Durham University

# Durham E-Theses

# *An investigation of dermatoglyphic variation among ethnic populations in Kenya*

P. J. Rosa

**How to cite:**

Rosa, P. J. (1981) An investigation of dermatoglyphic variation among ethnic populations in Kenya. Doctoral thesis, Durham University.

**Use policy**

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/7665/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

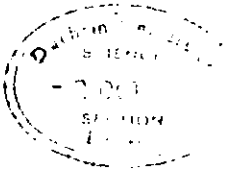
The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

# AN INVESTIGATION OF DERMATOGLYPHIC VARIATION AMONG ETHNIC POPULATIONS IN KENYA

P.J. ROSA



Thesis submitted for the degree of
Doctor of Philosophy

Department of Anthropology
University of Durham

1981

The copyright of this thesis resis with the author. No quotation from it should be published without his prior written consent and information derived from it should be acknowledged.

## ABSTRACT

In 1973 and 1975 hand prints were collected from 6,235 Kenyan schoolchildren drawn from 61 ethnic populations, of which 57 male and 53 female sampl es are suitable for analysis. The study describes dermatoglyphic variation in these samples, and attempts preliminary explanatory analysis.

Part One is concerned with background (Chapter 1) and the nature of biological populations, sampling and fieldwork (Chapter 2).

Part Two discusses methodology, and inter-sample dermatoglyphic variation. Chapter 3 introduces the approach to the descriptive and explanatory investigations, and Chapter 4 outlines methodological procedures adopted for dermatoglyphic attribute measurement, coding and the minimisation of errors. The descriptive investigation begins in Chapter 5, by examining the "basic structure" of dermatoglyphic variation in Kenya. The results show that the variation is discordant, and that no distinctive clusters are in evidence when multivariate measures of inter-sample affinity are applied in Chapter 6.

A feature of the descriptive results is the failure of the male and female series to exhibit similar taxonomic rankings for equivalent dermatoglyphic traits. The explanatory investigation in Part Three begins with an analysis of sex differences to try and account for this discordance (Chapter 7).

Chapters 8 and 9 are devoted to the examination of major types of variation. Chapter 8 explores the extent to which observed dermatoglyphic resemblance between population

samples corresponds to patterns of linguistic and ethnic affinity. In Chapter 9 ecotypic variation is considered. Specifically investigated is the question whether systematic and non-random dermatoglyphic variation exists when samples are pooled into altitude and rainfall zones. The positive results obtained are amongst the most interesting of the study.

CONTENTS


| | | Page |
|---|---|---|
| INTRODUCTION | | 1 |
| I.1 | Physical anthropology as a historical discipline.. | 1 |
| I.2 | Objectives of the research | 9 |
| I.3 | Plan of the present study | 13 |
| PART ONE: | BACKGROUND DESCRIPTION, THE POPULATIONS, SAMPLING AND FIELDWORK | 18 |
| CHAPTER 1: | INTRODUCTION TO KENYA AND ITS PEOPLES | 19 |
| 1.1 | The physical environment | 19 |
| | 1.1.1 Relief and drainage | 19 |
| | 1.1.2 Climate and rainfall | 25 |
| | 1.1.3 Ecological zones and traditional modes of livelihood | 28 |
| 1.2 | The ethnic and linguistic diversity of the peoples of Kenya and their detailed distribution | 38 |
| | 1.2.1 Peoples of the Niger-Congo family | 39 |
| | 1.2.2 The Chari-Nile family | 51 |
| | 1.2.3 The Afroasiatic family | 61 |
| | 1.2.4 Hunting, gathering and fishing groups | 65 |
| 1.3 | Historical outline of the peopling of Kenya | 69 |
| | 1.3.1 Early pre-colonial history | 69 |
| | 1.3.2 History of the Coast and the immediate pre-colonial period | 76 |
| | 1.3.3 Colonial and post-colonial history | 86 |
| 1.4 | Major demographic trends since 1948 | 94 |
| | 1.4.1 Present distribution of population: (Republic of Kenya 1969 Population Census). | 94 |
| | 1.4.2 Population growth since 1948 | 97 |
| | 1.4.3 Patterns of movement | 99 |
| CHAPTER 2: | THE POPULATION UNITS, SAMPLING AND FIELDWORK | 106 |
| 2.1 | Overview | 106 |
| 2.2 | The nature of "tribal" populations in Africa | 108 |
| 2.3 | Endogamy in African tribal populations | 119 |
| 2.4 | The extent of historical and biological heterogeneity of African peoples | 124 |
| 2.5 | Sampling design and fieldwork | 129 |
| | 2.5.1 The population units | 129 |
| | 2.5.2 Fieldwork | 130 |
| | 2.5.3 Conclusion | 140 |
| PART TWO: | THEORETICAL APPROACH, METHODOLOGY AND DESCRIPTIVE ANALYSIS | 142 |
| CHAPTER 3: | ANALYTICAL APPROACHES TO THE INVESTIGATION OF THE PRESENT MATERIAL | 143 |

| | | Page |
|---|---|---|
| 3.1 | The usefulness of conventional approaches to the description of biological differentiation between human populations | 143 |
| | 3.1.1 The ability to interpret results from multivariate analysis in the light of sampling and other methodological inadequacies | 146 |
| | 3.1.2 Choice and objectivity in the use of multivariate statistics | 148 |
| | 3.1.3 Multivariate analysis simplifies patterns of differentiation | 150 |
| | 3.1.4 Towards a new approach to the description of human biological differentiation | 151 |
| 3.2 | Types of human biological differentiation | 157 |
| | 3.2.1 Major types of biological differentiation | 157 |
| | 3.2.2 Types of concordant and discordant variation | 160 |
| 3.3 | Principals of statistical structure of human biological variation and the analysis and measurement of biological discordance and concordance | 171 |
| | 3.3.1 The statistical structure of human biological variation | 172 |
| | 3.3.2 Types of correlations between biological attributes | 174 |
| | 3.3.3 Assessing relative discordance | 178 |
| | 3.3.4 Discordance and stability | 183 |
| | 3.3.5 Implications of comparing discordance of attributes within and between populations | 185 |
| 3.4 | Overall structure of human biological differentiation | 187 |
| 3.5 | Plan for the analysis of dermatoglyphic variation in Kenya | 189 |
| CHAPTER 4: | GENERAL METHODOLOGY OF DERMATOGLYPHIC DATA MEASUREMENT | 193 |
| 4.1 | Introduction | |
| 4.2 | Ridge discontinuities and their classification | 196 |
| | 4.2.1 Characteristics of ridge flow and configuration | 197 |
| | 4.2.2 Low resolution methodologies | 198 |
| | 4.2.3 Medium resolution methodologies | 201 |
| 4.3(1) | An appraisal of the topographical and topological approaches | 202 |
| | 4.3.2 Further criticisms of the topological method | 206 |
| 4.4 | Detailed rules and methods for the measurement and classification of the dermal features considered in the study | 209 |
| | 4.4.1 Ridge counting | 209 |
| | 4.4.2 The classification of digital patterns | 212 |
| | 4.4.3 The classification of palmar patterns | 215 |
| | 4.4.4 The measurement of the axial triradii | 220 |
| | 4.4.5 The atd angle | 222 |
| | 4.4.6 The mainlines | 224 |

| | | Page |
|---|---|---|
| 4.5 | Methods of print reading | 225 |
| 4.6 | Coding and computer variables | 227 |
| | 4.6.1 The computer coding sheet | 227 |
| | 4.6.2 Detailed coding procedures | 228 |
| | 4.6.3 Missing values | 233 |
| | 4.6.4 Procedures for defining and computing new variables in SPSS | 234 |
| 4.7 | Techniques for minimising errors and file editing. | 240 |
| | 4.7.1 Inconsistency in print reading | 240 |
| | 4.7.2 The misreading of attributes | 241 |
| | 4.7.3 Copying errors, punching errors and file errors | 243 |
| | 4.7.4 Procedures for detecting and correcting errors in the data files | 244 |
| CHAPTER 5: | BASIC STRUCTURE OF DERMATOGLYPHIC DIFFERENTIATION BETWEEN SAMPLES | 247 |
| | Introduction | 247 |
| 5.1 | Presentation of facts | 248 |
| 5.2 | An outline of the range and significance of dermatoglyphic variation at the minimal ethnic unit level of resolution | 250 |
| 5.3 | The taxonomic structure of differentiation between minimal population units: an assessment of relative discordance | 253 |
| | 5.3.1 Overview | 253 |
| | 5.3.2 Detailed approach | 254 |
| | 5.3.3 Descriptive structural analyses | 269 |
| 5.4 | The extent of true taxonomic concordance between dermatoglyphic attributes considered in this study | 303 |
| | 5.4.1 Approach and methods | 304 |
| | 5.4.2 Results | 306 |
| | 5.4.3 Discussion | 308 |
| CHAPTER 6: | OVERALL STRUCTURE OF DERMATOGLYPHIC DIFFERENTIATION BETWEEN MINIMAL ETHNIC UNIT SAMPLES | 310 |
| | Introduction | 310 |
| 6.1 | Methodological and statistical approaches | 311 |
| | 6.1.1 Choice of distance statistic and the calculation of the distances | 311 |
| | 6.1.2 Choice and implementation of techniques to analyse phenetic distances | 316 |
| 6.2 | General structure of overall differentiation between the minimal population samples | 319 |
| | 6.2.1 The nature of overall structure: the evidence from phenetic distances | 320 |
| | 6.2.2 The nature of overall structure: further evidence from principal component scores | 325 |
| | 6.2.3 Discussion | 327 |

| | | Page |
|---|---|---|
| 6.3 | Detailed multivariate relationships between population samples | 330 |
| | 6.3.1 Introduction | 330 |
| | 6.3.2 Analysis | 333 |
| 6.4 | Discussion | 339 |
| PART THREE: | EXPLANATORY ANALYSIS | 343 |
| CHAPTER 7: | SEX DIFFERENCES IN THE VARIATION OF DERMATOGLYPHIC ATTRIBUTES | 344 |
| | Introduction | 344 |
| 7.1 | Sex differences at the total sample level | 345 |
| | 7.1.1 Digital attributes | 345 |
| | 7.1.2 Palmar attributes | 348 |
| 7.2 | Sex differences at the macro population level | 348 |
| | 7.2.1 Summary of inter-population trends between ethnic categories and differences between the male and female series | 349 |
| | 7.2.2 Relative sex differences in magnitude and direction in selected attributes, betwen ethnic categories | 352 |
| 7.3 | Discussion | 357 |
| CHAPTER 8: | AN EXPLANATION OF THE EXTENT TO WHICH POPULATION SAMPLES IN KENYA TEND TO VARY DERMATOGLYPHICALLY ACCORDING TO LINGUISTIC AND CULTURAL PATTERNS OF AFFINITY | 359 |
| | Introduction | 359 |
| 8.1 | The degree of correspondence of dermatoglyphic affinity and linguistic and cultural affinity at the macro population level | 361 |
| | 8.1.1 General trends of dermatoglyphic relationship between ethnic categories | 363 |
| | 8.1.2 The relationship between dermatoglyphic variation and linguistic and cultural affinity in ethnic groups | 364 |
| 8.2 | The extent of overall dermatoglyphic heterogeneity of ethnic groups: an examination of affinity between minimal ethnic population samples, within and between ethnic groups | 374 |
| | 8.2.1 Approach | 374 |
| | 8.2.2 Univariate analysis of ethnic group homogeneity and affinity from the evidence of minimal ethnic population sample differentiation | 377 |
| | 8.2.3 Direct comparison of relative homogeneity between ethnic groups as depicted by correlation between variation for minimal population samples and variables that measure ethnic affiliation | 388 |

| | | Page |
|---|---|---|
| 8.3 | Overall trends of relative homogeneity and affinity: the evidence from PCO vectors ..... | 393 |
| | 8.3.1 Structural dermatoglyphic similarity and patterns of affinity between minimal population samples within ethnic groups.... | 398 |
| CHAPTER 9: | THE EXTENT OF DERMATOGLYPHIC VARIATION EXHIBITED BY MINIMAL ETHNIC POPULATION SAMPLES.. | 417 |
| | Introduction ..... | 417 |
| 9.1 | Approach and methods ..... | 418 |
| | 9.1.1 Choice of environmental and climatic variables ..... | 418 |
| | 9.1.2 Sources for the distribution of altitude and rainfall in Kenya used in the following analysis ..... | 419 |
| | 9.1.3 Allocation of individuals and populations to altitude and rainfall zones ..... | 420 |
| | 9.1.4 Categorising altitude and rainfall zones... | 421 |
| 9.2 | Analytical approach ..... | 424 |
| | 9.2.1 The extent of the relationships between dermatoglyphic taxonomic variation and altitude and rainfall..... | 425 |
| | 9.2.2 Correlations of dermatoglyphic summary taxonomic attributes with altitude and rainfall..... | 425 |
| | 9.2.3 Correlations of dermatoglyphic non summary taxonomic attributes with altitude and rainfall..... | 427 |
| 9.3 | The nature of the relationship of the variation of dermatoglyphic taxonomic attributes with altitude and rainfall..... | 430 |
| | 9.3.1 The nature of the relationship of summary dermatoglyphic taxonomic attributes with altitude and rainfall..... | 431 |
| | 9.3.2 The nature of the relationship of the variation of non summary dermatoglyphic taxonomic attributes with altitude and rainfall..... | 433 |
| 9.4 | Summary and discussion ..... | 438 |
| CONCLUSION | ..... | 442 |
| C.1 | Problems and issues associated with the nature of populations in Kenya, and sampling ..... | 442 |
| C.2 | Problems and issues associated with the methodology of dermatoglyphic attribute measurement and of minimising errors ..... | 449 |
| C.3 | Problems and issues of analysing and describing the variation ..... | 453 |
| C.4 | Problems and explanatory issues following on from the descriptive analysis ..... | 457 |
| C.5 | Concluding remarks ..... | 465 |

| | | Page |
|---|---|---|
| TABLES: | ..... | 467 |
| Table I.1: | Major human biological surveys of Kenyan populations up to the commencement of the present survey (1973) ..... | 468 |
| Table 1.1: | The peoples of Kenya and their linguistic classification ..... | 470 |
| Table 1.2: | Population distribution and growth by tribe and nationality. (From the Kenya Population Census 1969, Vol.IV, Table 3.3) ..... | 473 |
| Table 1.3: | Percent distribution of population by whether or not born in same district of residence by tribe and nationality. (From the Kenya Population Census 1969, Vol.IV, Table 6.8)..... | 475 |
| Table 1.4: | Number and percentage of people from each ethnic population living inside their traditional home provinces and districts... | 476 |
| Table 1.5: | Distribution of those sections of ethnic populations who do not inhabit traditional home provinces. (Qualification >10 per cent outside home province)..... | 477 |
| Table 1.6: | Distribution within home province of those sections of ethnic populations who do not inhabit traditional home districts. (Qualification >7 per cent movement within province)..... | 478 |
| Table 2.1: | List of minimal populations surveyed and numbers sampled ..... | 479 |
| Table 2.2: | Sampling itinerary during fieldwork 1973 and 1975..... | 480 |
| Table 3.1: | Elementary sub-types of concordant variation ..... | 483 |
| Table 4.1: | The analysis of variance of the minimal atd angle by age ..... | 484 |
| Table 5.1: | Analysis of variance of minimal population samples (males and females) by all major dermatoglyphic digital and palmar attributes: and Chi Square results for nominal finger pattern attributes ..... | 485 |
| Table 5.2: | Percentage of significant attributes for various attribute groupings - from information in Table 5.1..... | 487 |
| Table 5.3: | Range of variation of minimal sample unit means and frequencies ..... | 488 |
| Table 5.4: | Pearson correlation matrix for male and female taxonomic attributes ..... | 490 |
| Table 5.5: | Eigenvalues for principal components analysis based on 26 male and 26 female summary traits ..... | 492 |

| | | Page |
|---|---|---|
| Table 5.6: | Principal components analysis (unrotated components) - analysis based on 26 male and 26 female traits ..... | 493 |
| Table 5.7: | Principal components analysis (varimax rotated solution) - Analysis based on 26 male and 26 female traits ..... | 494 |
| Table 5.8: | Eigenvalues for principal components analyses (males and females): based on 26 summary traits ..... | 495 |
| Table 5.9: | Principal components analysis (unrotated solution) - Analysis based on 26 summary traits ..... | 496 |
| Table 5.10: | Principal components analysis (varimax rotated solution) - Analysis based on 26 summary traits ..... | 496 |
| Table 5.11: | Eigenvalues for principal components analyses: based on 80 digital traits ..... | 497 |
| Table 5.12: | Principal components analysis (unrotated solution) - Based on 80 digital traits - (males)..... | 498 |
| Table 5.13: | Principal components analysis (unrotated solution) - Based on 80 digital traits - (females)..... | 500 |
| Table 5.14: | Principal components analysis (varimax rotated solution) - Based on 80 digital traits - (males)..... | 502 |
| Table 5.15: | Principal components analysis (varimax rotated solution) - Based on 80 digital traits - (females)..... | 504 |
| Table 5.16: | Eigenvalues for principal components analyses: based on 32 palmar traits..... | 506 |
| Table 5.17: | Principal components analysis (unrotated solution) - Analysis based on 32 palmar traits - (males)..... | 507 |
| Table 5.18: | Principal components analysis (unrotated solution) - Analysis based on 32 palmar traits - (females)..... | 507 |
| Table 5.19: | Principal components analysis (varimax rotated solution) - Analysis based on 32 palmar traits - (males)..... | 508 |
| Table 5.20: | Principal components analysis (varimax rotated solution) - Analysis based on 32 palmar traits - (females)..... | 508 |
| Table 5.21: | Statistical (Spearman) correlations (SC) compared to taxonomic (Pearson) correlations (TC): Summary attributes (excluding nominal digital attributes) - Males lower matrix, females upper matrix... | 509 |
| Table 5.22: | Statistical (Spearman) correlations (SC) compared to Taxonomic (Pearson) correlations (TC): Digital radial and ulnar ridge counts (males lower matrix, females upper matrix)..... | 510 |

| | | Page |
|---|---|---|
| Table 5.23: | Statistical (Spearman) correlations (SC) compared to taxonomic (Pearson) correlations (TC): Palmar ridge counts (male lower matrix, females upper matrix) | 511 |
| Table 5.24: | Statistical (Spearman) correlations (SC) compared to taxonomic (Pearson) correlations (TC): Palmar patterns and triradii (males lower matrix, females upper matrix) | 512 |
| Table 5.25: | Analytical comparison of statistical and taxonomic correlations through pairwise t-tests, for 4 major groups of attributes | 513 |
| Table 6.1: | Concordance between principal vectors and components as measured by Spearman's correlation coefficient | 514 |
| Table 7.1: | T-test probabilities for differences in attribute means between the sexes and $X^2$ probabilities for comparisons involving qualitative attributes | 516 |
| Table 7.2: | T-test probabilities for univariate comparisons between ethnic categories (males) | 517 |
| Table 7.3: | T-test probabilities for univariate comparisons between ethnic categories (females) | 518 |
| Table 7.4: | Summary of inter-category t-test probabilities: number of significant results for types of attribute group | 519 |
| Table 7.5: | Differences in means, standard deviations or frequencies between the sexes (males minus females) in Kenyan ethnic categories for a representative sample of attributes | 520 |
| Table 8.1: | Principal coordinates analysis of Kenyan ethnic groups (males): Analysis based on 26 summary traits | 521 |
| Table 8.2: | Principal coordinates analysis of Kenyan ethnic groups (females): Analysis based on 26 summary traits | 522 |
| Table 8.3: | Analysis of variance for mean ethnic group distances between and within ethnic categories for males and females | 523 |
| Table 8.4: | T-test for mean inter-ethnic group distances between (0) and within (1) ethnic categories for males and females | 524 |
| Table 8.5 | Qualitative breakdown of the distribution of inter-ethnic group distances between and within ethnic categories | 525 |

| | | Page |
|---|---|---|
| Table 8.6: | Distribution of Δg distances between populations at various sample sizes..... | 526 |
| Table 8.7: | Analysis of variance for male ethnic group minimal ethnic population samples (non summary and summary attributes)..... | 527 |
| Table 8.8: | Analysis of variance for female ethnic group minimal ethnic population samples (non summary and summary attributes)..... | 528 |
| Table 8.9: | Number and percentage of significant differences from analysis of variance of minimal population samples by ethnic group..... | 529 |
| Table 8.10: | Pearson correlations between ethnic groups: (with minimal sample populations recoded as belonging to a particular ethnic group (1) or not belonging to that group (0)) and 26 summary traits (males)..... | 530 |
| Table 8.11: | Pearson correlations between ethnic groups: (with minimal sample populations recoded as belonging to a particular ethnic group (1) or not belonging to that group (0)) and 26 summary traits (females)..... | 530 |
| Table 9.1: | Altitude and rainfall zone values and details, for minimal ethnic population samples..... | 531 |
| Table 9.2: | Significant Pearson correlations for summary taxonomic attributes with altitude; and partial correlations for summary taxonomic attributes with rainfall: controlling for altitude..... | 532 |
| Table 9.3: | Significant Pearson correlations for taxonomic non summary attributes with altitude..... | 533 |
| Table 9.4: | Significant partial correlations for taxonomic non summary attributes with rainfall: controlling for altitude..... | 534 |
| Table 9.5: | Analysis of variance of summary taxonomic attributes with altitude: displaying altitude zone means. (Only attributes that correlate significantly with altitude are included)..... | 535 |
| Table 9.6: | Analysis of variance of summary taxonomic attributes with rainfall: displaying rainfall zone means. (Only attributes that correlate significantly with rainfall are included)..... | 536 |
| Table 9.7: | Analysis of variance of non summary taxonomic attributes with altitude: displaying altitude zone means. (Only attributes that correlate significantly with altitude are included) - Males..... | 537 |

| | | Page |
|---|---|---|
| Table 9.8: | Analysis of variance of non summary taxonomic attributes with altitude: displaying altitude zone means. (Only attributes that correlate significantly with altitude are included) - Females..... | 538 |
| Table 9.9: | Analysis of variance of non summary taxonomic attributes with rainfall: displaying rainfall zone means. (Only attributes that correlate significantly with rainfall are included) - Males..... | 540 |
| Table 9.10: | Analysis of variance of non summary taxonomic attributes with rainfall: displaying rainfall zone means. (Only attributes that correlate significantly with rainfall are included) - Females..... | 541 |

FIGURES: Page


Figure 4.1: The computer coding sheet..... 543

Figure 6.1: Distribution of the 5050 distances between 101 populations of sub-Saharan Africa. (Hiernaux 1972:Figure 2)..... 544

Figure 6.2: Distribution of Δg distances between 57 male and 53 female minimal ethnic population samples: based on 26 summary traits..... 544

Figure 6.3: Distribution of Δg distances between 57 male and 53 female minimal ethnic population samples: based on 110 non summary traits..... 544

Figure 6.4: PCO analysis of Kenyan minimal samples (males): Sample distances from centroid in Δg units - Analysis based on 26 summary variables..... 545

Figure 6.5: PCO analysis of Kenyan minimal samples (females); Sample distances from centroid in Δg units - Analysis based on 26 summary variables..... 545

Figure 6.6: PCO analysis of Kenyan minimal samples (males): Sample distances from centroid in Δg units - Analysis based on 110 variables ..... 546

Figure 6.7: PCO analysis of Kenyan minimal samples (females): Sample distances from centroid in Δg units - Analysis based on 110 variables..... 546

Figure 6.8: Distribution of eleven principal components: total sample (males & females). 547

Figure 6.9: Distribution of six principal components: total sample (males)..... 548

Figure 6.10: Distribution of six principal components: total sample (females)..... 548

Figure 6.11: Principal coordinates analysis of Δg distances between Kenyan minimal samples: (males) - 26 summary variables. Vector 1 versus Vector 2..... 549

Figure 6.12: Principal coordinates analysis of Δg distances between Kenyan minimal samples: (males) - 26 summary variables. Vector 1 versus Vector 3..... 549

Figure 6.13: Principal coordinates analysis of Δg distances between Kenyan minimal samples: (males) - 26 summary variables. Vector 2 versus Vector 3..... 550

Figure 6.14: Principal coordinates analysis of Δg distances between Kenyan minimal samples: (females) - 26 summary variables. Vector 1 versus Vector 2..... 550

| | | Page |
|---|---|---|
| Figure 6.15: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (females) - 26 summary variables. Vector 1 versus Vector 3..... | 551 |
| Figure 6.16: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (females) - 26 summary variables. Vector 2 versus Vector 3..... | 551 |
| Figure 6.17: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (males): - 110 variables. Vector 1 versus Vector 2..... | 552 |
| Figure 6.18: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (males) - 110 variables. Vector 1 versus Vector 3..... | 552 |
| Figure 6.19: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (males) - 110 variables. Vector 2 versus Vector 3..... | 553 |
| Figure 6.20: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (females) - 110 variables. Vector 1 versus Vector 2..... | 553 |
| Figure 6.21: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (females) - 110 variables. Vector 1 versus Vector 3..... | 554 |
| Figure 6.22: | Principal coordinates analysis of $\Delta_g$ distances between Kenyan minimal samples: (females) - 110 variables. Vector 2 versus Vector 3..... | 554 |
| Figure 6.23: | Principal components analysis of Kenyan minimal sample variation: (males) - 26 summary variables. Component 1 versus Component 2..... | 555 |
| Figure 6.24: | Principal components analysis of Kenyan minimal sample variation: (males) - 26 summary variables. Component 1 versus Component 3..... | 555 |
| Figure 6.25: | Principal components analysis of Kenyan minimal sample variation: (males) - 26 summary variables. Component 2 versus Component 3..... | 556 |
| Figure 6.26: | Principal components analysis of Kenyan minimal sample variation: (females) - 26 summary variables. Component 1 versus Component 2..... | 556 |
| Figure 6.27: | Principal components analysis of Kenyan minimal sample variation: (females) -26 summary variables. Component 1 versus Component 3..... | 557 |

| | | Page |
|---|---|---|
| Figure 6.28: | Principal components analysis of Kenyan minimal sample variation: (females) - 26 summary variables. Component 2 versus Component 3..... | 557 |
| Figure 8.1: | Distribution of the first six principal coordinates - Cushitic groups (males). Analysis based on twenty six digital and palmar variables..... | 558 |
| Figure 8.2: | Distribution of the first six principal coordinates - Cushitic groups (females). Analysis based on twenty six digital and palmar variables..... | 558 |
| Figure 8.3 | Distribution of the first six principal coordinates - Nilotic groups (males). Analysis based on twenty six digital and palmar variables..... | 558 |
| Figure 8.4: | Distribution of the first six principal coordinates - Nilotic groups (females). Analysis based on twenty six digital and palmar variables..... | 558 |
| Figure 8.5 | Distribution of the first six principal coordinates - Bantu groups (males). Analysis based on twenty six digital and palmar variables..... | 559 |
| Figure 8.6: | Distribution of the first six principal coordinates - Bantu groups (females). Analysis based on twenty six digital and palmar variables..... | 559 |
| Figure 8.7: | Distribution of first four principal coordinates. Cushitic samples (males) Analysis based on twenty six variables..... | 560 |
| Figure 8.8: | Distribution of first four principal coordinates. Cushitic samples (females) Analysis based on twenty six variables..... | 560 |
| Figure 8.9: | Distribution of first six principal coordinates. Cushitic samples (females) Analysis based on one hundred and ten variables..... | 560 |
| Figure 8.10: | Distribution of first four principal coordinates. Dorobo samples (males). Analysis based on twenty six summary variables..... | 561 |
| Figure 8.11; | Distrubtion of first six principal coordinates. Dorobo samples (males). Analysis based on one hundred and ten variables..... | 561 |
| Figure 8.12; | Distribution of first four principal coordinates. Non Kalenjin Nilotic samples (males). Analysis based on twenty six summary variables..... | 561 |
| Figure 8.13: | Distribution of first four principal coordinates. Non Kalenjin Nilotic samples (females). Analysis based on twenty six summary variables..... | 561 |

| | | Page |
|---|---|---|
| Figure 8.14: | Distribution of first four principal coordinates. South-Western Kalenjin samples (males). Analysis based on twenty six summary variables..... | 562 |
| Figure 8.15: | Distribution of first four principal coordinates. South-Western Kalenjin samples (females). Analysis based on twenty six summary variables..... | 562 |
| Figure 8.16: | Distribution of first four principal coordinates. North-Eastern Kalenjin samples (males). Analysis based on twenty six summary variables..... | 562 |
| Figure 8.17: | Distribution of first four principal coordinates. North-Eastern Kalenjin samples (females). Analysis based on twenty six summary variables..... | 562 |
| Figure 8.18: | Distribution of first four principal coordinates. Southern Luyia samples (males). Analysis based on twenty six summary variables..... | 563 |
| Figure 8.19: | Distribution of first four principal coordinates. Southern Luyia samples (females). Analysis based on twenty six summary variables..... | 563 |
| Figure 8.20: | Distribution of first four principal coordinates. Central Luyia samples (males). Analysis based on twenty six summary variables..... | 563 |
| Figure 8.21: | Distribution of first four principal coordinates. Central Luyia samples (females). Analysis based on twenty six summary variables..... | 563 |
| Figure 8.22: | Distribution of first four principal coordinates. Northern Luyia samples (males). Analysis based on twenty six summary variables..... | 564 |
| Figure 8.23: | Distribution of first four principal coordinates. Northern Luyia samples (females). Analysis based on twenty six summary variables..... | 564 |
| Figure 8.24: | Distribution of first four principal coordinates, Kikuyu and Kamba samples (males). Analysis based on twenty six summary variables..... | 565 |
| Figure 8.25: | Distribution of first four principal coordinates. Kikuyu and Kamba samples (females). Analysis based on twenty six summary variables..... | 565 |
| Figure 8.26: | Distribution of first four principal coordinates. Ndia, Embu, Mbere samples (males). Analysis based on twenty six summary variables..... | 565 |

| | | Page |
|---|---|---|
| Figure 8.27: | Distribution of first four principal coordinates. Ndia, Embu, Mbere samples (females). Analysis based on twenty six summary variables..... | 565 |
| Figure 8.28: | Distribution of first four principal coordinates. Southern Meru sample (males). Analysis based on twenty six summary variables..... | 566 |
| Figure 8.29: | Distribution of first four principal coordinates. Southern Meru samples (females). Analysis based on twenty six summary variables..... | 566 |
| Figure 8.30: | Distribution of first four principal coordinates. Northern Meru samples (males). Analysis based on twenty six summary variables..... | 566 |
| Figure 8.31: | Distribution of first four principal coordinates. Northern Meru samples (females). Analysis based on twenty six summary variables..... | 566 |
| Figure 8.32: | Distribution of first four principal coordinates. Mji Kenda samples (males) Analysis based on twenty six summary variables..... | 567 |
| Figure 8.33: | Distribution of first four principal coordinates. Mji Kenda samples (females). Analysis based on twenty six summary variables..... | 567 |
| Figure 8.34: | Distribution of first four principal coordinates. Pokomo and Taita samples (males). Analysis based on twenty six summary variables..... | 567 |
| Figure 8.35: | Distribution of first four principal coordinates. Pokomo and Taita samples (females). Analysis based on twenty six summary variables..... | 567 |

APPENDICES


| | | Page |
|---|---|---|
| APPENDIX 1: | UNIVARIATE TABLES OF DERMATOGLYPHIC VARIATION | 568 |
| Appendix 1.1: | The Total Sample | 569 |
| Table Ap.1.1.1 | Distribution of finger ridge counts in the Kenya total sample by sex | 569 |
| Table Ap.1.1.2 | Distribution of finger pattern intensity in the Kenya total sample by sex | 570 |
| Table Ap.1.1.3 | Distribution of digital patterns in the Kenya total sample by sex | 570 |
| Table Ap.1.1.4 | Distribution of palmar ridge counts in the Kenya total sample by sex | 571 |
| Table Ap.1.1.5 | Distribution of palmar patterns and triradii in the Kenya total sample by sex | 571 |
| Table Ap.1.1.6 | Distribution of some rare palmar patterns and triradii in the Kenya total sample by sex | 571 |
| Appendix 1.2: | Ethnic categories | 572 |
| Table Ap.1.2.1 | Distribution of digital ridge counts and finger pattern intensity in Kenyan ethnic categories (males) | 572 |
| Table Ap.1.2.2 | Distribution of digital ridge counts and finger pattern intensity in Kenyan ethnic categories (females) | 573 |
| Table Ap.1.2.3 | Distribution of digital patterns in Kenyan ethnic categories (males) | 574 |
| Table Ap.1.2.4 | Distribution of digital patterns in Kenyan ethnic categories (females) | 575 |
| Table Ap.1.2.5 | Distribution of palmar ridge counts in Kenyan ethnic categories | 576 |
| Table Ap.1.2.6 | Distribution of palmar indices of triradial intensity and the atd angle in Kenyan ethnic categories | 576 |
| Table Ap.1.2.7 | Distribution of palmar patterns in Kenyan ethnic categories | 577 |
| Table Ap.1.2.8 | Distribution of palmar triradii in Kenyan ethnic categories | 577 |
| Table Ap.1.2.9 | The frequency of interdigital triradii in Kenyan ethnic categories | 577 |
| Appendix 1.3: | Ethnic Groups | 578 |
| Table Ap.1.3.1 | Distribution of digital ridge counts in Kenyan ethnic group samples (males) | 578 |
| Table Ap.1.3.2 | Distribution of digital ridge counts in Kenyan ethnic group samples (females) | 579 |

| | | Page |
|---|---|---|
| Table Ap.1.3.3 | Distribution of total finger ridge counts and palmar ridge counts in Kenyan ethnic group samples (males and females) | 580 |
| Table Ap.1.3.4 | Distribution of digital patterns in Kenyan ethnic group samples (males) | 581 |
| Table Ap.1.3.5 | Distribution of digital patterns in Kenyan ethnic group samples (females) | 582 |
| Table Ap.1.3.6 | Distribution of basic total digital patterns in Kenyan ethnic group samples (males and females) | 583 |
| Table Ap.1.3.7 | Distribution of total summary finger patterns in Kenyan ethnic group samples (males and females) | 583 |
| Table Ap.1.3.8 | Distribution of finger and palmar indices of triradial intensity and the atd angle in Kenyan ethnic group samples (males) | 584 |
| Table Ap.1.3.9 | Distribution of finger and palmar indices of triradial intensity and the atd angle in Kenyan ethnic group samples (females) | 585 |
| Table Ap.1.3.10 | Distribution of common palmar patterns and triradii in Kenyan ethnic group samples (males) | 586 |
| Table Ap.1.3.11 | Distribution of common palmar patterns and triradii in Kenyan ethnic group samples (females) | 587 |
| Table Ap.1.3.12 | Frequencies of the number of interdigital triradii in Kenyan ethnic group samples (males and females) | 588 |
| Table Ap.1.3.13 | Significance of ethnic group variation for a range of dermatoglyphic attributes | 589 |
| Appendix 1.4: | Minimal Ethnic Population Samples | 590 |
| Table Ap.1.4.1 | Distribution of radial ridge counts in Kenyan minimal ethnic population samples (right hand males) | 590 |
| Table Ap.1.4.2 | Distribution of radial ridge counts in Kenyan minimal ethnic population samples (left hand males) | 591 |
| Table Ap.1.4.3 | Distribution of radial ridge counts in Kenyan minimal ethnic population samples (right hand females) | 592 |
| Table Ap.1.4.4 | Distribution of radial ridge counts in Kenyan minimal ethnic population samples (left hand females) | 593 |
| Table Ap.1.4.5 | Distribution of ulnar ridge counts in Kenyan minimal ethnic population samples (right hand males) | 594 |

| | | Page |
|---|---|---|
| Table Ap.1.4.6 | Distribution of ulnar ridge counts in Kenyan minimal ethnic population samples (left hand males)..... | 595 |
| Table Ap.1.4.7 | Distribution of ulnar ridge counts in Kenyan minimal ethnic population samples (right hand females)..... | 596 |
| Table Ap.1.4.8 | Distribution of ulnar ridge counts in Kenyan minimal ethnic population samples (left hand females)..... | 597 |
| Table Ap.1.4.9 | Distribution of unilateral ridge counts in Kenyan minimal ethnic population samples (right hand males)..... | 598 |
| Table Ap.1.4.10 | Distribution of unilateral ridge counts in Kenyan minimal ethnic population samples (left hand males)..... | 599 |
| Table Ap.1.4.11 | Distribution of unilateral ridge counts in Kenyan minimal ethnic population samples (right hand females)..... | 600 |
| Table Ap.1.4.12 | Distribution of unilateral ridge counts in Kenyan minimal ethnic population samples (left hand females)..... | 601 |
| Table Ap.1.4.13 | Distribution of absolute ridge counts in Kenyan minimal ethnic population samples (right hand males)..... | 602 |
| Table Ap.1.4.14 | Distribution of absolute ridge counts in Kenyan minimal ethnic population samples (left hand males)..... | 603 |
| Table Ap.1.4.15 | Distribution of absolute ridge counts in Kenyan minimal ethnic population samples (right hand females)..... | 604 |
| Table Ap.1.4.16 | Distribution of absolute ridge counts in Kenyan minimal ethnic population samples (left hand females)..... | 605 |
| Table Ap.1.4.17 | Distribution of total ridge counts in Kenyan minimal ethnic population samples (males)..... | 606 |
| Table Ap.1.4.18 | Distribution of total ridge counts in Kenyan minimal ethnic population samples (females)..... | 607 |
| Table Ap.1.4.19 | Distribution of finger pattern intensity in Kenyan minimal ethnic population samples (right hand males)..... | 608 |
| Table Ap.1.4.20 | Distribution of finger pattern intensity in Kenyan minimal ethnic population samples (left hand males)..... | 609 |
| Table Ap.1.4.21 | Distribution of finger pattern intensity in Kenyan minimal ethnic population samples (right hand females)..... | 610 |
| Table Ap.1.4.22 | Distribution of finger pattern intensity in Kenyan minimal ethnic population samples (left hand females)..... | 611 |

Page

Table Ap.1.4.23 Distribution of total arches and total loops in Kenyan minimal ethnic population samples (males)..... 612
Table Ap.1.4.24 Distribution of total arches and total loops in Kenyan minimal ethnic population samples (females)..... 613
Table Ap.1.4.25 Distribution of total whorls and ulnar loops in Kenyan minimal ethnic population samples (males)..... 614
Table Ap.1.4.26 Distribution of total whorls and ulnar loops in Kenyan minimal ethnic population samples (females)..... 615
Table Ap.1.4.27 Distribution of concentric whorls and double loops in Kenyan minimal ethnic population samples (males)..... 616
Table Ap.1.4.28 Distribution of concentric whorls and double loops in Kenyan minimal ethnic population samples (females)..... 617
Table Ap.1.4.29 Distribution of radial loops in Kenyan minimal ethnic population samples (males and females)..... 618
Table Ap.1.4.30 Distribution of basic digital patterns on all fingers by hand in Kenyan minimal ethnic population samples (males).. 619
Table Ap.1.4.31 Distribution of basic digital patterns on all fingers by hand in Kenyan minimal ethnic population samples (females)..... 620
Table Ap.1.4.32 Distribution of total patterns on all fingers by hand in Kenyan minimal ethnic population samples (males)..... 621
Table Ap.1.4.33 Distribution of total patterns on all fingers by hand in Kenyan minimal ethnic population samples (females)..... 622
Table Ap.1.4.34 Distribution of indices of triradial intensity and the atd angle in Kenyan minimal ethnic population samples (right hand males)..... 623
Table Ap.1.4.35 Distribution of indices of triradial intensity and the atd angle in Kenyan minimal ethnic population samples (left hand males)..... 624
Table Ap.1.4.36 Distribution of indices of triradial intensity and the atd angle in Kenyan minimal ethnic population samples (right hand females)..... 625
Table Ap.1.4.37 Distribution of indices of triradial intensity and the atd angle in Kenyan minimal ethnic population samples (left hand females)..... 626
Table Ap.1.4.38 Distribution of palmar ridge counts in Kenyan minimal ethnic population samples (males)..... 627

Page

Table Ap.1.4.39 Distribution of palmar ridge counts in Kenyan minimal ethnic population samples (females)..... 628
Table Ap.1.4.40 Distribution of palmar patterns and triradii in Kenyan minimal ethnic population samples (right hand males)..... 629
Table Ap.1.4.41 Distribution of palmar patterns and triradii in Kenyan minimal ethnic population samples (left hand males)..... 630
Table Ap.1.4.42 Distribution of palmar patterns and triradii in Kenyan minimal ethnic population samples (right hand females).... 631
Table Ap.1.4.43 Distribution of palmar patterns and triradii in Kenyan minimal ethnic population samples (left hand females)..... 632
Table Ap.1.4.44 Frequency of interdigital triradii in Kenyan minimal ethnic population samples (males and females)..... 633

APPENDIX 2: PROGRAM FOR THE COMPUTATION OF HIERNAUX'S STATISTIC (Δg) ..... 634

APPENDIX 3: MULTIVARIATE VARIATION IN KENYAN MINIMAL ETHNIC POPULATION SAMPLES ..... 635

Table Ap.3.1 Δg distances between Kenyan minimal ethnic population samples (males): analysis based on 26 summary attributes.... 636
Table Ap.3.2 Δg distances between Kenyan minimal ethnic population samples (females): analysis based on 26 summary attributes.... 638
Table Ap.3.3 Δg distances betwen Kenyan minimal ethnic population samples (males): analysis based on 110 non summary attributes..... 640
Table Ap.3.4 Δg distances between Kenyan minimal ethnic population samples (females): analysis based on 110 non summary attributes..... 642
Table Ap.3.5 PCO analysis of Δg distances between minimal ethnic population samples (males): analysis based on 26 summary attributes.... 644
Table Ap.3.6 PCO analysis of Δg distances between minimal ethnic population samples (females): analysis based on 26 summary attributes..... 645
Table Ap.3.7 PCO analysis of Δg distances between minimal ethnic population samples (males): analysis based on 110 non summary attributes..... 646

| | | Page |
|---|---|---|
| Table Ap.3.8 | PCO analysis of Δg distances between minimal ethnic population samples (females): analysis based on 110 non summary attributes..... | 647 |
| Table Ap.3.9 | Unrotated component scores (males): based on 26 summary traits..... | 648 |
| Table Ap.3.10 | Unrotated component scores (females): based on 26 summary traits..... | 649 |
| Table Ap.3.11 | Rotated component scores (males): based on 26 summary traits..... | 650 |
| Table Ap.3.12 | Rotated component scores (females): based on 26 summary traits..... | 650 |
| Table Ap.3.13 | Unrotated component scores: based on 26 male and 26 female summary traits..... | 651 |
| Table Ap.3.14 | Rotated component scores: based on 26 male and 26 female summary traits..... | 652 |
| Table Ap.3.15 | Unrotated component scores (males): based on 80 digital traits..... | 653 |
| Table Ap.3.16 | Unrotated component scores (females): based on 80 digital traits..... | 655 |
| Table Ap.3.17 | Rotated component scores (males): based on 80 digital traits..... | 657 |
| Table Ap.3.18 | Rotated component scores (females): based on 80 digital traits..... | 658 |
| Table Ap.3.19 | Unrotated component scores (males): based on 32 palmar traits..... | 659 |
| Table Ap.3.20 | Unrotated component scores (females): based on 32 palmar traits..... | 660 |
| Table Ap.3.21 | Rotated component scores (males): based on 32 palmar traits..... | 661 |
| Table Ap.3.22 | Rotated component scores (females): based on 32 palmar traits..... | 662 |

BIBLIOGRAPHY..... 663

MAPS (see folder)

MAP 1 Topography of Kenya
MAP 2 Distribution of mean annual rainfall (inches)
MAP 3 Major vegetation zones of Kenya
MAP 4 Major ethnic populations in Kenya
- Map $4_1$ Distribution of the Luyia, Bantu and Elgoni Kalenjin
- Map $4_2$ Distribution of Central Bantu
- Map $4_3$ Distribution of Mji Kenda Bantu
- Map $4_4$ Distribution of Pokomo Bantu
- Map $4_5$ Distribution of Luo "tribes" and territorial divisions.

MAP 5 Sampling locations

## ACKNOWLEDGEMENTS

Many African governments today have a restrictive and very selective policy regarding academic research. In view of this, the open policy pursued by the Government of Kenya is welcome and I am grateful for the permission to conduct research in that country.

In addition to the general permission from the Government, my presence in each administrative province and district had to be cleared and registered. I am indebted to the administrative officials for the efficient and speedy fashion in which they dealt with the red tape. Many in fact, went beyond the call of duty, in offering me advice, local information and in two occasions hospitality in their homes.

My deepest gratitude, however, is reserved for the schoolchildren, who cooperated so enthusiastically, in spite of being considerably perplexed as to why I should require their palm and finger prints. I also greatly appreciated the tolerant and warm welcome of the headmasters and staff of the schools visited, who often proved to be hospitable, witty and congenial companions in what was often a lonely business.

The task of collecting hand prints from so many schoolchildren would have been daunting without the help of volunteer assistants from time to time. My deep thanks are due to Sammy Maraka, Penny Robie, Josiah Njare, Anne Olivecrona and Jim Gillespie who helped me in 1973, and to John Murgatroyd and Anne Rush in 1975.

During my field trips I could not have lasted the pace without hospitality from many friends, and acquaintances met along

the route. They are too numerous to mention all by name. I would, however, like to especially thank Dean and Helen Oyen, Lew and Nancy Hurxthall, John and Liz Baggalley, Tony and Honour Gasston and Luciano and Louisella Marani for their outstanding kindness.

Throughout the project I have received substantial and continuing support from Professor E. Sunderland, who has patiently supervised by study since 1973. I am grateful for his considerable help, especially during some difficult financial periods when his sympathetic encouragement was the vital factor in persuading me to continue.

I am also grateful to my colleagues in the Durham Department, particularly Roger Dennis, for offering advice and active criticism during the earlier stages of analysis. My thanks are also due to the staff of the Computer Unit, in the University of Durham, for their patient technical advice, and to Mrs. Elves and her team of computer card punchers who prepared over 19,000 cards of data. At the University of Stirling my thanks are due to Graham Cameron and Les Oswald for allowing me to use some of their graphical programs, and to Richard Bland who corrected statistical misconceptions at a crucial stage in the analysis.

The typing of the manuscript is the work of Mrs. Margaret Buchanan, who has efficiently plodded away in the face of constant distractions and revisions. She has my appreciation and gratitude.

To Dr. Alison Bowes fell the unenviable task of correcting my wayward grammar and spelling. Her reward is the dubious distinction of being one of the few people who is ever likely to read this work in its entirety. I also thank

Georg Feuerstein, who tested friendship with some meticulous and pertinent corrections. Finally I am indebted to Irene Wilson for the hours she spent on the cartography of the tribal maps, in spite of having the busiest schedule of any person I know. Her unselfish help is greatly appreciated.

In spite of all this help, however, I would end by stressing that errors in the present study are all my own.

I could not have conducted the research without the financial support of official bodies and foundations. My thanks are due to the Sigma XI Foundation, Connecticut, and to the Horniman Fund of the Royal Anthropological Institute for providing funds for fieldwork; and not least to the University of Durham for awarding me a Durham Studentship.

# INTRODUCTION

## 1.1 Physical anthropology as a historical discipline

Physical anthropology is and has been primarily concerned

> "with describing and interpreting the patterns of geographical variety which are found to be so striking in the human species"
>
> (Harrison 1977:1)

In the pre war and early post war years, the main physical anthropological approach was associated with the construction and comparison of racial taxonomies based on anthropometric and other polygenic attributes. In recent years, however, the emphasis has changed to the detailed study of the operation of the laws of population genetics in human populations (especially micro populations), and the collection of unifactorial genetic data has been a central concern of workers in the field.

The rise of what Weiner calls the new "bio-anthropology" (Weiner 1979:4), has seen increasing criticism of the pre war taxonomical approach, which Weiner in the 1978 Huxley Memorial Lecture accuses of:

> "landing physical anthropology into the cul-de-sac of racial taxonomy, of racial classification and craniometry pursued relentlessly for its own sake .. an era of unrestrained evolutionary interpretation, a multiplication of speculative phylogenetic trees."
>
> (Weiner 1979:4)

Harrison, however, reminds us that the pre-war taxonomic approach was not just a matter of classification,

Durham University SCIENCE Site - 2 OCT 1981 SECTION Library

> "for on the Darwinian premise that similarity indicates phylogenetic affinity, it was taken that taxonomies attempted to express the evolutionary history of the species. Ascertaining this history has always been and remains one of the main goals of physical anthropology."
>
> (Harrison 1977:1)

In spite of this affirmation of the traditional historical role of physical anthropology, the study of the phylogenetic origins of human populations is much more uncertain today than it was in the pre-war era. The reason for this is not so much antipathy towards taxonomy and phylogeny, or a preference for concentrating on synchronic population studies, but rather the complexity of defining human taxonomic populations objectively.

Since physical anthropologists began seriously to examine serological genetic variation, and to investigate the complexities of localised patterns of variation within major geographical populations, it has become clear that:

a) there is often considerable overlap in the ranges of biological means and frequencies between populations that were once thought to be distinct taxons:

b) patterns of variation for different attributes are often not concordant between major populations:

c) many attributes are more adaptive than was previously suspected, thus increasing the probability that close resemblance between populations for these attributes is superficial, and does not, therefore, imply common origins.

Hiernaux in particular (e.g. Hiernaux 1972), has done much to illustrate the difficulty of delineating distinctive taxons within major geographical populations. In sub-Saharan Africa, an

area once thought to contain a multiplicity of taxons (e.g. see Oschinsky 1954), Hiernaux failed to find any distinctive population clusters that could be linked hierarchically to depict phylogenetic ancestry. Hiernaux wonders from his findings whether the situation he found in Africa might not apply in the rest of the world too (Hiernaux 1972:106-7).

The experience of physical anthropologists since the war in studies using both multifactorial and unifactorial attributes, has been that only major continental populations show any truly taxonomically distinct characteristics. As Harrison remarks:

> "broad continental patterns emerge, and taxonomies of man, based on polymorphic systems, are broadly in harmony with those based on anthropometric traits."
>
> (Harrison 1977:2)

It is because of the non-taxonomically distinct range of variation found between populations within continents, and the fact that even continental populations show an overlap in ranges of attribute means and frequencies (e.g. for dermatoglyphic digital patterns (Leschi 1950)), that there have been few recent phylogenetic studies of human macro populations. Even the elegant phylogenetic reconstructions of Howells (1973) using craniometric traits, and Cavalli-Sforza and Edwards (1967) using serological traits, are more methodological exercises than serious attempts at evolutionary reconstruction.

The modern trend is rather to concentrate on the study of populations at a smaller regional scale. Biological history is still an important interest in such studies, but the emphasis is on non-taxonomic similarity between biological populations as measured multivariately by coefficients of genetic distance. As

in the taxonomic orientation, relative similarity still implies a common origin, but it is no longer expected that phylogenetic relationships should be hierarchical in structure. The non-taxonomic historical approach has been advocated by Hiernaux in particular (Hiernaux 1972). One of his most notable successes has been to confirm the common origins of the Bantu populations of sub-Saharan Africa through the analysis of relative distance calculated from a series of mostly polygenic attributes (Hiernaux 1968b).

Another important advance has been the use of genetic similarity to reconstruct the phylogenetic history of small isolated populations, where the influence of complicating factors such as systematic selection and mixture is slight. This phylogenetic model was developed by Cavalli-Sforza and Edwards (1966-1967) and is based on the assumption that differentiation over time results mainly from random stochastic processes. Through the probabilistic estimation of rates of divergence, it is possible not only to link populations hierarchically, but also to calculate the time elapsed since the divergence of each branch. This model was adapted with some success by Malyutov et.al. (1972) to reconstruct the phylogenetic relationship of isolated East Sayan populations from Siberia. It remains to be seen however, how far this approach can be profitably adapted to reconstruct the biological history of populations in areas where the biological situation is much more complex.

In interpreting evolutionary history to be mainly concerned with the reconstruction of phylogenetic origins, in

terms just discussed, physical anthropologists have followed the traditional historical aspirations of the zoological sciences. There is, however, a more general historical ambition implicit in physical anthropology, and that is the reconstruction of the pre-documentary history of human peoples, in particular of third world peoples who usually lack any documentary historical sources before the pre-colonial period.

In a "zoological" orientation, human populations are viewed as biological populations, or natural breeding units in which matings between individuals within a unit are much more common than those between units. During the course of generations, stable, relatively endogamous populations will become biologically homogeneous and relatively distinct from other biological populations. The onus on the physical anthropologist therefore, is to identify such units, sample them and explore relative resemblance between them in order to deduce some phylogenetic scheme. In the second, "ethnological" approach however, the historical units of interest are peoples who are defined according to linguistic and cultural criteria rather than biological ones.

In fact peoples' own views of their identity are independent of academic definitions - for example the history of the Kikuyu or Welsh is one of self-defined communities. The fact that ethnological history is that of "real" populations means that the task of the physical anthropologist is much more complex than it would be were he to focus on groups like "Australoids", or "Hamites", defined not by the peoples themselves but by scientist outsiders. To assume, for example, that people who call themselves

"Kikuyu" or "Sonjo" or "Benin" constitute a biological population, and to embark on a phylogenetic study, would be dangerously simplistic. It is more than likely that the "Kikuyu", "Sonjo" and "Benin" are each composed of some subsections of diverse origins.

The initial task of a physical anthropological approach to ethnological historical reconstruction is to explore how far the populations of interest do in fact represent homogeneous biological populations. It is of paramount historical interest for instance, to ascertain whether homogeneous linguistic and cultural populations are composed of heterogeneous biological elements, or conversely whether peoples who are linguistically and socially heterogeneous are biologically homogeneous. In practice, this requires a shift of emphasis from dealing directly with known peoples, to examining the biological history of geographical regions. In other words, the greatest profit is to be derived when the questions asked are not for instance, "what are the origins of the Kikuyu and Meru, have there been previous migrations of Kikuyu from the coast?" but rather "are the people currently living in the Kikuyu area of the same origin, and if not, are there any clues which might help us explain their heterogeneity? Is it reasonable to suppose that the Kikuyu area was originally populated by people from the coast, followed by later waves from different areas? How can an examination of the biological evidence help illuminate these questions?".

The ethnological historical role of physical anthropology has never been overtly distinguished from the "zoological" orientation. There are for instance, no treatises specifically devoted to ethnological issues in the physical anthropological

literature. It is in fact only recently that any sophisticated treatment of ethnological issues has been attempted, (e.g. Hiernaux 1968b). It is thus not surprising that professional historians who specialise on the pre-colonial past of third world peoples, have not viewed physical anthropology with much enthusiasm. Kiernan for instance, in a review of the role of the historian in East Africa, remarks that the

> "results of the work of physical anthropologists, biochemists and biologists have been disappointing. Attempts to classify peoples of Africa have not been successful.... Human remains provided by archaeology have not been numerous enough to enable any worthwhile conclusions to be reached ... and it was hoped at one time that the investigation of man's genetic inheritance would provide more information than has been the case".
>
> (Kiernan 1974:14)

He unflatteringly adds that "genetic studies are most useful from a historical point of view when they relate to plants" (Kiernan 1974:15) - here; he is referring to studies concerned with mapping the introduction of various crops into Africa. By contrast, the contribution of linguists, cultural archaeologists and oral historians is highly praised.

Kiernan expresses the hope that "human biology and related sciences will produce in the course of time some more sophisticated tools for detecting the effect of history" (Kiernan 1974:14). It is however, not so much a question of developing techniques, but rather one of collecting adequate data and of asking the correct questions. The future of physical anthropology as an ethnological science lies in the short term at least, not so much in setting up major historical schemes as linguists have done, but in

integrating its material with the findings of linguists, archaeologists and oral historians, as Hiernaux did in his assessment of Bantu expansions (Hiernaux 1968b). We must examine hypotheses constructed by historians in these other disciplines, and decide which of them can be tested by using biological data, and test them by collecting relevant material.

Looked at in this way, the potential range of historical questions that can be examined by physical anthropologists is large. The more extensive the data available from any region, both in terms of the number of regional sampling points and of the attributes measured, and the more detailed the synchronic explanatory analysis of regional variation of attributes, the more sophisticated the treatment of historial questions can become. The ethnological potential of any regional biological study therefore, is highly dependent on the quality of conventional biological explanatory analysis. It is not just a matter of obtaining enough data to be able to compute relative resemblance between samples. For example, an assessment of relative resemblance based on non adaptive attributes can tell us how far sampled micro-populations in an area can be said to have a common origin. If we know that differentiation is primarily a consequence of random stochastic processes, then some form of dating of divergence of branches can be attempted using the Edwards and Cavalli-Sforza models mentioned earlier. If selection, whether through adaptation or not, is characteristic of specific attributes, then different problems can be investigated. For example, in the case of sickle cell trait, if an African population shows a very low incidence of the trait in a highly malarial environment where other populations

manifest high sickling rates, we could assume that the group has only recently migrated to the zone.

## 1.2 Objectives of the research

In 1972, I decided to study how far human biological data from sampled Kenyan peoples could be utilised to further the reconstruction of their pre-colonial history. From the outset, there were three options open to me. Firstly, I could use the material from existing surveys to develop analytical techniques to test historical hypotheses. Secondly, I could collect my own detailed data from a limited area of Kenya and use the material for a case study in historical analysis. Thirdly, I could attempt a national survey for a limited range of attributes and pursue historical analysis at different levels of resolution.

The study of pre-existing data held some promise. As Table I:1 shows there had been by 1972 some extensive surveys for ABO blood groups, sickle-cell haemoglobin and a limited range of anthropometric measurements. Further data in the form of Rhesus, MN and other serological systems, PTC tasting and colourblindness were available for several scattered samples. I decided against confining my study to pre-existing data however, for several reasons. Firstly, the range of attributes surveyed extensively is extremely limited. In the case of sickle-cell and stature, nasal and head length and breadth, (the five anthropometric measurements with adequate coverage), moreover, their well known adaptive characteristics make them unsuitable for the assessment of common origins. Secondly, the ethnic groups sampled in most

detail are the Coastal and Central Bantu. Samples of other groups are either absent, too small or too generalised, that is, large scale ethnic units were treated as single populations. Leys and Joyce (1913) for example took only two Luyia samples; the Luyia from the Kakamega area and the other Luyia. In reality, the Luyia are composed of 17 once autonomous sub-tribes which should all be sampled individually. This generalised treatment of large scale populations such as the Luyia and Kikuyu is in direct contrast with that for the Coastal peoples where individual sub-tribes such as the Giriama, Ribe, Kauma, Jibana and Digo have been individually sampled in several surveys (see Table 1.2.1). Another disadvantage of the pre-existing data is the non-concordance of populations sampled in different surveys. There are for example very few populations sampled for ABO, sickle-cell and anthropometric measurements.

Owing to these drawbacks, I decided that it was preferable to collect my own data, although the pre-existing material can still be useful for particular types of analysis. For instance, the potential for comparison of populations for serological attributes has increased significantly since the work of Corrain (see bibliography). In any case it would have been a pity to have by-passed the opportunity to add to the number of biological markers from Kenya.

The option of collecting detailed data from a local region of Kenya seemed particularly attractive when I designed the study. I could for instance, have taken regional samples within an ethnic group for a large range of serological, anthropometric, and possibly dermatoglyphic attributes, with some comparative material from neighbouring populations. This was the approach of Corrain above.

Corrain's approach, however, is not ideal for an ethnological study, where a lack of comparative material from a wide geographical range might have severely limited the potential of ethnological investigation. I tended to favour a large scale survey in the first instance, followed in the future by more detailed regional studies. There was one additional constraint. As a beginner in physical anthropological research, it seemed inadvisable to try and survey serological attributes in a country where the necessary and expensive back-up facilities were not at hand. It consequently seemed unprofitable to pursue a detailed regional study with any less precise genetic measures.

This left me with the option of trying to achieve some sort of wide-scale comprehensive coverage of Kenya for fewer attributes. The attributes I chose to survey were dermatoglyphic, as the collection of prints is relatively uncomplicated, and providing that standards of printing are high, are not prone to methodological errors of measurement in the field. Mistakes in print reading for instance, can be remedied by rereading the permanent records, but inexperienced errors in measuring anthropometric traits or collecting serological traits cannot be rectified after fieldwork. Secondly, prints can be obtained with a minimum of equipment, and on materials that are cheap to buy, an important consideration with the funds I had available for fieldwork. Finally hand and finger prints contain many genetically inherited attributes, which, when I began my study, appeared to have the advantage of being non-adaptive.

The final research design therefore, had four stages. The first was fieldwork, during which I planned to obtain dermatoglyphic samples of both sexes from as many high resolution ethnic and geographical populations as I could cover in the time available. The second stage was measuring the digital and palmar prints and preparing the data for statistical analysis. The third entailed conventional descriptive analysis to determine the extent and nature of dermatoglyphic variation in Kenya, and to try and determine possible causes for the variation. The final stage was to use the results of the analysis to explore relevant historical hypotheses.

It became evident however, as the research progressed, that the project was too big to include all of it in a PhD thesis. When it came to writing up the results, I found there was so much information and analysis that it would have taken several volumes to cover all aspects of the research. For the PhD study therefore, I decided to limit the report to those aspects of the project that relate to describing and accounting for the observed dermatoglyphic variation between the sampled populations. In essence therefore, the present work is a conventional human biological study of dermatoglyphic variation. Even so, in this more limited field of inquiry there are still many areas that have not yet been explored in detail. I hope that aspects not covered here will be written up in the future in the form of articles. The historical analysis, to which I have devoted many hours, is so complex that it will take some time to explore in sufficient depth, and requires detailed consultation with

authorities in other historical fields. This over-riding aim of the research, therefore awaits future discussion.

## I 3 Plan of the present study

The study begins in Chapter 1, with a background description of Kenya and its peoples, including sections on topography, rainfall, vegetation and ecological zones, a description of the peoples and their ethnic and linguistic affinities, a brief sketch of their history, and finally macro demographic characteristics drawn from the Kenya Population Census 1969. The chapter is primarily intended to put the reader into the picture, especially if he or she is unfamiliar with Kenya and its peoples. Various analytical chapters also draw on the basic information outlined in this chapter, in particular Chapters 2, 8 and 9.

The investigation begins in earnest in Chapter 2, which is divided into two parts. In the first I discuss the traditional physical anthropological view of African ethnic units as constituting natural biological populations, on the assumed characteristics that they are highly endogamous and relatively homogeneous biologically. I criticise this standpoint, and outline the complexities of the African societies and peoples. In particular I draw on the experience of social anthropologists to advance a view that pre-colonial African peoples were far from cohesive units, being rather structured on small scale semi-autonomous primary groups loosely linked to varying degrees of common identity, interaction and purpose. It is, in fact, only since colonial times that identity has broadened and cohesiveness increased. As

pre-colonial African peoples consisted of loosely linked primary groups, they were relatively unstable structures, being liable to disintigration in times of disaster such as famine or warfare. Many African ethnic units, therefore, far from being stable cohesive societies over time, are more than likely recent aggregations of primary group sections of diverse origins. The nature of African populations today, therefore, are highly complex, and if my appraisal is correct, a sampling design should be based on small scale populations that will be able to pick up any heterogeneity within ethnic units. The ideal unit for study, I suggest, is the minimal population unit, which is the unit of maximal identity and common purpose in pre-colonial times.

The second part of Chapter 2 considers the correct units to sample in Kenya specifically, and follows with an account of fieldwork and the sampling procedures adopted.

In Part Two I consider the descriptive analysis of the survey. Chapter 3 is a theoretical discussion of the correct approach to adopt in the analysis of regional biological variation, and provides the framework for subsequent investigation. I argue that in order to explain differentiation, it is first necessary to describe it adequately (Jardine 1971). The conventional practice of viewing the description of variation as falling into two parts, univariate and multivariate, with explanatory analysis being associated mainly with the latter, is not only wasteful, but potentially aimless. In a dermatoglyphic study, when a large number of attributes and populations is involved, the detailed description of univariate and multivariate elements of variation can be very lengthy. It is necessary therefore, to formulate

carefully the aims of description and analysis prior to the investigation. In order to preserve a correct balance between the openess of survey discovery and the rigidity of pre-formulated aims and hypotheses, one must draw from the experience of previous human biological studies. These show that regional variation usually conforms to a limited number of types. It is the assessment of which of these types best fits the data that gives survey description flexible purpose. The remainder of the chapter outlines various types of biological differentiation, and tries to show that they are best explored through seeing variation not as a series of independent elements, but as having a structure. The structure of variation is discussed in detail, and statistical methods for its investigation are suggested.

Following this theoretical framework, the descriptive analysis of the Kenyan variation commences in Chapter 4 with an outline of the methodology of attribute measurement and the setting up of the data matrices for statistical analysis. After a short section to introduce the basic evidence in the form of univariate tables, Chapter 5 is devoted to the description of structural relationships between attributes, and is in three stages. The first is designed to establish whether there is sufficient univariate significance to warrant a more sophisticated structural analysis. Having shown that there is a great deal of significant variation, I go on in the second stage to examine the relationship between taxonomic attributes, and assess the extent of taxonomic discordance between attributes. In the third stage I examine how far taxonomic discordance is related to underlying patterns of statistical discordance between the same attributes in individuals.

Chapter 6 introduces the concept of "overall" structure and investigates through various techniques the nature of multivariate variation. The results of the analysis of structure through the use of principal component analysis are compared to those of the analysis of genetic distance, to help gauge the potential usefulness of the distance strategy as an analytical rather than a descriptive tool. One of the major aims of this chapter is to see whether the overall structure of dermatoglyphic variation in Kenya is similar to that found by Hiernaux for other sub-Saharan populations in resembling a complex network of relationships rather than a series of distinct clusters (Hiernaux 1972).

Following the description of variation, in Part Three I move on to explore some explanatory issues. In Chapter 7 I concentrate on trying to explain why the sexes, while showing considerable replication in structural trends, nevertheless demonstrate marked taxonomic discordance in attribute variation. In Chapter 8 , patterns of linguistic and cultural affinity between Kenyan populations at three levels of resolution are contrasted with patterns of biological resemblance. At the highest level of resolution, that for minimal population samples, the concept of structural similarity is introduced to investigate biologicl relationships within ethnic groups, and to ascertain whether these relationships approximate to known patterns of linguistic divergence. Chapter 9 moves on to consider whether systematic dermatoglyphic relationships exist with climate, in particular altitude and rainfall zones, and the possibility that selection may be operating on some dermatoglyphic attributes is briefly discussed.

The explanatory analysis, cannot be considered comprehensive. I simply explore the possibilities of certain key areas. The number of possible explanatory issues is far from exhausted in the present work.

# PART ONE: BACKGROUND DESCRIPTION, THE POPULATIONS, SAMPLING AND FIELDWORK.

# CHAPTER 1

## Introduction to Kenya and its Peoples

### 1.1 The Physical Environment

#### 1.1.1 Relief and drainage

Part of the charm of Kenya is the diversity of its topography. Coral strands, white beaches and the Indian Ocean give way to thousands of square miles of low lying plains, which gradually rise to meet the central Highlands, a region of mountain ranges, plateaux and lakes. Of outstanding beauty is Mt. Kenya, permanently covered in snow and at 17,058 feet, Africa's second highest mountain. Bisecting the highlands is the spectacular Rift Valley, with its picturesque lakes and imposing escarpments. Finally in the far west lies the Winam Gulf (formerly Kavirondo Gulf), part of Lake Nyanza, the largest lake in Africa.

a) The coastal strip and low lying plains

The coast as Map 1 shows, runs in an irregular north easterly direction. Prominent features are Mombasa Island and its channels, the Kilifi inlet, the Galana/Sabaki and Tana Rivers estuaries, and the two big islands of Lamu and Pate.

The coastal plain (below the 500 ft. contour), is very narrow in the south, rising rapidly south of Mombasa to the Shimba Hills which just exceed 1,000 ft. North of Mombasa it widens rapidly, leaving an island of high ground, Mengea Hill west of Malindi. It broadens northwards west of the Tana River for about 150 miles, and then swings north eastwards parallel to the coast into Somalia. North of Lamu therefore, the coastal plain is very wide.

The land between 500 ft. and 4,000 ft. is low lying plains - the increase in altitude is gradual and not normally broken by any notable feature of relief. From Map 1 it will be seen that there are two altitude gradients. The first is the rise in altitude from the coast to the highlands in the west. The 3,000 ft. contour runs from south to north, parallel to the 4,000 ft. contour forming a narrow strip of high ground, and the 2,000 and 1,000 ft. contours form a wider band to the east. This concourse of lower contours swings gradually westwards in a circle as it approaches the escarpments round Lake Turkana.

The second gradient is an increase in altitude from south to north towards the Ethiopian Highlands. The 4,000 ft. contour runs just off the map, following the Ethiopian side of the Kenya/ Ethiopian border. In the east of the country the 500 to 1,000, and 1,000 to 2,000 ft. bands which run for some way from the Somali border in an east to west course, are very wide, and increase in altitude is gradual. Only near the Ethiopian border does altitude start to increase more rapidly.

The coastal and low lying plains cover some two thirds of Kenya in area, and are mostly featureless. The only relief lies close to the highland areas in the north and west. Of the few islands of higher ground the most spectacular is Mt. Marsabit in the central north, which rises to over 5,000 ft. from the surrounding plains.

The plains are intersected by hundreds of water-courses, which are dry for most of the year, and sometimes for years at a time. In this vast area there are only four permanent rivers:

a) the Galana Daua which rises in the Ethiopian Highlands and forms part of the Kenya/Ethiopian border in Mendera; b) the Ewaso Ngiro which rises in the highlands north of Mt, Kenya and disappears in the Lorian Swamp, never to reach the sea; c) the Tana River which has its source in the Kikuyu Highlands and flows into the sea some 50 miles north of Malindi (at 440 miles Kenya's longest river); and finally d) the Athi/Galana/Sabaki River which rises near Nairobi and reaches the sea just north of Malindi. There are also two other semi-permanent rivers, the Tiva and the Voi rivers, the former north of the Galana, and the latter to its south. Neither reaches the sea.

b) The Highlands and Rift Valley

The detailed relief of the Highland and Rift Valley regions is extremely complex. The region can best be envisaged as a series of escarpments and off-lying plateaux following the eastern and western ridges of the Rift Valley, which runs from north to south between the $36/37^{O}$ longitudinal meridians.

b.i) The northern Rift Valley and Lake Turkana Region

Lake Turkana occupies the Rift Valley floor for some 150 miles from the Ethiopian border. Unlike the more southern sections of the Rift Valley however, the section around Lake Turkana is not associated with very high land masses. The escarpment on the eastern shores of the lake is low, barely rising above 3,000 ft. On the west side the escarpment is not continuous, but a broken series of small hill ranges which do not exceed 4,000 ft. Beyond these immediate hills to the west the Turkana Plains stretch for some eighty to a hundred miles westwards in a north south direction

parallel to the lake.

The Turkana plains like those to the east of Kenya are low lying, mostly between 1,000 to 3,000 ft. Unlike the eastern plains however, they are frequently interspersed with ranges of hills. For example, the Karasuk Hills in south-western Turkanaland are particularly high, reaching 9,000 ft. Turkanaland also enjoys two permanent rivers, the Turkwell and the Kerio which drain into Lake Turkana after journeys of over 200 miles from the highlands to the south and west.

The western boundaries of Turkanaland , and of Kenya in the northwest, are marked by a long series of escarpments which exceed 5,000 ft. The direction of the escarpments is in a slight north west to south east direction, and they merge into the western Rift Valley land masses towards the south.

The highlands immediately south of Lake Turkana, and as far east as Mt. Marsabit, are of recent volcanic origin. The whole area is littered with volcanic craters, some of which are still smouldering. The volcanic activity has been responsible for the creation of large expanses of lava desert east of Lake Turkana.

b.ii) The Central Rift Valley and associated land masses

The Rift Valley floor south of Lake Turkana is a very complicated area, with a major range of broken hills within it. The higher land of the Loriu Plateau, which exceeds 4,000 ft. in places, gradually rises towards the Tugen Hills, which are higher than 7,000 ft. in some parts. East of the Loriu Plateau, the Rift Valley is called the Suguta Valley, and is one of the hottest places in Kenya.

To the south of Lake Turkana is Mt. Nyiro, the apex of a triangle of plateau lands which widens as it nears the equator. The western escarpments of the plateau areas are those of the eastern Rift Valley, and the eastern escarpments are a series of mountain ranges, the Ndoto and Mathews Ranges. The plateaux occupy the whole of the Samburu and Laikipia Districts, and vary between 3,000 and 6,000 ft. in height, with the mountain ranges reaching 7,000 ft. in places.

The western escarpment of the Rift Valley is extremely dramatic just north of the equator, rising from under 2,000 ft. along the Kerio Valley west of the Tugen Hills, to over 11,000 ft. above the Elgeyo Marakwet escarpment in the Cherangani Hills. To the west of these hills, the land falls away in a series of high plateaux in the south, bordered by the bulk of Mt. Elgon on the Uganda border, which at 14,170 ft. is Kenya's second highest mountain.

The Rift Valley floor rises rapidly as it approaches the equator. From about 3,000 ft. around the Lake Baringo and Lake Hannington areas, it rises to over 6,000 ft. in the Eldama Ravine, Nakuru and Njoro region, is still 6,000 ft. at Lake Naivasha, and only drops down to low altitudes again south of the Kedong Valley, south-west of Nairobi.

The highland areas bordering the Rift Valley near the equator are perhaps the most spectacular in Kenya. To the west there is the great Mau range/escarpment which is over 10,000 ft. at its highest point, and the high ground continues up to the Lake Nyanza basin to the west and south. To the east there is the Aberdare Range, which reaches 13,000 ft. Almost a continuation of

the Aberdares is Mt. Kenya and the Nyambeni Range, which border the Eastern plains.

The Rift Valley south west of Nairobi falls in altitude, and contains Lake Magadi, an almost lifeless water containing large deposits of soda. To the west the Nguruman Escarpment gives way to the Loita Hills and the Loita plains, which at 6,000 ft. form a high plateau area. The land rises again slightly at the Esoit Oloololo Escarpment, and then gradually falls again towards the Lake Nyanza basin. To the east of the Magadi Rift Valley there is another series of high plains, the Kadjiado and Athi plains, which reach as far as Nairobi in the north east and drop in altitude towards Mt. Kilimanjaro in the south east. East of these plains is the southward continuation of the Mt. Kenya land mass, which continues well into Kamba Land, exceeding 5,000 ft. in places. The southernmost hills of this land mass are the Chyulu Range and the Taita Hills, which are close to the Mt. Kilimanjaro, Pare mountains. The most bizarre feature of the highlands south west of Mt. Kenya is the Yatta Plateau, a massive and ancient lava flow, over two hundred miles long, ten to twenty miles wide, and flat at the top, falling a few feet every mile. At the western base of the Yatta flows the Athi or Galana River, which flows eventually into the sea at Malindi.

The Kenya Highlands are well watered, and there is a multitude of permanent rivers and streams. Major drainage systems are the Kerio and Turkwell in the north west, Ngwaso Ngiro, Tana and Athi/Galani in the east, Ewaso Ngiro and Mara in the south and south west, and the Nzoia and Yala in the west.

c) The Lake Nyanza basin

Lake Nyanza lies at 3,718 ft. and in Kenya the prominent Winam Gulf is surrounded by plains ranging from 4,000 to 5,000 ft. On the northern shores of the gulf there is a large area of swampland in the Saiya district, formed by the Yala estuary, and south of Kisumu the Kano plains stretch southwards, an area of poor drainage and black soils that are often flooded. South of these plains the relief becomes more broken, with hills in the Homa Bay vicinity reaching 7,000 ft. in the case of the Gwasi Hills. On the whole, the Winam basin is narrow, with altitude rising rapidly at all points eastwards, towards the Kisii highlands and the Mau in the south and the Nandi Escarpment in the centre and north.

North west of the basin lies the Luyia country, a flat plateau extending into Uganda, surrounded on the east and north by the Nandi, Uasin Gishu and Mt. Elgon highland areas. The slope of the Luyia plains decreases from over 5,000 ft. close to these highland areas, to under 4,000 ft. near the Uganda border south of Busia.

It is important to note that the Luo and Luyia areas are separated by prominent and near continuous precipitous escarpments from the higher plateaux that surround them, and are thus relatively isolated by these topographical barriers.

1.1.2 Climate and Rainfall

Kenya is famous for the temperate climate of its highlands. It is sobering to point out however, that 72 per cent of its surface area receives a mean rainfall of less than twenty inches a year and

only 15 per cent of the land enjoys regular rainfall of more than 30 inches. The proportion receiving more than 50 inches is only 3 per cent. (Griffiths 1962:79).

Rainfall in Kenya is dependent on the South East trade winds, which blow moist air currents over the country from April to December. Atmospheric moisture is also supplemented by Lake Nyanza, the surrounding areas of which receive additional rainfall.

The distribution of mean annual rainfall is given in Map 2. As much rainfall is orographically induced in Kenya, the distribution of rainfall contours follows closely the distribution of altitude contours. The higher the altitude, the greater the rainfall is likely to be. The highest rainfall zones tend to be at altitudes of over 6,000 ft. and regular rainfall is thus concentrated in the central highlands.

In the highlands east of the Rift Valley, particularly Mt. Kenya and the Aberdare Ranges, and to some extent on the Mau Range to the west, there is a pronounced rain shadow on the north eastern slopes, with the south eastern flanks receiving the highest rainfall.

The Kenya coast is exceptional in that it receives up to fifty inches of rainfall a year in spite of its low altitude. The highest rainfall occurs south of Mombasa, but as we proceed north towards Lamu it gradually falls to an annual average of 30 inches. Throughout the coastal strip moreover, rainfall falls rapidly to less than 20 inches a year beyond ten to twenty miles west into the interior from the coast.

The joint contribution of the South East Trades and the moist air streams from Lake Nyanza produces very high rainfall

indeed on the highlands bordering the Lake. In the Kakamega area over 70 inches a year is usual, and in the Kisii zone rainfall often exceeds 60 inches.

The high plateaux and plains south and west of Nairobi (Athi and Loita plains), though often above 5,000 ft., do not receive more than 30 inches a year, with an average nearer twenty. This is due to the fact that these areas lie at the periphery of the main moisture-bearing currents, and there is an absence of very high mountains to help cause the rainfall.

The low lying plains to the east and north are extremely arid, often receiving less than 20 inches a year and in the north and east, under 10 inches. The low zones east and south east of Lake Turkana, and in the Lodwar area of Turkanaland, receive under 5 inches of rainfall a year, and can be classed as true desert. Rainfall gradually increases in the far north, approaching the southern Ethiopian Highlands.

Mean annual rainfall however, is only a rough guide. Even when rainfall appears to be adequate (e.g. 30-40 inches), it can be erratic, with years of abundance interpersed by drought. Moreover rainfall can be very localised, with farmers suffering from a surfeit of rain while their neighbours, suffer drought. Every two or three decades furthermore, there can occur deluges in areas which normally receive little rain, causing disastrous flooding and erosion. The reliability of rainfall tends to increase as altitude increases, but it is only above 5,000 ft. on the pluvial side of mountains that regular yearly rainfall becomes a rule. The only exception is the Lake Nyanza basin, where drought is

virtually unknown.

In addition to affecting rainfall, altitude has a noticeable effect on temperature. The higher the altitude, the cooler it becomes. Frost becomes possible at altitudes greater than 7,000 ft., and snow above 14,000 ft. Permanent snow occurs only above 15,000 ft. and is confined to Mt. Kenya. In the low lying areas temperatures can occasionally exceed 120$^{o}$ Farenheit, and temperatures in the nineties are common during the day. In high altitude areas however (above 6,000 ft.), the maximum temperatures seldom exceed 80$^{o}$F., and the mean temperature is usually close to 60$^{o}$F.

### 1.1.3 Ecological zones and traditional modes of livelihood

The major ecological zones in Kenya can be delineated on the basis of the plant communities that characterise them. The composition of each plant community is primarily determined by the climatic conditions prevailing in an area, although soil conditions are also important for the detailed distribution of species within each vegetation zone.

In spite of its generalised character, the classification by Edwards and Bodgan (1951) and Edwards (1956) of vegetation zones in Kenya is one of the most useful, as it is accurate enough for general use without being over-detailed. Edwards distinguishes eleven major vegetation communities, shown in Map 3. These are described below, and the agricultural and pastoral potential of each zone is outlined.

1) Highland forest and grassland

This vegetation zone usually occurs between an altitude of 6,500 and 9,000 ft., and is characterised by large tracts of montane forest, interspaced by areas of rolling grassland and moorland where the forest has been cleared. Rainfall varies from 40 inches in the lower altitudes and areas of rain shadow, to 90 inches in the best watered slopes. Rainfall is supplemented by frequent mountain mists which play an important role in ensuring an adequate moisture level during periods of low rainfall. Temperature rarely exceeds $75^{o}$ or falls below $45^{o}$ Farenheit, although frost sometimes occurs at the higher altitudes.

The distribution of the zone (see Map 3), corresponds with the elevated land masses that flank both sides of the Rift Valley in central Kenya. East of the Rift, the forest zone predominates along the whole length of the Aberdare Range above 6,000 ft. in the east and 7,500 ft. in the west (where the rain shadow is). In the western slopes of the Aberdares, especially in the Kinankop and Thompson Falls regions, large areas of highland grassland exist where forest growth has been inhibited by the rain shadow, or where it has been cleared. Mt. Kenya is also surrounded by a broad belt of forest. In the eastern slopes which are particularly well watered, the forest, after extensive clearing now begins at 5,000 ft. In the north west of the mountain however, the rain shadow is pronounced, and the forest belt is thin. A large tract of grassland and moorland lies between 6,000 and 9,000 ft. In the north east, the Mt. Kenya forest continues uninterrupted along the Nyambeni range for a distance of over 60 miles.

The forest zone flanking the western escarpments occupies an even greater area. It stretches almost continously from the northern Cherangany and Pokot Hills to the Tanzanian border, with a short break in the Narok and Magadi localities. The width of the zone varies from 20-30 miles in the Elgeyo escarpment to over 70 miles in the Mau-Kisii region. An isolated but sizeable forest tract also occurs on the higher slopes of Mt. Elgon. Highland grassland and moorland is especially prevalent in the Uasin Gishu area where the forest belt was narrow. Extensive forest clearing has occurred in patches throughout the western zone, especially in the Tinderet, Molo, Nandi, Londiani and Kericho regions. In these places, grassland interspersed with patches of forest is still common.

Whilst the localities mentioned above form the main highland forest zone, there are also smaller mountains above 6,000 ft. dotted about Kenya (for example Mt. Kulal which have patches of forest on the higher slopes).

The character of the forest varies according to locality and prevailing soil and moisture conditions. On the higher mountains, it eventually gives way to bamboo at around 9,000 to 11,000 ft., followed by moorland and alpine vegetation which can continue up to the snow line. The large tracts of grassland are usually a direct consequence of forest clearing or persistent burning, accompanied by erosion, which removes the top layer of soil. Where burning and agriculture predominated, red oat grass and Kikuyu grass, in association with Kenya White Clover are the dominant species. These grasses are important, as they provide excellent grazing and a "productive sward" which maintains a high

level of soil fertility (Edwards 1956:95).

The forests are now a natural resource protected by the creation of forest reserves, which prohibits encroachment by cultivators. The highland forest zone however, has never been the most fertile or most productive agricultural zone, in spite of its high rainfall, except at the lower altitudes (5,000 to 7,000 ft.). This is because the cooler climate and the danger of frosts inhibit crop production. The allocated agricultural ex-forest zones are today predominantly devoted to cash crops such as tea, coffee, and pyrethrum. The grassland and moorland however, is valuable grazing, and traditionally forest clearings have always been used to graze cattle, even as far as the high moors above the tree line. Today, many of these areas are devoted to intensive rearing of European breeds of sheep, and cattle are still to be found grazing inside forest reserves.

2) Coastal forest and coastal bush and open high grass glades

This zone stretches from the coast up to thirty to fifty miles inland. The predominant conditions are seasonal rainfall of 30 to 50 inches with periods of high humidity, which encourage the growth of tropical forest of a distinctly West African type. Where rainfall is lower, or soil conditions are not right, thick bush interspersed by sizeable trees occurs, with Baobabs, species of Terminalia and Combretum a characteristic feature. The forest and bush zones are frequently interspersed by large grassland clearings, which have resulted from persistent burning. North of the Tana River, a great deal of the grassland is natural and marshy in character.

The forest is mostly confined to areas within twenty miles of the coast and is located in two main areas, the Shimba Hills south of Mombasa, and the Arabuko Sokoke forest south west of Malindi. The main areas of coastal bush woodland are east of Malindi and north of the Tana River to the west of Lamu. The immediate coastal strip is largely clear of woodland and bush. Towards the interior, the thick woodland bush degenerates rapidly into scrub as rainfall diminishes.

The coastal zone has a high agricultural potential, and people living in the zone grow a wide variety of agricultural produce, concentrating on maize, tropical fruits, sorghum, millet and palms. Cash crops include cashew nuts, sisal and kapok. Cattle are also reared in some localities. The main disadvantage to cultivation in the well watered zones is the diversity and richness of the insect and epedemiological fauna. The coastal zone is thus particularly prone to the scourges of agricultural and cattle diseases, and the debilitating effect of human disease, especially malaria.

3) Scattered tree grassland (low tree, high grass)

This zone is mainly situated between 3,000 and 6,000 ft., and consists of undulating high grass up to 8 ft. high and scattered woodland. The trees, which rarely exceed 15 ft. in height, are broad-leafed and deciduous. Of the many species that characterise this zone, the most numerous are those of the Combretum family, with those of Terminalia, Ficus, Faura and Heeria also common. The climatic conditions which help to maintain this plant community are high rainfall varying from 35 to 70 inches, and a warm, mean annual

temperature, rarely exceeding 85° Farenheit and seldom falling below 50°.

The grassland depends for its existence on periodic burning and clearing for agriculture. When burning is not practised, woodland rapidly engulfs the grassland. Where rainfall and ground water levels are exceptionally high, taller trees occur, up to 90 feet high. In some limited areas, forest of a tropical moist West African character is in evidence. In other favourable areas, especially between the 4,000 and 6,000 foot elevations south and east of Mt. Kenya and the Aberdares, there used to be a belt of thick low tree forest (patches still survive), forming a transitional zone between the Combretum woodland and the highland forest.

This plant community has an extensive distribution in Kenya. In western Kenya the Combretum woodland extends in a broad belt from the Endebess and Kwanza plains flanking the north eastern slopes of Mt. Elgon to the Lake Nyanza basin. A large portion of this area is composed of rolling grassy plains, which are poorly drained in parts, resulting in swampland. Belts of lush riverine forest also occur along the banks of rivers that drain the area. In the Kakamega and Bungoma districts, rainfall exceeds 70 inches, resulting in tropical forest around Kakamega and Malava. East of the Rift Valley, the distribution of this vegetation community is confined to the lower eastern and southern slopes of Mt. Kenya and the Aberdares. The woodland, from the evidence of remnant patches, must once have been extensively dense before it was cleared by the ancestors of the present Bantu inhabitants.

Where rainfall is regular and plentiful, and soil conditions are fertile, the zone is extremely productive for staple agricultural crops. This explains why it contains the highest density rural population in Kenya, in both the Western and Eastern areas. At the lower altitudes (3,000 to 4,000 ft.), when rainfall falls below 40 inches a year, agriculture becomes a risky enterprise. In good years a high yield can be obtained, but only too often drought afflicts these areas. These marginal areas are to be found in the lower slopes of Mt. Kenya and the Aberdares below 4,000 ft., and in the Busia area of Western Kenya.

4) Scattered tree grassland (Acacia-Themeda)

The Acacia-Themeda grasslands are the best known of the East African vegetation zones, being frequently featured on television wildlife films, where grassy plains with grazing gnus and zebra are interrupted by the sight of a flat-topped acacia. The Acacia-Themeda community occurs at medium altitudes of 4,000 to 6,500 ft., and is associated with erratic rainfall of 20 to 30 inches, with frequent droughts. The grassland is characterised by a dominant species, Themeda triandra, which is a grass seldom exceeding 5 feet in height. Scattered amongst the grass are species of acacia, of which the "whistling thorn" acacia (A. drepanolobium), 6 to 8 feet high, is particularly common. Many species of Acacia are to be found in the community, but they seldom grow higher than 20 feet, being stunted by the frequent burnings that maintain the grassland, and by poor drainage in the black cotton soils on which they commonly grow. In areas of good drainage and no burning, mostly around rivers, the acacias grow to a considerable size, exceeding fifty feet in height.

The distribution of the community is extensive. As Map 3 shows, the main areas are concentrated in and around the central and southern Rift Valley. One major zone is north west of Mt. Kenya in the Laikipia district. A smaller belt also occurs north of Mt. Elgon in Karasuk. The biggest zone however, is south of the equator, stretching down the Rift Valley from Nakuru to Suswa, branching out to the west over the Loita and Mara plains, and to the east over the Kaputie and Athi plains well past Machakos.

The erratic rainfall makes this vegetation zone unprofitable for cultivation, although sections around Machakos are used for crop production where the soil is better drained and rainfall more reliable. For the most part however, the characteristics of the zone makes it ideal for raising cattle, and large herds are found throughout. The droughts when widespread periodically result in the death of many cattle, through starvation and weakened resistance to disease caused by impoverished pasture rather than an absence of drinking water. Nevertheless, a few good years quickly recuperate the loss, as the graze is rich and abundant, given an adequate rain supply.

5) Desert and semi-desert zones

Edwards (1956) distinguishes four desert and semi-desert plant communities (see Map 3). "Desert Grass-Bush (dry bush with trees)" occupies by far the largest area of the low-lying plains, occupying most of the plains and low highlands south of the Tana River, and most of the north eastern plains near the Somali border. This community consists of large areas of thick bush, with deciduous Commiphora species predominating, accompanied by

species of Acacia and Terminalia in some areas. The bush, usually from 8 to 15 feet high, also contains larger trees in some localities. The distribution of this vegetation type is associated with rainfall from 10 to 15 inches, with extremely localised rainfall, and droughts lasting several years are not uncommon. Dessiccating winds also blow through a large part of the year. In the areas north east of the Tana the bush is interspersed with large areas of perennial grassland. South of the Tana the clearing of bush and the development of grassland has only begun in the last ten years, except in the areas east and north east of Mt. Kilimanjaro, which have been converted to grassland for a much longer period.

As rainfall decreases to under ten inches, desert scrub appears, consisting of a thin covering of stunted trees, still mainly Commiphora and Acacia, interspersed by large tracts of bare ground. The presence of grasses is ephemeral, the grasses depend on the erratic rainfall for the stimulation of growth. Where rainfall falls below 5 inches, the scrub and grass cover becomes very thin, and where soil and ground conditions are very unfavourable, there is very little vegetation.

The distribution of these desert communities is confined to the central and western northern areas. Most of Turkanaland west of Lake Turkana is semi-desert, and there is a large belt of desert and semi-desert east of Lake Turkana stretching from the Lake in a south easterly direction as far as the Tana River. The most arid zones are south, north and west of Mt. Marsabit, where the combination of lava and very low rainfall is extremely hostile to plant growth.

These arid environments are for the most part totally unsuitable for any form of agriculture. The soils however, are generally fertile, and cultivation is possible along the banks of permanent rivers such as the Tana and Turkwell. Traditionally, cultivation along the banks of these rivers has depended on flooding during the rainy season, but today irrigation is also beginning to play an important role.

Outside these very limited areas only pastoralism is possible. Paradoxically, the better watered dry Commiphora woodland, which provides excellent browse for camels and goats also encourages the breeding of tsetse flies, and is thus mostly uninhabited where the bush is very thick. South of the Tana River moreover, the absence of permanent water away from the Tana and Galana Rivers is also a considerable limitation to pastoralism. North of the Tana the bush is associated with much larger areas of relatively fly-free grassland, and sources of water are more plentiful. The dry bush zone in the north thus supports a much higher density of cattle, sheep, goats and camels than in the south.

The semi-desert scrub only supports cattle in the better watered areas, such as the Merti Plateau. For most of its distribution cattle are uncommon, and the predominant economy is based on small stock and camels. East of Lake Turkana it is so arid that the local Rendille people keep no cattle at all. In Turkanaland, cattle are also virtually absent from the semi-desert and desert plains, and are kept only on the higher slopes of the mountain ranges. Camels, small livestock and donkeys however, are a common sight on the plains.

## 1.2 The Ethnic and Linguistic Diversity of the Peoples of Kenya and their Detailed Distribution

The indigenous peoples of Kenya can, on the basis of their linguistic characteristics, be classified into three African language families, the Niger-Congo, Chari-Nile and Afroasiatic following Greenberg's nomenclature. (Greenberg 1963). These language families are analogous to "orders" in zoological taxonomy, and encompass many sub-divisions at different hierarchical levels, before we reach the minimal unit which is a named tribal population (e.g. the Nandi).

These linguistic families are, as their names imply, distributed widely in Africa. The Afroasiatic group extends even outside Africa to the Middle East (the Semitic languages are part of the family). In spite of the widespread distribution however, each language family is associated with peoples who follow distinctive lifestyles. The Niger-Congo speakers tend to be agriculturalists and are associated with lush savannah and tropical forests, the Chari-Nile speakers are predominantly cattle pastoralists in more arid Savannah, and the Afroasiatic speakers are agriculturalists and/or pastoralists, associated with highland and arid semi-desert. A fourth family, the Khoisan, which also occurs in East Africa, though it is not confirmed in Kenya, consists of the ancient languages of Bushmanoid hunting and gathering peoples. East Africa owes a great deal of its linguistic diversity to its geographical position, which is peripheral to the main centres of all these families, and also to its topographical and climatic diversity, which produces conditions favourable to the preferred lifestyles of peoples associated with each family.

The precise linguistic affinities of Kenya's indigenous peoples are outlined in Table 1.1. Their general distribution is given in Map 4, and their detailed distribution and characteristics are discussed more fully below.

1.2.1 Peoples of the Niger-Congo family

As Table 1.1 reveals, the Bantu linguistic group is the only Kenyan representative of the Benue-Congo division, which is in turn a major sub-division of the Niger-Congo family. Of the many Bantu sub-groups, which can be found from the Cameroons right down to South Africa, three are to be found in Kenya. These are the Interlacustrine, Highland and Coastal Bantu sub-groups.

The Bantu are traditionally agriculturalists through much of their range, practising small scale slash and burn cultivation in the wooded and forest environments that they usually inhabit. As is to be expected given the vast area of their distribution, there is a great deal of social and cultural variation. For instance, some Bantu peoples, like the Gisu of Uganda, are a decentralised collection of villages based on patrilineages, while at the opposite extreme are peoples such as the Lozi of Zambia who are organised in villages based on matrilineages, under the central authority of a king. In Kenya however, the Bantu peoples tend to follow the Gisu type of primary organisation, modified in various areas by cultural features borrowed from non-Bantu neighbours such as age sets, circumcision and clitoridectomy.

1.2.1.1 The Inter-Lacustrine Bantu

The central distribution of this large Bantu division is in Uganda, where it represented predominantly by large, sophisticated

and powerful kingdoms such as those of Nyore, Ganda and Toro. The Kenyan members of this group lie at the eastern periphery of the distribution, and unlike many of the Ugandan representatives of the group are not centralised kingdoms - except for one people, the Wanga.

The linguistic similarity of the Inter-lacustrine Bantu in Kenya and the close regional proximity to each other in the area north of the Winam Gulf, have led them to assume a common identity in recent years. They are now known as "Luyia", although in the past they were 18 autonomous peoples believed to be of heterogenous origins.

The distribution of the Luyia peoples is shown in Map 4.1. To the south are the Isukha, Idakho, Tiriki, Maragoli, Bunyore and Kisa. The population density for these peoples is one of the highest in Kenya, over 400 people a square mile through most of this area. The high rainfall makes it possible for this large concentration of peoples to eke out a living from the few acres they possess, in spite of the fact that the boulder strewn soils are not as fertile as other areas of the Luyia country. The Tiriki stand out culturally from their neighbours in having age-sets, a custom borrowed from the bordering Nyangori and Nandi (who are Nilotes - see below).

In the centre of the Luyia country lies the small former kingdom of Wanga, and its neighbours the Marama and Butsotso. The Wanga area is bisected by the Nzoia river, and is relatively swampy. As a result, the population density is lower than in the land further south, and in recent years, the traditional subsistence

agriculture has been supplemented by the cultivation of sugar cane for the Mumias sugar refinery.

To the west and south near the Uganda border the country becomes more arid and population density declines appreciably. The Southern Bunyala and Busonga live on the estuary of the Yala, and the country is very swampy, permitting the cultivation of rice as well as maize. Fishing is also an important source of revenue and food. The Samia and Bukhayo to the north grow cotton as a cash crop, as well as the usual staples, and as they overlap into Uganda, smuggling thrives.

Finally, the northern Luyia are represented by the Bakusu (by far the largest Luyia population), the Tachoni, Kabras and Northern Bunyala. In the northern regions the land is both fertile and relatively unpopulated, permitting a greater acreage for each family. Cattle moreover, are far more in evidence than further south. The Bukusu are closely related to the Gisu of Uganda, and in recent times they have expanded into the Trans Nzoia District as a result of settlement schemes on former white-owned land. In addition, development in the form of a multi-million dollar paper mill on the Nzoia River at Webuye has come to the Bukusu area. The mill lies close to the ruins of Chetambe Fort where Bukusu resistors put up a hopeless last ditch stand against the British in 1865. The Tachoni are somewhat anomalous, being predominantly of Kalenjin ancestry, although they have been culturally and linguistically assimilated by the Luyia. (Were 1967: 63).

1.2.1.2 The Highland Bantu

Of the 16 major Highland Bantu populations listed by Sutton (1974:82), half are in Kenya and the rest in northern Tanzania. In Kenya the centre of their distribution is around Mt. Kenya and associated highlands, but the sub-group is also represented on the south Kenya coast by the Segeju, and in the south west by the Kuria, Gusii and Suba (See Map 4).

i) The Central Bantu

This sub-group refers to the main concentration of Highland Bantu around Mt. Kenya, and is more a geographical group than a strict linguistic sub-division. Like the Luyia, the Central Bantu peoples share a strong sense of common identity and represent one of the major ethnic alliances in the present day. In pre-colonial times however, the Central Bantu comprised 15 autonomous peoples. Their distribution is outlined in detail in Map 4.2.

Demographically, the Central Bantu can be divided into three main groups. The first group consists of those peoples who live mainly between the 4,000 to 6,000 ft. contours. These are the Kikuyu on the southern flanks of the Aberdares, the Ndia, Gicugu and Embu on the southern slopes of Mt. Kenya, and the Chuka on the south east of Mt. Kenya. All these populations inhabit the best watered parts of the Combretum high grass zone, and enjoy some of the most fertile and productive lands in Kenya. Population density in this zone is extremely high, and a wide range of staple crops are grown, including beans which are very characteristic of this area. Today moreover, cash crops (coffee, pyrethrum, tea and citrus fruits) are a common sight, as are European breeds of cattle in the southern Kikuyu areas near the highland forest zone.

The second group of populations inhabit the lower altitude and rainfall zones (2,000 to 4,000 ft., 20 to 40 inches). These are the Mbere and Tharaka, who show a relatively low population density, and are regularly afflicted by droughts. The grazing of small livestock is particularly important, and in times of hunger, hunting, fishing and gathering are also practised. The government is currently encouraging the growing of cotton and, where irrigation is possible, rice as cash crops.

The third group are those populations who have sections in both the high and lower altitude zones. These are the Kamba, Muthambi, Mwimbi, Igoji, Imenti, Tigania and Igembe. The Imenti, Tigania and Igembe are particularly favoured, in that not only do they have access to the lower and high slopes of the Nyambeni Range, where low and high altitude crops are grown, but they also possess considerable grazing concessions in the plains south and east of Isiolo. The grazing of cattle is thus an important aspect of their economy.

The Kamba deserve a special mention, as they cover by far the largest area of any Central Bantu population (see Map 4). The Kamba themselves distinguish three major divisions of their land, the Machakos area, with high hills, medium rainfall where population density is high, and intensive cultivation is practised, the Kitui section where access to the dry bush country permits the grazing of cattle as well as marginal agriculture, and southern Kamba where hunting and gathering used to be essential to supplement the meagre yield of the land. Most of the Kamba country is unfortunately marginal for both agriculture and grazing. It is only comparatively recently that the Kamba have begun to share the richer grazing areas

of the Athi plains that border Machakos, which were denied to them by firstly Maasai and subsequently European occupation.

The Kikuyu are also a complex people, who today at over 2,000,000 are Kenya's largest "tribe", much bigger than any other Central Bantu people except the Kamba. Even before British rule, the Kikuyu country was thickly populated, and there was a growing difference between the people south and north of the Chania River. (Lambert 1949:4). North of the Chania there was also some degree of separateness between the southern and northern inhabitants. These three regional divisions are known as the Karuara, Metume and Gaki sections by the Kikuyu, and correspond in area to the present Kiambu, Muranga and Nyeri Districts.

The recent history of the Kikuyu is one of a high rate of expansion and development. Rural development accelerated in the early 1950's when the individual ownership of land was permitted within the "reserve" by the Colonial administration, giving landowners the incentive to grow cash crops with enthusiasm. Their proximity to the urban and industrial areas of Nairobi and Thika also precipitated considerable participation in industry and urban entrepreneurship. Chronic land hunger led Kikuyu families to spread en masse into the Nyandarua District west of the Aberdares, first as labourers to European farms, and after independence as occupiers under settlement schemes. Kikuyu farmers are now to be found in pockets throughout Kenya, and there is no urban centre without a significant Kikuyu population of administrators, traders and entrepreneurs.

From the point of view of linguistic and geographical affinity, the Central Bantu fall into two blocks, the Kikuyu and

Meru clusters. Closely related to the Kikuyu are the Ndia and Gicugu in Kirinyaga District, who call themselves Kikuyu now if asked to reveal their identity. The Embu, Mbere and Kamba are also closely allied to the Kikuyu in language and identity, although they preserve their separate status. Of the nine Meru peoples, the Muthambi, Mwimbi, Igoji, Imenti, Tigania and Igembe seldom call themselves anything other than Meru nowadays. The Chuka however, have always been proud of their separate identity, having often been in a state of war with the other Meru peoples. They are the only Meru people to claim aboriginal origins (Lambert 1949;3). The Tharaka are also distinct from the other Meru, especially in their lifestyle which resembles more that of the Kamba. Lambert does not consider them to be true members of the "Meru proper" (Lambert 1949:3).

The Central Bantu peoples, though socially organised on an acephalous patrilineal basis, have borrowed many customs from their Maasai and Cushitic neighbours. Particularly distinctive for Kenyan Bantu is the high state of development of age-sets and traditional military organisation, and the prevalence of circumcision and clitoridectomy.

ii) The Coastal and Western Highland Bantu

The Segeju on the Kenya coast near the Tanzania border (see Map 4.3) have distinct linguistic affinities with the Kikuyu cluster. In culture and lifestyle however, they are very similar to the other coastal peoples described in the next sub-section.

The western Highland Bantu are represented by the Gusii, Kuria and Suba (see Map 4), a group that extends well into Tanzania.

The Kuria and Gusii occupy highland south west of the Winam Gulf, and much of the area is typical of the highland grassland and forest zone. Population density is high and subsistence agriculture is the prevalent form of livelihood. The area is relatively under-developed compared to other Bantu regions, and cash crops are not nearly so common as in the Kikuyu and Meru country. The Suba however, are a small people, and live on the Islands of Lake Nyanza (Mufinga, Rusinga, Takawiri and Bukula Islands). They can be considered a remnant of the Gusii-Kuria peoples which were cut off by the Luo penetration of the Lake Nyanza basin in the nineteenth century. Their relative isolation on the islands and the low Luo density on the lake shore opposite the islands, which is mountainous and infected with tsetse, saved them from otherwise inevitable assimilation.

The Gusii and Kuria, though linguistically related to the Central Bantu, are more similar culturally to the Luyia in the north. Oral tradition, as Were (1974:189) points out, even states an origin in the Luyia area for the Gusii, but it is difficult to reconcile this with the linguistic evidence.

1.2.1.3 The Coastal and Hinterland Bantu

This is another large Bantu group, which overlaps well into Tanzania. In Kenya, it is represented by four major divisions, the Mji Kenda and Swahili peoples of the coast, the Pokomo of the Tana River, and the Taita and Taveta of the highlands east of Mt. Kilimanjaro.

i) The Swahili group

The Swahili are the most ancient Bantu inhabitants of the coast, and are distributed in a thin strip of coastline near Mombasa and Malindi. In Lamu, there is a closely allied population, the Bajun.

The Swahili language is a hybridised blend of a Bantu grammatical and vocabulary base, supplemented by hundreds of Arabic words. The language reflects the long and close relationship between the Swahili peoples and Arabs, who settled and traded along the coast for centuries. During its long contact with the Arabs, the Swahili community became increasingly hybridised both culturally and physically, and was expanded with the assimilation of ex-slaves. Today the Swahili are highly Islamicised, and tend to be predominantly urban, though they also practice some agriculture in the vicinity of the old Arab and Swahili towns such as Lamu, Malindi and Mombasa.

ii) The Mji Kenda Peoples

The Mji Kenda are a heterogeneous cluster of nine peoples, the Digo, Duruma, Kauma, Kambe, Ribe, Rabai, Jibana, Chonyi and Giriama, who assumed a common identity in colonial times. Their detailed distribution is given in Map 4.3. It should be noted however, that the boundaries indicated on the map are somewhat arbitrary, and there is considerable inter-mingling of peoples in close proximity.

Most of the Mji Kenda, apart from the Duruma and western Giriama, inhabit the best-watered sections of the coastal plain. The picture presented up to 15 to 20 miles inland is a series of

small-holdings and maize fields, under a canopy of mango, coconut, kapok and cashew nut trees. The coconuts, as well as providing a steady source of revenue from servicing the tourist coastal hotels and the sale of copra, are used for making palm wine which is of considerable social significance (see Parkin 1972). Mangos, kapok and cashew nuts are also valuable cash crops. In some localities cotton is also grown, and recently there have been experiments with dairy farms to provide milk for the tourist hotels. Fishing along the coast is also practised in some areas. Particularly valuable is the collection of shells which fetch good prices in the local tourist markets.

The Mji Kenda peoples, owing to the proximity of the sophisticated Arab/Swahili towns, have enjoyed the benefits of trade for many generations, an economic opportunity denied to most Bantu peoples of the interior. As well as servicing the food requirements of the Arab towns, they used to act as middle men for ivory, procured by the hunting groups of the coastal interior.

The Duruma and western Giriama live in the more arid parts of the coastal plains, and their environment is mostly dry bush and woodland. For these people, drought and the loss of crops are frequent hazards, and in times of emergency hunting, trapping, gathering and fishing are necessary for survival. The range of animals eaten by the Giriama for instance, is a source of amazement for visiting Bantu peoples from the interior.

Considering the length of time the Mji Kenda have been in contact with the Arab and Swahili peoples, they have been remarkably unaffected by Islamicisation. The most affected are the Digo south of Mombasa, large numbers of whom are now Muslims. North of Mombasa

however, especially in Giriama land, most people still practise their traditional religions and customs.

iii) The Pokomo peoples

The Pokomo are usually classed as one tribe in the literature. They consist however, of thirteen peoples who live exclusively on the banks of the Tana River from its estuary up to about 200 miles inland. Their identity and distribution is given in Map 4.4. The map only gives their rank position along the river, as I have not been able to find any information regarding precise boundaries.

The Pokomo live for most of their inland range in a harsh, arid climate, with rainfall erratic and seldom exceeding 15 inches a year. The plains flanking the Tana are amongst the most arid in Kenya. The Pokomo nevertheless are successful agriculturalists. Cultivation is possible during the rainy season in those parts where the river floods, and the Pokomo are proficient at maximising production during the growing season without resorting to irrigation. This is achieved through careful selection of fields. In many parts of the Tana, large tracts of riverine forest occur, and these play an important part in retaining moisture in the soil.

Apart from limited agriculture however, the Pokomo have had little opportunity to try out other forms of livelihood. Intermingled with the Pokomo are the Cushitic Orma on the south bank of the Tana and Somali on the north, who herd cattle on the surrounding plains and on those parts of the river shore that are not cultivated. Their superior military organisation and their fanatical love of cattle has ensured for them a complete monopoly of cattle ownership.

The close proximity of the Orma has resulted in considerable intermixture in the northern areas. The Malakote are probably the most hybridised of the Pokomo Bantu speaking peoples, and north of the Malakote, the Korokoro, once Pokomo speakers, now speak a dialect of Orma. One consequence of contact with the Somali has been the spread of Islam, which is now almost complete. One manifestation of this is the large numbers of Koran charms to be found impaled in fields of growing crops. In spite of the spread of Islam, the traditional culture is still very much alive. For instance, Pokomoland is one of the few places in Kenya where drums may beat throughout the night, accompanying traditional ceremonies.

iv) The Taita and Taveta

The Taita and Taveta are often classed together in one group, but in reality the similarity between them is superficial. The Taveta live on the eastern lower slopes of Mt. Kilimanjaro, and claim to be related to the Tanzanian Chagga, whilst the Taita have a long history of uninterrupted occupation of the Taita Hills near Voi, an oasis of highland in a sea of bush.

The Taveta occupy a strategic point along the Tanzania/Kenya border, which was an important staging post for caravans and traders in the nineteenth century. The present inhabitants are therefore most probably a heterogeneous amalgam of groups from the neighbouring areas, predominantly Chagga and Pare, with some Maasai (Dundas 1924) During colonial times, Taveta Town became an important railway link with Tanzania, and numbers of Luo and Kamba labourers settled there to work on Greek-owned sisal estates.

The Taita are composed of three regional sections, the Dabida, Sagalla and Kasigau, which inhabit respectively the upper slopes of the Taita, Sagalla and Kasigau Hills, which exceed 5,000 ft. Each range of hills is separated from the others by low lying, arid plains. The largest of the sections is the Dabida, which also displays the highest population density. The valley between the Dabida and Sagalla is intersected by the Mombasa/Nairobi and Voi/ Taveta railways. Voi, at the base of the Sagalla Hills, is an important town and contains a sizeable population of migrants from up-country. Luo, Kamba and Luyia labourers are also established on the large sisal estates near Voi and Mwatate. The migrants however, are on the plains, and have little contact with the main population concentrations on the upper slopes.

The smallest and most isolated of the Taita peoples is the Kasigau, who live on the upper slopes of Kasigau Hill some fifty miles to the south east. They are one of the most isolated communities in Kenya, surrounded on all sides by miles of uninhabited dry bushland.

In the last ten years, some attempt has been made to develop the uninhabited plains around the Taita Hills. Experimental cooperative Taita Ranches have now been established in the area, and the building of Taita Hills Salt Lick Lodge has brought in tourists. The effects of these developments however, are still marginal for the majority of Taita.

### 1.2.2 The Chari-Nile Family

This complex linguistic family is represented by only one division in Kenya, the Eastern Sudanic, and of this division, only

one sub-branch, the Nilotic Group occurs. (See Table 1.1). All three of the Nilotic sub-groups that are represented in East Africa are found in Kenya. These are the River-Lake Nilotes, Plains Nilotes, and Highland Nilotes. Until recently, the Plains and Highlands Nilotes were not considered true Nilotes, but were referred to as Nilo-Hamitic, and when the term Hamitic fell into disrepute (see Sutton 1974:96), were sometimes referred to as Paranilotes. The justification for the "Nilo-Hamitic" classification was the observation that, both culturally and physically, the peoples encompassed by the term seem to display a hybridised mixture of Nilotic and Hamitic characteristics. Sutton however, states the modern view when he points out that

> "recent studies of their languages show that they are basically Nilotic, and that any Cushitic word borrowings are superficial. Some of them (notably Kalenjin and Maasai, have assimilated numbers of Cushites and have also made cultural borrowings from them, but this is irrelevant to a strictly linguistic classification that we insist on using here." (Sutton 1974:97).

Indeed there is no choice but use a linguistic classification, as there is no really objective method of classifying peoples on the basis of cultural characteristics. Besides, the Central Bantu have also borrowed many Cushitic cultural characteristics, but nobody refers to them as Bantu-Cushites, or Parabantu.

1.2.2.1 The River-Lake Nilotes

The classic type populations of this group are the Nuer and Dinka that inhabit the Nile savannahs in the Sudan. They are predominantly cattle people who also hunt and fish at favourable times of the year. They are traditionally organised into acephalous segmentary lineages, which help to define territorial groups (Evans-Pritchard 1940).

The only representatives of the River Lake-Nilotes in Kenya are the Luo, who belong to the Lwoo linguistic sub-group. They occupy a large part of the Lake Nyanza basin around the Winam Gulf (see Map 4), a climatic zone of moist grasslands, woodland and swamp not too dissimilar to their ancestral environments in the Sudan.

The Luo though considered as one "tribe", are in reality a collection of over 20 units that operated autonomously in the nineteenth century. The precise identity and number of these "tribes" is difficult to establish from the literature. Ogot (1967) for instance, does not give a complete list in his "History of the Southern Luo", by far the most detailed account of this people. The list compiled in Table 1.1 is thus provisional, and needs verifying. The distribution of these "tribes" is given in Map 4.5, and the boundaries are based on those of Luo Locations given by Morgan and Shaffer (1966:30).

The predominant mode of livelihood for the Luo used to be cattle-keeping. Population density however, is now so high around the Winam Gulf basin that there is little land for large-scale grazing. The dominant mode of livelihood has therefore been replaced by agriculture, with cattle grazed where there is room. In the south moreover, tsetse flies are common, and bovine sleeping sickness has discouraged cattle keeping. Occasional outbreaks of human sleeping sickness have also been recorded in the Luo area.

Around the Lake shore, fishing is also an important activity. Since independence, several fishing cooperatives have been established to exploit the fish commercially. Sugar cane is another common cash crop, supplementing large scale company

plantations in the Nyando Valley north east of Kisumu.

The Luo are Kenya's second largest "tribe" after the Kikuyu, and are to be found in all large urban centres, where they fill a multiplicity of occupations. On the whole, their greatest contribution to the modern nation has been in the field of labour, where most trade unions are organised by Luos. They are also prominantly active in the railways.

1.2.2.2 The Plains Nilotes

The Plains Nilotes traditionally occupy the drier savannahs, which can support a fair number of cattle, but where rainfall is erratic enough to discourage agriculture. The scattered tree grassland (Acacia-Themeda type) is a typical environment, although they also occupy semi-desert bush if not too dry. Of the three clusters found in East Africa the Bari are exclusively in the southern Sudan and northern Uganda, the Itunga/Karimojong straddle eastern Uganda and the Sudan, overlapping into Kenya, and the Maasai cluster is in Kenya and northern Tanzania.

i) The Karimojong Cluster

Kenya has only two representatives of this group, the Iteso south west of Mt. Elgon, an offshoot of the Uganda Teso, and the Turkana west of Lake Turkana (see Map 4).

The Iteso occupy the lush grassland and wooded savannah that is typical of the Luyia area. They migrated to the area in the mid and late nineteenth century from the main Teso group north west of Mt. Elgon. Although an enthusiastic cattle people, they live in a rich agricultural environment, and have not hesitated to exploit agricultural methods of subsistence and revenue. The

high potential of the Iteso environment for both pastoralism and agriculture is the reason why the population density for the Iteso area, though lower than that of their Bantu neighbours, is by far the highest of any other Plains Nilotic people.

Turkanaland in contrast has a population density which ranks amongst the lowest in Kenya, consisting of a vast area of low lying plains, interspersed by areas of high ground (see Map 1). The Turkana prefer to live near sources of water on the hot plains, where they subsist mainly on milk from goats, camels and donkeys, with their diet supplemented with edible wild vegetables and roots which they gather and any mammals they can hunt. The Turkana are organised into family villages, collections of picturesque bee-hive huts normally inhabited by the elder generation, women and children. The younger men however, have a much more nomadic existence. A minority are detailed to herd the camels, donkeys and goats on the low plains, while the majority look after and guard the considerable herds of cattle, which as they cannot find adequate graze on the plains, are confined to the better watered higher areas. Most of the men therefore, are often to be found many miles away from their home villages.

The Turkana territory is on the whole so hostile that protracted droughts can cause major human tragedy. Loss of livestock so common during droughts means not just impoverishment, but death to a man and his dependents. Most poor Turkana today have a chance of safety when they lose their herds through either migrating as labourers to the towns and highland farms, or by becoming fishermen on the shores of Lake Turkana. Some Turkana

families have even begun to cultivate along the banks of River Turkwell.

In spite of the poverty of the environment and the extremely decentralised social system, one of the least cohesive in Kenya (Mair 1962), the Turkana are remarkably successful as a people. They successfully expanded their territory at the expense of the seemingly far more organised Samburu (Maasai) and the Kalenjin Pokot in the nineteenth century, and since independence in 1963 have succeeded in infiltrating and possessing a great deal of traditional Samburuland on the eastern Rift Valley south of Lake Turkana. Moreover their seemingly undeveloped and traditionally picturesque lifestyle is misleading. Many Turkana fought for the Kings African Rifles in Burmah, and have spent time as labourers on European farms before returning to Turkana. Some Turkana (mostly disaffected soldiers of the King's African Rifles who resented their Kikuyu officers after independence), have even formed a sophisticated bandit movement called the Ngoroko, which specialise in cattle rustling over four international frontiers and are reputed to possess modern rifles and machine guns. The Turkana on the whole are not nearly so out of touch with the twentieth century as their traditional appearance suggests.

ii) The Maasai Peoples

The Maasai peoples in the nineteenth century occupied all the floor of the Rift Valley and much of the plateau land off it from the shores of Lake Turkana as far down as the miombo woodland of southern Tanzania. At the height of their power however, around the 1840s, different Maasai peoples embarked on a series of disastrous

civil wars, which saw the annihilation of the Uasin Gishu agricultural Maasai and a large section of the Laikipia Maasai by the Purko and other southern sections. Their population was further decimated in the late nineteenth century by epidemics which afflicted both humans and cattle. By the time the British took on the administration of Kenya, there were large empty areas in the northern Rift Valley, which were alienated for European settlement, with the remaining Maasai pushed backwards to the south. In 1913 the remaining Laikipia Maasai were "persuaded" peacefully to evacuate the Laikipia plateau, and were re-settled in the area west of the Mara Game Reserve near the Gusii and Kuria. The Maasai were thus split into two geographical groups, a northern group consisting of the Samburu and small Maasai speaking Dorobo groups (Ndigri, Momonyot, Luasi, Ngwesi), and a larger southern group consisting of the dispossessed Laikipiak, the Purko, Loita, Isiria, Damat and Moitanik from Narok District, Kaputie, Keekonyokie, Kelani and Kankere from Kadjiado District (see Table 1.1 and Map 4).

With the extinction of the Uasin Gishu Agricultural section, all the present Kenya Maasai are pastoralists, relying mostly on cattle. Although deprived of some of their best grazing lands, they use the plains south and west of Nairobi which are still some of the best grazing areas in Kenya. Only the Samburu in the north have substantial areas of arid bush within their area, although this is compensated by lush grazing in the higher latitudes around Maralal.

The Maasai have often been described as pround and resisting change and development in favour of their traditional lifestyle. This however, is not really true. The Maasai were resistant to

change advocated by the colonial adminstration, as they distrusted their actions after the early land alienation. Since independence, the Maasai have encouraged development, and schools and roads are slowly blossoming in all areas. The Kaputie Maasai have participated in a government-sponsored ranching scheme, and the Purko Maasai have leased acres of prime agricultural land near the Mau Forest for wheat growing, with the local council in Narok directly benefitting. As Ole Sankan, former lecturer in History at the University of Nairobi and himself a Maasai points out,

> "attitudes have begun to change rapidly, with the Maasai in certain areas accepting the fact that ultimately they will have to settle down permanently in definite areas, and properly demarcate their land, accepting title deeds." (Ole Sankan 1971:xxxix).

1.2.2.3 The Highland Nilotes

The Highland Nilotes are so called because they tend to live at high altitudes in the Highland Forest and Grassland zones. There are two linguistic branches of this family, the Dadog group which is now confined to Tanzania, and the Kalenjin peoples which are all in Kenya except for one population, the Sebei who occupy the northern slopes of Mt. Elgon in Uganda.

The Kalenjin are traditionally confined to the highland areas on the western escarpments of the Rift Valley and associated land masses, and in the northern reaches are also found on the Rift Valley floor. Geographically and linguistically it is possible to distinguish three clusters, the Sabaot on the slopes of Mt. Elgon, the Pokot, and the Nandi group.

The Nandi group is the largest cluster, and contains eight peoples, which from north to south are the Endo, Marakwet, Tugen,

Keyo, Nandi, Nyangori, Kipsigis and Ogiek. The Endo, Marakwet, Keyo and Tugen inhabit the highlands of the Cherangani and Tugen Hills, where they herd cattle in the forest glades and grow staple crops along the forest edges. The Tugen however, stand out in occupying a large area of the Rift Valley floor south and west of Lake Baringo, which they mainly use for grazing cattle. The general pattern is for the main population to live in the higher altitude zones, while the young men herd cattle below in the plains. On Lake Baringo at Kampi-ya-samaki there is now a Tugen fishing village, originally sponsored by the Government to alleviate the plight of landless Tugen.

The Nandi and Kipsigis, who occupy respectively the Nandi Plateau and the western slopes of the Mau Range below the forest, are the most developed Kalenjin in Kenya. The Kipsigis area borders the great Brook Bond tea estates around Kericho, and there are also large tea estates near the Nandi Hills zone in the Nandi country. A visitor to the Nandi and Kipsigis areas would be struck too by the number of European breeds of cattle to be found there, as well as the mounds of potatoes on sale on the road sides.

The Nandi and Kipsigis had a formidable military system in the nineteenth century, based on age sets and a system of alliances between territorial sections (Peristiany 1939). In the first decade of this century, the Nandi in particular terrorised the Uganda railway, and it took the British twelve years to subdue their resistance. Even today they are natural soldiers, and form a considerable proportion of the Kenya Rifles. In an uneasy alliance with the Kikuyu, they have come to dominate the Nandi

sections of the Rift Valley previously allocated for White settlement (especially the Uasin Gishu and Cherangany portions of the Trans Nzoia District). With the other Kalenjin peoples they are now a powerfulnational political bloc, and have produced Kenya's second president Mr. Daniel Arap Moi (who is a Tugen). Finally they have also produced the majority of Kenya's Olympic athletes.

The Nyangori to the north west of the Nandi plateau are unusual in that they have inter-bred significantly with the Luyia, especially the Teriki whom they now resemble closely culturally. Scattered about the Mau and Tinderet forests moreover, are bands of hunters and gatherers called the Ogiek group, who speak Kalenjin (these will be discussed in section D below).

The Pokot occupy a large area of the northern Cherangany Hills, and the surrounding plains north of Lake Baringo and in the Sigor and Karasuk areas. Although they speak a Kalenjin language, they have borrowed many of the customs of the Turkana and Karimojong peoples whom they border. Like the Tugen they have a highland agricultural section, in the Kapenguria District, and a pastoral section in the surrounding plains. Unlike the Tugen however, the lowland zone is inhabited permanently by families who do not migrate backwards and forwards from the hills. The Pokot like the Turkana were until recently one of the least changed peoples of Kenya. In the last ten years however, they have responded (not without trauma), to vigorous government propoganda to develop. In spite of this however, theirs is still a very poor area.

Finally, the Sabaot peoples live on the southern and northern slopes of Mt. Elgon. In Kenya the main representatives are the Kony, Pok and Bungomek, small isolated groups on the forest

edge of the mountain where they mainly cultivate staples such as maize. All these peoples are surrounded by Bukusu of the Luyia group, and it is probable that there has been considerable inter-marriage in recent years. The Kony, Pok and Bungomek however, (see Map 4.1), remain poor and underdeveloped, receiving little of the national cake.

1.2.3 The Afroasiatic family

This family in East Africa and the Horn is predominantly represented by five divisions of Cushitic speakers, whose main concentration is in Ethiopia and Somalia. The Cushites are more like black Caucasoids than negroes in external appearance, possessing fine facial features and long hair. The ancient Christian Empire of Ethiopia and the strong Islamic Somali culture testifies to their long proximity to the Middle Eastern and Eastern Mediterranean civilisations.

In Kenya and Tanzania there are two Cushitic divisions, the Southern and Eastern Cushites. The Southern Cushites were the first Neolithic inhabitants of East Africa, and pre-date all the Bantu and Nilotic peoples. Today they have been reduced to small groups in the northern Tanzania Rift Valley, with only one representative in Kenya, the Dahalo of the lower Tana River (see Map 4.4). All except the Iraqw are in danger of dying out through being absorbed by their more numerous neighbours (Sutton 1974:84).

The Eastern Cushites however, are relatively recent migrants to Kenya, having arrived in the last few hundred years, and occupy most of the low lying plains east of Lake Turkana and north of the Galana River. Their traditional environments are arid dry bush and

semi-desert zones, where they herd camels and small livestock and cattle where the conditions are favourable. Although they display the decentralised political and social structures so typical of Kenyan pastoralists, their military organisation based on age sets was once highly developed, and these peoples had a reputation of being ferocious fighters. Their Cushitic customs of circumcision, clitoridectomy and a dislike of eating fish have been adopted to varying extents by many of the Maasai and Kalenjin Nilotes, and by the Central Bantu.

In Kenya it is possible to distinguish four main Eastern Cushitic sub-groups, the Galla, the Somali, the Mukogodo and the Burji. The latter two are small isolated groups.

1.2.3.1 The Galla Cluster

The Galla group consists of four peoples, the Gabbra from the north of Kenya, the Boran from south and north-east of Mt. Marsabit, the Sakuye from the same general areas as the Boran, and the Orma south of the Tana River. (See Map 4). The Sakuye are only cautiously included here in the Galla group, as I have not been able to find any accurate account of their affinities.

The Galla peoples all share a mutually intelligible language. Before colonial rule, they were suffering badly from Somali encroachment, and were saved by the colonial administration which demarcated and enforced a boundary between the two peoples. Once peace was established, the Galla were left much to themselves, and it is only very recently that development has begun to affect them significantly. One noticeable change has been the spread of Islam amongst them in recent decades.

1.2.3.2 The Somali Group

The Somali are often referred to as if they were one homogeneous people. This is far from the case, for Lewis (1955), lists and maps the distribution of over three hundred Somali "tribes" in Somalia, Ethiopia and Kenya, all linked together in a complex hierarchy of linguistic affinities. One consequence of the practice of seeing them as just "Somali" has been that it is difficult to find any reliable map accurately listing and demarcating Somali peoples in Kenya. Even Lewis's survey mentioned above suffers from a natural concentration on the Somalia groups at the expense of the peripheral populations. The Somali peoples listed in Map 4, therefore, only approximate to their real identity and distribution.

Although a majority of Somali in Kenya are still to be seen tending their flocks and herds, many have exploited the commercial potential of the wider Kenya economy. In the northern and eastern areas of the country, Somali traders and shopkeepers are a common sight. The richer and more enterprising traders send their lorries as far afield as Uganda and the Sudan. The Somali have also been prominent in recent years in the poaching business, and have successfully exterminated most of the valuable game on the eastern and northern plains.

The Somali peoples of Kenya represent one of the five points on the star on the flag of the Somali Republic. The Somali claim to this part of Kenya led to a protracted guerilla war in the 1960s and early 1970s, which occupied sizeable detachments of the Kenya army through most of the period. The Kenya army is still

deployed in the Somali country, and is certain to have brought major changes to the local Somali peoples. It is by no means clear, however, what the extent of this change is, or how far the Kenya Somali participated in the guerilla offensive.

Linguistically related to the Somali, but not part of the true Somali group are the Rendille, who occupy the very arid deserts and semi-desert plains east and south east of Lake Turkana (see Map 4). Their land is so barren that they do not keep any cattle, and rely for most of their subsistence on camels and goats. Culturally, they have much in common with the Samburu Maasai with whom they have enjoyed a close alliance within living memory. A substantial proportion of the surplus Rendille population, moreover, migrates and is absorbed by the Samburu, whose much lusher country and cattle economy can support a much greater population (Spencer 1973:138-142). The Rendille are amongst the poorest and least developed of the Kenya pastoralists, and even today development is occurring only at a very slow pace.

1.2.3.3 The Burji Group

This group consists of the Burji and Konzo, two small populations living on the higher slopes of Mt. Marsabit, where they subsist exclusively by cultivation. They are in fact, the only Cushitic peoples in Kenya that practice any form of agriculture. The Burji and Konzo are very recent migrants, having arrived in the first two decades of this century from the southern Ethiopian Highlands, where the Ethiopian Konzo and Burji are sizeable peoples. According to informants, the colonial administration permitted their settlement on Mt. Marsabit around 1912, but I have not been able to confirm this date.

1.2.3.4 The Mukogodo

The Mukogodo are Cushitic speaking hunters and gatherers, and will be discussed below.

1.2.4 Hunting, gathering and fishing groups

Scattered over much of Kenya are small populations who subsisted until recently by hunting and gathering or fishing. These peoples are usually called the "Ndorobo" by the Maasai, meaning "the poor ones", or those without cattle, and the "Sanye" by the Orma, meaning "blacksmiths", as they used to be the original iron smelters before the Bantu arrived. These hunting groups tend to adopt the language of the more powerful peoples that surround them. Thus the Dorobo of the Mau speak Kalenjin, those of Laikipia speak Maasai and those of the Tana River area speak Cushitic languages. Their linguistic instability is best demonstrated by the El Molo, a fishing group on the southern shores of Lake Turkana, who used to speak a Cushitic language fifty years ago, and today speak exclusively Samburu, in spite of the fact that there have been very few cases of inter-marriage between the El Molo and Samburu (Spencer 1973: 213-214). Spencer expresses the modern view that the present Dorobo are the "residue of the existing peoples of the area than the descendants of an aboriginal race" (Spencer 1973:218). This however, may not be true for some of the Dorobo, in particular some of the Sanye and Boni groups who appear to have some remnant Khoisan features.

1.2.4.1 The Kalenjin Groups

These groups are scattered throughout the margins of the Western Mau forest, and in the Tinderet and Eldama Ravine forests.

Their collective name is the "Ogiek". In the nineteenth century they used sophisticated hunting techniques, including traps, weighted spears and booby traps to kill forest antelope and elephant. Today however, they are increasingly settling down to cultivation.

1.2.4.2 The Laikipia and Samburu Dorobo

These groups used to be found in pockets all over the Nanyuki, Laikipia and Samburu Districts. Spencer (1973:201-203) lists 17 of these groups, who lived in the area during the nineteenth century. Of these, some have since been completely absorbed by their Bantu and Nilotic neighbours. Today there are two main concentrations, the Sueie and Laidotok groups who inhabit the Mathews and Ndoto Mountains respectively, and the Leuaso, Lngwesi, Ndigiri, Mukogodo and Mumonyot who occupy the Dol Dol "reserve" in the northern part of the Nanyuki District. The Dol Dol groups all speak Maasai and have taken to pastoralism in recent years. Only the Mukogodo still enthusiastically pursue their traditional gathering and bee keeping. The Mukogodo differ from the others in that they are presently bi-lingual in Maasai and their ancient Cushitic language which they spoke exclusively until recently. The Mukogodo language is yet to be studied in detail, but it is believed to be an archaic form of Eastern Cushitic, with some resemblance also to Southern Cushitic. The Ngwesi moreover, are relatively recent arrivals from the Meru area, whence they were removed by the colonial administration. Their stay in Meru had been relatively brief, as they had been moved there from Laikipia in the late nineteenth century.

In the north of Samburu District on the eastern shore of Lake Turkana, there used to be another group of Cushitic hunters and fishers. Of these, only the El Molo are left today. In the Rift Valley on the southern shores of Lake Baringo there is also a Maasai fishing population called the Njemps. They are believed to be on heterogeneous origins having originated from northern Maasi sections dispersed during the nineteenth century civil wars.

1.2.4.3 The Kikuyu Hunters and Gatherers

These are strong oral traditions in the Kikuyu areas and parts of Meru that speak of the existence of a people called the "Gumba", who were small and dwarf-like in appearance. These were gradually absorbed during the eighteenth and nineteenth century migrations of the Kikuyu into the Aberdare footslopes. There was also another group of hunters who were reputed to have inhabited the plains around the Athi River. The descriptions of the Athi are more in keeping with accounts of present Dorobo groups (Lambert 1949:Chapter IV). Today there are no hunting peoples left in Kikuyuland. They have all been assimilated. In his evidence to the 1932 Land Commission however, L.S.B. Leakey was still able to state that

> "up behind Kijabe (on the southern Aberdare forest near the Rift Valley escarpment) there are innumerable families, half true Dorobo and half Kikuyu ... They call themselves Kikuyu, but live like Dorobo...." (quoted by Lambert 1949:58).

1.2.4.4 The Sanye and Boni Groups

In the thick, dry bush country north and south of the Tana River, there are two related groups of Cushitic speaking hunters

called the Boni (north of the Tana) and the Sanye or Lliangulo south of the River. They are famous for their massive bows, with pulls of over 150 pounds, and their mastery of poison for arrows. Their bows however, are technologically far in advance of those of any other Kenyan people, and are an improvement of the sophisticated Galla design. (Parker, private communication). Their culture is centred on elephant hunting, and for many generations they have paid tribute in ivory to the Orma, in exchange for commodities of which tobacco is nowadays the most important.

There is some evidence that these groups may have been in their present areas for a very long time. Their bushcraft for instance, is far more detailed than other Kenyan hunting groups, and they have none of the strict food prohibitions which other groups have borrowed from their neighbours. Moreover some groups, the Dahalo in particular, have recently been found to have clicks in their language, a typically Khoisan trait (Nurse and Parker, private communication). This suggests that they might have been a Khoisan people before assuming Cushitic customs.

The Sanye and Boni are in an unenviable position today. Their lands are slowly being leased (without their consent) to ranchers, and they are not permitted to hunt. Nor is their culture compatible with the demands of the twentieth century. Like the American Indians in the late nineteenth century, it is likely that they will suffer as change comes.

## 1.3 Historical outline of the peopling of Kenya

### 1.3.1 Early pre-colonial history

The Middle Stone Age (50,000 B.C. to 15,000 B.C.) in East Africa saw Homo sapiens becoming increasingly specialised in the manufacture of tools. During this period, three stone industries can be distinguished. The Songoan and the later Lupemban industries are associated with wooded zones, and tools from major sites such as Nsongezi and Sango Bay (Uganda) and the Kinangop (Kenya) show specialisation in wood working and digging tools (Posnansky 1974:64). The third industry, the Stillbay, occurs on sites in more open savannah areas such as the Rift Valley (e.g. Gil Gil River) and is characterised by small leaf-shaped artefacts.

By 15,000 B.C. these industries had given way to the Wilton and Kenya Capsian industries, in which the development of new stone working techniques led to the creation of sophisticated composite tools, such as arrows with barbed heads, and the manufacture of microliths. The new technology permitted for the first time the exploitation of hard stones such as jasper and rock crystal. The Wilton industries, which are found over most of Africa south of the equator, are associated with skeletons of proto-Bushman or Boskopoid type. This suggests that Khoisan speaking hunters were well established in East Africa. The existence of remnant Khoisan speakers in northern Tanzania to this day, such as the Hadza and Sandawe, would seem to indicate that Bushman type hunters must have survived for some time in Kenya too.

The Kenya Capsian culture is associated with Caucasoid people. The excavation of Gamble's Cave near Lake Elementeita in Kenya, for

example, revealed skeletons distinctly similar to those of modern Cushites. The Capsian industries have been found as far north as North Africa, in areas inhabited by modern Afro-Asiatic speakers. The Caucasoid Capsian hunters reached a high level of sophistication in the pre-Neolithic period, especially in the manufacture of fishing harpoons and spears. The later Wilton cultures moreover, are associated with rock paintings, many examples of which survive in East, Central and South Africa. In East Africa the main concentration is near the present Khoisan speaking area around Lake Eyasi, where over a thousand sites are known.

The diffusion of pastoralism and agriculture into Kenya began between 1,500 and 1,000 B.C. (Posnansky 1974:68), and is linked with the later Capsian and Elementeita industries. From linguistic analysis, it now seems that these early pastoralists and agriculturalists were Southern Cushitic speakers. During the course of time they cooperated with and assimilated many of the surrounding hunting peoples, although pockets of hunters survived in the more remote areas. Some of the hunters adopted the Southern Cushitic language without changing their mode of livelihood (e.g. the Dahalo), and others adopted the lifestyle without changing their Khoisan language (e.g. the Sandawe) (Ehret 1974:155).

The end of the first millenium B.C. and the first few centuries A.D. saw the advance of two new peoples into Kenya. The first wave was the ancestors of the present Highland Nilotes, who migrated down from the Lake Turkana region, and gradually completely assimilated the Southern Cushitic peoples in the Kenya Highlands west of the Rift Valley, adopting many of their customs in the process. This can be deduced from the significant proportion of

Cushitic words in the present Highland Nilotic dialects, and the prevalence of customs such as circumcision and clitoridectomy, and an aversion to eating fish. Many of these customs however, were also acquired earlier through contact with proto-Eastern Cushitic peoples in the Ethiopian vicinity before the southward migration (Ehret 1974:154).

The second wave was the Bantu peoples, who first began to arrive in Tanzania about 2,000 years ago. They began to intermingle with and assimilate the Southern Cushites on the southern Tanzanian Rift Valley, gradually adopting many Cushitic customs in the process. The ancestors of the present Highland Bantu were well established in the northern Tanzania highlands by the middle of the first millenium A.D. As they expanded, the Bantu introduced iron to East Africa between about 300 and 500 A.D.

During this early period, East Africa was not completely isolated from the Roman Empire to the north. There exist two documents from the Imperial Roman period, the first an anonymous Greek commercial handbook called "Periplus of the Erythean Sea" written in 110 A.D. and the "Geography of Ptolemy" from about the fourth century A.D. These which show that the East African coast was part of an extensive network of trade between the Roman Empire, Persia and India. There was a number of "emporia" or trading ports down the Red Sea and the Somalia, Kenya and Tanzanian Coasts, which acted as transit ports for shipping from India. The southernmost was Rhapta, which was a thriving port exporting ivory and importing iron implements. The "Geography of Ptolemy" even mentions the existence of snow-capped mountains in the hinterland, implying some exploration of the interior (Mathew 1963). The

East African ports however, were not directly under the control of Rome, but under the suzerainity of "Charbael, ruler of Himyar in south-west Arabia" (Chittick 1974:104), with effective control residing in the people of Mouza from the Yemen. After the decline of the Roman Empire, the Arab connection was maintained, although the extent of Arab involvement varied considerably from one era to another.

About the end of the first millenium A.D. the ancestors of the present Maasai branches of the Plains Nilotes began to move into Kenya from the Sudan/Ethiopian cradle lands. The Maasai by this stage had already been strongly influenced by the Eastern Cushites, borrowing as the Highland Nilotes did before them many of their cultural traits. The Maasai expansion was much more rapid than that of the earlier peoples, with movement channelling down the Rift Valley well into Tanzania by the first three centuries A.D. The Maasai gradually split off into two sections, the agricultural Maasai represented by the Uasin Gishu in Kenya and the Arusha peoples in Tanzania, and the more widespread pastoral peoples on the Rift Valley floor and surrounding plains. During the course of this expansion they came into contact with both the Highland Nilotes and the Bantu. It is possible that the highly evolved military system based on age sets, and military paraphernalia such as oval shields and ostrich head dresses were borrowed from the Highland Nilotes. Their interaction with the Bantu peoples was considerable. As Ehret remarks, "former Maasai speakers probably constitute a significant element of the physical ancestry of the Kikuyu and Chagga" (Ehret 1974:162), these being the peoples closest in proximity to the Maasai areas.

While the Maasai were expanding southwards down the Rift Valley, the first Eastern Cushites migrated westwards from the northern Somali/Ethiopian regions, around the turn of the millenium. Little is known about these early invaders, but their descendants are represented by the modern Rendille people and remnant hunting groups of which only the Mukogodo now retain the language. The Rendille are closer linguistically to the Somali, suggesting an origin in the Somali areas of the Horn. These early Eastern Cushites however, were either assimilated or pushed back by the Galla peoples around 1600 A.D. The Galla occupation of the eastern plains drove a wedge between the Rendille and Somali peoples. By the nineteenth century however, Somali power had increased, and the Somali re-conquered much of the eastern portion of the country, a process that was not completed by the time of the establishment of colonial rule.

Meanwhile the Bantu had expanded into all their present eastern and general localities by the sixteenth century. By the tenth century, the ancestors of the Mji Kenda and Pokomo had probably already settled along the Kenya coast. At one time it was thought that they had reached as far as the Juba and Shebelle Rivers in Somalia, where apparently remnant negro populations are still to be found (Murdock 1959). There are also strong oral traditions that speak of an origin for the Coastal and Central Bantu in a mythical place called "Shungwaya", situated somewhere along the southern Somali Coast. Turton (1975) and Lewis (1966) however, have pointed out that there is no substantive evidence that the present Coastal Bantu ever reached north of Lamu, as the "remnant negroes" are in fact very Cushitic in character, and the Shungwaya Myth

has been reported by Morton (1972) as being of colonial origin. By the eighteenth century, the Coastal Bantu, especially the Pokomo had come into extensive contact with the southwards Galla expansion, resulting in the linguistic assimilation of the Korokoro and considerable impact on other northern Pokomo groups such as the Malakote.

The linguistic similarity between the Kenya Central Bantu and the Tanzanian north eastern Highland Bantu suggests that the original migration routes emanated from those areas, either via the Kenya Coast and across the dry plains following the Tana River, or from the south via the Taita and Kamba Hills. Very little is known about the early expansion into the Mt. Kenya Highlands in the early centuries of this millenium and possibly earlier. By 1800 however, the character of the present peoples was firmly established, with subsequent movement confined to the gradual expansion of the Kikuyu south of the Aberdares, and their assimilation of the Gumba and Athi peoples (Lambert 1949). One of the important aspects of the Bantu occupations of these regions was that there was much contact between the coast and the hinterland. The Kamba for example, had established a significant trading network with the coast by the early nineteenth century (Kimambo 1974:207).

In Western Kenya, the peopling of the Luyia and Lake Nyanza areas by Bantu was accomplished between the fifteenth and eighteenth centuries. The patterns of movement are extremely complex, and there is still much that is unclear. The peopling of the south-west by the Gusii/Kuria peoples is especially obscure. It is clear however, that migrations of Inter-lacustrine Bantu from the Uganda shores of

Lake Nyanza were a major source of this expansion, and that many of the present Luyia peoples originate from that region. The Kingdom of the Wanga in particular was established by the sixteenth century and was to dominate the region until colonial times.

The expansion of the Bantu into Western Kenya resulted in the assimilation of all remaining Southern Cushitic groups in the area, and the complete absorption of the whole Kitoki Highland Nilotic branch, a process complete by 1800. In the north the Teriki,Gisu and Bukusu Bantu peoples had been so affected by Highland Nilotic contact that they preserve many characteristics of that culture today. The Kalenjin Highland Nilotic peoples however, escaped large scale contact with the Bantu, except in the case of the Nyangori who still show many Bantu features today. The Bantu migrations into the Luyia area also effectively cut off the Mt. Elgon Sabaot branch of the Kalenjin from the main groups on the western Rift Valley escarpments. These remaining Kalenjin representatives of the Highland Nilotes effectively resisted any further Bantu encroachment into their areas (Ehret 1974).

The last major population migrations into Kenya were those of the Karimojong Plains Nilotes and the Lwoo branch of the River-Lake Nilotes. The Turkana entered the Rift Valley in about the eighteenth century, from the southern Sudan, and gradually displaced the Maasai Samburu who inhabited the whole valley floor west of Lake Turkana (Spencer 1973:Chapter 5). In the south, the Turkana came into contact with the Pokot Kalenjin speakers, and the Pokot acquired many Turkana customs, such as the wearing of mud head packs and the practice of removing the lower incisors, making a hole in the lower lip, and inserting lip plugs.

In the meantime, the main Karimojong cluster had expanded southwards in eastern Uganda, and in the latter decades of the century the Teso started to migrate around the west of Mt. Elgon into Kenya in the Malakisi area. This migration was halted by the arrival of the British.

The seventeenth and nineteenth centuries also saw an intense drive southwards by the Lwoo section of the River-Lake Nilotes, from central Uganda into Kenya around the Winam Gulf. The Kenya Luo migration was marked by many battles, and they drove the Bantu inhabiting the Lake shores and any remaining Highland Nilotic sections into the hills east of the basin. As Ogot points out however, the Luo migration consisted of the independent movement of many autonomous sections, who although were fierce fighters, nevertheless for the most part achieved their conquest through assimilation rather than extermination (Ogot 1967:153-154). By the late nineteenth century the Luo's natural southern progress was halted by the coming of the British.

### 1.3.2 History of the Coast and the immediate Pre-Colonial Period

From the eighth century onwards the history of the Coast is documented in Arab chronicles and geographies. The early accounts up to the tenth century however, are few and sketchy. The true scale of Arab involvement with the Coast in early times is not really known, but is not thought to have been great. In the eleventh century however, there occurred a migration of people from the Persian Gulf and South Arabia called the "Shirazi". They spread southwards from Mogadishu, and Muslim Sultanates and settlements arose at points down the Coast as far south as the Comoro Islands

(Chittick 1974). The greatest Sultanate was that of Kilwa on the south Tarzanian coast, which held hegemony as far north as Zanzibar. By the fifteenth century however, its power was on the decline, and the power of settlements in Zanzibar, Mombasa, Malindi, Gedi and Pate increased considerably.

The impact of the Islamic penetration on the indigeneous negro population is of considerable interest. The coastal negroes or Zenj as the Arabs called them, though not mentioned in the Roman accounts, werely clearly established on the coast by the late Roman period. As Islamic power increased, one might have expected considerable assimilation of these peoples by the Arab invaders. Chittich however, thinks that the Arab impact was not as significant as one might think:

> "The impact of this civilisation on much of the mainland coast was slight, and inland non-existant. It is unlikely that any Moslems went into the interior, save on an occasional war-like raid. Their religion never penetrated beyond the shore of the mainland, nor did their impressive skills in building have any influence in the hinterland. Buildings in stone and the burning of lime for mortar were unknown even 5 miles from the coast."
> (Chittich 1974:113).

The sixteenth century witnessed the arrival of Europeans to East Africa for the first time since the Roman era. In a period from 1505 to 1512 the Portuguese conquered Kilwa and established a garrison there, and the occupation of the southern territorial possessions of Kilwa was to form the nucleus of Portugal's East African Empire. The Portuguese expansion into the north however, met with persistent resistance. On the Kenya Coast, Mombasa, which was by then the dominant power, proved

especially stubborn, and was sacked three times (1505, 1528 and 1589) before finally relinquishing its independence (Berg 1974:120).

The Portuguese presence in East Africa was motivated by strategic considerations, designed to expand and preserve Portugal's European maritime trading monopoly in the Indian Ocean. As long as strategic ports such as Mombasa could be kept firmly under control through the presence of a garrison and the exacting of tribute, there was no need for further involvement. Portuguese rule in East Africa therefore, had only a minimal impact on the traditional lifestyles of the peoples in its sphere of influence.

By the eighteenth century, Portugal's monopoly and power in the Indian Ocean was undermined by English and Dutch commercial incursions. As Portuguese power declined, the garrisons in East Africa found it increasingly difficult to suppress frequent uprisings. In the 1720s they were finally ousted from Mombasa and the other Kenyan ports and settlements, with the help of the Imam of Oman, who had designs on the area himself.

The replacement of Portuguese authority with Omani suzerainty never materialised, owing to the invasion of Oman by Persia in the middle decades of the eighteenth century, which saw the downfall of the ruling Omani Yarubi dynasty (Berg 1974:125). Nevertheless the involvement of Omani individuals (migrant families and traders), led to a

> "process of re-Arabisation ... in which Arab kinship, values and some elements of material culture gained prestige at the expense of the Swahili culture. In the long run Swahili society was considerably modified by the process." (Berg 1974:126).

Around 1750 however, an Omani family, the Mazrui, effectively established a ruling dynasty in Mombasa and made Mombasa the most powerful town on the coast. By the 1820s, their power declined as direct Omani authority was re-asserted, and the Muscat Bu-Saidi family established itself on Zanzibar. By the 1830s Zanzibar was the central power of the coast and Mombasa became subject to its rule. With the backing of Indian finance, Zanzibar became very prosperous, the chief port for the east African slave and ivory trade.

The early nineteenth century was a period of considerable penetration of the East African interior by Arab traders, who established a virtual trading monopoly in ivory and slaves. At first, trading was done with intermediaries in the interior (for example the Kamba in 1825 organised a caravan to the coast to trade). It was not long however, before the Arabs and Swahilis organised their own caravans, which "traded" as far as Zaire.

The slave trade in East Africa in the early decades of the nineteenth century was the main factor that stimulated British interest in the area. Following the abolition of slavery from British possessions in 1833 and its elimination in the West Indies and West Africa, humanitarian forces in Britain caused the British Government to focus on the East African slave trade.

The use of force was considered logistically impractical and too costly. Instead the ideas was born to "regenerate" Africa, to introduce trade and commerce that would, through overwhelming competition, stifle and gradually eliminate the slave trade as an economically viable proposition. In the light of its commitment

to abolish slavery, the policy of the British Government in East Africa was to create the conditions necessary for missionary and commercial expansion. This policy was implemented by strengthening the Sultanate of Zanzibar and applying pressure, rather than by direct colonisation. In 1840, a British Consul was appointed to Zanzibar and he pressurised the Sultan to reduce the slave trade and accept compensation. In 1875 military aid was "accepted" by the Sultan and an army equipped with modern weapons was established on the mainland under the generalship of a British Naval Officer, with the purpose of reinforcing and extending the Sultan's authority and to help British commercial interests.

The imposition of indirect control on Zanzibar provided a convenient base from which expeditions to the interior could be mounted. By the 1880s the days of the great explorers were over. Livingstone, Speke, Burton, Grant, Thomson, Von Honell and Stanely had all started from Zanzibar, and it was due to their efforts that East Africa was put on the European map. Of major interest was the discovery of the large and sophisticated interlacustrine kingdoms around Lake Victoria, which were seen as being of great commercial potential and provided a stimulus for expansion of European settlement and enterprise.

In the wake of the great explorers came the missionaries. The writings of Livingstone in the 1840s did much to draw attention to the barbaric attrocities of the Arab slave trade. The pioneering efforts of the Church Missionary Society in the 1840s laid the foundations for efforts to penetrate inland in the 1860s and 70s. The Society by the mid 1870s had founded missions in Uganda, the primary focus of interest, with Kenya viewed primarily as a route

to Uganda. In 1878 the French Catholic White Fathers had also left for Uganda, the first signs that Britain was no longer the only party interested in Africa.

With its vital interest in India the British policy was to safeguard against any strategic threat to this interest. The discovery of the source of the Nile and the building of the Suez Canal meant it was vital to strengthen the British position in Egypt and the Red Sea. In particular, control of the source of the Nile was considered of considerable strategic importance. The annexation of the Cameroons and South West Africa by Germany in the 1880s started the "scramble for Africa". In East Africa the fanatical imperialist zeal of Karl Peters led to the declaration of the German Protectorate of Tanganyika, and parts of the Kenya Coast around Witu were also claimed. The Italian drive to annex Ethiopia in the north had also started. The subsequent British backlash led to the treaties of 1888 and 1890 which partitioned East Africa into three spheres of influence, with Britain enjoying the prize of Uganda and the Nile, and the intervening interior - most of what is now Kenya.

To exploit its sphere, Britain created the Imperial East Africa Company, which received its charter in 1888. It was granted wide powers to govern under concessions or treaties obtained from local rulers. As a delegate of British power it was initially successful, with treaties negotiated with sections of the Taveta, Taita, Pokomo, Galla, Kamba, Kikuyu, and Wanga to open up a safe route for caravans to Uganda. Stations were also built along the caravan route, and kept well stocked with provisions. On the other hand, as a commercial venture the Imperial East Africa Company was

an abysmal failure, finding little in the interior to offset its considerable expenses. In 1894 the charter was revoked, and the Ugandan Protectorate was declared, followed soon afterwards by the East African Protectorate in 1895, which encompassed all of modern Kenya east of the Rift Valley.

The new administration faced two strategic problems. Firstly it had to set up an efficient administration and persuade the peoples of the hinterland of the benefits of Pax Britannica. Secondly it had to realise some of the economic potential of the Protectorate to help pay its way. As a matter of urgency, the Uganda railway was built to alleviate chronic transport difficulties and to link the coast with Uganda, the area thought to be of highest economic potential. The railway was started in 1895, and its progress was made possible by successful pacification of the Maasai, who were advised by their leaders not to resist. As the railway progressed towns sprang up at strategic points, including the future capital of Kenya, Nairobi.

In order to provide labour for the railway building, large numbers of Indians were recruited. By the time the railway was completed in 1900, many of these Asians had successfully started shops and busineses both in the coast and at the new towns. The Asians were the first people to exploit the commercial potential of the new Protectorate.

The early 1900s saw the final pacification of all the peoples near the railway except the Nandi, who resisted stubbornly for 12 years. In 1901 and 1902 successful punitive expeditions were mounted against the northern Kikuyu, and in the next six years

the Embu, Tharaka, Kikuyu and Gusii were brought under control. The way was paved for the opening up of the large tracts of seemingly uninhabited country in the Rift Valley and the Nairobi area to settlement.

The early administrators however, were not as a rule enthusiastic about the prospects for European settlement. As Sorrenson notes

> "Ainsworth in 1896 pointed out that there was not much room for European colonisation, certainly not in Kikuyu, which is a small country, thickly populated... Lugard too, was struck by the density of Kikuyu cultivations when he arrived in Dagoretti in 1890. Although he subsequently passed through unoccupied land in the highlands beyond the Kikuyu escarpment, it did not occur to him when writing up his diary that this might be settled by Europeans." (Sorrenson 1968:32).

Indeed in 1901 and 1902 very serious consideration was being given to opening up areas firstly for Indian settlement and secondly for the repatriation of Eastern European Jews. It was only the efforts of a small but vociferous and influential number of pioneer British settlers led by Lord Delamere, and championed by Eliot, the commisioner, that finally persuaded the Foreign Office to consider European settlement more favourably.

The alienation of land for European settlement began in earnest in 1903. By 1915, 3,168 holdings comprising a total of 5,275,000 acres were set aside for European use (Sorrenson 1968: Appendix 2). The bulk of this land was for pastoral use, and came from sparsely inhabited Maasai lands in the Rift Valley north of Nairobi. In 1913 the whole of the Rumuruti Plateau was also claimed, and the Laikipiak Maasai were "persuaded" to move ignominiously over five hundred miles to a new area in the south west of the country

(Trans Mara). The agriculturalists, in particular the Kamba, Kikuyu and Gusii, lost comparatively little, but in the case of the crowded Kikuyu, the loss of their land was deeply resented.

The growth of the settler community had a profound influence on the direction of government policy, in particular on the allocation of resources for economic development. In other British Colonies and Protectorates with no settler communities, the growth of the African rural small-scale cash economy was actively encouraged. In the E.A.P. however, it was the large scale European farming economy that received the bulk of the aid, especially in the form of government marketing and buying of agricultural produce and cattle, which catered almost exclusively for the needs of the European farms.

Meanwhile the indigenous peoples were allocated to reserves with fixed boundaries, and gradually came under direct local administration in the form of Provincial and District Commissioners and District Officers. In 1902, Hut Tax was introduced, provoking for the first time a need for cash. To help enforce this law in the face of a traditional absence of any indigenous form of centralised authority, Government Chiefs were appointed.

By 1914, the character of the future colony was firmly established. Any moves towards an internally self governed settler colony on the Rhodesia model however, were suspended on the outbreak of the first world war. East Africa proved to be an active theatre of war. By 1917 the small German army under Von Lettow had proved so difficult to pin down that the number of

British and colonial troops reached 30,000, a staggering figure for a colonial army in Africa. This however, was not enough. The ravages of disease were depleting reinforcements so rapidly that the British commander, General Smuts, who had thought that native East Africans would be unsuitable troops, changed his mind. Within weeks, the Kings African Rifles were expanded from a few thousand to 35,000 (Miller 1974:260). Even more dramatic was the decision to supplement the costly methods of truck and horse transport with African porters. The East African Carrier Crop was founded, and the number of porters raised through "volunteering" from about 7,000 to 135,000. The bulk of these recruits came from the tribes surrounding the Uganda Railway - they could not have numbered more than 1½ to 2 million in all at that time. For the first time, a very significant proportion of the adult male generation of these peoples was thrown into close contact with hundreds of white men, and with peoples from other tribes and countries. People who had often not walked further than a few miles from their homes in the course of their lives were suddenly aware of the diversity of the outside world. Above all, the myth of the superhuman qualities of the white man was dispelled. These factors were to be important in changing the awareness and aspirations of Africans after the war.

In 1920, after a period of post-war recovery, the East African Protectorate was upgraded to colonial status and was renamed Kenya Colony and Protectorate. A substantial area west of the Rift Valley, Uganda's eastern province, was added to the colony, although some regions such as Karasuk remained under Ugandan administration

until the 1950s. The Coastal strip however, remained a Protectorate, as the land was nominally still under the suzerainity of the Sultan of Zanzibar, and was leased from him by the Kenyan Administration.

### 1.3.3 Colonial and post-Colonial History

The period of economic recovery after the first world war was marked by the expansion of the road and railway network within the Colony to service the White settlement areas, and the further development of agricultural marketing. The most serious handicap to the development of the White farming areas however, was a chronic shortage of labour. The survivors of the East African Carrier Corps when they returned home were poor advertisements for the benefits of service with the white man due to their unhappy experiences, and there was additionally little incentive for men who could make an adequate subsistence living in the manner to which they were accustomed, to embark on a new career.

The attitude of the Administration to the Africans under their charge was one that saw little ultimate advantage in leaving the "natives" to their own devices. They had to be persuaded eventually to become westernised citizens of the new Colony. In the short term however, the Government pursued a dual policy. Those peoples who lived far from the areas of White settlement around the railway were left very much to their own devices, provided that they broadly observed the peace. Those peoples in the highlands close to the railway, however, were earmarked for participation in the wider economy, and this meant participation in supplying labour. In 1921 the hut and poll taxes were sharply raised to encourage the need for cash (Middleton 1965:356), and this resulted in a

gradual increase in the supply of labour, and succeeded in extending cash awareness to the rural areas. A lot of the labour however, was recruited using more dubious means, including the bribing of chiefs to employ coercion in their areas. The Government, while not approving of these measures, did not succeed in stamping them out until the end of the decade.

By the 1930s, the labour crisis was eased by the rapid rise in population amongst the peoples in the prime agricultural areas around Mt. Kenya and the Luyia and Luo countries. As land became increasingly shorter in supply, and traditional expansion was blocked by strict policing of boundaries, the surplus population had little choice but either to move to the towns and cities or to become labourers on white farms. By this time too, some aspects of European consumerism had taken deep root in the peoples bordering the railway. Bicycles, clothes, textiles and even gramophone records became increasingly sought after, stimulating a need for cash that could only be satisfied by fuller participation in the settler economy. As land became scarcer, even traditional aspects of life such as the paying of bridewealth and tribal fines became more cash-orientated.

This enforced participation in the cash economy was accompanied by an expansion of education mainly through the establishment of mission schools. By the 1920s, the missionary educated boys of the pre-war period were already mobilising political activity. In 1920 the Kikuyu Association was formed, followed by the Young Kikuyu Association in 1921 under the leadership of Harry Thuku. Following his organisation of active protest against the increase in the poll tax and the establishment of registration for

adult males (the kipande system), he was arrested in 1922 and became the first Kenyan political martyr. The Rev. J.W. Arthur wrote in 1922:

> "The development of the native peoples in Kenya in one short year is simply past thinking. They are almost now able to safeguard themselves against oppression and exploitation. What is much more to be feared is the native uprisings led by young educated Kikuyu .... Thuku has now got hold of the majority of young Kikuyu."
> (quoted in Middleton 1965:358).

In 1925 the more mature Kikuyu Central Association was formed, under the leadership of Joseph Kangethe and Jesse Kariuki, who were to become prominent politicians, and was joined by a proliferation of political associations in Kikuyuland. During the 1920s and early 30s these Associations became the voice of Kikuyu political protest against particular Government measures.

The Kikuyu example was followed quickly by other peoples bordering the railway. In the Nyanza area, the Kavirondo Tax Payer's Welfare Association was instigated in 1923, under the direction of an Anglican Archdeacon Owen. This split up into Luo and Luyia sections, and in 1934 the first political Luyia organisation, the North Kavirondo Central Association was formed modelled on the Kikuyu Central Organisation. This pattern of growing political awareness was followed by the Kamba, Meru, Mji Kenda and Kalenjin (Middleton 1965). In all these tribal areas, previously autonomous peoples were forming ethnic alliances for political ends.

The Nyanza (Luo and Luyia) Associations however, had functions other than political protest. Their branches in the urban centres provided welfare, local community leadership and a focus of identity for incoming rural migrants. Their importance for the

social organisation of Luo and Luyia communities in big cities such as Kampala and Nairobi is described in considerable detail by Parkin (1969, 1978).

By the end of the 1930s, a significant proportion of the Highland and western Bantu peoples, and the Luo, had been integrated into the settler rural and urban economies, either as labourers, or small scale entrepreneurs and workers (in light industries, ports, roads and railways). By 1948 half of Kenya's African labour force was to become non-agricultural (Bennett and Smith 1976:111).

There was moreover, an emergence of what Middleton terms "individualization":

> "The concept of "individualization" implies the appearance of the individual as the unit of production and residence. A single man, often helped by his wife and children, came to be the producer of cash crops, or the labourer or squatter in urban and rural areas alike. Women also moved to the towns to work as individuals. The traditional lineage and neighbourhood work-groups, although still functioning in many areas, were losing importance. Ties and obligations of kinship and neighbourhood were irksome to the men who could earn money and aspired to a new way of life and standard of living. Individual ownership, especially in money and in some places of land, was developing, with consequent changes in the respective roles of senior and junior men in the traditional system. Distinctions of wealth and power appeared within tribal units where traditionally there had been few or none. There was a marked increase in social mobility, both in spatial terms, with widespread labour migration to the towns and European farming areas, and also in class terms, with the growth of incipient classes and class distinction."
>
> (Middleton 1965:385-386).

By the second world war therefore, Kenya was developing most of the characteristics of a complex society.

The second world war, like the first, saw the recruitment of a large number of men into the Kings African Rifles. Altogether there were 75,000 men from Kenya, which Middleton calculates was

some 20 per cent of the adult male population (Middleton 1965:386). The Kenyan soldierswere not confined to the African theatre, but saw service in the Middle East and the Far East. The major consequence of the conscription of such a large percentage of the male population was not so much the broadening of horizons through seeing strange lands and being in contact with foreign peoples, though this was an important factor, but the considerable education that the soldiers received in the army. Not only were they taught English and basic literacy, but many received technical training in a variety of fields.

After the war, this army of semi-trained and trained men, with the incentive of a great deal of army back pay which was saved for them, was to have a considerable impact on the more progressive areas where the temptation to lapse back into a traditional lifestyle was less strong. They were able to take full advantages of the lifting of restrictions on retail trade and growing of cash crops previously imposed on Africans. As Middleton remarks:

> "For the first time Africans entered into retail trade in large numbers in direct and open competition with Indian traders; they also entered the market as buyers of local produce and in both Nyanza and Kikuyu they captured a large share of the trade."
>
> (Middleton 1965:387).

The post war period saw growing political unrest, with the successor of the Kikuyu Central Association, the Kenya African Union at the forefront, one of whose leaders was Jomo Kenyatta. The leaders of the KAU and other prominent African politicians began to press hard for greater and elected representation on the Legislative Council.

These moves coincided with a growing resentment in the rural areas of Government economic reform, involving conservation

measures such as cattle culling to prevent erosion, and improved agricultural techniques such as contour ridging. This reform was seen as an essential prerequisite of future political progress by the Government, (Bennett and Smith 1976:119), but was regarded as aribitrary and irrelevant by Africans in the light of the great differential in land allocation between Europeans and Africans.

Growing grievances in Kikuyuland, both political and economic, led to the Mau Mau rebellion in 1950. This traumatic experience for the Kikuyu peoples led to major changes, such as land consolidation, the concentration of the population into policed and controlled villages, and eventually, to the individual ownership of land.

In the short term, the Mau Mau rebellion hampered rather than aided the move to independence. Political parties were banned under the state of emergency, and money that could have been poured into rural development was used for the war. In the long term, however, it probably contributed significantly to the growing British view that early independence for the African possessions was a necessity.

Any signs of Kikuyu nationalism were held firmly in check by the Government in the middle 1950s. This gave the other large ethnic groups, particularly the Luo, who had not participated in the rebellion, to put forward their own political claims. In 1957 the elections to the Legislative Council, the first in which Africans voted, saw all the elected African posts going to non Kikuyus.

In the late fifties and early sixties the Kikuyu were

permitted back into the political arena. With independence in the air two rival parties were established, the Kenya African National Union which was supported by the Central Bantu peoples and was under Kikuyu leadership, and the Kenya African Democratic Union which was a mainly western Kenya Luo, Luyia and Kalenjin alliance under Luo leadership. The future of Kenya was negotiated by these two parties with the British at the Lancaster House Conferences in 1960 and 1962. Although KADU, and in particular the numerically smaller Kalenjin section, won important concessions, it was strongly outvoted in the 1961 and 1963 elections.

In June 1963 the leader of KANU, Jomo Kenyatta, became the first prime minister of Kenya and led the country to independence on the 12th December 1963. The following year he declared a Republic, and Jomo Kenyatta became the first president. He also rapidly banned all opposition parties and made Kenya a one party state.

Since independence there have been many far reaching changes. The rural areas have developed out of all recognition, as the resources of the country have been concentrated on this previously neglected sector. Today cash crops and individual land ownership have brought the opportunity of wealth to all the agricultural peoples, and the pastoral peoples are slowly being encouraged to settle down. With the population almost doubled since colonial times, more and more surplus population has had to leave the rural sector. All urban populations have expanded significantly during this period, as have the industries to provide employment. In the rural areas, moreover, there has grown a large

network of traders, shop keepers, bar and hotel employers and staff, entrepreneurs and taxi (Matatatu) drivers).

On the uglier side is the inevitable increase in unemployed landless peoples as land shortage increases. The many settlement schemes on previously white owned agricultural and pastoral land have only been partially successful in alleviating the problem, and at best are only a short term remedy. The ultimate hope is the expansion of industry, as unused good agricultural land is rapidly disappearing.

Finally, the greatest success has been to bring education at least on the primary level to all the agricultural rural areas, and to parts of the pastoral ones. In 1976 primary school fees were abolished, making this level accessible to the children of all parents who can afford the school uniform. The full impact of this great increase in literacy however, will only be known in the future. As Kenya enters its second post independence era under the presidency of Daniel Arap Moi, it is fast becoming a highly complex nation.

## 1.4 Major demographic trends since 1948

The first population census of Kenya was carried out in 1948. Since then there have been three further censuses, in 1962 just before independence, in 1969 and in 1979. The 1979 Census however, has not been analysed yet, and is not relevant to this study, as it post-dates my last fieldwork expedition in 1975 by four years. The results of the 1948, 1962 and particularly of the 1969 Censuses, provide detailed information on many important aspects of demography in the country, including age, sex, tribe or nationality, birthplace, marital status, urbanisation and mortality and fertility. The reliability of the Census varies considerably from one area to another - the most remote eastern pastoral areas being the most unsatisfactory. Self reporting of age for instance, becomes increasingly suspect in these areas.

### 1.4.1 Present distribution of population: (Republic of Kenya 1969 Population Census)

As Table 1.2 shows, the population of Kenya in 1969 was just under 11 million, of which 97 per cent were Kenyan Africans. The most numerous tribal populations are the Kikuyu, (2¼ million), the Luo (1½ million), the Luyia (just under 1½ million) and the Kamba (1 million). The Kikuyu alone account for 20 per cent of the total population, and these four peoples in total comprise 58 per cent of the total. Other Bantu populations (Meru, Kisii, and Mji Kenda) approach the half million mark, and these are joined by the Kalenjin speaking Kipsigis. The remaining populations usually number less than 200,000. Particularly small are the Taveta, Swahili, Boni/Sanye, Njemps, Sakuye, and Gosha who number under 10,000. In

fact of the 42 Kenyan African populations specified in the census, 34 populations in total comprise only 28 per cent of the population and of these 23 contributed individually less than one per cent. Populations size therefore, is very unevenly distributed. The largest populations are Bantu, and Nilotic Luo, followed a long way behind by the Maasai and Kalenjin Nilotes and the Somali Cushites, with the smallest populations being the Galla Cushites and the hunting Sanye and Dorobo groups.

This distribution of population size is closely related to two factors. The first is the scale of definition for each tribal population in the Census. For instance the Luyia are strictly speaking not truly comparable to the Nandi, Kipsigis and Tugen, as the Luyia are not differentiated by minimal unit, while the Kalenjin are. The collective Kalenjin actually number over a million. If minimal units were differentiated then only the Kamba would figure as disproportionately numerous. The second factor is mode of livelihood, and the percentage of arable land occupied by each people. The most numerous populations are agriculturalists who occupy the most fertile lands, or who have large proportions of urban dwellers.

Populations size is also highly related to population density. The areas with the highest population densities are found in two zones. The first is in western Kenya in the high rainfall and fertile agricultural zones bordering the Lake Victoria basin. Population density rises from just over 100 persons per square kilometre around the Lake shore reaching very high densities in the Kisii, Kakamega and South Elgon areas, ranging from about 300

to over 500 a square kilometre. The second high density zone is the footslopes of the Aberdare and Mt. Kenya Highlands, where the average density is nearly 200 a square kilometre. In the Kikuyu Meru and Kamba districts, up to 500 people a square kilometre are found in the most favourable localities, and in the Kiambu Kikuyu District the maximum density is nearly 800. The areas of highest density therefore, coincide with the homelands of the Kikuyu, Luyia, Luo, Gusii and Meru, the most numerous ethnic populations.

The coastal and high forest and grassland areas support medium population densities, ranging from about 50 to 200 people a square kilometre outside the forest reseves. These densities reflect the fact that the land they occupy, though fertile and well watered, is either underdeveloped or not totally suitable for intensive agriculture. For example, the high forest and grassland areas suffer from erosion and occasional ground frosts. These zones of medium density are mostly occupied by the Coastal Bantu and the Kalenjin Nilotes.

The areas of lowest density are the arid eastern and northern plains, where the average district densities if Isiolo, Marsabit, Garissa, Mandera, Wajir, Tana River, Samburu and Turkana are well under 5 to the square kilometre. These are the districts inhabited by the Cushitic peoples and the Turkana and Samburu Plains Nilotes. The Maasai pastoralists from Narok and Kadjiado Districts also manifest a very low density of under 10 a square mile. The overall pattern of population density therefore, is highly correlated with rainfall reliability and intensity.

1.4.2 Population Growth since 1948

In the first decades of this century, the population of Kenya was roughly estimated to be around 2½ to 3 million (Middleton 1968:337). By the time of the first population Census in 1948, it had nearly doubled to 5¼ million. Since then, there has been a dramatic rise to 8½ million in 1962, and to 11 million in 1969. Unofficial estimates of the 1979 Census speak of a current figure of around 15 million. The population therefore, had doubled between 1948 and 1969, and had increased threefold between the early decades of the Century and 1969, and probably five-fold up to 1980. The most dramatic symptom of this increase is the very large proportion of young people in the population, with 76 per cent of the population under 25 years of age, and 19 per cent under 5 years (Kenya 1969 Population Census Vol. IV: Table 1.2).

Nobody really knows why the population has increased so dramatically. The increased availability of medicines, improved standards of living and hygiene, and the introduction of high yield hybrid maize have undoubtedly had a considerable impact. Perhaps more important however, has been the elimination of the locust plagues that were so prominent in the 1930s, the success of anti-pest measures for crops and general dipping of cattle, and above all twenty years in the sixties and seventies of adequate rainfall and the absence of famine in most areas. Linked with these trends has been the expansion of urban opportunities, and the opening up of large areas of previously white owned agricultural land, to absorb the surplus population. The favourable conditions of the post-independence years are reflected in a decrease in the infant mortality

rate of 13 per cent between 1962 and 1969. (Kenya Population Census 1969: Vol. IV: 43).

The population increase or decrease of each tribal population is detailed in Table 1.2. Two populations, the Bajun and Sakuye have increased by over 100 per cent since 1962. These figures however, probably reflect a redefinition of tribal identity since 1962. The Bajun are a bit of a mystery in this regard; possibly many were put down as Swahili or Pokomo in 1962. The Sakuye however, may have increased almost entirely because a large number called themselves Boran in 1962, and have reverted to a more specific identity in 1969.

Excluding these anomalies, the largest increase of 40 - 60 per cent have occurred in the Kuria, Nandi, Njemps, Dorobo, Orma and Gurreh. The Nandi increase coincides with a major expansion into the neighbouring previously white owned Uasin Gishu District. The increase of the Dorobo and Njemps is probably due to rapid transformation from a predominantly hunting, gathering and fishing mode of livelihood to agriculture and pastoralism. The Orma and Gurreh increases almost certainly reflect a significant amount of "tribal redefinition".

Most of the major Bantu populations and the Luo have increased in the 30 per cent range. In view of the original sizes of these populations it means that they have shown by far the greatest numerical increase. The Kikuyu have even improved their overall percentage composition of the total Kenya population by 1 per cent.

The peoples in the pastoral areas have on the whole either registered a negligable increase (for example the Maasai increased by

only 0.5 per cent), or sharp decreases. Particularly dramatic decreases are those for the Boran (-41.6%), the Gosha (-59.7%). The Sanye/Boni and many of the Somali groups also show a marked reduction in population size. Although the Cushitic peoples do show a higher rate of infant mortality and a lower rate of fertility than other Kenyan populations (Kenya Population Census Vol.IV), this cannot possibly account for such large decreases. The most likely explanation is that the exhibited decreases in population size are a reflection of the unreliability of the 1969 Census in the Cushitic areas; involving some element of tribal re-definition and above all a serious underestimation of numbers. It cannot be coincidental that the populations most affected by sharp decrease in size lie in the areas involved in the war against Somalia in the mid 1960s.

1.4.3 Patterns of Movement

Since the imposition of colonial rule, there has been an increasing migration of people from their traditional homelands to seek employment and new opportunities in areas of large scale farming and in the cities. Since independence migration has accelerated for several major reasons. Firstly colonial laws restricting and monitoring movement no longer apply. Secondly there is the shortage of land caused by the large population increase in the highlands which has forced large numbers of people to seek alternative opportunities. Thirdly there has been the development of new opportunities. Prominant among these have been the opening up of large areas previously allocated to large scale farming to re-settlement by African families; the widespread development of many sectors of the Kenyan economy both rural and urban; and finally the expansion

of Government posts. The 1969 Census provides detailed figures from which migratory patterns can be estimated, although not as precisely or reliably as one would wish.

One method of estimating the extent of population movement is by examining the proportion of people from each population have birth places that do not coincide with their present districts of residence. In Table 1.3 I have reproduced the figures given in the 1969 Census analysis (Kenya Population Census Vol. IV: Table 6.8). The results reveal a very static picture, with 15 of the 42 populations listed having over 90 per cent of their population born in their district of residence, and only 11 with over 20 per cent outside their district of birth.

The largest proportion of "migrants" occur in the Kikuyu, Tharaka, Pokomo, Taveta, Taita, Bajun, Keyo, Turkana, Dorobo, Rendille and Boran. The Tharaka, Turkana and Rendille are particularly striking in that over 40 per cent were found to live outside their district of birth. The Rendille are also remarkable in that over 40 per cent appear to be living outside their province of birth.

These figures however, are curious to say the least. In Vol.I of the 1969 Census for instance, the number of Rendille living in Marsabit District, the traditional district of residence, is put at 17,931 which is 95.7 per cent of the total number of Rendille in Kenya. in marked contrast to the above figures from Vol.IV. Similarly it is difficult to believe that 30.1 per cent of Taita were not born in their District of residence when Vol.I lists 80.1 per cent living in their traditional home district in

1969. Finally the figure of 41.9 per cent of Turkana having been born outside their district of residence is equally curious seeing that 86 per cent of Turkana still live in Turkana District. Even allowing for movement back from non-tribal areas and considerable mobility for those outside their traditional areas of origin, there is still a considerable discrepancy.

Furthermore the comparison of birthplace with residence cannot distinguish clearly between migration from traditional home areas to the cities and developed large scale farming districts, and movement within non-traditional areas of residence. In view of this I have constructed a different set of figures from the 1969 Census Vol.I, which compare the proportion of each people still living in their traditional home districts and in their home province. As in the period just preceding colonialisation all but a very low percentage of each people lived in their traditional home areas, the percentage now living outside them can be used as a rough measure of the extent of migration into the cities and areas alienated in early colonial times.

Of the 36 populations listed in Table 1.4 (not 42 as in the Census as I have amalgamated the Somali populations), 29 (80.5 per cent) have over 90 per cent of their populations living in their "home" provinces and only one population, the Swahili, fall under 80 per cent. Moreover 21 populations (58 per cent) have 90 per cent or more of their population still living inside their traditional home district or districts. The notable "migrants" are the Kikuyu, Swahili, Bajun, Nandi, and Keyo over 20 per cent of which live outside their traditional home districts. Of this list

Durham University
SCIENCE
- 2 OCT 1991
SECTION
Library

however, only the Kikuyu and Swahili can be termed large scale migrants, as the majority of the movement for the Bajun, Nandi and Keyo is within the home province, (see Table 1.4). If we take movement outside the home province as the greatest indicator of movement, we get ten populations only who have more than ten per cent of their population living outside their traditional home province. These are the Kikuyu, Kamba, Luyia, Taveta, Swahili, Luo, Iteso and Boran. Numerically the greatest contributors to migration have been the Kikuyu, Kamba, Luo and Luyia, the largest of the Kenyan peoples, and the peoples most affected by land shortage in their densely inhabited home areas.

In Table 1.5 I have broken down the people living outside their home province for these ten peoples according to where they are now living. It shows that the Kikuyu although to be found in all of Kenya's provinces, have only migrated in very large numbers to Rift Valley Province (84.7 per cent of 403,095). Of the 341,480 Kikuyu who have migrated to Rift Valley Province 49 per cent have moved to Nakuru District, 17 per cent to the Uasin Gishu/Trans Nzoia Districts, and 11 per cent to Laikipia District and 8 per sent to Kericho District. Eighty-five per cent of Kikuyu in Rift Valley Province therefore, have moved to non-traditional Districts occupied in colonial times by large scale European farms, and where in the case of Nakuru and the Uasin Gishu Districts, considerable light industry has been developed since independence.

The importance of these areas can be seen from the fact that Rift Valley Province contains a significant percentage of migrants from all other peoples listed in Table 1.5 except for the

Kamba. A significant proportion of migrants for all peoples except the Iteso moreover, is accounted for by Nairobi, especially in the case of the Kamba and Luo. The Kamba instead of having moved significantly to Rift Valley Province, have tended to move to Coast Province, especially to Mombasa. A surprising number of Luo (16 per cent) have also moved to the coast, particularly Mombasa. The Kikuyu however, are noticeably under-represented at the coast as are the Luyia.

The important aspect of these results is the low proportion of migrants to traditional home areas of other peoples. For example North-Eastern, Nyanza and Western Provinces were mostly unaffected by European land alienation and are the traditional home areas of the Somali, Luo and Luyia peoples respectively. The proportion of migrants of each people for these areas in Table 1.5 is minimal. Only the 10 per cent of Luo migrants in Western Province contradict the trend, which is not surprising as the many of these apparent migrants are in reality Luo who have land overlapping into Western Province at the Luo land borders, and who are difficult to precisely screen, to eliminate this bias.

The extent of overall migration outside home provinces, therefore, is proportionally small compared to the numbers of each people who still inhabit their home areas. In Table 1.6 I have broken down movement within home provinces outside traditional home districts. Only the Bajun, Taita, Nandi, Kipsigis and Keyo have more than ten per cent of their population living outside traditional home districts within their home province, and when the threshold is lowered to 7 per cent, they are joined by the Kikuyu,

Pokomo, Tugen and Turkana. The major Nandi migration has been to the Uasin Gishu District, where large numbers have been resettled on ex-European farms. The Uasin Gishu District moreover, borders the traditional Nandi District and is today considered as an extension of the Nandi country by them. Similarly 91 per cent of the Keyo migrants have also settled in the Uasin Gishu District, which borders the traditional homeland in the north east. The Kipsigis have moved in numbers to Narok District, encroaching on the southern Mau forest slopes bordering Kipsigis country which was historically disputed with the Maasai, and secondly to Nakuru District which is also close to the Kipsigis Home area. The Tugen have also joined their Kalenjin neighbours in sharing the spoils of the Nakuru and Uasin Gishu ex-European farming areas. The Turkana in contrast have migrated more evenly to the ex-European Districts of Laikipia, Trans Nzoia, Uasin Gishu and Nakuru, where they worked as herdsmen for European farmers.

The other major centre of internal migration is at the Coast, where a majority of the Bajun, Pokomo and Taita internal migrants have moved to the city of Mombasa. The Bajun and Pokomo presence in Kilifi District is accounted for by two factors. Firstly, the area is closest to their traditional areas, and they do not have so far to migrate. Secondly, Kilifi contains the major town of Malindi which attracts migrants not only because of its reputation as a trading centre, but also it is the centre of the central Kenya Coast tourist industry.

Finally, the Kikuyu also show some 7 per cent of "internal" movement, a figure which rises to 25 per cent if we class Nairobi

as lying outside the home area. Of these migrants 47 per cent have moved to the ex-white areas on the western Aberdare slopes (Nyandarua District) and 53 per cent to Nairobi.

# CHAPTER 2

## The Population Units, Sampling and Fieldwork

### 2.1 Overview

At the beginning of any physical anthropological study, previous experience tells us which types of human populations are likely to correspond to biological units. According to Hiernaux (1972), the ideal units are "those human groups formed from that way of dividing the population which comes nearest to a model of contiguous isolates, or genetically closed communities." (Hiernaux 1972:97).

In complex industrial societies where population movement is fluid, and social barriers to mating are not institutionalised beyond a minimal prohibition on marriages between incenstuous categories of relatives and a legal enforcement of monogamy, the main barrier to mating is geographical distance. Other factors that could inhibit mating, such as social disapproval of marriage partners, are difficult to separate and quantify in any social situation. Most models seeking to delineate "contiguous isolates" therefore, are based on geographical and demographic assessments of mating patterns. (See Dennis 1977[a].Chapter 3).

In third world countries however, there appears to be less difficulty in identifying "genetically closed communities". Indian castes for instance, are almost totally endogamous in many areas of India. Hiernaux moreover, argues that in sub-Saharan Africa:

> "It is relatively easy to come near to such an optimum division. In many parts of the sub-continent people live in highly endogamous tribes or castes. For example the Shilluk and Dinka, two adjacent tribes in the Sudan, intermarry at a frequency of only 0.02 in both directions (Roberts 1962). In the Bedik, an ethnic group of East Senegal, the males marry within the group at a frequency of 0.975 (Gomila 1971). The Moyen Chari Province in the Tchad is inhabited by ten ethnic groups; in these the fathers of the present young adults took their wife in their own group at an overall and statistically uniform frequency of 0.975 (Hiernaux 1969). These endogamy rates from three distinct areas of middle Africa are nearly identical and very high. The ethnic groups concerned are sufficiently closed to make them acceptable as units."
>
> (Hiernaux 1972:97)

High endogamy however, is not the only condition that must be satisfied for tribal populations to be treated as natural breeding units Hiernaux argues that:

> "They must also satisfy a condition of genetic homogeneity. This condition was satisfied by the largest (in demographic size and area occupied) of the Moyen-Chari ethnic groups, the Majingay; when the tribe is sub-divided into three regions, regional exogamy amounts to 0.10, four times higher than inter-group exogamy. No regional differentiation is observed in those metrical features in which different tribes of the area show significant differences for similar samples sizes. (Hiernaux 1969).
>
> So far there have been no combined demographical and anthrobiological survey of suitable units in sub-Saharan Africa. In any anthrobiological study on so large a geographical scale, we have to accept as units those groups which are described as endogamous and homogeneous, as long as contrary evidence is not produced. A rough estimate of the number of such units in Africa south of 22°N is 1,000."
>
> (Hiernaux 1972:97).

Hiernaux's views on the relation of sub-Saharan African tribal populations to natural biological populations are very important, as they reflect the general attitude of physical anthropologists to these populations. The conformity of most physical

anthropologists to these views is best illustrated by the numerous examples in the literature of researchers who have procured samples of named tribes, for example the Hehe (Roberts et.al., 1976), the Kalahari Bushman (Tobias 1961) and the numbers of biological surveyers in Africa reviewed in Tobias (1966). By procuring samples of say Kikuyu, Nandi or Dinka, we are basically acknowledging through the ethnic definition of such populations that endogamy is maintained by socio-cultural barriers to mating. By assuming that the sample is typical of the wider unit, we also assume genetic homogeneity for the total population.

In the section which follows I will question Hiernaux's views on the nature of the correspondence between so called "tribal" African populations and biological units. I will question the use of the term "tribal" to designate a particular type of population in Africa whether now or in the past. I will argue that the ethnic units used in Hiernaux's studies are more a creation of the colonial and post colonial eras than of tradition. In particular I will examine the "high rates of endogamy" listed by Hiernaux, and suggest that genetically homogeneous populations are probably uncommon given the nature of pre-colonial African history. Whilst demonstrating how Hiernaux's approach is unsatisfactory, I will suggest a more appropriate and useful definition of the units of study which physical anthropologists can use.

## 2.2 The nature of "tribal" populations in Africa

The New Century Dictionary and Cyclopedia (quoted in Fried 1975:7) defines a tribe as a "primary aggregate of people in a primitive

or barbarous condition". No anthropologist today would associate him or herself with the underlying sentiments of such a definition. Nevertheless, peoples living in Africa and other post-colonial areas of the world are invariably called "tribal" peoples, in an effort to differentiate them from the more complex societies of Europe, America and some parts of Asia. The word "tribe" through its vagueness is useful, because in an all-embracing symbolic sense it encapsulates the radical differences between "civilised" societies and less sophisticates ones.

The symbolic power of the concept of "tribe" however, is much greater than its scientific usefulness, as it is difficult to find concrete characteristics that define a "tribal" society. Indeed, the sheer diversity of peoples encompassed by the term is immense. What have small hunting and gathering groups such as Pygmies and Negritos in common for instance, with pre-literate centralised kingdoms such as are found in Uganda and Nigeria? Yet both types of societies have been called tribes in the past. We can well sympathise with Uchendu when he remarks:

> "the word tribe ... lacks a precise meaning. When large population groups like the Ibo and Yoruba, each numbering five million, possessing a common language, a similar sense of historical experience, a common world view, and sharing a strong sense of common identity in relation to other population groups - characteristics associated with national entities - are not terminologically distinguished from small descent groups, one wonders how absurd conventional labels can become."
>
> (Uchendu 1970:53)

Social anthropologists today rarely use the word "tribe" any more, and prefer the phrase "small-scale" to describe the types of societies that inhabited many third world countries in pre-colonial

times. Small-scale societies according to Burton Benedict (1969) are distinguished from more sophisticated societies by being composed "chiefly of primary groups", a primary group being a localised collection of families occupying a minimal territorial section to provide them with land for subsistence. Primary groups moreover, fall into two types. In the first the "total social field is small", (Benedict 1966:24) or in other words the primary groups are highly autonomous and self sufficient with little interaction between them, while the second comprises "a series of interlocking similar small groups .... Island societies such as Tikopia and Dobu are examples of the first type, and segmentary societies (e.g. Nuer, Tiv, Tallensi) are instances of the second type" (Benedict 1966:24). Nadel, quoted by Benedict, provides a model for the second type:

> "Think for example, of a "tribe" divided into a number of extended families; these all duplicate each other, both in their structure and in their modes of action; each is relatively self-contained, and such relations as obtain between them (inter-marriage, economic cooperation etc.) do not follow from their constitution or "statutes", but are contingent upon circumstances and outside interests. Though such segments may in fact combine to form the society at large, they could exist without each other and in any number; one could add or subtract from it without affecting the working either of each segment, or of the embracing group."
>
> (Nadel 1951:178).

Nadel's model explains (to paraphrase Benedicts next point), why large African societies such as those mentioned by Uchendu earlier, who number millions, can still be regarded as small scale. The primary groups (usually a village or lineage territorial segment, or in the case of hunters, a band) functioning semi-autonomously, are the basis of practically all sub-Saharan traditional societies. Even in centralised kingdoms such as those

in Uganda, the primary group prevailed at the local level, and often annual tribute to the king was the only effective contact with the wider society (Beattie 1971).

The extent to which primary groups interact and participate in the wider society can be used as a criterion to type sub-Saharan African societies, as follows:

1) Hunting Bands

The hunting and gathering peoples fall into the first type of primary group outlined by Benedict (1966:24). The band for most of the year operates as a completely autonomous subsistence and social unit. Individual members of bands however, sometimes move to neighbouring bands, and there is exchange of marriage partners between bands, which is necessary given that members of a band are usually highly related, and rules of exogamy have to be complied with. There was however, before European contact,no overall conception of the extent of the distribution of the wider hunting people, and contact with more distant bands was virtually non-existent. The classical example of this type of society are the Ituri Pygmies (Turnbull 1965, 1972). Many African hunting peoples moreover, (e.g. the Pygmies, Sanye, Ik), interact for some of the year with neighbouring pastoralists or agriculturalists with whom they trade, exchanging primary produce (ivory, meat), for iron implements and other necessary tools. In the case of the Pygmies for instance, their relationship with their Bantu neighbours is intricate, and one way inter-marriage of Pygmy women to Bantu men is not uncommon (Turnbull 1972). Each hunting band however, has independent contact with the Bantu and there is no overall political cooperation between bands in opposition to the agriculturalists.

2) Acephalous agriculturalists and pastoralists

This group is very common and widespread in Africa, and conforms to the model described by Nadel earlier, (Nadel 1951). The organising principle of this type of society is that the semi-autonomous lineage based primary groups are linked in an hierarchical fashion to wider lineage territorial groups. The split of each lineage branch is associated with a common ancestor, who existed more remote in time the larger the scale of the lineage. At the top of the pyramid is the mythical founding ancestor of the whole people, who provides the focus of the widest bonds of common identity in the society. In practice however, social interaction and relationships become very weak and ephemeral beyond the primary group and minimal lineage levels of affiliation. For instance any dispute involving members of different maximal lineages is often associated with violence and a lack of effective means of enforcing dispute settlements, owing to the absence of close ties between the primary lineages of the two sections. Classical accounts of segmentary lineage societies are those of Evans-Pritchard (1940) of the Nuer, Bohannan (1953) of the Tiv, and Middleton (1970) of the Lugbara.

3) Acephalous societies with developed age-set systems

Although the local primary groups are organised along the lines of a lineage based system similar to that of segmentary lineage societies, the ties between primary groups and lineage segments are greatly re-inforced by the operation of an age system of organisation which cross-cuts the segmentary lineage structure. Throughout the territory of a people, young adolescents are simultaneously initiated into a named age grade every ten years or so. At any time there are

usually four age sets in existence, junior and senior warrior grades, and junior and senior elder grades. In the most sophisticated of these societies, such as the Karimojong (Dyson Hudson 1966), the warrior and elder grades are formally part of generation grades. When a new age grade is initiated, the previous age grades move up one grade in the cycle.

In the nineteenth century, the age grade system was recognised as a highly effective organisation for mobilising for warfare and raiding. Small-scale raiding was normally carried out independently by the warrior grades of local lineage territorial segments, supervised by the elders. For bigger ventures however, the local lineage segments would unite according to wider lineage based groups to form small armies. In rare emergencies, when the whole people was threatened, the warrior age grades from all territorial sections would unite.

The age set system thus provided a much stronger focus of overall identity than the segmentary lineage system alone. The very fact that the initiation of age grades was carried out simultaneously for the whole people meant that all primary sections were aware of the precise extent of the wider society, and that elders from all sections exercised procedures to cooperate in deciding when the new age set was to come into being. Amongst the Karimojong for instance, delegates from the whole people would meet at a central locality roughly every thirty years to instigate the new generation grade (Dyson-Hudson 1966). Other classical accounts of age-set societies are those of Spencer (1965), for the Samburu and Gulliver (1963) for the Arusha.

4) Centralised acephalous based societies

In these societies the traditional segmentary structure predominated at the local level, but the system had been fundamentally changed by the centralisation of the society, through the formation of a kingdom . The focus of identity was provided by the king and his ministers, who ruled from a central capital, their power administered through local chiefs who exact tribute (food, cattle and women) from the local villages. Although local disputes and everyday social and economic business was the independent preserve of the local lineage groups, more serious legal and economic decision making had to be referred to the chief and sometimes to the king for settlement. The de facto semi-autonomy of the local groups moreover, was often interrupted by forced participation in centralised policy implementation such as building roads and military service. Classical accounts of these societies are those of Roscoe (1911) for the Ganda and Beattie (1971) for the Nyoro.

The relative cohesiveness of primary groups and the strength of feelings of identity of their members with the wider society thus varies considerably in Africa. For hunting bands like the Pygmies, there was little awareness of distant Pygmy groups, and no overall conception of a wider Pygmy society. Highly segmented peoples like the Lugbara, as Middleton points out "recognise themselves to be one people, although they rarely know the boundaries of their own country." (Middleton 1970:204). Peoples with age sets however, not only recognise the boundary of their own country, but occasionally act together to forward mutual interests. Finally in centralised kingdoms, not only is there a strong identity with the

whole society, but also institutionalised administration of all districts by the central body.

Even in the most centralised kingdoms however, the local primary groups were never fully integrated into the wider society. In times of weak central authority there was nothing to impede primary sections from migrating to other areas, as fully viable subsistence units. In fact it is possible to visualise all African societies as having been unstable over time, with disasters such as famines, epidemics and major wars leading to rapid fragmentation of wider bonds of identity, and the migration and reassembly of local primary group sections. There are many examples of African peoples such as the Mandari (Buxton 1970), the Luyia (Were 1967), and the Kuria (Ogot 1967), who are recent fusions of incoming migrant groups. The scale of disaster or population pressure however, needed to break up a society or parts of a society would have had to have been much greater for the more cohesive types of society.

The sense of common identity, that was so weak in pre-colonial peoples in Africa has become significantly strengthened as a result of colonialism. The effect of colonial administration was to transform a fluid situation to one where groups and territorial boundaries of existing peoples were mapped and rigidly defined. This did not prevent disadvantaged families from carrying out the traditional practice of migrating to seek better land or opportunities, but it meant that the movement of such families was carefully controlled. In Kenya for instance, during the 1920s groups of families without land due to the overcrowding in some of the "reserves" squatted on European farms. Whereas in the past they would probably have been assimilated by the people in possession of

the land, they were instead either evicted or tolerated on condition that they worked for at least 180 days a year for the landowner (Leys 1975:47).

One consequence of the definition of "tribal" lands was the tendency of the European administrators to view as "tribes" peoples which corresponded to cultural blocs rather than to pre-colonial autonomous units. In his ethnography of the Nuer, Evans-Pritchard (1940) was careful to point out that the Nuer were not one people but a collection of distinct and autonomous "tribes". By tribe he meant something very specific. Nuer tribe was a unit that had "a common and distinct name , a common sentiment, a common and distinct territory, and a moral obligation to unite in war." (Evans-Pritchard 1940:123). As Tait puts it a "tribe" is "the largest unit of common values" (Tait 1970:68). The unit referred to by these anthropologists is the maximal unit of traditional identity for segmented and acephalous peoples, and distinguishes one autonomous society from another WITHIN culturally very similar peoples. Unfortunately, the colonial administrators did not make any such fine distinctions when they labelled their maps. They tended to refer to the "Maasai tribe", not the autonomous units within the Maasai such as the Purko, Keekonyokie or Kaputie, or to the "Meru tribe", not to the Chuka, Mwimbi, Igoji etc. To make matters even more confusing, other peoples were defined as they used to be in pre-colonial times (e.g. the Nandi). Thus if we compare the Maasai with the Nandi, we are in reality comparing a cultural and linguistic bloc, with an autonomous unit within a linguistic and cultural bloc. The Maasai ought to be compared to the Kalenjin

group to make comparison compatible, or at a higher level of resolution a Maasai sub-unit, say the Purko, should be compared to the Nandi.

For practical purposes however, the identification of "tribes" with linguistic and cultural blocs did not really matter, as scales of identity rapidly widened for such peoples. Migrants to the towns and European farms in particular tended to identify themselves with the lowest common denominator of linguistic and cultural similarity, which was usually the ethnic bloc, not the previously autonomous ethnic unit within the bloc. For the first time for instance, Luo from different ethnic units began to see themselves as one people, the Luo, in opposition to other cultural blocs. The nine coastal Bantu peoples slowly began to call themselves Mji Kenda rather than Digo, Duruma, Giriama or Rabai. Only in the most remote regions did previous patterns of identity remain, small-scale and static. The broadening identity saw the creation of cohesive peoples that were never unified before. The classical example is the Luyia, who "came into existence in 1935. Before that time no such group existed, either in its own or anyone elses estimation." (Southall 1970:35).

The broadening and consolidation of ethnic blocs is above all a political process. Initially, the cohesion of culturally similar peoples was achieved by the efforts of politically aware Africans, in order to mobilise grass root support not only for opposing the colonial administration, but also for consolidating power bases in rivalry to other similar power blocs. This phenomenon, known as "tribalism", is to be found all over sub-Saharan Africa, and has greatly accelerated in the post-colonial era as political

rivalry has intensified. Its pre-eminence in politics can be explained by the fact that alliance by ethnic similarity consolidates traditional loyalties and identities, without sacrificing modern aims and ambitions. As Lloyd remarks (for Nigeria),

> "The large ethnic units are united by cultural features which are also what sharply distinguishes them from neighbouring groups. The factors which have united the various Yoruba speaking peoples cannot, at a higher level unite them with the Ibo. The particular segmentary structure of Ibo society which has engendered such strong sentiments of cohesion, cannot be replicated at a level which would unite all Nigerian peoples. Ethnic exclusiveness provides a ready basis for political separatist movements."
>
> (Lloyd 1969:302).

These ethnic groups are thus very much a creation of the colonial and post-colonial eras (Leys 1975:199), and whilst the elements that characterise them may be drawn from their traditional past, they cannot for the most part be equated with any pre-colonial groups. Many of the 1,000 or so ethnic groups that Hiernaux (1972: 97), discerns in sub-Saharan Africa therefore, have a very recent history as cohesive structures.

This factor however, does not invalidate the view that such large ethnic units correspond with breeding populations. Indeed the trend of these units to become increasingly cohesive and culturally homogeneous could be regarded as a factor reinforcing this attitude. The recent origins of the units however, diminishes their value for physical anthropological studies, as they only indirectly mirror pre-colonial structures. The Luyia for instance, may be developing into a homogeneous cultural unit, but to treat them as such in a physical anthropology sampling scheme would be to ignore the reality that the previously autonomous sub-units may have diverse origins.

## 2.3 Endogamy in African tribal populations

The extent of endogamy apparent in African tribal societies (Hiernaux 1972:97) is much greater than one would expect from the operation of geographical barriers to mating alone. It is reasonable to infer therefore, that socio-cultural factors are of considerable importance in helping to maintain breeding isolation. African tribal societies, differ significantly from tribal pooulations in Asia and America, and from castes in India, in that there are no rigid rules restricting the choice of marriage partners to certain categories of kin (see Levi Strauss 1970), or rules prohibiting marriage to outsiders. In African tribal societies the marriage rules, though often complex, are mostly concerned with exogamy. There are no rules prohibiting inter-marriage with other tribes. Endogamy therefore, when strong, is maintained by preference. In other words, people will prefer to marry one of their own kind owing to incompatibility with or without distaste for linguistically and culturally different peoples.

The power of cultural prejudice however, is not normally all-pervasive. We should note for instance, that the traditional cultural antipathy of the British towards certain other European peoples has not prevented inter-marriage when the opportunity has arisen. For example the fact that Britain was at war with Germany and Italy during the second world war did not stop many British soldiers from returning home with German and Italian wives, even though the Germans and Italians were not well regarded at the time. In Africa, there is no reason to suppose that tribal Africans are any less inclined to ignore traditional prejudice in these matters.

Far more important than cultural prejudice in helping to maintain endogamy are the positive social and economic advantages of marrying within "tribes". For instance in many parts of Africa payment of bridwealth is a central feature of marriage (Mair 1971). The more distant social and territorial ties become between the bride's family and that of her husband, the more difficult it is to enforce the payment of bridewealth. This can have serious consequences for the future marriage and re-marriage prospects of the bride's brothers and her father as they depend on her bridewealth to help pay for their own wives. Any girl who married a man from another tribe would be effectively an economic loss, as her father would only have good will to enforce the payment of bridewealth. In the absence of legal ties across tribal societies, considerable force would have to be used as a sanction, a usually impractical alternative (see Evans-Pritchard 1940).

The disadvantages of a man leaving his home to marry outside the "tribe" can be even more serious, as in many African societies it is his patrilineal kin group that provides him with the bulk of the economic and social support that he might need. It helps raise brideprice for his marriage, helps support any legal actions he might instigate, helps him to organise his herds or fields, and with the support of his kin, he may become a politically influential stateman in later life. By migrating to go and live in another "tribe" a man will virtually forsake all this essential support and his inherited rights. Although he might recuperate some of these benefits by attaching himself to his affinal relatives, it is a poor alternative. Both men and women therefore, would

appear to gain little from marrying outside the tribe, and have much to lose.

In spite of these reasons for preferring intra-tribal marriages however, the fact remains that there are significant instances in the literature that indicate that inter-tribal marriages are perhaps not as uncommon as is implied by the very high rates of endogamy mentioned by Hiernaux (1972:97). Turnbull for example, mentions that

> "there is a limited amount of intermarriage (of Pygmies) with all the villager groups and it maintains a single outward direction, village males taking Pygmy females, the offspring of which are said to never return to the forest, but remain, with their offspring as villagers. Thus the pygmy population contributes to the village gene pools but is not itself affected."
>
> (Turnbull 1972:296).

A more extensive example of a process of one-way inter-marriage is given by Spencer (1973:137-142). Spencer shows that there is considerable intermarriage and assimilation of Rendille by their Samburu neighbours, an arrangement that has mutual advantages. The Cushitic speaking Rendille, because they live in a very arid environment which permits only camel herding, can only support a very low density of population. The surplus population migrates to Samburuland, which is far less arid and able to support a much higher population being based on a cattle economy, and is readily assimilated by the Maasai speaking Samburu. Of crucial importance is the fact that in Rendille society marriage is always late and generally monogamous, and this tends to create a surplus of women. The Samburu however, suffer from a shortage of marriageable women owing to the practice of polygyny by the elders.

Rendille women when they migrate are thus very welcome additions to Samburu society, and their assimilation is facilitated by the presence of previous Rendille migrants. In spite of linguistic and cultural differences Rendille/Samburu intermarriages is advantageous to both peoples, and has resulted in a considerable change in the genetic composition of the Samburu. Spencer (1973:138) shows moreover, that the extent of this intermarriage is considerable. Out of a sample of 149 Samburu men, only 83 per cent married endogamously, and of the rest, only 1 per cent were not accounted for by inter-marriage with Rendille.

An example of assimilation of a different kind is described by Evans-Pritchard who notes that "over the generations the Nuer have absorbed many Dinka groups, and war fare between the two peoples was a normal state of existence, even in comparatively recent times". (Evans-Pritchard 1940:221-228). Captured Dinka were not killed but adopted and assimilated into Nuer society, where they had little difficulty finding wives or husbands. The scale of this assimilation was considerable. Evans-Pritchard also states that "a very large number of Dinka have been incorporated by adoption into Nuer lineages." (Evans-Pritchard 1940:223). Moreover, capture and adoption were not the only sources of assimilation and intermarriage. "There are also many Dinka lineages in Nuer Society descended from men who came of their own free will to settle in Nuerland, either to escape famine, to visit captured sisters, or to occupy sites from which Nuer raids had ousted them." (Evans-Pritchard 1940;223). It is not inconceivable that many of these more peaceful processes of assimilation are still taking place in Nuerland.

These examples show that assimilation and intermarriage may have been widespread at one time, and that they probably still are in many areas. For instance during my own fieldwork I noted considerable intermarriage between Kikuyu and Keekonyokie Maasai around the Ngong area west of Nairobi. How then, can the very high endogamy rates given by Hiernaux (1972:97) be explained? Firstly, I should point out that some of the peoples he mentions (the Dinka and the Moyen Chari ethnic groups) are not really "tribal" units in the Tait and Evans-Pritchard sense, but cultural blocs who occupy very large territories. We would therefore expect most marriages to be endogamous through geographical factors of distance alone. In other words, anybody who does not live close to the boundaries of the bloc will have little opportunity to meet anybody of another people, all but a small percentage of the total population would be in this position. Secondly, it can be very difficult to identify people who may have married in from other peoples, as such people will tend to identify themselves with their adopted people when asked "what is your tribe", and can be reluctant to divulge their origins. This especially is true if there is a history of assimilating bordering peoples. Obviously it would be much easier to detect intermarriage if the foreign marriage partner is from an area far from the borders.

What is probably more relevant however, is to ascertain how much intermarriage there is between previously autonomous units within cultural blocs. Unfortunately there have been no surveys done at this level of population definition, but one can surmise that intermarriage could be considerable between members of adjacent units.

Therefore Hiernaux's practice of seeing African ethnic units as natural biological populations owing to the high rate of endogamy they seem to exhibit, is misleading for several reasons. Firstly, most of this endogamy is in reality significantly maintained by geographical distance as well as socio-linguistic factors. Secondly, there is contemporary and historical evidence that shows that a very high rate of endogamy is by no means universal in African societies. Above all, ascribing high rates of endogamy to large cultural units implies they have always been stable and cohesive single societies, rather than recent integrations of diffuse autonomous units.

## 2.4 The extent of historical and biological heterogeneity of African peoples

African historians have long recognised the fact that some African populations (if not the majority), are a heterogeneous mixture of successive waves of incomers. For example Buxton remarks in the case of the Mandari of the southern Sudan:

> "Historically the Mandari regard themselves as being composed of different levels of incomers, superimposed on earlier populations. This building up has been the result of immigration and absorption, which is clearly shown in their histories."
>
> (Buxton 1970:69).

There are plenty of other examples in the literature for example

> "The Lango-Omiru are not a homogeneous people with a single past history, but a composite mixture of elements."
>
> (Crazzolara 1960:176).

> "The Lango are an amalgam of Luo, Teso elements with the former predominant."
>
> (Ogot 1967:62)

> "History of the Acholi is the history of the clans. Each was a political entity with its own history."
>
> (Bere 1947:1-8).

It could be argued that African oral historians (e.g. Ogot 1967, Were 1967, Muriuki 1974), concentrate too much on the history of micro units (clans, lineage territorial segments). There is a danger for instance, that small-scale shifts of positions by minor groups might feature prominently in presenting a much less stable picture than was actually the case. Nevertheless, given the weak links between primary groups it is not inconceivable that the heterogeneous picture holds good for many peoples.

Physical anthropologists have been slow to appreciate the consequences of the potential instability of nineteenth century acephalous societies. For example, in the classic study by Talbot and Mulhall of Nigerian peoples, which included an analysis of intra-tribal variation amongst the Ibo, initially conformed rigidly to the conventional paradigm that the Ibo are a homogeneous tribal unit. They remark that

> "The Ibo have long been regarded as culturally homogeneous ... and in the course of their history appear to have kept very much to themselves; and intermarriages between them and other peoples, when it occurred was a rare event."
>
> (Talbot and Mulhall 1962:4-5).

When however Mulhall analysed the results of anthropometric differences between 20 Ibo sub units, he was surprised to discover that "the results suggested an unexpected variation of physical characters among the Ibo groups examined" (Talbot and Mulhall 1962:94). The analysis of a $D^2$ matrix between the Ibo groups and neighbouring peoples revealed a close relationship between the Central Ibo and the Ekoi. We can infer from this that the Northern and Central Ibo

groups either originate from their respective neighbours or have heavily intermarried with them over time. The Ibo are thus clearly a phenetically heterogeneous people.

A more recent study of intra-tribal variations is that of Gomila (1972) on the Bedik of Senegal. His analysis of $C^2H$ distances between six Bedik villages and neighbouring populations showed that the range of distances for the villages was greater than most of the distances between the villages and the outside populations. This picture of heterogeneity is particularly interesting, as the Bedik are not a large people (Gomila 1972:133).

Following these examples, we may ask how far the so-called homogeneity of ethnic groups in Africa is correlated with a general absence of empirical observations by physical anthropologists, and also to an over-reliance on the notion that stated high endogamy is a reflection of genetic and phenetic homogeneity. Both the historical and available biological empirical evidence shows that Hiernaux was mistaken in recommending that in the absence of such evidence "we have to accept as units those groups which are described as homogeneous and endogamous (Hiernaux 1972:97). I feel, on the contrary, that we should not accept any apparently endogamous or homogeneous African population at its face value, and that we should assume heterogeneity unless proved otherwise. In stating this position, I do not wish to imply that no African peoples are homogeneous, or stable endogamous units over time. I am only saying that it is unlikely that a majority will be so, and therefore we should not assume these characteristics prior to sampling. Indeed a central problem of any study must be to discover which populations are homogeneous and which are not. If we do assume

heterogeneity when sampling, then it follows that no single sample from a named people can be considered typical of the whole people but only of the geographical locality from which it is drawn within the territory of that people, unless of course it is a pooled sample incorporating elements from all geographical localities.

The evidence presented earlier in this Chapter, demonstrates that many of the ethnic units mentioned by Hiernaux of which there are about a thousand in Africa (Hiernaux 1972:98), are not really "natural" population units, but creations of the colonial and post-colonial eras. They represent cultural blocs who have only become cohesive units this century. The natural sampling units therefore, are those that Evans-Pritchard (1940) and Tait (1970) call "tribes", or what Ogot (1967) and Were (1967) refer to as "sub-tribes" or what I term "minimal ethnic units", in other words those sub-divisions of present major "tribes" and linguistic/cultural blocs which in pre-colonial times represented independent autonomous peoples. Peoples such as the Luyia, Dinka, Nuer, Luo and Ibo contain many previously autonomous sections.

If we define units at this level of resolution it means that there are thousands of such units in sub-Saharan Africa. Although any survey would greatly benefit by sampling only units at this level, in the large continental survey that Hiernaux undertook (Hiernaux 1968[a]), which also depended on other peoples' results, he is obviously justified in treating the wider cultural/ linguistic blocs as the sampling units for practical purposes. Nevertheless it must still be borne in mind that these large ethnic units, though practical, are not really the "natural" units to sample.

In the present study, which is limited in scope to one country, it would be quite inappropriate to define sampling units at the level of resolution Hiernaux uses. The correct sampling design is to clearly define the identity of minimal ethnic units in Kenya and to try and obtain a representative sample from each unit. In pursuing this type of design, physical anthropologists in Africa, are fortunate that a majority of members of most minimal ethnic units will reside in their pre-colonial localities, and people will still tend to identify themselves with such units in opposition to others at a similar scale. A Purko Maasai for instance, will readily tell you that he is not a Kaputie Maasai, although in wider contexts he might identify himself generally as a Maasai in opposition to non Maasai.

The tendency of people to identify themselves in terms of the minimal ethnic unit they belong to however, is a phenomenon nowadays confined to traditional rural areas. When members of such units move to the cities, they tend to identify themselves with wider ethnic units. For example a person who sees himself as an Idakho in Kakamega District in Kenya, in opposition to other minimal ethnic units such as Tiriki, Maragoli or Bunyore, will quickly change his allegiance to the Luyia cultural unit (of which the Idakho, Tiriki, Marogoli and Bunyore are divisions today), when he moves to the city. His children and grandchildren born in a city miles from the traditional Idakho homeland will sooner or later forget their roots. For them their historical association with Idakho will be irrelevant compared to the social and political realities of being a Luyia in opposition to urban members of other macro ethnic units such as Meru or Kalenjin.

In view of this it is essential when sampling in Africa to obtain rural samples from traditional areas. It is quite unreliable to obtain urban samples in the hope that they will typify minimal ethnic units of origin. Such samples can of course, be used as representing wider ethnic units, but since the physical anthropological value of such macro samples is limited, urban sampling is clearly an unprofitable practise except for exploratory and very general surveys.

## 2.5 Sampling design and fieldwork

### 2.5.1 The Population units

The Kenya Population Census 1969 lists 41 "tribal" populations in Kenya. Some, such as the Luyia, Maasai, Sabaot, Luo, Meru are ethnic groups, while others, such as the Nandi, Tugen, Pokot, Rendille and Taveta correspond to minimal ethnic units. When the scale of population definition is made compatible, there are well over a hundred and forty minimal ethnic units. These are listed in Table 1.1 in relation to the linguistic groups and divisions to which they belong. Even this list however, is an underestimate. For instance the Gusii minimal units are not given as I have not yet found a complete list. It is in fact remarkably difficult to obtain detailed and accurate information of the distribution of minimal ethnic units for some peoples that are popularly treated as a homogeneous "tribe", (e.g. the Somali).

A majority of Kenya's peoples (Maasai, Cushitic, Kalenjin, some Western and the Highland Bantu, and some coastal Bantu, Karimojong and Luo Nilotes), have some sort of age set organisation.

They are therefore, traditionally more cohesive structures than the classical segmentary lineage societies. This is especially true for the Maasai and Highland Nilotes, the Galla and Somali Cushites, and the Highland Bantu which have well developed age set systems. Classical segmentary societies are relatively uncommon. The Luo with poorly developed age sets resemble the Nuer in many points of organisation (Evans-Pritchard 1949), and western Luyia and some of the coastal Bantu group (e.g. the Taita), were unaffected by the diffusion of age organisation. The Kenyan hunting groups all have a band structure, the most autonomous groups being the Sanye and Boni, and some of the Laikipiak and Samburu Dorobo (e.g. the El Molo). Many of the Dorobo groups however, especially the groups in the Dol Dol area and the Kalenjin speaking Ogiek have begun to adopt the age grade system (e.g. the Ngwesi, Luasi, Momonyot). The centralised type of society was represented chiefly by the old Wanga Kingdom, but it has long ceased to exist as an autonomous society and today is one of the 18 Luyian minimal units.

### 2.5.2 Fieldwork

When I arrived in Kenya in February 1973, I decided to adopt a simple sampling design. This was to obtain a dermatoglyphic sample from as many of the Kenyan minimal ethnic units as I could cover in the time available. I was to achieve this by visiting the traditional home territory of each minimal ethnic unit, and by sampling schools within the unit.

The use of schools is one of several options that I could have adopted. I could for instance, have decided to concentrate

on obtaining village samples, samples from hospitals, European farms and plantations or urban meeting centres such as churches. Village samples however, though essential if a micro-genetic study is being conducted, are too impractical for the large scale survey I had in mind. It is extremely difficult to persuade adults to participate in a dermatoglyphic survey. Firstly they usually cannot see the point of it, and tend to confuse the researcher with the police. Secondly it is not easy to gather sizeable groups, and to put up with impatience as they await their turn. Above all, however, is the slow negotiations needed to get the permission of the local chief and sub-chiefs, and the leading elders, to conduct the survey in their area. All these problems can be overcome with patience, but it would take a long time to procure each sample.

Obtaining the cooperation of farm workers on European farms is not difficult. On the other hand such samples, in common with urban and others obtained in non traditional areas, tend to be inadequate. Firstly, as I mentioned earlier (page 128), these migrants or descendants of migrants tend to identify themselves with wider ethnic groups, and the areas of origin in their traditional homelands are often imprecisely stated. They are thus not very reliable samples of minimal ethnic units. For some peoples (e.g. the Nandi, Tugen) this does not really matter, as they are minimal units in their own right. On the other hand the fact that the sample is not from the home area means that it becomes difficult to accommodate them in fugure analysis if a geographical or ecological orientation is taken rather than an ethnic one.

Samples from plantations and urban meeting centres also have the disadvantage that the researcher had to deal with

sophisticated and consequently very inquisitive people, who are particularly suspicious of "finger prints." My one attempt at testing plantation adults almost caused a strike.

Finally thanks to the interested attitude of the medical staff, it is easy to obtain samples from hospitals if full permission is obtained. People everywhere always seem to cooperate in the presence of doctors and nurses. On the other hand hospital samples can be biased, as people who fall ill are not necessarily a random sample of the people at large. Sampling therefore, should be confined to pregnant mothers, mothers who have just given birth, or to outpatients with accidents or minor ailments. Hospital samples are on the whole not to be recommended.

Schools therefore, are the only really practical alternative for a large scale survey. Kenya moreover, is one of the countries whose education programme has expanded out of all recognition since 1963, and not only has every district an official District Secondary School for both sexes, but most districts have additional Government or missionary secondary schools. Moreover there is scarcely a locality in the less remote districts without a primary school. Absence of schools therefore, is not a problem.

Of the two types of schools, secondary schools provide the best samples. Their catchment area is much larger than that of primary schools, and pupils come from a large portion of the area covered by the people of interest. Secondly owing to the fact that most Kenyan parents cannot afford to send more than one or two children to secondary school, the likelihood of bias through the inclusion of close relatives is greatly diminished. Moreover

the generally much better discipline and education in secondary schools makes it easy to weed out any relatives. Secondary schools however, have one disadvantage. They are serious hard-working establishments, and I found headmasters often reluctant to disrupt their school schedules. They usually did not mind when I tested for colour-vision defects, as the turnover of subjects was rapid. The taking of hand and finger prints however, is a much slower business, and some headmasters became uncooperative after the first hour or so. A few went as far as to refuse altogether. As a result, and for fear that I might be reported as a nuisance to the District Authorities, and also in order to avoid unnecessary offense, I concentrated mainly on primary schools for dermatoglyphic collection, although I used secondary schools for colour-vision as much as possible, and for dermatoglyphic collection when the changes seemed favourable.

Primary schools were almost always under the direction of headmasters who were only too pleased to welcome me. This advantage however, has to be weighed against several problems. Firstly, the catchment area is much smaller than that of secondary schools. Secondly, as primary school fees were much lower than secondary school fees (primary school fees were abolished in 1976 after my fieldwork was completed), parents could afford to send more of their children to school. Consanguineous relatives are thus much more common in primary schools, and more difficult to identify. Thirdly, primary schools tended to be far less disciplined, although this varied considerably. In some primary schools the teachers would

leave me with all the children to control, with others (a majority) were orderly but spirited. At the extreme end of the discipline scale was a school where some trivial misbehaviour resulted in 36 boys and girls being caned en masse. (One particular primary school in the Meru area was striken with panic when I arrived. The headmaster and I watched in amazement as children fled from classrooms and disappeared screaming into the surrounding fields. It turned out that they thought I was from the blood transfusion unit.)

Bad discipline at schools can result in serious bias as well as hampering the speed of data collection. The most obvious source of bias is from mischievous children who think it amusing to have their prints taken twice. In the poorly controlled schools (a small minority), I overcame this problem by marking the hand with some endorsing ink, after printing was completed. Secondly, bad discipline can result in inefficient screening of relatives. When a whole school was left in my charge I ordered them to line up behind their brothers and sisters, or half-siblings, and daubed all but the eldest of each file with endorsing ink. At my first badly disciplined school, I tried to employ as a punishment a stiff neat solution of PTC, which I would invite the offender to drink. Unfortunately it did not work, as all the other children also wanted to try it. The only improvement in discipline occurred when I told the children that they could only try it if they lined up quietly.

For the majority of organised schools, where children were kept in their classroom, and teachers accompanied me while I worked, I found it less disruptive to visit each class in turn rather than call the children to an empty room. The screening of relatives was

achieved by getting teachers to point them out, and subsequent confirmation from the children. I would start by visiting the youngest children first, and asking all siblings and half-siblings and uncles and aunts to stand up. (Uncles and aunts the same age as their nephews and nieces were occasionally encountered). Cousins were not called out as the definition of a cousin can be very complex in lineage societies. The eldest relative was then asked to sit down, and the remaining relatives were demarcated to isolated desks. In the next classroom the same procedure was repeated, except that they were also asked if they had relatives in the class already done. Quite often by the time I reached standard seven there were few eligible subjects left. The eldest classrooms were visited last as they were better able to remember relatives in the host of lower forms, while all the youngest children had to remember was the relatives in their own or contemporary classrooms.

Teachers were also asked as a matter of course whether there were any pupils from different "tribes" other than from that being sampled. Those found were subsequently carefully segregated. Even in the most isolated areas odd individuals from other tribes were sometimes encountered, usually the sons and daughters of Government employees who worked in the area.

Dermal prints of palms and fingers were taken using for the most part Kleenprint paper and ink. Owing to the fact that I was often short of this paper, I had to devise very economical methods for using it. I found that with practise it is possible to easily fit two palm prints on to each sheet, with rolled finger prints round the edges. For young children or those with small hands, prints from two individuals could also be fitted with comfort

on one sheet of paper. Nevertheless, I did run out on several occasions. When this happened I used police finger print ink, and when that ran out, washable printing ink, with hands printed on bond typing paper. I could detect no real loss of quality for this latter method. (About a thousand people were printed with ink.)

I was very careful to ensure that those dermal areas associated with ridge counting were clearly printed. This often necessitated taking more than one finger print of the same digit, or of the interdigital area alone. In fact I would recommend the separate printing of the interdigital area if there is time, as it usually gives a much clearer print. The best method is to overlap the top of the palm on to the paper over the edge of a table. In spite of these precautions however, some individuals still came out badly, owing to sweaty hands that just would not dry, or because of very indistinct ridges. In order to standardise atd angles, all palm prints were taken with fingers together. Finally sponges were not normally used, as I found that the prints came out clearer without them if the centre of the palm was pressed firmly, and not too much ink was applied. Only in the case of individuals with bony and hollow hands was a pad used. Even then failure often resulted, and in extreme cases the paper was applied to the hand, a delicate operation if double printing is to be avoided.

Hand printing is a slow and laborious process. At top speed I found I could do about 40 per hour, but usually the figure was in the lower thirties. Speed improved significantly when I was fortunate enough to be accompanied by assistants, who took care

of the clerical organisation and the tidying up.

The total time I spent in the field was 15 months, from February to November 1973, and from June to October inclusive 1975. Of this period however, I only managed to spend eight months collecting data full time. In 1973 I was not able to start until the end of June, as the Office of the President took its time to grant my research permit. Moreover I could not visit schools in August either in 1973 or 1975, as the schools are on holiday for a month. This meant that to achieve my immediate target of sampling 60 populations I had only 3 days for each one (taking into account weekends). In practice however, it was less than three days, owing to the frequent breakdowns of an old and unreliable car in 1973, and a shortage of money after August 1975 which meant I had to use slow local transport to get from one area to another. The data collection therefore, was conducted in an atmosphere of urgency.

In 1973 I decided after the initial delay, that I would have to limit sample size to 100 from each population (50 of each sex). When I arrived amongst the Luyia however, in September 1973 (See Table 2.1), I reduced the same size to 50 (25 of each sex) to save time. This was a mistake that I have regretted ever since (except during print reading). After the Luyia expedition I came to my senses and realised that 25-50 of each sex was not enough. Some of my largest samples (Maasai females, Luo and Somali) were obtained in the last weeks of fieldwork. In 1975 I had learnt my lesson and greatly increased the sample sizes. This can be seen in the results of the two trips. In 1973 I sampled 36 populations with an average sample size of 54 for males and 41 for females, whilst

in 1975 I samples 21 male and 18 female populations for an average sample size of 76 for males and 55 for females (the above figures do not include those samples of less than ten individuals).

The precise composition of the samples is given in Table 2.1. A majority of the samples that show a marked imbalance between the sexes in size (mostly biased in favour of males), are accounted for by not being able to find enough of that sex in the schools visited (after relatives had been taken out). This is particularly true for the Chonyi, Ngwesi, Mukogodo, Turkana, Samburu, Cushitic populations, Purko Maasai and Bukusu. The much larger Keekonyokie Maasai female sample happened because I thought at that time that a surplus of females for that population would counterbalance the very low Purko female sample. In other words I thought that the males of the two populations could be pooled when it came to analysis. Only later did I discover that this is theoretically unjustified.

Inexperience also accounts for the lower Idakho, Kikuyu Imenti, Tigania and Igembe female samples (all from the 1973 fieldwork). Owing to haste and fatigue I did not count the sexes until I had left the area, and did not notice the imbalance.

The precise itinerary of the two fieldwork trips and the geographical location of the samples is presented in Table 2.2. Of the 85 sampling points, only 6 were secondary schools, 74 were primary schools, 3 were farms, one was a mission, and one a mission hospital. The sample from Syokimau Farm (24 Kamba) was my first practise before I started in earnest, as I was living there while I waited for research clearance. The two remaining farms, the mission at Kiminini and Ortum Hospital were visited during August in 1973,

while the children were on their holidays. The Bukusu, Pokot and Turkana samples which were obtained from the first three localities are not strictly geographical samples of the minimal ethnic units they represent, as they were located in the Trans Nzoia District which is an ex-White farming area. The Bukusu sample taken at Kiminini Mission however, is not far from the borders of the Bukusu home territory, and can thus be considered as reasonably representative. The Turkana sample from the farms on the other hand, is not so satisfactory, as the members of the sample originated from all parts of Turkanaland, with a majority from the southern and central districts. Only 28 Pokot were obtained from these farms, which in any case were situated close to the Trans Nzoia/Pokot District boundaries. The rest of the Pokot sample came from Ortum Mission Hospital which lies in the heart of the agricultural Pokot area. It consists totally of women who have just given birth and outpatients with minor injuries and ailments.

There is also another category of unsatisfactory samples. These are those taken in schools situated in urban areas. All but 50 of the Luo for instance, were sampled in Nairobi in a last desperate attempt to raise the sample numbers for this important people before I left Kenya in 1973. They are thus just a general sample of the Luo group, and no minimal Luo ethnic units can be distinguished. The Somali from Garissa Town and the Pokomo from Hola Town though sampled in the home territories for these peoples nevertheless are generalised ethnic group samples which are not differentiated into constituent minimal ethnic units. The reason for this is that the children could not remember precisely enough

from which parts of Somaliland (Kenya) or Pokomoland their parents had originated. The final sample worth of mention is the Samburu one from Laisamis, which is situated in an area of overlap between the Samburu and Rendille. The Samburu from that area are typical of the main body who inhabit the higher plateaux around Maralal. Firstly they are situated in a much more arid environment, and secondly they belong to that section of Samburu that have assimilated most Rendille migrants (Spencer 1973). It would be wrong therefore, to think of the present Samburu sample as being typical of the Samburu as a whole

2.5.3 Conclusion

From my fieldwork performance we can learn four main lessons:

1) It is important to carefully demarcate and map all minimal ethnic units, before sampling commences. During fieldwork in 1973 for instance, I was not aware that the Luo were a group of such units, not a single unit.

2) Minimal ethnic units should be sampled only in the areas of their pre-colonial distribution. Samples from urban areas or from lands alienated by the Government in colonial times should be avoided, unless such areas are additionally included in the sampling design.

3) Ideally there should also be regional sub-samples from within each minimal ethnic unit. At the very least, more than one primary school in different localities should be sampled.

4) Sample sizes should be a lot larger than most of those I managed to obtain. The minimum should be a hundred of each sex, with

a norm of at least two hundred. It is better in my opinion to measure less traits and maintain large sample sizes than to measure many traits at the expense of sample size.

# PART TWO: THEORETICAL APPROACH, METHODOLOGY AND DESCRIPTIVE ANALYSIS.

# CHAPTER 3

## Analytical approaches to the investigation of the present material

### 3.1 The usefulness of conventional approaches to the description of biological differentiation between human populations.

> "prerequisite for the formulation and testing of explanations of human differentiation is precise and unambiguous description of what is to be explained. Biologists and physical anthropologists cannot hope to explain differentiation of human populations until they know how to describe it."
>
> (Jardine 1971:1)

Every researcher faced with reporting the results of a human biological survey is inevitably faced with problems in presenting and describing the observed differentiation between human populations. These problems are especially acute in large dermatoglyphic surveys (for example those conducted in England and Wales by Dennis (1977[a]), Williams (1978) and Smith (1979), which testify to the large amount of descriptive results that can be produced from dermatoglyphic data). The approach of these researchers to the task of describing and analysing dermatoglyphic variation has been to organise the investigation into univariate and multivariate sections. The univariate analysis describes statistical trends of variation for individual attributes and, where possible, for groups of phenetically highly related attributes, whereas the multivariate approach considers "overall" trends of relationships between populations. In practice, the univariate anlysis corresponds to the stage of preliminary description, and the multivariate analysis to the phase where explanatory exploration is attempted.

The division of analysis into univariate and multivariate sections is a policy which is well-established in physical anthropology. For instance, Talbot and Mulhall (1962) describe univariate trends before their analysis of multivariate inter-population relationships. Gomila too (1972:118) mentions he performed a "classic and quite laborious univariate analysis" before the multivariate investigation he reported. We could say to follow a univariate analysis with a multivariate one is today a common approach for the exploration of human biological regional surveys.

In recent years, however, there has been some disenchantment with univariate approaches, and a growing emphasis on multivariate forms of analysis can be seen. Howells, for instance, condemns the conventional univariate approach "where populations have been compared (normally two populations at a time)" as "inefficient" and "not especially objective" an approach where

> "personal judgement is called upon in a high degree for the interpretation of the meaning of the differences in different measurements."
>
> (Howells, 1973:3)

He states that

> "multivariate analysis, though late on the scene, is indeed the appropriate and correct method of handling populations"
>
> (Howells 1973:2)

This view is also shared by Dennis, who finds after his conventional univariate analysis of dermatoglyphic variation in the Yorkshire Dales that

> "the univariate summaries do not lend themselves to simple univariate description"

and that

> "the nature of the univariate results necessitates a multivariate approach to identify general situations"
>
> (Dennis 1977[a]:344)

Dissatisfaction with the univariate approach, however, is not new in physical anthropology. In fact it was the recognition of its limitations which provided the stimulus for the development of multivariate measures of association. The first such statistic was devised as early at 1909 by Czekanowski (Constandse-Westermann 1972:23). Until the advent of electronic computers, however, multivariate analysis was too laborious to calculate for it to be practical for human biologists engaged in sizeable regional surveys. The emphasis on univariate analysis that has persisted until recently, therefore, can be seen not so much as a testimony to its intrinsic value, but as a pragmatic response to the fact that the desirable multivariate approaches were too difficult to implement in most research projects. Consequently the aims of most multivariate analyses in the early post-war period were directed more towards the testing of techniques than the pursuit of biological discovery (e.g. Mahalonobis et al. 1949; Trevor 1950).

Now that the wide availability of computers has rendered it practical to apply sophisticated multivariate techniques to the analysis of human biological differentiation, leading physical anthropologists have been adopting such techniques with enthusiasm (see e.g. Jardine 1971; the authors in Weiner and Huizinga 1972; Howells 1973; Rightmire 1970; and in the field of dermatoglyphics Coope 1971; Dennis 1977[a]; Williams 1979). In spite of this,

however, there remains some doubt whether such enthusiastic advocation of the multivariate approach at the expense of the conventional univariate one, is in the best interest of the subject. As Kowalski remarks:

> "There does exist another side of the story, and that much research needs to be done before multivariate analysis can fulfill the promise it holds for physical anthropologists."
>
> (Kowalski 1972:119)

In view of Kowalski's comments it is necessary to evaluate how far the multivariate approach is successful in simplifying and describing patterns of differentiation between human populations, and in making the interpretation of results more objective. This issue is best approached by examining some of the disadvantages of using multivariate statistics.

### 3.1.1 The ability to interpret results from multivariate analysis in the light of sampling and other methodological inadequacies.

In many human biological surveys it is logistically difficult to procure large sample sizes with detailed demographic information and complete data records. Moreover, there is sometimes error in the measurement of attributes, especially in the field of dermatoglyphics which is beset with ambiguities in procedure for data measurement and with complex rules for the definition of patterns and the counting of ridges, which are difficult to stringently apply in all circumstances. The investigator of human biological differentiation can only ever hope to minimise sampling and methodological errors, never to eliminate them altogether.

In view of this it is realistic during analysis to expect approximate solutions, rather than precise and optimimum ones. Unfortunately multivariate statistics are sophisticated tools that provide extremely precise solutions. Of course, this is also a characteristic of univariate statistics too. For example, means can be measured to an accuracy of many decimal places. What differentiates the two types of statistics is their comparative ability to take into account the approximate nature of the subject matter.

To begin with univariate statistics have highly developed and well-tested methods of measuring standard errors and assessing significance. In contrast the development of such measures is still at a pioneering stage for many multivariate statistics, particularly for the majority of distance measures that are so important for physical anthropologists. Even where they are developed to a satisfactory extent such as in discriminant function analysis, multiple regression and Mahalonobis's $D^2$, they are often not implemented in computer packages, or implemented in a format not quite suitable for the specific requirements of the problem in hand.

Secondly, and perhaps more serious, is the fact that whereas univariate statistics are well known to be "robust"(Blalock 1960), many multivariate procedures, especially the more sophisticated varieties - such as discriminant function analysis, operate under strict assumptions where insufficient is known about the consequence of failing to meet them (see Kowalski 1972:127). The real value of solutions derived from the application of such statistics can never be precisely assessed, as we are not sure

of (a) the extent of the failure to meet the assumptions and (b) the extent of distortion reflected in the solution.

### 3.1.2 Choice and objectivity in the use of multivariate statistics.

One of the apparent advantages of multivariate analysis is the great choice of techniques available to measure and analyse variation. Theoretically there is a technique to cater for the precise needs of almost every particular aspect of analysis. In practice, however, the availability of a wide choice of techniques is a distinct disadvantage, as there is no universally recognised method of deciding which of a large range of possible solutions is the most correct.

The range of possible solutions is large in whatever multivariate approach is adopted. If a factor or principal components analysis is pursued, there is the initial problem of deciding which of a wide choice of basic factoring techniques is most appropriate. There is also the problem of deciding whether to rotate the initial solution, and if so, what method of rotation to use. If discriminant function analysis is being adopted, there is also a wide choice of methods (see Nie et al. 1975, chapters 24 and 23 for a detailed listing of choices in factor and discriminant analysis). In the analysis of statistical distance between populations, considerably different detailed solutions can be obtained depending on which of many measures of similarity or dissimilarity is used, whether or not a clustering of ordination strategy is used to analyse the distance matrix, which particular clustering or ordination

technique is employed and, in the case of ordination, which and how many dimensions are considered to graphically display population points in space. Finally, compounding the whole area of choice, is the question of which variables to include in an analysis. A comprehensive multivariate analysis repeated for different combinations of variables could produce a staggering number of different solutions.

There are two ways of overcoming the problem of choice. The first is to clearly focus the basic aims of the investigation and the second is to decide from close familiarity with, and experience of the mathematical nature of the techniques and the inherent qualities and flaws in the data, which technique or group of techniques is the most appropriate. Unfortunately people trained in the biological sciences seldom have the necessary mathematical background to objectively evaluate the technical merits of multivariate techniques and, even if they do, there is no multivariate technique without some disadvantage. To take full advantage of multivariate analysis it is essential to choose techniques where the disadvantage in one is offset by advantages in another. Even then we are still often left with the task of deciding which one of several solutions is the best.

Unless very carefully handled, therefore, the multivariate analysis of human biological differentiation can produce so much descriptive information - a multitude of variations on a theme - that it can become even more lengthy and laborious to report than univariate findings. Moreover, the interpretation of a whole series of possible solutions can rapidly degenerate to a level of subjectivity transcending even that normally associated with the evaluation of univariate results.

3.1.3 Multivariate analysis simplifies patterns of differentiation

Multivariate techniques have the apparent property of being able to clarify and simplify the complex web of inter-relationships between operational taxonomic units (OTUs) that result from univariate analysis, and this is one of the main reasons why they are being increasingly employed in physical anthropology and human biology. Clifford and Stephenson, however, make the important point that what such techniques really do is not so much simplify relationships, but condense them:

> "It is very tempting to assume that just because ordination reduces the space in which the data are to be presented, the system has been simplified. This is far from being true, for while the space has been reduced, the complexity of the axes has been increased. It is as if one has moved the problem from one of the understanding of a complex space defined by many simple variables, to a simple space defined by several complex variables."
>
> (Clifford and Stephenson, 1975:389).

Although in this instance the above authors were mainly concerned with ordination, their point is equally valid for other types of multivariate procedures, particularly factor and discriminant function analysis.

If the aims of physical anthropologists when analysing human biological differentiation were confined purely to the objective description of overall inter-relationships between populations, the process of condensing simple variables to fewer complex ones would be just a by-product leading to the desired end result. Inter-population relationships, however, need to be interpreted and explained in terms of evolutionary processes that have caused the observed differentiation. Such interpretation would be difficult without some understanding of the characteristics

of multivariate condensed space, and this understanding can only be achieved by exploring the relationship between multivariate variation and the variation of the univariate attributes on which the multivariate analysis is based. Consequently a univariate analysis is an essential prelude to the types of multivariate analysis engaged in by physical anthropologists.

### 3.1.4 Towards a new approach to the description of human biological differentiation.

I conclude from the preceding discussion that neither the conventional univariate approach nor the multivariate approach that is increasingly supplanting it, is without fundamental disadvantages when applied to describing and analysing survey data of human differentiation. Both approaches, as I have tried to show, can produce pages of laborious description and the interpretation is not necessarily made more subjective by the application of more sophisticated techniques. What then is wrong?

A clue to where the basic trouble lies is provided by Howells, when he bemoans the fact that in physical anthropology

> "we gather data first and ask questions afterwards, an old anthropological habit learned from justly revered masters, who adopted it because of the rush of ethnic and ethnographic change. Our publications are apt to consist of a report of the data, followed by a discussion section in which evolutionary factors are browsed upon and ruminated, but are not or cannot be submitted to test"
>
> (Howells 1966:531)

The implications from Howell's statements is that there must be some preconceived direction which can produce meaningful results, both at theearly design stages of a survey and in the analysis of the data that emerges. A similar implication is also

inherent in Kowalski's concluding remarks concerning which of the two approaches, univariate or multivariate, is most appropriate for physical anthropological analysis:

> "in the final analysis the decision of which approach to use in a given situation rests with the individual investigator. In my view there is a case to be made for using BOTH univariate and multivariate analyses ... We should never lose sight of the fact that statistics is a tool, and it seems wise to choose the particular tools or tool which are best suited to the job in hand. Nobody questions the use of multivariate techniques when they add something to the results of simpler analyses, but everybody should question their use when they impede communication among anthropologists and focus attention on mathematical artefacts instead of biological truths."
>
> (Kowalski, 1972:128)

The debate, therefore, boils down to an old and difficult problem regarding survey analysis. If the investigator follows the survey approach, he is concerned with a purpose of discovery. He is there to find out what is previously unknown. In this approach it is the "facts", the "data" that are of central importance, and the principal task is how to describe the "facts". Once the data have been described, advocates of this approach would say that the results themselves will suggest promising avenues of investigation, and to embark on a survey with pre-conceived notions of what is to be discovered is not only narrow-minded, but is likely to inhibit and prejudice the nature and explanation of novel discoveries. Explanation, therefore, must follow after the facts have been presented.

In contrast to this is the scientific practice of clearly formulating the nature of the problem before the study is designed. In this approach the collection of the data is totally conditional on the requirements of testing the pre-formulated

question or questions. Scientists advocating this line of thinking would argue that description of data is in itself unfruitful, unless firmly directed towards some testable aim. Even if hypotheses can be applied following the process of description, it is unlikely that data gathered with no specific purpose will be coherent or complete enough to test such ad hoc hypotheses.

As I have shown, authorities such as Howells and Kowalski are advocating a shift of direction from the open "survey" approach to the analysis of human biological differentiation, to one determined more by pre-formulated biological questions. Nevertheless, I do not think that they would be prepared to go so far as to espouse the inherent rigidities of the "hypothesis" approach. In effect, the real issue is to find a middle ground where an investigator ceases to enter a survey with only the most general expectations of what he ought to find but one where, at the same time, the directions and aims of the survey are not so rigid as to prejudice or ignore the importance of unexpected and new discoveries (in that they may not be strictly relevant to the investigation of the pre-determined questions under test).

Justifying and reconciling these apparently opposed approaches can only be achieved by re-examining one of the most basic questions in physical anthropology and human biology. What in fact is the purpose of studying human biological differentiation? In this regard few human biologists would disagree with Jardine when he states that:

> "One of the main goals of the study of physical diversity, both in human local populations and of other species is the discovery of ways in which the interaction of historical, ecological and geographical factors have determined the observed patterns of differentiation."
>
> (Jardine 1971:1)

In pursuing these goals, human biologists have tended to pursue three kinds of empirical approach:

A) The "demonstration" approach in which the main aim is to demonstrate the operation of fundamental genetic and evolutionary processes acting on human populations, by asking such questions as: "Does genetic drift operate on human beings? If so, where and how can I test and demonstrate it?" Examples of this type of study are those of Roberts on the operation of drift and the "bottle neck" effect in Tristan da Cunha (Roberts 1967), or Siniscalco et al. (1966), testing the malarial hypothesis in Sardinia. This type of empirical studies are firmly in the traditions of the "hypothesis" scientific practice of testing pre-determined questions.

B) The second type is the detailed regional approach, in which the aim is to find out the extent of biological differentiation amongst the populations of a specifically defined geographical region, and to try and explain how the observed variation has occurred. Examples of this type are Hiernaux's investigation of human differentiation in sub-Saharan Africa, and the British Isle studies of Sunderland and his colleagues in Durham. Such studies are in the tradition of the "survey" approach, characterised by an emphasis on data collection and description.

C) The last type,the least systematic, is the "hopefully useful" data collection approach, in which data is collected in the hope that it might prove useful some day to some other investigator. The collection of data in this way can be useful for people investigating issues in a comparative manner. For

example the studies on intermixture in the Americas (see Morris 1972, Chapter vi) rely significantly on comparative West African material, often collected by people with the above aim in mind.

The growing disenchantment with the data gathering and descriptive approach in physical anthropology can consequently be viewed from two points of view. Firstly, we could complain that there has not been enough emphasis on seriously tackling the challenges of demonstrating fundamental biological tenets (not enough studies of the "demonstration" empirical type). Secondly, we could postuate that the conventional way of tackling regional studies has been inadequate and approached in too vague a manner. It is not clear from the dissatisfaction of Howells and Kowalski which of these two areas of criticism they have in mind. For the purpose of the present study however, it is the second that is of overriding importance.

When embarking on a regional biological study, the investigator usually knows nothing about the differentiation for those attributes he is going to measure and collect. It is thus natural for him or her to regard the survey as discovery, and to orientate the analysis and report around the description of the data collected. Where a fundamental mistake can be made, however, is to fail to examine the assumptions which underly the collection and analysis of the data. In actual fact, an investigator undertaking a human biological survey assumes a great deal about the nature of the variation to be discovered in the area of interest even before the study is begun. This prior knowledge is the heritage of the theoretical and empirical experience of the

discipline of human biology. Consequently, before the study commences the investigator knows that (a) there will be biological variation in the region, that no two individuals and no two populations will be identical in their genetic and phenetic constitution; (b) that the displayed variation can be explained in terms of a finite number of factors, evolutionary and genetic as well as environmental and methodological; and (c) that the people in the region are not spatially and socially independent units, but tend to cluster into aggregates that condition the nature of mating patterns and form the basis of biological populations that can be sampled and analysed. What the investigator does not know is the precise nature of the variation, the precise identity and combination of factors that explain the variation, and the precise composition and biological characteristics of the biological populations.

Without analysing the underlying framework of this assumed knowledge, the investigator is in danger of succumbing to vague and over-detailed description. In answering the question "What is the nature of the observed variation,", there is a risk of interpreting the variation in terms of very detailed inter-sample relationships - for example for trait A, Unit X is closer to Unit Y than Unit Z, or in a multivariate context, "overall" Unit Y is closer to Unit X than is Unit Z. There is an underlying lack of purpose. Describing detailed relationships in this way is like trying to discover the shape of a house just by counting and relating individual bricks.

By analysing the underlying assumptions of the excercise, however, description ceases to lack purpose. In analysing the nature

of variation we are no longer thinking in terms of detailed inter-relationships between sampled populations in the study area, but in terms of expectations conditioned by previous experience in the subject. For example in studying a region in Africa we could say that previous biological studies of "tribal societies" in Africa show that variation is of a type that does not show clustering of populations, and that ecological factors are important explananda for the variation, whereas inbreeding and drift are not. Is this also true of our present study area. In other words the emphasis changed from looking at variation on a series of described inter-population relationships, to one of seeing how far observed PATTERNS of variation fit expected models and TYPES of variation as formulated from previous studies of similar areas. The emphasis has switched from the study of elements to a study of structure.

## 3.2 Types of human biological differentiation

### 3.2.1 Major types of biological differentiation

The biological differentiation exhibited by populations (either within or between species) in any region is always unique. Nevertheless, the experience of numerical taxonomists in zoology and botany has shown that patterns can be discerned in all biological differentiation, and that these patterns conform to a limited number of definable types. The identification of which type or types of differentiation apply to an predominate in the observed differentiation of any study region not only helps to describe the differentiation, but also provides powerful clues as to the factors accounting for the observed differentiation, and helps pave the

way for a more intensive investigation of specific causative agents.

There are basically two main categories of biological differentiation that is distinguished by taxonomists. The first category is when patterns of relationship between populations produce highly distinctive clusters or marked gradients between cluster groups. The relationship between population clusters, moreover, can either by phylogenetic and hierarchical, with the closest relationships being between populations which have only recently diverged in time and space from a common ancestor; or clustered but non-hierarchical where close resemblance is determined by the action of similar selective forces on the gene pools. The second category is where patterns of variation are neither clustered, clinal or hierarchical in structure. Instead, the relationships between populations resembles a network or web of quasi-random links between populations. This second categorial type is called "mosaic evolution" by Jardine and Sibson, and is defined by them as:

> "a pattern of evolution in which rates of divergence of populations with respect to different attributes or complexes of attributes differ widely"
>
> (Jardine and Sibson 1971:272)

Jardine and Sibson associate the first category of variation types with the differentiation found when species or larger scale taxonomic groupings (genera, families etc.) are compared. The second category of types is usually to be observed when populations within species are biologically investigated and contrasted (Jardine and Sibson 1971:272). It is

important to realise, however, that the association of discrete clusters and marked clines with inter-specific or supra-specific evolution, and of "mosaic" differentiation with intra-specific evolution, refers to patterns of differentiation between whole gene pools, not between specific attributes or groups of attributes. Marked clusters and clines for example, can often be detected between intra-specific populations for specific attributes. These clines and clusters, however, will tend to disappear once multivariate comparison is undertaken based on a wide range of phenetic and genetic characters.

The reason for the absence of multivariate clusters and clines in intra-specific variation is due to the fact that differentiation between intra-specific populations is usually characterised by "discordant variation", defined by Jardine and Sibson as

> "a pattern of variation in which the relative extents which populations are differentiated do not coincide for different sets of attributes"
>
> (Jardine and Sibson 1971:268)

Where discordant variation pertains the relative similarity between populations and the ranking of inter-population distances will change dramatically from one attribute to another. The combining of different attributes through multivariate analysis will thus result in an averaging out of inter-population distances on particular attributes, and in a complex mosaic-like configuration of population points in multidimensional space (Hiernaux 1972:106).

Where concordant variation is in evidence, however, the relative rankings of inter-population distances will be preserved from one attribute to another. Similar and dissimilar populations

on one attribute will also be respectively similar or dissimilar on a majority of other attributes. This will naturally result in discrete clusters of closely related populations, with clusters inter-related at various levels of affinity.

Between complete discordance and complete concordance there exists a multiplicity of possible situations. For instance, where concordance is appreciable but far from complete it is possible to get what Jardine (1971) calls "overlapping clusters". Here the "mosaic" of population points contains denser concentrations of points which are not completely distinguished from other concentrations. In other words there is overlap between the extreme populations of related clusters.

The analysis of relative discordance, therefore, is a central aspect of any investigation of biological differentiation. It can tell the investigator in the first instance whether the variation contains discrete clusters, or is mosaic in structure. This in turn narrows down appreciably the specific types of variation he should investigate to shed some light on the probable evolutionary causes of the observed differentiation. These possibilities are discussed in detail below.

3.2.2.1 Sub-types of concordant variation

In the above discussion I have equated concordant variation with a clustered structure. This, however, is not always so. Concordant variation will produce discrete clusters when differentiation is considerable and natural "breaks" occur in the continuum of the ordering of the differentiation. Where the ordering is a gentle continuum, the variation will be clinal

rather than clustered. Finally there is a further type of concordant variation which actually shows up as discordant. This is when the samples are differentiated so little that the variation is almost certainly due to chance; in other words where homogeneity is observed.

The validity of these statments can be shown by analysing the content of Table 3.1, where the three types of situation are illustrated in a very basic way. In section (a) the nine populations show complete concordance for three attributes. The breaks in the continuum allow the discernment of three distinct clusters. In section (b) the nine populations also show complete concordance, but there are no natural breaks. The variation is thus clinal in nature. In section (c) the three attributes show considerable discordance. The magnitude of the differences, however, are so small as to indicate homogeneity.

In studying human intra-specific biological differentiation however, the likelihood of finding complete concordance is remote. It is possible, nevertheless, to find high concordance in the observed biological variation of some regions. Where this is shown, we would expect the variation to resemble in general terms one of the above sub-types, or to show characteristics of a combination of the three. The demonstration of a high degree of concordance, moreover, gives the investigator an advantage in that a situation of high observed concordance is very unlikely if the underlying causation of the variation is not biological. In other words, if sampling and methodological factors are strong in

determining the observed variation, the resulting structure will tend to be discordant. The chances that random error will produce the same general ordering of populations for a range of attributes is infinitessimal. Consequently, if strong concordance in the form of a clustered or clinal structure is demonstrated, the investigator is almost toally safe in assuming that the observed structure makes biological sense. The next step, therefore, is to conceive of the type of situations which might produce clines, clusters and rigid homogeneity.

3.2.2.1.1 Homogeneity

Homogeneity occurs when the samples drawn from a region display no statistically significant variation given reasonable sample sizes. There are two possible explanations for biological homogeneity; firstly that the samples all represent parts of a single biological population, or secondly that they represent several biological populations which (a) have only recently dispersed and have had little opportunity to significantly differentiate or (b) which have co-existed in close proximity for such a long time that admixture/gene flow has effectively obliterated any significant variation. Homogeneity is obviously one of the first types of variation to be tested.

3.2.2.1.2 Clustering

The most important characteristic of clustering is that the biological differences for populations between clusters will be much greater than those within clusters. Concordance will thus

be strong in the ordering of relationships in different traits between populations or samples drawn from different clusters, but it is not necessary for this type of concordance to pertain between populations drawn from the same cluster. Within clusters the dominant type of concordance can be homogeneity, which, as I have demonstrated, can manifest itself as a small differences of random causation - discordant in character. The measurement of concordance in a clustered structure is thus often both a qualitative and quantitative excercise, in which areas of homogeneous discordance (in reality overall concordance) are contrasted with strong "ordered" concordance between such areas. The types of situations that best fit this type of structure are when the variation is either topotypic (geographical) ecotypic or "ethnotypic" in character.

3.2.2.1.2.1 Topotypic or geographical variation

This type of variation is characterised by relative homogeneity within populations in geographical regions, and significant variation between such regions. The dominant evolutionary model which explains geographical variation is where any significant variation within regions is due to random differentiation through the slow dispersal of local populations over time. The further two populations are from each other, geographically, the greater is the time interval between their dispersal from a point of common origin, and thus the greater is the biological difference between them.

The clustering of populations when topotypic variation obtains should approximate closely in pattern to the true geographical locations of the sampled populations.

The demonstration of topotypic variation depends to a large extent on being able to show that the action of natural selection in differentiating populations is minimal or uniform when compared with the effects of random drift. Geographical variation can hence be expected to predominate when the nature of the study area is relatively homogeneous regarding climate, topography, ecology and patterns of settlements. The researcher must also be able to demonstrate that patterns of inter-population contact leading to gene flow are relatively stable and conform in the main to a pattern of contact due to geographical proximity, rather than to unstable patterns of contact such as sudden migration of groups into geographically distant areas where rapid assimilation takes place. The occurrence of such discontinuous migration occurs in most parts of the world over time. The geographical model, however, can accommodate such events as long as they occur on a small scale and are relatively rare. Finally, in order for geographical clustering to occur, patterns of inter-mixture must be irregular; that is to say, there must be genetic barriers to mating to break up the continuum of overlapping regional micro-populations. Such genetic barriers do not necessarily mean a total absence of inter-marriage, but imply a significant reduction of the incidence of such marriages.

The empirical verification of the existence of geographical variation in the human species is perhaps most advanced in the British Isles. Kopéc's analysis of the distribution of

the ABO blood system is a classic exposition of the principle, although the limited range of attributes considered rendered the exercise inconclusive (Kopéc 1973 in Roberts and Sunderland 1973). Kopec's conclusions, however, were confirmed by a more exhaustive multivariate analysis by Sunderland and Cartwright (1975), based on a much wider range of traits.

3.2.2.1.2.2 Ecotypic variation

In ecotypic variation the biological differences between populations within a particular ecological zone are relatively similar compared with those between populations drawn from different ecological zones. If significant barriers to mating exist between such zones, a clustered structure will pertain.

The overall pattern of variation exhibited by ecotypic variation is very similar to that manifest by topotypic variation. The major difference, however, lies in the biological model explaining it. In ecotypic variation it is not random drift that causes the differentiation, but natural selection as determined by factors or forces applying relatively constantly within each zone, and differentially between zones. As in topotypic variation, moreover, intermixture between populations should be constant and stable. Discontinuous migration would greatly reduce the discreteness of clusters.

Ecotypic variation has never been adequately demonstrated in the human species for a large range of concordant attributes. (At the univariate level, however, it is relatively common) Hiernaux (1972;111), however, has gone some way to remedying this

state of affairs by noting that in sub-Saharan Africa the mean biological distances (based on a large range of attributes) are higher between habitats than within habitats.

Finally, we should note that ecotypic variation, unlike topotypic differentiation, produces clusters which are not easily related in a phylogenetic sense. This is because differentiation is not linked to the concept of progressive divergence in time and space from a common dispersal point. Where ecotypic differentiation is in evidence, relative similarity between clusters is more likely to reflect similar trends in adaptation than common ancestry.

3.2.2.1.2.2 Ethnotypic variation

The word "ethnotypic" has been adopted here as a label for the type of variation that is implied when physical anthropologists talk of affinity between linguistic, socio-cultural, caste or tribal populations.

Where strong ethnotypic variation occurs, the pattern of differentiation is one where distances between populations within an ethnic division will differ less than distances between divisions. The general structure will be concordant, clustered and often hierarchical, and will tend to mirror linguistic and socio-cultural patterns of relationship, patterns which, as I demonstrated in chapter 1 in the case of the Kenyan ethnic situation, are also strongly hierarchical. Thus, for instance, if we could demonstrate a high correlation between linguistic and biological affinity in a particular area, this would be strong evidence for the existence of ethnotypic variation.

The biological model explaining this type of variation is almost identical to that described for topotypic variation, with a strong emphasis on processes of gradual dispersal and drift. The chief difference is that patterns of mating and genetic barriers are not geographical but socio-culturally defined and maintained. Hence it is possible to have patterns of ethnotypic variation showing little correlation with geography. This is particularly true in the case of cities where endogamous ethnic communities residentially intermingle but do not intermarry to any great extent.

The main difficulty with demonstrating ethnotypic variation (in addition to those already described for the drift model in general in the discussion of topotypic variation) in many areas of the world is that "ethnic units" are also geographical and ecological units too. For example the Dinka of the Sudan occupy a consolidated and continuous territory which is also a highly characteristic and unique habitat. This situation holds in many areas of the world. Considerable care is needed, therefore, correctly to distinguish ethnotypic variation from the other two types.

3.2.2.1.3 Clinal variation

Clinal variation is defined as

> "a pattern of intra-specific variation in which there is a trend in one or more attributes correlated with some external factor such as geographical location (topoclinal variation), soil-type, altitude or climate (ecoclinal variation)".
>
> (Jardine and Sibson 1971:266)

Where strong concordance between attributes exist, clinal variation can exist in two contexts both in the topological and ecological

forms of variation. The first is when there are no significant barriers to mating between regions or ecotypes, and the cline is seen as a gradual gradation in biological characteristics following (a) an ecological gradient or (b) emanating from a central point of dispersal or (c) constituting the beginning of a radiating chain of overlapping gene flow within a region. The second is when barriers to mating are incomplete between the boundaries of regions or ecotypes and is caused by overlapping gene flow over the boundary. In the former context there will be no significant clustering, but rather a series of gradations or chains. In the latter context clustering will exist, but the boundaries between clusters will be indistinct owing to the existence of boundary populations that could fit into either neighbouring cluster. Where clinal variation is significant, the search for phylogenetic origins become very difficult.

3.2.2.1.4

There exists a fourth sub-type of concordant variation where the pattern of differentiation neither correlates with geography nor with ecology. This is known as "reticulate variation" which is defined by Jardine and Sibson as

> "a pattern of differentiation in which populations show significant differences which cannot be represented adequately by a hierarchic classification and in which clinal variation is not involved"
>
> (Jardine and Sibson 1971:273).

Although this type of variation has been recognised in botanic species (see Jardine and Sibson, 1971;147), the chances that this type of variation exists in the human species alongside heavy concordance is remote. For this type of variation to exist in a

human situation, the local populations would have to be simultaneously very isolated or endogamous, and prone to considerable discontinuous migration without intermixing significantly in the process. Very few parts of the world would produce the conditions to make this likely.

3.2.2.2 Sub-types of discordant variation

The chief characteristic of discordant variation is the failure of sampled populations to vary in the same way from attribute to attribute. The best known example of discordance, for instance, is the difference in patterns of differentiation manifested by

human populations between anthropometric and some serological traits, (e.g. the sickle-cell trait).

The sub-types of variation discussed earlier in the section on concordance variation apply with equal validity to discordant variation. The crucial difference is that while in the case of concordant variation the overall multivariate genetic structure will tend to conform to one or other of the main types, or to subtle combinations of them, in the case of discordant variation it is only individual attributes that will be affected. Thus, for example, where discordant variation is demonstrated, it could be found that some attributes will show homogeneity, while others might differ appreciably; some might show strong associations with climate while others vary according to quite different ecological criteria; some might vary in a manner that suggests a topocline while others show a random patterning of relationships, and so on.

This characteristic web of cross-cutting types and patterns of variation makes the analysis and interpretation of discordant variation particularly problematic. It is true that much can be learned from studying the type of variation given by individual attributes in turn. A string of such findings, however, will do little to resolve the fundamental problem of interpreting the observed integration of these discordant elements and the particularised explanations accompanying them. Experienced investigators of discordant variation such as Hiernaux and Jardine have tended to try and resolve the problem of integrating structure in terms of improving statistical techniques which not only measure and display overall structure in a way that minimises the considerable distortion inherent when differences are discordant, but which permit an improvement in interpretation from the point of view of biological meaning (see e.g. Jardine 1971). These attempts, however, though valuable, do not compensate for what is really missing, viz. a series of general theoretical models that might explain the operation of different types of discordant variation; models that could provide a standard by which to interpret empirical observations.

So far I have been treating discordant variation as purely a biological problem: that is to say I have presumed that discordant variation is true biological and genetic variation. In a real situation, however, it is observed variation that is being examined, and the process of analysis and observation is beset with sampling and methodological errors. One of the most significant potential explanations for observed discordance is precisely these types of errors, where true patterns of relationships

are hidden by random fluctuation of means and frequencies due to inadequate sample sizes, and distorted by bias due to inconsistent data measurement. It is thus possible that observed discordance could be no more than an error-ridden representation of underlying biological concordance, or alternatively a distorted picture of true biological discordance. The assessment of the extent of sampling and methodological error is thus a crucial aspect of analysing the nature of discordant variation.

## 3.3 Principles of the statistical structure of human biological variation and the analysis and measurement of biological discordance and concordance

I conclude from the preceding discussion that one of the most relevant preliminary tasks in the analysis and description of the nature and structure of human biological differentiation is the assessment of the extent and nature of biological discordance. A successful demonstration of the characteristics of discordance or concordance can considerably narrow down the field of the types of variation that could explain the observed data, and enable the researcher to focus attention on those types of variation that are most likely to be operating at the discerned level of discordance or concordance. In the following sub-section I shall be assessing methods which could be applied to the analysis of discordance.

In pursuing this aim, however, I am still some way from discussing the relative merits of specific statistical techniques. Before embarking on this course it is essential to form some understanding of the nature of the statistical structure of biological variation because in applying statistical techniques we

are directly analysing a mathematical abstraction of supposed biological reality, and a major task in interpreting statistical results is the translation of mathematical results into meaningful biological inferences. The evaluation of the relationship between statistical structure of variation and biological structure is consequently a critical prelude to the assessment of the appropriateness of specific statistical techniques.

8.3.1 The statistical structure of human biological variation

On the completion of sampling and measurement of attributes the data consist of a series of records of measured attributes for individual members of samples drawn from local biological populations. In terms of statistical structure this can be conceived as a rectangular matrix of attribute scores in which the individuals form the rows and the attributes scores the columns. The affiliation of individuals to their respective samples is expressed either in terms of an additional column vector where each individual is assigned a number according to a nominal scale identifying the sampled populations, or as "subfiles" or defined groupings of the matrix rows where each row is an individual case and each group a population sample. This type of matrix I will call a first order or type A data matrix.

This matrix can be considerably condensed to form a second order (type B) summary data matrix, where each row represents a sampled population and each column represents the summarised statistical score for each attribute state (usually the mean for quantitative traits and frequency for qualitative traits, but for specialised analyses scores can also be standard deviations, medians,

standard scores and so on). The classic univariate summary tables in human biological reports can be viewed as printed and labelled versions of this type of second order matrix.

3.3.1.1 Structural relationships

These two types of basic matrix form the foundation of all advanced analysis of human structural variation in human biology. The individual and population attribute scores for each respective type of matrix represent the structural elements and are the smallest divisible units at each respective level of resolution. The means of discerning structural trends from the consideration of these "elements" lies in analysing the way they relate to each other.

The analysis of structural relationships in human biology has tended to take two paths. The first is the "correlation" strategy in which the common variance between two attribute variables is precisely quantified and subjected to further structural analysis through some sort of factor or principle component analysis. This approach has tended to be mainly applied to type A matrices to evaluate relationships of attributes in individuals. (For example Coope 1971 and Dennis 1977[a] adopted this approach to try and identify genetic components of variation for dermatoglyphic attributes, and Howells 1973 used a similar approach to delineate functional relationships for craniometric measures.

The second path is the "distance" strategy in which the relative resemblance between units is precisely measured through

some form of appropriate distance statistic based on the common summation and averaging of all attribute differences between two units. This is the common strategy applied by human biologists and numerical taxonomists to the analysis of type B matrices where the populations form the row units. It is, however, statistically possible to apply a correlation strategy to type B matrices too, and conversely to apply a distance strategy to type A matrices, although this is not often desirable.

Of the two types of approaches that relate attribute states, I regard the "correlation" type as the most important. This is because it is possible to analyse correlations without having to take any "distances" into account, but it is extremely unwise, as most textbooks describing distance techniques point out, to measure distances between units without exploring the characteristics and implications of the correlations between the attributes that go into the calculation of the distances. It is thus important to discuss correlations between biological attributes before evaluating the broader implications of the "correlation and distance" strategies.

### 3.3.2 Types of correlations between biological attributes

Jardine and Sibson distinguish "at least five distinct senses in which taxonomists have written of 'correlation' of attributes and characters" (Jardine and Sibson 1971:25). These are discussed with particular emphasis to correlations between dermatoglyphic attributes (Jardine and Sibson 1971, Chapter 4:3), and are listed as follows:

a) "Logical correlations" called by the above authors "the phenomenon of conditional definition". This type of correlation is a common feature in dermatoglyphics, especially when patterns and triradii are defined and related according to the topological principles of Penrose (see Chapter 4). The most extreme dermatoglyphic correlations of this type occur between patterns and triradii which are absolutely conditional to each other's presence. Thus, for example, a thenar pattern I or $I^r$ has to be topologically accompanied by a thenar triradius e/f. The correlation between the two features is thus, barring error, unity. Similarly only in the rare event of a parathenar loop being present, or the occurrence of a hypothenar radial arch, is a hypothenar border triradius $t^b$ not associated with a hypothenar central loop $\overset{\wedge}{H}$. In practice, the correlation between these two latter features is extremely high. Again there is a total correlation between the occurrence of a true arch on a digit, and a zero ridge count. The axial triradius t is also highly negatively correlated with t', as one usually replaces the other, except in the relatively infrequent occasions when both are present (where a hypothenar peripheral loop H occurs) or when t" is present. There are also many other features that are partially conditionally related through the topological scheme - for instance palmar ridge counts and interdigital triradii. The issue of how far such high logical correlations imply the same genetic feature has not yet been resolved in dermatoglyphics.

b) "Functional" Correlations of attributes; In the words of Jardine and Sibson,

> "this kind of correlation is that which attributes have when their states can describe parts of organisms which are jointly involved in the performance of the same function. Thus the attributes SIZE OF CANINE TEETH, SIZE OF MASTOID PROCESS, SIZE OF STERNAL PROCESS were functionally correlated in the sabre-toothed tigers, in which the powerful sternomastoid muscles were used to wield the head as a whole in striking prey with the massive canine teeth. Functional correlations may have some predictive power, since given a valid inference about function from some attribute states it may be possible to predict the states of other attributes".
>
> (Jardine and Sibson, 1971:26)

The evaluation of functional relationships are clearly critical in the analysis of anthropometric and morphological attributes. In the field of dermatoglyphics, however, the functional relationships of dermatoglyphic attributes are not known. Indeed it is difficult even to conceive of any possible function for dermatoglyphic configurations, in contrast to the dermal ridges themselves which are known to be related to sensitivity of touch and sweating processes. It is possible, however, that the "logical correlations" listed earlier might also be functionally related, but this has yet to be proved. From the state of present knowledge, therefore, we can effectively discount the importance of functional correlations among dermatoglyphic attributes, except in the case of the atd angle, which is clearly related to hand length and breadth, and hence to age (see Chapter 4).

c) Statistical correlation of attributes within populations: This type of correlation is to a large extent self explanatory, referring to the calculation of correlations between quantatitive attribute states in individuals (such as Pearson's r, Spearman's Rho, Kendall's Tau). The separate consideration of "statistical correlations" by Jardine and Sibson, as a type different from logical

and functional correlations, is puzzling. I would see "statistical correlations" rather as a generic type encompassing the sub-types "logical" and "functional" correlations. This is, in fact, how Jardine did regard them in a contemporary publication (Jardine 1971:7).

If the type "statistical correlation" is accepted as generic, the issue then becomes a question of whether other types exist of correlation of attributes within populations that are not logical or functional. In this regard it is possible to distinguish "genetic" correlations, where high correlations between attributes implies that the attributes are determined by the same set of genes. The decipherment of underlying genetic trends through the analysis of correlations between dermatoglyphic attributes in individuals has been undertaken in some detail by Coope (1971) and Dennis (1977a). The dangers of this approach, however, is that it is difficult to be sure that observed inter-attribute correlations really do reflect a common genetic base in reality, and secondly there is an added complication that the observed patterns of statistical correlations in one population might differ appreciably when other populations are considered. The available evidence, however, (see Coope 1971) suggests that statistical correlations between dermatoglyphic attributes in individuals are remarkably uniform in different world populations.

d) Taxonomic Correlations: Jardine (1971:7) identifies taxonomic correlations as being the other major generic type, to be contrasted with "correlations within populations". A taxonomic correlation is the extent to which attributes relate to each other between populations, or as Jardine points out, "when the relative extents to which populations are discriminated by different attributes are similar" (Jardine 1971:7).

e) Redundancy: Redundancy refers to the phenomenon where attributes are both very highly correlated in individuals and are also highly correlated taxonomically, especially when they are either genetically the same feature, or different scales of measurement for the same feature. (E.g. measuring stature both in inches and centimetres). In dermatoglyphics redundancy can be seen in those attributes that show high "logical" correlations.

### 3.3.3 Assessing relative discordance

The assessment of relative discordance is a progression of the evaluation of the statistical structure of human biological differentiation. From the preceding discussion it is now possible to distinguish two levels at which discordance can be analysed. The first is at the level of attribute relationships within populations, and the second is at the level between populations. The contrasting of patterns of discordance between these two levels of structure are vital pointers to consequent explanatory analysis.

a) Discordance within populations

The analysis of discordance within populations chiefly entails the calculation of statistical correlations between attributes in individuals, and the further analysis of the resultant correlation matrix through some type of R factor or principle components analysis, or through other techniques that can accurately reproduce representations of attribute relationships from a correlation matrix, such as non-metric multidimensional scaling. The work of Coope and Dennis in pioneering this type of analysis for dermatoglyphic attributes has already been discussed, although

much work has still to be done to correctly evaluate the separation of the effects of logical correlation from those of genetic and environmentally determined correlations.

In addition to the assessment of attribute discordance in individuals, there exists another area of potential discordance within populations. This is discordance in patterns of variation for attributes between the sexes in a population. Such discordance can be analysed in two ways. The first is by comparing summary statistical scores for each attribute between individuals pooled into male and female sample groups. (For example by comparing means frequencies between the sexes through probabilistic tests of significance). The second method is to contrast the relative patterns of correlations between attributes within each sex. This entails separating the total population sample into two type A data matrices by sex, and performing an independent analysis of statistical correlations in individuals for each sex division.

The first method above measures the size differences between the sexes for each attribute - the magnitude and significance of differences between means and frequencies, with discordance between attributes - and the second measures the shape of differences - differences in the pattern of structural relationships.

b) Discordance between populations

Multivariate relationships in physical anthropology are usually associated with the measurement of "genetic" or "phenetic" distance which is the interpretation given to statistical distance between population samples calculated from a range of measured biological attributes. It is possible to get many different

estimates of "genetic" distance according to the combination of attributes used as input. The measurement of discordance can consequently be conducted through comparing distance matrices calculated with different combinations of attributes. Dennis, for example (1977[a], Chapter 10) produced several different matrices of dissimilarity for combinations of palmar and digital dermatoglyphic attributes (using Edward's $E^2$ for qualitative traits and Mahalonobis's $D^2$ for quantitative traits) and assessed the extent of discordance through correlations and scattergrams of the distances for each pair of matrices.

A much more sophisticated treatment of discordance, however, is given by Jardine (1971:9-11). He proposed a scale free measure of discordance between dissimilarity coefficients $d_1$ and $d_2$ where:

"

$$\Delta(d_1,d_2) \doteq \Sigma\left|\frac{d_1(a,b)}{\Sigma d_1(a,b)} - \frac{d_2(a,b)}{\Sigma d_2(a,b)}\right|$$

summation being over pairs of elements of the set of populations P."

The measure has a value of zero for perfect concordance and an upper value of two for complete discordance. In applying the method Jardine suggests the calculation of a dissimilarity matrix for every attribute being considered in turn, followed by the application of the above technique between each pair of distance matrices, to produce a symmetrical matrix of discordance values with "$\frac{1}{2}$ n(n-1)" entries. The matrix of discordance values can then be subjected to further investigation through cluster or ordination analysis to discern clusters of taxonomically related

attributes, and to determine between which groups of attributes the maximum discordance lies. The effectiveness of the method is well illustrated by Jardine when he analysed discordance in attributes recorded by Hiernaux (1956) from fifteen Congolese and Ruandan populations (Jardine 1971:10).

Although this method of analysing discordance is effective and mathematically elegant, it is nevertheless cumbersome to put into practise. Firstly there is the calculation of a separate dissimilarity matrix for each attribute. Then the calculation of the discordance matrix, and finally the by no means simple stage of applying cluster and ordination analysis. This process is relatively problem free when the number of attributes and population units is small . Distance matrices, however, possess the unfortunate property of growing exponentially as population units increase, and in addition to this basic problem is obviously compounded as the number of attributes considered increases at the first stage of analysis. The use of Jardine's method thus becomes progressively unwieldy and costly in terms of computational effort the more attributes and population units are included.

The key concept for exploring discordance between populations using a correlation approach is the appreciation of the possibilities of "taxonomic correlations" as defined earlier in this subsection. A matrix of taxonomic correlations is easily calculated by setting up a type B data matrix where populations (rows) have a summary attribute score (means or frequences) for each attribute being analysed, and by correlating the attributes through an appropriate correlation statistic. It should be noted that

attribute vectors of mean and frequency scores are all quantitative measures, and broadly compatible scales of measurement. For example it is quite permissable to correlate mean scores for ridge counts with frequency scores for whorls, as the question being asked is do populations vary in the same way for ridge counts and whorls? Do populations with high ridge count means also have a high frequency of whorls? (and so on). The approach thus gains in flexibility in being able to include attributes that are both quantitative and qualitative in individuals. This is something that is not usually possible in the dissimilarity approach, as distance statistics are usually only designed to accommodate either qualitative or quantitative traits, not both. The few exceptions are statistics that are relatively unsophisticated. (E.g. Hiernaux's statistic, see Hiernaux 1965).

The setting up of a matrix of taxonomic correlations is a prelude for more sophisticated analysis such as R factor or principle components analysis, or ordination techniques such as non metric multi-dimensional scaling which can simplify the correlation matrix and identify structural trends. Clusters of taxonomically related attributes can readily be discerned through the application of such techniques, and the extent of discordance assessed.

Analysing taxonomic correlations has the added advantage that the computational facilities available to undertake all stages of analysis are readily available in computer packages such as SPSS and, are easily accessed and are well tested. Moreover unlike in the Jardine method, the computation of taxonomic correlations becomes increasingly efficient as population units considered increase,

as the correlations are based on more cases. On the other hand the number of attributes that can be included in analyses such as R factor analysis is still limited (SPSS for example, has a limit of one hundred variables - see Nie et.al., 1975:Chapter 24), although the advantage of being able to simultaneously analyse taxonomic correlations for as many as a hundred attributes is a considerable improvement on the dissimilarity strategy.

### 3.3.4 Discordance and Stability

The stability of a matrix of genetic or phenetic distances is achieved when the addition of further attributes in the calculation or the distances does not significantly re-order the relative relationships between the population units. In a situation where high concordance and strong taxonomic correlations are predominant, it takes relatively few attributes to attain stability. Where high discordance exists, on the other hand, stability might theoretically never be achieved until a majority of genetic traits in a series of populations are included. The measurement of stability is thus an empirical problem than can only be resolved by examining the extent of discordance between a wide range of attributes.

The desirability of achieving stability is taken for granted in human biological studies. There is nothing more annoying than to have a carefully constructed taxonomic or phylogentic scheme thrown out of balance by the consideration of new attributes. Discordance, however, when prominent is a fact of life and as Jardine states,

> "there is no justification for the view that dissimilarity between populations is a parameter which can be estimated with increasing accuracy as more attributes are considered. This view would only be correct if the attributes selected in the study of population differentiation formed random samples from the same population of attributes. In practise taxonomists often select first the more easily recorded attributes which are good discriminators and then those which are less good discriminators or are less easily recorded."
>
> (Jardine 1971:11)

Stability therefore, should not be viewed as an essential methodological prelude to multivariate analysis of human biological differentiation. Rather it should be regarded as a property of distance matrices that reflects the extent of concordance or discordance between attributes. A failure to achieve stability is not a methodological tragedy, except to those human taxonomists and phylogeneticists who are anxious to find hierarchical divergence or rigid taxonomic relationships in human populations. An absence of stability is a challenge in that it forces us to consider the implications of discordance and mosaic evolution in human populations.

### 3.3.5 Implications of comparing discordance of attributes within and between populations

The comparison of these two aspects of discordance helps to provide valuable indications of the underlying characteristics of the true biological variation as opposed to statistical variation that can so easily reflect sampling and methodological error. The patterns of correlation of attributes within populations can be cautiously interpreted as reflecting genetic relationships between attributes, with high correlations implying a common genetic causality for the attributes involved. If this is so, one would expect these

same attributes to show high taxonomic correlations too. Failure to do so could be interpreted as showing that either the statistical correlations in individuals are not related significantly genetically or that the relationships of the attributes between the populations is inefficiently expressed (Jardine and Sibson 1971;28), probably due to poor sampling. The existence of high statistical correlations and high taxonomic correlations for the same attribute sets, on the other hand, is a symptom of redundancy in the attributes.

To the investigator of biological differention between populations however, the most interesting implications occur when high discordance exists in attributes at the level of individuals, and strong concordance is seen in attributes between populations. This means that attributes that are not genetically related are still varying in a similar fashion between populations. The way is thus opened for a confident appraisal of types of concordant variation. In contrast if the only high taxonomic correlations are between attributes that are also strongly correlated statistically within populations, a situation of discordant variation is presented for further investigation.

### 3.3.5. Implications of inter-sex discordance

The analysis of inter-sex discordance both within and between populations is an extremely complex excercise as it adds further permutations to an already involved series of comparisons. The interpretation of the analysis is made even more difficult by the lack of systematic research in this field. It is significant that while most researchers differentiate the sexes meticulously

during univariate analysis, they seldom do so at the multivariate level. (Howells 1973 and Dennis 1977a are notable exceptions, although the emphasis in interpreting sex differences in both studies was at the "within" populations level rather than at that between populations.)

At the level of individuals, the statistical correlations between the sexes should be almost identical in pattern if the attributes are strongly genetically determined, unless sex-linked genes predominate. Similarities in general structure ("shape") but marked "size" differences between the sexes implies, from a genetic viewpoint a polygenic similarity between the sexes influenced by some sex-linked component.

Given a similar genetic structure between attributes, and cognisance of the fact that individuals from both sexes share the influence of common evolutionary forces if drawn from the same population and geographical locality, we would expect that the pattern of taxonomic correlations between populations in one sex should be replicated in the other. Failure to do so is an obvious pointer to the overriding influence of sampling and other errors, although it is theoretically possible that it could reflect inconsistent patterns of mating.

The really difficult area, however, occurs when the patterns of discordance within and between populations are very similar when assessed independently in each sex, but where taxonomic correlations of the same attributes between the sexes are low, This indicates that the magnitude of differences of means or frequencies between the sexes within populations varies appreciably

from one population to another, and if these differences are not caused by insufficient sample numbers, a highly complex biological situation is implied.

## 3.4 Overall structure of human biological differentiation

In spite of its quality of clarifying descriptive structural trends in the variation of attributes, the analysis of basic structure is still too detailed. For instance, strong discordance in the variation of a range of attributes could indicate that the pattern of overall inter-population relationships will resemble a network. Analysis of discordance alone, however, will not tell us the precise nature of the network. Similarly, while the demonstration of considerable concordance in a range of attributes increases the chances of finding multivariate clusters as a feature of variation, the analysis of basic structure cannot tell us the precise composition of the clusters, their relative strengths and distinctiveness.

The analysis of overall structure, therefore, is a natural progression from the analysis of basic structure, in which the stress is not the structural relationships of attributes, but of populations. In other words, overall structure can be defined as the way populations structurally relate to each other with respect to their generalised multivariate biological characteristics.

The investigation of multivariate biological relationships between human populations has been a central concern of human biology for some time. I am suggesting here, however, that structural multivariate description should be sharply distinguished

from the role of multivariate biological analysis as an explanatory tool. Overall structural description, moreover, should precede any attempt to explain variation.

The description of overall structural trends entails looking for general characteristic patterns of multivariate relationships between populations. It is not an exercise in looking at specific population inter-relationships (e.g. observing that populations sample A is closer to sample B than either A and B is to C). The analysis of basic structure tells us to some extent what we might expect to find in terms of overall structural patterns. The task of the analysis of overall structure is to try and establish whether these expectations are justified, and if they are, to detail the nature of the structural patterns found.

The first task in the analysis of overall structure is to examine the question of whether the general pattern is one which contains a series of clusters, or whether it is one which contains no discrete clusters, resembling rather a network or mosaic. From the answer to this basic question, it is possible to progress to increasingly detailed issues. For example, if clusters are found, we could go on to discover the composition of each cluster in terms of populations, the relative homogeneity and distinctiveness of each cluster, which specific populations are excluded from any cluster and so on.

Only when structural description if finished, are we really in a sound position to start investigating explanatory issues, such as the search for athnotypic, topotypic or ecotypic types of variation.

## 3.5 The plan for the analysis of dermatoglyphic variation in Kenya.

During this chapter, I have pointed out the problems in investigating regional human biological variation, particularly when there are a large number of population samples and attributes to consider. The conventional practice of dividing the investigation into univariate and multivariate sections is especially unsatisfactory. As an alternative, I have discussed the possibilities of a descriptive structural approach, in which patterns of relationships between attributes form a preliminary base for analysis and patterns of relationships between populations a secondary descriptive phase. Both descriptive stages lead on to the investigation of specific types of variation that may be of explanatory significance. The ideas of Hiernaux (1972), Jardine (1971) and Jardine and Sibson (1971), have been particularly influential in the preceding discussion.

I now turn to the task of outlining specifically how I propose to implement this approach in the analysis of the Kenyan dermatoglyphic data:

1) The analysis begins in Chapter 4, with a methodological discussion concerning the measurement and reading of dermatoglyphic features, and an outline of procedures adopted to minimise errors, both during the "print reading" stage and the setting up of the raw data matrices for computer analysis.

2) Following this necessary methodological chapter, the analysis begins in earnest in Chapter 5, which deals with basic structure. The chapter commences with an outline description of the basic univariate data tables, which are presented for purposes

of reference in Appendix 1. The next stage establishes whether there is sufficient significant differentiation between samples to warrant more sophisticated analysis. Should there be no significant variation, then a highly homogeneous situation would be suggested, and there would be little point in pursuing a detailed structural analysis.

Having established that there is considerable variation, I then explore detailed relationships between taxonomic attributes, through the application of correlation and principal components analysis on type B matrices of taxonomic dermatoglyphic attributes (i.e. matrices where population attribute means or frequencies form the columns, and population samples the rows). This stage of analysis should establish how much discordance exists between variation groups of attributes.

The final part of Chapter 5 seeks to establish how much true taxonomic concordance is actually manifested between attributes. "True taxonomic concordance" refers to the concordance that is left once phenetic and logical correlations in individuals are controlled for. Of particular significance is the extent to which attributes that show low correlations in individuals may show higher taxonomic correlations.

3) The main issue under investigation in Chapter 6 is the nature of overall structure, with a special focus on the question of whether significant population clusters can be discerned. In implementing these aims, I compare (1) the results of applying statistical procedures based on phenetic distances and their further analysis through ordination analysis with (2) those obtained from applying principal components analysis on type B data matrices,

and graphically portraying resulting factor scores. The meaningfulness of existing methods for graphically displaying ordination solutions are also assessed.

4) Chapter 7 concerns differences in patterns of variation between the male and female series that arise from the descriptive analyses in Chapter 5 and 6. Should considerable discordance exist in patterns of variation between the sexes, then an attempt to explain the differences is crucial, as we might theoretically expect considerable concordance in taxonomic trends of variation given that males and females in a population are a product of a common genepool.

5) In the last two chapters (8 and 9), I investigate specific types of variation that might provide indications of the forces that have been or are influencing dermatoglyphic differentiation in Kenya:

In Chapter 8, I examine "ethnotypic variation", or the extent to which dermatoglyphic affinity between samples corresponds with known patterns of ethnic affinity between the populations from which the samples are drawn. In Chapter 9, I examine "ecotypic" variation, or the possibility that some dermatoglyphic attributes may vary systematically with climate and environment.

Although each chapter concentrates on a specific set of issues, there are some questions that run as themes throughout the analysis. In view of the theoretical discussion in this chapter, the following questions are of special interest:

i) Does considerable discordance in the variation of taxonomic attributes necessary result in the absence of distinct clusters? Conversely does concordance mean that strong

multivariate clusters or clines will be a feature of the overall variation?

ii) In a situation of discordance can we really expect to find significant overall patterns of geographical, ecotypic or ethnotypic variation?

iii) Is the examination of the "structure" of biological differentiation as a descriptive prelude to explanatory analysis, any more profitable than the more well tried univariate and multivariate approaches outlined in the first part of this chapter?

# CHAPTER 4

## General Methodology of Dermatoglyphic Data Measurements

### 4.1 Introduction

The pioneering work of Galton (1892, 1895), Henry (1900) and Bonnevie (1924, 1929) was largely responsible for the formulation of methods for the measurement of dermatoglyphic attributes. The original methods of these authors, with some amendments and improvements, are summarised by Cummins and Midlo (1943).

The development of a consistent and detailed methodology for dermatoglyphic attribute measurement was initially inspired by the demands of police forensic work and subsequently also by the requirements of genetic investigation of dermal features. Although the value of dermatoglyphics in population studies was appreciated at an early date however, the specialised requirements of this subject were not reflected in the development of the methodology of attribute measurement. For example there is an emphasis in the early traditional methods on detailed definition of pattern sub-types (e.g. types of whorls) which while essential for genetic investigation, tends to be too refined for population studies where limitations on sample sizes renders the consideration of rarer pattern types problematic. In contrast there was relatively less effort devoted to tightening up ambiguities in the definition of common and more general major features which are of considerable potential value in population studies (for example axial triradii and mainlines).

The majority of physical anthropological studies which consider dermatoglyphic differentiation in human populations have

traditionally used the Cummins and Midlo method, more often than not without any attempt to assess its shortcomings. This is especially true of non-specialists who sometimes supplement fieldwork in other fields with a short dermatoglyphic survey. The main drawbacks of the traditional methods of sampling and collecting prints have been the imprecise definitions of some attributes which have resulted in considerable inter-observer discrepancy (for example for axial triradii). Moreover a lack of a standard method of collecting prints has handicapped the task of data collation and comparison. For example some people have collected finger prints only, or just right or left hands, or in extreme cases just thumbs and index fingers. Other undesirable practises involves the collection of small samples, one sex only (usually males) and unrandom samples poorly screened for close relatives. Even when such practises are less prominent, the report is often marred by the inclusion of too few attributes. (For example ridge counts are often unreported owing to the effort needed to read them). Finally the adoption of recent non-traditional methodologies by some investigators (e.g. Plato and Wertelecki 1972) has rendered comparison with earlier surveys particularly difficult.

The recent computer revolution has for the first time made the analysis of large dermatoglyphic data series possible. On the other hand techniques for the ordering, accessing and filing of data matrices in the new medium have been independently developed by an unknown number of investigators. As yet there has been little effort to exchange ideas between different departments engaged in the computer analysis of dermatoglyphic variation. Eventually the

collation and comparison of dermatoglyphic data on an inter-regional basis will only progress through the ready interchange of data stored on computer tape. For this to become practicable the first priority is to devise a stringent methodology for dermatoglyphic attribute measurement which will be accurate, consistent and above all accepted and applied on a standardised basis by all dermatoglyphic investigators. The second stage involves the computer archiving of data presented in a format common in all future surveys, thus permitting easy interchange of data for inter-regional analysis.

As this Utopian state of affairs shows little sign of materialising, the dermatoglyphic workers in the Durham Department of Anthropology have pursued an independent approach. The methodological procedures outlined in this chapter were originally drawn up by Dennis and myself in the period 1974-5 as a response to the need to organise the extensive Yorkshire Dales and Kenya series. The procedures (Dennis 1977a; 1977b) have subsequently also been adopted by Williams and Murray for large series of Welsh data. A satisfactory degree of inter-observer consistency has been manifest in the Durham Department, with attributes read and formatted for computer analysis in the same manner. The linking up of the Welsh and Yorkshire sets of data for comparative analysis has already been achieved with success (Dennis and Williams; private communication).

The methodology for the measurement of attributes contains little that is original regarding the basic elements. It is the blend of traditional and modern methods that is of particular interest. The traditional methods for ridge counting, mainlines and topographic definition of digital patterns have been retained, with amendments for some attributes which were too ambiguously defined. The

topographic description of palmar patterns has been replaced by the theoretically sounder methodology of Penrose and Loesch (1969, 1970) which also permits the topological treatment of digital patterns. The topological methodology has been amended in a few places as a result of fundamental disagreement with the points involved. The methodology for formatting the attributes for computer analysis is the work of the Durham Department, in particular of Dennis.

In the present Chapter, therefore, I shall discuss and describe in some detail the dermatoglyphic methodology adopted by the Durham Department generally, and as implemented in this study specifically. The first part of the Chapter deals with important general issues, in particular those concerned with the relative merits of the traditional Cummins and Midlo (1943) methodology for the definition and treatment of digital and palmar patterns and the alternative proposed by Penrose and Loesch (1970). Following this discussion I shall outline detailed rules and methods employed during the reading of the prints, and of the detailed coding procedures for computer analysis. Finally, I shall discuss the precautions taken for the elimination of errors, and will assess the areas of measurement most prone to errors.

## 4.2 Ridge discontinuities and their classification

The flowing concourse of epidermal ridges that traverse the hands and soles is periodically interrupted by imperfections in the ridges, or by discontinuities caused by deviations of the ridges

from their parallel course resulting in the formation of various types of configuration. The initial recognition by Galton (1895) of typical ridge arrangements and the realisation that such patterns were probably inherited provided the stimulous for subsequent detailed classification.

4.2.1 Characteristics of ridge flow and configurations

High resolution observation of the ridges reveals that imperfections, breaks and bifurcation are common phenomena. Galton (1895) described them in detail and termed them "minutiae". Although the study of minutiae has been vital in the development of police forensic procedures, their value to other dermatoglyphic investigators has been minimal. Genetic investigation by Grunenberg (1928) and Steffens (1965) on monozygotic twins did not produce promising results, as no evidence of heritability was found for individual types of minutiae. Okajima (1966; 1967) however, has indicated that the quantity of minutiae present in an individual is inherited, and he demonstrated marked variation in the incidence of forks between the sexes, different finger patterns and dermal regions. The study of minutiae however, is of doubtful relevance to population studies owing to the impracticability of measuring them in more than a small series of individuals.

During observations at a low level of resolution, minutiae can be ignored except in cases where they can influence the measurement of a feature (e.g. in ridge counting). Ridges therefore, can be viewed as continuous lines whose flow results in the formation of various arrangements. At their simplest manifestation, ridges form "open fields" of parallel ridges with no significant interruption of

the flow. They can however, curve to form arches; converge and then spread out in a series of bifurcations to form "fan" configurations; twist sideways from the general flow terminating eventually at a cross flowing ridge producing a "ladder" effect (most commonly found in thenar areas); converge from three directions onto a single point to form triradii or deltas; curve sharply from the radiants of deltas to form loops; alternatively the curves from two deltas can converge to form spiral patterns (whorls) or asymmetrical spirals (double loops) and finally can result in ridge arrangements which resemble embryonic patterns which are termed "vestiges". All these types of configuration are illustrated in detail in Penrose (1968).

4.2.2 Low resolution methodologies

The traditional methodology summarised by Cummins and Midlo (1943) is a topographical approach, with ridge configurations defined according to characteristic shape and location on various areas of the dermal surface. The demarcated areas are the upper pads of the ten digits, and zones on the palm defined by the usual presence of primary triradii underneath each finger and at the base of the palm. The dermal features described in the preceding sub-section are specified in detail and an elaborate classification and nomenclature exists based on the location of each type of feature in the various demarcated areas. The classification of patterns is especially detailed where in addition to the primary types (arches, loops, whorls and accidentals) a large number of sub-types are recognised and defined according to core and directional peculiarities. The resulting possible combinations of configurations in the traditional system are numerous.

Also central to the methodology is the close examination of characteristics of ridge flow, through the tracing of mainlines emanating from the radiants of the primary palmar triradii and terminating at numerated exit points at the perifery of the dermal surface. The study of mainlines is considered to be an essential complement to that of configurations.

Apart from some theoretical inadequacies the traditional methodology can only be criticised in its imprecise definition of a few important features and its perhaps over-detailed typology. The overall methodology, however, is practical and thorough and is still relevant to all aspects of dermatoglyphic research.

The main theoretical failing of the traditional methodology has been a lack of cohesion caused by the failure to appreciate and quantify the evident relationship that exists between deltas, patterns and ridge flow. It was not until 1965 that Penrose published an account of the topological relationship between these features. He noted that:

> "If an area is covered with ridges parallel to one another in small fields, it is topologically necessary that at least two loops are present on it. For every additional loop there must be a compensating triradius. Thus the total number of loops will always be two more than the total number of triradii."
>
> (Penrose 1965:545).

This topological scheme is in evidence in all phenomena where parallel line fields occur, such as zebra strips or human hair on the back, but is nowhere clearer than on the hands and soles:

> "On the fingers and toes a single loop is accompanied by one triradius and the same applies to a symmetrical whorl in which two loops have been fused. Curvature of the lines, which is insufficient to make a loop is called an arch.
>
> On the palms the minimal number of triradii is four. These are usually in the positions designated a, b, c, d and t. For every loop which occurs on the palm there is another triradius. Moreover it can be shown empirically that over the whole hand, including the fingers, the number of triradii exceeds the number of loops by four.
>
> An exceptional arrangement occurs at the apex of each digit where the ridges lie nearly parallel to the edge of the nail; this is equivalent to the presence of a loop pattern, the core of which is replaced by part of the nail. The wrist is also exceptional and corresponds topologically to a loop."
>
> (Penrose 1965:546).

Thus if wrist and fingertip loops are ignored, the following general formula applies to the hand:

$$T + 1 = L + D$$

where T is the number of triradii, L is the number of loops and D is the number of digits. This formula can be reduced to:

$$L = T - 4$$

The topological theory was developed into a concise methodology by Penrose and Loesch (1969; 1970) and Loesch (1971; 1974a; 1974b). Although several aspects of the reported nomenclature are insufficiently defined and several points need amending (see section 3), the topological method is a significant theoretical advance and confers several practical advantages over the topographic approach. Firstly in specifying only triradii and loops as opposed to the intricate topographic pattern typology, the number of possible features is greatly reduced. Secondly it is a more precise method as the topological formula enables one to predict the presence of extralimital triradii when they are not

fully printed. The relative merits of the two methodologies are discussed in more detail in section 3.

4.2.3 Medium Resolution Methods

The accepted methods for the measurement of ridge intensity were initially formulated by Henry (1900) and described in detail by Cummins and Midlo (1943). The importance of ridge intensity, in particular the digital ridge counts, was immediately appreciated by forensic scientists, but it was not until the 1920's that Bonnevie discovered their high heritability and their importance to dermatoglyphic genetic studies. In the post-war period the digital ridge counts have been intensively investigated by Holt in many papers,by Coope (1971) by Roberts and Coope (1972; 1975) and Dennis (1977a). The palmar ridge counts have received less detailed consideration by Fang (1950), Pons (1964), Sciulli and Rao (1975) and Dennis (1977a).

The accepted method for measuring ridge intensity is that of ridge counting, in which triradii and loop cores are used as reference points. On the fingers the number of ridges between the deltas and loop cores are counted. Thus arches have no ridge count, loops one and whorls two. On the palm the primary interdigital triradii are used. As well as the ab, bc and cd ridge counts, the bd and ad counts are sometimes adopted.

A second method for the measurement of ridge intensity is sometimes used, in which an arbritary length, usually 1 cm., is demarcated and the number of ridges counted within that distance. This is mainly used for the measurement of mean ridge breadth (Penrose 1968:1). The method, however, is open to criticism as such arbritary measurement does not allow for individuals with

different sized hands, especially during growth. The location of the count is also critical, as the intensity of ridges can vary dramatically in different parts of the hand.

The traditional scheme of digital and palmar ridge counts is followed in this study, and the detailed rules are outlined in section 4.1

4.3.1 An appraisal of the topographical and topological approaches

It is curious that in supplanting the traditional topographical method Penrose and Loesch have confined their criticisms to isolated paragraphs. Perhaps this can be viewed as a tribute to the general thoroughness of the traditional methodology. At the outset of the discussion the distinction has to be made between the theoretical contribution of the topological theory, which is immense, and its implementations as a formal methodology by Penrose and Loesch which is inadequate on many points. Many of the ambiguities in the traditional approach have not only failed to be amended but have even been included in the topological scheme. Moreover the topological method has introduced ambiguous features of its own which are not present in the traditional system. Workers in the Durham Department take the view that the two methods can be regarded as complementary in many respects. We have attempted to utilise the best aspects of both methods as some of the deficiencies in each method are amended by the other.

One of the more detailed criticims of the topographical methodology is made by Loesch (1971):

> "The genetics of non-measured dermatoglyphic characters could not be specifically studied, because the traditional (topographical) classification of pattern elements was not suitable for their genetical analysis, for the following main reasons:
>
> Firstly, it was based chiefly on main line exits; such characters could not be formulated precisely enough, particularly on the soles; moreover described together with true patterns, they give overlapping information.
>
> Secondly, it combined both true patterns and local disarrangements in the parallel course of ridges which, effectively, resulted in approximately 3,500,000 possible combinations of elements in individual descriptions; this made any genetic analysis impossible."
>
> (Loesch 1971:132)

These assertions can be criticised on the following points:

a) The topographical approach was not "based chiefly on main line exits". The mainline system is just one of the aspects of the method with the main purpose of enumerating pattern direction and general ridge flow. The extensive typology of patterns and other configurational features, especially on the digits can be treated independently of the mainline system, although the topological theory has since shown that they are integrally linked.

b) In eliminating the mainlines Penrose and Loesch have not given enough consideration to whether their alternative scheme adequately compensates for their loss. Their willingness to do away with mainlines resulted from the fact that the topological method was initially developed as a system for sole classification, (Loesch personal communication). The enumeration of mainlines on soles is extremely difficult, but it does not follow that the palmar mainlines present the same problems. On the palms the interdigital mainlines (B, C and D) are as easily enumerated as most patterns. The problematic exit points are 2, 3, 4, 5' and 5'' which lie outside

the interdigital area. The only mainline that is seriously affected by the ambiguous notations (if the 5' and 5'' distinction is eliminated) is the A mainline. Yet the A mainline is the only one which Penrose and Loesch felt was worth retaining in clinical studies. (Penrose and Loesch 1970).

The assertion that the mainlines give "overlapping information" is not proven by Loesch. Extensive multiple regression and factor analysis by Dennis (1977 and private communication) reveals that although a great deal of their variance is explained by the interdigital patterns, a significant portion of the variance is not explained. The A mainline, moreover, comes out as a distinct and independent element.

The redundancy of the interdigital mainlines can only be fully achieved if amendments are made in the interdigital pattern nomenclature. If third and fourth loops are identified separately as formed by either 1) the b triradius, 2) the c triradius, 3) the d triradius and 4) by the third or fourth accessory triradii. only the variance when the B mainline goes to positions 5, 4 or 3 would be lost. This amended nomenclature however, is tantamount to reading the mainlines.

The reading of mainlines, or the adoption of a method of pattern notation that includes all their information is essential if all the dermatoglyphic variance is to be taken into account. Loesch's nomenclature is too general, and she has encountered problems in how to include missing interdigital deltas in her system. Her answer, that of zygodactylous triradii, is ambiguous and confusing, as two distinct phenomena are incorporated by the term. (Discussed later).

c) The criticism that the topographical methodology results in an impossible number of element combinations is a valid observation. In a sense, however, the number of possible combinations is irrelevant. Most of the topographical pattern types listed by Cummins and Midlo are extremely rare, and can be amalgamated into major categories. Thus Loesch's point that there are too many types for genetic study is spurious. For example the many types of hypothenar combinations listed by Cummins and Midlo would require enormous samples if the genetics of each type is to be investigated. However it is equally valid to investigate the genetics of undifferentiated hypothenar loops, which in essence is what Loesch has done. In another instance Holt (1975[a]) conducted a meaningful genetic investigation of the hypothenar radial arch without adopting the topological methodology. The traditional method in fact offers a choice of genetic work on detailed and rare patterns as well as general pattern types.

The study of pattern combinations in individuals is moreover, relatively unprofitable due to the enormous samples needed to obtain significant results for the rarer combinations. The elaborate dictionaries of palmar pattern combinations compiled by Loesch rarely transcend the descriptive level due to the effort involved in compiling them. In addition there is little point in combining features that could turn out not to be genetically linked at all. It is both more practical and more meaningful to treat each feature as a separate variable in the first instance, and to explore the intervariable relationships through correlation, principal components, factor and multiple regression analysis, rather than through descriptive dictionaries.

d) The main consequence of objecting to the combination of "true patterns with local dissarrangements in the parallel course of ridges" has been the total elimination of arches, vestiges and other configurations which do not conform with true topological patterns as defined by Penrose (1968). In doing this, it is worth considering that these elements are probably genetically important. The topological theory, with its clear cut definition of what constitutes a true pattern (e.g. a delta has radiants of 90 - 120$^{o}$) does not reflect the full genetic or development spectrum. Vestiges, for instance, in many cases resemble embryonic loops and deltas which have not achieved full penetrance. Their true place in the topological scheme has not been investigated. Moreover some of the patterns and deltas listed in the topological method clearly do not conform to the exacting definitions of Penrose (1968). Central interdigital loops and thenar loops are rarely accompanied by classical looking deltas with radiants of 90 - 120$^{o}$. The topological theory urgently needs revising to account for these anomalies.

4.3.2 Further criticisms of the topological Method

a) As was pointed out in the section on pattern and delta location, the definition of the positioning of the axial triradii t, t' and t'' is extremely difficult due to the lack of precise reference points on the palm. The ambiguous status of these features has been appreciated for a considerable time. Several methods have been suggested to define the features more precisely (these will be discussed in the next section), but Penrose and Loesch, in retaining

these features in their methodology, have failed to define the features any more accurately than Cummins and Midlo. The definition of the boundary between t and t', and t' and t'' is left to the whim of the observer. In addition they have created variables which are equally ill-defined, which are not present in the topographical system. There is no objective method by which the thenar triradii e and f can be accurately differentiated, and their central hypothenar triradius $t^u$ is difficult to separate from $t^b$. These distinctions have been eliminated in the present study with no noticeable loss of information, and at the same time a substantial improvement in accuracy has resulted.

b) It has already been indicated that the interdigital nomenclature is too general, and that loops should be specified as to which triradius forms them if mainlines are to become redundant. In addition a serious loss of information results from the failure of Penrose and Loesch to specify accessory triradii, the enumeration of which is standard practise in the topographical nomenclature. A general criticism of the topological method is that it seeks to specify thenar, axial and hypothenar patterns and deltas in considerable detail when the means for precise definition is lacking, while at the same time their interdigital classification is generalised inspite of the ready availability of means for precise classification.

c) The final criticism surrounds the confusing status of zygodactylous triradii as defined by Penrose and Loesch. Once again the cause of the confusion has been the direct adaptation of the topological sole classification for the palms without considering

closely enough their specialised characteristics. On the soles zygodactylous triradii are common and occupy a unique position in the central inter-toe V skin area often overlying loops. On the palm, however, they are rare in their classic form, and can be identified when two of the interdigital triradii are missing from their usual positions, and are compensated by a single triradius lying halfway between the normal location points - under the interdigital V. Penrose and Loesch, however, have extended the definition to a different situation when only a single primary delta is missing, usually c. A single missing delta is merely missing when all the other deltas are in their normal positions, and there is no logical basis for classifying this phenomenon as a 'fused' zygodactylous triradius. The rarity of true zygodactylous deltas means that Loesch in her results is mostly reporting the frequency of single missing triradii, features which are fully accounted for in the traditional mainline classification.

The topological principles outlined by Penrose 1965, nevertheless constitute a considerable contribution to dermatoglyphic studies as, for the first time, the mathematical relationship between deltas, loops and ridge flow has been evaluated, bringing overall cohesion to the subject. The mathematical linking of loops with triradii expressed in the topological formula $T = L + (D - 1)$ is especially advantageous in detecting errors, during data cleaning (see section 7). Moreover the formula permits a very accurate compilation of hypothenar and thenar patterns. In the topographical system peripheral deltas such as $t^b$ were often missed if they lay in areas that are usually not printed. In the

topological system the presence of a central hypothenar loop automatically requires the presence of $t^b$ even if it is not visible on the print.

The criticisms of the topological method were primarily directed at Penrose and Loesch's empirical classification rather than at the principles themselves. In many respects the topological classification is similar to the topographical method, as it uses the same areas for delta and pattern location, areas which are topographically rather than topologically defined, (e.g. the II, III and IV interdigital loops have the same general notations in both systems). Apart from the disagreement regarding mainlines and vestigial configurations the main difference is that different types of loop and whorl cores are genetically unimportant and thus the lack of recognition of these features in the topological classification has to be treated with caution.

In the methodology adopted for this study the toplogical classification has been used for the palmar patterns, a decision which was made easier by the rarity of palmar compound patterns. Amendments have overcome many of the main criticisms. For the digital patterns there seems to be no overiding objection why the patterns cannot be treated both topologically and topographically, as each approach yields different but complementary information.

## 4.4 Detailed rules and methods for the measurement and classification of the dermal features considered in the study

### 4.4.1 Ridge Counting

a) The finger ridge counts

The finger ridge count can be defined as the number of

ridges that intersect a straight line drawn from the centre of a triradius to the centre of a loop core. The rules, originally formulated by Henry (1900) are described by Cummins and Midlo (1943) and Penrose (1968). Arches have a ridge count of zero, loops one ridge count and whorls two. There are two basic ridge counts, a radial count when an ulnar loop is present or an ulnar count for a radial loop. Whorls have both a radial and ulnar count. The basic counts, however, have been rarely used in the past. The accepted form is the unilateral ridge count which in the case of whorls the larger of the two counts is taken. The absolute ridge count is sometimes described, when both ulnar and radial counts are summed. The Total Ridge Count (TRC) has been the main measure used in genetic analysis, particularly by Holt, and is the summation of all ten unilateral counts. The total radial, ulnar and absolute counts can also be computed.

Ridge counts are best read under a microscope of X10 magnification. The following rules apply during counting:

1) Where the central point of a triradius is a ridge, the ridge is not included in the count. Where the central point is a space, penultimate ridges before the space are included.

2) As in (1), the last ridge at the centre of a pattern core is not included. In the case of double loops, moreover, where two cores are present, the core nearest to its respective triradius is taken, and the lines of the two counts must not intersect.

3) In the case of lateral pocket loops, where two triradii lie on the same side of the pattern, the greater of the two counts is taken.

4) In the case of extralimital triradii where the delta lies off the printable area of skin, an estimated count is made.

5) Ridges resulting from forking which cross the straight lines are both included in the count.

6) Ridges which fall short of the line are excluded from the count.

7) Short ridges or islands are included in the count.

8) Incipient or interstitial ridges are excluded from the count. These can be easily identified as they are much smaller than true ridges, lie in between ridges and show no sweat pores.

9) Tented arches have zero ridge counts, as do arches, fan formations and occasionally some loops with only the single central core ridge.

b) Palmar ridge counts

The palmar ridge counts are defined as the number of ridges that intersect a straight line drawn between the primary interdigital triradii a, b, c, and d. The reading of palmar ridge counts present more problems than the finger ridge counts owing to the fact that the distal transverse crease often intersects the counting line and also as a result of occasional absence of the primary deltas, which is relatively common in the case of c but rare for the other deltas, especially a and b. The ab ridge count is thus the most reliable one, and is the one most frequently used. To overcome the problem of missing c, the bd ridge count is often

used (e.g. Chai 1972) and the ad count is sometimes adopted. In the present study the ab, bc and cd ridge counts are all counted. In the case of missing c the bc and cd counts are coded as missing, and the bd count is taken. In the case of missing a, b and d the affected count is coded as missing, but no substitute count is presented. The rarity of missing a, b and d triradii ensures that the results are not significantly affected. The total palmar ridge count can also be computed by summing the ab, bc, cd and bd counts on each hand. The rules for ridge counting are similar to those of the finger counts especially rules 1, 4, 5, 6, 7, and 8. The following additional rules also apply:

1) Ridges partially suppressed by white lines are included.

2) Where flexion creases cross the line no estimate is made to compensate for the crease. In essence the crease is jumped.

3) In the case of shifting 'a' triradii, where the a triradius shifts inwards from its normal position the count is still valid.

4) When c is missing the bc and cd counts are coded as missing and the bd count is taken. The missing notation is 00, as opposed to 99 for counts which are impossible to read or totally absent.

5) Missing a, b and d triradii are coded as missing.

4.4.2 The classification of digital patterns.

The digital patterns can be treated both topologically and topographically. The topological classification recognises only the presence of loops. Thus loops can be identified by the

presence of one triradius, whorls are treated as two loops and accidentals as three. Arches have no deltas and are not differentiated. The frequency of triradii on each finger corresponds to the topographical arches, loops and whorls distinction and the total frequency of loops is synonomous with pattern intensity. The pattern intensity of each digit and total left and right pattern intensity are considered in detail in the subsequent population analysis.

A topographical classification has also been included as it yields valuable additional information. Moreover as previous population surveys have universally followed the topographical methodology, and the digital patterns are the most common variables reported, it is necessary to include topographical results for the purpose of data collation and comparison.

Of the many pattern types listed by Cummins and Midlo (1943:60-64), the following frequent types are differentiated. 1) Arches; 2) tented arches; 3) ulnar loops; 4) radial loops; 5) concentric whorls; 6) double loops; 7) central pocket loops; 8) lateral pocket loops. All the remaining rare compound types with three triradii are classified as accidentals.

The patterns are classified according to characteristics of core shape and the presence of deltas. Penrose (1968:4) describes a loop as when the ridges 'turn through an angle of $180^{o}$ and a triradius as a "point where three different fields of almost parallel ridges meet .... resulting in three spokes, the angles of which are greater than $90^{o}$ and are typically $120^{o}$." In practise many triradii have radiants which fall below $90^{o}$ and the rule has been

relaxed in the present study. Vestigial formations which resemble deltas are readily identified as they are usually formed by very few ridges.

The following rules have been followed for the classification of the main pattern types utilised in the study:
Amended definitions and innovations are indicated by *.

1) Arches are easily identified by the absence of deltas and the gentle curving of the ridges.

2) Tented arches have a triradius at the centre of the pattern with the distal radiant terminating at the centre of the pattern core.

3) Two types of loop are differentiated. When the loop core opens out towards the ulnar side of the finger, it is termed an ulnar loop. The core of a radial loop points in the opposite direction towards the radial side of the finger.

4) Patterns with two triradii are known generally as whorls, but six main sub-types are differentiated in the study:

a) Concentric whorls are easily identified when i) * the ridge counts are roughly of the same magnitude ii) when only one core is present.

b) Double loops have two cores whose centres point in opposite directions. Whorls with two visible cores even if the cores are only separated by a minimum of two ridges, are still classified as double loops.

c) & d) When two cores are present that point in the same direction, the patterns are termed lateral pocket loops. If the cores follow an ulnar direction, they are termed ulnar lateral pocket loops or radial lateral

pocket loops if they follow a radial direction.

e) & f) When two deltas are present, but one of them lies very close to a single core, the pattern is termed either a radial or ulnar central pocket loop according to which side of the finger the smaller delta lies. Central pocket loops are difficult to identify in some cases when the delta is small and indistinct. When the core is only one ridge they are classed as loops if the delta is barely discernable. * To render the definition more precise a concentric whorl is a central pocket loop if one of the ridge counts less than half of the other.

5) Compound patterns with three triradii are known as accidentals, and are no differentiated.

4.4.3 The classification of palmar patterns.

The Penrose and Loesch classification has been adopted, with amendments indicated by an asterix (*). As compound patterns such as whorls are rare, no topographical information has been included.

The principles of classification have been outlined in Penrose and Loesch (1970) and Loesch (1974b). The following list is directly quoted from the latter, page 133.

"1) The principles are uniform for palms and soles.
2) Only loops and triradii are included in the description. Loops are named according to the anatomical areas in which they are situated and the main direction of their cores, which can be distal (also called peripheral, i.e. pointing upwards towards the digits in the case of the interdigital loops or outwards towards the ulnar

periphery in the case of hypothenar loops) or proximal (also called central when the loop cores point towards the centre of the palm). If their are two possible types of loops in one and the same area, which are both distal or proximal, one of them is separately specified as radial or ulnar.

Triradii are named according to traditional nomenclature (Penrose 1968) with some modifications."

(The explanations in the parentheses are my own).

Loesch identifies and defines the following patterns and triradii:

"PALMAR PATTERNS

| | | |
|---|---|---|
| I | - | peripheral loop on area I. |
| $I^r$ | - | radial loop on area I. |
| II | - | peripheral loop on interdigital area II. |
| $\overset{\wedge}{II}$ | - | central loop on interdigital area II. |
| III | - | peripheral loop on interdigital area III. |
| $\overset{\wedge}{III}$ | - | central loop on interdigital area III. |
| $III^t$ | - | tented loop on interdigital area III. |
| IV | - | peripheral loop on interdigital area IV. |
| $\overset{\wedge}{IV}$ | - | central loop on interdigital area IV. |
| $IV^u$ | - | ulnar (peripheral) loop on interdigital area IV. |
| H | - | $peri^{ph}_{\wedge}$eral loop on hypothenar area. |
| $\overset{\wedge}{H}$ | - | central loop on hypothenar area. |
| $H^r$ | - | radial loop on hypothenar area, which corresponds to traditional carpal loop. |
| $T^c$ | - | tented loop on the hypothenar area with its core pointing carpally. |
| $T^r$ | - | tented loop on the hypothenar area with its core pointing radially. |
| $T^u$ | - | tented loop on the hypothenar area with its core pointing towards the ulnar side of the palm. |

PALMAR TRIRADII

e - lies in distal part of area I.

f - lies in central or proximal part of area I.

t, t', t'', t''' - axial triradii as traditionall described (Penrose 1968 ).

$t^b$ - border (or extralimital) hypothenar triradius.

$t^r$ - very rare triradius deviating to radial side of palm.

$t^u$ - triradius situated near the centre of the hypothenar area.

z - zygodactylous (fused ab).

z' - zygodactylous (fused bc), traditionally known as the absence of c.

z'' - zygodactylous (fused cd)."

(Loesch 1974b:136).

The reader can find these patterns represented diagramatically in the same paper and in Penrose and Loesch (1970).

(N.B. In this paper figures 10v and vi contain errors. HH should read $HH^r$).

4.4.3.1 Further rules and amendments * in the Penrose and Loesch classification

1) Tented loops are only common on the c triradius, and can be defined as a loop whose core does not display assymetry of direction. The triradius associated with the loop has a radiant terminating at the apex of the loop curve. The extreme rarity of tented loops on the a, b and d triradii, means that they do not have to be specified in the system. On the very few instances when these loops were encountered a and b tented loops were coded as II and d tented loops as IV. *If mainlines are retained these rare tented loops are fully accounted for, but a comprehensive

interdigital nomenclature to be used without mainlines should include spearate notations for these loops.

2) The presence of compound hypothenar patterns such as whorls or double loops could, in some instances, present difficulty. The topological notation in practise, however, is clearly defined. A whorl or double loop usually is noted $\overset{\Lambda}{HH}$ and occasionally $HH^{r}$.

3) *The peripheral hypothenar loop is clearly defined and usually easy to identify. In some cases, however, a peripheral hypothenar loop occurs formed by the ulnar proximal radiant of the axial triradius and curving proximally so that it merges with the wrist crease where a second axial triradius occurs. This type of H loop is extremely difficult to spot at times in print reading, and failure to appreciate its relatively frequent occurrence could result in considerable inter-observer discrepancy. This type of pattern should, perhaps, be coded separately to avoid confusion in future surveys.

4) *All hypothenar patterns whose cores point towards the centre of the palm are classified as hypothenar central loops by Penrose and Loesch. A case, however, exists for the retention of the traditional parathenar loop which differs from the classical $\hat{H}$ in that its apex is not in the hypothenar area associated with triradius $t^{b}$, but in the axial area classically formed by the central lower radiants of two parallel axial deltas. In some populations they are relatively common and should be differentiated from $\hat{H}$. In this survey they are not coded separately as they can be readily computed from existing variables. (See next section).

5) *The hypothenar tented loops $T^c$, $T^r$ and $T^u$ are all extremely rare in the present populations. These variables have been amalgamated into one general variable, $H^T$.

6) *The distinction between the thenar triradii e and f has been eliminated as there is no objective criteria by which they can be defined precisely.

7) The axial triradii are discussed in detail in the next sub-section. The t, t' and t'' distinction has been retained with the divisions coded as defined by Penrose (1968). Triradius t''' has been eliminated as it cannot be usefully differentiated from t''.

8) *$t^r$ has also been eliminated as it is easily incorporated into the traditional nomenclature. Its rarity, moreover, makes it an unprofitable variable for population analysis. $t^u$ is not sufficiently well defined to keep its distinction, and there is no real objection why it should not be coded as $t^b$. The centre of the hypothenar area, which is the crucial criteria for the definition of $t^u$, can only be vaguely demarcated, especially when the periphery of the area is badly printed.

9) *The main objections of Penrose and Loesch's classification of zygodactylous triradii has already been discussed. It must be stressed that the absence of c does not result in the 'fusion of triradii bc'. Classical zygodactulous formations are rare, on the palms, and occur when two primary triradii are missing and are compensated by a triradius midway between their normal positions; just underneath the V between the fingers. Ridge counts can be used to define whether the triradius is acceptably midway, but this is easily achieved by eye in most cases. In this study only classic

zygodactylous triradii are recognised. Normal missing triradii are accounted for in the mainlines only.

10) The interdigital accessory triradii are not individually specified, although it would be advantageous to do so. Unfortunately this latter fact was not fully appreciated until after the prints had been read. In this study the recommended pooling of the interdigital triradii by Penrose and Loesch to form one variable 'D' has been adopted. D corresponds to interdigital delta intensity.

4.4.4 The Measurement of the Axial triradii

The positioning of the axial triradii has already been shown to be extremely problematic to define. The general definition of axial triradius positioning is given by Penrose (1968:7).

> "In the proximal region of the palm where the triradius tends to have the form of a λ, it is simply called t. If the triradius lies near the centre of the palm, where it resembles a γ, it is termed t''. When it is in an intermediate position and two radiants are nearly perpendicular, it is called t'."

The classic delta shapes that Penrose lists are in practise not very helpful as defining criteria as exceptions occur. Moreover the boundary between the low, intermediate and central areas are not defined. In the past considerable inconsistency resulted between dermatoglyphic workers as a result of the practise of allocating axial triradii by eye.

To overcome this problem several methods have been suggested, but all have their disadvantages. Of the quantitative measures the atd angle is inadequate as it reflects lateral positioning of the interdigital a and d triradii as much as the longitudinal position of t. Moreover it is affected by environmental factors, primarily age. The t to d ridge count (Berg 1968) is not only too lengthy and tedious to read, but suffers from having to cross the major transverse creases, thus introducing a significant environmental component. Sharma's t-index (Sharma 1962) is relatively unknown, and confers few additional advantages due to its reliance on creases as reference points. (These methods are criticised in detail by Dennis 1977[a]:127-132).

The method adopted in this study is outlined by Penrose (1968:9).

> "A measurement can be made of the axis length, i.e. the distance between the most distal wrist crease and the most proximal crease on digit IV; the distance between the triradius and the distal wrist crease is then measured and expressed as a percentage of the axis length. If this percentage exceeds 14 and is less than 40, the triradius is termed t'; if it is 40 or more the term is t''. Otherwise it is termed t."

Inspite of the fact that the method is superior to subjective measurement by eye, it still has disadvantages. Firstly inaccuracy results from having to use creases as reference points. Wrist creases are rarely so distinct that the 'feather edges' can be located exactly, and vary considerably both in type, dimension and location amongst individuals. Moreover creases vary according to age. The 14 and 40% points are entirely arbitrary, and a detailed methodological study would be desirable to determine

the best transition points. The positioning of the axial triradii can be considered the least accurate measures in this study, but nevertheless t and t' display consistent variation in the larger sample levels.

4.4.5 The atd Angle

The atd angle exists in two forms. The maximal atd angle is measured from the most distal axial triradius when more than one are present. The minimal atd angle, however, is taken from the most proximal axial triradius. The maximal atd angle has been the usual form adopted by previous dermatoglyphic workers (e.g. Barnicot and Woodburn 1972) and has been found to be an important diagnostic feature in the diagnosis of mongolism. (Penrose 1954 ). The minimal atd angle, however, is more suitable as a quantitative measure as it does not reflect the relative incidence of peripheral hypothenar loops. The minimal atd angle is as much a measure of the relative lateral positioning of triradii a and d than the longitudinal position of t.

There are two main sources of inconsistency in the atd angle which seriously reduces its effectiveness as an objective and meaningful variable for population analysis. Firstly the angle can vary considerably depending on whether the individual's fingers are open or closed during printing. Mainigi and Sharma (1971:48) found that the difference between maximal and minimal abduction could be as high as $11^{o}$. Great care was taken in this study during printing that the fingers were close together. The error was thus reduced to less than $1^{o}$ in most individuals.

Secondly, the atd not only reflects the relative positioning of triradii but also an anthropometric component, the relative breadth and length of the palm. Unlike the other variables which are purely dermal in character the atd angle is affected by age. Penrose (1968:9) recognised the influence of age as a serious objection and this was confirmed by David (1971) who noted that maximal atd angles of over $50^{o}$ are common in neonates.

The decrease in the atd angle as age increases is clearly shown by the Kenyan samples in table 4.1. To minimise the affect of inter population variation the sample was divided into the three major ethnic cateogires, the Cushites, Nilotes and Bantu. The results are very consistent in all the tribal categories, but the differences are slightly less marked in the females. The results show that the angle decreases markedly and progressively through the first four age categories, but is less marked between the teenage and adult categories. The higher values in the female Nilotic 16 to 30 age group is likely due to random factors as the sample size is relatively small. The relationship with age is clearly strong, as the F probabilities are at the 0.000 level in all the ethnic categories for both sexes except for the Bantu females on the left hand.

The minimal atd angle is obtained by drawing straight lines between the at and td triradii. Accuracy depends on the correct location of the exact centre of each triradius as stipulated in rule 1 of the finger ridge count directives. The angle of the lines where they intersect at the central point of the proximal t triradius is measured using a protractor.

In the rare cases where a hypothenar radial arch is present, when the axial deltas are absent and replaced by $t^{b}$, the angle is coded as missing.

4.4.6 The Mainlines

It was originally intended to include the mainlines in the study for reasons discussed in section 3 of this chapter. Due to lack of time however, the reading of the mainlines were not completed, and have had to be left for future analysis.

The omition of the mainlines has inevitably resulted in the loss of variance. However the results of a correlation factor and multiple regression analysis on the mainlines by Palmar patterns by Dennis (1977[a] and personal communication) shows that a large portion of the variance of the interdigital mainlines is accounted for by the interdigital pattern. The A mainline comes out as a separate factor in the factor analysis, but the ambiguous terminations on the ulnar border of the palm makes the measurement of the variable problematic. The omition of the mainlines is thus less serious than the loss of other variable sets except for the atd angle and the positioning of the axial triradii.

The methodology of mainline measurement is comprehensively discussed by Dennis (1977[a], Chapter 4).

4.5 Methods of print reading

At the first stage each sample was prepared by drawing the lines for ridge counting and for measuring the atd angle. While the lines were being drawn small thenar and hypothenar patterns which could be missed in reading were marked. Samples were read at random in order to minimise the development of preconceived trends that might have led to inconsistency.

When reading commenced it was considered more accurate to read related variables only throughout each sample before proceeding to the next variable set, rather than read all the variables consecutively for each individual. The underlying rationale was that by doing this errors could be crosschecked in the subsequent reading of different variables. The order of reading was as follows:

a) The finger patterns with each type written above the print.

b) The finger ridge counts: These enabled the finger patterns to be simultaneously checked, especially central pocket loops which are in part defined by ridge count. The counts were also written above the print.

c) The palmar patterns: the notation was written by each hand.

d) the palmar ridge counts: written above each count.

e) The mainlines: (in all but 1,500 prints). These enabled the direct verification of palmar interdigital patterns.

f) The atd angle: written by each print providing a chance

to reread the thenar, hypothenar and axial deltas and patterns.

In addition ridge counts which displayed marked assymetry with their homologues were recounted.

Finally, when all the prints were read the figures were copied onto the computer coding sheets which provided a final rough crosscheck. These methods are discussed in more detail in the section on error identification, later in this chapter.

The use of a stereoscopic microscope at 10x magnification was essential for the accurate counting of ridges and the tracings of mainlines, especially the A mainline. A sharp steel pointed needle was initially used in ridge counting, but sharp pencils were later utilised to permit rapid recording of ridge scores.

Smudged, obscured and scarred prints, as well as some missing digits and hands, inevitably resulted in having to code variables as missing. The digital and palmar patterns were rarely affected. Digital and palmar ridge counts however, presented more problems, as any imperfection can lead to errors in counting. The practise of reprinting the palmar interdigital area and digits when printing was imperfect, saved a significant portion of the counts. The digital ridge counts presented the most serious problems, as whorls and loops with large counts were more likely to be partially obscured than patterns with small counts. If the percentage of patterns with large counts coded as missing is substantially greater than that of small counts, serious distortion results. Therefore unless a count was impossible to read, an uncommon occurence, an estimate was made of obscured patterns. A

comparison of counts of imperfect patterns with repeated perfect ones, done unconsciously on many occasions when a repeated pattern was not immediately identified as such, revealed close similarity in most cases. Errors of estimation are thus less serious than the disproportionate allocation of missing values to patterns with high ridge counts.

The total time required to read all the variables for one individual and the transfer of information onto computer sheets took approximately fifteen minutes.

## 4.6 Coding, and computer variables.

### 4.6.1 The computer coding sheet

The computer coding sheet used in the present study was designed by Dennis in a general format to permit direct comparison with other projects in the Durham Department. The coding sheet is reproduced in Figure 4.1.

There are three cards for each individual with an optional fourth card for more specialised variables. The case number and card number are the first variables on each card. The first card is concerned with demographic variables. In the case of the present survey only SEX, AGE and POP are relevant. The variable STATE has been included for the purpose of future data collation between surveys in different countries. Dennis (1977b) has allocated a code number for each country in the world as listed in the United Nations Statistical Yearbook 1973.

The second card is mainly concerned with the digital and palmar ridge counts and the digital patterns. The third card contains one overlapping finger pattern variable (RP5), the palmar

patterns, the mainlines and the atd angle. Not all the variables used in the analysis have been specifically coded. The coded variables represent primary attributes and the excluded variables can be computed from the existing traits, (e.g. the right absolute count on digit 1 can easily be calculated by summing the respective radial and ulnar counts). The mainlines have not been filled in, resulting in the loss of zygodactylous triradii which cannot be computed without them. The hypothenar tented loops are all included in UHTR, and $t^u$ and $t^r$ have not been utilised.

4.6.2 Detailed coding procedures

a) Variable CASE

CASE is the identity number for each individual and occurs on each card as a precaution in case they were dropped, and for easy sorting. The case numbers run sequentially through each sample, with the males of each tribal sample enumerated first. In retrospect it would have been easier if the females had been listed following all the males, as they were segregated into separate files during analysis for convenience.

b) Variable CARD

It was essential to number each card to enable computer sorting which is much quicker than manual sorting. The need for sorting resulted from the fact that the punchers preferred to punch each card independently, rather than each case separately.

c) Variable SEX

Males were coded as 1 and females as 2. Males and females were separated initially during print reading and filling in the computer sheets, and they were also punched at different

times. There was thus no possibility that punching errors would result in the mis-allocation of individuals. The male and female series were copied separately onto two independent files. The variable SEX was only utilised when analysis between the sexes took place. In these instances the two files were temporarily amalgamated.

d) Variable AGE

The importance of this variable was reduced by the fact that some individuals had to be aged subjectively owing to the fact that they were ignorant of their year of birth. The only occasion that it was utilised was in relation to the atd angle, (see section 4). The persons assumed age was entered in the two boxes provided.

e) Variable POP

POP is the most important variable in the survey, as it enumerates the tribal samples at there highest level of resolution. These could subsequently be recoded into larger units.

The ordering of units was on linguistic and cultural affinity as discussed in Chapter 1. Firstly the samples were divided into non Bantu and Bantu. The non Bantu had on one polarity the Somali, the least negroid of the Kenyan peoples, and on the other extreme the Luo, the least hybridized Nilotic speakers in Kenya. The Somali and Cushitic populations were followed by the Maasai and members of the Dorobo, Kalenjin and Karimojong groups, and the Luo. The placing of the Laikipiak Dorobo after the Maasai resulted from the fact that they have recently adopted the Maasai language. They are, however, a most distinct group, and no preconceived biological affinity is implied. Within these linguistic

groups the units have been ordered as far as is practicable on geographical proximity.

The Bantu groups followed the non Bantu and were much simpler to order as they occupy three distinct geographical regions. A west to east orientation was chosen. The Gusii and Luyia were listed first starting with the most westerly the Gusii, and finishing with the most easterly, the Bukusu. The Central Bantu were also ordered from west to east in a general sense. The Tharaka, however, were placed last owing to their marked distinctiveness from other Meru peoples. The Coastal Bantu were ordered from north to south, thus maintaining a geographical link with the central Bantu tribes. The Pokomo group was listed first, followed by the Mji Kenda and Taita groups.

The ordering format has thus preserved some degree of linguistic and geographical affinity, an ordering which greatly facilitated the analysis. The following detailed coding was applied to the samples:

Cushitic Category;

Somali Group:

01 - Somali
02 - Rendille

Galla Group:

03 - Gabbra
04 - Boran
05 - Orma
06 - Konzo
07 - Burji

Nilotic Category: (1)

Maasai Group:

08 - Samburu
09 - Keekonyokie (Kadjiado)
10 - Purko (Narok)

Laikipiak Dorobo (separate category)

11 - Ngwesi
12 - Mukogodo

Nilotic Category: (2)

Kalenjin Group:

13 - Kony
14 - Nandi
15 - Kipsigis
16 - Tugen
17 - Keyo
18 - Marakwet
19 - Pokot

Karimojong Group:

20 - Iteso
21 - Turkana

Luo Group:

22 - Luo

Bantu Category;

Luyia Group:

23 - Gusii
24 - Tiriki
25 - Maragoli
26 - Bunyore
27 - Idakho
28 - Marama
29 - Wanga
30 - Marach
31 - Bukhayo
32 - Samia
33 - Butsotso
34 - Bunyala
35 - Kabras
36 - Bukusu

Central Bantu:

37 - Kiambu Kikuyu
38 - Muranga Kikuyu
39 - Nyeri Kikuyu
40 - Nairobi Kikuyu
41 - Ndia
42 - Kamba
43 - Embu
44 - Mbere
45 - Chuka
46 - Muthambe
47 - Mwimbi
48 - Igoji
49 - Imenti
50 - Tigania
51 - Igembe
52 - Tharaka

Pokomo Group:

53 - Pokomo
54 - Malakote
55 - Korokoro

Mji Kenda Group:

56 - Giriama
57 - Chonyi
58 - Mid Mji Kenda
59 - Digo

Taita Group;

60 - Taita
61 - Taveta

f) Finger ridge counts RFR1 to LFU5

The value of each ridge count is directly entered into the two boxes provided. The radial ridge count on each finger precedes the ulnar count. Arches, tented arches and loops with no ridge count are coded as 00 for both the ulnar and radial counts. In the case of ulnar loops the radial count is not coded as missing, but as 00. The converse situation applies for radial loops.

g) Palmar ridge counts RAB to LBD

The value of each ridge count is directly entered into the two boxes provided. When missing c occurs the bd ridge count is entered, and the bc and cd counts are coded as 00. Missing a, b, and d triradii are given the missing value notation of 99.

h) Digital patterns RP1 to LP5

Two boxes are provided for each digit. The following coding values for each pattern type apply:

| | |
|---|---|
| 00 - Arch | 06 - Ulnar Lateral Pocket Loop |
| 01 - Tented Arch | 07 - Radial Lateral Pocket Loop |
| 02 - Ulnar loop | 08 - Ulnar Central Pocket Loop |
| 03 - Radial loop | 09 - Radial Central Pocket Loop |
| 04 - Concentric Whorl | 10 - Accidental |
| 05 - Double loop | |

i) Palmar patterns PTR to TDL

One box for each pattern type is provided. The absence of a loop or triradius is coded as 00, the presence a particular pattern type is coded by enumerating the number of patterns of that type. The presence of two patterns is only common in the case of IV loops. The number of interdigital triradii, specified in TDR and TDL vary from 3 to 8. The palmar patterns and deltas

are thus coded on an interval scale as opposed to the digital patterns which are coded nominally.

j) The mainlines and mainline index, MAINAR to MAINADL will not be discussed as data for the mainlines is not included in the survey. Coding details however, can be found in Dennis (1977b).

k) The minimal atd angle MATDR and MATDL

The value of each angle is entered into the two boxes provided. When the axial triradius is absent[1], due to the presence of a hypothenar radial arch accompanied by $t^b$, MATDR and MATDL are coded as missing (99).

### 4.6.3 Missing values

If missing values are to be completely free from inaccuracy, they have to be well outside the range of variation of a particular trait. As a rule combinations of 9 are ideal. The missing values for AGE, POP, RFR1 to LFU5, RAB to LBD, RP1 to LP5 and MATDR, MATDL are all 99. For SEX, and PTR to TDL the missing values are 9.

The only variables which use two missing value notations are the palmar ridge counts. In statistical analysis of the bc and cd counts, 00 when missing c occurs is coded as a missing value in addition to 99. The reason for this is that when missing c occurs ridges still traverse the region and it is only the absence of a reference point that prevents the ridge count being read. The ridge

---

1. Absent features should not be confused with missing features. A feature is absent when it does not occur on the hand, i.e, there is an absence of patterning in the affected region. A feature is missing when either the hand has not been printed or when, due to imperfections in printing, the variable cannot be distinguished with sufficient accuracy.

count is thus missing rather than absent. In computing the total palmar ridge count 00 is no longer coded as missing, as the bd count is substituted for bc and cd.

4.6.4 Procedure for defining data and computing new variables in SPSS.

The Statistical Package For the Social Sciences (Nie et. al., 1975) offers unique facilities not only in the range of statistical procedures available for the analysis of data but also for computing new variables from the original variable list. Although the bulk of the analysis in the present study was performed using SPSS, recourse was also made to other packages and private Fortran programmes. Facilities for computing new variables in other packages however, are limited and difficult to handle in comparison to the SPSS procedures. In using other packages the scores for computed variables were outputted onto a new raw data file, using SPSS. The definition of data and the computing of new variables in SPSS is thus an integral part of the general methodology.

4.6.3.1 Variables and attributes

The primary variables listed on the coding sheet (Fig 4.1) were used to produce compound variables, and all variables which could be computed from the existing variables were not coded seperately on the computing sheets.

SPSS permits two types of variable transformation. Categories within a variable can be altered and amalgamated using the RECODE command and existing variables can be computed to form new variables using COMPUTE, IF, SELECT IF commands. (SPSS Manual: Nie et.al., 1975).

The following new variables and categories were computed using these procedures:

a) Ethnic categories and groups

The three major tribal categories, the Cushites (1), Nilotes (2), and the Bantu (3), were computed by the use of the RECODE command applied to POP. The Mukogodo and Ngwesi Dorobo (11, 12) were coded as missing as they were of problematic, linguistic and ethnic status, (see Chapter 1).

```
col 1                16
    RECODE           POP(01 THRU 07=1)(08 THRU 10=2)(11,12=99)
                     (13 THRU 22=2)(23 THRU 61=3)
```

The ethnic groups, the Somali, Galla, Maasai, Dorobo, Kalenjin, Karimojong, Luo, Luyia, Central Bantu, Pokomo, Mji Kenda and Taita were computed in a similar manner. The male groups are identical to the female groups except for the Dorobo which are coded as missing (99) in the females due to the small sample size of the latter. The order of the numbers below corresponds to the order of the listed groups above.

```
col 1                16
    RECODE           POP(01,02=1)(03 THRU 07=2)(08 THRU 10=3)(11,12=4)
                     (13 THRU 19=5)(20,21=6)(22=7)(23 THRU 36=8)
                     (37 THRU 52=9)(53,54,55=10)(56 THRU 59=11)
                     (60,61=12)
```

b) Unilateral ridge counts (RF1 to LF5)

The unilateral ridge count is commonly used by dermatoglyphic researchers. It differs from the basic radial and ulnar ridge counts in that a count is always taken for all pattern types except arches,

If a loop occurs a count is always recorded, irrespective of whether the loop is ulnar or radial. If a whorl is found, the greater of the two counts is noted. The unilateral ridge counts were computed in this analysis using the SPSS IF procedure:

e.g.

```
col 1                16

    IF                (RFR1 GE RFU1)RF1=RFR1
    IF                (RFU1 GT RFR1)RF1=RFU1
    ASSIGN MISSING    RF1(99)
```

c) Absolute finger ridge counts (RABS1 to LABS5)

Absolute ridge counts follow the same rules as the unilateral ridge counts, except that when whorls are encountered, the radial and ulnar counts are summed.

e.g.

```
col 1                16

    COMPUTE           RABS1=RFR1+RFU1
    ASSIGN MISSING    RABS1(99)
```

d) Total finger ridge counts (radial: RTRADRC, LTRADRC; ulnar: RTULNRC, LTULNRC; unilateral: RTRC, LTRC; absolute: RTABSRC, LTABSRC).

Total finger ridge counts were computed by summing the respective individual counts.

e.g.

```
col 1                16

    COMPUTE           RTRADRC=RFR1+RFR2+RFR3+RFR4+RFR5
    ASSIGN MISSING    RTRADRC(999)
```

e) Total palmar ridge count (RTPRC, LTPRC)

Total palmar ridge count is the sum of all the individual palmar ridge counts. When the c triradius is absent, or when zygodactylous triradii are present, the bd ridge count is substituted for the bc and cd counts. In contrast to the statistical treatment of the individual bc and cd counts, when absent triradii resulted in the missing notation of (00), in the calculation of the total count (00) is counted as a valid score as it is replaced by the bd count. The missing value notation of (99) however, is retained for the individual counts during the calculation of the total count, as (99) refers to counts that were impossible to read owing to imperfect impressions rather than to absent triradii.

e.g.

```
col 1                16

    COMPUTE          RTPRC=RAB+RBC+RCD+RBD
    ASSIGN MISSING   RTPRC(999).
```

f) Finger pattern intensity (RD1 to LD5)

Finger pattern intensity is the number of triradii present on each digit, and can be calculated by using the RECODE procedure on the raw finger pattern variables RP1 to LP5:

e.g.

```
col 1                16

    COMPUTE          RD1=RP1(0=0)(1,2,3=1)(4 THRU 9=2)(10=3)
```

g) Total finger pattern intensity (RTFPI, LTFPI).

Right and left total finger pattern intensity is the sum of RD1 to LD5;

e.g.

```
col 1                16

    COMPUTE          RTFPI=RD1+RD2+RD3+RD4+RD5
    ASSIGN MISSING   RTFPI(99)
```

h) Total palmar pattern intensity (RPPI, LPPI)

Right and left total palmar pattern intensity is the sum of all the triradii on the hand minus 4. This is to be distinguished from total palmar triradial intensity (RTPLMI, LTPLMI) where the four primary topological triradii are included.

e.g.

```
col 1                16

    COMPUTE          RPPI=TEFR+TR+T1R+T2R+TBR+TRR+TDR - 4
    ASSIGN MISSING   RPPI(99)
```

i) Interdigital pattern intensity (INTOR, INTOL)

This corresponds to interdigital triradial intensity (TDR, TDL) minus the three triradii unnaccompanied by loops specified by the topological formula.

j) Hypothenar pattern intensity (HYPOR, HYPOL)

Hypothenar pattern intensity is the sum of all the hypothenar and axial triradii minus one.

e.g.

```
col 1                13

    COMPUTE          HYPOR=(TR+T1R+T2R+TBR) - 1
    ASSIGN MISSING   HYPOR(99)
```

Hypothenar triradial intensity (RTOTHI,LTOTHI) is just the sum of the axial and hypothenar triradii.

k) The following palmar patterns and triradii can also be calculated from existing variables,

i) Parathenar loops (PAR, PAL)

```
IF          (CHR EQ 1 AND TBR EQ 0)PAR=CHR
```

ii) Hypothenar radial arch (HRAR, HRAL)

```
COMPUTE     RAX=TR+T1R+T2R+CHR
IF          (RAX EQ 0 AND TBR EQ 1)HRAR=TBR
```

iii) Zygodactylous triradii (TZR,TZL,TZ1R,TZ1L,TZ2R,TZ2L)

The mainlines are required to compute these variables. Z needs mainlines A and B,Z' B and C and Z'' C and D:

e.g. TZR

```
col 1                16
    IF                (MAINAR EQ 0 AND MAINBR EQ 0)TZR=1
    IF                (MAINAR EQ 1 AND MAINBR EQ 1)TZR=0
    IF                (MAINAR EQ 1 AND MAINBR EQ 0)TZR=0
    IF                (MAINAR EQ 0 AND MAINBR EQ 1)TZR=0
```

l) Total variables

Right and left hand variables are usually considered as separate, but homologue variables can also be summed by using the COMPUTE procedure.

For example to obtain total unilateral ridge count (TRC) RFRC and LFRC would be summed by a COMPUTE statement,

e.g.

```
col 1                16
    COMPUTE           TRC=RFRC+LFRC
```

## 4.7 Techniques for minimising errors and file editing

Errors inevitably arise in the compilation and manipulation of large data series, not only as a result of human error, but occasionally through computer deficiencies. Dermatoglyphic series are especially prone to error, owing to imprecisions in the methodology, indistinct prints and the fatigue which saps concentration when reading prints over a long period of time. In this section the main sources of error are examined, and the techniques used in the present study to prevent, identify, correct and minimise errors are discussed. The major types of error are as follows:

### 4.7.1 Inconsistency in print reading

The present series of 6332 individuals took over three years to read and to transfer onto computer sheets. The danger of inconsistency in the measurement of the attributes was a possibility that had to be constantly borne in mind during print reading.

The most important procedure to reduce inconsistency was the amending of previous methodologies by seeking to minimise the effect of ambiguity and subjectivity in the identification of patterns and the measurement of problematic attributes. (The details of the final methodology were elaborated earlier in this chapter). The ambiguous and subjective elements, however, were reduced rather than eliminated, and judgements in the case of borderline patterns, indistinct ridges, problematic delta midpoints and shifting interdigital triradii, and above all the decipherment of exact positioning of the wrist and digital creases in a typical individual, vital for the correct assignment of certain attributes, are to a variable extent still dependent on the subjective idiosynchracies of the investigator.

At the early stages of the project the problem of inconsistency was acute, as the final methodology evolved over a period of time as fresh problems were encountered. It became apparent that the early prints were not read in the same manner or as precisely as the later ones. I calculated that roughly 800 of the early prints were suspect. After that figure the methodology was applied with uniform consistency throughout the remaining prints. To make the total series methodologically compatible, the first 1,000 prints were reread for all variables.

Any inconsistencies that persisted probably affected those elements which could not be sufficiently defined in the methodology (e.g. the wrist and digital creases). It is difficult to ascertain how far fatigue and changing moods contributed to this type of inconsistency. It is unlikely however, that inconsistency ever reached large enough proportions to significantly distort the results, except perhaps in the case of t and t' in the smaller samples. There is little one can do to amend this type of error, as it is liable to recur on rereading the prints. The only remedies are to tighten ambiguities in the methodology further, or to seek the aid of other investigators to check the prints, an impractical proposition in the case of this study.

4.7.2 The misreading of attributes

Apart from expected methodological problems, it is a common practise among beginners in the practise of print reading to fail to spot the presence of obvious patterns, and to miscount ridges. A practised and competent campaigner can be expected to be accurate most of the time, but lapses in concentration are liable

to afflict anyone. In short data series this is not a problem, as there is sufficient time to reread the prints to amend errors. In large data series, however, the prints can only be read once. Moreover the reading of the prints can take months to complete, and accumulated pressure, fatigue and boredom can eventually play havoc with concentration.

The misreading of variables was reduced by the adoption of a method of print reading which enabled a certain amount of cross-checking of the data. Instead of reading all the variables for one individual at a time, a section of the variables were read for a series of individuals, and the process was repeated for the remaining variable groups. On the fingers, the patterns were read first for at least each sample, and then the finger ridge counts were read. The ridge counts checked any errors in the patterns, especially central pocket loops which are in part defined by the ridge counts. The palmar patterns were read next, followed by the mainlines which were of considerable help in verification, especially for third tented loops. Further verification was provided by the separate reading of the palmar ridge counts and the atd angle. For the several hundred cases which were not read for mainlines, verification had to depend on the atd angle especially. The digital and palmar ridge counts could not be directly rechecked. Any homologues which did not closely correspond, however, were reread. The atd angle was read last, as it is a measure which is not liable to serious misreading. Finally when all the variables had been read and the results written on the prints themselves, the final stage of copying the results onto computer sheets provided a final opportunity to cross-check the data. Several errors were detected

at this final stage of preparation, thus proving the value of the methods of verification outlined above.

In spite of all precautions, errors as a result of misreading are still bound to be present in the file. I do not think, however, that they would significantly alter the results, as a trial rereading of 100 prints revealed no errors of a magnitude liable to cause concern.

4.7.3 Copying errors, punching errors and file errors

The copying of the data from the prints to the computer coding sheets produced the largest proportion of the errors that were subsequently detected. The most common copying error was the placing of a score in a wrong adjacent blank box. Secondly the boxes for the palmar ridge counts and the atd angle, which coded subsequently to the other variables, were sometimes not filled in. The copying of wrong values was rare, and no cases were detected in a trial check of 100 sheets.

The coding sheets were presented to the computer punchers, who won my admiration by punching over 18,000 cards in less than a month. The incidence of punching errors was stated to be 5% in the introductory computing course I attended, but when the cards are verified and repunched, the rate drops to 0.1%. Punching errors for the present data occured at a much lower rate than 0.1% and a considerable portion of these were attributable to the mispunching of ambiguous numerals on the coding sheets.

The cards were fed into the computer and stored as a magnetic disk file. The card reader has a reputation of skipping cards when large numbers of cards are input, but this did not happen

in my case. The computer also is not entirely infallible, and after prolonged storage on disk, numbers sometimes alter in the file. Only one case has come to my notice since my data was archived.

4.7.4 Procedures for detecting and correcting errors in the data files

The computer cards were organised into their proper case and card order by the M.T.S. *SORT procedure. Not all the cases were ordered correctly, due to punching errors in the case numbers and several missing cards which were inadvertently overlooked by the punchers. An OSIRIS programme, MERCHECK, was adopted to detect misranking and the precise location of aberrant cards. The file was finally corrected using standard editing facilities. All misranked cards were directly checked with the original computer sheets.

The second stage of the operation was to detect errors in the variables. Initially the *SPSS CONDESCRIPTIVE sub-programme was applied to all variables. CONDESCRIPTIVE outputs the maximum and minimum values of each variable, and enables one to clear all variables whose range of values lie within credible limits; (e.g. if the minimum atd angle value is zero, there is obviously errors, as the atd angle always exceeds twenty). Unfortunately over 80% of the variables required more intensive screening.

For detailed error detection, the *SPSS SELECT IF and WRITE CASES procedures were used.

e.g.

```
SELECT IF       (MATDR LT 30)
WRITE CASES     (5x,F5.0,F2.0)CASE,MATDR
```

In this instance any value of MATDR which is less than thirty is printed along with the case number. The erroneous value can be thus easily accessed, and checked with reference to the original coding sheet and print.

Apart from checking extreme values topological agreement was also verified.

e.g.

SELECT IF (TOTLOOP - i.e. total number of palmar loops - NE (TOTPLMI-4) - i.e. total number of palmar triradii-)

Most topological disagreements resulted from the placement of values in the wrong boxes during copying onto the coding sheets. This was almost totally confined to the palmar patterns, where adjacent blank boxes rendered misplacement possible. All errors were checked with the orginal prints.

The procedures used eliminated all major errors from the file. It was not possible, however, to identify errors that lay within the credible range of values, except for the variable POP where all values were sequentially coded according to ethnic affiliation. I discovered during early analysis that the *SPSS sub-programme AGGREGATE reveals any misranking of the POP identity values, and the whole statistical analysis up to that time had to be repeated due to the fact that enough punching errors were discovered to significantly alter the results in some of the smaller samples.

Errors occured at a rate of 8% and 9% of cases in the male and female files respectively, but averaged only 0.09 and 0.11% for each variable.

Any errors that remain undetected are unlikely to seriously distort the results. The main indications that the files have been adequately cleaned and that general errors have been confined to acceptable limits are firstly that the univariate and multivariate results reveal consistent and meaningful variation even in the smaller samples, and secondly that the factor and correlation analyses correspond closely to the results obtained by other researchers in the Durham Department, who have employed the same methodology.

# CHAPTER 5

## Basic structure of dermatoglyphic differentiation between samples

### Introduction

In keeping with the guidelines outlined in chapter 3 this chapter seeks to analyse the basic structure of inter-sample dermatoglyphic differentiation in Kenya. This approach centres around the assessment of relative discordance between taxonomic attributes and involves three stages of analysis.

The first concerns the tabulation and presentation of the summarised facts which provide the evidence on which all subsequent analysis is based. This stage requires little descriptive writing, but is nevertheless a bulky excercise in terms of tables.

The second stage is designed to determine whether it is necessary to proceed to a sophisticated analysis of discordance. As I pointed out in chapter 3, section 2.2.1, the most basic form of overall concordance is homogeneity between samples. If significant difference exist between the samples in the attributes examined at no greater frequency than one might expect from chance, it is clear that all the observed variation is most likely explained by random sampling error and one can assume, until larger samples are obtained, that either a) all the samples are biologically homogeneous or b) that they are too small for us to detect any underlying biological differences with any confidence.

The third stage requires that marked and significant differences have been found in stage two; justifying a more detailed appraisal of relative discordance. The approach I shall pursue at this stage of analysis is based on the direct investigation

of taxonomic correlations between attributes through principal components analysis, and necessarily entails separate analyses of the male and female series. The analytical approach is discussed in section 5.3.

## 5.1 Presentation of the facts

It is essential in any study of human biological differentiation to present at some stage the basic evidence on which subsequent analysis is based. Dermatoglyphic surveys are notorious for the amount of tabulated results they can produce, and this survey is no exception. The justification for presenting this mass of facts is two-fold. Firstly the results represent the smallest divisible structural elements at each population sample level of resolution considered, and from the evidence it will be possible to check subsequent analytical results and conclusions. At the very least, a perusal of the figures will provide some subjective insight into the structure of variation, and underlying trends of relationship. Secondly, they are included for other researchers to use as comparative material. The value of the material is that it represents the most complete and concentrated dermatoglyphic survey of negro populations to date.

In order to avoid a lengthy interruption in the thread of the analysis the results are to be found in Appendix 1. They are, moreover, organised according to three population levels of resolution. At the macro level, they are given for the three maximal ethnic and linguistic divisions to be found in Kenya, the Cushites, Nilotes and Bantu. Although the value of these results

at this level for taxonomic purposes is limited, they nevertheless provide some indication of overall ethnic affinity, and the large sample sizes permit the testing of more general issues such as those presented for inter-population variations in sex differences in chapter 7. The second level is the ethnic group level which provides comparative data between macro regional as well as ethnic divisions. This level of resolution is on the whole more compatible with that used by Hiernaux in his comparative survey of sub-Saharan African populations (Hiernaux 1968[a]). The final level is the minimal ethnic unit level, comprising samples of the smallest definable ethnic units. These samples, in spite of the much lower sample sizes, are by far the most important, as they permit an examination of differentiation unprejudiced by preconceived schemes of ethnic and geographical affinity. Unlike the first two population levels, a minimal sample does not require any significant pooling of potentially widely diverse biological population, and this is therefore the level that will provide most of the subsequent analysis.

The results are given for all attributes measured at all population levels, except for rare traits. Right and left hand attributes are considered separately, with independent series for males and females. To save space, no data are presented for both hands and for pooled sexes. In any case, in view of the heterolateral and sex differences found and discussed in chapter 7, such results are of dubious value. Comparative workers interested in them can nevertheless work them out from the raw data provided.

Most data tables in Appendix 1 for the extensive ethnic group and minimal sample series were produced by computer procedures,

Means, standard deviations, frequencies and valid numbers were initially calculated and written into a binary file using SPSS AGGREGATE procedures (Nie et.al., 1975: Chapter 15), and converted into an image file using the WRITE CASE facilities of SPSS. I then used FORTRAN programs to link the attribute results to population and sample labels and attribute headings and to write them in table format onto a file. These were then printed using the heavy printer for maximum legibility. The tables were condensed further by scissors and paste onto A3 paper and reduced. Thus, the maximum possible accuracy was ensured.

## 5.2 An outline of the range and significance of dermatoglyphic variation at the minimal ethnic unit level of resolution

The purpose of this section is to ascertain whether there is sufficient significant variation exhibited by the minimal ethnic samples for the range of dermatoglyphic attributes considered, to justify a more sophisticated analysis of the structure of dermatoglyphic variation.

In order to test the extent of significance, I applied a one-way analysis of variance to all quantitative attributes, and chi-square to nominal qualitative attributes. Palmar patterns and triradii were treated as quantitative for two reasons. Firstly they are not nominal variables, and many of the palmar patterns and triradii can occur more than once in the hand of an individual (for instance IV loops, peripheral and central hypothenar loops, thenar loops and triradii, and axial triradii). As many of the occurrences greater than one are uncommon (except for IV loops), it means that many of these less common categories have unacceptably low cell

frequencies when chi square is applied. To avoid this distortion through recoding them on a presence and absence basis (for example a person on the right hand has either at least one PH loop or no PH loop) is not satisfactory in that the full extent of the variance is not being tested. Secondly one-way analysis of variance is perhaps a better measure of significance in that it is not proportionally related to sample size as chi square is (Blalock 1960:292-295). The only attributes tested with chi square here therefore, are the finger patterns, recoded to the three basic types of arches, loops and whorls. All the results of significance given here were calculated using SPSS CROSSTABS and ONEWAY procedures (Nie et.al., 1975).

In Table 5.1 are listed the results of tests of significance applied to 60 digital and 32 palmar non summary attributes, and 18 summary attributes. Of the total of 110 attributes considered, 89 or 80.9 per cent of attributes were significant for the males, and 72 or 65.5 per cent for the females.

Breaking down the results into digital and palmar non summary categories we can see from Table 5.2 that the males show a consistently high proportion of significant differences in both categories of attributes (78 per cent for digital and 75 per cent for palmar attributes). The females however, although showing only a slightly lower percentage of significant digital attributes than the males (72 per cent) demonstrate a much lower percentage of significant palmar attributes (50 per cent).

In both sexes, the largest proportion of significant differences is manifested in the radial, unilateral and absolute digital ridge counts, and in the palmar ridge counts and thenar patterns and deltas. Only the finger patterns (females only), the

interdigital patterns (both sexes) and the hypothenar patterns and deltas (females only) show a majority of non significant attributes.

In the case of summary attributes, all the male tests proved significant as did all the female ones for digital attributes. Only 37.5 per cent of the female palmar attributes however, were significant, reflecting the trends in the non summary palmar traits.

In Table 5.3 are also presented the minimum, maximum and range of attribute means and frequencies for the minimal ethnic units sampled. The results reveal that the range of variation is considerable for all attributes. Moreover the attributes with the greatest variation are the palmar and finger patterns which, as we have just seen, are the least likely to be significant. For example, the female range of frequencies for P4L is 39.4 and not significant, in contrast to the range of means for say LFR1 of 6.2, which is highly significant. We can interpret these results as showing that the palmar and finger patterns are much more likely to reflect random fluctuation of frequencies, particularly in the units with the smallest sample sizes. For instance an examination of the percentage frequencies of interdigital patterns in Appendix 1: Tables AP.1.4.40 to AP.1.4.43, reveals that populations with the most extreme upper and lower values are commonly those with sample sizes of under thirty.

On the whole however, the results reveal that there is considerable non random differentiation exhibited by the minimal ethnic units for a majority of dermatoglyphic attributes considered in this study. A more sophisticated analysis of variation is thus necessary, as a situation of simple homogeneity is not apparent.

## 5.3 The taxonomic structure of differentiation between minimal population units: an assessment of relative discordance.

### 5.3.1 Overview

In this section, I shall be examining the structure of dermatoglyphic differentiation between the minimal population units based on the analysis of taxonomic correlations. In doing so, I shall be trying to assess the relative complexity of attribute relationships which form the basic components of the structure of dermatoglyphic variation (see chapter 3).

As I showed in chapter 3, the assessment of relative discordance is a descriptive exercise which focuses attention on likely avenues of exploration for underlying causes that may explain the variation. The greater the concordance, the more likely it becomes that the overall structure is highly ordered and compatible with multivariate clustering or clinal types of variation, such as topotypic, ethnotypic, topoclinal and ethnoclinal variation. In contrast, the more discordant the basic structure, the more likely it becomes that the causative processes associated with "mosaic" evolution are involved (chapter 3).

The analytical treatment of taxonomic correlations and relative discordance involves three avenues of investigation. First, taxonomic correlations between the sexes must be examined as the sex samples from each population are part of a common gene-pool, one would expect them to vary in the same way relative to other populations. If they do not, the most likely explanation is sampling error. If however, the sexes do not vary in the same way and sampling error cannot account for this, we must conclude that differential evolutionary or genetic processes are acting on

the sexes, pointing to a highly complex situation.

The second avenue of investigation is the independent assessment of taxonomic correlations for each sex, which is the main analysis of discordant or concordant structure. The final avenue is one of refinement, in which taxonomic correlations between attributes are compared to statistical correlations between the same attributes. In other words, it is necessary to differentiate taxonomic concordance due to the fact that attributes are related logically or genetically, from concordance that exists in spite of the fact that attributes are not related in indivuduals. This latter type of concordance will be called "evolutionary" concordance, as it implies that evolutionary processes are causing attributes to vary the same way between populations even though there is no phenetic or genetic relationship between the attrbutes. Taxonomic concordance reflecting underlying concordance in individuals is termed "phenetic" concordance to differentiate it from "evolutionary" concordance.

5.3.2 Detailed approach

In chapter 3 I discussed at some length why I considered it more efficient to base an analysis of relative discordance between attributes on a correlation strategy (i.e. the direct analysis of taxonomic correlations) rather than one of distance. From a technical point of view, the analysis of taxonomic correlations involves several stages. The first is the setting up a type B matrix of population means and frequencies for each attribute. The second is the calculation of the correlations. The third is the analysis of the correlation matrix through either principal component or factor analysis, or through some form of cluster or ordination

analysis. Although cluster analysis can be effective in delineating clusters of related attributes (Jardine 1971), a principal component or factoring approach is preferable, as it provides a much more sensitive portrayal of weaker and less obvious relationships between attributes, and also provides a ready means to picture attribute clusters graphically through the ordination of the components.

The main practical difficulty in a correlation approach to the measurement of discordance occurs when a large number of attributes are to be considered, as is the case in this study. The aggressive solution is to input all the attributes into a single very large correlation and/or principal components analysis, and to interpret what comes out. One is precluded from undertaking such a direct assault however, by the facts that a) such a large correlation matrix would be too difficult to interpret directly and would take up far too much space in a report; and b) the number of variables that can be coped with by existing factoring programs is always limited. SPSS factoring procedures for example, can handle only a maximum of one hundred variables in any analysis. We are thus left with the choice of either searching for another program with less stringent limitations or tackling the analysis in sections of different attribute groupings. It is the latter policy that I intend to pursue here.

It is possible to break down the total range of non-rare dermatoglyphic attributes into several divisions. The most basic generic division I propose is to divide the traits into two groups; a) summary composite traits such as total ridge counts, arches, loops and whorls pooled on all fingers of each hand etc.; and b) ordinary non-summary traits from which the summary traits are

composed (for instance the ulnar ridge counts on each digit). These groups can then be divided if the need arises into two further divisions on morphological grounds; digital traits and palmar traits. The summary traits are the most reliable from the point of view of stability of sample size (for example, counting arches on five digits effectively increases sample size by five), but have the disadvantage of potentially losing variation that exists in the ordinary attributes from which they are constructed. The strategy I shall adopt regarding choice of traits at each stage of analysis is to start with the summary traits to examine general trends, and then to break them down into constituent groupings based on morphological divisions to ascertain how much variation has been lost at the summary stage.

As I stated above, a direct correlation analysis becomes increasingly difficult to interpret, the more variables are considered, unless the underlying structure is very simple indeed. (For instance a correlation matrix of over a thousand variables all correlating at well over 0.95 will not be difficult to interpret, but is unlikely ever to be encountered). It is tempting therefore, to by-pass direct correlation analysis in favour of some form of principal component or factor analysis.

This short cut procedure would be totally acceptable if principal component (PCA) or factor analysis (FA), were entirely objective and free of distortion. In practice however, there are many difficulties associated with using PCA or FA. The two procedures are in fact groups of related techniques which produce subtley different solutions according to the methods used. It is wasteful and extremely time consuming to try out every possibility,

and this does not ultimately lead to any easier interpretation. A selection of a limited range of factoring techniques based on careful a priori judgement is preferable not only on the grounds of efficiency and practicality, but also because it forces one to consider the relative limitations of various techniques in more depth.

No matter how well reasoned the basis of choice however, its ultimate justification lies in its effectiveness in clearly depicting underlying trends with minimal distortion. The only practical way of assessing this is to compare solutions derived from the factoring methods adopted with the direct correlation matrix on which they acted. PCA and FA therefore, should ideally be used to complement direct correlation analysis, not as a substitute for it, although it is seldom practical to pursue this policy when correlation matrices are large. At the very least, when the number of attributes is too large to permit a full direct correlation analysis, a pilot study should be performed to compare the direct solution with the factoring one.

The following investigation then will consist of nine stages of analysis:

1) A general correlation analysis of summary attributes both male and female, with an emphasis on delineating broad trends both between and within the sexes. This stage will provide a basis from which to check the efficiency of subsequent principal component analyses.

2) This will consist of a PCA on the same summary traits considered in stage 1, with a focus on inter-sex taxonomic correlations.

3) & 4) Will investigate taxonomic correlations within each sex for summary attributes.

5),6),7), & 8) The use of summary traits is a neat and concise way of examining variation, and has the advantage of keeping output within reasonable limits. On the other hand, in concentrating analysis on summary traits, a doubt remains on how well they really do summarise the overall variation. The purpose of performing PCA on the non-summary traits is to see how much more complex structure really is than the analysis of summary traits suggests.

In considering the non-summary traits, the major problem is that there are too many of them to include in a single analysis. By the time this stage of analysis is reached, however, the summary analyses ought to have clarified relationships between major divisions such as males and females and digital and palmar categories of attributes satisfactorily. (For instance, if digital traits show little relationship with palmar traits in the summary analyses, it is unlikely that they will in the analyses of non-summary traits). The main interest therefore, will be on taxonomic correlations within each morphological category. In stages 5 and 6 I shall thus consider the components of digital variation for males and females respectively, and in 7 and 8 the intra-sex components of palmar variation.

9) This last stage of analysis will separate as far as possible "evolutionary" and "phenetic" taxonomic correlations, by selectively comparing statistical correlations with taxonomic correlations. The issues involved at this stage will be elaborated later.

5.3.2,2 Technical methodology

a) Setting up type B matrices

The correlation and principal component analyses that follow are based on seven type B data matrices. Six of these are for the analysis of intra-sex correlations and one for inter-sex correlations.

The intra-sex matrices are relatively simple, consisting of the means and frequencies for 57 male and 53 female samples of a) 26 summary traits, b) 80 digital traits and c) 32 palmar traits. The precise attributes included at each stage of the analysis are listed in the tables. The surplus of four male populations is accounted for by the fact that the Ngwesi, Mukogodo, Narok Maasai and Malakote female samples were much too small for meaningful analysis. The matrices were drawn up using SPSS AGGREGATE procedures.

The inter-sex matrix is much more complex. In order to compare taxonomic correlations between the sexes (for example to compare how far arches in the male samples vary in the same way as arches in the female samples) it is necessary to differentiate each attribute separately in the male and female series. Thus for instance the Somali population has two attributes for arches on the right hand, one for males and one for females. The structure of the data matrix can be illustrated as follows;

POP1 : ($\bar{X}$ or %) att1(m), ($\bar{X}$ or %) att2(m)..($\bar{X}$ or %) $att_n$(m)..,($\bar{X}$ or %) att1(f)
($\bar{X}$ or %)att2(f),.....($\bar{X}$ or %) $att_n$(f) etc.

POP2 : Ditto

POP(n): Ditto

Where: ($\bar{X}$ or %) att1 = mean or frequency of POP1 for attribute 1.
(m) = males and (f) = females.

In essence, the equivalent attributes for males and females are integrated into one single data matrix, rather than the more conventional format of treating males and females as units of different data series, in effect as separate populations.

The main drawback for this procedure is the fact mentioned earlier that there are four more male populations than female ones. In order to equalise the series, those populations with no female samples have to be omitted.

5.3.2.2 Statistical methodology: the correlation matrices

The taxonomic correlations on which the following analyses are based are Pearson's product moment correlations which measure the extent of common variance between the attribute population means and frequencies that form the columns of the type B matrices described earlier. More technically, Pearson's r is "the ratio of the covariation to the square root of the product of the variation in X and the variation in Y" (Blalock 1960:378) and can be expressed in terms of the formula

$$r = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\left|\Sigma(X - \bar{X})^2 \quad \Sigma(Y - \bar{Y})^2\right|} = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}} \qquad \text{(Blalock 1960:378)}$$

Pearson's r is generally recognised as one of the most powerful correlation measures. Its adoption however, is conditional on the data having certain properties. The most basic of these is that the data must be on an interval or ratio scale, and that a state of bivariate normality and homoscedasticity pertains (Blalock 1960,369). Failure to meet the latter two requirements can lead to

significant distortion, especially if variances about the slope are low and a small number of "outliers" or extreme points at a distance from the slope exist (Blalock, 1960:383). Perhaps the most important limitation of r however, is that it is a measure of linear relationship, and considerable error can result if the bivariate distribution is at all curved.

The attributes of population means and frequencies used in the following analyses are all quantitative measures, and clearly meet the most basic requirement that the data must be at least on an interval scale. Bivariate normality, homoscedasticity and linearity were tested, as Blalock recommends (1960:370) by examining scattergrams. A selection of scattergrams between all combinations of ten representative digital and palmar summary traits showed no tendency for the scatter of points in any of the diagrams to depart from the expected norms of linearity. In other words, in the case of high correlations the points scattered evenly and closely along the length of the slope with the greatest concentration near the middle, while the points became progressively more random and dispersed as correlation decreased. As a further check, I plotted the frequency distributions of all the summary traits and a small selection of the non-summary traits, and found after means and frequencies had been rounded off to whole numbers, a distinct central tendency in all attributes with the mode near the centre of the distribution. (Owing to the great amount of space the above analyses take up, it is impossible to report them here. An examination of the distribution of means and frequencies in Appendix 1 however, will give an approximate verification of these statements).

In view of the fact that the data seemed to be compatible with the requirements of Pearson's r, and that a pilot comparison with Spearman's rank order correlation, which is less stringent in its assumptions, revealed only minor differences, I saw no reason why the Pearson's correlations that are the default option in SPSS factoring procedures should not be adopted. In the correlation analysis that follows, the values are all taken from the preliminary output of SPSS FACTOR procedures (Nie et.al., 1975: Chapter 24) which are laid out in a more condensed and attractive manner than the specifically designed correlation procedures (Nie et.al., 1975: Chapter 18). The FACTOR correlations, however, are less accurate as they are not calculated in double precision. This potential source of error however, is unlikely to be significant, and I did not feel that it was justified to go to the cumbersome and time consuming procedure of calculating seven separate matrices using PEARSON CORR in double precision, and subsequently having to input the matrices into principal components analysis.

5.3.2.4 Principal components analysis

A particularly difficult problem when using FA or PCA in the detailed investigation of taxonomic correlations is to decide which of the several factoring and rotational methods readily available in computer packages such as SPSS should be adopted. This choice is not only theoretically important (as we shall see), but also of practical significance, as different methods can give different terminal solutions. To try several methods is not only time consuming with possibly considerable replication of results, but also an invitation to subjective interpretation in the

sense that there is a danger of concentrating on that solution that conforms best to prior expectations. In order to avoid this temptation, it is better to evaluate choice in some detail before commencing analysis. In the discussion that follows, I shall be examining the issues surrounding choice in two primary areas, a) choice of basic factoring technique, and b) whether it is worth rotating the factor solution, and if so which rotation technique is the most apt for the present investigation.

a) The choice of factoring method

Although the two types of factoring (PCA and FA) can be very similar in their results and are often not clearly differentiated by researchers, they nevertheless differ fundamentally at a theoretical level. As Kowalski stresses:

> "In component analysis we begin with the observations and look for components in the hope that we may be able to reduce the dimensions of variation and ascribe to them some biological meaning. We procede from the data towards the model. In factor analysis we work the other way round: we begin with a model and check to see if it agrees with the data, and if so we use the data to estimate the parameters of the model."
>
> (Kowalski 1972:124)

The model being referred to is a mathematical model, not an experimental one. In PCA the derived components are exactly "defined". As there is no underlying model, we are primarily asking

> "What is the best linear combination of the variables - best in the sense that the particular combination of variables would account for more variance in the data as a whole than any other linear combination of variables. The first principal component therefore, may be viewed as the single best summary of linear relationships exhibited in the data. The second component is defined as the second best linear combination of variables,

> under the condition that the second component is orthogonal to the first. To be orthogonal to the first component, the second component must account for that portion of the variance not accounted for by the first one. Thus the second component may be defined as the linear combination of variables that accounts for the most residual variance after the effect of the first component is removed from the data. Subsequent components are defined similarly until all the variance in the data is exhausted."
>
> (Nie et.al., 1975:470)

In FA however, we are making an assumption that there are r underlying regularities or "determinants" in the data. As Nie et.al. put it

> "The part of a variable that is influenced by shared determinants is usually called common, and that part that is influenced by the shared determinants is called unique. Under this assumption, the unique part of a variable does not contribute to relationships among variables. It also follows from the preceeding assumption that the observed correlations must be the result of the correlated variables sharing some of the common determinants. The implicit faith on our part is that those assumed common determinants will not only account for all the observed relations in the data, but will also be smaller in number than the variables ... Therefore, if there is any correlation between two variables, it is assumed to be due to common factors."
>
> (Nie et.al., 1975:471)

The difference between FA and PCA is thus fundamental when it comes to interpreting factor loadings. In FA we are saying that causal "Factor X" accounts for so much per cent of the variance of variable A, so much of variable B and so on. In PCA on the other hand we are saying that "component X" is so much per cent of the variance of variable A, so much of variable B etc. The causal factors in FA are presumed to exist and there is an assumption that they can be readily identified from the pattern of factor loadings. Thus for example in psychology, factors such as "neuroticism" and "aggression" could be interpreted as accounting for high loadings on

certain types of behavioural attributes. This would be a theoretically incorrect process of reasoning if a PCA had been applied. In the present analysis however, there is no attempt to ascribe a causal model to the data - indeed it would be impossible to do so at the present stage of the investigation. The main aims are predominantly descriptive, to ascertain the nature of basic taxonomic structure as a directional foundation for subsequent causal analysis. As such, PCA is theoretically the most suitable method to achieve these aims.

The choice of SPSS factoring procedures (Nie et.al., 1975: Chapter 24) in this investigation was determined by general convenience and practicability. Although less accurate than BMD procedures which unlike SPSS calculate all factoring statistics at double precision, SPSS gains not only in being easier to access and manipulate, but also because its facilities to output factor scores are much easier to implement. Of the factoring methods available in SPSS, namely PA1, PA2, Rao's canonical factoring, Alpha and Image factoring, only PA1 provides a classic principal components solution. The other methods are all factoring techniques. The use of PA1 is thus, in view of the previous theoretical discussion, the only real choice. The closest of the factoring methods to PA1 is PA2, which differs mathematically from PA1 mainly in that communalities are estimated through an interactive process rather than treating them as unity. The similarity of the two methods can in practice lead to very similar results provided that no causal interpretation is put on the derived factors from PA2. PA1 however, has the practical advantage of being far less costly in terms of computer time used.

b) Choice of rotation technique

In FA or PCA there is one other fundamental area of choice that must be considered. That is whether to "rotate" the initial factors or components, and if so, what method of rotation to adopt. The main justification for rotation in PCA is that the initial components are so strictly conditioned by the requirements of orthogonality that they can be viewed as still too complex, and to some extent, too artificial. Thus, for example, a component could represent an in-between point between two unrelated clusters of variables. Similarly it could be complex, in that many variables load high on the component, while at the same time the same variables also load high on other components.

Owing to the strictures of orthogonality, and the fact that components are extracted in the order of importance, there is a general tendency in PCA for the first component to be a general size factor, loading high and positively on every variable, while the second and remaining components tend to be bipolar, with half the variables loading positively and the other half negatively, and it is consequently difficult to interpret them with confidence. By rotating the components the object is to relax consequences of this artificial process and to achieve a "simplification" of the initial factors or components, and produce much tighter clusters of variables. The desired effect of rotation is to reduce complexity so that variables that load fairly high on several components or factors are rearranged (or viewed from a different angle) so that they load high on fewer components, hopefully on only one. Rotation thus produces a conceptually simpler solution.

SPSS provides four methods of rotation. QUARTIMAX, VARIMAX and EQUIMAX are all orthogonal methods of rotation, in which the initial factors are rotated at 90 degrees to preserve the total independence of each component or factor. QUARTIMAX is a method that seeks to maximise the loadings on one factor or component; it simpliefies the rows of a factor matrix, while VARIMAX seeks to simplify the columns by maximising the variance of the squared loadings in each column. In other words, in QUARTIMAX the focus of interest is on individual variables, while in VARIMAX it is on column combination of variables. EQUIMAX is a compromise between the two methods (Nie et.al., 1975:485). VARIMAX is the best method to adopt in the present study, as it is combinations of variables in "columns" that are of predominant interest.

The fourth method is OBLIQUE, which permits the analyst to relax orthogonality, and to rotate the initial factors at whatever angles he may choose. From a theoretical point of view, oblique rotation has several advantages over orthogonal rotation, as there is no sound reason to suppose that underlying structures in a correlation matrix are completely independent, (i.e. from my point of view, it is artificial to regard components of dermatoglyphic variation as related linearly and additively). From a practical viewpoint, however, oblique rotation has disadvantages. It takes a great deal of experimentation of rotating initial factors at different angles before the "optimum" solution can be found. Secondly, there is no criterion other than subjective intuition to decide on which rotation is the most appropriate. Thirdly, in relaxing orthogonality, the rotated factors or components are no longer decorrelated. This obviously makes the interpretation of

factor scores more difficult, even though SPSS does output a rotated factor matrix which gives the correlations between the rotated factors. Considering the predominantly descriptive purpose of PCA in this study, I decided that the interpretative disadvantages of oblique rotation outweigh the theoretical advantages. VARIMAX rotation is consequently the only method of rotation used in the analyses that follow.

Rotation fundamentally determines the nature of any factor scores that are produced, as the factor score coefficients that form the weightings from which factor scores are calculated, are produced from the rotated pattern matrix rather than the unrotated one. The relationship of attributes mirrored by the factor scores are thus those of the rotated terminal solution. Factor scores therefore represent, if a rotated solution is used, the most condensed and simplified summary elements of differentiation between the populations.

To complete this discussion, I would like to clarify subsequent nomenclature concerning principal component analysis. In order to differentiate my use of PCA from FA at all stages of analysis, the word "factor" will be dropped henceforth. Thus I shall refer from now on to unrotated and rotated components and component scores, rather than the more conventional rotated "factors" and "factor scores".

## 5.3 Descriptive structural analyses.

### 5.3.3.1 Analysis 1: A preliminary examination of taxonomic correlations between male and female summary attributes.

This preliminary analysis examines directly a matrix of taxonomic correlations between both male and female summary attributes. There are two reasons for starting the analysis of the basic structure of variation in this way. Firstly the results of the investigation should provide initial insight into the nature of attribute discordance. Secondly the depiction of these trends directly from the raw correlation matrix should provide an independent series of observations from which to assess and interpret the results produced by PCA on the same attributes.

The correlation matrix on which this analysis is based is reproduced in Table 5.4 It consists of taxonomic correlations between 52 attribute variables, 26 male and 26 female summary attributes, with each taxonomic attribute consisting of the means or frequencies for that attribute by sex of 53 minimal ethnic population unit samples (henceforth referred to as minimal samples). As mentioned earlier in the previous section, the number of minimal samples for each sex have been equalised by omitting those male samples which have only very small female samples (i.e. the Mukogodo, Ngwesi, Narok Maasai and Malakote). There are thus three primary areas of interest. Firstly the nature of taxonomic correlations of attributes within the males and within the females. Secondly how far the trends within the males are similar to the trends within the females. Thirdly the nature of the correlations between male and female attributes.

This third area of interest can be divided into two parts. The first is the nature of correlations between different attributes

across the sexes. (For example between say right total radial ridge count males vs. left total arches females). The second is the type of correlations exhibited between EQUIVALENT male and female attributes. (For example between right total radial ridge count males, and right total radial ridge count females).

A) Within sex trends

1) Digital attributes

All digital attributes show some degree of association within each sex. Homologues show high correlations, being highest in the radial, ulnar, unilateral and absolute total ridge counts (usually well over 0.7) and lowest in the digital patterns (in the 0.5 to 0.7 range). These trends are very similar in both males and females.

i) Total ridge counts and finger pattern intensity

The total ridge counts and finger pattern intensity are all highly inter-correlated within each sex, especially the homolateral counts (i.e. those on the same hand). Within this high range we can distinguish two groupings. The first is the unilateral and radial total ridge counts which inter-correlate very highly (over 0.8). The second is the ulnar ridge count and total finger pattern intensity, which show higher correlations with each other than with the attributes in the first group. Right and left total absolute count in contrast correlates very highly with all the total counts and finger pattern intensity.

ii) The digital patterns

Inter-correlations of digital patterns are in the main only moderate (except for homologues). Total arches show the

least association with other patterns, especially within the males where the correlations are as low as 0.1. The ulnar loops however, show relatively high inverse correlations with concentric whorls, especially on the right hand in both sexes (in the -0.5 to -0.7 range). In contrast correlations between ulnar loops and double loops are lower (less than -0.5) in both sexes. Surprisingly the correlations between double loops and concentric whorls are the lowest of all digital attributes within the sexes (0.2 or less in the females, and reaching as low as 0.05 in the males). The taxonomic distinctiveness of these two types of whorl is interesting, in view of the common procedure of not distinguishing them in dermatoglyphic reports.

iii) Digital ridge counts and pattern intensity versus digital patterns

The radial and unilateral ridge counts only correlate moderately with digital patterns. Moreover there is one exception to the usually strong replication between the sexes in that the radial and unilateral counts correlate very moderately with left ulnar loops (around 0.1) in the males, whereas in the females the same correlations are over 0.5. Arches and ulnar loops correlate negatively with all ridge counts and pattern intensity, whilst concentric whorls and double loops correlate positively with them. Double loops show the lowest correlations of all the digital patterns with the ridge counts.

The ulnar ridge count, total finger pattern intensity and total absolute ridge count show moderate heterolateral correlations with all patterns except for double loops which in contrast show low correlations; homolateral correlations of over

0.7 however, are common between these above attributes. Inter-correlations between homolateral ulnar ridge counts and ulnar loops and concentric whorls are particularly high approaching 0.9.

2) Palmar attributes

Palmar attributes within each sex show a much more variable range of inter-correlations. Homologues exhibit very high values only for total palmar ridge count, which is around 0.9 in both sexes. Interdigital, hypothenar and total palmar triradial intensities by contrast exhibit moderate homologue correlations in the 0.5 to 0.6 range.

Palmar ridge counts show no correlations higher than 0.2 with any other palmar attribute, demonstrating an unusual degree of taxonomic independence. Interdigital intensity (TD) exhibits homolateral and heterolateral correlations of around the 0.5 to 0.7 range with total palmar intensity, but no relationship with total hypothenar intensity (all inter-correlations being less than 0.2 in both sexes). The relationship between hypothenar intensity and total palmar intensity is complex. In the males, both right and left hypothenar intensity correlate at around 0.5 with right total palmar intensity. With left total palmar intensity however, they correlate at 0.3 or less. In the females only right hypothenar intensity versus right total palmar intensities intercorrelate at around 0.5. The other combinations are all low.

3) Digital versus palmar attributes

Correlations between palmar and digital attributes are in the main, low in both sexes. In the males however, interdigital

intensity shows some moderate correlations reaching 0.3 with some digital ridge counts and with finger pattern intensity, but not with any digital patterns. In the females in contrast, interdigital intensity shows very low correlations with the digital attributes that correlate moderately with this feature in the males, but total palmar ridge count, which in the males showed no associations with any digital attribute, correlates up to 0.3 with many digital attributes in the females.

4) Replication of trends between the sexes is strong. Areas where disagreement occurs have already been pointed out.

B) Correlations between attributes across the sexes

1) Across non-equivalent attributes

The large part of the correlation matrix (Table 5.4) which depicts cross correlations between male and female attributes is not easy to interpret. The clearest observation is that no very high correlations are evident. All but two of the cross-correlations are less than 0.5, and a majority are less than 0.3

In general intra sex trends described previously are also mirrored (in a relative rather than obsolute sense) in the inter-sex correlations. For example male digital attributes correlate more highly with female digital attributes than to female palmar attributes. Male ridge counts and total finger pattern intensity correlate negatively with female total arches and ulnar loops, and positively with female concentric whorls and double loops.

In spite of this tendency for the same types of relative relationships to be reflected in the inter-sex as well as the

intra-sex comparisons, in an absolute sense there is some variation in the magnitude of correlations between different sets of inter-sex attributes. Between the ridge counts across sexes for instance, the magnitude of the correlations are approximately half the size of the same set of correlations within each sex. (For example right total radial ridge count correlates with right total absolute ridge count at 0.91 within males, and 0.94 within females, but only 0.42 between males and females). In other sets of comparisons however, the inter-sex correlations are actually higher than the within sex equivalents. For instance right radial ridge count (males) correlates only at 0.15 with left double loop (males), but correlates at 0.35 with left double loop (females).

On the whole however, the cross sex correlations are too complex to be objectively interpreted without the more precise tool of principal components analysis.

2) Inter-sex correlations for the same or equivalent attributes

The inter-sex correlations between equivalent attributes have been underlined in Table 5.4 to make them more easy to pick out. All these correlations are moderate to low, with many falling below 0.3. The digital ridge counts display a similar range of values, correlating in a range from 0.37 to 0.48. The equivalent correlations for right and left finger pattern intensity however, are much lower at only 0.2. The digital patterns show large differences in equivalent correlations, from practically zero for arches (less than 0.1 in both hands) to the relatively high correlations of 0.47 and 0.51 for left and right double loops

respectively. Ulnar loops and concentric whorls however, are closer to the figures for arches, at around 0.2. Double loops are thus distinctive in showing much higher values than the other digital patterns.

Of the palmar comparisons, the palmar ridge counts are very distinctive in showing the highest equivalent inter-sex correlations (0.61 and 0.58 for the right and left hands respectively). All other palmar attributes correlate much less highly at around 0.2, except for left hypothenar triradial intensity which correlates between the sexes at only 0.08.

Discussion

The analysis of taxonomic correlations of attributes within the sexes shows that there is a lack of relationship between digital and palmar attributes, with correlations within each sex across the two groups usually very low. The digital traits reveal the simplest structure, all being at least moderately inter-correlated, and a majority being highly inter-correlated. The most complex associations are the radial/unilateral ridge counts versus the digital patterns and between the digital patterns themselves. The ulnar ridge counts and total finger pattern intensity are in contrast highly related to concentric whorls, and negatively to ulnar loops and arches. Radial ridge counts correlate fairly highly with the "ulnar" variables, particularly the ulnar ridge counts but not nearly as highly with unilateral and absolute ridge counts. The absolute ridge count can be considered a perfect summary of ulnar and radial counts as it correlates very highly with both. Finally, the digital attributes that stands out as displaying only low to moderate correlations with all others, are double loops.

The palmar variables show a complex pattern of inter-correlations, the highest correlations being with homologues only. Apart from homologue comparisons, all the correlations are moderate to low.

The most striking feature of the within sex analysis is the strong replication of trends between the sexes. The trends are not, of course identical, but very similar. On the whole, the female inter-correlations, especially for digital variables, tend to be slightly higher than those for the males.

The inter-sex results are much more complex than the intra-sex ones. There are no high correlations exhibited between any traits between the sexes. At first sight the moderate to low inter-sex correlations support the theory that sample size is having some effect in reducing the correlations between equivalent traits in the sexes. A closer look at the results, however, reveals that the range of correlations for equivalent traits is fairly large, from 0.6 to less than 0.2. The sample size hypothesis would again fit if it could be demonstrated that the traits with the higher correlations are quantitative, and those with the lower correlations are qualitative. In other words, in theory, continuous traits such as ridge counts ought to have means with much lower standard errors than qualitative traits such as digital patterns which would be expected to reveal much larger standard errors in their frequencies for the same overall sample numbers. Therefore the highest correlations between the sexes ought to be for continuous traits, as the lower standard errors would be expected to preserve relative rankings in each sex better. Unfortunately the results

show that not all the qualitative traits correlate lower than the quantitative traits. Double loops in particular have inter-sex correlations higher than any other trait except for palmar ridge counts. Thus the structure of variation between the sexes is fairly complicated and deserves more intensive analysis.

5.3.3.2 Analysis 2: PCA based on 26 male and 26 female summary traits

The aims of this analysis are to firstly clarify the inter-sex trends provisionally described in the previous correlation analysis, by employing a PCA analysis on the correlation matrix presented in Table 5.4; and secondly to compare the results of the present analysis to the correlation analysis in order to assess the reliability and nature of the PCA unrotated and varimax rotated solutions.

The component eigenvalues and the percentage of the variance accounted for by each component are presented in Table 5.5 (The percentage variance of a component is the eigenvalue divided by the number of variables in the analysis). The results show that only 11 out of 52 components have eigenvalues exceeding one, and that these 11 components account for a total of 87.2 per cent of the variance. The highest portion of the variance (30.1%) is accounted for by component 1, followed by component 2 (13.3%). The remaining components all account for less than 10 per cent of the variance, and by component 12 the percentage drops to under two. From component 17 onwards the variance account for is residual at under one per cent.

In deciding which components to describe and display in detail I have followed the conventional practise of putting the cut-off

point at an eigenvalue of one, although to some extent this is an arbritary procedure. The 11 unrotated components with an eigenvalue exceeding one are listed in Table 5.6

Unrotated solution

Component 1 is a conventional size component which loads heavily on the ridge counts and total finger pattern intensity of both sexes and also moderately on ulnar loops, concentric whorls and to a lesser extent on double loops. Loadings on palmar variables are very low. Component 1 is thus a general digital component.

Component 2 loads highly on no single variable. Moderate loadings, however, occur on all the digital variables, with a slight emphasis on those that express ulnarity (i.e. the ulnar ridge counts, total finger pattern intensity, ulnar loops and concentric whorls). Some of the palmar pattern intensity traits also load moderately, especially hypothenar intensity in the males. Component 2 is thus a moderate general component.

Component 3 is a palmar component, loading highly on all palmar variables except the palmar ridge counts and left hypothenar intensity, females. Some moderate loadings (0.1 to 0.3) are also in evidence for some digital patterns, although these are not too consistent between the sexes. Component 3 is thus mainly one of interdigital and total palmar intensity, with hypothenar intensity loading moderately as well.

Component 4 loads highly on the palmar ridge counts (especially in the females), with hypothenar intensity also loading significantly in the females. Arches and double loops also load moderately in both sexes. Component 4 is mainly a palmar ridge count component.

Component 5 is a mixed component, loading highest (0.6) on right total palmar ridge count, males and left double loops, males. Many digital variables also load moderate at around 0.4 to 0.1, as do the palmar intensity variables in the males.

Components 6 to 11 show no high loadings and few middle range loadings. Although it is still possible to discern trends, they are rather weak, and do not warrant detailed consideration.

The unrotated solution is thus highly summarised and most of the clear trends are concentrated in the first four components. The results, moreover, do not clearly mirror the detailed trends (especially the inter-sex trends) described in the preceding correlation analysis. The probable reason for this is that the unrotated solution is too condensed concentrating on generalised trends rather than more subtle particular trends. There is thus strong justification for trying out a rotated solution to simplify the variance.

Rotated solution

The results of the varimax rotated solution is given in Table 5.7.

Component 1 loads highly on female digital variables, in particular the ridge counts, finger pattern intensity, ulnar loops and whorls. The same male attributes load moderately at less than 0.3.

Component 2 is a reverse of component 1, with the traits listed loading highly in the males instead, and only moderately in the females.

Component 3 loads very highly (0.8+) on the palmar ridge counts in both sexes. All other loadings are very low. Component 3 is thus a pure palmar ridge count component.

Component 4 loads high on double loops (slightly higher in the males), and low on other variables. Component 4 is clearly a double loop component.

Component 5 loads high (0.8+) on arches in the males with radial and unilateral counts and ulnar loops in the males also loading appreciably (0.3 to 0.5). Component 5 is mostly an arch component, males.

Component 6 loads highly on interdigital and total palmar intensity in the females and moderately on hypothenar intensity, females.

Component 7 is the male equivalent of component 6.

Component 8 is the female equivalent of component 5.

Component 9 loads highly on hypothenar intensity, males and moderately on right total palmar intensity

Component 10 is the female equivalent of component 9. It is interesting that in both these components there is a moderate relationship between hypothenar intensity and some digital patterns, notably whorls and double loops, in contrast to the usually very low associations manifest between palmar and digital traits.

Component 11 is a weak component with no clear trends.

Discussion

The components in the rotated solution follow closely the trends already pointed out in the preceding correlation analysis. They also provide a clearer picture of what the correlation matrix

contains than it was possible to glean just by looking at the raw correlations.

It is interesting that the inter-sex trends are sharply highlighted in the rotated solution, whereas they were comparatively hidden in the unrotated components. For example in the unrotated solution both male and female ridge counts all load heavily on component 1, masking the fact that in reality the within sex correlations for these traits are generally twice as large as those between the sexes for the same type of traits. The rotated solution, however, closely delineates this important difference by loading highly on male and female ridge counts respectively in different components.

From this evidence it appears that the unrotated solution is only of limited value. It is too summarised, complex and insensitive, whereas the rotated solution is able to depict the subtlest trends fully. Nevertheless the two types of solution must be regarded as complementary, and must be considered together if the raw correlation matrix is omitted from the discussion (as is desirable if it is a large matrix). The rotated solution, by loading highly on only a few variables in each component, tends to imply that the rest of the variables do not really correlate very highly with the variables that show high loadings. This is not necessarily so. By loading highly on specific traits, the rotated solution is concentrating on the most related variables only. The unrotated solution, however, gives a much better idea of the general magnitude of the correlations in the matrix. Significant loadings on the first "size" component in particular tell one that all the traits with such loadings are at least moderately inter-

correlated, as opposed to variables which do not load highly, which show no relationship to the major cluster picked out in component 1. Lesser clusters of moderately to highly inter-correlated traits appear as loading highly on subsequent components.

In essence, therefore, the unrotated solution picks out general clusters of related variables in decreasing order of importance. The rotated solution elucidates trends within each cluster. In describing and interpreting principal components analysis therefore, there is no need to discuss the unrotated solution in much detail, whereas it is important to look at the rotated solution with particular attention.

Interest in the present principal component analysis is primarily focused on the clarification of inter-sex trends, as the intra-sex trends will be the subject of the next analysis. The low to moderate inter-sex correlations and the relatively high intra-sex relationships between certain traits are sharply displayed in the rotated solution. If one excludes the eleventh weak component, eight out of the ten components can be divided into a series of alternating male and female components. Thus component 1 loads high on certain digital traits in the females, component 2 loads high on the same traits in the males, component 5 loads high on arches in the males, component 8 loads high on arches in the females, component 9 loads highly on hypothenar intensity in the males, component 10 loads high on hypothenar intensity in the females. Only in components 3 and 4 do the males and females load highly on the same equivalent traits, namely palmar ridge counts and double loops respectively. This clearly reflects what was observed in

the correlation analysis, where palmar ridge counts and double loops exhibited the highest inter-sex correlations. With the evidence from principal components analysis also available it is now possible to clarify the interpretation of this difference in inter-sex relationships between the traits.

In the discussion of the correlation analysis, I pointed out that although low inter-sex correlations for equivalent traits are normally symptomatic of sampling or methodological error, the difference in the magnitude of such correlations (from less than 0.1 to 0.6) implied that biological factors were also important in accounting for these associations. The principal components analysis has confirmed the observed trends, and clearly differentiated palmar ridge counts and double loops as being relatively much more highly correlated between the sexes than any other traits.

Table 7.5 lists differences between means and frequencies (males minus females) for a selection of traits for each ethnic category. It shows that the relative associations of equivalent traits between the sexes in the present analysis are linked to the degree to which the magnitude of sex differences varies from population to population. Thus double loops and palmar ridge counts which exhibit here the highest inter-sex associations are traits which in Table 7.5 show the least tendency to vary regarding sex differences in means or frequency between the ethnic categories. The other traits, especially the ridge counts, have sex differences which vary appreciably between the ethnic categories. Thus the traits in the males do not vary in exactly the same way as their equivalent traits in the females not only because of the influence of low sample numbers (which must be a significant factor), but

also because of major biological differences whose character is as yet unknown.

5.3.3.3 Analysis 3: PCA based on 26 male summary attributes
Analysis 4: PCA based on 26 female summary attributes

In these two analyses I shall consider trends of association between summary attributes within each sex in more detail than was possible in the basic correlation analysis in Analysis 1. The female correlations are identical as those shown in Table 5.4 The male correlation matrix however, differs from that presented in Table 5.4 in that the populations removed to equalise the male and female series have now been returned.

The component eigenvalues are displayed in Table 5.8 for each sex. In each analysis six components are distinguished as having an eigenvalue greater than one, accounting in toto for 86.1 per cent of variance in males and 85.3 per cent of variance in females. The relative proportion of the variance accounted for by each of the six components is very similar in both sexes, with component 1 displaying the largest percentage variance at just over 40 per cent in both males and females, with components 2 and 3 showing just over 10 per cent of the variance, and the remaining components under 10 per cent.

The unrotated components and communalities are given in Table 5.9 and the rotated components in Table 5.10 for each sex.

The unrotated solution

Component 1 in both males and females is a general digital component, loading very highly on all digital ridge counts, finger pattern intensity, concentric whorls and ulnar loops. Arches

and double loops load more moderately. Palmar attributes in contrast load very low.

Component 2 is largely a palmar component, loading significantly on all palmar attributes in females and all except palmar ridge counts in the males. The magnitude of the loadings are lower than those for digital attributes in component 1, reflecting the lower inter-correlations between palmar attributes in the basic correlation matrix (Table 5.4).

Component 3 in males loads highest on arches and ulnar loops, with moderate loadings on other digital attributes, particularly double loops. The nearest female equivalent is component 4 which also loads significantly on ulnar loops and arches, but also, surprisingly, on palmar ridge counts (0.4).

Component 4 in males loads primarily on the palmar ridge counts, and to a lesser extent on double loops, especially on the left hand which loads at -0.49. The female equivalent is component 5, which differs from the males in that loadings on double loops are much lower, and loadings with left hypothenar intensity on the left hand are higher.

Component 5 in males loads most significantly on palmar ridge counts too, and on right double loops. There is no female equivalent.

Component 6 in males is distinguished by low to moderate loadings, with the highest being on hypothenar intensity, and left interdigital intensity. Again there is no female counterpart.

Component 3 in females is similar to component 2 females, and component 2 males, loading highly on interdigital and total palmar intensity. It differs in not loading on hypothenar intensity of either hand.

Component 6 in females, like in males, is weak, and shows only low to moderate loadings. However, it differs in loading highest on double loops.

The unrotated solution thus presents a picture of strong replication of trends in the first two components between the sexes, but a weakening of sharp relationships and replication between the sexes in the last 4 weakest components. Complexity is evident in the way some attributes load significantly on several components (for example palmar ridge counts on components 4 and 5 in males, and 2, 4 and 5 females; and ulnar loops and arches on components 1 and 3 in males, and 1 and 4 in females). The rotated solution is required to simplify these relationships.

The rotated solution

Component 1 in the males loads highly on radial and unilateral ridge counts; and also on arches, with total absolute ridge count and finger pattern intensity for the right hand loading moderately. The closest female counterpart is component 3, which also loads very high on arches, but unlike the males shows only moderate to low loadings on other traits. This component, therefore, represents a major difference between the sexes.

Component 2 in the males loads highly on all the digital traits except for arches and double loops. The female counterpart is component 1, which loads almost identically to that of the males, the main difference being the fact that the females load slightly higher on the ridge counts.

Component 3 in the males loads almost exclusively on interdigital and total palmar intensity. This component is almost identical to component 2 in the females.

Component 4 in the males loads highly only on double loops, and also very moderately on ulnar loops and concentric whorls for the left hand. The female counterpart is component 5 which also loads highly on double loops, but which differs from the males in loading appreciably higher on ulnar loops, ulnar ridge counts and total finger pattern intensity.

Component 5 in the males loads high on hypothenar intensity and on total palmar intensity for the right hand. In the females hypothenar intensity also loads highly in component 4, but total palmar intensity not as highly as in the males. Another notable difference is the relatively high loadings on the ulnar loops, concentric whorls and total finger pattern intensity in the females.

Finally Component 6 is the same in both sexes, and loads almost exclusively and very highly on the palmar ridge counts.

The advantage of this principal components analysis over the preceding correlation analysis is its superior ability in delineating trends inherent in the correlation matrix. In particular, the rotated solution has succeeded in highlighting the subtle differences between the sexes in intra-sex relationships which were not very obvious when looking at the raw correlations (e.g. component 1, males compared to component 3, females).

The above rotated solution confirms the virtual independence of digital and palmar traits, except for hypothenar intensity which is linked to some extent with digital patterns. It also emphasises the close relationships between digital ridge counts, total finger pattern intensity, double loops and concentric whorls. On the other hand, although it picks out a unique relationship between arches and

radiality, particularly in the males, it does not delineate any unique ulnarity component. Double loops are again shown to be relatively unrelated to other digital traits, and form a component in themselves in both sexes. The palmar variables form three distinct components, the most discrete being palmar ridge counts which load highly on only one component in each sex, interdigital and total palmar intensity which are highly related, and thirdly hypothenar intensity which has more complex relationships with a) total palmar intensity and b) some digital patterns.

The communalities are also interesting in ascertaining the uniqueness of each trait in relation to others. A noticeable feature is the higher communalities on the whole in the males. In the main, the highest values in both sexes are displayed by the digital and palmar ridge counts and total finger pattern intensity. Palmar intensities in contrast have lower communalities, especially in the females where the values drop to under 0.7 for hypothenar intensity. Double loops on the left hand also show lower values in both sexes. We can conclude from the communalities that all traits show significant relationships with at least one other trait and that the portion of unique variance is highest in the palmar intensity traits.

The intra-sex trends of relationships are thus relatively clearcut and point to comparatively simple patterns of relationships. Nothing has been revealed in the present analysis which should change the conclusions reached in the correlation analysis regarding the significance of intra-sex trends. The main gain has been the clarification of components of variation, so that the precise structure of each component is known.

5.3.3.4 Analysis 5: PCA based on 80 digital attributes, males
Analysis 6: PCA based on 80 digital attributes, females

The present analysis seeks to ascertain a) whether significant variation has been lost in displaying the variation of digital summary traits, and b) whether components of variation "hidden" in the analysis of summary traits will be picked out by principal components analysis of the raw digital traits.

The traits included in the following two analyses are the ulnar and radial ridge counts for each digit, pattern intensity for each digit, arches, ulnar loops, radial loops and double loops for each digit. Absolute and unilateral ridge counts have been omitted, owing to the fact that they are in themselves constructs based on radial and ulnar ridge counts, and tented arches, lateral and central pocket loops have been discounted on the grounds of general rarity.

The eigenvalues for the male and female analyses respectively are presented in Table 5.11. Eighteen components have been distinguished as having an eigenvalue exceeding one in both sexes, accounting in toto for no less than 87 per cent of the variance in each sex. Of the eighteen components, component 1 accounts for 27.7 per cent of the variance in the males and 26.5 per cent in the females. Components 2 to 5 each account from c.10 to c.5 per cent of the total variance and all other components fall below 5 per cent. The results are consistent between the sexes.

The unrotated solution is presented in Table 5.12 for the males and Table 5.13 for the females, and the rotated solution in Table 5.14 for the males and Table 5.15 for the females.

Component 1 in the unrotated solution is very similar in both sexes. Moderate to high loadings appear on a majority of the attributes, with only radial loops and double loops showing low loadings. Radial loops and double loops, however, fail to load significantly as groups in any other component, but tend to load very moderately in many components. The general structure is thus one where radial loops and double loops are differentiated from the remaining variables which are as a group at least moderately correlated. The relationship of double and radial loops with each other and with the remaining variables is complex. Of the remaining variable arches, while loading higher than double or radial loops in general, still load appreciably lower than the remaining traits. Arches can thus be considered as partially differentiated.

More detailed trends can be observed by looking at the rotated components:

Component 1 in the males is an ulnarity component loading high on digits 4 and 5 (especially on the left hand) on ulnar ridge counts, finger pattern intensity, ulnar loops (negatively) and on concentric whorls. There is no precise counterpart to this component in the females. Component 4 in the females loads very highly on ulnar ridge counts, fourth digit on the left hand, (LFU4) on LD4, ULOOPL4 and WHORLL4. Moderate loadings are also in evidence on the same types of variable on digits 2 and 3, but digit 5, unlike in the males, loads low. Component 6 in the females also resembles male component 1, loading moderately on digit 4 and quite high on digit 5, on the right hand in particular, though loadings on digit 4 are low on the left hand. Component 1 in the males is thus expressed in two separate components in the females.

Component 2 in the males is very similar to component 1 in the females. These components load highly on digit 1 in ulnar ridge counts, finger pattern intensity, ulnar loops, concentric whorls and double loops. Digit 1 as expressed in these components is one of the very few instances when concentric whorls and double loops load relatively highly together.

Component 3 in the males is almost identical to component 3 in the females. They can be interpreted as generalised ridge count components loading at least moderately on most ridge counts and finger pattern intensity, but only low to moderately in digital patterns. The emphasis in both sexes is on the radial ridge counts on digit 4 and 5, with high loadings on both hands.

Component 4 is an arch component in the males, loading highly and negatively on digits 2 and 3 on both hands. The female counterpart is component 11, which is nearly identical in its main trends.

Component 5 in the males is very similar to component 5 in the females, loading heavily on radial ridge counts and arches on digit 1 on both hands.

Component 6 in the males loads significantly on double loops only, on the second and fourth digits. The nearest female counterpart is component 15, which differs by loading highly on digit 3 as well, especially on the left hand.

Component 7 in the males loads on radial loops, digits 4 and 5 on the left hand. Component 9 in the females also loads highly on these two traits, but unlike the males, a) loads positively instead of negatively and b) also loads very highly on RADLR3.

Component 8 in the males is a radial loop component loading relatively highly on digit 2, right hand, and on digits 1, 2 and 3 on the left. The highest loadings are, moreover, all negative. There is no exact counterpart for this component in the females.

Component 9 in the males likewise has no close female parallel, and is a weak component, loading moderately on WHORLR4, RADLL5 and DLOOPR1 only.

Component 10 in the males is another weak, inconsistent component, loading moderately only on RADLR1, RADLR3 and DLOOPL5. Component 18 in the females also loads on DLOOPL5, but the loading is higher than in the males, and there is no hint of any significant loadings on radial loops.

Component 11 in the males loads highly on RADLR5, and moderately on WHORLL1. In the females, component 17 also loads on RADLR5, but there is no loading worth mentioning on whorls.

Component 12 in the males shows a strong trend of ulnarity on digit 2, left hand only, with high loadings to be seen for this digit in LFU2, LD2, ULOOPL2 and less high on concentric whorls. Component 8 in the females also loads highly on digit 2 for the same traits, but unlike the males, also loads highly on the equivalent traits for the right hand.

Component 13 in the males loads highly on the same traits on the right hand too. It thus takes two male components to express the same relationships inherent in component 8 in the females.

Component 14 in the males is a strong arch component (digit 2 excepted), loading highest on digits 5 on both hands, and less so on digit 4, with only moderate loading on digit 1. Component 7 in the females is very similar, with even higher loadings on ARCHR5, and substantially higher ones for digit 4.

Component 15 in the males loads high only on DLOOPR5 and moderately on ULOOP R & L5. There is no female counterpart for this component.

Component 16 in the males is a weak component loading on digit 5 in the RFU5, ULOOPR5, and WHORLR5. The component is weakly related to component 6 in the females.

Component 17 loads moderately on digit 3 on the left hand (LFU3, ULOOPL3, LD3) and highest on DLOOPR3. The closest female equivalent is female component 2, where loadings on digit 3, particularly on the right hand, are particularly high. Component 17 is thus a pretty weak version of female component 2.

Component 18 loads high on only one trait, RADLR4, in the males. There is no precise female counterpart.

This account following a predominantly male orientation with matching female counterparts, still leaves unaccounted for some female components which have no precise equivalents in the males. These are:

a) Component 10, which loads moderately on RADLL5 and moderately high on RADLL2.

b) Component 14, which is a weak component loading moderately on WHORLR1 and DLOOPR1.

c) Component 16, which loads moderately on RADLL3 and ARCHR5.

d) Component 17, which loads high on $\overset{\wedge}{\text{RADLR5}}$ and moderately on RADLR4.

Discussion

The above account shows plainly that detailed trends are far more complex when individual digital traits are considered than they are in the case of the summary traits. Nevertheless the unrotated components show that apart from double loops and radial loops, all the other traits are at least moderately inter-correlated. As was the case in the summary traits, there is also a strong relationship between ulnar ridge counts, finger pattern intensity, ulnar loops (negative) and concentric whorls in the rotated components. Another broad similarity is the relative uniqueness of arches and double loops, which load less readily in association with other digital types of traits.

A consideration of the more detailed trends shows that the most consistent relationships are between ulnar ridge counts, pattern intensity, ulnar loops and concentric whorls on particular digits. A high loading in any one of these traits in a particular digit is almost invariably accompanied by significant loadings on the others. Replication between the sexes also tends to be high when this pattern of relationship occurs, although differences occur in respect of the relative contributions of left or right hand traits. The strongest trends are exhibited by digit 1 where both hands show the above patterns in the same component in both sexes, followed by digits 4 and 5 which are more complicated in that they show different degrees of joint association between the sexes. Digits 2 and 3 display the weakest components of the above type in the males, but are relatively strong in the females.

A second strong relationship is expressed in component 3 which is very similar in both sexes. This is mainly a radial

ridge count component which is not strongly linked with any patterns. Radial ridge counts show strong relationships to the patterns only in the case of component 5 in both sexes, where radial ridge counts are related to arches on digit 1 on both hands.

Arches show strong relationships to each other by loading high on component 14 in the males and component 7 in the females. Radial loops and double loops, however, show very patchy distributions and weak replication between the sexes. The components in the two analyses which have no counterpart in the opposite sex invariably load unsystematically mainly on double loops and/or radial loops.

5.3.3.5 Analysis 7: PCA based on 32 palmar attributes, males
Analysis 8: PCA based on 32 palmar attributes, females

The aim of analyses 7 and 8 is to see how many components are required to account for a major part of the variation of unsummarised palmar attributes, and whether the nature of the components is similar for both sexes. Secondly, the analyses are designed to find out whether the variation exhbited by the non-summary traits is appreciably more complex than that already displayed by summary palmar traits in analyses 1 to 4.

The 32 traits included in the analysis are mostly conventional, such as palmar ridge counts, interdigital and hypothenar patterns and thenar triradial intensity (TEF). AX(RL) comprises the means for the minimal samples of the summation (in individuals) of axial triradii t, t' and t''. It is thus a measure of axial triradial intensity. Finally, the interdigital triradii are included in a quantitative form in traits TDR and TDL, (i.e. the mean of the total number of interdigital triradii on right and

left hands respectively), and also in qualitative form in variables TDR3 to TDL6, (i.e. they measure the percentage of people with 3, 4, 5, or 6 triradii in each population in the analysis on each hand).

The eigenvalues for both the male and female analyses are given in Table 5.16. Unlike in analyses 3 to 6, the number of components extracted with eigenvalues above unity is not the same in males and females. In analysis 7 in the males, ten components are extracted, whereas in analysis 8 in the females the number is eleven. Nevertheless this difference is not really significant, as the eleventh male component fails only by a fraction of a per cent to reach unity, and is little different from its female equivalent which is just above unity. If one compares the percentage of variance accounted for by each component, the results are very similar for both sexes, with eleven components account for 85.9 per cent of the variance in the males and 86.4 per cent in the females. Each male component, moreover, is extremely similar to its female counterpart in the relative portion of variance accounted for. The eigenvalues also point towards some degree of complexity as no component accounts for more than 23 per cent of the variance, and all but the first three in the males and the first two in the females account for less than ten per cent of the variance.

Unrotated solution

The unrotated solution is presented in Table 5.17 for the males and Table 5.18 for the females. Component 1 is almost identical in both sexes, loading highly on TDR and TDL and on TDR4 to TDL6. Fairly high loadings are also present on P2R and P2L and moderate loadings also appear on some interdigital patterns and

AXR and AXL. In the males, moderate loadings also appear on the bc ridge counts, while in the females the loadings are low for these variables, but moderate for the ab ridge counts instead. Component 1 is thus a size component in each sex, with a strong emphasis on interdigital triradial intensity. Particularly interesting is the lack of association of TDR3 and TDL3 with the other interdigital triradial traits in both sexes.

Component 2 is also very similar in both sexes, loading highly on the palmar ridge counts, especially on the ab counts. Loadings on other variables are at best moderate. Some of the higher remaining loadings are on interdigital patterns. It is surprising however, to find consistent, if moderate, loadings on hypothenar patterns and axial triradii in the males, traits which are normally not associated with palmar ridge counts. In the females, however, all loadings on hypothenar and axial traits are low.

From component 3 onwards, the loadings start decreasing significantly in magnitude, and tend to load relatively highly on particular traits, rather than clusters of traits. Moreover, there is no longer strong replication of components between the sexes. To clarify detailed trends, it is better to consider the rotated components.

Rotated solution

The rotated components are presented in Tables 5.19 and 5.20 for males and females respectively.

Component 1 in the males is similar to component 2 in the females, and loads highly on P2R, TDR, TDR4 and TDR6 with more

moderate loadings on TDR5. It is thus a right hand interdigital component. The left hand equivalent traits also load significantly but moderately at under 0.5. One enigma is the high loading of 0.67 for TEFL in the males, which is not replicated in the females.

Component 2 in the males is very similar to component 1 in the females and is almost exclusively a palmar ridge count component.

Component 3 in the males loads very highly on PHR and PHL, and on AXR and AXL. This component is paralleled by components 3 and 4 in the females. In component 3 the females load very highly on PHL and AXL; and on PHR and AXR in component 4. This shows that there is much greater asymmetry in the female means and frequencies for these traits. The association between PH and AX reflects the strong topological relationship between PH and AX. The axial area can only have more than one triradius if it is accompanied by a peripheral hypothenar loop. The variation in axial intensity thus should closely follow variation in peripheral hypothenar loops.

Component 4 in the males loads highest and negatively on P4R, and slightly less highly and positively on P3TL and P3L. Left hand equivalents load very moderately. The nearest female counterpart is component 9, which loads highly on P4R, and more moderately than in the males on P3TR and P3R. In the females, however, it is P4R that is positively loaded.

Component 5 in the males is the left hand equivalent of component 4, but has no noticable female counterpart.

Component 6 in the males is very similar to component 5 in the females, and loads highly only on central hypothenar loops on both hands. Central hypothenar loops are thus distinguished as relatively unique traits.

Component 7 in the males has no exact female counterpart and loads highly on P2L, TDL, and TDL6. These traits are separated from the other interdigital equivalent traits which load highly on component 1. The absence of a precise female counterpart can be explained by the fact that P2L, TDL and TDL6 load appreciably higher in relationship to their right hand equivalents on component 2 than they do in the males in component 1.

Component 8 in the males and component 10 in the females load very highly on TDL3 only.

Component 9 in the males loads on TDL5, TDL4, and TDL, and is a relatively weak counterpart of component 8 in the females where the loadings on these traits are appreciably higher. In the females P3R also loads moderately high.

Component 10 in the males loads highest on TDR3 with lower loading on TDR5. Component 7 in the females is similar, differing only in that TDR4 is also significantly loaded.

The only female component with no male counterpart is component 6, which loads exclusively on TEFR and TEFL. In the males TEFR does not load together with TEFL on any component. Moreover, TEFR in males is weakly expressed, never loading higher than 0.4. TEFL in addition loads highly on component 1 in the males (0.67), in conjunction with interdigital traits, a relationship which has no counterpart in the females. The absence of a clear thenar component in the males and its sharply defined existence in the females is a notable sex difference.

## Discussion

In both sexes there are two components of variation that override all others in importance (important in the sense that they

clearly express clusters of traits that are most strongly inter-correlated). They are not only very similar in both sexes, but are picked out in the first two components both in the unrotated and rotated solutions. They are a) a strong interdigital component loading highly on certain interdigital patterns, interdigital intensity in its quantitative form (TDR and TDL) and interdigital intensity in its qualitative form (TDR4 to TDL6), and b) a strong palmar ridge count component loading highly on all the palmar ridge counts.

The remaining components tend to be more particularised in that they tend to load on only a few specific traits, and replication between the sexes is not so strong. The most consistent of the remaining components are those loading on central hypothenar loops, (component 6 in both sexes; rotated solution). Peripheral hypothenar loops and AXR and AXL are also strongly differentiated in both sexes, although in the females, right and left hands for these traits are expressed in two different components (rotated components 3 and 4), whereas in the males (rotated component 3) attributes of both hands load together on one component. There is no apparent relationship between the variation of central and peripheral hypothenar traits.

The strong female thenar component (rotated component 6) and its absence in the males has already been briefly discussed. In order to understand its apparent complexity in the rotated solution, it is essential also to look at the unrotated components. In the males TEFR fails to load to any extent on any of the unrotated components either, but TEFL unlike in the rotated solution, does not load to any great extent on any single unrotated component. Instead

it loads moderately on unrotated component 1 (0.37) and unrotated component 8 (-0.44). On both these components, TEFL is associated with higher loadings on some interdigital traits. The male thenar variation, therefore, is one where thenar variation on each hand shows only moderate correlations at best with other interdigital traits, and low correlations between the homologues. In the females however, homologue correlations are not low, as is shown by their loadings on unrotated component 4, and of course on rotated component 6. The homologue correlations are so much closer to each other than any other trait, that they load highly together on one rorated component only, with no visible relationships with other traits. The unrotated female component 4, however, shows that in reality appreciable correlations do exist between the variation of thenar traits, and that of some interdigital and hypothenar traits, as such traits also load markedly on unrotated component 4. Thus if one compares thenar variation in both types of solution in each sex, it becomes apparent that the peculiar trait is TEFR in the males, which is alone responsible for low correlation with TEFL and the absence of a single male thenar component. This can be inferred from the fact that TEFR in the males is the only thenar trait in both sexes that fails to show loadings alongside other palmar interdigital and hypothenar traits.

The remaining undiscussed components deal with the by no means considerable interdigital variance which is not expressed in the main interdigital components discussed earlier (i.e. unrotated component 1 and rotated component 2 in the females, and unrotated and rotated components 1 in the males). The "residual" interdigital

variance not accounted for in the aforementioned primary components is complex, and occupies four rotated components in both sexes (i.e. components 7 to 10 in both sexes). Clusters of related interdigital traits on these components are sharply delineated into right and left hand components, showing that homologue and heterolateral correlations are low compared to homolateral correlations. The general homolateral loadings tend to, with exceptions, stress relationships:

a) P2 with TD-6.

b) P3- and P4- with TD-4 and TD-5 (P3 negatively associated with P4 and TD-4 negatively associated with TD-5).

c) TD-3 largely loading on its own, but sometimes and inconsistently loading with TD-4 and TD-5.

The general picture of the relationship of variation expressed by interdigital traits (not including palmar ridge counts) is a highly complex one.

It is clear from this discussion that the individual palmar traits contain much more information than the palmar summary traits. Nevertheless, the summary traits have succeeded in summarising the most important aspects of the general variation. The most successful summary traits are the palmar ridge counts, whose individual constituent traits load highly on one component only in each sex. There is thus little information lost by using total palmar ridge count to summarise the individual palmar traits. Moreover, as in the summary analyses, the palmar ridge counts show little association with the other palmar traits. Hypothenar intensity is also a reasonable summary trait, in the sense that the

hypothenar and axial traits are clearly differentiated in the present analysis. Unfortunately hypothenar intensity fails to differentiate between two distinct elements of hypothenar variation, namely between the hypothenar central loops (and the topologically identically related triradius $t^b$), and peripheral hypothenar loops and AXR and AXl. Interdigital intensity in contrast only succeeded in summarising a portion of the manifested interdigital variation (although it is an important portion as shown by unrotated component 1 in both sexes). The true interdigital variation is in reality fairly complex. Finally the absence of a summary trait for thenar variation is somewhat mitigated by the fact that some association exists between thenar and interdigital variation. Thenar associations however, are complex, and are no more than partially expressed by total palmar intensity.

## 5.4 The extent of "taxonomic" concordance between dermatoglyphic attributes considered in this study.

In considering structural relationships between attributes the study so far has been based purely on taxonomic attributes. As we have seen, though there is considerable discordance displayed between many attributes, particularly between attributes belonging to different areas of the hand (e.g. between digital versus palmar attributes), attributes within the same areas often show high taxonomic correlations. This is especially evident for digital attributes, particularly homolateral ones.

As an appraisal of the extent of taxonomic concordance, however, the analysis to this point is misleading, as I have not taken into account the statistical correlations between attributes

in individuals. If, for example, the presence of whorls is strongly correlated with ulnar ridge counts in individuals, this would imply that they are problably phenetically, logically or even genetically related. It would thus be very surprising if they were not also highly related taxonomically. To assess the true extent of taxonomic concordance, therefore, we must control for statistical correlations in individuals. If taxonomically related attributes show no more concordance than we would expect from their phenetic relationships, then they are not really taxonomically concordant.

In the following investigation, I shall attempt to estimate the extent of taxonomic concordance by comparing correlations between attributes both in individuals and in populations. Of particular interest is the question, not so much of whether attributes that show high statistical relationships are even more related taxonomically, but rather of whether attributes that show low or moderate statistical relationships, are more highly related taxonomically. In addition, it is of interest to find out whether the extent of true taxonomic concordance differs appreciably between groups of attributes belonging to different areas of the hand.

### 5.4.1 Approach and methods

One of the major problems in comparing statistical and taxonomic correlations for dermatoglyphic data is that it is impossible to correlate all attributes in individuals meaningfully because some are nominal variables. In this study, the purely nominal attributes are the digital patterns.

Although it is possible to convert nominal attributes into a series of dichotomous variables on a present and absence

basis (for example the individual has a whorl, or does not have a whorl on digit 5 of his left hand), thus rendering them suitable for inclusion in correlation analysis, I felt that this was impractical, given the large number of permutations involved. Thus the digital patterns are omitted in the analysis here. This in a sense is unfortunate, as some of the most interesting relationships occur between digital patterns and ridge counts in the analyses in section 5.3.

Nevertheless this still leaves all the remaining digital attributes (ridge counts and pattern intensity), as well as all the palmar attributes. The number of attributes left, moreover, is far too large to permit an analysis of correlations between every attribute. As a result I have organised attributes into four groups. In the first are non nominal summary attributes, digital and palmar. An examination of these attributes should reveal the nature of general trends, for instance, whether taxonomic correlations are higher than statistical correlations in digital rather than palmar attributes. In the remaining three groups, non summary attributes are considered. The first of these three consists of digital ridge counts (radial and ulnar), the second of palmar patterns and triradii, and the third of palmar ridge counts. The separate consideration of palmar ridge counts is desirable for they showed the highest inter sex taxonomic correlations of any group of attributes in section 5.3.3.1 and, therefore, might be more concordant taxonomically than other attributes.

The basic evidence for this investigation is correlation matrices between attributes, in which statistical correlations

between attributes in individuals are displayed side by side with taxonomic correlations between the same attributes. The statistical correlations were calculated using Spearman rank order correlations on sub sets of the main sample of individuals, using SPSS procedures (Nie et.al., 1975). Owing to the fact that the workspace required to calculate Spearman correlations with SPSS using the full male and female matrices of individuals was enormous, I had to limit the sample size to 1,500 randomly selected individuals for each sex.

The taxonomic correlations were calculated as described earlier in section 5.3

The basic matrices comparing statistical and taxonomic correlations, however, are too complex to permit more than a very general appraisal of trends. The main analysis consists of comparing means of statistical and taxonomic correlations through pairwise t-tests, (test for paired samples), and through correlating (Pearson's r) the two sets of correlations (Nie et.al. 1975:Chapter 17).

### 5.4.2 Results

The basic correlation matrices are presented in Tables 5.21 to 5.24 for the four main groups of attributes described earlier in section 5.4.1. It is plain from a cursory look at two types of correlations that they are very similar. If we compare, for instance, correlations between RTRADRC and other summary attributes in Table 5.21, both statistical and taxonomic correlations are almost identically high in both sexes with other

digital attributes, and very low with palmar attributes. The main fluctuations in fact, occur when both types of correlations are low. For instance RTRADRC males vs TDR and TDL shows higher taxonomic correlations of .2, compared to statistical correlations of .02. In females, however, the equivalent correlations are close to zero for both statistical and taxonomic correlations. These fluctuations at the low polarity are almost certainly due to random effects.

For more precise interpretation , we must turn to the results in Table 5.25, which compare means of statistical and taxonomic correlations in different groups of attributes. The most outstanding trend is that the mean taxonomic correlation is higher than the mean statistical correlation in every attribute group or sub-group examined. The differences between the means, moreover, are significant in 16 of the 20 t-tests performed, and highly significant in all 8 t-tests for the undivided 4 major groups of attributes, A, B, C and D.

The non significant differences in mean occur in the summary attribute sub-divisions; in the inter-digital correlations in both sexes, and in the inter-palmar and palmar vs. digital correlations in the females only.

In spite of the significance of these results, however, it is noticeable that the mean of taxonomic correlations is only slightly higher than that of statistical correlations in all attribute groups except for palmar ridge counts. The difference in means is usually well under .1. Palmar ridge counts, in contrast, have strikingly higher taxonomic correlations than statistical ones, with differences in mean of .19 in both sexes.

The essential overall similarity between statistical and taxonomic correlations is portrayed by the results of Pearson's r in Table 5.25. The correlations between statistical and taxonomic correlations are well over .9 for summary attributes and palmar summary attributes in both sexes, and are usually well over .7 for all sets of inter attribute correlations except two. These are for B1 (radial versus ulnar ridge counts), where r is moderate (.4 in males and .6 in females); and for A3 (digital versus palmar summary attributes), where r is very low (.002 in males and .336 in females). In A3 the low correlations are associated with very low means for both statistical and taxonomic correlations. It is obvious that all correlations between attributes in this group, vary randomly within the group to a large extent. Nevertheless it should be emphasised that all statistical and taxonomic correlations in the A3 group are universally low, and are significantly low when compared to the correlations in sub-groups A1 and A2.

### 5.4.3 Discussion

The overall similarity between statistical and taxonomic correlations as implied by the results of Pearson's r above, show that the high taxonomic correlations found between many attributes in the analyses in section 5.3, predominantly reflect the fact that such attributes are highly related phenetically in individuals in the first place. In other words total taxonomic patterns of concordance closely reflect underlying patterns of phenetic concordance. The only notable exception is in the case of palmar ridge counts, where there is much more taxonomic

concordance than we would expect from the displayed figures of phenetic concordance.

Having said this, it is also plain that there is some small, but significant taxonomic concordance in all the other major groups of attributes, which is not accounted for by phenetic concordance. We can quantify this pure taxonomic effect if we use as a constant the fact that theoretical total concordance is one.

The difference in the means between taxonomic and statistical correlations multiplied by a hundred and divided by one, is the percentage of "pure" taxonomic concordance (assuming that the mean for taxonomic correlations is greater). In Table 5.25 we can see (multiplying the differences in mean by 100), that "pure" taxonomic concordance varies from 1 to 9 per cent in groups A to C, but is 19 per cent for palmar ridge counts. Pooling attributes in all the groups in Table 5.25, we get an overall figure of "pure" or adjusted concordance of 6 per cent in males and 6 per cent in females.

We can infer from this that in Kenya, the basic structure of dermatoglyphic variation is taxonomically very discordant. I would predict, therefore, that the overall structure of variation should not contain any strong multivariate clusters or clines, and that the strength of any types of variation such as ecotypic or ethnotypic that might be operating in Kenya, will at best be moderate on the overall gene pool, though of course, they might be important on some individual attributes.

# CHAPTER 6

## Overall structure of dermatoglyphic differentiation between minimal ethnic unit samples

### Introduction

The analysis of basic structure in Chapter 5 has revealed that taxonomic correlations between attributes are only high in those attributes that are also correlated significantly in individuals. This discordant taxonomic situation implies that overall population inter-relationships will be patterned according to a web or mosaic of cross-cutting relationships, rather than to a series of mulivariate clusters of clines.

The notion that significant discordance in the basic structure will also reflect an absence of multivariate clusters or clines however, needs to be empirically demonstrated, as so far it has only been put forward as a theoretical proposition (see Chapter 3). The first part of this chapter therefore, examines the overall structure of dematoglyphic variation between the minimal population unit samples to see whether the multivariate variation is structured in a manner resembling a quasi-random network or web, rather than as a series of clusters or clines.

The second stage of analysis investigates structure in detail, considering the nature of the characteristics of the scatter of specific populations in multidimensional space as condensed through ordination analysis. If the overall multivariate structure does indeed resemble a mosaic, the aim of the detailed consideration of the scatter of population points will be to ascertain whether meaningful regularities can be discerned in the affinity of population points. By meaningful is meant whether the regularities

are consistently demonstrated in various ordination solutions obtained from different methods.

In contrast if the overall structure is not complex, but shows distinct clusters or clines, the way is opened up for a detailed investigation into the exact nature of these clusters or clines, and of their relative homogeneity, discreteness and diagnostic characteristics.

## 6.1 Methodological and statistical approaches

The analysis of overall structure that follows is based on two approaches. The first involves the calculation of phenetic distances between the population units and the analysis of these distances through the application of ordination techniques. The second approach is based on the analysis of principal component scores, resulting from the PCA analyses examined in Chapter 5.3. The first approach, the analysis of distances, considers the nature of inter-sample relationships based on attributes that have not been decorrelated, and can hence be viewed as a measure of overall phenetic affinity. The examination of PCA scores in contrast is based on attributes which are decorrelated during the calculation of the PCA scores, and therefore can be viewed as measuring adjusted phenetic affinity, with a likelihood that the depicted relationships reflect genetic affinity more closely.

### 6.1.1 Choice of distance statistic and the calculation of the distances

The most important requirement of a distance statistic that is applied to measure the overall structure of variation (see Chapter 3) is that it should be able to accommodate a full spectrum

of attributes, unrestricted by the statistical nature of the attributes being considered. Distance statistics that can only cope with quantitative attributes (for example Penrose's $C^2H$, the CRL, Mahalonobis's $D^2$) or those that can only accommodate qualitative traits (for example Edward's $E^2$, Sanghvi's $G^2$, Spuhler's R - see Constandse Westermann 1972 for details of all these measures) cannot examine more than a limited range of dermatoglyphic attributes, and therefore fail to meet the requirement.

Constandse-Westermann (1972: Chapter VI) describes two measures of affinity that can accommodate both quantitative and qualitative attributes. These are Hiernaux's $\Delta g$ (Hiernaux 1965) and Oliver's $X^2_g$ (Oliver 1970). The latter measure however, is not a true measure of distance, involving the calculation and summation of pair-wise t-test and chi square probabilities. Hiernaux's $\Delta g$ therefore, is an obvious choice for calculating distances between populations based on both quantitative and qualitative attributes.

Hiernaux's $\Delta g$ acts directly on a type B matrix of population attribute means and frequencies, where each attribute is standardised by the world gamma of variation for that attribute (in other words gamma is the known maximum minus the known minimum mean or frequency for a trait, recorded in human populations). Its formula is as follows:

$$\Delta g = \frac{10{,}000}{r^o} \left\{ \sum_{j=1}^{r^1} \left( \frac{M1j - M2j}{aj} \right)^2 + \sum_{j=1}^{r^{11}} \sum_{k=1}^{sj} \left( \frac{Pkjk - P2jk}{ajk} \right)^2 \right\}$$

where: $a_j$ ($a_{jk}$) is the world gamma of variation for the $_j$th quantitative attribute (M) or for the $_k$th class of the $_j$th qualitative attribute (P): r' is the number of quantitative attributes, r'' the number of qualitative classes/attributes, and $r^o$ = r' + r''. (Constandse-Westermann 1972:127-8).

Unlike Mahalonobis's $D^2$, $\Delta g$ has the disadvantage that it does not decorrelate its attributes before the distances are calculated. Nevertheless Hiernaux (1965 and 1972) has demonstrated that distances from $\Delta g$ based on quantitative traits correlate highly with distances from $D^2$ and $C^2_H$ based on the same attributes, implying that its failure to decorrelate attributes is not serious in practise.

The problems encountered in implementing $\Delta g$ in the analyses that follow were threefold. The first and most pressing was that there was no available computer program. This problem was overcome by writing a program in FORTRAN 77 based on the formula presented by Constandse-Westermann (1972: Chapter VI). The reliability and accuracy of the program was tested by a) running the example provided by Constandse-Westermann in the same chapter to see if the same result was obtained; and b) by comparing five distances as computed by the program from the current analyses with the same distances calculated by hand. The program is listed in Appendix 2.

The second problem concerned the calculation of a world gamma of variation, which is used to standardise the attributes. As the range of dermatoglyphic attributes covered in this study is much greater than has been the practice previously (excepting Dennis's (1977a) Yorkshire Dale survey), and given that the use of the Penrose and Loesch (1970) method makes comparison of patterns

in this series strictly incompatible with those measured traditionally in the bulk of previous studies, it became plain that the calculation of a world gamma of variation would be a dubious exercise. In any case, the prime purpose of gamma is to adjust for differential arithmetical weighting of means and frequencies (e.g. so that a mean range of 135 to 250 is not considered in a raw form alongside one of say 1.456 to 3.567 ). In effect, gamma has the role in $\Delta g$ that the standard deviation has in quantitative measures of distance such as the CRL and Penrose's $C^2H$. The desirability of a world gamma of variation is that all distances in all future studies using the same attributes will be totally comparable, and relative homogeneity of distances of populations in any study area can be easily measured in relation to the maximum theoretical distance of 10,000 (achieved if two populations lie at the opposite extremity of world variation for every attribute).

In this study I have used instead the Kenya gamma of variation, which is the mean or frequency range of variation for a particular attribute at the minimal popualation sample level (MALE and FEMALE). Thus for example, the Kenya gamma for total arches for the left hand is 12.3 (Kabras females) minus 0.0 (Gabbra males). Thus all the distances, both male and female, are strictly comparable when based on the same attributes.

A further problem associated with working out gamma in the present study is the relation of variation to sample size. Very small sample sizes are obviously going to show a very large range of variation by sheer chance, especially for dermatoglyphic attributes which show considerable variability between individuals.

I have not adjusted for sample size in the calculation of the Kenya gamma, as I could not decide on what constituted an acceptable lower limit for sample size. The main effect of this is that the gammas are likely to be greater than the true underlying biological range of variation, especially for the qualitative traits. This means that the distances are correspondingly smaller than would have been the case if smaller gammas had been used. Nevertheless relative distance is not directly related to the arithmetical size of the distances. Indirectly, however, there could be differences in relative distance caused by the fact that those attributes most susceptable to sample size fluctuation will produce the largest gammas, and will correspondingly weight less in the final distances. As sample size improves, and gamma becomes increasingly a better reflection of true biological variation, the attributes most sensitive to sample size progressively weigh more in the distances. In the distance matrices in the present study, qualitative attributes such as digital and palmar patterns, show the greatest range of variation (see Table 5.3) and therefore individually weigh less in the distances vis a vis quantitative attributes.

The final problem concerning the calculation of the $\Delta g$ matrices is choice of attributes. The primary consideration that must be kept in mind is that the analysis of overall structure should reflect the total spectrum of phenetic variation, and that one of the primary reasons for choosing $\Delta g$ is that it is one of the few measures that can incorporate all attributes irrespective of whether they are quantitative or qualitative. In the analyses that follow I adopted the basic procedure outlined in the previous

chapter of considering both summary and non-summary variations. The reasons for preserving this division are the same as those listed in section 2 of Chapter 5.

In the subsequent investigation, I calculated four matrices of $\Delta g$ distances. The first two correspond to distances based on the same 26 summary attributes as in the PCA's 3 and 4 in Chapter 5.3, one matrix for each sex. The third and fourth matrices are male and female distances based on 110 non-summary attributes, consisting of all the digital and palmar attributes included in the non-summary PCA analyses (Chapter 5.3, Analyses 5 to 8) excepting two, $III^{t}$ on each hand, which are methodologically imprecise attributes. The atd angle moreover, as before, was also excluded owing to its propensity to vary significantly with age (see Chapter 4).

### 6.1.2 Choice and implementation of techniques to analyse phenetic distances

A matrix of genetic of phenetic distances is a complex range of inter-population comparisons, forming a representation of points in multidimensional space. If distances are calculated between more than a very low number of populations, trends of relationship between the populations become increasingly difficult to interpret with any objectivity as the number of dimensions rises. The purpose of techniques that analyse similarity or dissimilarity matrices is to portray the number of population points in as few dimensions as possible, and to delineate trends of affinity clearly and objectively.

There are many statistical techniques now available that can analyse matrices of distances. Conventionally the range of techniques are divided into two divisions, a) clustering techniques (both hierarchical and non-hierarchical), and b) ordination techniques. Cluster analysis basically seeks to partition the units into progressively smaller groupings through divisive or agglomeritive algorithms. Divisive algorithms begin by treating all units as one cluster, and split them progressively into smaller clusters until the original units are unique. Agglomeritive algorithms begin with the unique units and progressively amalgamate them into bigger clusters, linked hierarchically, until only one cluster is left which incorporates all units, (see Sneath and Sokal 1973; Everitt 1974 for detailed descriptions of the range of possible techniques). Ordination in contrast is the condensation of the multidimensional scatter of points onto few dimensions without partition.

In a descriptive analysis of the overall multivariate structure of biological variation, cluster analysis is not the best type of technique to use. As Hiernaux points out:

> "the fact will remain that cluster analysis tends to find clusters even when the data do not lend themselves to division into discrete groups. Cluster analysis therefore, may not be applied automatically: it only makes sense if the points in the hyperspace are grouped into clusters in the totality of the scatter. If the scatter of points is random, or clustered only in some sectors, cluster analysis is misleading."
>
> (Hiernaux 1972:105)

Ordination analysis is preferable as a preliminary descriptive tool, as it does not treat the units as automatically forming clusters. If clusters exist, ordination analysis will show them as concentrations of points in close proximity separated from other points or groups of

points. If no clusters exist, the scatter of points will appear as a continuous scatter with no discrete concentrations of points.

Ordination solutions can be obtained by producing scattergrams of canonical variates in discriminant function analysis, or scattergrams of component scores when PCA is applied to type B data matrices. When the data is in the form of raw distances however, techniques are needed to transform the distances into equivalents of canonical variates or principal components. The best known techniques that do this are a) metric scaling or principal coordinates analysis (Gower 1966) and b) non metric multidimensional scaling (Kruskal 1964a; 1964b).

a) Principal coordinates analysis (PCO)

PCO analysis is similar to PCA, in that the distances are transformed into latent roots and vectors, the elements of the latent vectors giving the coordinates which when graphically scattered produce an ordination solution, (Everitt 1978:18-19). Like in PCA the first latent vector accounts for the greatest portion of variance, the second latent vector the next highest portion of variance and so on. Thus the greater the variance accounted by the first two vectors, the less distorted is the low dimensional solution of the scatter of coordinates for these vectors. Detailed description of the method is given by Gower (1966).

b) Non-metric multidimensional scaling (NMMS)

NMMS acts on ranked distances rather than the absolute values of distances, and reduces dimensionality through an iterative function that minimises a measure of "stress" and seeks to maximise the goodness of fit of the low dimensional solutions (Kruskal 1964b).

The detailed descriptions of this mathematically complex technique are given by Kruskal (1964a; 1964b).

Both PCO and NMMS have been used with success by physical anthropologists. (For instance Hiernaux 1972 for PCO; Rightmire 1976; Dennis 1977a for NMMS). Gower (1972), in assessing the two methods states that:

> "Although (NMMS) uses less information than metric methods like principal components and canonical variates analysis, and might, therefore be thought superior to other methods of reducing dimensionality, it has been my experience, and that of many other workers, that the resulting low-dimensional configurations differ little from their metric counterparts. In this respect, a comparison of the two types of technique is similar to the comparison of the CRL distance with Mahalonobis's distance, where the choice of method has been found to make little difference in practise, although the agreement cannot be adequately explained."
>
> (Gower 1972:14-15).

In view of Gower's assessment, it is clear that both PCO and NMMS are both adequate techniques for analysing matrices of biological distance, although NMMS is theoretically preferable. In the present study the choice of PCO over NMMS was purely pragmatic. The computer facilities at the University of Stirling, where the analyses were performed, included access to a program to compute PCO, but not NMMS. PCO was thus calculated through access to a program written by Gower in the GENSTAT package (Gower, 1977).

## 6.2 General structure of overall differentiation between the minimal population samples

The examination of general structure in this study is as we have seen, based on two approaches. A) The analysis of the matrices of $\Delta g$ distances, mainly through principal coordinates

analysis and B) through the graphical display of component scores. The evidence is contained in the following tables (Appendix 3):

| | |
|---|---|
| Tables AP3:1 & AP3:2 | The Δg matrices for males and females based on 26 summary attributes. |
| Tables AP3:3 & AP3:4 | The Δg matrices for males and females based on 110 non-summary attributes. |
| Tables AP3:5 & AP3:6 | The results of principal coordinates analysis on the summary Δg matrices. |
| Tables AP3:7 & AP3:8 | The results of principal coordinates analysis on the non-summary matrices. |
| Tables AP3:9 to AP3:12 | The component scores* for males and females based on the summary traits. |
| Table AP3:13 & AP3:14 | The component scores for males and females together based on 52 male and female summary traits. |
| Table AP3:15 to AP3:18 | The component scores for males and females based on 80 digital non-summary traits. |
| Table AP3:19 to AP3:22 | The component scores for males and females based on 32 palmar non-summary traits. |

6.2.1 The nature of overall structure: The evidence from phenetic distances

In order to find out whether significant clusters are likely to exist in the data, Hiernaux adopted a technique first proposed by Flight (1967), in which the distribution of all the distances in a matrix is plotted. As Hiernaux points out:

> "if the scatter of points is random, the distribution of the distances will be unimodal. In the case of clustering, on the contrary, there will be an excess of small or high distances, and the distribution will be bi- or multi-modal."
>
> (Hiernaux 1972:105)

* Component scores for unrotated and rotated solutions.

Hiernaux found the distribution of 5,050 distances between 101 sub-Saharan populations to be strongly uni-modal, and he concluded that "there is no evidence of a systematic clustering of points", (1972:105).

The simplicity and effectiveness of this method clearly commends it for the analysis that follows. However it is no more than a rough and preliminary guide to the nature of the multivariate structure implied in the distance matrix. To begin with, the plotting of the distribution is to some extent an arbitrary procedure, as raw distances tend to be relatively unique. (That is to say that, for example, very few $\Delta g$ distances in a matrix will be, say, exactly 236). In order to draw up a meaningful distribution, the raw distances must be pooled in some way. To a significant extent, the shape of the distribution depends on the criterion of pooling. The larger the units of pooling, the more homogeneous the distribution will become, as the boundaries between antimodes are eradicated. The conventional practise of pooling in units of fives, or tends, fifties, hundred etc., is arbritrary and far from satisfactory. Nevertheless, the empirical juggling needed to overcome this problem is seldom worth the effort.

The application of this method to the four $\Delta g$ matrices (AP3:1 to 4) turned out to be simpler than anticipated. The $\Delta g$ program was amended to output the distances sequentially one per line into a file, which was then treated as a normal SPSS data matrix. The distances were then transformed using SPSS COMPUTE procedures. Firstly, they were square rooted to obtain the true distance rather than the squared distance ($\Delta g$ like all distance measures is a squared distance). Secondly, they were rounded off into whole

numbers to produce "unit" scores. Thirdly, the square rooted distances were divided by ten and rounded off to produce pooled categories. Frequencies were then calculated for the unit and pooled category scores, and the histograms were compared. I found that the distributions of the pooled category scores differed little from those of the unit scores in shape, though of course the unit score distributed was more flattened. I therefore decided to present the results of the category scores only, as this is the level of resolution used by Hiernaux (1972:105). His graph of true distances between 101 populations from sub-Saharan Africa is reproduced in Figure 6.1. It should be noted that his X axis labels cannot possibly be the square root of distances as he claims, as the squares of the numbers at the middle to upper end of the range exceed 10,000. Finally, my graphs are calculated from percentages, not real numbers, to permit an exact comparison of my four series of distances.

The distributions for the male and female series of distances based on summary traits are given in Figure 6.2, and the equivalent for non-summary traits in Figure 6.3. The general shapes of all four distributions are almost identical with these of Hiernaux's distribution, showing negatively skewed unimodal curves. The presence of only one mode and the relatively narrow spread of the majority of distances either side of it suggests a marked central tendency with an absence of clusters. The rapid rise to the mode testifies to the very small proportion of low distances. The rapid fall from the mode at the positive end of the curve, followed by a gradual and lengthy "tailing off", points to the existence of a few populations that have large distances from the majority of the other populations.

A comparison of the four detailed distributions (Figures 6.2 and 6,3) reveals some subtle differences between them. The two "summary" distributions for instance show a greater range of distances than the two "non summary" distributions, but parodoxically, a mode located at a much lower distance. (That is to say, in the summary distributions the distances range from less than 10 to seventy compared to less than ten to fifty in the non summary ones, while the summary modes lie at a distance of 20 compared to 30 in the non summary ones). There are moreover, some differences between the sexes in the summary distributions, though not in the non-summary ones. As Figure 6.2 shows, the females exhibit a more condensed central tendency around the mode which is higher and more peaked than in the males, and the "tail" of large distances is longer in the females. Thus the summary traits serve to homogenise distances at the centre, and emphasise peripheral distances, especially in the females, compared to non summary traits.

Owing to the arbitrary nature of pooling however, the trends manifest in the plotting of the distribution of distances cannot be totally relied on. In order to obtain an alternative perspective, I decided to make use of distances from the centre of the scatter of points calculated during principal coordinates analysis. These are listed in Tables AP3:5 to AP3:8 for the four distances matrices. The distances from the centroid are all in $\Delta g$ unit scales, directly equivalent to the squared distances contained in the basic matrices. The distributions of distances from the cnetroid were plotted as circles radiating from the central point of the distribution using a private program utilising GHOST sub-routines.

The resulting plots are presented in Figures 6.4 to 6.7 for each series of distances. The results confirm the trends of the previous analysis of histograms, but are a lot clearer. Firstly, it is noticable that in all four distributions there is a marked central tendency with a majority of the samples radiating close to the centre. There is no marked separation of very low distances from the centroid from the main central mass, but at the opposite extreme there exists a small number of samples that lie far from the centre and which are clearly isolated from the central group. It is these that obviously account for the "tails" in the histogram distributions.

The non summary distributions show little sex differences in shape, and are distinguished from the summary ones by a) exhibiting a much lower range of distances, b) being more spread out from the centre and c) demonstrating much lower distances from the central mass to the extremes. The summary distributions however demonstrate a marked sex difference. The females are remarkable in showing an extremely "tight" central mass representing a majority of populations, while at the same time the extremes are more separated than in the males.

Finally, in the plots I have specified who the peripheral populations are. There is a noticeable absence of concordance between the four plots, especially between the males and females. Only one population, the Kony, lie at the extremity on all four plots. This pattern of discordance is further evidence of significant overall complexity of structure. (The precise significance of these isolates will be discussed later).

6.2.2 The nature of overall structure: further evidence from principal component scores

Before discussing the significance of the findings in 6.2.1, I wish to depict overall structure from a different angle. One could argue for instance, that as the preceding distributions are based on the same type of technical results (i.e. Δg distances), they might well present a distorted picture, in particular as the attributes are no decorrelated in the calculation of the distances. The use of principal component scores provides an alternative approach picturing structure in terms of rotated decorrelated components.

The principal component scores as they are presented in Tables AP3:9 to Ap3:22 closely resemble univariate variables in format. Each component emphasises the variation of only those raw attributes that weigh significantly on it. Each different component thus only relates to a particular aspect of the total variation. In order to picture overall structure, it is necessary to somehow combine all components that result from a particular analysis.

I resolved this problem through regarding it primarily as a graphical rather than a computational problem. The component scores possess an essential property: they are all standardised scores whichare on the same scale of measurement from component to component. In order to picture overall structure therefore, all that needs to be done is to link the scores of each population by a straight line on a graph with the number of the component on the X axis and the range of scores on the Y axis. The resultant plot will thus consist of a series of lines linking components, each line representing a population or sample. If clusters exist,

the pattern of lines will consist of a series of separated bands of zig-zag lines, rising and falling together, each band being a cluster. Overlapping but nevertheless discernable clusters will manifest themselves as dense bands separated by less dense webs of lines at the boundaries. An absence of clustering will show itself as an untidy mess of criss-crossing lines.(1)

One limitation of this graphical method is that it is impractical to picture more than a limited number of components on a single graph. This means that it is not really feasible to present all 28 digital and palmar non summary components. Not to do so however, would mean that only a portion of the total variation would appear on a single graph. A pilot run on the males however, revealed that the general structure was not radically different from that pictured by summary traits. I decided therefore, to include only plots of summary distribution in this analysis.

The results then consist of three plots. The first depicts the most complete picture of overall structure, being based on the unrotated component scores produced from the analysis of 52 male and female traits (see Chapter 5.3). The second and third are based on unrotated component scores for independent male and female series based on twenty six summary traits. The graphs were plotted using private programs using GHOST sub-routines.

---

(1) This graphical method is an unsophisticated and basic version of Andrews plots (Andrews 1972), illustrated by Everitt (1978:81-87) on Mahalonobis et.al. (1949) Indian data. Although Everitt sent me a program to calculate and graph Andrews plots, it proved incompatible with the graphical procedures in use at the University of Stirling's computer, and proved too difficult to adapt with the time at my disposal.

The three distributions (Figures 6.8 to 6.10) reveal a marked absence of distinct bands. The lines are rather inter-woven resembling an untidy web. The greatest concentration of points on each component is at the central portion of the range, although some multi-modality is in evidence by gaps between concentration of points. It should also be noted that few populations run across the central areas on all components. The usual pattern is one of either a zig zag from periphery to centre of conversely from centre to periphery, and from periphery to the opposite extreme. The outermost points on each component are few and represent isolates. No sample however runs consistently from extremity to extremity.

A comparison of the male and female plots also shows evidence of the trend observed in the previous analyses based on distances, that the females exhibit a stronger central tendency in the scatter of points. In Figure 6.10 this can be seen as a thicker band of lines near the centre. Extreme points indicating outlying populations are not as pronounced relative to males as they are in the previous distributions from the analysis of distances.

### 6.2.3 Discussion

The principal conclusion to be drawn from the anlyses in 6.2.1 and 6.2.2 is that the results are consistent with Hiernaux's (1972) findings on the pattern of differentiation in sub-Saharan populations which were based on a much wider variety of attributes. The structure here, like that observed by Hiernaux, resembles a mosaic rather than a series of distinct clusters. Moreover, it is also possible to infer from the strong central tendency of the distributions and the relative rarity of very small and very large

distances that overall multivariate clines are not apparent in the structure. This does not mean that clines are totally absent. For instance they could be in evidence between populations from restricted localities or zones. All that the present results show is that macro-clines involving majority of populations (such as east to west or north to south gradients) are not present.

The analysis of structure based on distances demonstrates a strong unimodal central network of relationships with a few separated peripheral points. This indicates relative homogeneity in the central swarm and the presence of a few isolated populations. The analysis of PCA scores however, qualifies this picture. The isolates in reality are extreme on relatively few components. The explanation for this seeming contradiction lies in the relative weighting of raw attributes in the principal components and distance approaches. In the calculation of $\Delta g$ distances attributes are not decorrelated. Thus for example, extreme raw scores for attributes such as radial ridge count will be replicated to a varying but significant extent in a large range of other digital attributes such as ulnar ridge counts, digital pattern intensity, whorls arches and loops. Moderate raw scores on other unrelated attributes such as palmar attributes will not compensate for the heavy weighting of the extreme digital scores. The population affected will thus have large distances from other populations and will appear as a clear isolate. In the principal components analysis however, extreme scores on the large range of digital attributes will produce a heavy score on only one or two digital components when the attributes are decorrelated. The analysis of

principal component scores introduces a cautionary note into any attempt to explain the isolates in terms of drift. It shows that the fundamental question that should be asked is not why the population is isolated, but by rather why it exhibits extreme values on specific components.

Finally, as I have already remarked, the distributions of distances show differences a) between the series based on summary traits compared to those based on non-summary ones and b) in the case of the summary distributions, between the males and females. The greater homogeneity of samples at the centre of the distribution of distances based on summary traits, particularly in the females, is a particularly puzzling trend. One would expect the distributions of the summary series to be more homogeneous at the centre and more isolated at the extremities, because of the greater stability of the summary traits regarding sample size. In other words, true biological directional trends would be expected to be emphasised in the summary distances. Thus for instance similar populations should become more homogeneous as random fluctuation lessens and correspondingly dissimilar populations, more distant. This hypothesis however is contradicted when one compares the males and females "summary" distributions. The male sample sizes are significantly larger than the females, yet it is the females who show the most condensed homogeneous relationships at the centre and the most diffuse at the periphery. This is yet further evidence of the complex relationships between the male and female series in this study.

## 6.3 Detailed multivariate relationships between population samples

### 6.3.1 Introduction

So far, the descriptive analysis has shown not only that there is considerable taxonomic discordance in patterns of differentiation between phenetically unrelated attributes, but also that, probably as a result of this, the overall structure is of a nature that does not contain any major clusters or clines. According to Hiernaux, this type of differentiation is incompatible with inquiry that seeks to apply a taxonomic or phylogenetic interpretation on the data. He emphasises however, that a non-taxonomic approach can be profitable in helping to "shed light on the evolutionary processes which led to the present diversity." (Hiernaux 1972:110).

In applying the non-taxonomic approach to the analysis of the sub-Saharan data, both Hiernaux (1972) and Rightmire (1976) have explored the possibilities of further analysing the basic $\Delta g$ matrix of distances through the application of ordination techniques. Thus Hiernaux (1972) applied principal coordinates analysis to the distances between 15 Kivu/Rwanda populations, and Rightmire (1976) applied non-metric multidimensional scaling to a set of distances between a wider selection of 60 of Hiernaux's original 101 populations (Hiernaux 1968[a]).

The rationale for applying ordination analysis is that although the genetic or phenetic distances between populations give the most detailed and complete picture of multivariate relationships, they nevertheless contain "much more information than can be digested through simple inspection." (Rightmire 1976:445). In a more

abstract sense the distances can be viewed as points scattered in multidimensional geometric space. The purpose of ordination is to "represent the multidimensional scatter in a space of any desired low number of dimensions", (Hiernaux 1972:103). The fewer the dimensions that the data can be condensed to, the easier it becomes to discern and interpret trends. A most effective demonstration of the power of ordination analysis is given by Everitt (1978). He applied principal coordinates analysis to a matrix of real geographical distances between 48 British cities. When this complex and indigestible scatter of distances was reduced to a low number of dimensions, and the first two principal vectors were plotted as a scatter diagram, the resulting configuration was a close approximation to the map of the British Isles, (Everitt 1978:20).

From the examination of the way population points scatter at a few dimensions both Hiernaux and Rightmire have investigated many detailed issues pertaining to the diversity of the respective areas examined, including the deduction of zonal clines, common ancestry, recognition of isolates and relating linguistic and biological affinity. Rightmire in particular presents a very detailed analysis of the implications of the observed scatter of populations at I, II, and III non-metrically scaled dimensions, (Rightmire 1976).

From the evidence of precedents, therefore, it seems that ordination analysis is an essential approach. Given that the basic overall structure of differentiation is here very similar to that pictured by Hiernaux, one should expect that the examination of the low dimensional scatter of points is going to be as illuminating

in the present study as it has apparently been for Hiernaux and Rightmire.

I have however some reservations. Firstly, there is the problem that in reducing the scatter of points to few dimensions, considerable distortion can result. In applying non-metric multidimensional scaling, distortion can be readily assessed through Kruskal's measure of stress which can be interpreted on a scale of adequacy (Kruskal 1964). Similarly, in principal coordinates analysis the amount of variance accounted for by the first few dimensions can be viewed as an indication of the power of the low dimensional solution. The implications of stress however, are never dwelt upon in anthropological investigations. I have for instance, not come across a study which states that the stress or distortion found was "unnacceptable" for the objective interpretation of the low dimensional solution. The question of what is actually being represented by the first few dimensions when distortion is high is yet to be seriously tackled.

Secondly, there is the problem of different solutions given by various ordination techniques. When distortion is low (as in the Everitt example given earlier, where over 90 per cent of the variance was accounted for by the first two vectors) different techniques will usually give very similar results. When distortion is significant however, it is generally the case that solutions from differing ordination techniques will often differ radically from one another. Which is one to believe, the ordination of discriminant functions, principal components (rotated or unrotated), principal coordinate vectors or non metrically scaled vectors? In

addition, there is the possibility that male and female series will not be concordant in their low dimensional representation. Even subtle differences between these possible solutions are detrimental to objective interpretation.

I take the view that before one can start interpreting low dimensional solutions to further the investigation of specific hypotheses that seek to explain the observed differentiation, it is necessary to assess how far solutions as given by different techniques and between males and females are concordant. If sufficient concordance is not forthcoming, then it is pointless to attach too much credence to the results of such procedures.

### 6.3.2 Analysis

In the initial stages of the investigation I explored the possibilities of several ordination and clustering techniques. The $\Delta g$ matrices of distances were subjected to principal coordinates analysis, and to Ward's hierarchical clustering. Secondly, component scores for all principal components analyses described in Chapter 5.3, both for the rotated and unrotated solutions were produced. For each analysis, the PCO vectors and PCA components were graphically ordinated using private graphical programs utilising GHOST subroutines. All permutations were obtained for the first four vectors or components in each analysis (e.g. vector 1 vs 2, 1 vs 3, 1 vs 4, 2 vs 3, 2 vs 4, 3 vs 4).

All these permutations resulted in 108 scatter graphs (each a low dimensional solution) and 4 Ward's cluster diagrams (one for each male and female $\Delta g$ matrix based on summary and non summary attributes). Furthermore, I experimented with three

dimensional graphical representations of the first 3 PCO vectors for each male and female analysis based on 26 summary attributes. The three dimensional graphs were produced using PLUTO, a program devised by chemists to portray 3 - D molecular structures. Even this condensed procedure however, was prolific in its output, as each 3-D solution can be viewed from many angles (above, below, X, Y or Z axis etc). Moreover, with 57 male and 53 female samples the graphs were messy, and complicated central clusters had to be magnified and reproduced individually. Most of these combinations were looked at interactively on a TEKTRONIX VDU, and no hard copy was made.

It is evident from this review of results that it is possible to produce a vast array of different low dimensional graphical solutions, each graph giving a unique representation of the ordering of population points for those dimensions considered. A sample of these graphs is presented in Figure 6.11 to 6.28. An inspection of the graphs reveals general concordance in only a very few areas in the extreme peripheries. Thus for example, the Kony (pop 13) are extreme on the first dimension in both male and female series in both PCO and PCA solutions, but are not strongly differentiated in other dimensions. The Digo (59) are very extreme in the females at the opposite polarity to the Kony in the females, but are not extreme in the males. The Somali (1), Rendille (2), Gabbra (3), Burji (7) and Samburu (8) occupy the same general areas in the first two dimensions of both male and female PCO and PCA solutions. On the whole however, there is marked discordance in the ordering and scatter of specific populations at different

dimensions and solutions. Even concentrating on the ordination of the first two vectors or components (which account for the largest portion of variance), there is marked lack of concordance particularly between male and female configurations.

For instance, in the male PCO analysis on the distance matrix based on 26 summary traits vector 1 vs vector 2 (Figure 6.11) (in short male PCO(26)v1:v2), there is a tight cluster at the lower middle portion of the graph consisting of pops 34, 52, 22 and 57. In male PCO(110)v1:v2 (Figure 6.17), the tight cluster disappears, although the populations still scatter in the same relative area of the graph. In contrast in both female PCO(26)v1:v2 (Figure 6.14), the female PCO(110)v1:v2 (Figure 6.20), the populations in the above male configurations that cluster closely are completely dispersed. This discordance between different solutions is typical for a majority of seeming clusters on individual graphs, especially for those populations in the central area of the graphs which are least discriminated.

It became obvious while I tried to make sense of this wealth of different discordant configurations, that a more objective method other than direct inspection was needed to compare the affinities of populations on different vectors and components from male and female and summary and non-summary attribute series. I assessed discordance by calculating Spearman rank order correlations between all combinations of:

a) The first three PCO vectors from analyses based on both summary and non summary attributes males and females: (PCO(26)v1 to v3; PCO(110)v1 to v3; males and females).

b) Unrotated component scores for the six significant components derived from PCA analyses based on summary traits (males and females) as described in Chapter 5.3 (UNC1 to UNC6 for each sex).

c) The equivalent of (b) for rotated component scores. (C1 to C6 for each sex).

These vectors and components were chosen to provide a representative picture of possible discordance within and between the methods of ordination. It would have been a very lengthy exercise to have inter-correlated all vectors and components for both summary and non summary attributes.

The results containing inter-correlations for 36 variables (3 PCO vectors for summary traits, 3 PCO vectors for non summary traits, 6 rotated and 6 unrotated PCA component scores for summary traits only, doubled for the male and female series) nevertheless take up too much space to report in full. They are however, summarised in Table 6.1. It should be carefully noted that all unspecified combinations exhibit very low correlations.

i) Intra-sex correlations between PCO vectors from distances based on 26 summary (PCO26) and 110 non summary (PCO110) attributes

In both sexes PCO(26)v1 correlates very highly with PCO (110)v1 (nearly 0.9). In other words in spite of the fact that PCO(26)v1 accounts for twice as much of the variance (c40%) as PCO(110)v1 (c20%) in both sexes, the relative ordering of population is very similar. Only 20 per cent of the variance between the vectors is not in common. This however is enough to produce subtle differences in the ordinated graphs.

In contrast, there is less concordance between PCO(26)v2 and v3 and PCO(110)v2 and v3. In the females moreover, PCO(26)v2 is not equivalent to PCO(110)v2 but to PCO(110)v3 and conversely PCO(26)v3 is equivalent to PCO(110)v2. There is a crossover in the females but not in the males. In both sexes however, it is PCO(26)v2 that shows the lowest correlation with its equivalent (only 0.5 in the males and -0.5 in the females), while PCO(26)v3 correlates higher at -0.7 in the males and 0.7 in the females. When graphs between v1 and v2, 1 and 3 etc., are plotted, it is thus not surprising that significant differences exist in the configurations between the summary and non summary series as the combined effects of the lack of concordance between each pair of vectors is considerable. Nevertheless, there is sufficient concordance to permit the discernment of broad similarities between the two series.

ii) Intra-sex correlations between PCO vectors and unrotated component scores (UNC1 to UNC6).

In this section, I shall concentrate on comparing the unrotated components just with the PCO(26) vectors as both types of solution are based on the same data (summary attributes). The relationships between UNC1 to UNC6 and PCO(110) vectors are moderate and complex, except for the first components and vectors, as can be seen in Table 6.1

There is almost total concordance between PCO(26)v1 and UNC1 approaching unity in both sexes. Concordance for vectors 2 and 3 with UNC2 and 3 respectively are also extremely high, particularly in the males. Sex differences are evident for

PCO(26)v3 vs UNC3 which is lower and positively correlated in the females. The results thus show that there is a very strong similarity between PCO analysis on a matrix of Δg distances and unrotated PCA analysis based on the same attributes, but that subtle differences exist for the lesser components. It should be noted moreover, that the percentage variance accounted for by each component (see Table 5.8) is almost identical to that explained by each PCO equivalent vector (see Tables AP3:5 and AP3:6). Consequently we should conclude that the whole complicated business of calculating Δg and applying PCO analysis was superfluous, as all but a small portion of variance is identical in the two solutions. Moreover PCA analysis is not only simpler to compute, but also indicates which attributes are directly loading on each component.

iii) Intra-sex correlations between PCO(26) and rotated PCA component scores (C1 to C6), and between rotated and unrotated component scores.

The rotated solution as we saw in the previous chapter differs appreciably from the unrotated one. This is clearly reflected in Table 6.1, where PCO(26) v1 to v3 and the largely equivalent unrotated components UNC1 to UNC6 exhibit complex patterns of relationship with C1 to C6. At best, the correlations are lower than 0.8 and are in some cases considerably lower. As the order of female and male rotated components differs (e.g. C1 males is really equivalent to C2 females - see Table 5.10 in the previous chapter), the apparent lack of agreement between the male and female combinations (Table 6.1) is spurious. There is in fact strong equivalence between the sexes.

The comparison of unrotated components and PCO vectors with rotated components shows that there is far more discordance between the unrotated and rotated components than there is between unrotated components and PCO vectors. The main difference between the rotated solution and the others is that it is far less generalised, being more sensitive to subtle differences between groups of taxonomically related characters (see previous chapter).

iv) Inter-sex correlations between PCO vectors and PCA components

There are none listed in Table 6.1 as all combinations between the sexes exhibit very low correlations indeed, usually below 0.1. The highest correlation is between PCO(26)v2 males versus UNC3 females and is only -0.34. All ordination scatter graphs will thus differ radically between the sexes.

## 6.4 Discussion

The above analysis raises some interesting points. Firstly, the strong equivalence between PCA and PCO unrotated components and vectors puts into question the advantage of calculating and analysing genetic and phenetic distances. In the present study, it is now obvious that the PCA analysis not only reproduces all essential information given by the analysis of the $\Delta g$ distances, but has the additional advantages of easy computation and of being able to relate components to individual attributes.

Secondly, there is the implication of distortion and discordance. In a non-metric multidimensional scaling solution, which produces very similar results to a metric solution such as PCO, (Gower 1972:14) distortion is seen as "stress", or the extent

of monotonicity between the new distances in low dimension and the original rank order (Rightmire 197 :446). Here however, distortion can be envisaged rather in terms of how much variance is accounted for by the first few components or vectors.

In a sense, PCO and PCA low dimensional solutions are not really distorted representations of multidimensional space, but undistorted condensed aspects of the variance. This can be seen by considering what is actually being measured by each PCO vector. As each vector is closely equivalent to an unrotated component, then we can assume that a PCO vector accounts for the same variance as a PCA unrotated component. We can moreover, easily interpret what each component is measuring. Unrotated component 1 for summary attributes has been described in detail in the previous chapter, and is a generalised digital component in both sexes. Hence PCO(26)v1 in both sexes also is a general measure of digital variance, as the correlation between PCO(26)v1 and UNC1 is close to unity. UNC2 in contrast is a generalised palmar component, with an emphasis on total and hypothenar pattern intensity. It also accounts for much less variance than UNC1. PCO(26)v2 is consequently mostly a generalised palmar component as it correlates highly with UNC2 in both sexes. What are we measuring then when we produce an ordination scatter graph of PCO(26)v1 vs v2? We are effectively portraying the scatter of population points in two dimensional space for components of variance expressing some, but by no means all, digital variance on the X axis and some, but not all palmar variance on the Y axis. It is almost a compressed univariate scattergram.

This can hardly be viewed as a faithful representation of "overall" affinity between populations. This is a very important point, as when we see populations such as the Kony (13) appearing as isolates on the scatter of the first two vectors, it does not mean that they are necessarily true isolates. A look at the first PCO vectors in Tables PA3:5 and AP3:6 reveals that the Kony appear as isolates only on the first component. As this is a digital component, then it means that they just have unusual means and frequencies for some digital attributes. (In fact the Kony have unusually low ridge count means, and very low frequencies of whorls in both sexes, see Appendix/1). For palmar attributes, they are quite undistinguished. In explaining the multivariate differentiation therefore, the scattering of vectors and components, rather than giving an accurate "overall" picture, is only marginally more helpful than looking at scattergrams between univariate attributes.

The realisation of the basic relationship between PCO vectors and unrotated PCA components, and by extension, between PCO vectors and specific attributes, brings us round back to considering the implications of the discordance between taxonomic attributes. The discordance between the ordination solutions based on summary and non summary attributes, and between the sexes, is directly a consequence of the discordance between the underlying taxonomic attributes that are reflected in the PCO vectors. Thus in asking why do populations not scatter in the same way when the first two vectors are being examined, we are really asking a) when are the rankings of digital means and

frequencies so different between the sexes and b) when are the palmar rankings[9] so different? These are questions that to a large extent can only be answered univariately.

I have thus effectively shown that if the basic structure of attribute relationships is heavily discordant, as I demonstrated in Chapter 5, then the overall structure will reflect this discordance, by showing not only an absence of major clines and clusters, but also discordant patterns of points when they are represented in fewer dimensions. This discordance makes it impossible in my opinion, objectively to utilise low dimensional solutions for the pursuit of detailed explanatory hypotheses as Rightmire and Hiernaux have done, unless the precise causes and implications of the discordance are known. In effect, I have gained very little in trying to explain the differentiation between the minimal population units by examining the detailed "overall" structure.

I do not wish to suggest however, that the analysis of "overall" distance through ordination is necessarily always a waste of time. My essential point is that it only assumes meaning if it can be demonstrated beforehand that significant taxonomic concordance between attributes exists over and above that normally expected through phenetic concordance between attributes.

# PART THREE: EXPLANATORY ANALYSIS.

# CHAPTER 7

## Sex Differences in the Variation of Dermatoglyphic Attributes

### Introduction

Sex differences in dermatoglyphic attributes are usually regarded as being primarily of methodological importance in a regional human biological survey. Dennis (1977[a]:218ff) for example, sees the presence of marked sex differences as an important reason why the sexes cannot be pooled to increase sample size. The treatment of the sexes as separate series is thus a necessary methodological inconvenience. In the present study there are also appreciable sex differences in most dermatoglyphic attributes. The separate consideration of male and female samples in the descriptive analysis was thus a direct acknowledgement of the methodological necessity for teating them as separate. This was especially important owing to the imbalance of sample sizes between the sexes for many samples. The pooling of the sexes could have been achieved through the equalising of sample sizes, which would have resulted in equal weighting. This would have resulted however, in a drastic reduction of overall sample numbers, and there would have still been some imbalance owing to the differential occurrence of missing values.

The methodological importance of sex differences however, is not the primary point of interest in this chapter. Sex differences here are rather regarded as a fundamental characteristic of dermatoglyphic variation between sampled populations that needs further analysis and discussion. The descriptive analyses (Chapter 5 and 6) have raised two important questions. Firstly, there is the

enigma that while the sexes show very similar intra-sex patterns of phenetic and taxonomic concordance between attributes, they nevertheless are radically different in the inter-sex rankings of populations for equivalent attributes, (see Chapter 5.3). Secondly, in the general patterning of overall structure, the females tend to be more homogeneous in the clustering of central populations and more heterogeneous in the scatter of peripheral populations (see Chapter 6.2), and there is considerable overall discordance in the position of populations from the centre of the multidimensional distribution (see Chapter 6.2). To seek an explanation for these trends it is necessary in the first instance to examine more closely the characteristics of sex differences for individual attributes at a macro level, where random fluctuation due to low sample sizes is not influential.

## 7.1 Sex differences at the total sample level

The total sample means and frequencies are listed by sex for all attributes in Appendix 1.1. The figures reveal marked sex differences for a majority of attributes. An objective assessment of the differences however, can only be achieved by testing for significance. The results of t-tests for quantitative attributes and chi square for qualitative ones are summarised in Table 7.1. They reveal the following trends;

### 7.1.1 Digital attributes

All digital attributes show sex differences that are highly significant on the right hand. On the left hand however, there is some variability. Radial, unilateral and absolute ridge counts and

finger patterns all show significance, mostly high significance. The ulnar ridge counts however, show significance only on LFU4 (P = 0.04) and for LFU1 there is very strong homogeneity (P = 0.774). For finger pattern intensity LD1, LD2, and LD5 are not significant, while LD3 and LD4 are very significant.

These trends of significance for the left hand contrast appreciably with the results found by Dennis (1977[a]) for the Yorkshire Dale population. He reported that all digital variables showed highly significant sex differences on both hands (Dennis 1977[a]:220).

The explanation for the differences between the Kenyan and Yorkshire series lies in the fact that asymmetry is reversed on certain digits between the sexes in the Kenyan population, but not in the Yorkshire one. As Table AP1.1.1 shows, the males have higher means on the right hand than the left for ulnar counts on the first and fourth digits, whereas the female means are much higher on the left hand on all digits. The males on digits two, three and five have means on the right hand only slighly lower than on the left, but the female differences on those digits while in the same direction, are much greater. The overall trend is very clearly displayed by the total ulnar count. In the males total ulnar ridge count on the right hand is 18.73 compared to 18.49 on the left (almost identical) while in the females the right hand mean is 14.80 compared to 18.12 on the left, an enormous asymmetrical sex difference.

The reversed asymmetry shown by the ulnar counts is also more weakly demonstrated by finger pattern intensity. Except for digit one where the right hand mean is bigger, the females consistently show higher means on the left hand, whereas in the

males the right hand mean is greater (digits one, two, four and five) and only slightly less on digit three. Again the summary attribute shows the trend very clearly. Total finger pattern intensity is much higher on the right hand for males (6.392 compared to 6.230) while in the females it is higher on the left (6.049 compared to 6.114).

The trends of reversed asymmetry in ulnar ridge counts and finger pattern intensity contrast with those exhibited by radial, unilateral and absolute ridge counts, where right hand means are greater in both sexes. To a significant extent, however ulnar ridge counts and finger pattern intensity reflect the relative occurrence of whorls. In Table AP.1.1.3 the frequency of total whorls is greater in the males on the right hand on all digits except digit three, whereas the female frequencies are in contrast lower on the right hand than the left on all digits. Ulnar loops are also interesting, showing reversed asymmetry on digits one and four, and much greater homologue differences for the females on the other three digits.

The Yorkshire Dale material (Dennis 1977a), shows that means and frequencies for ridge counts, finger pattern intensity and whorls are higher in the right hand in both sexes, in contrast with the Kenyan total sample just reported. This remarkable population difference points to the fact that we could expect to find variable sex differences, in both magnitude and direction from one population to another, a factor, which I shall show, is of fundamental importance for the understanding of inter-sex discordance in the present study.

7.1.2 Palmar attributes

A majority of palmar attributes are also significantly heterogeneous for sex. As Table 7.1 shows only the bc ridge count on the right hand is not significant for the palmar ridge counts. All other palmar attributes show high significance except for thenar attributes on the right hand and hypothenar attributes on both hands. The hypothenar area is hence unique in the present spectrum of attributes as containing the only group of attributes that consistently shows no significant sex differences on both hands. (t'' and $III^{t}$ also show no significance, but they are uncommon traits).

7.2 Sex differences at the macro population level

I have established so far that not only are there highly significant sex differences at the total sample level for a majority of digital and palmar attributes, but also that there are major differences in the magnitude and asymmetry of some sex differences between the Kenyan and Yorkshire populations. The next issue to consider is whether sex differences vary between populations within Kenya. In order to achieve clear results free from the possibility of random flucturations in means and frequencies due to small sample numbers, I decided to concentrate the analysis on sex differences between the ethnic categories (see Chapter 5.1) which as Tables AP.1.2.1 to 1.2.9 reveal, show major differences for many attributes, especially between the Cushitic category and the Nilotic and Bantu Categories. For the purposes of the present analysis the outstanding characteristic of these pooled populations samples is their large sample sizes.

7.2.1 Summary of inter-population trends between ethnic categories and differences between the male and female series

If sex differences between the means or frequencies of attributes remain constant from one population to another, and male and female samples for each population are drawn from the same gene pool, one would expect, given adequate sample sizes, that rankings of inter-population differences for a particular inherited attribute would be very similar in each sex series. I will begin the analysis therefore, by examining how far the male and female series replicate trends of inter population variation at the ethnic category level, not only in overall taxonomic ranking, but also in the proportion of significant differences that occur. Should there be major differences between the sexes on one or both of these points, the analysis will procede to investigate fluctuations in the magnitude of sex differences between the populations for individual attributes.

It becomes evident from the findings of Tables A.P.1.2.1 to AP.1.2.9 which provide the raw univariate means and frequencies for each ethnic category by sex, that the Cushitic category is differentiated from the Bantu and Nilotic categories by having on the whole much higher ridge count means, higher frequencies of whorls, lower frequencies of loops and arches, lower means for palmar ridge counts, lower frequencies of thenar patterns, second peripheral loops and fourth peripheral loops, and triradius t', and lower means for interdigital and palmar triradial intensity; higher means for total hypothenar intensity. These trends are relatively consistent between the sexes. The Bantu and Nilotic categories are much less differentiated from each other than each are with

the Cushites, especially for digital attributes.

A more precise analysis however, can only be achieved by employing tests of significance. T-tests were consequently calculated between each category for all non-nominal attributes - the palmar patterns and triradii being treated as interval data. I felt that the t-test was more robust and reliable than a chi-square for these attributes, as the t-test is not affected by small expected frequencies arising in the case of those palmar patterns and triradii which occur rarely more than once in an individual on the same hand. Tests of significance for all attributes are presented in Table 7.2 for males and Table 7.3 for females. These tables however, contain too much information to be easily assimilated. To facilitate interpretation therefore, a summary of the main inter-population trends is provided in Table 7.4.

Table 7.4 lists the number of significant differences found for each inter-population combination (i.e. Cushites vs Nilotes, Cushites vs Bantu, Nilotes vs Bantu), for both male and female series, for a representative group of sixty digital and forty palmar attributes. The Cushitic category is highly distinctive with 73 per cent of attributes showing significance between Cushites and Nilotes and 78 per cent between Cushites and Bantu in the males; and 44 per cent and 53 per cent for the same respective combinations in the females. The Nilotes vs Bantu

in contrast show only 24 per cent significant differences in the males and 16 per cent in the females.

The relative rankings between the three categories (i.e. the Cushites are much more distinct from both the Nilotes and Bantu than the latter are from each other), are identical in both the male and female series. On the other hand for the full spectrum of attributes the males display roughly 30 per cent more significant differences than the females for each inter-population set of comparisons. The major difference between the male and female series thus lies in the proportion of significant differences in digital attributes for the Cushitic Category when contrasted to the Nilotes and Bantu. Hence Table 7.4 shows clearly that the proportion of significant differences in the Cushites vs Nilotes for digital attributes is over twice as big in the males (88.3 per cent) than in the females (41.7 per cent). Similarly when the Cushites are contrasted to the Bantu, the males show 90 per cent significant differences for digital attributes compared to 55 per cent in the females. In contrast for the Nilotes vs Bantu the proportion of significant differences is more even (23.5 per cent in males and 16.7 per cent in females).

For palmar attributes the discrepancy between the sexes in the number of significant differences is much less marked. When the Cushites are contrasted with the Nilotes the proportion of significant differences in males is 50.0 compared to 47.5 per cent in females, an almost identical distribution. When Cushites are compared to Bantu the males show 10 per cent more significant differences, as do the Nilotes vs the Bantu.

7.2.2 Relative sex differences in magnitude and direction in selected attributes, between ethnic categories

This greater overall homogeneity in the females particularly for digital attributes requires further investigation if some sort of explanation is to be found. In order to explore this phenomenon in more detail I decided to go back to the raw means and frequencies, and to calculate the differences in means and standard deviations (for quantitative variables) and frequencies (for qualitative variables) between the sexes for a selected representative sample of attributes, an excercise repeated for each ethnic category. (Thus for example the Cushitic female mean for right total radial ridge count was subtracted from the equivalent male mean , and the procedure was repeated for the Nilotes and Bantu). I anticipated that the direct contrast of such differences between the three categories would show how far it is safe to assume that differences between the sexes exhibited by dermatoglyphic variables are relatively constant from population to population. Should this assumption be incorrect, then some more specific explanation for the overall difference in variation between the sexes might emerge.

The results of this analysis are presented in Table 7.5. There are two small methodological points associated with this analysis however, that should be discussed before the results are explored in detail. Firstly, in all the calculations the female figures were always subtracted from those of the males. Thus the minus values in the table reveal that the female mean, standard deviation or frequency for that attribute is larger than that of the males. Such differences in direction are potentially as valuable as the magnitude of the differences in making sense of

the underlying trends. Secondly, in the case of the attributes that measure triradial intensity, where the whole spectrum of inter-population variation is measured in decimal points, the differences have been multiplied by ten to make the scores more equivalent to the range of values exhibited by the other variables.

Table 7.5 clearly shows that for many attributes, considerable variation in sex difference occurs between the categories. These differences will be discussed under two headings, a) differences in magnitude and b) differences in direction.

7.2.2.1 Differences in magnitude

The digital ridge counts display some of the strongest variation of sex differences between the categories. The general trend is for Cushites to exhibit substantially larger sex differences in both mean and standard deviation than the other two categories, who show broadly similar differences to each other. This trend is especially strong in the case of total absolute ridge count, and is least apparent in total ulnar ridge count.

Total finger pattern intensity follows the main trend, but the trend is much weaker than in the ridge counts, particularly for the right hand. When digital patterns are considered different trends become evident. Differences for arches are smallest in the Cushites and largest in the Bantu. For ulnar loops the Cushites show a much larger difference for the right hand, and the lowest for the left hand, with the Bantu displaying the largest difference. For true whorls the Cushites once again show the largest difference for the right hand, with the Nilotic and Bantu differences almost

identical, while for the left hand all three categories show very similar values.

One significant discernable trend in the digital variables is that those populations who exhibit a high mean or frequency, also tend to manifest the largest sex difference. Thus in the case of ridge counts it is the Cushites who show the highest means and standard deviations, and it is the Cushites who also show the highest sex differences. This is also the case for concentric whorls on the right hand. Similarly it is the Bantu who show the highest arch frequencies, and it is they who show the highest sex differences for arches.

In contrast to the digital variables, the palmar variables show more individualistic trends. Differences in the means for interdigital triradial intensity are marginally highest in the Bantu and lowest in the Cushites. The Bantu and Cushites, on the other hand, display similar high differences in standard deviation, while the Nilotic differences in standard deviation are half those of the other two categories. The differences in mean for total palmar intensity are highest in the Bantu and lowest in the Nilotes. The standard deviation differences are contradictory between right and left hands, with the Nilotes showing the highest value for the right hand and the lowest for the left. For total hypothenar intensity the Bantu show hardly any difference in either mean or standard deviation, while the Nilotes and Cushites show higher similar values.

Total palmar ridge count is also a relatively complex variable. For the right hand the Cushites show the highest sex difference in means and by far the largest difference in standard

deviation, with the Nilotes exhibiting the lowest differences. For the left hand, however, the Cushites manifest by far the lowest difference in mean, and the Bantu the highest, although the Cushites still exhibit by far the largest difference in standard deviation. The difference in standard deviation of the Bantu for left total palmar ridge count is particularly low.

Difference in frequency of triradii in the thenar area differ appreciably between right and left hand. For the right hand the Nilotes show the highest difference, but for the left it is the Cushites who exhibit the highest. The Bantu manifest hardly any difference on the right hand but a relatively large one on the left.

The description of sex differences between the categories for palmar variables so far has indicated a complex pattern of relationships. Like the digital variables, palmar variables also exhibit marked variations in sex difference between the categories, but unlike digital variables, the magnitude of the sex differences are not necessarily related to the magnitude of the original means and frequencies. Thus the Cushites show the lowest mean for right total palmar ridge counts, but they also display the highest sex difference for the trait and the highest difference in standard deviation. Similarly in the case of thenar triradial frequency for the left hand, the Cushites display by far the lowest frequency for the trait, but at the same time they also exhibit the largest sex difference. Thus the palmar variables do not support the nascent hypothesis that relative sex differences are associated with the magnitude of means and frequencies, as was suggested at the end of the discussion of the digital variables.

7.2.2.2 Differences in direction

The majority of the variables present no surprises regarding the direction of sex differences in mean and frequency, as the trends follow consistently those outlined for the total sample (see section 7.1). Thus the digital ridge counts show that as expected all the male means are greater than those of the females for all three categories, except for total ulnar ridge count on the left hand for the Bantu, where the females show a marginally greater mean. Similarly arch frequencies are consistently higher in the females, as are loops on the right hand, total hypothenar intensity, palmar ridge counts, fourth loops and hypothenar loops for the left hand. Whorls, double loops for the right hand, interdigital palmar intensity, total palmar intensity, peripheral third loops, central hypothenar loops for the right hand and triradius t all show means or frequencies greater in the males.

Inspite of the fact that a majority of variables conform to expectations, however, there are some differences between the categories, notably for left total ulnar count (Bantu), left true whorls (Bantu), left double loops (Cushites), right thenar loops (Cushites), left peripheral fourth loops (Nilotes), and right central hypothenar loops (Nilotes) which are opposite to the normal trends. The most spectacular differences, are to be found in the standard deviations of the ridge counts, where the Cushites display high values which are all positive (i.e. greater in the males), while the Nilotes demonstrate positive standard deviations for only right ulnar and absolute total ridge counts. The Bantu seem to occupy an intermediate position, being negative for the radial and unilateral counts and positive for the ulnar and absolute counts. Of the remaining variables where standard deviations are listed, the

Nilotes fail to follow the expected trend for total palmar intensity (left hand) and for hypothenar intensity (right hand), while the Bantu exhibit a marginally contrary trends for hypothenar intensity (left hand).

## 7.3 Discussion

The evidence from Table 7.5 shows therefore, that there are large variations in the magnitude of sex differences in means, standard deviations and frequency from one ethnic category to another. There are also variations in the direction of sex differences, especially in the standard deviations of the digital ridge counts.

It is now possible to work out why the females exhibit a smaller proportion of significant differences between the ethnic categories in Table 7.4. Where large and significant differences occur in males, but only marginal or non significant differences in females, this primarily reflects the differential magnitude in sex differences between the ethnic categories. In the case of digital ridge counts for instance, if the relative magnitude of sex differences was the same in all three ethnic categories, then the same degree of significance would prevail in both sexes. What is apparent however, is that while the relative magnitude of sex differences for digital ridge counts is the same for the Nilotic and Bantu Categories, in the Cushites sex differences are nearly double those for the Nilotes and Bantu, Thus the female Cushitic means for digital ridge counts are much closer to those of the female Nilotes and Bantu, than the male Cushitic means are to those of the male Nilotes and

Bantu. This accounts for the much greater proportion of significant differences in the males when the Cushites are compared to the Nilotes and Bantu.

There is still a great deal of work to be done to account for the discordant variation between the sexes. This investigation can be regarded as just a preliminary exploration of the problem. From the analysis so far however, it is possible to come to an important conclusion, that the assumption that sex differences for dermatoglyphic attributes are relatively constant from one population to another is not valid. This has major consequences for the interpretation of dermatoglyphic variation between the minimal population samples, as an absence of concordance between the male and female series can no longer be just attributed to low sample numbers, but should be regarded as a central characteristic of dermatoglyphic variation.

# CHAPTER 8

## An Examination of the Extent to which Population Samples in Kenya tend to vary dermatoglyphically according to Linguistic and Cultural Patterns of Affinity

### Introduction :

Spuhler (1972) remarks that:

> "Many anthropologists insist that there is no necessary close and permanent correlation between race (whether regarded as anatomical type or gene pool) language and culture."
>
> (Spuhler 1972:87)

This view has some empirical justification. Howells for instance, found only moderate correlations between linguistic resemblance and biological distance between populations from Bougainville in the Solomon Islands (Howells 1966). Hiernaux (1972) failed to find any distinctive taxonomic groups in sub-Saharan Africa, in spite of the fact that distinctive racial taxonomies were devised in the past which corresponded with large scale linguistic groups (e.g. Oschinksy 1954).

Nevertheless there is considerable theoretical justification for expecting a strong association between linguistic or cultural and biological variation. At any focus on the time continuum, few people would dispute that it is a normal process for children to inherit not only the genes, but also the language and culture of their parents. People also tend to choose marriage partners who speak the same language, share the same culture and who are usually living in relatively close geographical proximity to them before marriage. Transferring the emphasis from individuals to populations, we can infer that a breeding population which shares a common language and culture different from that of other populations,

will tend to preserve a close association between its gene pool, language and culture from one generation to another, given demographic stability and a degree of isolation over time.

If the simultaneous transmission of genes, language and culture from one general to another is the normal process in human peoples, then the fact that empirical evidence has often failed to show strong correlations between linguistic and biological variation can only mean that this normal process is often significantly disrupted. As Hiernaux points out in the case of sub-Saharan Africa:

> "As a result of many large-scale migrations and expansions, most living populations of this subcontinent result from multiple hybridization processes often involving very different components. .... The identity of population units is often ephemeral: many of them will disintegrate, and recombine after a few generations. As a result of such processes, a group of populations which claim a common origin and share features of culture and language, may be far from forming a biological cluster."
>
> (Hiernaux 1972:106)

Thus it only takes one period of major population movement followed by significant hybridization of different peoples to dilute markedly any strong relationship that may have existed between the gene pool, language and culture of the diverged peoples. The frequent occurrence of famine, epidemics, and warfare in many areas of the world makes Hiernaux's assessment highly plausible.

Although disruptive events may greatly alter the relationship between biological and linguistic or cultural affinity, is it reasonable to assume that such events are likely to eradicate all biological similarity between peoples who have diverged and who

might still share linguistic features? There is evidence to suggest that some similarity may still persist, even though the overall relationship between biological and linguistic affinity is not strong. Hiernaux for example (1968b) showed that the mean biological distance between all Bantu populations sampled was lower than that for distances between Bantu and non Bantu samples, a fact that suggested to Hiernaux that the Bantu had a common origin. Similarly Spuhler (1972) found a significant difference in the occurrence of smaller biological distances between Amerindian populations who belonged to the same linguistic group, when compared to distances between populations belonging to different linguistic groups. Some association between gene pool and language was found, in spite of the fact that considerable univariate heterogeneity existed within the linguistic groups examined.

In the investigation that follows I shall be seeking to assess how far dermatoglyphic patterns of variation between the Kenyan samples correspond to known patterns of linguistic and cultural affinity. In the first section, I shall examine the relative heterogeneity of ethnic categories as depicted by the dermatoglyphic variation of ethnic groups. In the second and third sections I shall explore the relative heterogeneity and the correspondence of dermatoglyphic resemblance with linguistic affinity of ethnic groups, from the evidence of minimal population samples.

## 8.1 The degree of correspondence of dermatoglyphic affinity and linguistic and cultural affinity at the macro population levels.

The peoples of Kenya, as I described in Chapter 1.2,

belong to three major linguistic families, the Cushites, Nilotes and Bantu. The precise allocation of minimal population samples to respective ethnic groups and categories was given in Chapter 4.5. Of the 12 ethnic groups distinguished in Chapter 4.5, two (the Somali and Galla) belong to the Cushitic Category, four (the Maasai, Kalenjin, Karimojong and Luo) belong to the Nilotic Category, and five (the Luyia, Central Bantu, Pokomo and Taita) belong to the Bantu Category. The two samples of traditional hunters and gatherers, the Ngwesi and Mukogodo Dorobo from Laikipiak, could be classified tentatively on linguistic grounds as belonging to the Cushitic and Nilotic Categories respectively. On cultural and regional grounds however, they are clearly a distinct ethnic group. Due to this fact and the uncertain nature of the linguistic classification of the Mokogodo, they have not been allocated to any of the three major ethnic categories (see Chapter 1.2).

The aim of this section is to examine how far patterns of dermatoglyphic relationships between the ethnic groups correspond to known patterns of linguistic and cultural affinity. The investigation of this issue begins with a brief appraisal of the nature of the dermatoglyphic variation between the ethnic categories, followed by an exploration of the relative homogeneity of the categories from the evidence of $\Delta g$ distances between ethnic groups. Detailed dermatoglyphic affinity between ethnic groups within ethnic categories is examined at the end of this section through a consideration of the evidence from a principal coordinates analysis of the $\Delta g$ distances. In view of the discussion at the end of Chapter 6, in which the conventional methods of delineating trends of resemblance from ordination analysis were criticised, the

graphical techniques used here to depict resemblance are unorthodox.

### 8.1.1 General trends of dermatoglyphic relationship between ethnic categories

Univariate differences between the three maximal ethnic divisions (the Cushitic, Nilotic and Bantu Categories), were discussed in another context in Chapter 7.2. It is clear from that discussion that the Cushitic Category is sharply differentiated from the other two, while the Nilotic and Bantu Categories in contrast show many fewer differences with each other. In fact as Table 7.4 demonstrates, over 70 per cent of attributes are significant when the Cushites are contrasted to the Nilotes and Bantu in the males, and over 40 per cent for the equivalent female comparisons. The Nilotes however, exhibit only 24 and 16 per cent significant differences for males and females respectively from the Bantu.

It should be noted however, that although there is a great deal of significant variation between ethnic categories, the range of variation for all attributes (see Appendix 1) at this level of resultion is very small compared to that manifested at the ethnic group and minimal population levels. This shows that the process of pooling populations on the basis of maximal linguistic and cultural affinity cancels out most of the variation. This is particularly true for the Nilotic and Bantu Categories, which display a large range of variation for their respective constituent populations, and which, when pooled, result in comparatively few significant differences when the categories are compared.

8.1.2 The relationship between dermatoglyphic variation and linguistic and cultural affinity of ethnic groups.

The examination of dermatoglyphic variation between ethnic groups, and the extent to which ethnic groups vary dermatoglyphically according to linguistic and cultural affinity, can be considered a prelude to the more detailed analysis involving minimal population samples. There are three reasons why the direct consideration of dermatoglyphic variation between ethnic groups is advantageous. Firstly, all ethnic groups are represented by large sample sizes, and any trends can thus be interpreted as being relatively free of the uncertainties surrounding the investigation of variation between the minimal population samples. Secondly, as the pooling of minimal population samples is normally associated with a considerable loss of variation (Hiernaux 1972:98-9), any marked trends that emerge when ethnic groups are analysed are particularly strong, and will thus be valuable guidelines for the further analysis of the minimal population samples that follows. Finally, the Kenyan ethnic groups in this study are units broadly compatible in scale with many units adopted by Hiernaux (1968a; 1972) in his sub-Saharan investigations. It is interesting to see what kind of trends can be associated with this scale of analysis, especially when contrasted with the findings from the analysis of the small scale minimal ethnic population samples.

8.1.2.1 Univariate descriptive evidence of dermatoglyphic variation between ethnic groups

Univariate tables showing variation for ethnic groups for a full spectrum of dermatoglyphic attributes are given by sex in Appendix 1.3.

Table AP.1.3.13 also list the results of applying analysis of variance and chi-square tests to univariate attributes, and reveal that most attributes demonstrate high significance for ethnic group differences in means and frequencies.

A detailed analysis of linguistic and cultural heterogeneity could include a full investigation into univariate differences between ethnic groups, within and between ethnic categories. Such an analysis however, would take up a great deal of time and space if it were to be fully reported. Moreover it would be comparatively difficult to interpret overall trends from the mass of results that would be the consequence of applying univariate tests of significance. The procedure followed here, therefore, is to investigate inter-ethnic group differentiation in a highly summarised form, through the calculation and further analysis of $\Delta g$ distances based on the twenty six summary attributes outlined in Chapter 5.3.2.

Tables 8.1 and 8.2 give the matrices of $\Delta g$ distances between ethnic groups for males and females respectively. Below each matrix are the results of principal coordinates analysis of each set of distances, consisting of a listing of latent vectors, squared distances from the centroid, latent roots and trace from which the percentage variance of each latent vector is calculated. These are the results on which all subsequent investigations in this section are based.

8.1.2.2 Relative linguistic and cultural homogeneity of ethnic categories from the evidence of $\Delta g$ distances between ethnic groups.

The $\Delta g$ distances between ethnic groups in Tables 8.1 and 8.2 show a great deal of variation in both sexes, ranging from 47

to 1665 in males and from 57 to 918 in females. In order to investigate the extent of dermatoglyphic homogeneity of each ethnic category as depicted by constituent ethnic groups, a one way analysis of variance was performed to see whether inter-group $\Delta g$ distances within ethnic categories (e.g. distances between Luyia, Central, Pokomo, Mji Kenda and Taita Bantu groups in the Bantu Category) are significantly smaller than distances of ethnic groups between categories (e.g. distances between Luyia or Pokomo Bantu versus Kalenjin Nilotes or Galla Cushites).

The findings are reported in Table 8.3. In neither sex are the F probabilities significant. This however, may be due to the fact that only one distance exists for the Cushitic Category within division, and the low numbers of distances in the Nilotic and Bantu within divisions. Nevertheless in the males, the "between all categories" mean distance of 433, is appreciably higher than any mean distances within either the Cushitic, Nilotic or Bantu Categories. This implies some degree of relative homogeneity for ethnic Categories. In the females however, there is in contrast no marked difference between the mean for distances in the "between all categories" division and those within each respective ethnic category, and hence implies a lack of relationship between dermatoglyphic and ethnic affinity.

To clarify these results, I pooled the within ethnic category distances and compared their mean with the mean for between ethnic category distances through a t-test. Table 8.4 shows that in the males, the mean of distances between ethnic categories is significantly higher than that for distances within ethnic categories, though in the females the two means are similar and not significant. There is thus a notable lack of agreement between the male and female results, making interpretation difficult. It would appear that only the males show any indication of a correspondence of dermatoglyphic affinity with linguistic and cultural affinity.

Owing to the contradictory nature of these results, I decided to look at the qualitative distribution of distances within and between ethnic categories to see if clearer relationships existed. The range of distances was divided into 3 divisions and the percentage occurrence of distances falling into each of the 3 divisions was examined for within and between ethnic category divisions. The results are contained in Table 8.5.

The first point to notice is that no less than 18 per cent of between ethnic category distances in males and 34 per cent in females are less than 200. This shows that an appreciable minority of small distances are between ethnic groups belonging to different ethnic categories. Similarly, 29 per cent of <u>within</u> ethnic category distances in males and 21 per cent in females are greater than 400, indicating some heterogeneity within ethnic categories.

Nevertheless it is also clear that a much larger percentage of small distances occur in the <u>within</u> ethnic category division in each sex. Thus in the males, 35 per cent of the within ethnic

category distances fall under 200, compared to 18 per cent of distances between ethnic categories; the equivalent comparisons in females are 64 per cent compared to 34 per cent respectively. It is also apparent that the females show a much larger proportion of small distances in toto than the males.

When chi-square tests of significance are applied to the male and female figures in Table 8.5, neither sex shows significance though the females approach significance at $p = 0.055$. This trend is a broad reversal of the trends of significance for the quantitative analysis in Table 8.4, where the females were far less significant than the males.

This reversal of trends of significance between the sexes can be accounted for by two factors. Firstly the males show a much greater ratio of large distances to small, accompanied by a much greater proportion of large distances in the between ethnic categories division (52 compared to 29 per cent). This contrasts with the females, where the percentage of large distances is only marginally greater in the between ethnic categories division (24 compared to 22 per cent). This accounts for the much higher male mean of distances in the between ethnic categories division, in Table 8.4, as the excess of large distances in the male "between" division carries considerable arithmetic weight. Secondly, the females are much closer to significance in the qualitative analysis in Table 8.5 because of the very large relative difference in proportion of distances in the between and within ethnic categories divisions that fall in the 200-399 division (42 per cent in the "between division" compared to 12 per cent in the "within division"),

This contrasts with a much more even distribution of distances in the males for the 200-399 division (29 compared to 35 per cent).

The qualitative analysis is, on the whole, perhaps a better indicator of the extent of variation along ethnic lines than the quantitative analysis. The preponderance of smaller distances in the within ethnic categories division in Table 8.5 is consistent in both sexes, unlike in the quantitative analysis in Table 8.4 where the female within categories mean differed little from that between categories. In spite of the fact that in both sexes there is a greater proportion of smaller distances within ethnic categories in the qualitative analysis however, the correspondence of dermatoglyphic with ethnic affinity implied by this trend is by no means strong, as there exists a minority of small distances in the between ethnic categories division in both sexes, and a minority of large distances in the within ethnic categories division.

The trends observed in this analysis of absolute distances are rather general. It has not been possible to assess the relative homogeneity of specific ethnic categories with any certainty, owing to the adverse affect of small numbers on the tests of significance used. For a more detailed appraisal we must consider the implications of the principal coordinates analyses in Tables 8.1 and 8.2.

8.1.2.3 Relative linguistic and cultural homogeneity of ethnic categories from the evidence of latent PCO vectors.

The principal coordinates analyses depicted in Tables 8.1 and 8.2 show that 88 per cent of the male and 83.5 per cent of the female variance is accounted for by the first three latent vectors,

and 98 and 92 per cent by the first six. An almost complete picture of ethnic affinity can be obtained therefore, by considering the first six latent vectors.

It is possible to approach the analysis of latent vectors either by considering them individually in turn in decreasing order of importance, or by producing scattergrams of each pair of vectors in turn. These methods however, as I pointed out in Chapter 6, suffer from difficulty of interpretation if discordance is manifested by different vectors. A series of scattergrams for instance, for every combination of six latent vectors, would produce thirteen different representations of relationships between ethnic groups, without dividing up the data by sex.

In order to obtain a more complete picture of inter-group affinity, it is necessary to combine the information contained in each vector on one graph for each population. As I showed in Chapter 6, this can be achieved by linking the plotted scores of each vector for each population with a straight line. Affinity can then be assessed not only by distances between points on each vector, but also by the "shape" of the population line. The "shape" in fact, might be more important than distance, as it is possibly more stable in the face of random fluctuation due to sample size. This method is described in more detail later in the analysis of minimal ethnic population variation (section 8.3)

The program used to plot the graphs that follow is a private one written by Graham Cameron of the Department of Psychology, University of Stirling. It not only provides full labelling, but also differentiates each population by a unique line and symbol.

For ease of interpretation, the number of lines on each graph is limited to six. The reader therefore, has to turn over pages to compare graphs of the same series of latent vector comparisons.

In the following analysis there are six graphs, three for males and three for females. Each graph portrays the linked vector coordinates for ethnic groups belonging to a specific ethnic category. Here the Dorobo group is included in the analysis, as it is interesting to see where it fits in, and it is included in the Cushitic graph.

a) Cushitic and Dorobo ethnic groups: (figure 8.1 (males); figure 8.2 (females)).

The outstanding characteristic of the Cushitic ethnic groups (Somali and Galla) is their strong similarity in the "shape" of their lines, implying considerable affinity. Moreover a comparison with the Nilotic and Bantu graphs reveals no other group with a similar shape. The Galla and Somali therefore, can be viewed as belonging to the same macro-taxonomic unit. Nevertheless the distance between the Somali and Galla in both sexes between the coordinates on the first vector in particular and less markedly on the second and third vectors, is quite appreciable. This possibly means that though structurally similar in shape, the two populations have nevertheless been separated for many generations. If this is the case, it would correspond to the relative linguistic affinity between the two groups.

The Dorobo group on the male graph are not only almost identical in shape to the Somali group, but also in distance for all but the second and sixth vectors. Dermatoglyphically therefore, the Laikipiak Dorobo are firmly allied to the Cushitic populations.

b) Nilotic and ethnic groups; (figure 8.3 (males): figure 8.4 (females)).

The male groups reveal marked differences both in shape and distance between two groupings. The first grouping consists of the Maasai and Kalenjin groups who who affinity in shape and distance on all six vectors. The second consists of the Karimojong and Luo groups who are very similar on the most important first three vectors, though not on vectors 4 to 6.

According to linguistic classification (Sutton 1974:82-3) the Luo are the sole Kenya representatives of the River-Lake Nilotes, who have closer affinities with the Plains Nilotes (Maasai group and Itunga (Kairmojong) group) than the Highland Nilotes of whom the Kalenjin are a branch. Culturally however, the Maasai have more in common with the Kalenjin than the Karimojong Group, the latter occupying a middle ground from the Luo who represent the true Nilotic cultural polarity. The dermatoglyphic resemblance here clearly reflects cultural rather than linguistic affinity.

The females however, show marked differences from the males. Firstly, there is much greater homogeneity in distance than for the males on all six vectors. Secondly, from the point of view of both shape and distance it is the Maasai and Karimojong groups who reveal the greatest affinity (on the first two vectors which account for most of the variance), with the Kalenjin and Luo lying on opposite polarities on the first vector. This clearly fits the linguistic classification better than the cultural classification. On the whole however, the general patterns of affinity in the females are unclear and inconclusive.

c) The Bantu groups: (figure 8.5 (males): figure 8.6 (females)).

Patterns of resemblance are very clear in the males. There is close homogeneous similarity in distance and shape between three groups, Luyia, Central Bantu and Pokomo. The Mji Kenda and Taita in contrast are not only very dissimilar from each other, but also from all the other groups. The Taita in particular are clear isolates on vector 1, and the Mji Kenda are distinct on vectors 2 and 3.

Like the Nilotes the Bantu females show greater homogeneity than the males. Only the Mji Kenda on vector 1 are sharply distinguished. Within the central mass however, the Taita are comparatively more differentiated (vectors 1, 3, 4 and 5) though not nearly as much as the males. The close affinity between the Luyia, Central Bantu and Pokomo Groups is evident on vector 1, but less apparent on the other vectors. With the exception of the Mji Kenda, therefore, the females show overall homogeneity in distance, though not in shape. Trends are not nearly so clear as in the males.

The alignment of dermatoglyphic affinities in both sexes is incompatible with linguistic and cultural affinity. According to these criteria, there ought to be three distinct groupings, one consisting of Mji Kenda, Pokomo and Taita, the second being the Central Bantu and the third the Luyia, with the Central Bantu in between the Coastal and Luyia groupings. The overall dermatoglyphic resemblance between Bantu Groups therefore, does not suggest any marked correspondence with known patterns of linguistic or cultural affinity.

## 8.2 The extent of overall dermatoglyphic heterogeneity of ethnic groups: an examination of affinity between minimal population samples, within and between ethnic groups.

If a close correspondence exists between dermatoglyphic gene pool and language and culture in Kenya, we can expect to find a) relative homogeneity of minimal sample means and frequences with ethnic groups and b) that trends of dermatoglyphic affinity between minimal samples within ethnic groups will follow patterns of known intra-group linguistic and or cultural affinity. These are the main issues that will be investigated in the following section.

### 8.2.1 Approach.

In the analysis of ethnic category homogeneity as depicted by ethnic groups, one of the main procedures used to test homogeneity was to compare the mean of the $\Delta g$ distances between ethnic groups within ethnic categories with that between ethnic categories. This method however, cannot be used in the analysis of minimal population samples, as the magnitude of $\Delta g$ distances is significantly dependent on sample size. In Table 8.6 are listed the mean $\Delta g$ distances at various sample sizes for male and female analyses based on a) twenty six summary attributes and b) one hundred and ten non-summary attributes. The clear trend to emerge from the table is the fact that the mean distances for the smallest sample size are much greater than those for the larger sample sizes. In fact, relative stability is only achieved when a sample of fifty is reached. As one of the ethnic groups, (the Luyia), is mostly represented by sample sizes well below fifty in each sex, the sample sizes in the thirties and forties

are liberally represented in many of the other ethnic groups, particularly in the females, any measure of relative homogeneity involving comparisons of wthin and between group variation of raw $\Delta g$ distances is bound to be distorted by sample size bias. The smaller the samples within an ethnic group, the higher the mean of inter-minimal population sample distances for that group is likely to be: from sample size effects alone.

When we look at the effect of sample size on $\Delta g$ distances therefore, we must take into account the fact that the extent of random fluctuation is disporoportionately greater in sample sizes in the twenties and thirties than in those in the fifties to and over a hundred. Random fluctuation however, will tend to be random in each attribute. In the calculation of the distances, the random variation will tend to cancel itself out when the distances from each attribute are averaged out. Nevertheless it will result in much bigger distances in absolute terms for those very small sample sizes where random fluctuation is proportionally more extensive.

The random variation in a distance therefore, can be termed as "random noise". This noise however, should not affect relative distance. Any systematic variation will tend to pull in one direction from attribute to attribute as the distance is calculated between two units. This relative systematic variation is what is mostly picked out during PCA and PCO analysis. That this is so can be seen by the fact that the PCO analyses on minimal population sample distance described in Chapter 6 produce latent vectors that strongly correlate with PCA component scores which load heavily on specific attributes whose inter-relationships are far

from random (see Chapter 5.3). PCA component scores and PCO vectors are thus much better representations of inter-population relationships than raw distances. Any analysis of within and between ethnic group affinity of minimal population samples should be based on such measures rather than raw distances.

PCA component scores and PCO vectors however, are not of course, free from the effects of sample size.But the issue is radically different from that adopted in the interpretation of raw distances. It is no longer a case of distortion due to a prevalence of random noise in distances of small samples, but one where the central point of interest is now how far a PCA score or PCO vector is a reliable estimator of the true population parameter. PCA scores and PCO vectors may be potentially unrepresentative measures, but they are not biased or badly weighted ones.

In the analysis of the linguistic and cultural homogeneity of ethnic categories according to the evidence from ethnic groups, I could assume from the fact that distances appear substantially to stabilise at a sample size around 50 to 60 (Table 8.6), that all ethnic category sample sizes were large enough to contain little random noise. The application of analysis of variance to raw distances of ethnic groups between and within ethnic categories was thus not only permissable, but also a useful summary method of exploring relative homogeneity. The fact that raw distances cannot be adopted in the analysis of minimal population sample homogeneity within ethnic groups means that a less summarised approach is necessary, involving not only the consideration of PCA component scores and PCO vectors but also univariate measures.

The choice of approach is influenced by one further factor namely the existence of marked discordance observed between phenetically unrelated taxonomic attributes described in Chapter 5.3. In Chapter 3, I proposed that from a theoretical standpoint, the demonstration of discordance is incompatible with the existence of strong overall trends of ethnotypic variation. This means that in the present group of population samples, it is unlikely that strong ethnotypic variation is going to be manifested in the analysis of overall resemblance. This does not preclude, however, the possibility that some specific attributes may vary according to ethnic patterns of affinity. A univariate approach therefore, is additionally advantageous in that it will permit the investigation of the notion that variations along ethnic lines may exist in certain attributes and not in others.

### 8.2.2 Univariate analysis of ethnic group homogeneity and affinity from the evidence of minimal population sample differentiation.

The procedure adopted to investigate this issue was to perform a one way analysis of variance for a representative sample of digital and palmar attributes on the constituent populations for each ethnic group in turn.

During this particular analysis the between group variance is the deviations of the sample means from the grand mean for the ethnic group under consideration, and the within group variance is the weighted average of the deviations of individuals (people) from within each sample mean. The ratio of the between to within group variances, the F ratio, has a known sampling distribution which

through the consideration of within and between group degrees of freedom gives a figure of probability from zero to unity. The greater the F ratio the less chance there is that the difference in sample means is random, and hence the lower the F probability.

What I am testing here therefore, is how far the distance between sample means within each ethnic group, is statistically random. The finding of significance however, cannot be directly interpreted as a sign of heterogeneity. It is possible for instance, to obtain high significance where very small differences in means exist if sample sizes are large enough. Significance cannot directly tell us therefore, whether differences in means within ethnic groups are relatively larger or smaller than differences in sample means between ethnic groups; or whether the differences of means in one ethnic group are greater than those in another. Indirectly however, significance can point to the existence of heterogeneity, in that in the present analysis the sample sizes are not very large. Consequently, a big difference in means is required to achieve significance. To find significance in the present analysis therefore, implies that the range of mean differences is large.

Although one way analysis of variance can imply the existence of large differences in means, it cannot indicate whether it is only the extreme samples that cause significance. It is possible for instance, to obtain significance when only one population out of many is very different. In this case, the ethnic group would be homogeneous for all but one member, but A.O.V. might produce the same figure of significance as a group where many

constituentpopulations differ significantly. The present analysis therefore, should only be regarded as a preliminary guide to ethnic group heterogeneity or homogeneity.

A possible source of distortion in the application of A.O.V. is the fact that in some attributes in certain groups assumptions of homoscedasticity could not be met (about five per cent of attributes in all). The application of sophisticated transformations to correct skewness through the use of an iterative function by Box and Cox (1964), on a pilot range of attributes, succeeded in slightly reducing heteroscedasticity, but failed to change the F probabilities to any significant extent. I decided therefore, to use untransformed data in the following analysis.

In view of the large amount of output that can be produced by one way analysis of variance, I limited the number of attributes to be considered in the analysis to 20 digital and 30 palmar non summary attributes; and 14 summary digital and palmar attributes (see Table 8.7 for a precise listing). Palmar patterns and triradii were treated as interval data for reasons already outlined in Chapter 5.2. Finally to reduce output further some ethnic groups were pooled. The Somali and Galla groups were pooled to form a general Cushitic group. The Luo, being on their own as a group, were included with the Turkana and Iteso to form a general non Kalenjin/Maasai Nilotic group (the Luo are in fact culturally more similar to the Turkana and Iteso than the other Nilotic populations). Finally, the Pokomo, Mji Kenda and Taita were amalgamated to form a general Coastal Bantu group - linguistically permissable as Sutton shows (Sutton 1973:83).

One way analysis of variance was applied separately on each ethnic group in turn for the attributes considered, with individual people providing the within variance and the minimal population samples the between variance for each ethnic group. The results in a condensed form are given in Table 8.7 and 8.8 for males and females respectively. Significance is indicated by an asterisk following the probability for the attribute, and is conventionally defined here as any probability equal to or less than 0.05.

Table 8.9 is a summary of Tables 8.7 and 8.8, and lists the number and percentage of significant differences found in each ethnic group for digital and palmar groupings and for undifferentiated summary attributes. The largest number and percentage of significant differences occur in the Coastal Bantu in both sexes, followed by the Central Bantu and the Luyia Bantu. The Kalenjin show a large proportion of differences in the males but not in the females, while the Cushites exhibit an opposite trends, with differences predominating in the females. The Maasai and Luo/Karimojong show few differences in both sexes, as do the Dorobo in the males.

One of the features of these results is the differential distribution of differences in digital and palmar attributes from one group to another. The Coastal Bantu and Kalenjin (males only) show far more differences in digital attributes, while differences in palmar attributes are far more common in the Central Bantu.

We thus have a situation where the variation in sample means varies statistically randomly in a majority of attributes for every ethnic group, the proportion of random variation ranging from 68 per cent of non summary attributes (Coastal Bantu males) to 100

per cent (Maasai males and Luo/Karimojong females). The non random variation moreover, tends to occur in clusters of phenetically related attributes. In the Coastal Bantu, it can be found in some radial and a majority of ulnar ridge counts (see Tables 8.7 and 8.8), and in hypothenar attributes in both sexes. In the Central Bantu, the largest differences are for palmar ridge counts in both sexes, on the hypothenar in the males and in the axial region in the females. The Luyia Bantu show some differences in radial ridge counts on the fourth and fifth digits and in ab and cd palmar ridge counts in both sexes. The Kalenjin exhibit large differences in radial counts in the males, and in the axial triradii in both sexes. Differences in the other groups are less common and systematic and show little replication between the sexes.

Having spotlighted which related groups of attributes are significant in each ethnic group, it is now possible to obtain preliminary indications of the extent of heterogeneity by examining the distribution of minimal population sample means and frequencies for significant attributes, using the univariate tables listed in Appendix 1.4. Of special interest is whether heterogeneity is a consequence of large differences between several populations; or between a homogeneous swarm ad one aberrant population. These issues are discussed below for each ethnic group ranked according to the largest proportion of significant differences listed in Table 8.9.

i) The Coastal Bantu

The Coastal Bantu show significance for digital ridge counts and hypothenar patterns and triradii in both sexes. In the males, there are three general levelsof variation for digital ridge

counts: low means in the Taita and Taveta, medium to medium high means in the Mji Kenda and Pokomo, and high to very high means in the Digo. Significance is therefore mostly a consequence of the large difference between the Digo at the high polarity and the Taita and Taveta at the low one.

In the females however, the Mji Kenda, Pokomo and Taita/ Taveta are relatively homogeneous for digital ridge counts showing low to medium values in all populations except for the Digo, who display very high means. Heterogeneity, is therefore due to the fact that the Digo are quite distinct from the rest.

Significance in the hypothenar area, particularly for central hypothenar loops, is also a pronounced feature for both hands in the females and for the right hand in males. In both sexes, very low frequencies for central hypothenar loops occur in the Chonyi and Mid Mji Kenda, while the other populations show medium to medium high values except for the Taita who exhibit very high frequencies. It is the Taita therefore, that contribute most to the observed significance.

ii) The Central Bantu

The overriding characteristic of the Central Bantu group is the very high significance exhibited by palmar ridge counts and axial triradii in both sexes. The distribution of the various palmar ridge count means is unusually concordance for the Central Bantu, even between sexes, with the range of means greater than the rest of the Kenya range. At the high polarity are the Muranga and Kiambu Kikuyu, Muthambi, Mwimbi, Tigania and Tharaka, while the very low polarity is represented by the Igoji and Imenti. The

rest show low medium to medium high means. For palmar ridge counts therefore, the Central Bantu are very heterogeneous, and the patterns of variation do not correspond to internal linguistic divisions. For example the Meru and non Meru samples are both heterogeneous.

A detailed analysis of the Central Bantu variation for axial triradii is not justified, owing to unreliability in the measurements of these attributes (see Chapter 4). A look at the frequencies for t and t'in Appendix 1.4 however, reveals that many Central Bantu samples differ appreciably in the frequencies for the traits.

iii) The Luyia Bantu

Considering the low sample sizes of the Luyia populations it is not surprising that most attributes show a lack of significance, even though, as the univariate tables demonstrate, there are very marked differences in means or frequencies for all attributes. The fact that some attributes are significant means that the range of means or frequencies must be particularly wide owing to the small size of the Luyia samples. For digital attributes, only radial ridge counts on the fourth and fifth digits in the males and only on the fifth in the females show significance. The only consistent pattern to emerge from the univariate means is that the Bukusu show very low means in both sexes, and the Marama very high ones. With other populations displaying a considerable but less concordant range of variation. The low Bukusu and high Marama values occur in most digital attributes, but only on the fourth and fifth digits is the difference in means large enough to produce significance.

Both sexes also manifest significance for palmar ridge counts. On the whole, most Luyia population samples in both sexes tend to show relatively high means - remarkably homogeneous considering the sample sizes. Once again however, it is the Bukusu who are outstanding, showing very low means in both sexes, especially in the females.

iv) The Kalenjin Nilotes

The Kalenjin show highly significant differences in the radial digital ridge counts in males, and in ab ridge counts and axial triradii in both sexes. This apparent sex difference for the digital ridge counts requires explanation, especially as sample numbers are broadly equal in both sexes except for the Kony which possess a much smaller female sample. The univariate tables in Appendix 1.4 for radial ridge counts demonstrate that all Kalenjin populations are broadly homogeneous, except for the Kony in both sexes, which have the lowest Kenyan digital ridge count means for most of these attributes, and the Kipsigis in the males only, which show relatively high means. The lack of significance in most female radial ridge counts is thus accounted for by the absence of the high Kipsigis polarity of the males, combined with a smaller sample size for the Kony. It should be noted however, that though the Kony do not differ significantly from the others in the females, their separation in terms of the absolute differences in means for digital ridge counts from the other Kalenjin populations is actually greater than in the males. The distinctiveness of the Kony moreover, is to be seen in all digital attributes, even where differences are not large enough to produce significance.

The significance of the ab ridge counts is due to less systematic differences than were manifest by the digital ridge counts. The population at the lowest and highest polarity differs from right and left hands, and from one sex to another. In general, however, it is the Nandi who show lower means, and the Keyo, Marakwet and Pokot high ones.

Finally, the axial triradii demonstrate that the Kipsigis have far the highest ad lowest frequencies of t' and t respectively in the males, and they are joined by the Nandi in the females.

v) The Cushites

The patterns of significance exhibited by the Cushites are remarkable in that although the male samples are generally much bigger than those of the females, it is the females who display by far the larger number of significant attributes. This implies the existence of considerable inter-sex discordance.

The females for instance, display no less than four significant differences for non summary digital attributes compared to only one for males. The univariate tables show however, that whereas in the males all Cushitic populations show remarkably uniform high means, in the females only the Somali, Rendille and to a lesser extent the Gabbra replicate the male trends. The Boran and especially the Burji, who showed high means in the males, reveal low to medium means in the females. It is thus the Boran and Burji females that are distinctive and atypical.

For palmar attributes, the main trend is the high female significance for peripheral hypothenar loops, a trend not replciated in the males who show no significance in any palmar attribute. For PHR and PHL in the females, the outstanding differences are between

the Rendille who who very high frequencies, and the others who show low to average ones. In the males, however, the Rendille fall well within the range of other Cushitic values, being if anything on the low side.

vi) The Maasai, Dorobo and Luo/Kariomojong

All these groups display very few significant differences in both sexes. The Maasai and Luo/Karimojong groups moreover, contain samples of respectable sample size, emphasising further the apparent homogeneity of the groups.

Discussion

From this analysis, we can provisionally distinguish four types of intra-ethnic group variation. The first type is the greatest heterogeneity-considerable differences between samples for a limited number of related attributes. The Central Bantu and Luyia Bantu fall into this category. The second type is where most intra-group samples show considerable similarity for attributes that show significance, but where one or two populations differ radically from the rest. The Coastal Bantu belong to this type, with the Digo differentiated from the rest for digital attribtues and the Taita for certain hypothenar attributes. The Kalenjin also conform, with the Kony showing radically lower digital ridge count means than theother Kalenjin samples. Thirdly, there is a type where overall homogeneity in one sex is not replicated in the other, but as in the case of the Kalenjin because sample sizes in the females are too low to significantly mirror the same trends significant in the males, but because there is considerable inter-sex differences

in some samples. The Cushites, where the Boran and Burji females show much lower means for digital ridge counts relative to the other samples than in the males, are the only example of this type.

The fourth type is where complete statistical homogeneity exists for all but a very few attributes, as is shown by the Maasai, Dorobo and the Luo Karimojong Groups.

In spite of these findings however, the analysis on the whole has failed to provide an objective criterion on which to measure relative homogeneity within ethnic groups. The procedure of comparing the number of significant differences between ethnic groups is open to distortion. If an ethnic group shows significance for an attribute, other attributes strongly related phenetically to it will also tend to show significance. For example the fact that the Coastal Bantu show significance for a digital ridge count means that it is likely to do so for other ridge counts. As there are more digital ridge counts than other kinds of phenetically related variable groupings, then showing significance on digital ridge counts will weight more than for instance, significance on thenar attributes.

Furthermore it is questionable to base an assessment of homogeneity on the extent of statistical significance. For instance the distinctive lower means for digital ridge counts of the Kony from other Kalenjin populations is evident not only in digital attributes that are significant in the males, but also in digital attributes that are not significant in both males and females. This example implies that type II error is probably important. In other words marked consistent heterogeneous trends of variation seem to exist without appearing as statistically significant.

8.2.3 Direct comparison of relative homogeneity between ethnic groups as depicted by correlation between variation for minimal population samples and variables that measure ethnic affiliation.

In view of the shortcomings of the previous analysis of inter-ethnic group trends of relative homogeneity, I employed a different approach that not only focuses specifically on each ethnic group in turn, but which also provides a standardised scale of measurement to assess comparative homogeneity and characteristic differentiation between groups on individual attributes.

For each ethnic group, the range of populations was recoded into two categories, the first consisting of all minimal population samples belonging to the ethnic group being assessed, and the second consisting of all samples not belonging to the ethnic group. A dummy variable was thus created in which members of the ethnic group were coded as 1, and the others as 0. The dummy interval variables, one for each ethnic group, were then correlated with each summary taxonomic attribute in turn using Pearson's r. Each correlation can thus be interpreted as a measure of the extent and nature of the relationship between ethnic group affiliation and attribute variation. The higher the correlation, the more distinctive and unique the range and magnitude of the within ethnic group attribute scores become, and hence the greater the chance of predicting the scores of a population belonging to that ethnic group. A high positive correlation furthermore, can be interpreted as showing that for that attribute, members of the ethnic group have high means or frequencies relative to the populations not in the ethnic group, and means and frequencies distinctively lower if the correlation is high and negative.

To save space the following analysis only reports the results of correlations with summary attributes. The results of the non summary analyses showed similar trends, and the magnitude of correlations were not any higher on the whole than those exhibited by the summary attributes reported here.

The results are contained in Table 8.10 for males and Table 8.11 for females. The first observation that emerges is that all correlations are low in both sexes, varying (ignoring signs) from 0.0 to 0.5. In other words the maximum amount of unique variance exhibited by any ethnic group for any attribute is 25 per cent. All ethnic groups therefore, are relatively heterogeneous and not very distinctive in relation to other groups.

Having made this general conclusion however, it is noticeable from the figures that there are some definite trends. The range of 0.0 to 0.5 though on the low side, is still an appreciable range, and higher correlations are not randomly scattered through the groups.

The Cushites show relatively high positive correlations in the 0.2 to 0.4 range in both sexes on all digital ridge counts, total finger pattern intensity and concentric whorls, though not on double loops. They are thus relatively characterised by high means and frequencies for these traits. On the opposite polarity, they exhibit moderately high negative correlations for interdigital and total palmar intensity, particularly in the males, and also for arches, implying low means and frequencies for these attributes.

The Maasai are one of the least distinctive groups, demonstrating only two correlations exceeding 0.2,* for left total

* This written analysis is based on figures at more than one decimal place, not on the summarised results of tables 8.10 and 8.11.

palmar intensity and left double loops in the males. Their apparent homogeneity in the previous analyses therefore, has hardly any taxonomic validity. The Dorobo on the other hand, follow the Cushitic trends for digital attributes, though not for palmar ones, for which the Dorobo are the least distinctive of all the male groups.

The Kalenjin are curious in that the male digital ridge counts are positive in the 0.2 to 0.3 range in the males, whereas in the females they show the same magnitude of correlation for these attributes, but negative. The females also exhibit negative correlations of -0.2 and -0.3 for TDR and TDL, but in the males only TDR follows the same trend. This apparent discordance between the sexes should not detract however, from the fact that the Kalenjin are relatively distinct. In fact they are the most distinct Nilotic group for digital attributes.

The Luo/Karimojong exhibit only two correlations exceeding 0.2, for TDL and LTPLMI, in the males. They are thus slightly more likely to have higher values for palmar triradial intensity measures. As in the case of the Maasai, moreover, the lack of distinctivness here contrasts notably with the lack of significance found for most attributes in the previous analysis.

For the Luyia, the highest correlations occur in palmar attributes in both sexes. In the males the most distinctive are right hand attributes for palmar ridge count, interdigital and total palmar triradial intensity, whereas in the females the highest occur on left hand palmar attributes. In both sexes, the correlations are positive for these attributes. Digital attributes in contrast demonstrate very low correlations in males and slightly higher

negative ones in females. Both sexes however, correlate at 0.3 for left hand ulnar loops.

Considering the low sample sizes of the Luyia, and the number of populations affiliated to this group, the correlations are surprisingly high, especially as they attain on some attributes as high a range of correlation as other ethnic groups whose populations have much larger sample sizes.

The Central Bantu are particularly distinctive regarding palmar attributes, especially in the males where negative correlations for TDR and RTPLMI are -0.4 and -0.5 respectively. These relatively high negative correlations for palmar triradial intensity measures are replicated in the females to a slightly lesser degree. Palmar ridge counts show the highest positive correlation of any group in the males (0.3; 0.2) but are very close to zero in the females. Sex differences are also apparent for double loops, which are positive in females (0.2; 0.3) and negative in males (-0.2; -0.3).

Finally, the Coastal Bantu show very similar trends to the central Bantu regarding the distribution of higher correlations The major difference is however, that correlations for palmar intensity triradial measures though slightly larger than those for the Central Bantu in both sexes, are positive rather than negative. Like the Central Bantu moreover, the males show much higher correlations for palmar ridge counts than the females, though unlike the Central Bantu males, the correlations are negative. Correlations for double loops are also relatively high in the Coastal Bantu, and positive in both sexes, while loops are highish and negative.

The results therefore, indicate that there is some relationship between ethnic affiliation and dermatoglyphic variation.

Some measure of the extent of this relationship can be gained by examining the proportion of significant correlations (those 0.3 or higher) in all groups. In the males there are 50 out of 208 attributes or 24 per cent while in the females there are 38 out of 182 attributes or 21 per cent. The extent of relative homogeneity and distinctiveness moreover, can be gauged by looking at the percentage of significant correlations in each group. With the sexes combined (i.e. the number of significant attributes in both male and female groups) we get, in order of the highest proportion: 1) The Chusties (48.0%); 2) the Dorobo (38.5%); 3) the Coastal Bantu (28.9%); 4) the Kalenjin (26.9%); 5) the Central Bantu (23.1%); 6) the Luyia Bantu (15.4%); and 7) and 8) the Maasai and Luo/ Karimojong (3.8%). From these rankings it appears that the Cushites are the most distinctive and homogeneous, followed closely by the Dorobo, with other groups showing much lower proportions, reaching a very low percentage in the Maasai and Luo/Karimojong. It should be remembered however, that although these results seem to imply strong relationships between ethnic affiliation and dermatoglyphic variation for groups at the higher end of the rankings, their distinctiveness is only relative. In absolute terms, no group is highly or exclusively differentiated, as there are no high correlations. The overall picture therefore, is one of relative heterogeneity, rather than of relative homogeneity. This contrasts markedly with the findings of the analysis of variance (8.2,1.1) where a majority of attributes in each ethnic group failed to achieve significance.

## 8.3 Overall trends of relative homogeneity and affinity: the evidence from PCO latent vectors

The general analysis in Chapters 5 and 6 showed that owing to the considerable discordance between attributes that are not phenetically related, the overall structure of variation is incompatible with the existence of strongly differentiated taxonomic clusters. The results of the analyses of the preceding section reflect the findings of Chapters 5 and 6, as a strong correspondence between ethnic patterns of affinity and dermatoglyphic variation was not found. Nevertheless, there is some evidence that at least some moderate and significant relationship does exist between ethnic and dermatoglyphic affinities on some attributes, being strongest in the Cushitic Group.

There are two basic models that could explain this situation of an overall absence of taxonomic distinctiveness, linked with a moderate tendency for patterns of variation to follow ethnic patterns of affinity. The first model is where populations who share similar languages and culture at the present, have common ethnic and biological origins. Divergence in time and space has not been phylogenetic in nature, owing to periodic hybridisation not only with each other, but also with populations outside the group. The extent of hybridisation however, has not been on a sufficient scale either to reverse divergence or to eradicate completely the basic overall similarity of members of the group when compared to members of other groups.

The second model is where splinter groups from different ethnic clusters have converged in the not too distant past, and have adopted the language and culture of one of their number, thus forming in time a new distinct ethnic unit. In spite of the fact that some biological similarity between the populations has resulted

over time owing to hybridisation between them, or to adaptation for some attributes to the new environment, the degree of hybridisation and adaptation has not been extensive enough to eradicate the fundamental biological differences of the original members of heterogeneous origins.

Between these two extreme models it is possible to conceive of situations in which some members of a present ethnic group might for instance, have common origins, while other members have only recently migrated to the area and adopted the language and customs of the group. In such a case, we could expect to find underlying biological similarity between the former populations, and only superficial similarity between them and the migrant population or populations.

The notion that underlying biological similarity will persist for a long time in populations who have a common origin or who have shared a common linguistic and cultural heritage for many generations, is based on the fact that divergence of gene pools is not normally constant for every attribute. This in a situation of divergence, coupled with periods of weaker convergence through hybridisation, some attributes might show large differences between the populations, and others small differences. We would expect therefore, an overall structural similarity between such populations to take generations to disappear completely, except in the case of small populations who are more liable to be completely dispersed and assimilated by other groups sometime in the course of history.

Conversely, where members of a present ethnic group have heterogeneous origins, we would expect it to take a long time for

the structural biological peculiarities that characterised each member at the time of convergence to disappear completely, unless processes of hybridisation were very strong indeed, or unless the formation of the ethnic group took place a very long time ago.

The essential concept underlying these two models therefore, is that of "structural biological similarity". Given that a proportion of attributes will be expected to show marked differences between members of an ethnic group in both models, how can we measure "structural biological similarity"? As I showed in Chapter 6, the traditional methods of assessing multivariate similarity between populations, involving the calculation of genetic distances, their analysis through cluster or ordination analysis, and the depiction of relationships either through dendograms or via pairwise scattergrams of ordinated vectors, runs into difficulty when discordance is strong. The scattering of vectors in a pairwise fashion during ordination analysis, for instance, tends to show different patterns of relationship from one pair of vectors to another. Moreover I also showed that each scattergram, even that between the first two factors that provides the strongest solution, is not really an overall assessment of patterns of relationship, but just an assessment of such patterns for a limited set of attributes that load on the two vectors being scattered.

In view of the shortcomings of the traditional methods of displaying overall affinity, I have employed in the following investigation the graphical analytical technique employed in the detailed analysis of ethnic group inter-relationships in section 8.1.2.3. In this technique, structural similarity can be interpreted

to exist between two populations, if the shape of the line formed when the PCO vector coordinates are linked, is similar in the populations. The rationale for this is that the likelihood of two populations displaying the same pattern ordering from one vector to another for a number of vectors purely due to chance is remote. This can be seen more clearly if we translate the vectors into biological terms. If say, the first vector is largely one expressing the decorrelated taxonomic variance for radial ridge counts, the second for palmar ridge counts, and the third for hypothenar patterns, then if two populations show a V "shape", it means that both show relatively high finger ridge count means, relatively low palmar ridge count means, and relatively high frequencies for hypothenar patterns. Thus, if several populations within an ethnic group were to show the same V shape, and others outside the group were not, then such a pattern of similarity is not likely to be coincidental. In practise, there are many possible shapes, the possibilities increasing considerably with every additional vector considered.

Similarity in shape however, can exist without the two populations being particularly close in the absolute coordinates they display on individual vectors. Using again the previous example, two populations could both exhibit characteristic V shapes for the three vectors, and at the same time, could lie in quite different areas of the graph: for instance one population could lie at the top of the graph and a second near the middle. The two populations would thus be structurally similar, but heterogeneous in "distance". Such a finding could thus be interpreted as possibly showing a common origin between the two populations, and

that they have been diverging for many generations. If two populations show the same shape with little distance on individual vectors, then total homogeneity is implied. These examples of course, are simplified to illustrate the thinking behind the method. In practise we would not expect exact similarities in either shape or distance.

In using this method to examine detailed dermatoglyphic affinity between populations within ethnic groups, I originally produced graphs for both sexes for PCO vector coordinates from $\Delta g$ distances based on both summary and non summary attribute series (see Chapter 6). I also produced some graphs from the PCA component scores to be found in Appendix 3. Owing to the large number of graphs produced, I have limited the analysis here to considering only those displaying affinities within ethnic groups based on PCO analyses based on summary attributes (see Appendix 3). Occasionally, I have also shown an equivalent graph for PCO analyses based on non summary attributes, if it clarifies a particular issue.

One of the problems encountered in graphically linking PCO latent vector coordinates is how many latent vectors to include in the graph. The fewer vectors that are included, the easier it becomes to interpret the figures. However, to include too few is to represent only a portion of the total variance. In the following figures, based on PCO analyses with summary attributes, I have included only the first four latent vectors. This is acceptable as they account for close to 75 per cent of the variance in both sexes. To have also considered vectors 5 and 6 would have improved the total variance by only just over ten per cent, while

at the same time adding greatly to the task of interpretation.

Owing to the fact that the number of populations in some of the larger groups exceeds the maximum of six permitted by the graphing program (see 8.1.2.3) these groups are represented by two or more graphs. The sub-division of affected groups was done on a regional basis. Thus for example, the Kalenjin were divided into south-western and north-eastern divisions. During analysis, intra-group similarities in different graphs for the same sex were gauged by superimposing the graphs. Unfortunately however, it is not possible for the reader to check the results in this way, as the pages are fixed. Nevertheless a close scrutiny by eye should justify any stated similarities. Finally, it should be stressed that there is no expected strong correspondence between lines in the male and female series, as vectors are not precisely equivalent or in the same order (see Chapter 6.3).

### 8.3.1 Structural dermatoglyphic similarity and patterns of affinity between minimal population samples within ethnic groups

In the following investigation, the detailed dermatoglyphic structural affinities between minimal population samples within ethnic groups will be described through the implementation of the techniques outlined in the preceding introduction, and contrasted with the patterns of linguistic and cultural affinities that pertain within each ethnic group. The format as each ethnic group is investigated is firstly to outline patterns of linguistic relationships within the group in greater detail than was given in Chapter 1.2, and secondly to present and discuss the dermatoglyphic evidence.

8.3.1.1 The Cushitic Group

The Somali, Rendille, Gabbra and Boran all belong to the Eastern Cushitic Group of Afro-Asiatic languages (Sutton 1974:83). The linguistic relationships between the Eastern Cushites are more precisely delineated by Spencer (1973:5) who uses as a measure the similarities they display in a sample of 100 commonly used words. The most similar on this basis are the Rendille and Somali, with 38 words in common, compared to less than 20 for the Boran when contrasted to the Rendille and Somali. The Gabbra, though not mentioned by Spencer, are linguistically very close to the Boran, being classified in the same Easter Cushitic sub-group: the Galla (Sutton 1974:83).

When we look at Figure 8.7, the males on the whole show an appreciable similarity in shape between all the Cushitic populations, particularly for the 2nd and 4th vectors, with only the Gabbra anomalous on vector 3. The general pattern moreover, approximates to that of the detailed linguistic evidence. On vector 1 in particular, which accounts for over 40 per cent of the variance, the Gabbra and Boran are clearly similar and differentiated from the Rendille and Somali who also show close similarity. The Burji on the other hand, are clearly separate from the others in distance.

The females however, demonstrate greater heterogeneity, and a different pattern of similarities. Unlike the males, where the overall shape was similar in all the populations, the females in Figure 8.8 appear to differentiate two groups on grounds of shape, the Boran/Burji and the Somali/Gabbra/Rendille, with the

Rendille abberrant on the 2nd and 4th vectors. This apparent heterogeneity in structural shape however, is not supported by Figure 8.9, which gives the distribution for the same female populations for vectors based on 110 non summary attributes. The structural similarity between all the female Cushitic populations is well demonstrated, especially on vectors 3 to 6, though the Burji and Boran are clearly separated on grounds of distance on the first two vectors. The females therefore, though showing overall structural similarity, do not as the males do, very closely according to internal linguistic affinity.

8.3.1.2 The Dorobo

From a socio-cultural perspective, the Mukogodo and Ngwesi are similar, being hunters and gatherers who in this century have adopted pastoralism and some significant aspects of Maasai culture (Spencer 1973: Appendix ). The Mukogodo however, have always apparently lived in their current location, but the Ngwesi are reputed to have moved to the Meru area in the late nineteenth century, from where they were resettled into the Dol Dol area in colonial times (Spencer 1973:202). An even more radical difference between the Ngwesi and Mukogodo spoke exclusively until very recently an archaic Cushitic language whose precise classification is problematic, but which has been provisionally assigned as Eastern Cushitic (Sutton 1974:88). Spencer mentions however, that today the Mukogodo are bilingual, speaking Maasai as well as their ancient Cushitic tongue.

One of the fascinating aspects of the Northern Kenya Dorobo groups is their apparent facility in adopting the language

and culture of their more powerful neighbours. In 1900 for example, the El Molo Dorobo from the shores of Lake Turkana spoke a Cushitic language. Today they have totally forgotten their ancestral language, and speak Maasai (Samburu) even though according to Spencer, there have been very few instances of intermarriage with outsiders, (Spencer 1973:214). It would seem therefore, that the Laikipiak Dorobo are good examples of small-scale peoples that can easily change their language and culture without necessarily radically altering the character of their gene-pool.

There are two conflicting theories as to the historical origins of the Laikipiak Dorobo. Early anthropologists such as Johnston (1902:857) and Dundas (1908:136-9) thought they represented the remnants of an aboriginal race. Spencer on the other hand, champions the more modern view that the Dorobo are

> "a residue of the existing peoples of the area. For as long as there have been historical records, there have been such groups, absorbing odd members of defeated and impoverished tribes, forming relations with neighbouring tribes and other Dorobo groups, and occasionally becoming members of richer tribes with whom they associated".
>
> (Spencer 1973:218)

If the "aboriginal race" hypothesis is correct, the Mukogodo and Ngwesi ought to show distinctive and radically different structural shapes from other Kenyan populations. If Spencer is correct, however, we should expect the Ngwesi to resemble the Samburu closely biologically and possibly also some of the Meru populations. The Mukogodo should also be similar to the Maasai/ Samburu, as they are reputed to have associated closely with the Momonyot in the last two generations, a Dorobo group reputed to be a defeated Laikipiak Maasai remnant from the late nineteenth century (Spencer 1973:202-3).

Figures 8.10 and 8.11 show however, that both the Mukogodo and Ngwesi are very similar in shape, especially for the first two vectors in Figure 8.10 and for the first four in Figure 8.11. They do however differ in distance and in the shape of the latter components, confirming that they are not identical types.

The shape of the Mukogodo in particular, moreover, is almost identical to that of the Burji, Somali and Rendille lines in Figure 8.7. When we compare the Ngwesi and Mukogodo shapes and distances to the Samburu and Maasai however, there is hardly any resemblance at all (see Figure 8.12). The only Nilotic male population shows the remotest resemblance is the Turkana to the Ngwesi. None of the Meru male populations show any similarity to the Ngwesi either. The inescapable conclusion therefore, is that the Mukogodo and Ngwesi are of Cushitic extraction as far as the dermatoglyphic evidence is concerned. It confirms the relative biological stability of these groups over time, in the face of considerable cultural and linguistic change.

8.3.1.3 The Maasai, Luo and Karimojong Groups

According to the linguistic evidence, the populations in the Maasai group (Samburu, the Keekonyokie from Kadjiado and the Purko from Narok), should be much more similar to each other than to the Itunga/Karimojong peoples represented here by the Iteso and Turkana. Both these groups moreover, are generically clusters of the Plains Nilotes, a linguistic division to be contrasted at the same scale with the River-Lake Nilotes divison, of which the Luo group is a branch. If linguistic and dermatoglyphic

affinity are compatible, we should expect to find the Luo out on a limb, sharply segregated from two sub-clusters; corresponding firstly to the Turkana and Itsso and secondly to the Maasai and Samburu.

In Figure 8.12 however, this is not the case. The male populations mostly show considerable heterogeneity in both shape and distance, with the closest resemblance being between the Narok (Purko) Maasai and the Samburu and between the Turkana and Luo. The Kadjiado (Keekonyokie) Maasai are in the polarity of the other Maasai populations, but do not resemble them significantly in shape. The Iteso resemble the Luo on the latter vectors, but differ radically on the first vector.

The picture presented by the females however, is clearer (Figure 8.13). At the high polarity the Kadjiado and Samburu Maasai resemble each other closely in both distance and shape on all four vectors, while the Turkana and Luo are equally similar at the lower polarity. The Iteso are more similar to the Luo/Turkana than to the Maasai, but their shape is "thrown out" by vector 2. On the whole therefore, there is some evidence of overall structural similarity between these Nilotic populations, but the patterns do not correspond closely to expected patterns of linguistic affinity. The dermatoglyphic evidence differentiates only two loose groups, the Maasai and Turkana/Luo/Iteso.

8.3.1.4 The Kalenjin Nilotes

The Kalenjin are a linguistic sub-division of the Highland Nilotes, a generic division at the same scale as "Plains" and "River-Lake" Nilotes. A detailed internal linguistic classification

for the Kalenjin is given by Heine (1971). From the evidence of an extensive vocabulary survey, he distinguished three sub-divisions, a) the Elgon group, represented by the Kony here, b) the Nandi group (Nandi, Kipsigis, Tugen, Keyo, Marakwet) and c) Northern Branch (the Pokot). The Elgon group is more closely related to the Nandi group than either group is to the Northern Branch.

These differences however, are relative. On the whole, the Kalenjin are linguistically fairly homogeneous, with the lowest percentage of common words between any two populations being sixty per cent. Within the Nandi group, over seventy-five per cent of words are held in common, with the Nandi and Kipsigis, and the Keyo and Marakwet having over ninety per cent of words in common. The internal affinities of the Nandi group therefore, are as follows: the Nandi/Kipsigis on one polarity with the Keyo/Marakwet on the other, with the Tugen occupying the middle.

The dermatoglyphic relationships are shown on four figures, Figures 8.14 and 8.15 for the Kony, Nandi and Kipsigis male and female samples, and Figures 8.16 and 8.17 for the male and female Tugen, Keyo, Marakwet and Pokot samples. The Kony are clearly very distinct in both sexes, especially on vector 1. In the females moreover, the Kony shape is similar to those for all populations except the Nandi and Kipsigis. In the males however, the Kony shape is quite different from that of other populations.

The Pokot, which are linguistically the most distinct of the Kalenjin, are not dermatoglyphically distinct, being similar in size and shape to the other north-eastern Kalenjin in both sexes.

The Nandi group demonstrates some heterogeneity in both size and shape. In the males, the Kipsigis are differentiated from a cluster (i.e. the Nandi, Tugen, Keyo, Marakwet and Pokot) who are relatively homogeneous in shape. The Tugen on vector 3 are the only population resembling the Kipsigis, in shape. Within this cluster moreover, the closest affinity is between the Tugen and Keyo, not between the Keyo and Marakwet as one would expect from the linguistic evidence. In the females however, the Nandi and Kipsigis are almost identical, both differing in shape though not size from the Tugen, Keyo, Marakwet and Pokot cluster, which shows strong resemblance in shape. In the females, the Keyo and Marakwet are the most similar.

The results therefore, are contradictory between the sexes. On the whole, the females exhibit a closer correspondence to linguistic affinities than the males. Except for the Kipsigis, the Kalenjin generally show overall regularities of shape, though there is some considerable fluctuation in size, particularly in the case of the Kony. In both sexes, the structurally most similar regional group is the north-eastern branch. The Pokot, which are the most linguistically separated, are not dermatoglyphically distinct from the north-eastern group, implying that geographical proximity transcends linguistic difference in this area.

Finally, when the Kalenjin are compared to the other Nilotic populations, the characteristic shape of the north-eastern Kalenjin group is similar to that of the Narok Maasai and Samburu in the males. In the females however, the shape is not similar to any other Nilotic group. The relationship between the Kalenjin and other Nilotes, therefore, is weak.

8.3.1.5 The Luyia Bantu

> "From the ethnic, linguistic and political points of view, the Bantu Kavirondo (Luyia) do not form a homogeneous group, although they are clearly distinct from the surrounding tribes".
>
> (Wagner 1949:19)

This opinion is shared by Were, who notes that the Luyia "are a hybrid community founded by people of different origins and culture" (Were 1974:187). The oral history of the Luyia as described by Were (1967) and Osogo (1966), is extremely complicated a picture of successions of small and large scale migrations of small groups and clans from diverse directions, and includes considerable assimilation of offshoots of neighbouring Kalenjin peoples. For instance, Were states that

> "between AD 1571 and 1652 a small column of Maasai or Nandi came from the eastern direction and settled in Idakho. They were later bantuised and lost their original language and culture".
>
> (Were 1974:188)

It is not surprising that the Luyia area should be so culturally heterogeneous. It is a natural topographical corridor, bordered in the south west by Lake Victoria and the Yala swamps, by the Nandi Plateau escarpment in the east and south east and by the Uasin Gishu, Trans Nzoia and Mt. Elgon highland areas in the North. Population movement into the Luyia area by migrating Bantu groups was only possible from the south, following the Lake Victoria shore and surrounding plains in the time preceding the Luo occupation of that region; or from the west and north west from Uganda. The eastern and north-eastern areas were blocked by generations of Kalenjin and Maasai occupation. The Luyia therefore, is a population melting pot, with routes admitting incomers, but with few options available to expand further.

In spite of this unstable history however, the Luyia peoples do form an ethnically distinct cultural block, and in recent times have assumed a common sense of identity. There is very little detailed linguistic information regarding the Luyia sub-tribes, but we do know that different dialects are mutually intelligible to a varying degree. On the basis of limited vocabulary lists, Wagner (1949) distinguishes four linguistic divisions: 1) The Wanga, Butsotso, Bunyore, Tiriki, Isukhe and Idakho; 2) Bukusu and Tachoni; 3) Bunyala and Samia; 4) Maragoli (Wagner 1949:26). Not all the sub-tribes sampled here however, are included in this classification. From the sketchy linguistic information of Osogo (1966:3) it is possible to ascribe the Marama to the Wanga group. This leaves the Bukhayo and Marach as linguistically unplaced.

The Luyia dermatoglyphic affinities are presented in six figures (Figures 8.18 to 8.23), comprising three for each sex, divided into three regional groups. The outstanding characteristic to emerge from these results is the great heterogeneity in both size and shape, and a lack of concordance between the sexes. This at first sight would appear to confirm the heterogeneous origins of the Luyia. This conclusion however, must be treated with some caution, as most of the Luyia samples are very small, usually around 25 for each sex.

It is interesting to note however, that the Bukusu, the only large Luyia sample in the males (Figure 8.22) have a shape that is similar to that of the Kony (Figure 14). This similarity between the Bukusu and Kony is also shown by the females (see Figures 8.15 and 8.23), even though both the Bukusu and Kony female samples are

small. From this evidence therefore, it appears that the Luyia Bantu speaking Bukusu have dermatoglyphically a much greater affinity with the Elgon Kalenjin speaking Kony, than with the other Luyia Bantu speakers. This tends to confirm Ehret's theory that the Bukusu are a Highland Nilotic people who adopted a Bantu language, as many of their culture features today are still typical of the Highland Nilotic peoples (Ehret 1974:160).

Finally, I have not mentioned the Gusii so far. This is because they are a problematic people to classify, showing many cultural similarities to some Luyia groups (the Maragoli in particular - Wagner 1949:25), but at the same time reputed to be linguistically similar to the Central Bantu (Sutton 1974:82). The Gusii here however (Figures 8.18 and 8.19) show some superficial similarity to the Maragoli and Tiriki in the males, and to the Tiriki and Idakho in the females (shape not distance). They do not show any resemblance to any Central Bantu population. The closest similarity of the Gusii is in fact to the Kadjiado Maasai, showing a very close resemblance in both sexes when the equivalent graphs are super-imposed.

8.3.1.6 The Central Bantu

The Central Bantu in the Mt. Kenya highland areas are relatively homogeneous in language and culture. There is for instance, a considerable degree of mutual intelligibility of dialects, and their general cultural similarity has led them to assume a common sense of identity in modern times.

The Central Bantu belong linguistically to the "Highland Bantu" division (see Chapter 1), which includes many northern

Tanzania Bantu populations and the Gusii and Kuria of south western Kenya, as weel as the Segeju of the south Kenyan coast. Little work however, has been done on the detailed linguistics of this large division (Sutton 1974:81). For the Central Bantu the major published source is still Lambert (1949). He states that

> "the languages or dialects .., divide themselves into to main groups. One includes the Kikuyu, Ndia, Gichugu, Embu and Mbere, and is closely related to Kamba ... The other group includes Meru (Igoji, Miutoni, Imenti, Tigania and Igembe), Muthambi and Mwimbe and shows closer affinity with the coast languages (Pokomo, Nyika (i.e. Mji Kenda)) than does the first .... Chuka perhaps comes somewhere between the two groups, and Tharaka is generally regarded as a conglomerate of similar dialects which are scarcely yet merged into uniformity; its present leanings are towards the Meru group, but the similarities with Kikuyu are marked and there has been some borrowing from Kamba".
>
> (Lambert 1949:5)

Bennett (1967) refines the classification of the Kikuyu group slightly by distinguishing between two Kikuyu dialects, a northern and southern one. The former corresponds to the present Kiambu and Muranga samples, and the latter to the Nyeri. In addition, Bennett examines the distribution of Dhal's law amongst the Central Bantu, a linguistic phenomenon confined to the East African Bantu. It is interesting that whereas the Meru and Kikuyu follow the Luyia and Gusii in displaying the phenomenon, the Kamba and Coastal Bantu do not. There is thus a suggestion of linguistic heterogeneity amongst the Central Bantu (Bennett 1967).

The oral history of the Central Bantu has been explored by Lambert (1949), Muriuki (1974), Fadiman (1973) and Saberwell (1967). These authors present a complex picture of population movement into the present areas of occupation. The general consensus is that all the population groups except for the Kamba probably originated

from the Kenya Coast, having migrated from there in a series of migrations following the River Tana. (Lambert and Fadiman are especially influential in proposing this hypothesis). The latest wave of migration was the peopling of the present southern Kikuyu areas around the Aberdares, which is believed to be recent (last two centuries or less). The Kamba are in contrast believed to have come from the Mt. Kilimanjaro region, following a south/north migratory route (Lambert 1949:28). The main exception to these oral historical traditions are those for the Chuka, who according to Lambert believed themselves to be the original inhabitants of their land, and who until colonial times kept themselves isolated from their neighbours, being

> "the butt of all the war-like tribes in the neighbourhood.., they came to regard all strangers with suspicion, and they therefore became more and more exclusive and isolated".
>
> (Lambert 1949:3).

The oral historical traditions therefore, present a picture of potential heterogeneity of origins, while the linguistic and cultural evidence is one of relative homogeneity. The dermatoglyphic evidence is presented in Figures 8.24 to 8.31. Owing to the number of populations in the group, four figures are necessary for each sex. Populations are grouped according to approximate geographical proximity.

In the males three types of characteristic shapes are discernable;

Type 1: is typified by the Imenti, showing a sharp climb from vector 1 to 2, a fall in vector 3 and a levelling off in vector 4. In addition to the Imenti the Tigania, Chuka and Nyeri Kikuyu fall

into this group. From the point of view of "distance" the most differentiated population in this group is the Chuka.

Type 2: is classically depicted by the Embu, showing a steep rise from vectors 1 to 2, a level progression from 2 to 3, and a fall from 3 to 4. Variants of this shape are the Ndia and Muthambi. A much more flattened version of this type is also shown by the Muthambi and Igoji, who are very similar to each other.

Type 3: is an inverted version of type 2, with a descent from vectors 1 to 2, a levelling off between 2 and 3, and a rise from 3 to 4. Classic examples are the Mbere and Tharaka. The Muranga Kikuyu also conform to this type, but are not so similar on the fourth vector. Other less classic variants are the Igem_be, Kiambu Kikuyu and Kamba. The latter two in fact are fairly distinct from the others.

In the females the most characteristic shape is one showing a rise from vectors 1 to 2, a fall from 2 to 3, and a rise from 3 to 4. Classical examples are the Ndia, Embu, Mbere, Muthambi Chuka, Tigania and Nyeri Kikuyu. The Kamba also fall into this group from the evidence of the last three vectors, but differ in showing a steep fall from vectors 1 and 2 rather than a rise.

Variants of Type 2 in the males (usually much flatter in the females) are the Muranga and Kiambu Kikuyu, Igembe, Imenti and Tharaka.

The Mwimbi females are in between the above two types, showing a horizontal straight line.

The dermatoglyphic affinities therefore, show no tendency to replicate linguistic affinities in either sex. Where is heterogeneity in shape in both sexes. On the whole, the females are much more homogeneous than the males, showing not only less variety in shape, but also much tighter central clusters of lines. In both sexes moreover, there is some evidence of both geographical and ecotypic affinity. In Type 3 in the males for instance, all populations listed, with the exception of the Muranga, occupy lower lying and more arid country. In the females there is a great deal of homogeneity for both shape and distance for the geographically contiguous Nyeri/Kikuyu, Ndia, Embu and Chuka. The discordant trends between the sexes however, make more detailed analysis problematic.

8.3.1.7 The Coastal Bantu

The coastal Bantu peoples (Taita, Mji Kenda, Pokomo), have strong oral traditions that speak of a recent common origin. They are reputed to have migrated from a place called Shungwaya, somewhere on the Somali coast, some 4-500 years ago.

The authenticity of these traditions has been disputed by Morton (1972), who argues that there is no mention of the Shungwaya myth prior to 1897, in the European and Arab literature. He concludes that the myth arose and spread amongst the coastal peoples of Kenya during the early years of the rule of the British East Africa Company (1888-1897), in response to a growing spirit of unity, particularly amongst the Mji Kenda.

Spear (1974, 1977), however, questions Morton's appraisal. He accuses Morton of being "ahistorical", for "arguing negatively from

a small and baised body of evidence" (Spear 1974:75), and claims, in reference to the supposed rapid spread of the Shungwaya myth in the late 19th Century, that:

> "such an intricate and coordinated manipulation of at least a dozen independent traditions during such a small period is extremely improbable - I would say impossible - even if the initial impact of colonial rule had been sufficient to warrant it."
>
> (Spear 1974:76)

That twelve independent bodies of oral tradition should agree so closely on the recent migrations from Shungwaya is seen by Spear (1974) as crucial evidence for the common origins of the Mji Kenda and the other coastal peoples. He notes, moreover, that from a cultural perspective, the Mji Kenda especially "share a total complex of cultural traits with one another" (Spear 1974:72). Finally, Spear suggests that the linguistic evidence also strongly supports a hypothesis of recent common origins. He refers to Guthrie (1959) and Bryan (1959), who group the Taita, Mji Kenda and Pokomo into one group of closely coordinate languages:

> "Linguistic thory holds that coordinate languages probably share a common origin in a single proto-language. In other words it is more probable that two coordinate languages evolved from a single ancestor than one borrowed extensively from another".
>
> (Spear 1974:71)

The work of Hinnebusch (quoted by Spear 1977), has considerably clarified the interrelationships of the coastal Bantu languages, since Spear's 1974 article. According to Hinnebusch's analysis, the proto north eastern Coastal Bantu languages fall into four groups: (a) Saghala (one of the Taita dialects), (2) Sabaki (Pokomo, Mji Kenda and Swahili), (3) Seuta and (4) Ruva (both latter groups confined almost totally to Tanzania). Interestingly, Saghala

is the least related of the four major groups.

The Sabaki group is further sub-divided into two main groups, the Swahili and Mji Kenda/Pokomo. Thus linguistically, the Mji Kenda and Pokomo are very closely related to each other, compared to the Saghala Taita. A substantial section of Taita, moreover, the Dabida, speak a language/dialect that is not even included in the north eastern Coastal Bantu family by Hinnesbusch.

Hence, having accommodated the fresh linguistic evidence, Spear (1977:240) produced a more elaborate reconstruction of the historical interrelationships and patterns of movement of the Kenya Coastal Bantu:

1) The north eastern Coastal Bantu dispersed from a primary site near Mt. Kilimanjaro to a series of secondary areas, each dispersal forming the nucleus of the present four major linguistic divisions.

2) The Sabaki group migrated up the Kenya Coast as far as Brava in Somalia, where they settled around the eleventh century.

3) By the mid sixteenth Century the Cushitic Galla had forced them to move southwards again. There were two migrations. The first was that of the Digo, who settled around Kwale south of Mombasa. The second involved the remaining Sabaki, who followed soon after. The ancestral Pokomo settled in the Tana River area, and the rest congregated south of Mengea Hill near Malindi, where they dispersed to the present Mji Kenda locations. An offshoot moreover, went to the Taita Hills, where they were assimilated by the resident Saghala and Dabida speakers.

4) Finally the Kauma (one of the Mji Kenda), split from the Ribe, and the Mji Kenda absorbed the Rabai who had previously

inhabited the area. The Duruma were formed by an amalgam of coastal slaves and immigrant Kamba and Digo.

The dermatoglyphic structural interrelationships are portrayed in Figures 8.32 and 8.33 for the Mji Kenda males and females respectively, and in Figures 8.34 and 8.35 for the Pokomo Taita and Taveta.

Looking first at the Mji Kenda, the male samples reveal a strong similarity in shape between the Digo, Chonyi and Mid Mji Kenda, although there is some distance between the Digo and the others on the second vector. The Giriama however, exhibit a different shape, in particularly failing to show the sharp fall from vector 1 to 2 characteristic of the other samples.

The females show similarity in shape for the Giriama, Chonyi and Mid Mji Kenda samples. The Digo, however, are very strongly differentiated on vector 1, but follow the characteristic shape of the other samples on the remaining three vectors.

The Pokomo and Malakote male samples do not appear as similar in Figure 8.34, which tentatively indicates that the Malakote have different origins from the main Pokomo group. It is interesting however, that the Malakote are much more similar to the Mji Kenda samples than the Pokomo are. Nor are the Malakote strikingly similar to the Cushitic samples (Figure 8.7), which we might expect, given the reputedly significant intermixture of the Malakote with the neighbouring Orma and Somali Cushites.

The Pokomo male sample is clearly different in shape from the Mji Kenda samples, withwhom we would expect close affinity given the historical situation outlined by Spear (1977). The female Pokomo sample, in constrast, shows a close similarity in shape with the other female Mji Kenda samples.

Finally, the Taita and Taveta in both sexes show little similarity with one another, confirming that although today they may be regarded as kindred peoples by outsiders, any such affinity is due to geographical proximity rather than any basic biological, linguistic or cultural similarity. Both the Taita and Taveta, moreover, show little similarity with either the Pokomo, or the other Mji Kenda samples

The interpretation of this evidence is made difficult by the contradictory evidence presented by the male and female series on three vital areas. The first is the anomalous position of the Giriama in the males but not the females; the second is the anomalous position of the Digo in the females but not the males; the third is the fact that the Pokomo are not similar to the Mji Kenda in the males, though they are in the females.

On the whole, however, there is sufficient agreement between the sexes to indicate that the hypothesis of common origins for the Mji Kenda and Pokomo is probably justified. The samples from these groups are much more homogeneous in shape than any of the other Bantu ethnic groups examined in this chapter. The main feature that at first sight strongly contradicts Spear's (1977) historical summary is the failure of the Taita to show any similarity to the Mji Kenda and Pokomo. The Taita sample here, however, is from the Dabida section, not the Saghala one, where the main group of original Sabaki speakers is supposed to have settled. The fact that the present sample of Taita is so different from the Mji Kenda and Pokomo in fact confirms the absence of close linguistic affinity between the Dabida Taita and the group of Sabaki speakers.

# CHAPTER 9

## The Extent of Dermatoglyphic Ecotypic Variation Exhibited by Minimal Ethnic Population Samples

### Introduction

In this chapter I shall undertake a preliminary examination of the possibility that dermatoglyphic variation in Kenya is related to climate and ecology. At first sight this is an unpromising area of inquiry. Firstly Hiernaux (1968[a]; 1977) failed to find any significant correlations between the variation of arches, loops and whorls of 54 sub-Saharan African populations and three climatic variables (rainfall, temperature and humidity). The highest correlation (-.1) was between rainfall and whorls, a figure which is not significant. In contrast Hiernaux found correlations as high as .55 between these climatic variables and anthropometric means.

Secondly there is an intuitive problem. It is difficult to rationalise why any dermatoglyphic configuration should be advantageous or disadvantageous in any given climate. Why, for instance, should a person with a whorl rather than a loop be favoured in an arid environment rather than a wet one? There is no speculative reason that easily springs to mind, in contrast to other polygenic attributes, for example, skin colour.

In spite of these unpromising indications, however, it is worth pursuing the possibility that dermatoglyphic variation in Kenya might be related to climate and environment. The justification for this view rests not so much on the premise that all factors that might account for the variation should be

investigated, if only to eliminate them, but rather on the fact that a close scrutiny of the univariate results in Appendix 1.4 seem to suggest some relationship with ecology for some attributes. For example, a cursory examination of the variation of total radial finger ridge counts (Appendix 1.4.1:4) reveals a rough trend that populations with higher means (e.g. Somali, Rendille, Gabbra, Digo) tend to be low altitude populations, while populations with low means (e.g. Kony, Bukusu, Taita, Pokot), are high altitude populations.

## 9.1 Approach and methods

### 9.1.1 Choice of environmental and climatic variables

The choice of appropriate climatic and environmental variables for this study is dictated primarily by practical constraints. For a climatic or environmental variable to be suitable for analysis, detailed records must be available from all the areas sampled for dermatoglyphic traits.

In Kenya, however, only a limited number of climatic and environmental variables are recorded, and of these, only some are measured throughout the country. Information relating to remoter and economically less developed areas is generalised, and often unreliable. In fact the only primary variables that meet minimal requirements of coverage are topography, soiltype, rainfall, and vegetation. Other variables such as humidity, temperature, evaporation, solar radiation are only measured in key areas. For example Griffiths notes in the case of solar radiation that:

> "it is unfortunate that measurements of incoming solar radiation energy are so rare in East Africa; only three stations, all in the Nairobi region, have a continuous radiation record extending over a few years."
>
> (Griffiths 1961:86)

In view of these shortcomings, I decided to limit this preliminary inquiry to two variables, altitude and rainfall, which are the most extensively and reliably covered. Although these variables are not in themselves likely to be direct selective agents, they nevertheless jointly correlate highly with climatic variables that might so be, such as temperature and humidity. In addition altitude and rainfall are major factors in determining the nature of major habitats, influencing not only patterns of human habitation, modes of livelihood and population density, but also the nature of the fauna and flora of regions - particularly fauna associated with the transmission of certain infectious diseases. If climate and environment do correlate at all with dermatoglyphic variation in Kenya, rainfall and altitude are the two variables most likely to show associations and to indicate areas of promising further investigation.

### 9.1.2 Sources for the distribution of altitude and rainfall in Kenya used in the following analysis

The most detailed information for altitude can be obtained from Survey of Kenya Topographical Maps which cover the country at two detailed scales. The populated southern, western and central areas are covered by well over a hundred maps at 1:50,000, and the eastern and northern areas by a similar number of maps at 1;100,000. I found it impossible however, to gain access to most of these maps as they are not in university libraries, and are too expensive to buy as a complete series. Instead I relied on the altitude contours (still very detailed) provided in the Survey of Kenya Topographical maps scale 1:250,000, of which forty

nine are needed for complete national coverage. I managed to obtain a complete series ensuring permanent access to their information.

Distribution maps for rainfall was more problematic. The only detailed records are to be found in the Survey of Kenya scale 1:250,000 Special Maps entitled "Climate and Vegetation". Unfortunately only the south western area of the country is covered in this series - only four maps available out of forty nine needed to achieve national coverage. Of these four maps moreover, I only succeeded in buying one owing to Government restrictions on their sale. I had to rely therefore, on less detailed sources: a) the Atlas of Kenya published by the Survey of Kenya which has a detailed map of rainfall distribution at a scale of 1:3000,000 and b) on the rainfall map given in Morgan and Shaffer (1966:32). All figures for rainfall in the following analyses refer to mean annual rainfall and are measured in inches.

### 9.1.3. Allocation of individuals and populations to altitude and rainfall zones

It is not possible to be entirely objective about the allocation of samples to rainfall and altitude zones. Some populations, for instance, (such as the Tugen) occupy territory which shows marked differences in altitude and rainfall. The procedure followed here is to allocate the samples according to the precise altitude and rainfall at the sampling location. Thus the Tugen for example, were sampled at the greatest population concentration in the high rainfall and altitude portion of their territory (the upper slopes of the Tugen Hills), and the sample is

thus typical of this section of Tugen. It does not claim to represent those Tugen who habitually inhabit the more arid lowland plains. Nor is it possible to take into account the possibility that some Tugen might regularly migrate from one zone to another.

Another factor that needs consideration concerns those samples of peoples that were collected in areas not typical of their normal habitat. The Turkana sample for instance, was collected in Kitale, a high altitude and rainfall area untypical of the harsh desert conditions of Turkana land, from whom the parents of the sampled children had migrated from. Similarly the Boran and Gabbra from Marsabit Town, and the Luo from Nairobi, are all comparatively recent and often temporary migrants from lower lying and hotter areas. For these migrant samples it makes more sense to allocate them to zones of immediate origin rather than to zones of residence, as it is the zone of origin that is most likely to have influenced their genetic constitution.

9.1.4 Categorising altitude and rainfall zones

Altitude and mean annual rainfall are continuous scales of measurement. Altitude in Kenya ranges from zero to over 17,000 feet, and mean annual rainfall from less than one to over 80 inches. Given detailed records, it would thus be possible to assign to a sample a very specific figure for altitude and rainfall. We could say, for example, that sample A was located at 3,553 feet in altitude with a mean annual rainfall of 42.3 inches. Such scientific precision, however, is inappropriate here, due not only to the absence of very detailed records, but also to factors such as local population mobility and differences in terrain within a

sample capture area. For example in the highland areas of Kenya the walk to school from home could result in a difference in altitude of several hundred feet.

Changing the perspective from the sample capture area to the broader population, it is apparent that significant local mobility can take place even in one generation. Pastoralists can move from one grazing area to another, which can result in a change of altitude and mean annual rainfall. Agriculturists can move from a densely populated high rainfall locality, to a more marginal agricultural area where the rainfall is less, but where more land is available for cultivation. Even change of residence through marriage can result in substantial changes in altitude and mean annual rainfall in some areas (particularly in the Meru and Taita Districts).

Much variation in altitude and rainfall, therefore, is bound to be non significant from a genetic point of view. Any genetic selective changes arising directly or indirectly from altitude and rainfall would be soon dissipated by local patterns of mobility. Nonetheless it is also apparent that local patterns of mobility will only significantly operate within a portion of the total altitude and mean annual rainfall ranges in Kenya. It is very unlikely, for example, that, historically, a significant proportion of high altitude, high rainfall dwelling agriculturalists will have moved temporarily to a desert low lying environment, or that desert dwelling pastoralists will have moved for a period to lush high altitude pastures. In fact until the colonial and post colonial eras, we can assume that movement from one major type of ecological zone to another by large numbers of people was very

uncommon, in contrast to movement within such zones, and in marginal areas of overlap.

In view of these arguments, I have categorised altitude and rainfall into four general divisions for analytical purposes, as outlined at the bottom of Table 9.1. In between the two extreme divisions (low and high altitude; low and high rainfall), there are two intermediate divisions. The cut off points are designed to roughly reflect the boundaries of zones that on the whole dictate the dominant mode of livelihood of the peoples encompassed by the divisions. Thus the very low altitude and rainfall divisions correspond to desert and semi desert conditions where small livestock, camels and some cattle form the main mode of subsistence. The lower intermediate rainfall and altitude divisions reflect conditions of arid savannah grassland where cattle predominate, and marginal agriculture takes place in some areas. The higher divisions encompass those peoples with prime agricultural land, with the highest population density, and include in the highest altitude zone those peoples which occupy or overlap into the less favoured high forest and grassland areas (see Chapter 1).

I would stress, however, that these ecological zones have only influenced the drawing up of demarcations between the altitude and rainfall divisions. It would be a mistake to interpret altitude and rainfall divisions as reflecting ecological zones in their own right, as the correlation between rainfall and altitude is far from being unity. In Table 9.1, for instance, the Luyia on the whole occupy the higher intermediate altitude zones, and the highest rainfall zone; and the Mji Kenda

though occupying the lowest altitude zone, occupy the higher intermediate rainfall zone.

In spite of these factors influencing the demarcation of altitude and rainfall zones, moreover, there is undoubtedly an arbitrary element in the demarcation of the zones. Such arbitrariness, however, is preferable to a policy of re-arranging the divisions to find an optimum relationship between dermatoglyphic variation and altitude and rainfall, as such a policy is open to subjective distortion. On the positive side I would submit that the divisions as allocated here represent a good starting point. As each division encompasses a broad range of populations, we would expect little differences between the zones when population means and frequencies are pooled if no relationship exists between dermatoglyphic variation and altitude and rainfall.

## 9.2 Analytical approach

The analysis of possible relationships between dermatoglyphic variation and altitude and rainfall in this section falls into two parts. In the first the extent of the relationships is explored through correlating taxonomic dermatoglyphic attributes with "categorised" rainfall and altitude variables. In the second, analysis of variance is applied to see whether variation of dermatoglyphic means or frequencies within altitude and rainfall zones differs significantly from the variation between zones. For those attributes that do show significance, the distribution of zone average means and frequencies will be presented and discussed. In effect the first part of the analysis measures strength of relationships, and the second the nature of relationships.

9.2.1 The extent of the relationships between dermatoglyphic taxonomic variation, and altitude and rainfall

The taxonomic attributes included in the analysis are the sample means and frequencies of all non rare dermatoglyphic attributes, both summary and non summary as already outlined in Chapters 5 and 6. Each taxonomic attribute is correlated with altitude and rainfall variables, consisting of the sample zone values for altitude and rainfall as enumerated in Table 9.1.

All correlations involving altitude are normal Pearson's correlations. In the case of rainfall, however, the correlations are partial ones, controlling for the effect of altitude. The employment of partial correlations is to avoid replication in view of the significant correlation between altitude and rainfall in Kenya (.5 in the case of the present study). Moreover the reason why partial correlations only apply in the case of rainfall is the fact that rainfall is affected by altitude, while altitude is not affected by rainfall. It would be nonsense, therefore, to produce partial correlations of dermatoglyphic attributes with altitude, controlling for rainfall. Finally all computations were performed using SPSS correlation procedures (Nie et.al, 1975, Chapters 18 and 19).

9.2.2. Correlations of dermatoglyphic summary taxonomic attributes with altitude and rainfall

The consideration of the summary attributes is to establish the nature of overall trends, in order to make it easier to interpret the more complex results of the non summary attributes.

The results are presented in Table 9.2. In order to save

space and to achieve a clear picture, only significant correlations are listed, and they are given in rank order within digital and palmar variable groupings. The main trends are as follows:

1) The strongest relationships tend to be with rainfall rather than with altitude. More correlations are significant with rainfall (10 to 8 in males and 9 to 5 in females); and correlations tend to be higher with rainfall (in males 7 out of 10 significant correlations exceed .3 with rainfall, compared to 3 out of 8 with altitude; and in females the equivalent figures are 5 out of 9 compared to 1 out of 5).

2) If we differentiate the attributes into digital and palmar groups, the digital attributes follow the main trend, that more correlations with rainfall are significant, and they tend to be higher than those with altitude. The palmar attributes however, show stronger relationships with altitude, especially in the females where three palmar correlations are significant with altitude, compared to none with rainfall.

3) On the whole the digital attributes that show significance with altitude tend to be "ulnar" ones; particularly right hand ulnar loops and total ulnar ridge count. Correlations with rainfall however, tend to be high on "radial" attributes, (total radial and unilateral ridge counts), and concentric whorls which directly affect radial ridge counts. The highest correlations moreover, are on the left hand in both sexes.

In the case of palmar attributes it is interdigital and total palmar intensity that tend to show the highest correlations with altitude in both sexes and with rainfall in males only.

4) Correlations both with altitude and rainfall are mostly negative, showing that the lower lying and more arid samples tend to have the higher means and frequencies. The only positive correlations are ulnar loops and RTPRC with altitude, and LTPLMI, TDR and TDL with rainfall. It is interesting that interdigital and total palmar triradial intensities correlate negatively with altitude, and positively with rainfall.

5) Trends between the sexes show some fundamental similarity. For instance both RTULNRC and RULOOP are significant for altitude in both sexes, right hand attributes predominate in both sexes for altitude, and left hand correlations predominate for rainfall in males, and left hand attributes are most significant for rainfall in females; "radial" attributes correlate significantly with rainfall in both sexes. Nevertheless there are also considerable differences between the sexes. For example LDLOOP correlates highest with altitude in males, but is not significant in females; correlations with both rainfall and altitude are higher for males, and even where the same attributes correlate significantly in both sexes, the rankings are not concordant.

### 9.2.3 Correlations of dermatoglyphic non summary attributes with altitude and rainfall

The results are presented in Table 9.3 for altitude and Table 9.4 for rainfall. The main trends are as follows:

1) The males follow the main trend exhibited by correlations with summary attributes; that correlations with rainfall are more numerous and tend to be higher. A comparison of the male correlations in tables 9.3 and 9.4 shows that 33 attributes correlated

significantly with rainfall, compared to 26 with altitude, and that 10 of these exceeded .4 with rainfall (maximum -.48) compared to only 4 with altitude (maximum -.43).

The females, however, manifest the opposite trend. It is significant correlations with altitude that are the most numerous and highest. No less than 42 attributes correlate significantly with altitude, compared to 29 with rainfall, and 12 of the correlations with altitude are .4 or greater, (maximum -.5) compared to 2 with rainfall (maximum -.41).

2) Segregating digital and palmar attributes reveals even more marked sex differences. The males show 14 significant digital correlations with altitude compared to 39 in females, and 12 palmar attributes correlate significantly with altitude in males compared to only 3 in females.

Correlations with rainfall exhibit a much less noticable disparity between the sexes. The males show 23 digital and 10 palmar significant correlations, compared to 20 digital and 9 palmar ones in females.

3) Detailed digital trends for altitude show very clearly in the males that it is mainly attributes of the "ulnar" type that correlate highly. The 7 highest correlations are all with digit 1 involving mostly ulnar ridge counts and ulnar loops. In the lower rankings WHORLR1 and DLOOPR1 also figure. The remaining significant correlationsare with attributes on digit 4 except for RFR5.

The females show much more complex trends. Of the 39 significant correlations with altitude 10 are attributes on digit 1, 9 each on digits 2, 3, 4, but only 2 on digit 5. Only digit 5,

therefore is markedly unrepresented. If we take, however, only the 12 most significant attributes (.4 or greater), the most prominant are those attributes on digit 1 (5 out of the 12), consisting as in the males, mainly of ulnar ridge counts and loops. The highest and most significant correlations, however, are LFR4 and RFR4 which correlate much lower in males. The remaining 5 attributes are on digit 3 and 2.

In summary altitude appears to affect mainly "ulnar" attributes on digit 1 and radial ridge counts on digit 4. The disparity between the sexes is due to the fact that altitude affects many attributes on digits 2 and 3 in the females, but not in the males.

4) Detailed trends of correlation of digital attributes with rainfall are complex in both sexes. In males 3 attributes are significant on digit 1; 6 on digit 2; 3 on digit 3; 5 on digit 4 and 6 on digit 5. In females 7 attributes are significant on digit 1, 5 on digit 2; 1 on digit 3; 5 on digit 4; and 2 on digit 5. Thus it is attributes on digit 3 in both sexes, digit 1 in males and digit 5 in females that are least affected by rainfall.

Attributes on digits 1 and 2 that are significant tend to be radial ridge counts, concentric whorls and double loops in both sexes, while those loading on digits 4 and 5 are predominantly "ulnar" attributes (ulnar ridge counts, ulnar loops). This contrasts with the correlations with altitude where attributes on digits 1 and 2 tended to be "ulnar" in type, while those on digit 4 tended to be "radial" in type.

5) In the case of palmar attributes, the fact that only 3 correlate significantly with altitude in females compared to 12 in males makes interpretation difficult. Only two attributes correlate significantly in both sexes, TDR6 and P4R. In the males all significant correlations are with interdigital attributes and palmar ridge counts, except for one hypothenar attribute (CHL). The significant palmar ridge counts (RAB, LAB, RCD),moreover, differ from the others in that they correlate positively.

Correlations of palmar attributes with rainfall tend to be higher than those with altitude. As in the case of altitude, most attributes correlating significantly with rainfall tend to be interdigital or palmar ridge counts. There are, however, some major differences. It is the bc rather than the ab ridge counts that correlate significantly with rainfall. Some interdigital attributes, furthermore, that correlate negatively with altitude (i.e. P2R, P2L, P4R), correlate positively with rainfall. Finally TEFL correlates significantly in both sexes with rainfall, unlike with altitude where all correlations with thenar traits are very low.

## 9.3 The nature of the relationship of the variation of dermatoglyphic taxonomic attributes, with altitude and rainfall

The investigation so far has shown that a notable number of dermatoglyphic taxonomic attributes correlate significantly in the 2.3 to 5.0 range with altitude and/or rainfall. Nineteen out of the 26 summary attributes correlate significantly at least once with either; (a) altitude males; (b) altitude females;

(c) rainfall males and (d) rainfall females. In the case of non summary attributes the equivalent ratio is 77 significant attributes out of the 112 attributes considered.

I now turn to examining the nature of the apparent relationships between dermatoglyphic variation and altitude and rainfall for those attributes that correlated significantly in section 9.3. The method adopted to achieve this is to examine the distribution of altitude or rainfall zone means (i.e. the zone mean of the population sample dermatoglyphic means or frequencies) through one way analysis of variance. In sub section 9.3.1 the analysis concentrates on summary attributes, and in sub-section 9.3.2 on non-summary attributes.

### 9.3.1 The nature of the relationship of summary dermatoglyphic taxonomic attributes, with altitude and rainfall

A) The nature of relationships with altitude

The results of analysis of variance are given in Table 9.5 for both sexes. Only those attributes that show significant correlations with altitude (Table 9.2) are reported, and the attributes are ranked as in Table 9.2 in decreasing order of relationship.

The results for digital attributes in both sexes show an overwhelming trend of a linear clinal relationship with altitude. In the case of RTULNRC and RTABSRC in males, and RTULNRC in females there is a progressive fall in mean as altitude increases, and for RULOOP in each sex (an attribute negatively related to ridge counts and whorls both phenetically and taxonomically), there is a sharp increase in mean as altitude increases. In females the

increase in mean is progressive and linear. In males, however, the first two low altitude zones are almost identical in mean, showing a sharp increase to the two high altitude zones which are also similar.

In the case of LDLOOP in males, an attribute that shows the strongest correlation with altitude, there is a marked drop in mean from the lowest to the highest altitude zone. The two medium zones, however, though exhibiting mean values that fall between those of the extreme zones, are relatively similar, and the means are a reverse of the dominant trend.

Palmar attributes in contrast do not show such clear clinal trends. TDR for instance, which in both sexes displays the strongest correlation with altitude, shows an alternating decrease and increase in mean from low to high zones, with the overall trend being a decrease in value as altitude increases. RTPLMI (males), TDL and LTPLMI (females) also show a similar pattern of variation to TDR. In males RTPRC exhibits a linear increase in mean as altitude increases, up to the third/fourth zones which are almost identical. LTOTHI in males is the only true linear clinal attribute, with means progressively decreasing as altitude increases.

B) Nature of relationships with rainfall

The results of analysis of variance are presented in Table 9.6, in the same format as those for altitude.

As in the preceding account concerning altitude, there are clinar trends manifested for the attributes which correlate significantly with rainfall. In males all digital attributes show

a tendency for means to decrease as rainfall increases, and all three palmar attributes demonstrate an increase in mean as rainfall increases. The trend for these attributes, however, is not one of regular increase or decrease from zone to zone. The cline rather resembles a Z, with the first two low rainfall zones homogeneous, and sharply distinguished from the also homogeneous two high rainfall zones.

In the case of females, however, there is more variety of trends. Although the Z trend is also in evidence, it is much weaker than that in males. The Z trend is strongest for LTRC, RTRC, RTFPI and RCWHORL. Nevertheless even for these attributes there is a clinar decrease in mean in the lowest two rainfall zones, before the main decrease occurs from the 2nd to 3rd zone. It is the two highest rainfall zones that are most homogeneous. In the case of LTRADRC, LTABSRC, LCWHORL and RTRADRC, the Z trend is very moderate, and there is a regular clinar decrease demonstrated.

9.3.2 The nature of the relationship of the variation of non summary dermatoglyphic taxonomic attributes, with altitude and rainfall

A) Nature of relationships with altitude

The results of analysis of variance are presented in Table 9.7 for males and Table 9.8 for females, with attributes ordered according to the correlation rankings listed in Table 9.3.

Of the 4 male and 12 female digital attributes that correlate at .4 or higher with altitude, the differences in zone means are all highly significant. Of these attributes, moreover,

RFU1, LFU1 and LD1 in males, and LFR4, RFR4, LFU1, RFU2 and RFU1 in females show a strong linear, clinal decrease in mean as altitude increases. The biggest decrease tends to be from the 3rd to the 4th and highest altitude zone. In addition ULOOPR3 and ULOOPR1 in females exhibit a strong linear increase in mean as altitude increases.

The attributes not yet mentioned that correlate higher than .4 show a less distinct clinal pattern of variation. Thus ULOOPR1 in males shows little difference between the first two altitude zones, but there is a sharp increase in mean in the two highest zones. A similar pattern is also demonstrated by LFR3 and RFU3 in females, only that there is a sharp decrease in mean in the two highest altitude zones. A different pattern is in evidence for ULOOPR2 and ULOOPL1 (females) where there is a sharp increase in mean from the lowest to the two medium zones, which differ little in mean, followed by a sharp increase again in the highest zone. Finally RD1 in females shows an increase in mean from the 1st to 2nd zones, followed by a strong decrease in the last two highest zones.

In the case of those digital attributes that correlate significantly at less than .4, only two in the males (RD1 and DLOOPL1) show significant differences between zone means. In the females, however, eleven attributes show significant differences (ULOOPL2, RD4, WHORLR1, RFU4, LFR5, RADLL1, WHORLL1, DLOOPR1, ULOOPR4, RADLR4 and WHORLR4). All the digital attributes correlating less than .4, whether significant or not, still show some evidence of the general trend of decreasing means as altitude increases for ridge counts, finger pattern intensity, concentric

whorls and double loops; and of increasing means for ulnar loops. In only one of these attributes, however, (LFU2 females), is the cline of means distinctly linear for all four zones. In the rest the trend is indistinct, with two or sometimes three zones showing little difference in means. For example in the case of DLOOPR2 females, there is a strong decrease in mean from the first to second zones, and relative homogeneity thereafter, or in the case of DLOOPR4 there is a Z type cline, with the lowest two zones showing homogeneous high means, followed by a decrease to the highest two zones which are also homogeneous.

Trends of variation for palmar attributes are more complex and less systematic than those for digital attributes. Dealing first with the males, which show by far the greatest number of significant correlations with palmar attributes (see Table 9.3) 6 out of 12 show significant differences between zone means, (TDR4, TDR6, TDL6, P2L, P2R and TDL4). The interdigital attributes TDR4 and TDL4 show a marked increase in mean from the 1st zone to the 2nd, followed by a decrease in the 3rd and an increase in the 4th. The reverse trend is apparent fro TDR6, TDL6, and TDR5 (i.e. decrease, increase, decrease). This alternating trend is also in evidence for P2R and P2L. Interdigital loop P4R and hypothenar loop CH1, however show a linear decrease in mean as altitude increases, with the 2 highest zones being relatively homogeneous. Finally the palmar ridge counts show a rough trend of increasing means as altitude increases. This is most evident for RCD where there is a rapid increase up to the relatively homogeneous highest two zones. In the case of RAB and LAB,

however, the notable increase in mean occurs from the 1st to the 2nd zone, with the higher 3 zones homogeneous in mean.

In the females the only 3 palmar attributes that correlate significantly with altitude (Table 9.3), TDR6 and TDR3, do not show significant differences in mean. The trends, however, are more systematic than in males, with a decrease in mean as altitude increases, with the middle two zone homogeneous, for TDR6 and P4R, and a similar increase in mean as altitude increases for TDR3.

B) The nature of relationships with rainfall

The results of analysis of variance are given in Table 9.9 for males and Table 9.10 for females. They show that 10 digital attributes in males (RADLL2, RFR2, RFU5, WHORLR5, LFR1, LFU5, LFR2, WHORLL5, ARCHL1, and ULOOPL5) and 8 in females (RFR2, LFR1, WHORLR4, WHORLL1, RFU4, ULOOPR4, LFR5, and RD4) have zone means that differ significantly.

Unlike relationships with altitude, where the attributes with the highest correlations were the ones which almost exclusively displayed strong linear clinal trends, relationships with rainfall are linear and clinal only for WHORLR5 in males and RFR2 in females for digital attributes that correlate at .4 or above.

In the males the strong linear clinal digital attributes are associated with ridge counts, ulnar loops and concentric whorls on digits 4 and 5, on the right hand irrespective of strength of relationship. Thus WHORLL 5, WHORLR4, RFU4, RD5 and RD4 all show a marked decrease in mean as rainfall increases.

The remaining attributes show weaker clines, positive in the case of RADLL2, RADLR2, ARCHL1, ULOOPL5, ULOOPL3; and negative for RFR2, RFU5, LFR1, LFU5, LFR2, WHORLL5, WHORLL3, WHORLL2, WHORLR2, WHORLL4 and WHORLL1. In each case the cline is weakened sometimes by irregularity in one zone (e.g. WHORLL5 where zone 2 has a higher mean that zone 1) or commonly by the fact that two contiguous zones are homogeneous in mean (e.g. ULOOPL3)

In females linear clines are only in evidence for RFR2, WHORLR4, ULOOPL4, RADLL2 and LFU5. Although this list stresses digits 4 and 5, unlike in the males, it is the left rather than the right hand that shows the strongest trends. As in the case of males, however, all remaining attributes except for ULOOPR2, show some evidence of a weaker cline with rainfall: negative for LFR1, WHORLL1, RFR1, DLOOPR1, DLOOPL1, RFU4, LFR5, RD4, LFR2, WHORLL3; and positive for RADLL1, ULOOPR4, DLOOPR2, ULOOPL4, ARCHR1, and RADLL2.

Palmar attributes in males show some consistent clines, though only for CHR is the trend linear (negative). RBC, TDR4, TDL4, and TEFL show a decrease in mean as rainfall increases, with the 3rd and 4th zones homogeneous, while P2R shows the same trend in the positive direction. TDL5 also shows a positive clinal trend, only it is the 2 medium zones that are homogeneous. P2L and TDR6 demonstrate a positive cline up to the 3rd zone, followed by a decrease in the 4th. Finally LBC is the weakest trend, differentiating effectively only zone 2, which has a much lower mean than the other zones.

Palmar attributes in females show that only TDL4 exhibits a strong clinal linear trend (negative) with rainfall. TDL5, TEFL and P4L show weaker clines, affected by homogeneity between contigous zones and in the case of TEFL, by a decrease in the 4th zone running contrary to the cline. LBC, RBC, P4R and CHL are very weak clines, showing a marked increase from zone 1 to the remaining 3 zones that are relatively homogeneous. Finally P3TR only differentiates zone 2 effectively from the other three.

In summary the analysis has shown that a notable number of dermatoglyphic taxonomic attributes correlate significantly with altitude and/or rainfall. Nineteen out of 26 summary attributes correlate significantly at least once with rainfall, altitude in the male or female series, and in the case of non summary attributes the equivalent ratio is 77 out of 112.

## 9.4 Summary and discussion

The analyses in sections 9.3 and 9.4 have demonstrated that there appears to be a moderate, but significant relationship between many dermatoglyphic taxonomic attributes, and altitude and rainfall. The nature of the relationships, moreover, are clinal in nature for a surprising number of summary and non summary attributes. The strongest relationships and clines tend to be for some digital attributes, particularly those on digit 1 and digit 4. On the whole digital ridge counts, concentric whorls and double loops tend to show a decrease in mean as altitude or rainfall increases, and ulnar loops an increase in mean. The least systematic relationships are with arches (only one attribute, ARCHL1 shows any significant relationship - see Table 9.4); and

with radial loops and double loops, which show less consistent clines when they are significant.

Palmar attributes also show significant correlations with altitude and rainfall, the strongest correlations being with rainfall non summary attributes. The nature of the variation is often clinal too. The strongest clines are with rainfall (see Tables 9.9 and 9.10). Generally it is measures of interdigital triradial intensity, second and fourth interdigital peripheral loops, thenar triradii and hypothenar central loops which show the strongest raltionships. Palmar ridge counts also figure as significant (bc with rainfall and ab with altitude). The higher zones, however, tend to be relatively homogeneous, sharply differentiated from the lowest and most arid zones which have markedly the lowest mean.

A theme familiar all through this investigation has also consistently manifested itself here. Patterns of variation by sex, though consistent in major trends, differ appreciably in detail. This is particularly in evidence in the number of non summary attributes that show significance with altitude, where far more digital attributes are significant in females than in males, whilst only 3 female palmar attributes are significant with altitude compared to 12 in males. One possible explanation for this apparent discordance is the fact that the males include four populations that have no female samples. Excluding these extra male samples, however, did not succeed in raising inter-sex concordance to any appreciable degree in an analysis (unreported here) to clarify this point.

In the analyses of variance, there are many attributes that appear as non significant. The reader, therefore, might be forgiven for deducing that the results are tenuous. It should be noted, however, that the analyses of variance here are all based on taxonomic attributes with population samples forming the cases. In reality the zone sample sizes (see Table 9.1) in terms of individuals are large. In a further stage of analysis (unfortunately not complete), in which individuals rather than populations were pooled according to altitude and rainfall zones, all but a few digital attributes were shown to differ highly significantly in zone means. The results here, therefore, should be rather regarded as being on the side of caution.

Holt (1975), discussing the possibilities of natural selection affecting total ridge count, states that:

> "There is no direct evidence of any action of natural selection, although it is suggested by the observed racial variation."
>
> (Holt 1975[b]:176)

If the results reported in this chapter really do reflect true trends of variation in Kenya, then they can be regarded as the first indications that some form of climatic or environmental selection is operating on dermatoglyphic attributes of the peoples of that country, and opens up the possibility that similar trends might be found in other parts of the world.

The results, however, do not give any real clue as to the nature of the selection operating. We could speculate, nonetheless, that altitude and rainfall are not likely to be, in themselves, direct selective agents. It is more plausible that

the association with these variables is due to the fact that altitude and rainfall are so prominant in determining the nature of ecology and environment in Kenya, which in turn determines the distribution of the real selective agent (some sort of disease?). If this should be true, then it might explain why Hiernaux (1968a) failed to find any significant correlations between digital patterns and climate at a continental scale of investigation, as the distribution of the true selective agent or agents might be determined by other factors in other parts of the African continent.

# CONCLUSION

The object of fieldwork in 1973 and 1975 was to collect dermatoglyphic data from a range of ethnic populations in Kenya, with the ultimate aim of using the findings from the analysis of this data to further the cause of reconstructing the pre-colonial history of the peoples of Kenya. In all 6,235 school children of both sexes from 61 ethnic populations were sampled for hand and finger prints. The present thesis has been concerned with the first stage of the overall project, that is reporting the findings of the survey, describing the extent and nature of dermatoglyphic variation of the sampled populations, and seeking to account for this variation as far as the available evidence permits.

In this conclusion I shall concentrate on problems and issues encountered during the study, and the extent to which I have succeeded in resolving them. Major problems are associated with four aspects of the study. These are: (1) the nature of populations in Kenya and sampling; (2) the methodology of dermatoglyphic attribute measurement and of minimising errors; (3) the analysis and description of the variation; and (4) explanation arising from this description. I shall proceed, therefore, by discussing these sets of problems in turn.

## C1: Problems and issues associated with the nature of populations in Kenya, and sampling.

The issues and problems in this area were discussed in detail in Chapter 2. When I embarked on fieldwork in 1973, the common practice of physical anthropologists in the field in

Africa was to obtain a single sample from a self-defined ethnic unit (e.g. a sample from the Dogon, Kikuyu, Dinka, Ibo, etc.). The justification for this method of sampling is outlined by Hiernaux (1972:97; 1974:50) who argues that sub-Saharan African ethnic units form natural biological populations owing to (a) the high rate of endogamy they display, and (b) their biological homogeneity, which we can assume, unless there are contrary indications (Hiernaux 1972:97).

My Kenyan experience, however, made me doubt the contention that African ethnic populations are "natural biological units". In Chapter 2, I presented a detailed critique of this traditional view of African populations. In particular I questioned the following points: (a) that culturally maintained endogamy is a central characteristic of African ethnic units; (b) that we should assume biological homogeneity for such units; and (c) that African ethnic populations were clearly demarcated and cohesive social and demographic groups in pre-colonial times.

An important fact to emerge from the discussion in Chapter 2 is that traditional African social systems were based on semi-autonomous primary groups (Benedict 1966), which were loosely linked by common identity and purpose to form wider social units. The total field of common identity could vary from very small in the case of hunting peoples, to relatively large in the case of centralised kingdoms. The pre-colonial situation, therefore, was one of aggregations of flexible primary groups, which were liable to rapid dispersion and re-integration in time of disaster such as war, drought and famine, rather than one of stable cohesion.

Paradoxically, the imposition of colonial rule has resulted in the stabilisation of African societies, of widening scales of identity, and of increasing cohesiveness and common purpose amongst peoples which were previously autonomous.

In addition, there has been considerable movement of people from traditional areas to urban centres and areas of large scale farming. In such areas, ethnic identity has been instrumental in providing a framework for group formation amongst migrants. The types of groups and societies that have arisen in such areas bear only superficial resemblance to pre-colonial social structures. Moreover, colonial times saw the spread of Christianity, the introduction of Western-type education, increasing occupational specialisation and the birth of incipient social classes, trends which have accelerated in the post-colonial era. Such factors are significantly influencing individual behaviour, especially choice of marriage partner in "progressive" urban and rural localities, and cross-cut traditional ethnic identity in determining group formation and identity.

The African situation, therefore, is highly complex, as complex as that of European and other large scale western and eastern societies. There is clearly a great deal of research to be done before physical anthropologists in Africa can know to any degree of certainty (a) the nature of the peoples they are sampling and (b) what their samples represent in detail.

The studies of Gomila (1972), Salzano (1964; 1972) and Chagnon (1972), have done much to reveal the nature of small-scale populations in the tropics. These studies, however, are micro

studies, concentrating on villages and local demographic description. In contrast, medium resolution demographic studies in tropical regions by physical anthropologists have been very much neglected, and present a challenge for the future. Of particular interest in Africa are (a) the nature of apparently traditional rural populations where change has been slow in recent decades; (b) the nature of populations in traditional areas where change has been substantial; (c) the nature of rural populations in non-traditional areas where migrants have had the opportunity to form new alignments and groups; and (d) the nature of populations in urban centres of different sizes. Also of importance are the patterns of movement of people between these four types of areas, and factors determining choice of marriage partner in societies characteristic of each areas.

Any sampling design in Africa must take these complexities into account, as far as is practical. Unfortunately, in the absence of detailed demographic information, it is seldom possible to adopt a sophisticated sampling scheme, especially if the range of populations being sampled is large. One important point that should be considered carefully, however, is how to sample populations at a scale of definition compatible with the range of the survey. As I pointed out in Chapter 2, African populations can be defined on a number of scales, according to ethnic or regional identity. Thus a man could, for instance, according to the definition, be an Aulihan, an Ogaden, Somali; or an inhabitant of Garba Tula, Isiolo District, Eastern Province, Kenya, East Africa, Africa. It is ethnic definition, however, which has posed problems in the past. It is not uncommon for large scale ethnic categories such

as "the Somali", to be sampled as a unit alongside small scale minimal ethnic units as "The El Molo".

In the present study, the range of the survey was a single African country, Kenya, and as such it would have been simplistic to adopt ethnic groups such as "Luyia", "Meru" or "Kalenjin" as the primary units to sample. I argued in Chapter 2 that the natural unit to sample is what I termed "the minimal ethnic population unit", being the maximal unit of common identity in the pre-colonial situation, before scales of identity broadened. The "minimal ethnic population unit" corresponds to what oral historians (e.g. Were (1967), Ogot (1967) call a "sub-tribe", or to what Evans-Pritchard (1940) defines technically as a "tribe". At this scale of definition, over 140 such units can be differentiated in Kenya.

The difficulty of conforming in a standard fashion to this scale of definition is well illustrated by this study. In spite of an awareness of the need to standardise, I still failed to obtain samples of Luo and Somali minimal ethnic units, and was forced to treat general samples of these peoples as corresponding to minimal units in opposition to the bulk of the remaining samples which were at the minimal scale of definition. Strictly speaking, this is bad practice. There is a clear need for physical anthropologists to be fully aware of the problem of sampling ethnic units which are not of a standard compatible scale of definition. Ultimately, standardisation can only be fully achieved when full knowledge is available of the identifiable sub-divisions of ethnic peoples as defined today.

In assessing the usefulness of the sampling approach I adopted, it is important to take into account the fact that sampling was geared to the requirements of the historical objectives which underlay the survey. The crucial difference between the historical sampling requirements and those of a more conventional human biological survey, is that in the historical approach it is essential to obtain samples in areas where the nucleus population has not moved residentially to any significant extent since the advent of colonisation, whereas in a normal human biological survey it would have been better to have sampled geographical points in a systematic manner irrespective of historial patterns of residential stability. Thus in the present study, the historical requirements dictated that the best sampling localities lay in rural areas which were removed from the main centres of recent immigration, the urban and ex-white farming areas. A non-historical human biological survey, however, should have sampled these non traditional areas as a matter of course, to obtain a comprehensive picture of regional variation.

The nature of populations from urban and non traditional rural areas is such a complex issue that the common practice of sampling ethnic units from such areas, as a "short cut" to visiting the traditional areas of origin, should be discouraged. Researchers should rather treat such samples as typical only of the non-traditional areas they were drawn from. The dangers of not fully complying with this recommendation are well illustrated in this study. The representativeness of the Luo who were mostly sampled in Nairobi, and the Turkana who were sampled in Kitale, a white farming area, has been a problem throughout the analysis.

Another important issue concerning sampling is the problem of getting the right balance between sample size and numbers of traits to be measured. Before this study was undertaken, the indications seemed to be that meaningful and significant variation could be detected even with samples of well under 30 of each sex, judging from the south Midlands dermatoglyphic survey of Roberts and Coope (1972). Since then, however, Dennis (1977a), considered that sample sizes of under 150 of each sex were inadequate (Dennis 1977a:436). Even allowing for the fact that the Yorkshire Dale populations are largely very homogeneous dermatoglyphically, thus requiring larger samples to detect any small differences, the adequacy of sample sizes of less than 50 of each sex must be considered suspect, except for comparison of continental populations where differences are large.

At the time of fieldwork I considered,having been influenced by the south Midlands study, that a sample size of 25 of each sex was likely to be adequate, and that 50 of each sex, the norm I adopted in 1973, was more than adequate. Faced with the choice of cutting down for practical reasons on one of three areas, (a) number of populations to be sampled, (b) sample size, and (c) number of traits to be measured, I decided to cut down on sample size. In retrospect it would have been better to have doubled sample size and to have cut downon the number of attributes measured. The reasons for increasing sample size are illustrated in Table 8.6, where it is apparent that $\Delta g$ distances do not stabilise until a sample size of 50 is reached. The justifications for cutting down on traits measured are twofold. Firstly, there

is the length of time needed to read all traits from a set of hand and finger prints. In this study it took over three years, even after mainlines had been excluded. Secondly, there is the fact that some traits are redundant or partially redundant, as the correlations in Chapter 5 show. With careful thought it should be possible for future researchers engaged in large dermatoglyphic regional studies, to substantially cut down the number of traits measured, without significantly losing major areas of variation, thus permitting more time to increase sample size.

C2: Problems and issues associated with the methodology of dermatoglyphic attribute measurement and of minimising errors.

These problems and issues were discussed in detail in Chapter 4. The main problem encountered before reading the prints was whether to adopt the traditional topographical methodology of Cummins and Midlo (1943) for palmar and digital patterns and triradii, or whether to employ the recent topological methodology of Penrose and Loesch (1970). The topological methodology is theoretically sounder in that through topological principles we can predict extralimital patterns and triradii, which are easily missed in the topographical methodology. This predictive property is especially relevant for triradius $t^b$ on the hypothenar area, which is often not printed when it lies at the extreme ulnar periphery.

Because of these advantages it became the policy of dermatoglyphic researchers in the Durham Anthropology Department to adopt the topological methodology. It became clear, however,

that the methodology possessed several inadequacies. Particularly serious was (a) the claim of Penrose and Loesch (1970) that all mainlines except for the A mainline were redundant in the topological methodology; (b) the imprecise definition of some palmar triradii (i.e. e/f; t,t',t''; $t^{u}$); and (c) the imprecise status of zygodactylous triradii, and the failure to distinguish between missing and zygodactylous triradii. Amendments were made which are fully described in Chapter 4, and in Dennis (1977a).

Digital patterns can also be treated topologically, by measuring the mean occurrence of triradii on each finger, and by calculating ulnarity/radiality indices (see Dennis 1977a). Dennis and I could see no reason, however, why the topological methodology should supplant the traditional classification of digital patterns. Firstly, most dermatoglyphic studies have reported frequencies of traditional digital patterns, and to not to report them would have rendered results useless for collation with those of other studies. Secondly, we felt that the topographical distinction between certain pattern types might be genetically important, and consequently not to report detailed pattern types might result in a serious loss of variation. In this study, therefore, I employed both a topographical and a limited topological treatment of digital patterns. This approach has been vindicated in Chapter 5, where an important aspect of the variation found is that patterns of variation for concentric whorls do not follow closely those for double loops.

The rules for defining topographical digital patterns, however, were not free from ambiguities. There is difficulty, for instance, deciding in marginal patterns whether a whorl is a concentric whorl or a double loop, or exactly when a central pocket

loop becomes a concentric whorl. The methodology was made more precise as outlined in Chapter 4.

The detailed rules regarding ridge counts were standard, as outlined by Cummins and Midlo (1943) and Penrose (1968). Considerable problems, however, were encountered with the precise definition of axial triradii. In spite of the more precise rules given by Penrose (1968), the distinction between t and t' is still unsatisfactory, as the main reference points, the distal wrist crease and proximal finger crease on digit 4, are imprecise, and are liable to subjective interpretation. The value of distinguishing axial triradii in this study, therefore, is questionable.

The minimal atd angle, with its variation with age, is also an unsatisfactory measure, and has been severely criticised by Brehme (1970). Although I have reported the summary results for the minimal atd angle in this study, my decision not to use it in the analysis is justified. The use of the atd angle in future dermatoglyphic population studies is clearly inadvisable.

The methods and detailed rules of dermatoglyphic methodology used in this study are consistent, and relatively free of serious ambiguities. Unfortunately the achievement of consistency has been at the expense of easy comparison of many of the present results with those of surveys outside the Durham Department. Firstly, the use of the topological methodology makes comparison difficult with these studies that have employed the Cummins and Midlo methodology. Secondly, even within the topological methodology, the amendments in the case of palmar triradii and zygodactylous triradii, mean that the results for these

features here will not be compatible with those of researchers who have strictly followed the Penrose and Loesch recommendations (Penrose and Loesch 1970). Thirdly, the unambiguous rules outlined here for detailed digital patterns (concentric whorls, double loops and central pocket loops), also partially hinder comparability with other studies. A majority of the most important traits, however, particularly ridge counts, total digital patterns, have no amendments which will adversly affect comparability.

In most dermatoglyphic reports, detailed procedures regarding the chores of actually reading the prints, are not usually described, in spite of the fact that such practical information could be valuable to novices. Of the lengthy details outlined in Chapter 4, I would stress the importance of not reading all attributes from each set of prints in one go. The practice of reading limited groups of attributes first throughout the sample, before tackling the next group of attributes, acts as a natural cross-check, and eliminates many errors.

Of much greater importance, however, is the problem of what to do when faced with partially obscured or indistinct features, especially indistinct finger prints when ridge counting. I would emphasise that the apparently correct procedure, of treating finger ridge counts as missing when a totally accurate count cannot be made, is liable to serious bias. This is because patterns with large ridge counts are much more likely to be partially obscured than those with small ridge counts. Ridge counting only very clear prints can result in a serious underestimation of ridge count means. The same type of bias also applies to palmar ridge counts.

Unless ridge counts are impossibly obscured, therefore, it is preferable to estimate them.

The last stage of the preliminary processing of the data was to set up the data matrices for computer analysis. After the prints had been read, there were two further sources of errors. The first was in transcribing the data from the prints to coding sheets, and the second was in punching the data from coding sheets onto computer cards. The overall rate of detected errors, described at the end of Chapter 4, was significant, emphasising the importance of rigorous data cleaning. It is interesting that all but a small proportion of errors arose from copying data onto coding sheets. The card punchers were almost perfect. All but a very few of "punching errors" were accountable by the fact that numerals on the coding sheet had been written ambiguously.

C3: Problems and issues of analysing and describing the variation.

Any comprehensive dermatoglyphic survey involves the consideration of a large number of attributes, well over two hundred if we count common digital pattern types on each digit as individual attributes. Compounding this wealth of attributes is the methodological necessity of treating the sexes as independent series, owing to the marked sex differences in a majority of attributes. Any study of dermatoglyphic variation, therefore, must face the problem of efficiently describing and presenting summary results for a large number of attribute, sex and population permutations.

In the present study, this problem was acute in that (a) the number of individuals in the total sample is larger than

usual (6,200+), and (b) the number of population samples exceeds 50. Moreover, it is not sufficient to present summary statistical results for population samples at the smallest scale of definition. To be compatible with other African studies, e.g. those of Hiernaux (1968, 1972), it is also desirable to examine the type of variation manifested by larger scale ethnic populations. Each different scale of population definition necessitates a different phase of description, with results doubled for each sex.

Owing to the multitude of basic descriptive statistical results that arises from the examination of these attribute/sex/population permutations, I felt that it was impractical to report in this study the detailed written univariate description that usually forms a major part of reports on dermatoglyphic regional variation (e.g. Williams 1978; Smith 1979). Moreover, I had fundamental doubts concerning the usefulness of detailed univariate and multivariate descriptive approaches in the first part of Chapter 3. I argued that the best way of coping with descriptive results is to focus the aims of a study in such a way that the results are directed towards some positive purpose, without the aims being so rigid as to stifle the role of open discovery that is the essence of survey analysis.

In the major part of Chapter 3, I proposed an approach in which description is orientated towards the investigation of specific types of variation, which the experience of human biology tells us might possibly be a feature of the area being examined, and which might provide a clearer insight into the types of evolutionary processes that might account for the observed

differentiation. Underlying this approach is the recognition that variation has a "structure", which can be analysed by examining the way attributes relate to each other in individuals (phenetic relationships) and in populations (taxonomic relationships). Statistically, these relationships can be analysed through correlating attributes in individuals (statistical correlations) or in populations (taxonomic correlations) (Jardine and Sibson 1971). If taxonomic relationships are discordant, then we can theoretically predict that overall patterns of relationship between populations (overall structure), will be complex, resembling a "network" (Hiernaux 1972), rather than a series of clusters or clines. Conversely, if taxonomic relationships between attributes are concordant, then we might expect significant clustering or multivariate clines to be a feature of the overall structure.

The investigation of the way attribute relationships relate to the nature of overall structure is the basic springboard from which more precise analysis follows. Thus the association of concordant taxonomic relationships between attributes and the presence of multivariate clusters in theoverall structure could indicate that major types of variation such as ecotypic, ethnotypic or topotypic are influencing the total measured gene pool of the populations sampled. Discordant variation, however, associated with an absence of clusters, could indicate that different types of variation are influencing specfic attributes, independently. From these indications arising from the basic structural description, the researcher has a better idea as to where to concentrate his attempts at explanation.

The analysis of the basic structure of dermatoglyphic differentiation of the minimal ethnic population samples in Chapter 5, established that most attributes varied significantly, and that consequently any displayed taxonomic discordance was not due to random fluctuation associated with a situation of strong statistical homogeneity. The examination of taxonomic correlations through principal components analysis showed that there was considerable concordance between some attributes, especially between those in the same area of the hand (e.g. between digital attributes) but an absence of concordance between many other attributes, especially between those drawn from different areas of the hand. At the end of Chapter 5, I examined the extent of "true" taxonomic concordance, by controlling for underlying phenetic concordance. The results showed that though there was some significant taxonomic concordance in evidence, it was nevertheless low for all attributes except palmar ridge counts, which displayed higher concordance (though still moderate). I concluded, therefore, that the overall situation was one of strong taxonomic discordance.

In Chapter 6, I demonstrated through the analysis of $\Delta g$ distances through principal coordinates analysis, that the overall structure resembled a network, with clusters totally absent. These results were very similar to those found by Hiernaux (1972) for major sub-Saharan African populations. The overall patterns of relationships here, moreover, showed a strong relatively homogeneous central tendency, with only a few populations radically differentiated, though only on one or two components or vectors of variation, when detailed graphs were examined. For a majority of components of variation, these outlying populations were not distinguishable from the main concentration of populations.

The situation in Kenya, therefore, is one of considerable complexity, as we might expect when the basic structure is taxonomically discordant. In Chapters 5 and 6, the results follow theoretical expectations, with taxonomic discordance associated with an absence of multivariate clusters and clines. It is premature, however, to regard these findings as confirming this association. We need to contrast empirical situations of both concordant and discordant variation before any link with characteristic patterns of overall structure can be reliably established.

Nonetheless, though the overall situation is one of discordance, there is some low but significant taxonomic concordance demonstrated at the end of Chapter 5. In view of this, the findings of Chapter 8, in which moderate but significant ethnotypic effects were found, and of Chapter 9, where ecotypic effects were shown, are especially interesting. Although these effects are only moderate, the fact that they appear in a situation of low but significant concordance tentatively reinforces the theoretical relationship between the extent of taxonomic concordance and the strength of the operation of major types of variation.

C4: Problems and explanatory issues following on from the descriptive analysis.

An outstanding issue throughout the analysis has been the comparison of trends of variation between the male and female series of samples. In the descriptive investigations in Chapters 5 and 6, it became apparent that while broad structural trends of variation

are at best only moderately concordant between the sexes. The fact that detailed trends are not strongly replicated makes any attempt to account for the variation between samples difficult, as evidence provided by samples of one sex is likely to be contradicted by that provided by the other.

In Chapter 5 for instance, patterns of relationship between taxonomic attributes within sexes were very similar in both males and females, as were trends of "adjusted" taxonomic concordance. In contrast, there was marked discordance in the ranking of population samples in equivalent taxonomic attributes between the sexes, as is shown by the inter-sex correlations in Table 5.4.

Similarly, in Chapter 6 both sex series of samples exhibited a characteristic absence of clusters, and a "network" pattern of overall structure with a marked central concentration of a majority of population sample points, and a few sharply separated outlying sample points. In spite of this similarity, however, the detailed overall variation is discordant between the sexes. For example, the outlying populations in Figures 6.4 to 6.7 are mostly different ones in each sex. Moreover the ordination PCO scattergrams (Figures 6.11 to 6.28) show quite different patterns of population sample points from one sex to another for equivalent combinations of vectors.

One possible explanation for the moderate to low concordance in taxonomic rankings between the sexes, is that sample sizes are inadequate, and that such discordance reflects random fluctuation in means or frequencies. Although this is likely to be so to a

certain extent, the analysis in Chapter 7 demonstrates that non-random factors are also important. Particularly relevant is the discovery that the magnitude of sex differences in mean or frequency can differ appreciably from one population to another. I showed that this was not only the case when the total summary results from Kenya were compared with those from the Yorkshire Dales (Dennis 1977), but also when results for major ethnic populations are compared within Kenya. Thus the Cushitic ethnic category, for example, exhibited nearly twice the magnitude of sex differences for digital ridge counts than that of the Nilotic and Bantu ethnic categories. The sample sizes for these populations were large, thus effectively removing random variation as a serious explanatory possibility for these differences.

Another interesting observation is that the attributes that demonstrated the highest inter-sex concordance in Table 5.4, (double loops and total palmar ridge count), showed little variation in sex differences between the ethnic categories relative to other attributes. There is thus a suggestion of some relationship between variation in sex differences between populations, and the extent of inter-sex concordance.

If sex differences were more or less uniform from one population to another, then it would be possible to combine the sexes to produce larger sample sizes by either (a) collecting equal numbers of each sex for each sample, or (b) by weighting the scores of one sex to equalise the differences. The fact that this study has demonstrated that sex differences are not uniform between populations means that the combining of the sexes cannot be justified under any circumstance.

Furthermore, this study shows that the analysis of different patterns of variation between male and female series of samples must be a central aspect of dermatoglyphic regional investigation. In particular, we have seen that dermatoglyphic taxonomic relationships between the sexes are complex, and there is still a great deal to be learnt before any comprehensive explanation can be put forward. In researching this issue, I suggest that systematic comparison of dermatoglyphic sex differences between major world populations would be an especially important field of investigation in the future.

In Chapter 8, I turned to considering the extent to which patterns of ethnic affinity corresponded to patterns of dermatoglyphic affinity for ethnic population samples. One issue under scrutiny was the possibility that if sub-Saharan African populations are highly endogamous as Hiernaux (1972:97) implied, and that endogamy is primarily maintained by socio-cultural barriers to intermarriage, then African ethnic populations should in time (a) develop distinctive biological characteristics, and (b) become relatively homogeneous, presuming that there are no significant barriers to intermarriage between local territorial groups within ethnic units.

The analysis in Chapter 8 showed, however, that there was some significant heterogeneity within Kenyan ethnic groups, and that even where attributes did not differ significantly within ethnic groups, heterogeneous trends were nevertheless often consistent between the sexes. (For example the Kony in both sexes consistently exhibit much lower digital ridge counts than other Kalenjin populations, though statistical significance is not achieved for a majority of attributes in each sex).

Moreover, it was also apparent in Chapter 8.2.1.2 that even though correlations between ethnic dummy variables and taxonomic attributes were sometimes significant in most ethnic groups, even these significant correlations were at best moderate, and typically low, indicating an overall lack of distinctiveness for any ethnic group. These results, together with those of Chapter 6 that failed to find any discernable taxonomic clusters, point to the fact that the "high endogamy" model is inapplicable to the Kenyan situation.

Although heterogeneity within major ethnic groups is incompatible with the "endogamy" model, I pointed out in Chapter 8.3 that the existence of heterogeneity does not necessarily disprove any underlying relationship between culture/language and genepool over time. It is possible that, in the absence of widescale migration and hybridisation with alien populations, patterns of differentiation in language in particular and genepool will preserve some correspondence over time. Thus although the differentiation in genepool may have been considerable after many generations of spatial separation of populations, nonetheless some similarity might still be detectable between populations of a present ethnic group when contrasted to populations outside the group. Any similarity, moreover, would be likely to be manifested structurally, rather than in individual attribute elements. Hence the finding of structural similarity within an ethnic group implies common ancestry and a correspondence between language and genepool over time, while an absence of such similarity implies different origins for the constituent populations of the ethnic group, and a more complex history.

The main problem with testing such a model is the question of how we can measure and detect "structural similarity". The difficulty of this problem is illustrated at the end of Chapter 6, where I questioned the usefulness of conventional ordination and graphical techniques as indicators of overall similarity in a situation of discordant variation. In Chapter 8.3, I tackled the problem experimentally, by adopting an unorthodox graphical technique, in which overall structural similarity between samples is depicted by the "shape" of lines when PCO vector scores are linked, and homogeneity is depicted by "distance" between lines of similar shape.

When the Kenyan minimal ethnic population samples were examined using this technique, the results were mixed. The Cushitic samples were very similar in shape, though not always in distance, the results consistent in both sexes. This implies a situation of common origins for the Cushitic samples, with heterogeneity in distance implying appreciable separation over time between some populations. The Mji Kenda show very similar trends to the Cushites, with population samples clearly related in shape in both sexes. Moreover the Ngwesi and Mukogodo Dorobo samples are not only very similar to each other in both distance and shape, but also very similar to neighbouring Cushitic populations. This implies a Cushitic origin for these peoples, rather than a Maasai or aboriginal origin.

In the remaining ethnic groups, however, difficulties arose in that patterns of similarity from one sex to another were contradictory. Nevertheless, some overall trends could be discerned.

Firstly, the Luyia and non-M_ji Kenda Coastal Bantu were very heterogeneous. Secondly, the Central Bantu samples showed a range of differing characteristic shapes in the males, and a smaller more homogeneous range in the females, implying perhaps that some ethnotypic diffusion had taken place among three or four major groups Of incomers in the past. Finally, the Nilotic ethnic groups tended to demonstrate a majority of samples with similar shapes within each group, and the presence of one or more population that differed substantially from the rest in the group. This type of situation perhaps implies that a central bloc of populations within an ethnic group share common origins, while one or two per_ipheral populations either originally migrated into the area and adopted the language and culture of the dominant ethnic group, or significantly hybridised with outsiders.

The investigation of structural similarity has been a useful approach in this study. The explanatory usefulness of the approach, however, has been limited by the familiar complications of inter-sex discordance in patterns of variation, and the inadequacy of some of the sample sizes in some ethnic groups. It can hardly be a coincidence, for instance, that the most heterogeneous ethnic group, the Luyia, contain the smallest sample sizes.

In Chapter 9 I moved on to examine the possibility that dermatoglyphic attributes vary systematically between environmental and climatic zones. This area of investigation is especially important, given that very little is known about the role of natural selection in determining dermatoglyphic variation. Indeed Holt states that:

> "dermal ridge patterns are presumably not markedly affected by natural selection to any marked extent."
>
> (Holt 1975:176)

In view of this comment, and the failure of Hiernaux (1977:193) to find any systematic climatic variation for digital patterns in Africa, it is remarkable that in this study I succeeded in demonstrating significant variation between altitude and rainfall zones in Kenya, for many digital and palmar attributes (see Chapter 9). Not only did a minority of attributes correlate as high as .4 to .5, with either altitude or rainfall, but the analysis of the nature of the significant relationships showed that clinal variation was common. On the whole, digital ridge counts, concentric whorls and double loops tended to relate negatively to rainfall and altitude, with means and frequencies often clinally decreasing as rainfall or altitude increased; and palmar ridge counts and ulnar loops tended to show a positive increase in means and frequencies as altitude and rainfall increased.

These results imply that some kind of climatically and environmentally linked natural selection is operating on dermatoglyphic traits in Kenya. In Chapter 9.4, however, I suggested that it was unlikely that rainfall and altitude had any direct selective influence. It is more probable that rainfall and altitude in Kenya have only an indirect effect, in that they affect the distribution of the real but unknown selective force.

In addition to finding that many dermatoglyphic attributes vary systematically with rainfall and altitude, I also discovered in a parallel investigation that the distribution of deuteranopia also shows a systematic relationship with these variables (Rosa 1981).

As none of these attributes have been linked with climatic and environmental variation in previous studies, two possibilities are suggested. Firstly, it is possible that similar associations will be found in other parts of Africa, or even outside Africa. In this case the failure to find associations in the past could be due to the fact that physical anthropologists have not seriously searched for them. Secondly, and alternatively, it is possible that the associations found in this study are merely local phenomena, confined to Kenya or maybe East Africa. If this is true, then the nature of natural selection acting on dermatoglyphic traits in Kenya is likely to be complex, and localised in its effects. It is impossible, however, in the light of present evidence, to judge which of these possibilities is more likely.

## C5: Concluding remarks

The survey reported in this study is the largest and most comprehensive undertaken in Africa in the field of dermatoglyphics. Perhaps its most valuable long term contribution lies in the detailed univariate results in Appendix 1, which can be accessed by researchers from many fields of human biological interest in the future.

It is clear from the present investigation that the potential provided by this data has only been partially realised. Firstly the field of dermatoglyphic variability in individuals has only been superficially touched upon here at the end of Chapter 5. It would be especially interesting to see how far patterns of dermatoglyphic variability in Kenya are similar to

patterns in populations elsewhere. A comparison with the compatible British series in Durham is an obvious area of future inquiry. Secondly the testing for major types of variation is still to be fully explored. A notable omission in this study has been the examination of geographical or topotypic variation. Thirdly even in the areas examined, there is still much more that can be done. In Chapter 9, for instance, it would be profitable (a) to investigate dermatoglyphic variation between altitude and rainfall zones from the point of view of pooled individuals, rather than from pooled population sample means and frequencies; (b) to re-examine the relationships with altitude and rainfall using different criteria for defining zones (e.g. in the case of altitude zones could be defined in 1000 ft. units rather than the more general divisions adopted in Chapter 9); (c) to explore relationships with other climatic variables such as humidity, solar radiation and temperature. Moreover rainfall zones could be redefined according to criteria based on marginal rather than mean annual rainfall. Finally, there is the exploitation of the findings for historical and ethnological inquiry. Something of the potential usefulness of the present results to this field is illustrated at the end of Chapter 8, particularly in the analysis of structural similarity in the Dorobo and Mji Kenda samples. In many respects, therefore, this present work is just a beginning.

