11/89
RAC05	"White"	Afrikaans	M	JHB	84/89
RAC06	"White"	Afrikaans	M	JHB	89/89
RAC07	"White"	Afrikaans	F	JHB	104/89
RAC08	"White"	Afrikaans	F	JHB	99/89
RAC09	"White"	Afrikaans	M	JHB	272/88
RAC10	"White"	Afrikaans	M	JHB	355/89
RAC11	"White"	Afrikaans	M	JHB	4/89
RAC12	"White"	Afrikaans	F	JHB	36/89
RAC13	"White"	Afrikaans	M	JHB	78/89
RAC14	"White"	Afrikaans	M	JHB	104/89
RAC15	"White"	Afrikaans	M	JHB	32/89
RAC16	"White"	Afrikaans	F	JHB	74/89
RAC17	"White"	Afrikaans	F	JHB	17/89
RAC18	"White"	Afrikaans	M	JHB	357/88
RAC19	"White"	Afrikaans	F	JHB	328/88
RAC20	"White"	Afrikaans	F	JHB	341/88
RAC21	"White"	Afrikaans	M	JHB	70/89
RAC22	"White"	Afrikaans	F	JHB	16/89
RAC23	"White"	Afrikaans	M	JHB	337/88
RAC24	"White"	Afrikaans	F	JHB	-
RAC25	"White"	Afrikaans	M	JHB	74/89
RAC26	"White"	Afrikaans	M	JHB	332/88
RAC27	"White"	Afrikaans	F	JHB	95/89
RAC28	"White"	Afrikaans	M	JHB	16/89
RAC29	"White"	Afrikaans	M	JHB	8/89
RAC30	"White"	Afrikaans	M	JHB	86/89
RAC31	"White"	Afrikaans	M	JHB	3/89
RAC32	"White"	Afrikaans	F	JHB	106/89
RAC33	"White"	Afrikaans	F	JHB	357/88
RAC34	"White"	Afrikaans	M	JHB	328/88
RAC35	"White"	Afrikaans	M	JHB	341/88
RAC36	"White"	Afrikaans	M	JHB	24/89
RAC37	"White"	Afrikaans	M	JHB	95/89
RAC39	"White"	Afrikaans	M	JHB	8/89
RAC40	"White"	Afrikaans	M	JHB	-
RAC41	"White"	Afrikaans	M	JHB	106/89
RAC42	"White"	Afrikaans	M	JHB	53/89
RAC43	"White"	Afrikaans	F	JHB	86/89
RAC44	"White"	Afrikaans	F	JHB	53/89
RAC45	"White"	Afrikaans	F	JHB	89/89
RAC46	"White"	Afrikaans	M	JHB	17/89
RAC47	"White"	Afrikaans	F	JHB	340/88
RAC48	"White"	Afrikaans	F	JHB	330/88
RAC49	"White"	Afrikaans	F	JHB	18/89
RAC50	"White"	Afrikaans	F	JHB	SAIMR/87
RAC51	"White"	Afrikaans	F	JHB	SAIMR/88
RAC52	"White"	Afrikaans	M	JHB	SAIMR/88

APPENDIX 2: Recipes and solutions

Analytical grade chemicals were purchased to prepare all solutions. The following stock solutions were made which were used to make up other solutions. Unless otherwise stated, water used was first distilled.

STOCK SOLUTIONS

1 M Tris-HCl (pH 7.5, 7.8 & 8.0):
Mix 121.1 g Tris Base (Sigma) in 800 ml of water. Adjust pH by adding concentrated HCl to that required, adjust volume to 1 liter, then autoclave.

0.5 M EDTA (pH 8.0):
Mix 186.1 g of EDTA with 800 ml of water. While stirring adjust the pH with a 10N NaOH solution and when the EDTA is completely dissolved, adjust the volume to 1 liter.

10% SDS:
Weigh 50 g of solid SDS and transfer to an autoclaved 500 ml bottle. Add sterile distilled water to 500 ml and stir over low heat until completely dissolved.

1 M $MgCl_2$:
Dissolve 101.7 g $MgCl_2$ in a final volume of 500 ml.
Aliquot in 100 ml volumes in bottles and autoclave.

5 M NaCl:
Dissolve 292.2 g of NaCl in a final volume of 1 liter.

3 M Sodium acetate (pH 5.2):
Dissolve 204 g of sodium acetate in 400 ml of water. Adjust volume to 500 ml and autoclave.

1 M DDT:
Dissolve 3.09 g of DDT in 20 ml of 0.01 M sodium acetate (pH 5.2) and dispense into 1 ml aliquots and store at -20°C.

1 M Glycine-KOH:
Prepare a 1 M glycine-potassium hydroxide stock (pH 9.4) by dissolving 7.5 g of glycine in 100 ml of water and adjusting the pH to 9.4 with 1 M KOH.

20X SSC:
Dissolve 175.3 g of NaCl and 88.2 g of sodium citrate in 800 ml of water. Adjust volume to 1 liter.

20X SSPE:
Dissolve 174 g of NaCl, 27.6g of $NaH_2PO_4.H_2O$ and 7.4 g of EDTA in 800 ml of water. Adjust the pH to 7.4 with NaOH and make up to 1 liter.

Ethidium bromide (10 mg/ml):
Add 1 g of ethidium bromide to 100 ml of water. Stir for several hours until completely dissolved and store wrapped in aliminium foil at 4°C.

Proteinase K (10 mg/ml):
Add 10 ml sterile distilled water to 10 mg proteinase K. When dissolved dispense into aliquots and store at -20°C.

BSA (10 mg/ml):
Dissolve 1 g of BSA (Sigma) in 10 ml water. Aliquot into 1 ml amounts and store at -20°C.

Salmon sperm DNA (10 mg/ml):
Allow salmon spern DNA (ratio of 1 g salmon sperm in 10 ml water) to dissolve properly by shirring for several hours. Shear thoroughly by passing through a gauge 22 needle. Denature by boiling for 10 min and cool rapidly on ice. Store at -20°C until required. Heat denature again before use.

Spermidine (1 M):
Dissolve 10 g of spermidine in 10 ml water and mix thoroughly. Store at -20°C and dilute to 0.1 M for use.

TCA (100%):
To a bottle containing 500 mg of TCA, add 227 ml of water. This yields a 100% (w/v) TCA solution.

100X Denharts:
Mix together 10 g Ficoll, 10 g polyvinylpyrrolidone and 10 g BSA in 500 ml of water. Dispense into 50 ml aliquots and store at 20°C.

Formamide:
Deionize required volume of formamide with 5% amberlite MB-6 resin by stirring for 30 min. Filter two times through Whatman No. 1 filter paper. Dispense into 50 ml Nunc tubes and store at -20°C.

SOLUTIONS

TE (pH 8.0):
Make up 1 liter by adding 10 ml of10 mM Tris.Cl (pH 8.0) and 2 ml of 1 mM EDTA (pH 8.0).

Triton X/NaCl solution:
Make 1 liter with 0.2% Triton-X 100 and 0.9% NaCl.

Genomic lysing buffer:
To 400 ml of water add 420 g of urea (7 M) and stir with low heat to help dissolve the urea. Then add 60 ml of 5 M NaCl (0.3 M), 20 ml of 0.5 M EDTA, pH 8.0 (10 mM) and 10 ml of 1 M Tris.HCl, pH 7.5 (10 mM) and adjust the volume to 1 liter.

Phenol (pH 8.0):
High purity phenol from Merck was used. Solid phenol (500 g) was dissolved by heating in a 65°C waterbath. Several extractions with an equal volume of 0.1 M Tris.HCl (pH 8.0) was needed to saturate the phenol with Tris to a pH of 8.0. This phenol mix was reduced by the addition of 0.5% 8-hydroxyquinoline and stored wrapped in aluminium foil at 4°C.

Chloroform:Isoamyl (24:1):
Chloroform used consisted of a chloroform: isoamyl mix in the ratio 24:1.

10X TBE:
Dissolve 108 g of Tris base and 55 g boric acid in 800 Ml of water. Add 40 ml of 0.5 M EDTA (pH 8.0) and make up to 1 liter. The pH of this stock should by approximately 8.3. Dilute as needed to obtain 1X working solutions (adjust pH to 8.0 if necessary) which has a final concentration of 89 mM Tris, 89 mM boric acid and 2 mM EDTA.

10X Ficoll loading dye:
A mix of 0.25% bromophenol blue, 0.25% xylene cyanol and 25% Ficoll (type 400) was made in distilled water and aliquoted in 1 ml eppendorf tubes and stored at 4°C.

Lambda DNA markers (invisible):
Mix together 100 µl ficoll dye, 1 µg denatured salmon spern DNA and 16.7 ng commercial lambda DNA molecular weight markers (Boeringer Mannheim), *Hind*III digested alone (marker III) and *Hind*III and *Eco*RI digested (marker IV) and make up to a volume of 1 ml.

Denaturing solution (1.5 M NaCl/0.5 M NaOH):
Appropriate volumes of denaturing solution were prepared by weighing out the equivalent of 1M NaCl and 0.5 M NaOH and dissolving in distilled water.

3 M Sodium acetate (pH 5.5):
To 500 ml of water add 408.2 g of sodium acetate and mix thoroughly by stirring. Adjust pH to 5.5 with glacial acetic acid and add water to 1 liter. Usually 5 liters of this was prepared and used as neutralization solution for Biodyne membranes.

mtDNA Extraction buffer:
Mix together 10 mM Tris.HCl (pH 7.8), 10 mM NaCl and 1 mM EDTA (pH 8.0) and make up to the desired volume.

MS Buffer (mannitol/sucrose):
Mix together 210 mM mannitol, 70 mM sucrose and 1 mM Tris.HCl (pH 7.5) and make up to desired volume with distilled water.

Biodyne prehybridization mix:
Make up a 50 ml mix by adding together 15 ml of 20X SSC, 5 ml of 100X Denhardts solution, 5 ml 1 M $NaPO_4$ buffer (ph 6.5), 3 ml denatured salmon sperm DNA and make up to 50 ml with water. Add 50 ml of deionized formamide before using.

Biodyne hybridization mix:
Same as the prehybridization mix, except that only 1.5 ml of salmon spern DNA is added to the mix. In addition, the mix contains 2.5 ml of 50% dextran sulphate.

Hybond-N hybridization mix:
Make up 25 ml of 2X Hybond-N mix by adding together 12.5 ml SSPE solution, 2.5 ml of 100X Denhardts solution, 2.5 ml of 10% SDS and 7.5 ml of water. Dilute to 1X with an equal volume of deionized formamide.

dNTP's:
Add 1-2 ml of 0.2M Tris.HCL to 25 mg quantities of each nucleotide (dATP, dCTP, dGTP and dTTP, Sigma) until a pH of 7.0 is obtained. Determine spectrophotometrically the concentration of these nucleotides using the following wavelengths and extiction coefficients:

dNTP	wavelength (nm)	extinction coefficient
dATP	259	1.54×10^4
dCTP	271	9.1×10^3
dGTP	253	1.37×10^4
dTTP	260	7.4×10^3

From these solutions prepare 10 mM stock solutions by diluting in double distilled water and store at -20°C. Mix equal amounts of all four nucleotides to prepare a 10X mix and store at -20°C.

10 μM primers:
Determine spectrophotometrically the concentration of the stock primers from an appropriate dilution and calculate its concentration in μM. Dilute this with double distilled water to 10 μM and store at -20°C. Store concentrated primers at -70°C.

10X Polynucleotide kinase buffer:
Mix together 500 mM Tris.HCl (pH 7.5), 100 mM $MgCl_2$, 50 mM DTT and 500 μg/ml BSA and make up to 1 ml and store at -20°C. The ATP that is required for driving the reaction is added freshly to the reaction mix to increase the longevity of the buffer.

10X λ exonuclease buffer:
Mix together 670 mM glycine-KOH (pH 9.4), 25 mM $MgCl_2$ and 500 μg/ml BSA and add water to 1 ml.

5X T7 buffer:
Composed of 40 mM Tris.HCl (pH 7.5), 1 mM EDTA (pH 8), 1 mM DTT, 50 mM NaCl and 0.1 mg/ml BSA. Store aliquots of this buffer at -70°C and discard after use.

5X Anneal buffer:
This buffer consists of 200 mM Tris.HCl (pH 7.5) 100 mM $MgCl_2$ and 250 mM NaCl.

Sequencing Stop solution:
Stop solution containing 95% deionized formamide, 20 mM EDTA (pH 8), 0.05% bromophenol blue and 0.05% xylene cyanol FF was dissolved in water and 1 ml aliquots were stored at -20°C.

10% Ammonium persulphate (APS):
Dissolve 1 g of APS in 10 ml of water and store at 4°C for a maximun of 2 weeks.

6% Acrylamide working solution:
Add 237 ml distilled water to 150 g premixed acrylamide/bis 19:1 (Biorad) and dissolve. This yeilds a 40% stock solution that should be stored wrapped in aluminium foil at 4°C. The 6% working solution is prepared by dissolving 420.4 g urea (Sigma) in 500 ml of water and then adding 100 ml of 10X TBE, 150 ml of 40% acrylamide stock and adjusting the volume to 1 liter. Store in a dark bottle at 4°C.

10% glacial acetic acid/ 10% methanol:
Make 4 liters by mixing 400 ml glacial acetic acid and 400 ml of methanol in 3 200 ml of distlled water. Store in a dark bottle.

APPENDIX 3: The restriction enzyme patterns (morphs) and the corresponding mtDNA type (mitotype) obtained for 795 individuals from southern Africa. MtDNA types were derived by combining the enyzme morphs obtained with the restriction enzyme *Hpa*I, *Bam*HI, *Hae*II, *Msp*I, *Ava*II and *Hinc*II (in this order) using the method described by Johnson *et al.* (1983) and Blanc *et al.* (1983).

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Zu38	3	1	1	7	1	2	132-2
Zu39	3	1	1	1	3	2	2-2
Zu41	3	1	9	1	3	2	133-2
Zu42	3	1	1	1	1	2	7-2
Zu43	3	1	1	1	3	2	2-2
Zu47	3	1	1	1	3	2	2-2
Zu48	2	1	1	1	2	2	21-2
Zu50	2	1	1	1	1	2	1-2
Zu14	3	1	1	1	3	2	2-2
Zu16	2	1	1	1	1	2	1-2
Zu18	3	1	1	1	1	2	7-2
Zu19	2	1	1	1	1	2	1-2
Zu20	3	1	1	1	1	2	7-2
Zu22	3	1	1	2	2	2	3-2
Zu24	3	1	1	2	2	2	3-2
Zu25	2	1	1	1	1	2	1-2
Zu26	3	1	1	1	3	2	2-2
Zu27	3	1	1	1	1	2	7-2
Zu28	2	1	1	1	1	2	1-2
Zu49	3	1	1	1	1	2	7-2
Zu50	1	1	1	1	1	2	8-2
Zu51	3	1	1	1	1	2	7-2
Zu52	3	1	1	2	2	2	3-2
Zu56	3	1	1	1	3	2	2-2
Zu58	3	1	1	1	3	2	2-2
Zu67	3	1	1	1	1	2	7-2
Zu69	3	1	1	1	1	2	7-2
Zu72	3	1	1	1	3	2	2-2
Zu73	3	1	1	1	1	2	7-2
Zu77	3	1	1	1	3	2	2-2
Sw01	2	1	1	1	3	2	47-2
Sw02	3	1	1	2	2	2	3-2
Sw03	2	1	1	1	1	2	1-2
Sw04	2	1	1	1	1	2	1-2
Sw05	2	1	1	1	1	2	1-2
Sw06	2	1	1	1	1	2	1-2
Sw07	3	1	1	1	3	2	2-2
Sw11	3	1	1	1	3	2	2-2
Sw12	2	1	6	1	1	2	26-2
Sw14	3	1	1	2	5	2	5-2
Sw15	1	1	1	1	1	1	1-1
Sw16	2	1	1	1	1	2	1-2
Sw17	2	1	1	1	1	9	1-9
Sw18	3	1	1	1	5	2	31-2
Sw20	2	1	1	1	1	2	1-2
Sw21	3	1	1	1	2	2	10-2
Sw22	3	1	1	1	1	2	7-2
Sw23	3	1	1	1	1	2	7-2
Sw25	2	1	1	1	1	2	1-2
Sw26	3	1	1	1	1	2	7-2
Sw27	2	1	1	1	1	2	1-2
Sw28	3	1	1	1	3	2	2-2
Sw30	2	1	1	1	1	2	1-2
Sw31	3	1	1	1	1	2	7-2
Sw32	3	1	1	1	3	2	2-2
Sw33	2	1	1	1	1	2	1-2
Sw34	3	1	1	1	1	2	7-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Sw36	2	1	1	1	1	2	1-2
Sw37	3	1	1	1	3	2	2-2
Sw38	3	1	1	1	3	2	2-2
Sw39	3	1	1	1	1	2	7-2
Sw40	3	1	1	1	2	2	10-2
Sw41	3	1	1	1	1	2	7-2
Sw43	3	1	1	1	3	2	2-2
Sw08	2	1	8	1	1	2	45-2
Sw10	2	1	1	1	1	2	1-2
Sw51	3	1	1	1	1	2	7-2
Sw52	2	1	1	2	1	2	124-2
Sw53	3	1	9	1	1	2	130-2
Sw54	3	1	1	1	1	2	7-2
Sw48	3	1	1	1	1	2	7-2
Xh01	3	1	1	1	3	2	2-2
Xh02	2	1	2	1	1	2	6-2
Xh03	3	1	1	1	5	2	31-2
Xh04	3	1	1	1	1	2	7-2
Xh05	2	1	1	1	1	2	1-2
Xh07	3	1	1	2	5	2	5-2
Xh10	3	1	1	1	1	2	7-2
Xh11	3	1	1	2	2	2	3-2
Xh20	3	1	1	1	1	2	7-2
Xh22	2	1	1	1	1	2	1-2
Xh14	3	1	1	1	1	2	7-2
Xh15	3	1	1	1	1	2	7-2
Xh16	3	1	2	1	2	2	131-2
Xh17	3	1	1	1	1	2	7-2
Xh18	2	1	1	1	1	2	1-2
Xh24	1	1	1	1	1	2	8-2
Xh27	3	1	1	1	3	2	2-2
Xh28	3	1	1	1	3	2	2-2
Tso01	2	1	1	1	1	2	1-2
Tso02	3	1	1	1	1	2	7-2
Tso03	3	1	11	1	1	2	125-2
Tso05	3	1	1	1	3	2	2-2
Tso07	3	1	1	1	3	2	2-2
Tso09	2	1	1	1	1	2	1-2
Tso10	3	1	1	1	3	2	2-2
Tso11	3	1	1	1	3	2	2-2
Tso12	3	1	1	1	1	2	7-2
Tso15	3	1	1	1	3	2	2-2
Tso16	3	1	1	1	3	2	2-2
Tso17	2	1	1	1	1	2	1-2
Tso18	3	1	1	1	1	2	7-2
Tso19	2	1	1	1	1	2	1-2
Tso20	2	1	1	1	1	2	1-2
Tso21	3	1	1	1	1	2	7-2
Tso22	3	1	1	1	1	2	7-2
Tso23	3	1	1	1	1	2	7-2
Tso24	3	1	1	1	3	2	2-2
Tso25	2	1	1	1	1	2	1-2
Tso20	3	1	1	1	3	2	2-2
Tso27	3	1	1	1	3	2	2-2
Tso28	3	1	1	1	1	2	7-2
Tso29	3	1	1	1	3	2	2-2
Tso30	2	1	1	1	1	2	1-2
Tso31	2	1	1	1	1	2	1-2
Tso32	3	1	1	2	2	2	3-2
Tso06	3	1	1	1	3	2	2-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Tso08	3	1	1	1	1	2	7-2
Tso37	3	1	1	1	3	2	2-2
Tso13	2	1	1	13	1	2	70-2
Tso39	3	1	1	1	3	2	2-2
Tso41	2	1	1	1	6	2	56-2
Tso42	2	1	1	1	1	2	1-2
Sot01	2	1	1	1	1	2	1-2
Sot02	3	1	1	1	3	2	2-2
Sot05	3	1	1	1	3	2	2-2
Sot06	2	1	1	1	1	2	1-2
Sot12	3	1	1	1	1	2	7-2
Sot26	3	1	1	1	3	2	2-2
Sot27	3	1	1	2	5	2	5-2
Sot29	2	1	1	1	1	2	1-2
Sot30	2	1	1	1	1	2	1-2
Sot44	3	1	1	2	2	2	3-2
Sot45	3	1	1	2	2	2	3-2
Sot46	3	1	1	1	1	2	7-2
Sot48	3	1	1	2	2	2	3-2
Sot03	3	1	1	2	2	2	3-2
Sot53	3	1	1	1	3	2	2-2
Sot54	3	1	1	1	1	2	7-2
Sot07	3	1	1	1	3	2	2-2
Sot11	2	1	1	1	2	2	21-2
Sot17	3	1	1	1	1	2	7-2
Sot19	3	1	9	1	1	2	130-2
Sot32	3	1	1	1	1	2	7-2
Sot33	3	1	1	1	1	2	7-2
Sot35	2	1	1	1	1	2	1-2
Sot36	3	1	1	13	2	2	135-2
Sot37	3	1	1	1	1	2	7-2
Sot39	3	1	1	1	2	2	10-2
Sot40	3	1	1	1	2	2	10-2
Sot43	2	1	1	2	1	2	124-2
Sot13	3	1	1	1	2	2	10-2
Ped02	3	1	1	1	5	2	31-2
Ped03	3	1	1	2	2	2	3-2
Ped04	3	1	1	1	3	2	2-2
Ped05	3	1	1	1	1	2	7-2
Ped06	3	1	1	1	1	2	7-2
Ped07	3	1	1	1	1	2	7-2
Ped08	3	1	1	1	1	2	7-2
Ped09	3	1	1	1	3	2	2-2
Ped17	3	1	1	1	1	2	7-2
Ped19	3	1	1	1	1	2	7-2
Ped20	3	1	1	2	5	2	5-2
Ped01	3	1	1	1	3	2	2-2
Ped11	3	1	1	1	1	2	7-2
Ped15	3	1	1	1	3	2	2-2
Ped21	3	1	1	2	2	2	3-2
Ped22	3	1	1	1	1	2	7-2
Ped23	2	1	1	1	1	2	1-2
Ped24	3	1	1	1	1	2	7-2
Ped25	3	1	1	1	1	2	7-2
Ped26	3	1	1	1	1	2	7-2
Ped27	2	1	1	1	1	2	1-2
Ped29	3	1	1	2	2	2	3-2
Ped30	3	1	1	2	2	2	3-2
Tsw01	2	1	1	1	3	2	47-2
Tsw02	2	1	1	1	1	2	1-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Tsw05	3	1	1	1	5	2	31-2
Tsw06	2	1	1	1	2	2	21-2
Tsw07	3	1	1	1	1	2	7-2
Tsw22	3	1	1	1	3	2	2-2
Tsw23	3	1	1	1	3	2	2-2
Tsw24	3	1	1	2	5	2	5-2
Tsw25	2	1	1	1	1	2	1-2
Tsw31	2	1	1	1	1	2	1-2
Tsw27	3	1	2	1	1	2	131-2
Tsw32	3	1	1	1	1	2	7-2
Tsw33	3	1	1	1	1	2	7-2
Tsw34	3	1	1	2	1	2	134-2
Tsw35	3	1	1	1	3	2	2-2
Ve01	3	1	1	1	3	2	2-2
Ve02	3	1	1	1	1	2	7-2
Ve03	2	1	1	1	1	2	1-2
Ve04	3	1	1	1	3	2	2-2
Ve05	3	1	1	1	1	2	7-2
Ve06	3	1	1	1	3	2	2-2
Ve07	3	1	1	1	3	2	2-2
Ve08	3	1	1	13	1	2	118-2
Ve09	3	1	1	2	2	2	3-2
Ve10	3	1	1	2	2	2	3-2
Ve11	3	1	1	1	1	2	7-2
Ve12	3	1	1	1	1	2	7-2
Ve13	3	1	1	1	1	2	7-2
Ve14	3	1	1	1	3	2	2-2
Ve15	3	1	1	1	3	2	2-2
Ve16	3	1	1	1	1	2	7-2
Ve17	3	1	1	2	2	2	3-2
Ve18	3	1	1	1	2	2	10-2
Ve20	2	1	1	1	3	2	47-2
Ve21	3	1	1	1	2	2	10-2
Ve22	3	1	1	1	3	2	2-2
Ve24	2	1	1	1	1	2	1-2
Ve26	2	1	8	1	1	2	45-2
Ve27	2	1	8	1	3	2	122-2
Ve28	3	1	1	1	2	2	10-2
Ve29	3	1	1	2	3	2	33-2
Ve31	3	1	1	13	3	2	123-2
Ve32	3	1	1	1	1	2	7-2
Ve33	3	1	1	1	2	2	10-2
Ve34	3	1	1	2	3	2	33-2
Ve35	3	1	1	1	3	2	2-2
Lem01	3	1	1	2	2	2	3-2
Lem05	3	1	1	2	5	2	5-2
Lem06	3	1	1	1	2	2	10-2
Lem07	3	1	1	2	2	2	7-2
Lem08	3	1	1	1	5	2	31-2
Lem10	3	1	1	1	1	2	7-2
Lem11	2	1	1	1	2	2	21-2
Lem12	3	1	1	1	3	2	2-2
Lem14	3	1	1	1	2	2	10-2
Lem16	3	1	1	1	2	2	10-2
Lem19	3	1	1	2	2	2	3-2
Lem20	3	1	1	1	3	2	2-2
Lem21	3	1	1	1	1	2	7-2
Lem24	3	1	1	1	5	2	31-2
Lem27	3	1	1	1	1	2	7-2
Lem31	3	1	1	1	1	2	7-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Lem32	2	1	1	1	1	2	1-2
Lem33	3	1	1	1	1	2	7-2
Lem35	2	1	1	1	1	2	1-2
Lem37	3	1	1	1	3	2	2-2
Lem38	3	1	1	2	31	2	115-2
Lem41	2	1	1	1	1	2	1-2
Lem42	3	1	1	1	3	2	2-2
Lem43	3	1	1	1	1	2	7-2
Lem45	3	1	1	2	2	2	3-2
Lem46	3	1	1	2	2	2	3-2
Lem47	3	1	1	1	1	2	7-2
Lem49	3	1	1	1	1	2	7-2
Lem53	3	1	1	1	3	2	2-2
Lem54	3	1	1	1	1	2	7-2
Lem55	3	1	1	1	1	2	7-2
Lem57	3	1	1	1	1	2	7-2
Lem59	2	1	1	1	1	2	1-2
Lem60	2	1	1	1	1	2	1-2
Lem30	3	1	1	1	3	2	2-2
Lem61	3	1	1	1	1	2	7-2
Lem70	3	1	1	1	1	2	7-2
Lem71	2	1	1	1	1	2	1-2
Lem75	3	1	1	1	3	2	2-2
Lem78	3	1	1	1	1	2	7-2
Lem79	3	1	8	1	1	2	136-2
Lem80	3	1	1	1	1	2	7-2
Lem81	2	1	8	1	1	2	45-2
Lem82	3	1	1	1	3	2	2-2
Lem83	3	1	1	1	1	2	7-2
Lem84	3	1	1	2	11	2	137-2
Lem85	3	1	1	1	3	2	2-2
Lem87	3	1	11	3	1	2	118-2
Lem91	3	1	1	1	1	2	7-2
Lem92	3	1	1	1	3	2	2-2
Lem94	2	1	1	1	1	2	1-2
Lem95	3	1	1	1	3	2	2-2
Lem98	3	1	1	1	1	2	7-2
He002	2	1	1	1	2	2	21-2
He003	2	1	1	1	2	2	21-2
He004	2	1	1	1	2	2	21-2
He006	2	1	1	1	1	2	1-2
He008	2	1	1	1	2	2	21-2
He009	2	1	1	1	1	2	1-2
He010	2	1	1	1	2	2	21-2
He012	2	1	1	1	2	5	21-5
He016	2	1	1	1	2	2	21-2
He017	2	1	1	1	2	2	21-2
He018	2	1	1	1	1	2	1-2
He019	2	1	1	1	1	2	1-2
He020	2	1	1	1	2	2	21-2
He021	2	1	1	1	2	2	21-2
He022	2	1	1	1	2	2	21-2
He023	3	1	1	2	5	2	5-2
He025	2	1	1	1	1	2	1-2
He027	2	1	1	1	2	2	21-2
He028	2	1	1	1	2	2	21-2
He029	2	1	1	1	2	2	21-2
He030	2	1	1	1	1	2	1-2
He031	2	1	1	1	1	2	1-2
He032	2	1	1	1	1	2	1-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
He033	2	1	1	1	1	2	1-2
He034	2	1	1	1	2	2	21-2
He068	2	1	1	1	2	2	21-2
He069	2	1	1	1	2	2	21-2
He070	2	1	1	1	2	2	21-2
He071	3	1	1	2	2	2	3-2
He072	2	1	1	1	1	2	1-2
He073	2	1	1	1	2	2	21-2
He076	2	1	1	1	2	2	21-2
He077	2	1	1	1	1	2	1-2
He078	2	1	1	1	2	2	21-2
He079	2	1	1	1	1	2	1-2
He080	2	1	1	1	2	2	21-2
He082	2	1	1	1	1	2	1-2
He083	3	1	1	1	2	2	10-2
He084	2	1	1	1	2	2	21-2
He085	2	1	1	1	1	2	1-2
He091	2	1	1	1	2	2	21-2
He093	2	1	1	1	1	2	1-2
He094	2	1	1	1	1	2	1-2
He095	3	1	1	1	15	2	117-2
He096	2	1	1	1	2	2	116-2
He097	2	1	1	1	2	2	21-2
He098	2	1	1	1	1	2	1-2
He099	3	1	1	2	2	2	3-2
He100	2	1	1	1	2	2	21-2
He101	2	1	1	1	2	2	21-2
He102	3	1	1	2	5	2	22-2
He103	3	1	1	2	2	2	3-2
He105	2	1	1	1	2	2	21-2
He106	2	1	1	1	1	2	1-2
Da01	3	1	1	1	1	5	5-2
Da07	2	1	1	1	2	2	21-2
Da10	3	1	1	13	1	2	118-2
Da11	2	1	1	1	1	2	1-2
Da12	2	1	1	1	2	2	1-2
Da13	2	1	1	1	2	2	1-2
Da14	3	1	1	13	1	2	118-2
Da15	3	1	9	1	8	2	138-2
Da16	2	1	1	2	2	2	116-2
Da17	3	1	9	1	8	2	138-2
Da18	2	1	1	1	1	2	1-2
Da19	2	1	1	1	1	2	1-2
Da20	2	1	1	1	2	2	21-2
Da21	3	1	1	2	2	2	3-2
Da22	2	1	1	1	1	2	1-2
Da23	3	1	1	13	1	2	118-2
Da24	2	1	1	1	1	2	1-2
Da25	2	1	1	1	2	2	2-2
Da26	2	1	1	1	2	2	2-2
Da27	2	1	1	1	1	2	1-2
Da28	2	1	1	1	2	2	21-2
Da29	3	1	1	13	1	2	118-2
Da30	2	1	1	1	2	2	21-2
Da31	2	1	1	1	1	2	1-2
Da32	2	1	1	1	2	2	21-2
Da33	3	1	9	1	8	2	138-2
Da34	2	1	1	1	2	2	21-2
Da35	3	1	1	1	8	2	119-2
Da36	2	1	1	1	2	2	21-2
Da37	3	1	9	1	8	2	138-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Da51	2	1	1	1	1	2	1-2
Da52	2	1	1	2	1	2	124-2
Da53	2	1	1	1	1	2	1-2
Da54	2	1	1	1	2	2	21-2
Da58	2	1	1	2	1	2	124-2
Da62	2	1	1	1	2	2	21-2
Da65	3	1	1	2	1	2	134-2
Da66	3	1	9	1	8	2	138-2
Da67	2	1	1	2	2	2	116-2
Da68	2	1	1	1	2	2	21-2
Da71	3	1	1	1	11	2	30-2
Da72	3	1	1	1	1	2	7-2
Da75	2	1	1	2	1	2	124-2
Ov01	2	1	1	1	1	2	1-2
Ov02	2	1	1	5	1	2	129-2
Ov03	3	1	1	1	3	2	2-2
Ov04	3	1	1	1	1	2	7-2
Ov05	3	1	1	1	1	2	7-2
Ov06	3	1	1	1	1	2	7-2
Ov07	3	1	1	1	1	2	7-2
Ov08	3	1	1	1	1	2	7-2
Ov09	2	1	1	1	1	2	1-2
Ov10	3	1	1	1	1	2	7-2
Ov12	2	1	1	1	1	2	1-2
Ov13	3	1	1	1	3	2	2-2
Ov14	3	1	1	1	3	2	2-2
Ov15	3	1	1	2	2	2	3-2
Ov16	3	1	1	1	1	2	7-2
Ov17	3	1	1	1	1	2	7-2
Ov18	3	1	1	1	1	2	7-2
Ov19	3	1	1	1	1	2	7-2
Ov20	2	1	1	1	2	2	21-2
Ov21	3	1	1	1	1	2	7-2
Ov22	3	1	1	1	1	2	7-2
Ov23	5	1	1	1	1	2	25-2
Om010	2	1	1	1	1	2	1-2
Om019	3	1	1	2	2	2	3-2
Om024	3	1	1	1	3	2	2-2
Om031	3	1	1	1	1	2	7-2
Om033	2	1	1	1	1	2	1-2
Om034	3	1	1	2	2	2	3-2
Om035	3	1	1	1	1	2	7-2
Om038	3	1	1	1	1	2	7-2
Om039	3	1	1	1	3	2	2-2
Om042	3	1	1	1	1	2	7-2
Om054	2	1	1	1	1	2	1-2
Om101	3	1	1	3	2	2	4-2
Om103	3	1	1	3	2	2	4-2
Om104	3	1	1	3	2	2	4-2
Om105	3	1	1	3	2	2	4-2
Om106	3	1	1	3	2	2	4-2
Om107	3	1	1	3	2	2	4-2
Om108	3	1	1	3	2	2	4-2
Om109	3	1	1	3	2	2	4-2
Om110	2	1	1	1	2	2	21-2
Om111	3	1	1	3	2	2	4-2
Om112	3	1	1	3	2	9	4-9
Om113	3	1	1	1	2	2	10-2
Om114	3	1	1	3	2	2	4-2
Om115	3	1	1	1	1	2	7-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Om116	2	1	1	1	2	2	21-2
Om117	3	1	1	1	2	2	10-2
Om118	3	1	1	2	5	2	5-2
Om119	3	1	1	1	2	2	10-2
Om120	3	1	1	3	2	2	4-2
Om123	3	1	1	3	2	2	4-2
Om124	3	1	1	3	2	2	4-2
Om125	3	1	1	1	1	2	7-2
Om126	3	1	1	3	2	2	4-2
Om127	3	1	1	3	2	2	4-2
Om128	3	1	1	2	2	2	3-2
Om129	3	1	1	1	2	2	10-2
Om131	3	1	1	3	2	2	4-2
Om132	3	1	1	1	2	2	10-2
Om133	3	1	1	1	1	2	7-2
Om134	3	1	1	1	1	2	7-2
Om135	3	1	1	3	2	2	4-2
Om136	3	1	1	1	1	2	7-2
Om137	3	1	1	2	2	2	3-2
Om138	3	1	1	1	1	2	7-2
Om139	3	1	1	2	2	2	3-2
Om140	3	1	1	2	2	2	3-2
Om141	3	1	1	3	2	2	4-2
Om142	3	1	1	1	1	2	7-2
Om143	3	1	1	2	2	2	3-2
Om144	3	1	1	12	1	2	113-2
Om145	3	1	1	1	1	2	7-2
Om147	3	1	1	3	2	2	4-2
Om148	3	1	1	3	2	2	4-2
Om149	3	1	1	1	1	2	7-2
Om150	3	1	1	2	2	2	3-2
Jm202	3	1	1	1	2	2	10-2
Jm31	3	1	1	1	1	2	7-2
Jm278	3	1	1	2	2	2	3-2
Na01	3	1	1	2	2	2	3-2
Na02	3	1	1	2	2	2	3-2
Na03	3	1	1	1	1	2	7-2
Na04	2	1	1	1	1	2	1-2
Na05	3	1	1	2	2	2	3-2
Na06	3	1	1	2	2	2	3-2
Na09	2	1	1	1	2	2	21-2
Na10	2	1	1	1	1	2	1-2
Na11	2	1	1	1	2	2	21-2
Na14	3	1	1	1	2	2	10-2
Na15	2	1	1	1	2	2	21-2
Na16	3	1	1	2	2	2	3-2
Na20	3	1	1	2	2	2	3-2
Na23	3	1	1	2	5	2	5-2
Na24	2	1	1	1	1	2	1-2
Na25	3	1	1	2	5	2	5-2
Na26	3	1	1	2	2	2	3-2
Na28	3	1	1	2	2	2	3-2
Na07	3	1	1	2	2	2	3-2
Na08	3	1	1	2	2	2	3-2
Na12	3	1	1	2	2	2	3-3
Na13	3	1	1	2	2	2	3-2
Na17	3	1	1	2	2	2	3-2
Na18	3	1	1	2	2	2	3-2
Na19	3	1	1	2	2	2	3-2
Na21	3	1	1	1	2	2	10-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Na22	3	1	1	2	2	2	3-2
Na27	3	1	1	1	2	2	10-2
Na29	3	1	1	2	2	2	3-2
Na31	2	1	1	1	2	2	21-2
Na32	3	1	1	2	2	2	3-2
Na34	3	1	1	1	1	2	7-2
Na35	3	1	1	2	2	2	3-2
Na36	3	1	1	2	2	2	3-2
Na37	3	1	1	2	2	2	3-2
Na38	3	1	1	2	2	2	3-2
Na39	2	1	1	1	2	2	21-2
Na40	3	1	1	2	2	2	3-2
Na41	7	1	6	13	1	2	139-2
Na42	2	1	1	1	2	2	21-2
Na43	3	1	2	2	2	2	3-2
Na44	3	1	1	1	6	2	140-2
Na45	3	1	1	2	5	2	5-2
Na46	7	1	6	13	1	2	139-2
Na47	3	1	1	2	2	2	3-2
Na48	3	1	1	1	1	2	7-2
Rv01	3	1	1	1	2	2	10-2
Rv02	3	1	1	1	2	2	10-2
Rv03	3	1	1	1	1	5	7-2
Rv04	3	1	1	2	2	5	3-2
Rv05	3	1	1	2	2	2	3-2
Rv06	3	1	1	2	2	2	3-2
Rv07	3	1	1	2	2	2	3-2
Rv08	3	1	1	2	2	2	3-2
Rv09	3	1	1	2	2	2	3-2
Rv10	3	1	1	2	2	2	3-2
Rv11	3	1	1	2	2	2	3-2
Rv12	3	1	1	2	5	2	5-2
Rv13	3	1	1	2	2	2	3-2
Rv14	2	1	1	2	1	2	124-2
Rv15	2	1	1	1	1	2	1-2
Rv16	3	1	1	2	2	5	3-5
Rv17	2	1	1	2	2	2	116-2
Rv18	3	1	1	2	2	2	3-2
Rv19	3	1	1	2	2	2	3-2
Rv20	2	1	1	1	1	2	1-2
Rv21	3	1	1	2	2	2	3-2
Rv22	3	1	1	2	2	2	3-2
Rv23	3	1	1	2	5	2	5-2
Rv24	3	1	1	2	5	2	5-2
Rv25	3	1	1	2	2	2	3-2
Rv26	3	1	1	2	2	2	3-2
Rv27	3	1	1	2	5	5	5-5
Rv28	3	1	1	2	2	2	3-2
Rv29	3	1	1	2	2	2	3-2
Rv30	3	1	1	2	5	2	5-2
Rv31	3	1	1	2	2	2	3-2
Rv32	3	1	1	2	2	2	3-2
Rv33	3	1	1	2	2	2	3-2
Rv35	3	1	1	1	1	2	7-2
Rv37	3	1	1	2	2	2	3-2
Rc01	3	1	1	2	2	2	3-2
Rc02	3	1	1	1	3	2	2-2
Rc03	3	1	1	1	3	2	2-2
Rc04	3	1	1	2	2	2	3-2
Rc05	2	1	1	1	3	2	47-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Rc06	3	1	1	1	1	2	7-2
Rc07	3	1	1	2	3	2	33-2
Rc08	3	1	1	2	2	2	3-2
Rc09	2	1	1	1	1	1	1-2
Rc10	3	1	1	2	2	2	3-2
Rc11	3	1	1	1	3	2	2-2
Rc12	3	1	1	1	1	2	7-2
Rc13	3	1	1	2	2	2	3-2
Rc14	2	1	1	1	1	2	1-2
Rc15	3	1	1	2	5	2	5-2
Rc16	2	1	1	1	1	2	1-2
Rc17	3	1	1	2	5	2	5-2
Rc18	3	1	1	2	2	2	3-2
Rc19	3	1	1	2	2	2	3-2
Rc20	2	1	1	1	3	2	47-2
Rc21	2	1	1	1	1	2	1-2
Rc22	3	1	1	1	3	2	2-2
Rc23	3	1	1	1	3	2	2-2
Rc24	3	1	1	1	2	2	10-2
Rc25	3	1	1	2	2	2	3-2
Rc26	2	1	1	1	1	2	1-2
Rc27	3	1	1	2	5	2	5-2
Rc28	3	1	1	2	2	2	3-2
Rc29	3	1	1	2	5	2	5-2
Rc30	3	1	1	2	5	2	5-2
Rc31	3	1	1	2	2	2	3-2
Rc32	2	1	1	1	1	2	1-2
Rc33	3	1	1	2	2	2	3-2
Rc34	3	1	1	2	5	2	5-2
Rc35	3	1	1	2	2	2	3-2
Rc37	3	1	1	3	32	2	120-2
Rc39	3	1	1	1	2	2	10-2
Rc40	3	1	1	2	5	2	5-2
Rc41	3	1	1	1	3	2	2-2
Rc42	3	1	1	1	1	2	7-2
Rc43	3	1	1	2	2	2	3-2
Rc44	2	1	1	1	1	2	1-2
Rc45	3	1	1	2	2	2	3-2
Rc46	3	1	1	1	1	2	7-2
Rc47	2	1	1	1	1	2	1-2
Rc49	3	1	1	2	5	2	5-2
Rc50	3	1	1	1	1	2	7-2
Rc51	2	1	1	1	1	2	1-2
Rc52	3	1	1	2	2	2	3-2
Rc53	3	1	1	2	5	2	5-2
Rc54	3	1	1	2	5	2	5-2
Rc55	3	1	1	2	5	2	5-2
Rc57	3	1	1	1	1	2	7-2
Rc58	3	1	1	2	2	2	3-2
Rc59	3	1	1	2	2	2	3-2
Rc38	2	1	1	1	1	2	1-2
Rc36	3	1	1	1	1	2	7-2
Rc66	2	1	1	1	3	2	47-2
Rc67	3	1	1	1	3	2	2-2
Rc68	3	1	1	2	31	2	115-2
Rc69	2	1	1	1	1	2	1-2
Rc70	3	1	1	1	3	2	2-2
Rc71	3	1	1	1	1	2	7-2
Rc72	3	1	1	2	5	2	5-2
Rc73	3	1	1	2	33	2	121-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
Rc74	1	1	2	1	1	2	9-2
Rc75	2	1	1	1	1	2	1-2
Rc76	1	1	1	1	1	2	1-2
Rc77	2	1	1	1	3	2	47-2
Rc78	3	1	1	1	1	2	7-2
Rc79	2	1	1	1	1	2	1-2
RW01	2	1	1	1	1	2	1-2
RW02	2	1	1	1	1	2	1-2
RW03	2	1	1	1	3	2	47-2
RW04	2	1	1	1	1	2	1-2
RW05	2	1	1	1	3	2	47-2
RW06	2	1	1	1	1	2	1-2
RW07	2	1	1	1	1	2	1-2
RW08	2	2	3	1	5	2	11-2
RW09	2	1	1	1	1	2	1-2
RW11	2	1	1	1	1	2	1-2
RW13	2	1	1	1	1	2	1-2
RW14	2	1	1	1	1	2	1-2
RW15	2	1	1	1	1	2	1-2
RW16	2	1	1	1	1	2	1-2
RW17	2	1	1	4	3	2	44-2
RW18	2	1	1	1	1	2	1-2
RW21	2	1	1	1	1	2	1-2
RW27	2	1	2	1	1	2	6-2
RW30	2	1	1	1	1	2	1-2
RW31	2	1	1	1	1	2	1-2
RW33	2	1	1	1	3	2	47-2
RW35	2	1	1	1	1	2	1-2
RW37	2	1	1	1	1	2	1-2
RW40	2	1	1	1	1	2	1-2
RW42	2	1	1	1	1	2	1-2
RW43	2	1	1	4	3	2	44-2
RW44	2	1	1	4	3	2	44-2
RW45	2	1	2	1	1	2	6-2
RW46	2	1	1	1	1	2	1-2
RW50	2	1	1	1	1	2	1-2
RW51	2	1	1	4	1	2	28-2
RW53	2	2	3	1	5	2	11-2
RW54	2	1	1	1	1	2	1-2
RW55	2	2	2	1	5	2	114-2
RW56	2	1	1	1	1	2	1-2
RW41	2	1	2	1	1	2	6-2
RJ01	2	1	2	1	1	2	6-2
RJ02	2	1	2	1	1	2	6-2
RJ03	2	1	1	1	1	2	1-2
RJ04	2	1	2	1	1	2	6-2
RJ05	2	1	2	17	1	2	127-2
RJ06	2	1	2	1	1	2	6-2
RJ07	2	1	2	1	1	2	2-2
RJ08	2	1	1	1	1	2	1-2
RJ09	2	1	1	1	1	2	1-2
RJ10	2	1	1	1	1	2	1-2
RJ11	2	1	1	1	1	2	1-2
RJ12	2	1	1	1	1	2	1-2
RJ13	2	1	2	1	1	2	6-2
RJ14	2	1	1	1	1	2	1-2
RJ15	2	1	1	1	1	2	1-2
RJ17	2	2	3	1	1	2	128-2
RJ18	2	1	2	1	1	2	6-2
RJ19	2	1	2	1	1	2	6-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
RJ20	2	1	1	1	1	2	1-2
RJ21	2	1	1	1	1	2	1-2
RJ22	2	1	1	1	1	2	1-2
RJ23	2	1	1	1	1	2	1-2
RJ24	2	1	2	1	1	2	6-2
RJ25	2	1	2	1	1	2	6-2
RJ26	2	1	1	1	1	2	1-2
RJ27	2	1	1	17	1	2	126-2
RJ28	2	1	1	1	1	2	1-2
RJ29	2	1	1	1	1	2	1-2
RJ30	2	1	1	1	1	2	1-2
RJ31	2	1	1	1	1	2	1-2
RJ32	2	1	1	1	1	2	1-2
RJ33	2	1	2	1	1	2	6-2
RJ34	2	1	2	1	1	2	6-2
RJ35	2	1	1	1	1	2	1-2
RJ36	2	1	1	1	1	2	1-2
RJ37	2	1	1	1	1	2	1-2
RJ39	2	1	2	1	1	2	6-2
RJ40	2	1	1	1	1	2	1-2
RJ41	2	3	1	4	9	2	18-2
RJ42	2	1	1	1	1	2	1-2
RJ43	2	1	1	1	1	2	1-2
RJ44	2	1	2	1	1	2	6-2
RJ45	2	1	1	1	1	2	1-2
RJ46	2	1	2	1	1	2	6-2
RJ47	2	1	1	1	1	2	1-2
RJ48	2	1	2	1	1	2	6-2
RJ49	2	1	1	1	1	2	1-2
RJ52	2	1	1	1	1	2	1-2
RJ53	2	1	1	1	1	2	1-2
RJ54	2	1	1	1	1	2	1-2
RJ55	2	1	1	1	1	2	1-2
RJ56	2	1	2	1	1	2	6-2
RJ57	2	1	1	1	1	2	1-2
RJ58	2	1	1	1	1	2	1-2
RJ16	2	1	1	1	1	2	1-2
RI01	2	1	1	1	1	2	1-2
RI02	2	1	1	1	1	2	1-2
RI04	3	1	1	2	2	2	3-2
RI06	2	1	1	1	1	2	1-2
RI07	2	1	1	1	6	2	56-2
RI08	2	1	1	1	1	2	1-2
RI09	2	1	1	1	1	2	1-2
RI11	2	1	1	1	1	2	1-2
RI12	3	1	1	1	3	2	2-2
RI13	2	1	1	1	3	2	47-2
RI14	3	1	1	1	5	2	31-2
RI15	2	1	1	1	1	2	1-2
RI16	2	1	1	1	1	2	1-2
RI17	2	1	1	1	1	2	1-2
RI18	2	1	1	1	1	2	1-2
RI19	2	1	1	1	1	2	1-2
RI20	2	1	1	1	1	2	1-2
RI21	2	1	1	1	1	2	1-2
RI22	2	1	1	1	1	2	1-2
RI23	2	1	1	1	1	2	1-2
RI24	2	1	1	1	1	2	1-2
RI25	2	1	1	1	1	2	1-2
RI26	2	1	1	1	1	2	1-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
RI27	2	1	1	1	1	2	1-2
RI28	2	1	1	1	1	2	1-2
RI29	2	1	1	1	2	2	21-2
RI30	2	1	1	1	1	2	1-2
RI31	2	1	1	1	1	2	1-2
RI32	2	1	1	1	1	2	1-2
RI33	2	1	1	1	1	2	1-2
RI34	2	1	1	1	1	2	1-2
RI35	2	1	1	1	1	2	1-2
RI36	2	1	1	1	1	2	1-2
RI37	2	1	1	1	1	2	1-2
RI38	2	1	1	1	1	2	1-2
RI39	2	1	1	1	1	2	1-2
RI40	2	1	1	1	1	2	1-2
RI41	2	1	1	1	1	2	1-2
RI43	2	1	1	1	1	2	1-2
RI44	2	1	1	1	1	2	1-2
RI45	2	1	1	1	1	2	1-2
RI46	2	1	1	1	1	2	1-2
RI47	2	1	1	1	1	2	1-2
RI48	2	1	1	1	1	2	1-2
RI49	2	1	1	1	1	2	1-2
RI50	2	1	1	1	1	2	1-2
RI51	2	1	1	1	1	2	1-2
RI52	2	1	1	1	1	2	1-2
RI53	2	1	1	1	1	2	1-2
RI54	2	1	1	1	1	2	1-2
RAC01	2	1	1	1	1	2	1-2
RAC02	2	1	1	1	1	2	1-2
RAC03	2	1	1	1	1	2	1-2
RAC04	2	1	1	1	2	2	21-2
RAC05	2	1	1	1	1	2	1-2
RAC06	2	1	1	1	1	2	1-2
RAC07	2	1	1	1	1	2	1-2
RAC08	2	1	1	1	1	2	1-2
RAC09	2	1	1	1	1	2	1-2
RAC10	2	1	1	1	1	2	1-2
RAC11	2	1	1	1	1	2	1-2
RAC12	2	2	1	4	3	2	141-2
RAC13	2	2	1	4	3	2	131-2
RAC14	2	1	1	1	1	2	1-2
RAC15	2	1	1	1	1	2	1-2
RAC16	2	1	1	1	1	2	1-2
RAC17	2	1	1	1	1	2	1-2
RAC18	2	1	1	1	1	2	1-2
RAC19	2	1	1	1	1	2	1-2
RAC20	2	1	1	1	1	2	1-2
RAC21	2	1	1	1	1	2	1-2
RAC22	2	1	1	1	1	2	1-2
RAC23	2	1	1	1	1	2	1-2
RAC24	2	1	1	1	1	2	1-2
RAC25	2	2	1	4	3	2	141-2
RAC26	2	1	1	1	1	2	1-2
RAC27	2	1	1	1	1	2	1-2
RAC28	2	1	1	1	1	2	1-2
RAC29	2	1	1	4	3	2	44-2
RAC30	2	1	1	1	1	2	1-2
RAC31	2	1	1	1	1	2	1-2
RAC32	2	1	1	1	2	2	21-2
RAC33	2	1	1	1	1	2	1-2

Sample	*Hpa*I	*Bam*HI	*Hae*II	*Msp*I	*Ava*II	*Hinc*II	mtDNA type
RAC34	2	1	1	1	1	2	1-2
RAC35	2	1	1	1	1	2	1-2
RAC36	2	1	1	1	1	2	1-2
RAC3	2	1	1	1	1	2	1-2
RAC39	2	1	11	1	1	2	51-2
RAC40	3	2	11	1	2	2	142-2
RAC41	2	1	1	1	3	2	47-2
RAC42	2	1	1	1	1	2	1-2
RAC43	2	1	1	1	1	2	1-2
RAC44	2	1	1	1	1	2	1-2
RAC45	2	1	1	1	1	2	1-2
RAC46	2	1	1	1	1	2	1-2
RAC47	2	1	2	1	1	2	6-2
RAC48	2	1	1	1	1	2	1-2
RAC49	2	1	1	1	1	2	1-2
RAC50	2	1	1	1	1	2	1-2
RAC51	2	1	1	1	1	2	1-2
RAC52	2	1	1	1	1	2	1-2

APPENDIX IV: Mitochondrial DNA control region sequence variation in southern African populations. Approximately 750 base pairs of sequence data were obtained by sequencing the two hypervariable regions contained within the control region of the mtDNA molecule from 144 individuals. From pairwise comparisons of sequences, 129 unique mtDNA types were identified and the nucleotide positions of the informative sites are listed according to the system in which the first base of the control region is nucleotide position 1 (see Fig. 1.2). Dots (.) indicate identity with the reference sequence "ANDE" (Anderson *et al.* 1981) which is shown at the top of each page. Missing data are indicated by a question mark (?) and the minus sign (-) indicates the positions where length variations (either insertions or deletions) were found.

POSITION				1	1	1	1	1	1	1	1	1	1	1	1	1	1
	4	7	8	0	0	1	2	2	2	3	4	4	4	4	5	5	5	5
	6	0	1	1	6	7	2	5	8	0	4	5	6	8	1	4	6	7
ANDE	C	T	C	T	G	T	G	C	T	G	C	C	C	T	C	C	T	C
XHOSA01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA03	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.
XHOSA05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA07	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA08	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
XHOSA14	.	.	.	.	A	.	.	T	.	.	.	T	.	C	.	.	.	.
XHOSA15	.	.	.	.	A	.	.	T	.	.	.	T	.	C	.	.	.	.
ZULU30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU33	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU34	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU36	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU37	.	.	.	C	.	.	.	.	.	.	?	?	?	?	?	?	?	?
ZULU38	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	C	T
VE31	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
VE27	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LE78	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.
LE84	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
TSW27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TSW34	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SW63	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TSC3	.	.	.	.	A	.	.	T	.	.	.	T	.	C	.	.	.	.
OV2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT03	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.
SOT05	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
SOT06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT17	?	?	?	?	?	?	?	?	?	?	.	T	.	C	.	.	.	.
SOT18	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT12	.	.	.	.	A	.	.	T	.	.	.	T	.	C	.	.	.	.
SOT26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT27	.	C	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT32	.	C	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
SOT33	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
SOT36	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	.	T
SOT37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT40	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT44	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT45	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT46	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
SOT47	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
SOT11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM19	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
OM28	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
OM29	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	T
OM30	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	T
OM31	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
OM33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM34	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	T

	4	7	9	10	10	11	12	12	12	13	14	14	14	14	15	15	15	15
	8	0	1	1	8	7	2	5	9	0	4	5	6	9	1	4	6	7
ANDE	C	T	C	T	G	T	G	C	T	G	C	C	C	T	C	C	T	C
OM38	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
OM41	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM42	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
OM43	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
OM44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM45	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM46	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
OM48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM54	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
OM82	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM101	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM103	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM104	T	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.
OM105	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	T
OM106	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
OM107	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM108	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM109	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM111	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM112	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM114	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM120	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM123	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM124	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM129	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM127	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM135	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM141	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM147	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
OM148	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
NA02	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
NA03	.	.	.	.	A	.	.	.	.	.	.	.	.	C	T	.	.	.
NA04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA06	?	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA07	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA09	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA10	?	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA11	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA13	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
NA14	?	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA20	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
NA21	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
NA22	?	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
NA23	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.
NA28	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
NA29	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
NA41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE09	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	4 6	7 0	9 1	10 11	10 16	11 17	12 22	12 26	12 29	13 30	14 44	14 45	14 46	14 49	15 51	16 64	16 66	16 67
ANDE	C	T	C	T	G	T	G	C	T	G	C	C	C	T	C	C	T	C
HE12	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE26	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
HE27	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE28	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE29	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA01	.	.	.	.	A	.	.	.	C	A	.	.	.	.	.	.	.	.
DAMA02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA04	.	.	.	.	A	C	.	.	.	.	.	.	.	.	.	.	.	.
DAMA07	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA08	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA09	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA10	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA11	.	.	.	.	A	.	.	.	.	.	.	.	.	C	T	.	.	.
DAMA12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA14	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA17	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA18	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA20	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA23	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA25	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA26	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA27	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA15	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
RC37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RC73	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
RV14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ17	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AFRIK40	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
AFRIK13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
	6	6	6	6	6	6	7	8	8	9	9	9	9	0	0	0	0	1
	2	3	5	6	7	8	0	1	7	0	1	2	6	0	1	2	8	2
ANDE	C	C	C	C	T	C	C	A	T	A	G	C	C	C	C	T	A	C
XHOSA01	.	.	.	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
XHOSA03	.	.	.	?	?	?	.	.	.	G	.	.	.	.	T	.	G	.
XHOSA05	.	.	.	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
XHOSA06	.	.	.	.	.	.	.	?	?	?	?	?	?	.	?	?	?	?
XHOSA07	.	.	.	.	.	.	.	.	.	G	.	.	.	.	T	.	G	.
XHOSA08	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
XHOSA14	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
XHOSA16	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
ZULU30	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
ZULU31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU33	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
ZULU34	.	.	T	.	C	.	.	.	.	G	.	.	.	.	T	.	G	.
ZULU35	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
ZULU36	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
ZULU37	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
ZULU38	.	.	.	T	C	.	.	.	.	.	.	.	.	.	T	.	.	.
VE31	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
VE27	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
LE78	.	.	.	T	C	.	.	.	.	.	.	.	.	.	T	.	.	.
LE84	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
TSW27	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
TSW34	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
SW63	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
TSO3	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OV2	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
SOT01	.	.	.	.	C	T	.	.	.	.	.	.	.	.	T	.	.	.
SOT02	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
SOT03	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
SOT05	?	?	?	?	?	?	T	.	.	.	.	.	.	.	.	.	.	.
SOT06	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
SOT17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.
SOT19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
SOT12	.	.	T	G	C	.	.	.	.	.	.	.	.	.	T	.	G	.
SOT26	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
SOT27	.	.	T	.	C	.	.	G	.	.	.	.	.	.	T	.	G	.
SOT32	.	.	T	A	C	.	.	.	.	.	.	.	.	.	T	.	G	.
SOT33	.	.	T	.	C	.	.	.	.	G	.	.	.	.	T	.	G	.
SOT36	.	.	.	?	?	?	.	.	.	.	.	.	T	.	.	.	.	.
SOT37	.	.	.	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
SOT40	.	.	T	.	C	.	.	.	C	.	.	.	.	.	T	.	.	.
SOT44	.	.	T	.	C	.	.	.	.	.	A	.	.	.	T	.	.	.
SOT45	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	G	T
SOT46	.	.	.	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
SOT47	.	.	T	.	C	.	.	.	.	G	.	.	.	.	T	.	G	.
SOT11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
OM18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T
OM28	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.
OM29	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OM30	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OM31	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.
OM33	.	T	.	.	.	.	.	.	C	.	.	.	.	.	T	.	.	.
OM34	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.

	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
	6	6	6	6	6	6	7	8	8	9	9	9	9	0	0	0	0	1
	2	3	5	6	7	8	0	1	7	0	1	2	6	0	1	2	8	2
ANDE	C	C	C	C	T	C	C	A	T	A	G	C	C	C	C	T	A	C
OM38	.	.	.	.	.	.	.	.	C	.	.	T	.	.	T	.	G	.
OM41	.	.	T	.	C	.	.	.	C	.	.	T	.	.	T	.	G	.
OM42	.	.	T	.	C	.	.	.	C	.	.	T	.	.	T	.	G	.
OM43	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
OM44	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
OM45	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
OM46	.	.	T	.	C	.	.	.	C	.	.	T	.	.	T	.	G	.
OM48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
OM60	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
OM64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
OM82	.	.	.	T	C	.	.	.	.	.	.	T	.	.	T	.	G	.
OM101	T	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM103	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM104	.	.	.	T	C	.	.	.	.	G	.	T	.	.	T	.	G	T
OM105	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OM106	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM107	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OM108	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OM109	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
OM111	.	.	T	.	C	.	.	.	.	.	.	T	.	.	T	.	G	T
OM112	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM114	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM120	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	T
OM123	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM124	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM126	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM127	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM135	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM141	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
OM147	.	.	T	.	C	.	.	.	C	.	.	T	.	.	T	.	G	T
OM148	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	T
NA02	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA03	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.
NA04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
NA05	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA06	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA07	.	.	?	?	?	.	.	.	.	.	.	.	.	.	T	.		.
NA09	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
NA10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA11	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
NA13	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA14	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA20	.	.	T	.	C	.	.	.	.	G	.	.	.	.	T	.	.	.
NA21	.	.	T	.	C	.	.	.	.	.	.	T	.	.	T	.	G	.
NA22	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.
NA23	.	.	.	T	C	.	.	.	.	G	.	.	.	.	T	.	G	.
NA26	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA28	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
NA41	.	.	T	.	C	.	.	.	.	.	.	T	.	.	T	.	.	.
HE06	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.
HE09	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.

	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
	6	6	6	6	6	6	7	8	8	9	9	9	9	0	0	0	0	1
	2	3	5	8	7	8	0	1	7	0	1	2	6	0	1	2	8	2
ANDE	C	C	C	C	T	C	C	A	T	A	G	C	C	C	C	T	A	C
HE12	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
HE16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
HE18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
HE23	.	.	T	.	C	.	.	.	.	G	.	.	.	.	T	.	G	.
HE26	.	.	.	.	.	.	.	.	C	.	.	.	.	.	T	.	.	.
HE27	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
HE28	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
HE29	.	.	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.
HE30	.	.	.	.	.		.	.	C	.	.	.	.	.	T	.	.	.
HE31	.	T	.	.	.		.	.	.	.	.	.	.	.	T	.	.	.
DAMA01	.	.	T	.	C		.	.	.	.	.	.	.	.	T	.	G	.
DAMA02	.	T	.	.	.		.	.	C	.	.	.	.	.	T	.	.	.
DAMA04	.	.	T	.	C		.	.	.	.	.	.	.	.	T	.	.	.
DAMA07	?	?	?	.	?		.	.	.	.	.	.	.	.	T	.	.	.
DAMA08	?	?	?	.	?		.	.	.	.	.	.	.	.	T	.	.	.
DAMA09	?	?	?	.	?		.	.	.	.	.	.	.	.	T	.	.	.
DAMA10	.	.	T	.	C	.	.	.	.	.	.	T	.	.	T	.	.	.
DAMA11	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.
DAMA12	?	?	?	.	?	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA14	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.
DAMA17	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA18	.	T	.	.	.	.	.	.	C	.	.	.	.	.	T	.	.	.
DAMA19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA21	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	G	.
DAMA23	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA24	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA25	?	?	?	.	?	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA26	?	?	?	.	?	.	.	.	.	.	.	.	.	.	T	.	.	.
DAMA27	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.
DAMA15	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
RC37	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
RC73	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
RV14	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
RJ41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ17	.	.	T	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.
RJ6	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ37	.	.	T	.	C	.	.	.	.	.	.	.	T	.	.	.	.	.
RJ24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
AFRIK40	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
AFRIK13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	1	2	2	3	3	4	4	4	4	4	5	5	6	6	6	6	7	7
	7	0	1	4	7	0	2	3	4	8	2	6	2	4	8	8	0	1
ANDE	C	C	T	C	C	C	C	A	C	C	G	C	A	C	C	C	C	A
XHOSA01	.	.	.	.	.	.	.	G	.	.	.	T	.	.	.	.	.	.
XHOSA03	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA05	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA06	?	?	?	?	?	?	?	?	?	.	.	.	.	.	.	.	.	.
XHOSA07	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA14	.	.	.	.	.		.	.	.	.	.	T	.	.	.	.	.	G
XHOSA16	.	.	.	.	.		.	.	.	.	.	T	.	.	.	.	.	.
ZULU30	.	.	C	.	.		.	.	.	.	A	T	.	.	T	.	.	.
ZULU31	.	.	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.
ZULU33	T	.	C	.	.		.	.	.	.	.	.	.	.	.	.	.	.
ZULU34	.	.	C	.	.		.	.	.	.	.	.	.	.	.	.	.	.
ZULU35	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU37	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
ZULU38	.	.	.	T	.	.	.	.	.	.	.	T	.	.	.	.	.	G
VE31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VE27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LE78	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LE84	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
TSW27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
TSW34	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
SW63	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TSO3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OV2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT02	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	T	.
SOT03	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT17	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
SOT18	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
SOT12	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	G
SOT26	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.
SOT27	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT33	.	.	C	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT36	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT37	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.
SOT40	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	G
SOT44	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT46	.	.	C	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
SOT48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT47	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT11	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
OM18	.	.	C	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
OM28	.	.	.	.	.	.	.	.	.	.	.	T	G	.	.	.	.	.
OM29	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM30	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM31	.	.	.	.	.	.	.	.	.	.	.	T	G	.	.	.	.	.
OM33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM34	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	1	2	2	3	3	4	4	4	4	4	5	5	6	6	6	6	7	7
	7	0	1	4	7	0	2	3	4	8	2	6	2	4	8	9	0	1
ANDE	C	C	T	C	C	C	C	A	C	C	G	C	A	C	C	C	C	A
OM38	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.
OM41	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.
OM42	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM43	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
OM44	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
OM45	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
OM46	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.
OM48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM60	.	.	.	.	.	.	I	.	.	T	.	T	.	.	.	.	.	.
OM64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM82	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.
OM101	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM103	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
OM104	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM105	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
OM106	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM107	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	T	.
OM108	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM109	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM111	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM112	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM114	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM120	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM123	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM124	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM126	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM127	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM136	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM141	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM147	.	T	C	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.
OM148	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA02	T	.	C	.	.	I	.	.	.	.	.	.	.	.	.	.	.	.
NA03	.	T	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA06	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA06	.	.	C	.	.	.	.	.	.	.	A	T	.	.	T	.	.	.
NA07	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
NA09	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
NA10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA11	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
NA13	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA14	.	.	C	.	.	.	.	.	.	.	A	T	.	.	T	.	.	.
NA20	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA21	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
NA22	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA23	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA28	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA29	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA41	.	.	.	.	.	.	.	C	.	.	.	T	.	A	.	T	.	.
HE06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE09	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.

	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	1	2	2	3	3	4	4	4	4	4	6	6	6	6	6	6	7	7	
	7	0	1	4	7	0	2	3	4	8	2	8	2	4	8	9	0	1	
ANDE	C	C	T	C	C	C	C	A	C	C	G	C	A	C	C	C	C	A	
HE12	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
HE16	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
HE19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
HE23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
HE25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
HE27	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
HE28	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
HE29	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
HE30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
HE31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA04	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA07	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA08	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA09	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA10	.	.	.	.	.	.	.	C	.	.	.	T	.	A	.	T	.	.	
DAMA11	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA12	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA14	.	.	.	.	.	.	.	C	.	.	.	T	.	A	.	T	.	.	
DAMA17	.	.	.	.	.	.	.	C	.	.	.	T	.	G	.	.	.	.	
DAMA18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA20	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA21	.	.	.	.	.	.	.	.	.	.	A	T	.	.	.	.	.	.	
DAMA23	.	.	.	.	.	.	.	C	.	.	.	T	.	A	.	T	.	.	
DAMA24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA25	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA26	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	
DAMA27	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DAMA15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?	
RC37	.	.	.	.	.	.	.	C	.	.	.	T	.	G	.	.	.	.	
RC73	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
RV14	.	.	.	.	.	.	.	C	.	.	.	T	.	G	.	.	.	.	
RJ41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	
RJ17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
RJ6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
RJ37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
RJ24	.	.	.	.	.	?	?	?	?	?	?	?	?	?	?	?	?	?	
AFRIK40	.	.	.	.	.	.	T	.	.	.	.	T	.	.	.	.	.	.	
AFRIK13	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	

	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
	7	7	7	7	8	8	8	8	8	8	8	0	0	3	3	3	4	4
	2	3	3	8	0	2	7	8	0	7	8	3	5	3	5	8	0	6
ANDE	C	C	C	A	A	T	A	T	A	G	C	T	C	C	T	C	T	T
XHOSA01	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
XHOSA03	.	.	.	.	.	.	.	C	.	.	.	.	.	?	?	?	?	?
XHOSA05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA06	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
XHOSA07	.	.	.	.	.	.	.	C	.	.	.	?	?	?	?	?	?	?
XHOSA08	.	.	.	.	.	.	.	C	.	.	T	.	.	.	.	.	.	.
XHOSA14	.	.	.	.	.		.	.	.	.	T	.	.	.	.	.	.	.
XHOSA15	.	.	.	.	.	.	.	C	.	.	T	.	.	.	.	.	.	.
ZULU30	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	.
ZULU31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU33	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU34	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
ZULU35	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ZULU36	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
ZULU37	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.
ZULU38	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	C
VE31	.	.	.	.	.	C	.	C	.	.	.	.	.	.	?	?	?	?
VE27	T	.	T	.	.	C	.	C	.	.	.	.	.	.	?	?	?	?
LE78	.	.	.	.	.	.	.	?	.	.	.	.	.	.	?	?	?	?
LE84	T	.	T	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
TSW27	.	.	.	.	.	.	.	.	.	.	T	.	.	.	?	?	?	?
TSW34	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
SW63	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
TSO3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
OV2	T	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
SOT01	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	?	?	?
SOT02	T	.	.	.	G	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT03	T	.	.	.	.	.	.	.	.	.	.	C	.	.	.	?	?	?
SOT05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT06	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	?	?	?
SOT17	.	.	.	.	.	.	.	C	.	.	T	.	.	.	.	?	?	?
SOT18	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	?	?	?
SOT12		.	.	.	.	.	.	C	.	.	T	.	.	.	.	?	?	?
SOT26	T	.	.	.	.	.	G	.	.	.	.	.	.	.	.	?	?	?
SOT27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT32	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	?	?	?
SOT33	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	?	?	?
SOT36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT37	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT40	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT44	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
SOT45	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	?	?	?
SOT46	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	?	?	?
SOT47	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	?	?	?
SOT11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?
OM18	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM28	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.	.	.
OM29	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM34	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3
	7	7	7	7	8	8	8	8	8	8	8	0	0	3	3	3	4	4
	2	3	3	8	0	2	7	8	0	7	8	3	6	3	6	8	0	6
ANDE	C	C	C	A	A	T	A	T	A	G	C	T	C	C	T	C	T	T
OM38	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM41	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM42	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM43	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM44	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.
OM45	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
OM46	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM48	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
OM50	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM54	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
OM82	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OM101	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM103	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM104	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
OM106	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM107	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM108	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM109	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM111	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM112	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM114	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM120	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
OM123	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM124	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM126	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM127	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM135	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM141	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM147	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
OM148	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
NA02	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	?
NA03	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	?
NA04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA05	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
NA07	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
NA08	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	?
NA10	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	?
NA11	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	?
NA13	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
NA14	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	?
NA20	.	.	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
NA21	.	.	G	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?
NA22	T	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
NA23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
NA28	T	.	.	.	.	.	.	C	.	.	.	.	.	T	?	?	?	?
NA29	T	.	.	.	.	.	.	C	.	.	.	C	.	.	?	?	?	?
NA41	T	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
HE06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
HE08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?

	27123	27338	27380	28227	28298	28301	28336	28388	33656	33836	34046
ANDE	C	C	A	A	T	A	T	A	G	C	T	C	C	T	C	T	T
HE12	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	?
HE16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
HE19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
HE23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
HE26	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	?
HE27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
HE28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
HE29	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	?
HE30	.	.	.	.	.	.	C	.	.	T	.	.	.	.	.	.	?
HE31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
DAMA01	.	.	.	.	.	.	C	.	.	.	.	.	.	C	.	.	.
DAMA02	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
DAMA04	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
DAMA07	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.
DAMA08	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.
DAMA09	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.
DAMA10	T	.	.	.	C	.	.	.	.	.	.	.	.	.	T	.	.
DAMA11	.	.	.	.	.	.	C	.	.	.	C	.	.	.	.	C	.
DAMA12	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
DAMA14	T	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	.
DAMA17	T	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.
DAMA18	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
DAMA19	.	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.
DAMA20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA21	T	.	.	G	.	.	C	.	.	.	.	.	.	.	.	.	.
DAMA23	T	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	.
DAMA24	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
DAMA25	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.
DAMA26	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.
DAMA27	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.
DAMA16	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
RC37	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	?	?
RC73	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
RV14	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
RJ41	T	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
RJ17	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
RJ6	T	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
RJ37	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
RJ24	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
AFRIK40	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?	?	?
AFRIK13	T	.	.	.	.	.	C	.	.	.	.	.	.	?	?	?	?

	3	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7
	6	2	3	3	4	4	4	5	6	7	9	9	0	0	0	0	1	2
	8	5	7	9	5	7	9	0	6	2	5	8	2	3	4	5	6	4
ANDE	G	A	G	A	A	A	G	C	C	C	G	T	C	C	T	A	C	T
XHOSA01	.	?	.	.	G	.	C	G	.	.	.	C	T	.	C	.	.	.
XHOSA03	?	G	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.
XHOSA05	.	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
XHOSA06	.	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
XHOSA07	?	?	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
XHOSA08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
XHOSA14	.	.	.	.	G	C	.	.	.	.	.	.	.	.	.	.	.	.
XHOSA16	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
ZULU30	.	G	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	A
ZULU31	.	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
ZULU33	.	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
ZULU34	A	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
ZULU35	.	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
ZULU36	.	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
ZULU37	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
ZULU38	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VE31	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VE27	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LE78	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LE84	?	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
TSW27	?	G	.	.	.	.	.	.	.	.	.	C	T	.	C	.	.	.
TSW34	?	G	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.
SW53	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TSO3	?	.	.	.	G	C	.	.	.	.	.	.	.	.	C	.	.	.
OV2	?	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
SOT01	?	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
SOT02	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
SOT03	?	?	?	?	?	?	?	?	?	?	.	C	.	.	C	.	.	.
SOT05	?	?	?	?	?	?	?	?	?	.	.	.	.	.	.	.	.	.
SOT08	?	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
SOT17	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
SOT19	?	G	.	.	.	.	.	.	.	.	A	C	.	.	C	.	.	.
SOT12	?	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
SOT26	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
SOT27	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
SOT32	?	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
SOT33	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
SOT36	?	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
SOT37	?	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
SOT40	?	.	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
SOT44	?	?	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
SOT45	?	?	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
SOT46	?	?	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT47	?	?	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
SOT11	?	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
OM18	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM28	?	G	.	.	.	.	.	.	.	.	.	.	T	T	C	.	.	.
OM29	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM30	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM31	?	G	.	.	.	.	.	.	.	.	.	.	T	T	C	.	.	.
OM33	?	G	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.
OM34	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.

	3	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7
	6	2	3	3	4	4	4	5	6	7	8	8	0	0	0	0	1	2
	8	5	7	8	5	7	8	0	6	2	5	8	2	3	4	5	6	4
ANDE	G	A	G	A	A	A	G	C	C	C	G	T	C	C	T	A	C	T
OM38	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM41	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
OM42	?	?	.	.	.	.	.	.	.	T	.	C	.	.	C	.	.	.
OM43	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
OM44	?	?	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
OM45	?	G	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.
OM46	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
OM48	?	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
OM50	?	G	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
OM54	?	G	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.
OM82	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM101	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM103	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM104	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM106	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM108	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM107	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM108	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM109	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM111	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM112	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM114	?	?	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM120	?	?	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM123	?	?	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM124	?	?	.	G	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM126	?		.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM127	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM135	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM141	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM147	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
OM148	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
NA02	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	T	.
NA03	?	G	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
NA04	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.	.	.
NA05	.	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
NA06	?	?	?	?	.	.	.	.	.	.	.	.	T	.	.	.	.	.
NA07	?	?	?	?	.	.	.	.	.	.	.	.	.	.	C	.	.	.
NA08	?	G	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
NA10	?	?	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
NA11	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
NA13	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
NA14	?	?	.	.	.	.	.	.	.	.	.	C	T	.	C	.	.	.
NA20	?	?	?	?	?	?	?	?	.	.	.	C	.	.	C	.	.	.
NA21	?	?	?	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
NA22	?	?	?	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
NA23	?	?	?	?	?	?	?	?	.	.	.	C	.	.	C	.	.	.
NA28	?	?	?	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
NA29	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
NA41	?	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
HE06	?	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
HE09	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.

	3	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7
	6	2	3	3	4	4	4	5	6	7	9	9	0	0	0	0	1	2
	8	6	7	9	5	7	9	0	6	2	5	8	2	3	4	5	6	4
ANDE	G	A	G	A	A	A	G	C	C	C	G	T	C	C	T	A	C	T
HE12	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
HE16	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
HE18	?	?	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
HE23	?	?	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
HE26	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
HE27	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
HE28	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
HE29	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
HE30	?	?	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
HE31	?	?	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.
DAMA01	?	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
DAMA02	?	G	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.
DAMA04	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
DAMA07	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
DAMA08	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
DAMA09	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
DAMA10	?	?	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
DAMA11	?	?	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
DAMA12	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
DAMA14	?		.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
DAMA17	?	?	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
DAMA18	?	?	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA19	?	?	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.
DAMA20	?	?	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
DAMA21	?	?	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
DAMA23	?	?	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
DAMA24	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA25	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA26	?	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
DAMA27	?	G	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
DAMA15	?	G	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.
RC37	?	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
RC73	.	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
RV14	.	G	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.
RJ41	?	G	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
RJ17	?	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ6	?	G	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.
RJ37	?	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
RJ24	?	G	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.
AFRIK40	?	G	.	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.
AFRIK13	?	?	?	?	.	.	.	.	.	.	.	.	.	T	.	.	.	.

	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	8	8	8
	3	3	3	3	3	4	4	4	5	5	5	5	6	7	8	0	0	1
	4	5	6	7	8	0	1	8	1	2	3	7	0	8	8	0	4	8
ANDE	C	A	G	G	C	A	A	T	C	T	A	T	G	G	T	G	G	A
XHOSA01	T	G	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.
XHOSA03	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
XHOSA05	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	G
XHOSA06	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
XHOSA07	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
XHOSA08	.	.	.	.	.	.	G	.	.	.	.	C	A	.	C	A	.	G
XHOSA14	.	.	.	A	.	.	G	.	.	.	.	.	.	.	C	A	.	G
XHOSA15	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	A	.	G
ZULU30	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
ZULU31	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	G
ZULU33	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
ZULU34	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
ZULU35	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
ZULU36	.	.	.	.	.	.	G	.	.	.	R	.	.	.	.	.	.	G
ZULU37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
ZULU38	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
VE31	T	.	.	.	.	.	C	C	.	.	.	.	.	.	.	A	.	G
VE27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LE78	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
LE84	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
TSW27	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
TSW34	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
SW53	T	.	.	.	.	.	G	C	T	.	.	.	.	.	.	A	.	.
TSO3	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	A	.	.
OV2	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
SOT01	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.
SOT02	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
SOT03	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
SOT05	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SOT06	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
SOT17	?	?	?	?	?	?	?	?	?	?	?	?	?	.	C	A	.	G
SOT19	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
SOT12	.	.	.	A	.	.	R	.	.	.	.	.	.	.	.	A	.	.
SOT26	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
SOT27	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	A	.	.
SOT32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	G
SOT33	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
SOT36	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
SOT37	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
SOT40	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	A	.	.
SOT44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
SOT45	.	.	.	.	.	.	.	C	.	C	.	.	.	.	.	A	.	.
SOT46	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
SOT47	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
SOT11	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G
OM19	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	A	.	.
OM28	T	.	.	.	.	.	.	C	T	.	.	C	.	.	.	.	.	G
OM29	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	A	.	.
OM30	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	A	.	.
OM31	T	.	.	.	.	.	.	C	T	.	.	C	.	.	.	.	.	G
OM33	.	.	.	.	.	.	G	C	.	.	G	.	A	.	.	.	.	G
OM34	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	A	.	.

	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	8	8	8
	3	3	3	3	3	4	4	4	5	5	5	5	8	7	8	0	0	1
	4	5	6	7	8	0	1	8	1	2	3	7	0	8	8	0	4	6
ANDE	C	A	G	G	C	A	A	T	C	T	A	T	G	G	T	G	G	A
OM38	.	.	.	.	.	.	G	C	T	.	.	.	.	.	.	A	.	.
OM41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
OM42	.	.	.	.	.	.	G	C	.	.	.	.	.	.	.	A	.	.
OM43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
OM44	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM45	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
OM46	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM48	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.
OM50	T	.	T	.	.	.	.	C	.	.	.	.	.	.	.	A	.	G
OM64	.	.	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.
OM82	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM101	.	.	.	.	.	.	.	C	T	C	.	.	.	.	.	A	.	.
OM103	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM104	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
OM105	.	.	.	.	.	.	G	C	.	.	.	.	.	.	.	A	.	.
OM106	.	.	.	.	.	.	G	C	T	.	.	C	A	.	.	A	.	.
OM107	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.
OM108	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM109	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
OM111	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
OM112	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.
OM114	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM120	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM123	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
OM124	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM126	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM127	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
OM135	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM141	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
OM147	.	.	.	.	.	.	G	C	T	.	.	.	A	.	.	A	.	.
OM148	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
NA02	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
NA03	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	G
NA04	.	.	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	G
NA06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
NA06	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
NA07	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
NA08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
NA10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA11	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	G
NA13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
NA14	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
NA20	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
NA21	.	.	.	.	.	.	G	C	T	.	.	.	A	.	.	A	.	.
NA22	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
NA23	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
NA28	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
NA29	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
NA41	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
HE06	.	.	.	.	.	.	G	.	.	.	G	.	A	.	.	.	.	.
HE08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	8	8	8
	3	3	3	3	3	4	4	4	6	6	6	6	6	7	8	0	0	1
	4	5	6	7	8	0	1	8	1	2	3	7	0	8	9	0	4	6
ANDE	C	A	G	G	C	A	A	T	C	T	A	T	G	G	T	G	G	A
HE12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
HE16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
HE18	.	.	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	G
HE23	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	A	.	.
HE26	.	.	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.
HE27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HE29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
HE30	.	.	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	G
HE31	.	.	.	.	.	.	G	C	.	.	G	.	.	.	.	.	.	G
DAMA01	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	G
DAMA02	.	.	.	.	.	.	G	C	.	.	G	.	A	.	.	.	.	G
DAMA04	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
DAMA07	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA09	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA10	T	.	.	.	A	.	C	C	T	.	.	.	.	.	.	A	.	G
DAMA11	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	G
DAMA12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA14	T	.	.	.	A	.	C	C	T	.	.	C	.	.	.	A	.	G
DAMA17	T	.	.	.	A	.	C	C	T	.	.	.	.	.	.	A	.	G
DAMA18	.	.	.	.	.	.	G	C	.	.	G	.	.	.	.	.	.	G
DAMA19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DAMA21	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.
DAMA23	T	.	.	.	A	.	C	C	T	.	.	C	.	.	.	A	.	G
DAMA24	.	.	.	.	.	.	G	C	.	.	G	.	A	.	.	.	.	G
DAMA25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
DAMA27	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	G
DAMA16	T	.	.	.	A	.	C	C	T	.	.	C	.	.	.	A	.	.
RC37	.	.	.	.	.	.	.	C	.	.	.	.	.	A	.	.	.	.
RC73	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
RV14	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.
RJ41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ17	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
RJ6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RJ37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
RJ24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
AFRIK40	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
AFRIK13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G

	8	8	8	8	8	8	8	8	8	8	8
	1	3	4	5	6	6	7	7	7	8	1
	8	8	7	1	4	7	4	6	8	3	6
ANDE	T	C	T	A	-	T	G	C	C	C	A
XHOSA01	.	.	.	.	.	.	.	.	.	T	.
XHOSA03	.	.	.	.	C	.	.	.	.	.	.
XHOSA05	.	.	.	.	C	.	.	.	.	.	.
XHOSA06	.	.	.	.	.	.	.	.	.	.	.
XHOSA07	.	.	.	.	.	.	.	.	.	.	?
XHOSA08	.	.	.	.	.	.	.	T	.	.	.
XHOSA14	.	.	.	.	?	.	.	.	.	.	.
XHOSA16	.	.	.	.	C	.	.	.	.	.	.
ZULU30	.	.	.	.	.	.	A	.	.	.	.
ZULU31	.	.	.	.	C	.	.	.	T	.	.
ZULU33	.	.	.	.	.	.	.	.	.	.	.
ZULU34	.	.	.	.	C	.	.	.	.	.	.
ZULU35	.	.	.	.	.	.	.	.	.	.	.
ZULU36	.	.	.	.	.	.	.	.	.	.	.
ZULU37	.	.	.	.	.	.	.	.	.	.	.
ZULU38	.	.	.	.	.	.	A	.	.	.	.
VE31	.	.	.	.	.	.	.	.	.	.	.
VE27	.	.	.	.	.	.	.	.	.	.	.
LE78	.	.	.	.	.	.	.	.	.	.	.
LE84	.	.	.	.	.	.	A	.	.	.	.
TSW27	.	.	.	.	.	.	.	.	.	.	.
TSW34	.	.	.	.	.	.	.	.	.	.	.
SW63	.	.	.	.	.	.	.	.	.	T	.
TSO3	.	.	.	.	.	.	.	.	.	.	.
OV2	.	.	.	.	.	.	A	.	.	.	.
SOT01	.	.	.	.	C	.	.	.	.	.	.
SOT02	.	.	.	.	.	.	.	.	.	.	.
SOT03	.	.	.	.	.	.	.	.	.	.	.
SOT05	.	.	.	.	.	.	.	.	.	.	.
SOT06	.	.	.	.	C	.	.	.	.	.	.
SOT17	.	.	.	.	C	.	.	.	.	.	.
SOT18	.	.	.	.	.	.	.	.	.	.	.
SOT12	.	.	.	.	C	.	.	.	.	.	.
SOT26	.	.	.	.	.	.	.	.	.	.	.
SOT27	.	.	.	.	.	.	.	.	.	.	.
SOT32	.	.	.	.	.	.	.	.	.	.	.
SOT33	.	.	.	.	.	.	.	.	.	.	.
SOT36	.	.	.	.	.	.	.	.	.	.	.
SOT37	.	.	.	.	C	.	.	.	.	.	.
SOT40	.	.	.	.	C	.	.	.	.	.	.
SOT44	C	.	.	.	C	.	.	.	.	.	.
SOT45	.	.	.	.	C	.	.	.	.	.	.
SOT46	.	.	.	.	.	.	.	.	.	.	.
SOT47	.	.	.	.	C	.	.	.	.	.	.
SOT11	.	.	.	.	.	.	.	.	.	.	.
OM18	.	.	.	.	.	.	.	.	.	.	.
OM28	.	.	.	.	.	.	.	.	.	.	.
OM29	.	.	.	.	C	.	.	.	.	.	.
OM30	.	.	.	.	C	.	.	.	.	.	.
OM31	.	.	.	.	.	.	.	.	.	.	.
OM33	.	.	.	.	C	.	.	.	.	.	.
OM34	.	.	.	.	C	.	.	.	.	.	.

	8	8	8	8	8	8	8	8	8	8	9
	1	3	4	5	6	6	7	7	7	8	1
	8	8	7	1	4	7	4	5	8	3	5
ANDE	T	C	T	A	-	T	G	C	C	C	A
OM38	.	.	.	.	.	.	.	.	.	.	.
OM41	.	.	.	.	.	.	.	.	.	.	.
OM42	.	.	.	.	C	.	.	.	.	.	.
OM43	.	.	.	.	.	.	.	.	.	.	.
OM44	.	.	.	.	C	.	.	.	.	.	.
OM45	.	.	.	.	C	.	.	.	.	.	.
OM46	.	.	.	.	.	.	.	.	.	.	.
OM48	.	.	.	.	.	.	.	.	.	.	.
OM50	.	.	.	.	.	.	.	.	.	.	G
OM64	.	.	.	.	.	.	.	.	.	.	.
OM82	.	.	.	.	.	.	.	.	.	.	.
OM101	.	.	.	.	.	.	.	.	.	.	.
OM103	.	.	.	.	.	.	.	.	.	.	.
OM104	C	.	.	.	.	.	.	.	.	.	.
OM106	.	.	.	.	.	.	.	.	.	.	.
OM108	.	.	.	.	.	.	.	.	.	.	.
OM107	.	.	.	.	.	.	.	.	.	.	.
OM108	.	.	.	.	.	.	.	.	.	.	.
OM109	C	.	.	.	.	.	.	.	.	.	.
OM111	.	.	.	.	.	.	.	.	.	.	.
OM112	.	.	.	.	.	.	.	.	.	.	.
OM114	.	.	.	.	C	.	.	.	.	.	.
OM120	.	.	.	.	.	.	.	.	.	.	.
OM123	C	.	.	.	.	.	.	.	.	.	.
OM124	C	.	.	.	.	.	.	.	.	.	.
OM126	.	.	.	.	.	.	.	.	.	.	.
OM127	C	.	.	.	.	.	.	.	.	.	.
OM135	.	.	.	.	.	.	.	.	.	.	.
OM141	.	.	.	.	C	.	.	.	.	.	.
OM147	.	.	.	.	.	.	.	.	.	.	.
OM148	.	.	.	.	.	.	.	.	.	.	.
NA02	.	.	.	.	.	.	.	.	.	.	.
NA03	.	.	C	.	.	.	.	.	.	.	.
NA04	.	.	.	.	.	.	.	.	.	.	.
NA05	.	.	.	.	.	.	.	.	.	.	.
NA06	.	.	.	.	.	.	A	.	.	.	?
NA07	.	.	.	.	.	.	.	.	.	.	.
NA08	.	.	.	.	.	.	.	.	.	.	.
NA10	.	.	.	.	.	.	.	.	.	.	.
NA11	.	.	.	.	.	.	.	.	.	.	.
NA13	.	.	.	.	.	.	.	.	.	.	.
NA14	.	.	.	.	.	.	A	.	.	.	.
NA20	.	.	.	.	C	.	.	.	.	.	.
NA21	.	.	.	.	.	.	.	.	.	.	?
NA22	.	.	.	.	.	.	.	.	.	.	?
NA23	.	.	.	.	.	.	.	.	.	.	?
NA28	.	.	.	.	.	.	.	.	.	.	.
NA29	.	.	.	.	C	.	.	.	.	.	.
NA41	.	.	.	.	.	.	.	.	.	.	.
HE08	.	.	.	.	C	.	.	.	.	.	.
HE09	.	.	.	.	.	.	.	.	.	.	.

	8	8	8	8	8	8	8	8	8	8	8
	1	3	4	5	6	6	7	7	7	8	1
	8	8	7	1	4	7	4	5	8	3	6
ANDE	T	C	T	A	-	T	G	C	C	C	A
HE12	.	.	.	.	.	.	.	.	.	.	.
HE16	.	.	.	.	.	.	.	.	.	.	.
HE18	.	.	.	.	.	.	.	.	.	.	.
HE23	.	.	.	.	?	.	.	.	.	.	.
HE26	.	.	.	.	.	.	.	.	.	.	.
HE27	.	.	.	.	.	.	.	.	.	.	.
HE28	.	.	.	.	.	.	.	.	.	.	.
HE29	.	.	.	.	.	.	.	.	.	.	.
HE30	.	.	.	.	.	.	.	.	.	.	.
HE31	.	.	.	.	.	.	.	.	.	.	.
DAMA01	.	.	.	.	.	.	.	.	.	.	.
DAMA02	.	.	.	.	C	.	.	.	.	.	.
DAMA04	.	.	.	.	.	.	.	.	.	.	.
DAMA07	.	.	.	.	.	.	.	.	.	.	.
DAMA08	.	.	.	.	.	.	.	.	.	.	.
DAMA09	.	.	.	.	.	.	.	.	.	.	.
DAMA10	.	.	.	G	.	.	A	.	.	.	.
DAMA11	.	.	C	.	.	.	.	.	.	.	.
DAMA12	.	.	.	.	.	.	.	.	.	.	.
DAMA14	.	.	.	.	T	C	.	.	.	.	.
DAMA17	.	.	.	G	.	.	A	.	.	.	.
DAMA18	.	.	.	.	.	.	.	.	.	.	.
DAMA19	.	.	.	.	.	.	.	.	.	.	.
DAMA20	.	.	.	.	.	.	.	.	.	.	.
DAMA21	.	.	.	.	.	.	.	.	.	.	.
DAMA23	.	.	.	.	.	.	.	.	.	.	.
DAMA24	.	.	.	.	.	.	.	.	.	.	.
DAMA25	.	.	.	.	.	.	.	.	.	.	.
DAMA26	.	.	.	.	.	.	.	.	.	.	.
DAMA27	.	.	C	.	.	.	.	.	.	.	.
DAMA16	.	.	.	.	.	.	.	.	.	.	.
RC37	.	.	.	.	.	.	.	.	.	.	.
RC73	.	.	.	.	.	.	.	.	.	.	.
RV14	.	.	.	.	.	.	.	.	.	.	.
RJ41	.	.	.	.	.	.	?	?	?	?	?
RJ17	.	.	.	.	.	.	A	.	.	.	.
RJ6	.	.	.	.	.	.	A	.	.	.	.
RJ37	.	.	.	.	.	.	.	.	.	.	.
RJ24	.	.	.	G	.	.	.	.	.	.	.
AFRIK40	.	.	.	.	.	.	.	.	.	.	.
AFRIK13	.	T	.	.	.	.	.	.	.	.	.